



HAL
open science

Construire les usages des données ouvertes en France : proposition d'un Open Data Canvas par l'approche du design science

Arthur Sarazin

► To cite this version:

Arthur Sarazin. Construire les usages des données ouvertes en France : proposition d'un Open Data Canvas par l'approche du design science. Gestion et management. Université Grenoble Alpes [2020-..], 2021. Français. NNT : 2021GRALG007 . tel-03496226

HAL Id: tel-03496226

<https://theses.hal.science/tel-03496226>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Sciences de Gestion**

Arrêté ministériel : 25 mai 2016

Présentée par

Arthur SARAZIN

Thèse dirigée par **Carine DOMINGUEZ-PÉRY**, professeur à l'**Université Grenoble-Alpes**,
et par **Khaled BOUABDALLAH**, professeur à l'**Université de Lyon**

préparée au sein du **Laboratoire CERAG EA 7521**
dans l'**École Doctorale Sciences de Gestion**

Construire les usages des données ouvertes en France : proposition d'un Open Data Canvas par l'approche du design science

Thèse soutenue publiquement le **1er Juillet 2021**,
devant le jury composé de :

Madame Françoise PAQUIENSÉGUY

Professeur des Universités, SciencesPo Lyon (Présidente)

Madame Amandine PASCAL

Professeur des Universités, Aix-Marseille Université (Rapporteur)

Madame Amel ATTOUR

Maître de conférences HDR, Université Côte d'Azur (Rapporteur)

Monsieur Samuel FOSSO WAMBA

Professeur titulaire à Toulouse Business School, HDR de l'Université d'Evry Val d'Essonne (Suffragant)

Monsieur Emmanuel COUX

Directeur des Moyens d'Information, Grenoble-Alpes Métropole et Ville de Grenoble (Membre invité)

Madame Carine DOMINGUEZ-PÉRY

Professeur des Universités, Université Grenoble-Alpes (Directrice de thèse)

Monsieur Khaled BOUABDALLAH

Professeur des Universités, Université de Lyon (Directeur de thèse)



Centre d'Études et de Recherches Appliquées à la Gestion de Grenoble

ÉCOLE DOCTORALE DE SCIENCES DE GESTION

Université Grenoble Alpes

Construire les usages des données ouvertes en France

*Proposition d'un Open Data Canvas par l'approche
du design science*

Thèse soutenue par Arthur Sarazin

Thèse financée par la région Auvergne-Rhône-Alpes

Sous la direction de Carine Dominguez-Péry (Université de Grenoble Alpes)
et Khaled Bouabdallah (Université de Lyon)



L'université n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Celles-ci doivent être considérées comme propres à leurs auteurs.

À Laurie...

« La douleur est une des choses les plus importantes de ma vie. Le mot “écrit” ne conviendrait pas. Je me suis trouvée devant des pages régulièrement pleines d’une petite écriture extraordinairement régulière et calme. Je me suis trouvée devant un désordre phénoménal de la pensée et du sentiment auquel je n’ai pas osé toucher et au regard de quoi la littérature m’a fait honte »

Marguerite Duras

Remerciements

Mes remerciements sont adressés en premier lieu à Carine Dominguez-Péry et Khaled Bouabdallah qui ont accepté de diriger cette thèse.

Ce travail n'aurait pas été possible sans le soutien de la Région Auvergne-Rhône-Alpes, qui m'a permis, grâce à une allocation de recherches, de me consacrer sereinement à l'élaboration de ma thèse.

Je tiens à remercier Anne Bartel-Radic pour son accompagnement et son soutien avant et pendant cette thèse.

Je tiens également à exprimer ma reconnaissance à l'équipe de Dataactivist pour leur ouverture et pour m'avoir encouragé tout au long de ces trois années. Merci enfin à l'ensemble des personnes rencontrées au cours de mes recherches de terrain pour leur disponibilité et leur confiance.

Merci aux chercheurs et à toute l'équipe administrative du CERAG pour leur accueil, leur accompagnement et leurs conseils. Être intégré à une vie de laboratoire a été une grande aide pour la réalisation de cette thèse. Merci à la team des doctorants qui a rendu ce travail plus chaleureux et convivial.

Résumé

Construire les usages des données ouvertes : une approche par le design science

Depuis l'entrée en vigueur de la Loi pour une République numérique en 2016, toutes les collectivités de plus de 3500 habitants ont l'obligation de publier leurs données de manière accessible, gratuitement et dans un format ne nécessitant pas l'utilisation d'un logiciel spécifique. Derrière la loi transparaît la promesse de voir émerger de nouvelles applications, services, informations issues de l'utilisation que le public ferait des données ouvertes. Or, si l'ouverture des données publiques se veut être un facteur qui encourage le développement de nouveaux usages, une telle démarche n'est pas nécessairement un gage de succès. En effet, depuis 2016, très peu d'usages utilisant l'ouverture des données ont vu le jour, à tel point que certains activistes définissent la période actuelle comme celle du "baby blues de l'open data". Dans ce cadre, l'objet de cette thèse est de construire un dispositif en Systèmes d'Information pour lever les barrières qui freinent la construction des usages. Une approche par le design science a été adoptée. Une première phase qualitative a permis de développer un modèle théorique de compréhension des phénomènes à l'œuvre dans la construction des usages des données ouvertes et des barrières socio-techniques qui freinent le processus. Une seconde phase de conception a permis de construire le design d'un dispositif en Systèmes d'Information appelé Open Data Canvas qui permet de lever ces barrières. Il s'appuie sur trois théories : la théorie de la construction sociale de la technologie, la théorie de la cognition distribuée et la théorie du middleground. Enfin, une dernière phase dite d'évaluation a permis d'interroger les usagers des données ouvertes sur l'utilité perçue de l'Open Data Canvas. Ces trois phases se sont affinées itérativement pour permettre au Canvas de s'intégrer dans le quotidien des usagers des données ouvertes.

Mots clés : Données ouvertes - Plateforme - Usage - Écosystème open data - Design science

Building new uses out of open data : a design science approach

Since the enactment of the Law for a Digital Republic in 2016, all public localities comprising over 3.500 inhabitants are required to publish their data in an accessible manner, gratuitously, and in such a format to not require the use of any specific software. Behind this Law lies the promise of new applications, services and information, resulting from the public use of open data. However, while opening up public data is meant as an encouraging measure for the development of new uses, it is not necessarily guaranteed to succeed. Indeed, since 2016, new uses of public data have been so infrequent, that activists define the current situation as the "open data baby blues". In such a context, this thesis aims to build an Information System device designed to remove the barriers hindering the development of new uses. We apply a design science approach to this problem. First, a qualitative phase develops a theoretical framework highlighting the various phenomena at work in the construction of new uses of open data, as well as addressing the socio-technical barriers preventing such construction. In a second time, we develop the Information System device, based on the design science approach, named the Open Data Canvas (ODC). The ODC is designed to remove the socio-technical barriers mentioned above. For doing so, it benefits from three theories: the Theory of the Social Construction of Technology, the Theory of Distributed Cognition and the Middleground Theory. Finally, we perform the concluding step of our process: the evaluation stage. This phase enables us to understand how to question open data users on their perceived usefulness of the ODC. The three steps presented above were refined iteratively in order to allow the ODC in being smoothly integrated into the daily life of open data users.

Keywords : Open Data - Platform - Reuse - Open Data Ecosystem - Design science

Sommaire

Introduction générale	12
1. Avant-propos	13
2. Problématique de recherche	15
3. Propositions de la recherche	21
<u>Chapitre 1. La revue de la littérature</u>	35
Section 1. Usages des données ouvertes	39
Section 2. Plateformes open data	53
Section 3. L'ouverture des données	68
Section 4. Les écosystèmes <i>open data</i>	79
Section 5. Identification du Gap	91
Section 6. Une absence de recherche utilisant la méthodologie de recherche design science	93
<u>Chapitre 2. Méthodologie de recherche design science</u>	154
Section 1. Présentation de l'architecture de recherche	156
Section 2. Présentation des cinq étapes du design de recherche	162
<u>Chapitre 3. Les barrières à la construction des usages issus des données ouvertes</u>	204
Section 1. Invention et conceptualisation des difficultés rencontrées dans l'innovation de services issue des données ouvertes. .	206
Section 2. Choix des observations à effectuer et réalisation	221
Section 3. Confrontation des observations avec les praticiens et formulation d'un nouveau cadre conceptuel .	272
<u>Chapitre 4. Les attentes des opérateurs de la donnée vis à vis d'un dispositif en Systèmes d'Information</u>	287
Section 1. Sélection et Présentation des cas	290
Section 2. Construction du protocole et instruments de l'étude de cas	296
Section 3. Analyse des données : codage des entretiens en utilisant un système permettant d'identifier les barrières socio-techniques et les moyens imaginés pour les lever	302
Section 4. Résultats et discussion : un besoin d'avoir des connaissances partagées entre opérateurs pour lever plus facilement les barrières	318
<u>Chapitre 5. Design d'un dispositif en Systèmes d'Information</u>	334
Section 1. Construction des propositions de design	339
Section 2. Construction des principes de design	352
Section 3. Construction du design de l'artefact	370
Section 4. Construction de scénarios d'usage compilant les pattern de coordination de l'artefact avec un contexte réel de développement des usages des données ouvertes	391
Section 5. Description des fonctions du design de l'environnement de conception des usages	

des données ouvertes.	401
<u>Chapitre 6.</u> Prototypage d'un dispositif en Systèmes d'Information	411
Section 1. Objectifs du prototype	412
Section 2. Sélection des fonctions	415
Section 3. Construction du prototype	420
<u>Chapitre 7.</u> Evaluation du dispositif en Systèmes d'Information	458
Section 1. Analyse de l'évaluation des prototypes	460
Section 2. Choix d'une stratégie d'évaluation	463
Section 3 Choix de la méthode d'évaluation	467
Section 5. Réalisation	467
Section 6. Implications	493
<u>Chapitre 8.</u> Discussion et conclusion générale	498
Section 1. Résultats marquants de la recherche	499
Section 2. Limites et ouverture de la recherche	516

Introduction générale

1. Avant-propos

1.1 L'ouverture des données : le nouvel or noir ?

Le 2 mai 2016, le groupe Oracle, géant américain de l'édition de logiciel, fait l'acquisition d'Opower pour un montant de 532 millions de dollars. Opower développe des applications pour faciliter la relation entre les distributeurs d'énergie et leurs clients. Derrière la valorisation impressionnante d'Opower se cache un fait surprenant : cette entreprise n'exploite que des données gratuites et accessibles à tous. En effet, ses applications sont basées sur les données produites par l'Agence d'Information de l'Énergie dans le cadre d'une enquête sur la consommation énergétique des ménages. Des données disponibles en accès libre.

Fait encore plus surprenant, l'exploitation de ces données gratuites et accessibles par Opower a permis de faire économiser au contribuable américain 200 millions de dollars de facture énergétique (US Department of Energy, 2013).

On l'aura compris, l'histoire d'Opower est surprenante à bien des égards : à partir d'une ressource gratuite, cette entreprise a su construire une suite d'applications nouvelles qu'elle a pu valoriser plus de 500 millions de dollars. Aussi a-t-elle su mettre en place, avec ces données gratuites, de nouveaux usages (i.e, le contrôle de la consommation d'énergie à partir d'une application servant de capteur) qui bénéficient à la fois aux citoyens qui font des économies d'énergie ; aux distributeurs qui fidélisent leur clientèle ; et au gouvernement américain qui accélère la transition écologique du pays. Enfin, elle fait partie de ces *success story* qui ont amené des gouvernements souverains à ouvrir les vannes d'accès à leur "nouvel or noir" : les données.

1.2 L'insuffisante ouverture des données pour créer de nouveaux usages

Ces nouveaux usages créés à partir de données gratuites et libres d'accès ont participé au développement d'un mouvement politique dit "mouvement *open data*". Né dans les années 2000, il demande aux gouvernements et aux collectivités d'ouvrir l'accès aux données publiques. Quelques années plus tard, les gouvernements s'emparent de la question et légifèrent en ce sens. En janvier 2009, Barack Obama signe deux mémorandums faisant acte du fort engagement de son administration dans les démarches d'ouverture des données. Cet acte marque le début des politiques d'ouverture des données publiques dans le monde. tandis qu'en France, il faut attendre la création de la mission Etalab en 2011. En 2015, la loi

pour une République numérique accélère la démarche en fixant une obligation d'ouverture des données de toutes les collectivités de plus de 3500 habitants.

Si l'ouverture des données publiques se veut être un facteur encourageant le développement de nouveaux usages, une telle démarche n'est pas nécessairement un gage de succès. Depuis l'instauration de la loi pour une République numérique, très peu d'usages utilisant l'ouverture des données ont vu le jour, à tel point que certains activistes ont parlé d'un "baby blues de l'open data" (Libertic, 2013). De nombreuses études abondent en ce sens et montrent que l'ouverture de données ne suffit pas à créer de nouveaux usages (Zuiderwijk et al., 2013, Whitemore, 2014, Kuhn, 2011, Jaakola et al., 2015).

En effet, pour générer de nouveaux usages à partir des données ouvertes, il ne suffit pas d'ouvrir les données. Les usages ne sont pas une conséquence indirecte de l'ouverture mais plutôt le fruit d'une démarche spécifique de construction (Goëta, 2016, Ruijter et al., 2017). L'usage des données ouvertes se définit dans ce cadre comme une construction socio-technique émergeant dans un cadre de compréhension partagée qui rassemble trois sphères sociales et en conséquence de leurs interactions, chacune réalisant une des trois activités clés à savoir la production, l'édition ou la réutilisation des données ouvertes (inspiré de Paquienséguy, 2007). Cette définition implique que, pour faire usage des données ouvertes, les acteurs doivent coordonner leurs pratiques, mettre en commun leurs compétences, connaissances et ressources et se mettre d'accord sur les résultats à atteindre. Dans ce cadre, il est courant que les acteurs s'appuient sur une plateforme open data.

Outre le fait qu'elles permettent de maximiser la densité des données ouvertes à un endroit donné et d'augmenter leur liquidité (i.e, capacité à circuler entre les acteurs), elles supposent d'instaurer une forme d'organisation qui régit les interactions entre acteurs.

Or, la mise en place d'une telle organisation s'avère très délicate pour deux raisons. La première provient du fait que les pratiques en lien avec l'usage des données ouvertes sont nouvelles et qu'aucune d'entre elles n'a été décrite et appliquée avec rigueur (Safarov et al., 2017). La deuxième raison est issue de l'hétérogénéité des usagers qui appartiennent à des sphères sociales qui n'ont pas l'habitude de travailler ensemble (Ruijter et al., 2017). *In fine*, la relative nouveauté des pratiques couplée à l'hétérogénéité des usagers rend les méthodes et outils traditionnels de gestion inefficaces pour générer de nouveaux usages et appelle au design de nouveaux outils (Zuiderwijk, 2015, p.8-9).

Aussi, la construction des usages est rendue difficile par le fait qu'elle soulève des tensions paradoxales entre les intérêts de l'organisation commune d'un côté, et des utilisateurs pris individuellement de l'autre.

Le premier et le principal de ces paradoxes concerne la mise à disposition des données dont il est question de faire usage. Si l'organisation a tout intérêt à faire circuler les données des uns et des autres dans le but de créer de maximiser la construction, les utilisateurs eux, n'ont aucun intérêt à mettre à disposition leurs données. Dans la mesure où ces données constituent pour eux une ressource clé de leur activité, ils risqueraient en les partageant de voir de nouveaux concurrents émerger, fort de l'exploitation de leurs propres données. Les créatures échapperaient à leurs créateurs ! Inversement, si les utilisateurs ont intérêt à garder leurs données pour leur propre compte, l'organisation, elle, perd toute son utilité si elle ne peut faire circuler les données.

Dans ce cadre, je peux dire que l'usage des données ouvertes soulève des problèmes dit pernicieux (Churchman, 1967, Rittel et Weber, 1973). Autrement dit, l'usage des données ouvertes réveille des tensions préexistantes au sein des organisations, entre les organisations mais surtout entre les différentes strates d'une organisation. On parle de tensions paradoxales qui sont définies comme les situations où "les différents niveaux d'analyse se confrontent les uns aux autres de manière contradictoire tout en étant nécessairement interdépendants" (K. Smith et al., 2017, p.305).

Cette thèse vise à proposer aux praticiens et aux chercheurs des pistes de réflexions pour résoudre ces problèmes.

2. Problématique de recherche

Pour illustrer les tensions paradoxales qui peuvent apparaître et introduire notre problématique de recherche, je présenterai un cas d'usage des données ouvertes.

2.1 Cas d'usage des données ouvertes

Je prendrai l'exemple d'un projet de réutilisation de données que j'ai suivi tout au long de la thèse. Je nomme ce projet : "Tri" afin de préserver la confidentialité de l'identité des porteurs. "Tri" est un projet d'application qui vise à aider les contribuables à trier leurs déchets. Dans le cadre de son développement, "Tri" s'est mis en relation avec une collectivité et une entreprise pour obtenir les ressources dont elle a besoin. Parmi elles, on trouve les données relatives aux lieux de collecte de déchets, aux consignes de tri et à la chaîne de recyclage des déchets.

Cependant, le développement d'usages à partir des données réveille des tensions préexistantes. De fait, les porteurs du projet ont été ralentis à la fois par des tensions au sein de la collectivité et de l'entreprise, par des tensions entre la collectivité et l'entreprise et même par des tensions situées en dehors de ces derniers. Par exemple, les tensions intra-organisationnelles se sont manifestées par un manque d'entente entre le service "déchet" et le service "données" de la métropole ; ce manque d'entente étant causé par l'incompatibilité des profils des agents et des outils technologiques présents au sein de ces deux services. Quant aux tensions interorganisationnelles, elles se sont exprimées à travers les intérêts divergents de la métropole et de l'entreprise sur le projet ; alors que l'entreprise comptait augmenter sa rentabilité à court terme, la métropole espérait faire des économies sur le long terme.

Du fait de ces tensions, l'application "Tri" n'a pas encore vu le jour et les données n'ont pas été utilisées.

2.2 Les écosystèmes open data comme espace-problème de la recherche

Ce cas soulève plusieurs points cruciaux que les utilisateurs doivent traiter s'ils veulent réussir à développer de nouveaux usages. J'ai choisi de les présenter sous forme d'un ensemble de questions pratiques que les utilisateurs peuvent se lancer dans une démarche de développement des usages :

- Comment s'assurer que les données qui seront utilisées pourront répondre aux besoins des futurs bénéficiaires de l'usage construit ?
- Comment s'assurer que les objectifs des parties prenantes du projet ne sont pas incompatibles ?
- Comment s'assurer que la diversité des profils au sein d'une organisation n'empêche pas la mise à disposition de la donnée ?
- Comment s'assurer de la complémentarité des actions indépendantes menées par chaque acteur ?
- Comment s'assurer et pérenniser l'utilité publique de l'usage construit au-delà des intérêts privés qu'il peut satisfaire ?
- Comment maintenir la collaboration entre les acteurs autour de ce projet dans le temps pour pérenniser l'usage construit ?

Les modes d'organisation et les pratiques d'usage de la donnée ouverte soulevées à travers ces questionnements s'inscrivent tous le cadre d'écosystèmes open data (Pollock, 2011). Ces écosystèmes sont des écologies informationnelles dont l'équilibre repose sur la circulation de jeux de données ouverts

qui vont venir structurer les systèmes qu'ils contiennent, de la même que les espèces ingénieur¹ dans les écosystèmes biologiques.

Ces écosystèmes constituent l'espace-problème (Newell et Simon, 1972) de mes recherches. Autrement dit, les écosystèmes open data constituent l'espace dans lequel j'identifierai les problèmes des praticiens, la problématique de recherche qui guidera cette thèse. Aussi c'est dans le cadre des frontières dans cet espace que je concevrai la solution à ces interrogations.

2.3 La question de recherche, un besoin à la fois théorique et pratique

Le concept d'écosystème s'est avéré capital pour concevoir l'organisation et les pratiques à mettre en place pour coordonner les efforts d'acteurs et d'outils très hétérogènes.

Par exemple, dans une unité de soin intensif d'un hôpital, l'organisation et les bonnes pratiques à mettre en œuvre peuvent se concevoir à travers ce concept. En effet, on y trouve une myriade d'individus avec différentes expertises (infirmières, administrateurs, médecins, chirurgiens, etc.) et une gamme d'outils technologiques très vastes (moniteurs, scanners, etc.). Malgré les tensions inhérentes à cette diversité, ces deux ensembles doivent se coordonner pour prodiguer des soins mais aussi s'adapter rapidement pour faire face aux fluctuations du volume de patients à traiter. La mise en place et le développement d'un écosystème permet de répondre à ces deux exigences et favorise l'atteinte des objectifs de l'unité.

De la même manière, construire un écosystème open data impliquant les porteurs du projet "Tri", la métropole et l'entreprise permettrait de lever les tensions paradoxales qui font obstacle à l'usage des données ouvertes et favoriserait la construction de nouveaux usages

La mise en place d'un écosystème open data aurait par exemple permis aux services "déchets" et "données" de la métropole de réfléchir à l'interopérabilité des outils technologiques qu'ils utilisent. Ainsi, ils auraient pu créer une chaîne automatique de production de données ouvertes, partant des applications qui récoltent et stockent les données "déchets" pour aller jusqu'aux outils du service "données" qui peut alors les libérer sur une plateforme. Cette chaîne est d'autant plus cruciale qu'elle permettrait au projet "Tri" de s'assurer que les données dont ils disposent sont fiables et actualisées.

Surprenamment, alors que les projets d'usage de la donnée ouverte se multiplient, se complexifient, et impliquent des outils et des individus aux capacités de plus en plus hétérogènes, la littérature demeure

¹ Aussi appelée "espèce clé de voûte", ce concept issu de l'écologie décrit les espèces qui structurent et assurent l'équilibre d'un écosystème (Byers et al., 2006).

relativement muette quant aux écosystèmes open data à mettre en œuvre dans de telles configurations. Écosystèmes qui permettraient de résoudre les tensions paradoxales au sein de chaque niveau d'analyse mais aussi entre les niveaux d'analyse.

Certes, plusieurs travaux antérieurs ont proposé des pratiques et des modes d'organisation qui permettraient de faciliter la construction de nouveaux usages. Or, très peu d'études prennent pleinement en considération les écosystèmes open data dans lesquelles ils s'insèrent ; peu d'auteurs tirent jusqu'au bout les conséquences méthodologiques qu'un tel espace-problème impose à la recherche.

Certains préconisent des solutions à mettre en œuvre à l'échelle d'un seul et unique niveau d'analyse. De nombreuses études offrent des solutions au niveau intra-organisationnel (Le Corf, 2016, Eckelberg et al., 2018, Loveluck, 2015). Parmi eux, Le Corf (2016) propose aux projets open data dans les organisations d'adopter le modèle de participation et de *crowdsourcing* comme mode de gestion. D'autres études ont mis en avant des solutions au niveau organisationnel (Courmont, 2015, McLeod et McNaughton, 2016, Zuiderwijk, 2015). Parmi eux, Courmont (2015) suggère de créer des instruments de gouvernance de la donnée, constitués d'opérateurs, de managers et de représentants de la donnée dans les organisations.

Je trouve également dans la littérature des études qui superposent les niveaux d'analyse et proposent des solutions dans lesquels ces derniers cohabitent (Danneels, 2017, Jetzek et al., 2014, Ojo et al. 2015). Parmi eux, Danneels (2017) propose un mode d'organisation à deux niveaux pour construire de nouveaux usages à partir des données ouvertes. Au niveau organisationnel, une entité unique se charge de mettre à disposition les données auprès des utilisateurs. Au niveau interorganisationnel, la même entité offre les moyens aux utilisateurs de communiquer entre eux.

Cependant, aucune étude ne propose aujourd'hui de solution qui prenne en compte l'espace-problème que constituent les écosystèmes open data, les tensions paradoxales entre les niveaux d'analyse et qui se proposent de les résoudre, pour favoriser la construction de nouveaux usages. De plus, les solutions offertes dans les études antérieures ne sont pas transposables parce qu'elles n'abordent pas la particularité des problèmes pernicieux. Par conséquent, ces solutions seraient immédiatement rendues inopérantes par les tensions paradoxales auxquelles elles seraient immédiatement soumises.

Dès lors, l'objectif de cette thèse est de proposer un mode d'organisation et des pratiques qui permettraient de prendre en compte ces tensions paradoxales et de les résoudre pour favoriser la construction de nouveaux usages. Notre question de recherche est donc la suivante :

Question de recherche : comment résoudre les tensions paradoxales inhérentes aux écosystèmes open data qui font obstacle à la construction de nouveaux usages à partir des données ouvertes ?

Cette problématique représente un besoin à la fois pratique et théorique. En effet, cette problématique appelle à une solution qui, d'un côté, permettrait aux utilisateurs de données ouvertes d'améliorer leurs performances et favoriserait la réussite des projets de construction de nouveaux usages qu'ils mènent conjointement. De l'autre, la solution viendra combler le gap dans la littérature sur l'usage des données ouvertes en proposant une méthodologie et un modèle de solution applicable à une nouvelle catégorie de problèmes en sciences de gestion : les problèmes pernicieux.

Je subdivise la question de recherche en trois sous-questions qui répondent à l'approche *design science* que j'ai choisi pour répondre à la question de recherche. Cette approche stipule au chercheur de mener de manière itérative trois cycles d'activités spécifiques.

Le premier cycle dit cycle de rigueur permet au chercheur d'identifier dans la littérature les connaissances antérieures qui vont venir informer la conception d'un dispositif. C'est dans ce cadre que je me demanderai :

Sous-question de recherche 1 : quel est l'état des lieux concernant la construction des usages ?

J'effectuerai une revue de littérature qui fera le tour d'horizon des champs de la littérature connexes à la construction des usages issus des données ouvertes.

Le deuxième cycle, dit cycle de pertinence, invite le chercheur à identifier les situations problématiques rencontrées par les praticiens et auxquelles la recherche peut apporter ses connaissances et son soutien. Il sera question dans ce cadre de savoir :

Sous-question de recherche 2 : quelles sont les situations problématiques qui freinent la construction des usages issus des données ouvertes ?

Enfin, le troisième cycle dit cycle de conception, suppose pour le chercheur d’alterner entre des activités conception d’un dispositif et d’évaluation de sa capacité à répondre à la problématique de recherche. C’est dans ce cadre qu’il s’agira de savoir :

Sous-question de recherche 3 : quels sont les éléments d’un dispositif en Systèmes d’Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ?

Ce dispositif en Systèmes d’Information, catégorisé ainsi car il s’agit d’un dispositif informatisé avec des données entrantes et sortantes, prendra la forme d’un framework. Autrement dit d’un cadre ou environnement (*frame*) permettant de faire advenir le travail de construction des usages (*open data work*). Le choix du framework comme modèle de solution s’explique par la spécificité de mon espace-problème : les écosystèmes open data. Je rappelle, avec Nardi et O’Day (1999) que :

“La dynamique des écosystèmes est une dynamique de flux et de mouvement - les personnes, les idées, les activités et les outils sont en mouvement alors que l’écosystème évolue continuellement sous la forme d’éléments qui s’ajustent et sont ajustés les uns par rapport aux autres, en essayant toujours et jamais en réalisant une adéquation parfaite.” Nardi et O’Day (1999, p.53)

Par conséquent, un dispositif visant à lever les barrières qui freinent la construction des usages ne peut qu’encourager le flux de mouvement permanent et ces ajustements entre les différents éléments du travail de construction des usages. De plus, ce dispositif doit-il pouvoir adresser les problèmes et solutions qui sont en perpétuelle évolution.

Le modèle du framework répond à ces deux exigences dans la mesure où il offre un espace permettant de générer de nouvelles combinaisons entre diverses composantes, donnant ainsi la possibilité aux acteurs clés de l’écosystème d’adapter en permanence la construction des usages.

3. Propositions de la recherche

3.1 Méthodologie adaptée

Afin de répondre à cette question, j'utiliserai dans cette thèse une méthodologie de recherche inspirée de la *design science research methodology* (Hevner et Chatterjee, 2010) appliquant de manière itérative des procédés de caractérisation de problème, de construction d'un dispositif et d'évaluation de celui-ci (Hevner et al., 2004).

Dans un premier temps, je me suis appuyé sur **une revue de littérature pour faire l'état des lieux sur la construction des usages et ainsi répondre à la première question de recherche** :

- **Revue de littérature** : une revue de littérature autour du concept d'usage des données ouvertes et des concepts périphériques a permis d'extraire, parmi les travaux antérieurs, tous les éléments qui permettent de comprendre les situations problématiques qui freinent la construction des usages. J'ai également identifié les facteurs qui peuvent permettre de comprendre le contexte dans lequel ces situations apparaissent. A l'issue de cette revue de littérature, j'ai conçu un modèle préliminaire schématisant l'usage des données ouvertes.

Dans un deuxième temps, j'ai adopté **une démarche qualitative afin de répondre à la deuxième question de recherche**. Il s'agit de caractériser les situations problématiques que rencontrent les usagers de données ouvertes et qui font obstacle à la construction des usages :

- **Études de cas exploratoires** : quatre cas d'usage des données ouvertes ont été explorés de façon à pouvoir discuter, compléter ou réfuter les conjectures théoriques issues du modèle préliminaire. J'ai pour cela mis en place une démarche d'exploration hybride (Giordano et Jolibert, 2012) basée sur des observations participantes. J'ai prêté une attention toute particulière à sélectionner des cas représentant la diversité des démarches d'usages des données ouvertes et des contextes dans lesquels ils peuvent advenir. À l'issue de cette partie, j'obtiens un modèle affiné de l'usage des données ouvertes qui représente les tensions paradoxales qui lui font obstacle. Ce modèle constitue une heuristique. Autrement dit, il s'agit d'une « représentation simplifiée destinée à expliquer/schématiser la situation réelle étudiée » (Charreire et Durieux, 1999, p.65).
- **Étude de cas comparative** : sur la base du modèle heuristique, une étude approfondie et comparative des cas de Grenoble et de Lyon a été menée de façon à pouvoir identifier les besoins

des utilisateurs pour pouvoir résoudre ces tensions. On parle de besoins fonctionnels². Ces éléments théoriques sont définis comme “ce que les futurs utilisateurs attendent d’un design” (Dym et Little, 2004). Pour cela, j’utiliserai la méthodologie d’Eisenhardt (1989) qui permet de mettre en évidence des éléments théoriques à partir d’une étude de cas.

Dans un troisième temps, j’ai adopté **une démarche de construction et d’évaluation d’un dispositif pour répondre à la troisième question de recherche**. Ce dispositif visera à résoudre les situations problématiques rencontrées dans la première partie :

- **Conception d’un dispositif** : la conception du dispositif se déroule en trois étapes. La première étape consiste à identifier dans la littérature les mécanismes permettant de répondre aux attentes des utilisateurs du dispositif. Selon la terminologie utilisée par les chercheurs en *design science*, ces mécanismes constituent des propositions de design, définis comme “les lignes directrices pour la création de solutions répondant à une catégorie particulière de problèmes empiriques” (Denyer et al., 2008, p;395). La deuxième étape vise à identifier les fondements théoriques³ qui renforcent la fiabilité des mécanismes identifiés lors de la première étape. Ces fondements théoriques sont définis comme : “les connaissances ou théories, venant des sciences naturelles, sociales ou de design, qui expliquent la logique de conception” (Gregor et Jones, 2007, p.322). Ces fondements théoriques constituent des principes de design, définis comme “les normes permettant au concepteur d’agir” (Bharosa, 2011, p.153). Enfin, la troisième étape consiste à schématiser l’ensemble du dispositif. Selon la terminologie utilisée par les chercheurs en *design science*, il est question de construire le design du système.
- **Opérationnalisation d’un dispositif** : l’objet de cette partie est d’assurer le passage de la dimension conceptuelle à la dimension matérielle. Autrement dit, il s’agit de mettre en oeuvre un processus d’opérationnalisation qui permet “d’implémenter des fonctionnalités dans un objet à partir d’une structure prédéfinie et pré-optimisée” (Cofer et Harding, 2006, p.120), qui se trouve être en l’occurrence le schéma du dispositif. Dans cette partie, j’appliquerai le processus en quatre étapes d’Hekmatpour (1997) : définir les objectifs de l’opérationnalisation relativement au projet de recherche, sélectionner les fonctions du dispositif à opérationnaliser, construire le dispositif, établir les mécanismes permettant de venir implémenter le dispositif dans un contexte d’usage réel.

² Termes originaux : functional requirements

³ Termes originaux : *kernel theories*

- **Évaluation du dispositif** : J'adapterai la méthode d'évaluation de dispositif établie par Venable et al. (2012) qui se déroule en sept étapes : analyse du dispositif à évaluer, choix d'une stratégie d'évaluation, choix des méthodes d'évaluation, conception de l'évaluation, réalisation des évaluations, implication de l'évaluation sur la prochaine itération.

Ainsi, notre méthodologie suppose d'adopter une démarche intrinsèquement itérative dans laquelle un premier dispositif est conçu avant d'être évalué. S'il est suffisamment pertinent empiriquement et rigoureux scientifiquement, alors la recherche s'arrête. Dans le cas contraire, la question de recherche est précisée et un nouveau cycle de recherche démarre. Cette nouvelle itération apporte un nouveau dispositif qui est à son tour évaluée, et ainsi de suite.

3.2 Plan de lecture de la thèse

Cette thèse est structurée en deux parties représentant au total sept chapitres.

La première partie permet de faire un *état de l'art* lié à ce travail de recherche et d'exposer les concepts méthodologiques que je solliciterai tout au long de ma réflexion. Dans cette partie, le suivi de mon raisonnement sera facilité par des encarts contenant des conclusions intermédiaires.

- Le chapitre 1 décrit un état de l'art des concepts qui permettent de comprendre l'usage des données ouvertes. En ce sens, par niveau d'analyse, je mobiliserai les concepts de plateforme, d'usage, de données ouvertes et d'écosystème *open data*. Ce chapitre présentera aussi le gap dans la littérature que cette thèse viendra combler.
- Le chapitre 2 expose les concepts méthodologiques issus du *design science* que j'importerai et qui seront mobilisés dans cette thèse en sciences de gestion.
- Le chapitre 3 présente la méthodologie inspirée du *design science* que je mets en place pour répondre à notre problématique ainsi que le design de la recherche.

La deuxième partie est consacrée à ma démarche empirique. Elle embrasse l'ensemble des étapes qui m'a permis d'aboutir à un dispositif en système d'information permettant de répondre à la problématique de recherche.

- Le chapitre 4 présente la démarche exploratoire qui m'a permis de construire un modèle heuristique permettant de comprendre l'usage des données ouvertes et le contexte dans lequel ce phénomène se met en place. Il expose dans un premier temps un modèle préliminaire articulant les concepts présentés dans la revue de littérature. Ensuite, il présente le périmètre des cas d'usage observés et des observations participantes effectuées avant de présenter un modèle affiné de l'usage des données ouvertes.
- Le chapitre 5 permet de présenter l'étude de cas comparative des métropoles de Grenoble et de Lyon. Tout d'abord, j'exposerai le contexte dans lequel se situent Grenoble et Lyon vis-à-vis de l'usage des données ouvertes. Ensuite, je mettrai en avant la démarche de collecte de données qualitatives par entretiens auprès des utilisateurs des données ouvertes. Par la suite, je présenterai le codage thématique de ces entretiens et le résultat de l'analyse subséquente. Ces résultats exposent les attentes des utilisateurs vis-à-vis du dispositif.
- Le chapitre 6 explique l'approche de construction d'un dispositif venant répondre à ma problématique de recherche. Dans un premier temps, je décrirai la démarche qui m'a permis de concevoir ce dispositif. Ensuite, j'aborderai les différentes étapes qui m'ont permis d'opérationnaliser le dispositif.
- Le chapitre 7 explique *in fine* la démarche d'évaluation du dispositif. Je commencerai par analyser le dispositif construit. Ensuite, je décrirai le détail du dispositif d'évaluation mis en œuvre. Enfin, j'introduirai les analyses des résultats de l'évaluation.

La discussion générale met en évidence les contributions théoriques et méthodologiques de ce travail de recherche et souligne ses principales limites. Aussi, met-elle en évidence la pertinence et les limites du dispositif construit telles que perçues par ses futurs utilisateurs. Il constitue la contribution managériale de cette thèse.

La conclusion générale est un récapitulatif des résultats itératifs des procédés de caractérisation de problème, de construction d'un dispositif et d'évaluation de celui-ci. Elle constitue une réponse synthétique à la question de la recherche.

La table 0.1 synthétise les chapitres de la thèse avec l'étape du design de la recherche correspondante, la question de recherche, l'approche analytique et les pattern de recherche choisis pour atteindre les objectifs du chapitre. La table 0.1 présente plus en détail la structure de la thèse. Les figures 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 et 0.8 présentent la structure de chacun des chapitres, à savoir les données en entrée du chapitre (*input*), les différents traitements effectués à l'aide de l'approche analytique et des *pattern* de recherche choisis ainsi que les données de sortie issues de ce traitement (*output*).

Table 0.1 - Aperçu de l'ensemble des chapitres de la thèse

Étapes du design de la recherche	Questions de recherche	Chapitre	Approche analytique/Pattern de recherche	Objectifs	Bases théoriques	Matériau empirique
0. Intérêt de la recherche	Question de recherche : comment résoudre les tensions paradoxales inhérentes aux écosystèmes open data qui font obstacle à la construction de nouveaux usages à partir des données ouvertes ?	Introduction générale	Descriptive	Présenter le phénomène empirique étudié, la pertinence managériale, théorique et méthodologique, la question de départ, l'objectif général de la recherche ; Présenter le design de la recherche		Une sélection de 252 articles basés sur les 4 construits choisis
1. Identification d'un problème empirique et du résultat souhaité	Sous-question de recherche 1 : quel est l'état des lieux concernant la construction des usages ?	Chapitre 1 : Revue de littérature des différents courants irriguant la recherche sur la construction de nouveaux usage à partir des données ouvertes	Qualitative/Développement d'un cadre conceptuel (Vaishnavi et Kuelcher (2015, p.115))	Définir les construits "Plateforme", "Usage", "Open data" et "Ecosystème open data" ; Identifier les associations entre les construits ; Identifier un gap théorique et méthodologique	Framework multi-niveau développé par Chesbrough and Bogers (2014) et affiné par Bogers et al. (2016) pour classer la littérature sur le concept d'open innovation (Chesbrough, 2003) selon le niveau d'analyse ; Théories du paradoxe et concept de tension paradoxale (K. Smith et al., 2017)	
2. Identification des caractéristiques d'un dispositif répondant au problème	Sous-question de recherche 2 : quelles sont les situations problématiques qui freinent la construction des usages issus des données ouvertes ?	Chapitre 3 : Démarche d'exploration hybride pour comprendre les difficultés des parties prenantes de la construction de nouveaux usages à partir des données ouvertes	Qualitative et abductive : démarche d'exploration hybride (Thiéart et al., 2014, p. 93) utilisant la méthode des cas (Hlady-Rispal, 2002) / Analyse d'un système complexe (Vaishnavi et Kuelcher, 2015, p. 107)	Proposer un modèle heuristique pour comprendre l'usage des données ouvertes et les 'situations problématiques' rencontrées par les praticiens.	Etude de 4 cas d'usage des données ouvertes grâce à des observations participantes ; 10 entretiens semi-directifs	
Idem	Sous-question de recherche 2 : quelles sont les situations problématiques qui freinent la construction des usages issus des données ouvertes ?	Chapitre 4 : Etude des cas grenoblois et lyonnais permettant de comprendre les attentes des 'opérateurs de l'open data' vis à vis d'une intervention en	Qualitative et abductive : méthodologie d'étude de cas basé sur les travaux d'Eisenhardt (1989) et de Glaser et Strauss (1967) pour le développement de	Etablir la liste des besoins fonctionnels auxquels un 'dispositif en Systèmes d'Information' doit répondre pour résoudre les situations problématiques qui freinent	Eisenhardt (1989) pour la méthodologie d'analyse et Glaser et Strauss pour le développement de théories ancrées	Etude de cas de l'usage des données au sein de la métropole Grenoble-Alpes et de la métropole du Grand Lyon ; 21 entretiens semi-directifs, 31 documents analysés, 23

		Systèmes d'Informations permettant de faciliter l'usage de la donnée	théorie ancrée / Identification et formulation de l'espace problème (Vaishnavi et Kuelcher, 2015, p.86-88)	l'usage des données ouvertes ?		observations participantes
3. Conception du dispositif	Sous-question de recherche 3 : quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ?	Chapitre 5 : Design d'un 'dispositif en Systèmes d'Information' permettant de résoudre les situations problématiques qui freinent l'usage des données ouvertes	Qualitative et inductive : revue de littérature spécifique à partir des besoins fonctionnels identifiés / Conception élégante (Vaishnavi et Kuelcher, 2015, p.132-133)	Etablir une liste d'éléments fonctionnels qui permettent de résoudre les situations problématiques (dits proposition de design); Identifier dans la littérature les cadres théoriques (dits **kernel theories**) qui expliquent en quoi les éléments fonctionnels permettent de résoudre les situations problématique (dits **principes de design**); Construire un modèle conceptuel dessine l'architecture du 'dispositif en Systèmes d'Information', intègre les éléments fonctionnels et explique les fonctions qu'il assurera dans un contexte réel (dit **design du système**)	Théorie de la cognition distribuée (Hutchins, 1995) ; Théorie de la structuration appliqué aux systèmes d'information (Orlikowski, 1992 ; DeSanctis et Poole, 1994) ; Théorie de l'interaction homme-données (Mortier et al., 2013)	Une sélection d'articles correspondant aux 12 besoins fonctionnels
4. Opérationnalisation du dispositif	Sous-question de recherche 3 : quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ?	Chapitre 6 : Prototypage d'un dispositif en systèmes d'information permettant de résoudre les situations problématiques qui freinent l'usage de données ouverte	Qualitative et déductive : processus d'instanciation (Ince et Hekmatpour, 1987) / Pattern de prototypage développés par Vaishnavi et Kuelcher (2015, p.119-156).	Présenter l'objet, instance du dispositif conçu		Une sélection d'outils logiciels permettant de construire l'objet
5. Evaluation du dispositif	Sous-question de recherche 3 : quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ?	Chapitre 7 : Evaluations visant à estimer la capacité du 'dispositif' à résoudre les situations problématiques qui freinent l'usage des données ouvertes	Qualitative et inductive	Evaluer la pertinence empirique et la rigueur scientifique du 'dispositif en Système d'Information' ; Identifier les manquements relatifs à ces deux aspects et à améliorer dans la prochaine itération		14 focus group inspirés des méthodes utilisées en ergonomie informatique

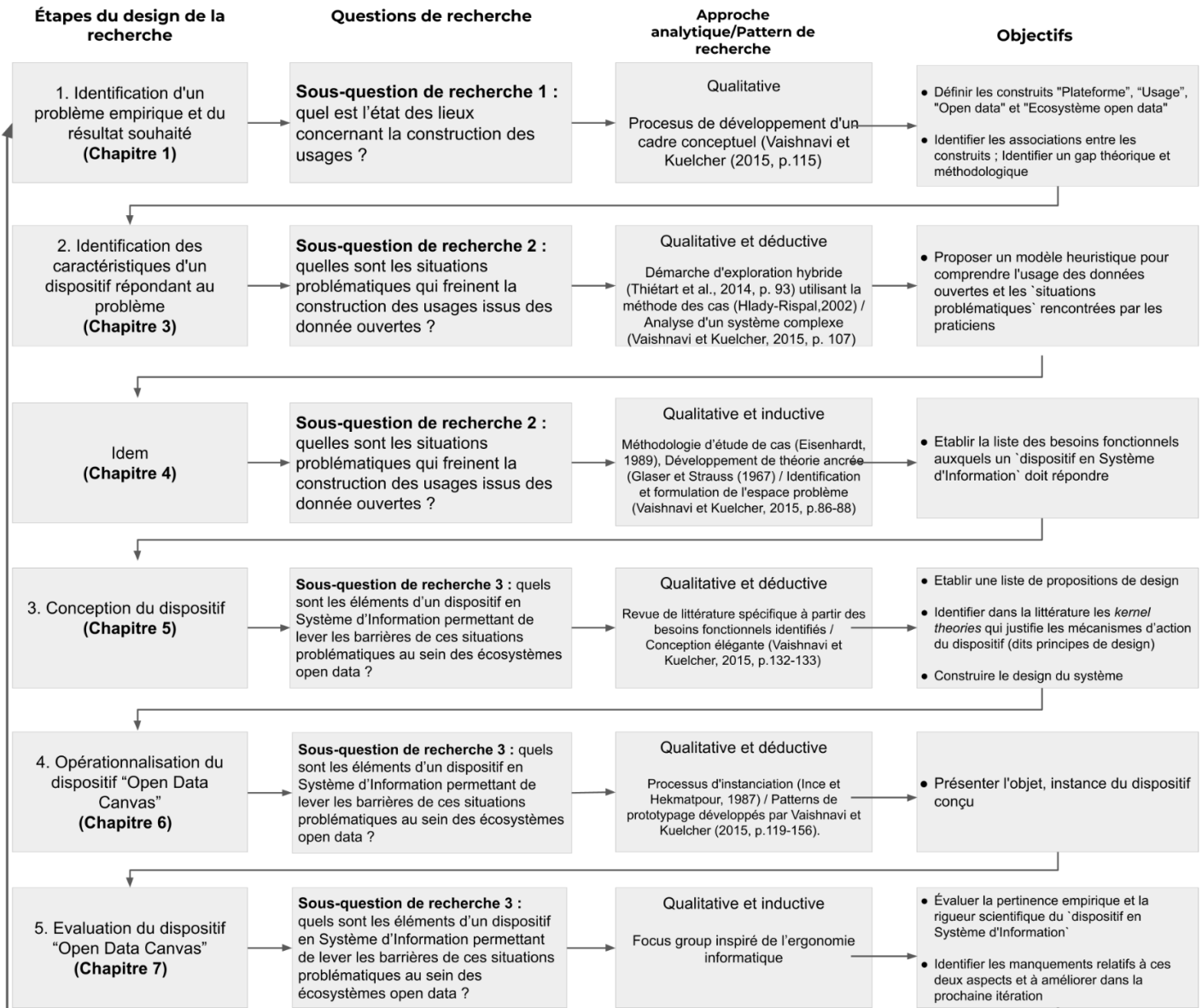


Figure 0.2 - Structure de la thèse

Cette thèse suit les différentes phases d'une recherche en *design science* décrites par Carlsson (2010). Il préconise de débiter par l'identification d'un problème empirique et des solutions désirées par les praticiens (chapitre 2,4,5), pour ensuite concevoir un dispositif et les connaissances (chapitre 6,7) et finalement, d'évaluer le dispositif auprès des praticiens (chapitre 8), ces trois phases s'affinant itérativement.

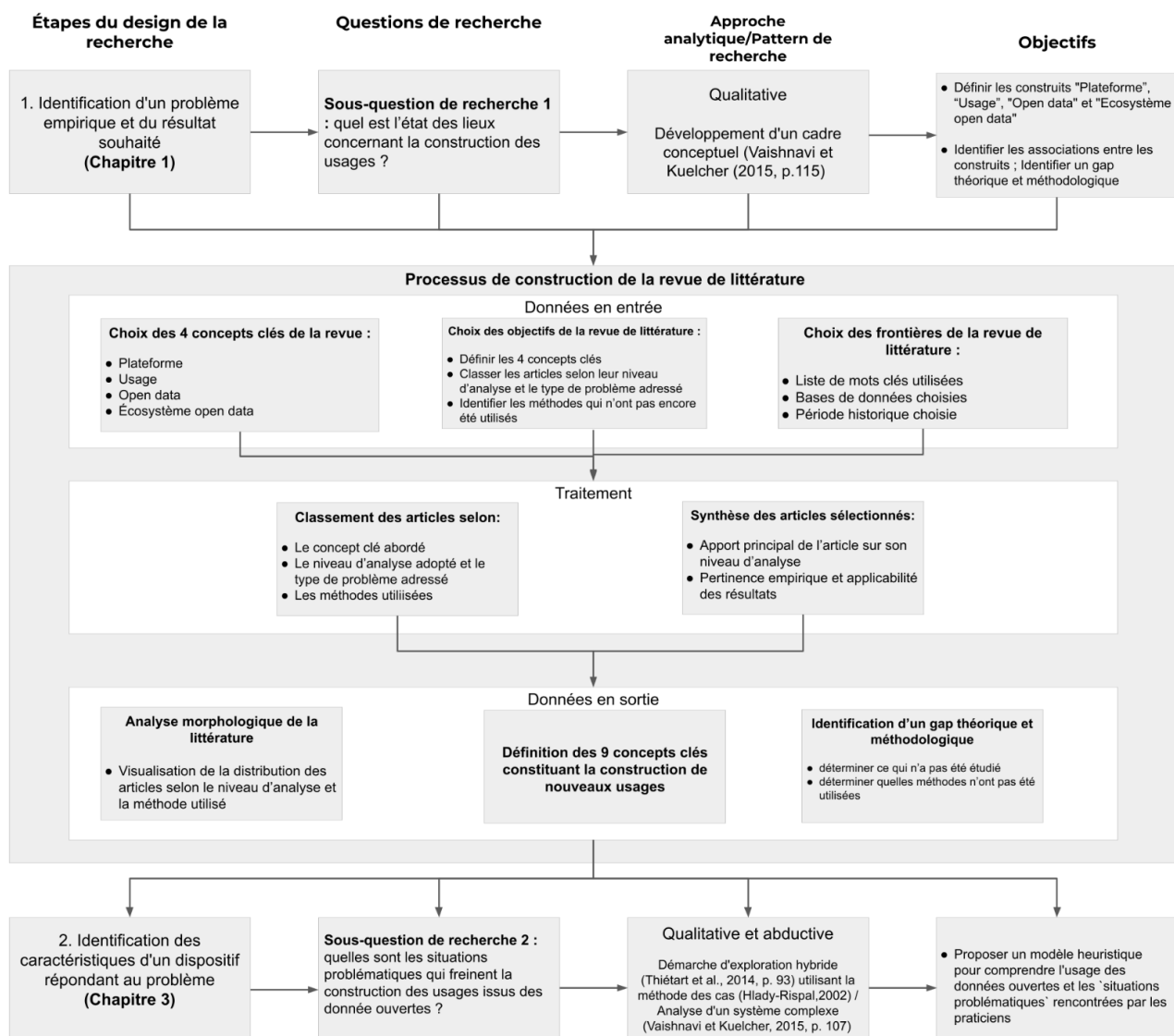


Figure 0.3 - Structure du chapitre 1

Le chapitre 1 suivra le pattern de développement d'un cadre conceptuel permettant de classer les différents champs de littérature connexe au sujet de cette thèse (Vaishnavi et Kuelcher, 2015). Ce cadre conceptuel permettra de classer et de donner forme à plus de 250 articles de recherche. *In fine*, ce cadre conceptuel facilitera l'identification d'un gap théorique et méthodologique dans lequel cette thèse s'engouffera.

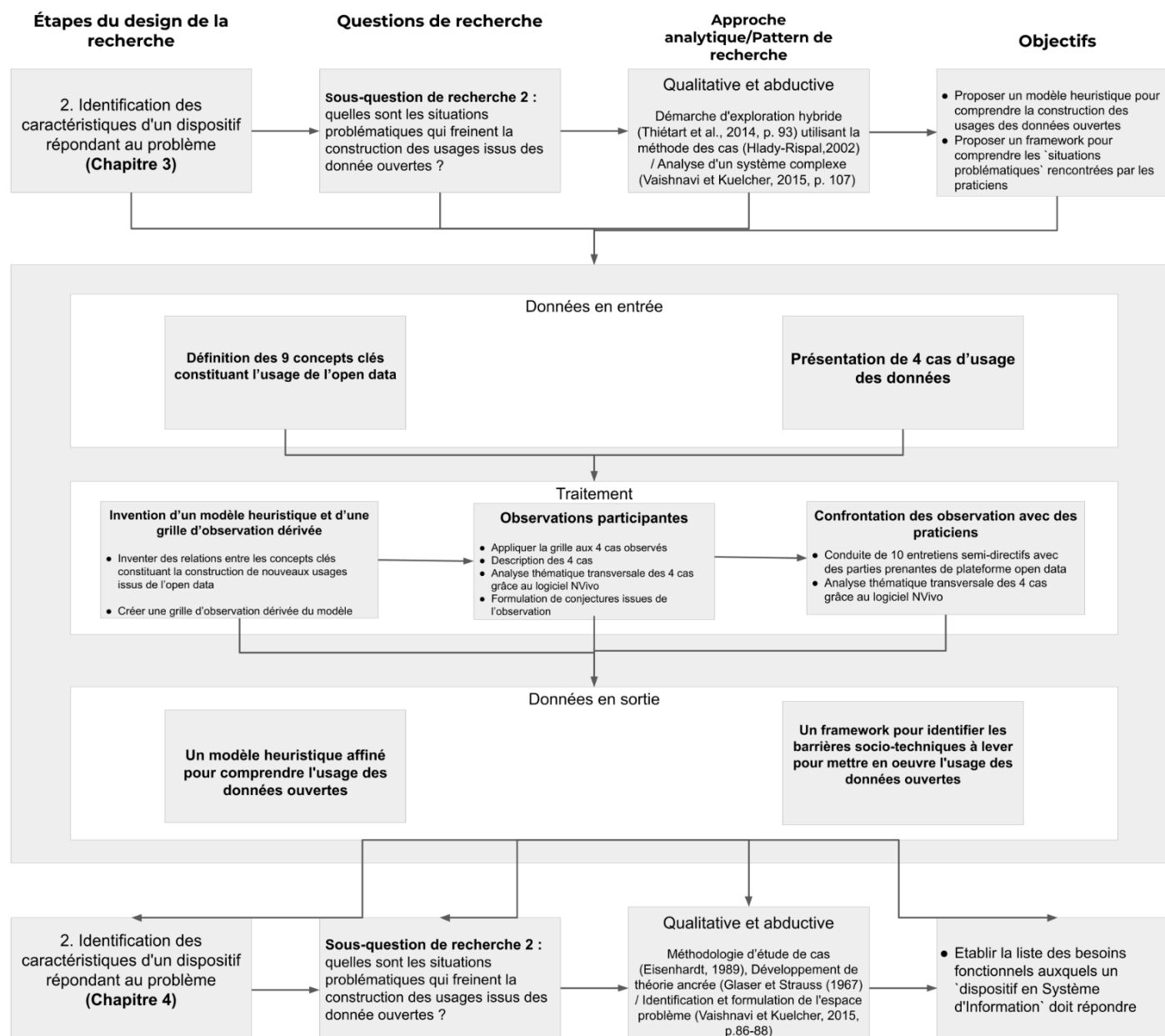


Figure 0.4 - Structure du chapitre 3

Le troisième chapitre suivra une démarche d'exploration hybride utilisant la méthode des cas pour développer un modèle heuristique de compréhension de la construction des usages issus des données ouvertes. Quatre cas de projet de développement des usages seront analysés, des entretiens semi-directif menés, leurs résultats triangulés avec la littérature aboutissant à un modèle heuristique affiné et un framework d'identification des barrières bloquant le processus.

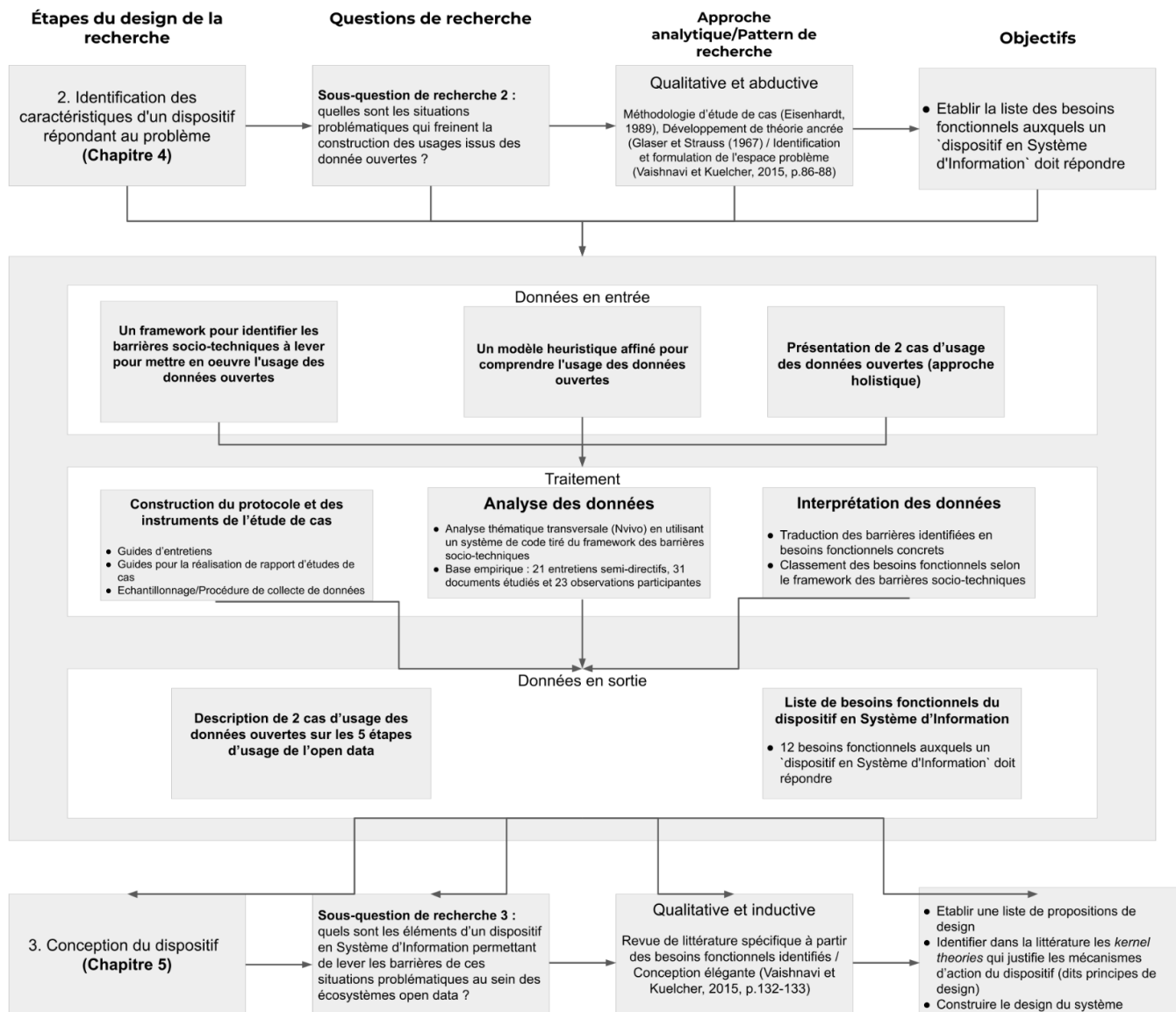


Figure 0.5 - Structure du chapitre 4

Le quatrième chapitre vise à identifier les caractéristiques d'un dispositif répondant au problème empirique. J'y mène une étude de cas dans l'écosystème open data de la métropole Grenoble-Alpes et de la métropole du Grand Lyon en utilisant la méthodologie d'Eisenhardt et les travaux de Glaser et Strauss (1967) sur le développement des théories ancrées.

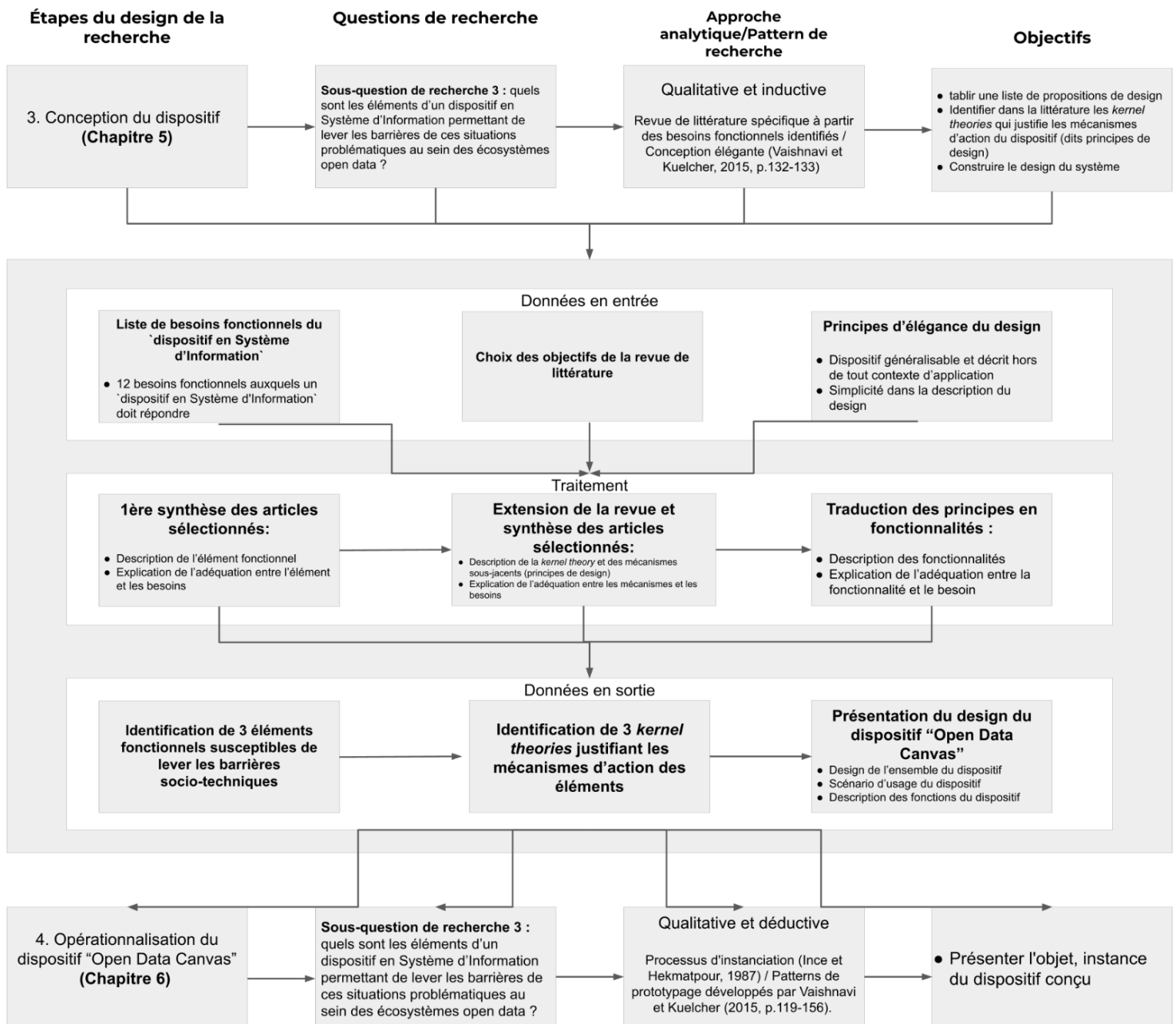


Figure 0.6 - Structure du chapitre 5

Ce cinquième chapitre traitera de la conception du dispositif et suivra pour ce faire, le pattern de conception élégante (Vaishnavi et Kuelcher, 2015). On trouvera comme données en entrées la liste des besoins fonctionnels, auxquels, lors de la phase de traitement, je ferai correspondre des éléments fonctionnels. Ces éléments fonctionnels seront finalement combinés.

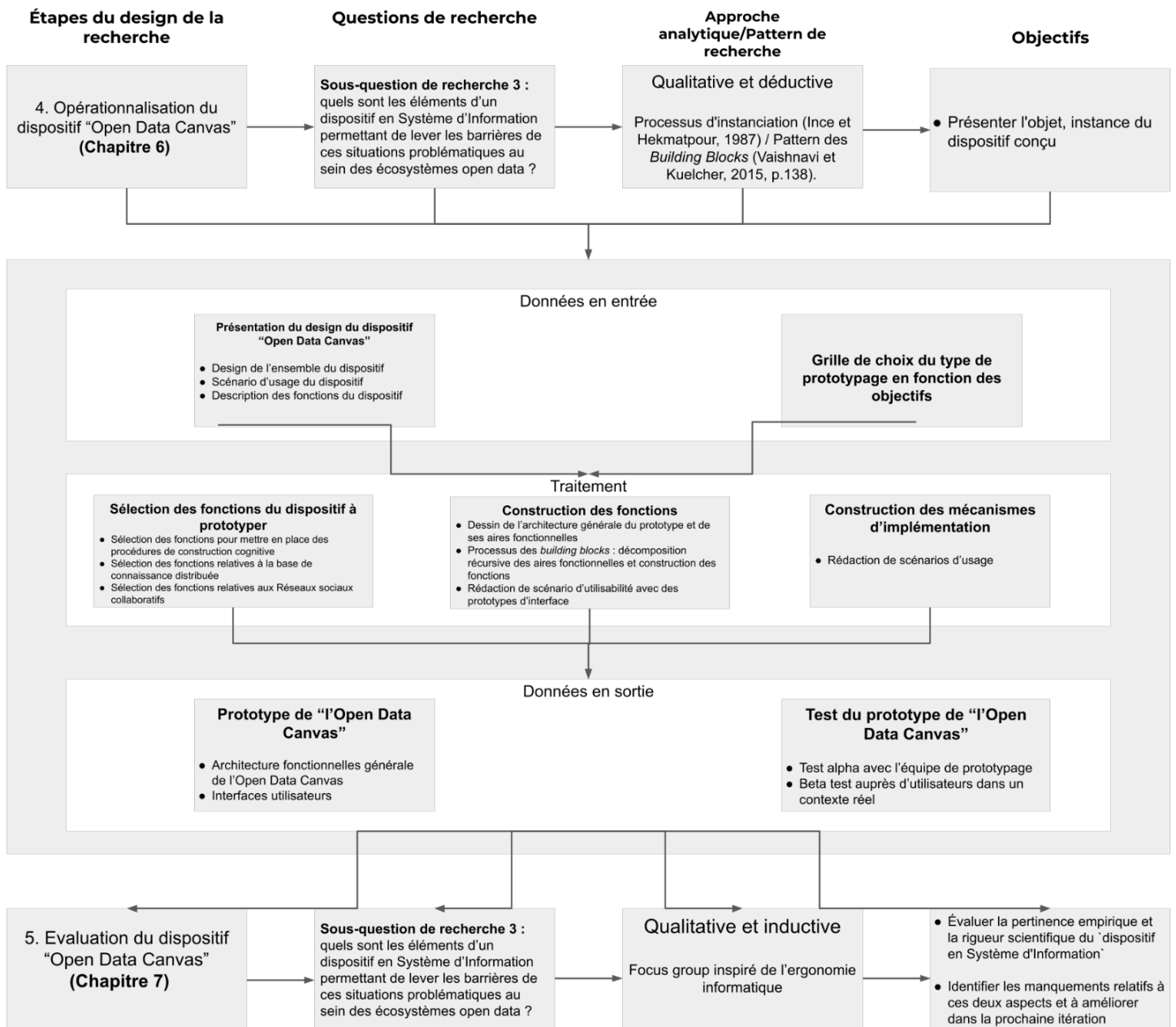


Figure 0.7 - Structure du chapitre 6

Ce sixième chapitre transformera le design conçu au chapitre 6 en prototype. Ce dernier sera construit en suivant le pattern des *building blocks* (Vaishnavi et Kuelcher, 2015) qui me permettra de sélectionner graduellement les fonctions à incorporer dans le prototype. Le prototype sera présenté sous la forme de cartographies d'architecture.

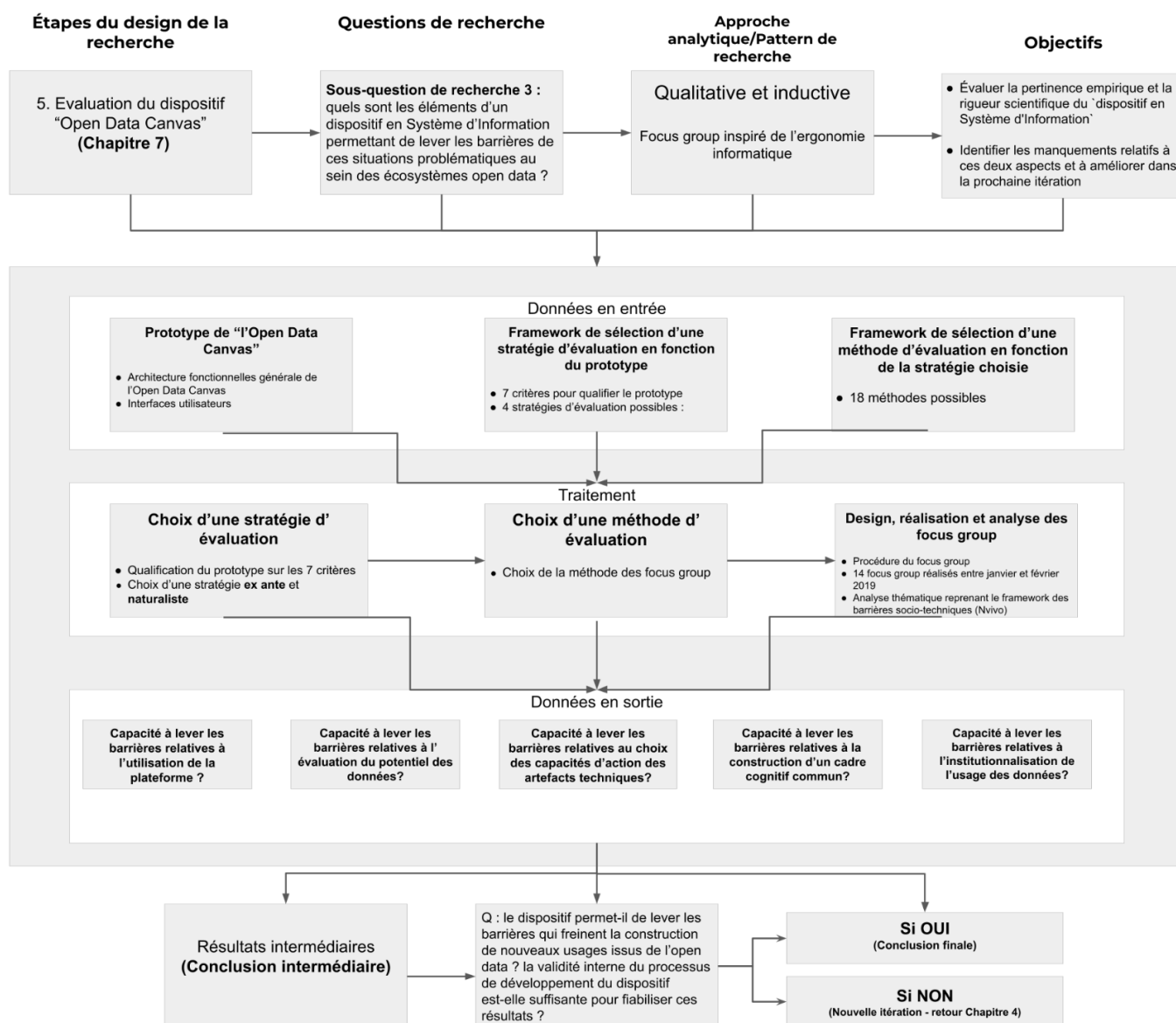


Figure 0.8 - Structure du chapitre 7

Ce septième chapitre sera dédié à l'évaluation du dispositif Open Data Canvas. Cinq focus group seront présentés. Ils évalueront dans quelle mesure il est en capacité de lever chacune des barrières identifiées au chapitre 3. Ces résultats seront interprétés de manière à pouvoir affiner la conception du dispositif et son prototype. De manière idéale, la recherche s'arrête quand le dispositif lève l'ensemble des barrières identifiées.

Chapitre 1

Revue de la littérature

Chapitre 1 - Revue de la littérature

Dans ce chapitre, je réaliserai une revue de littérature visant à fournir une image critique des différents courants qui irriguent la recherche sur la construction de nouveaux usages issus des données ouvertes. Cette revue de littérature sera guidée par ma première sous-question de recherche :

Quel est l'état des lieux concernant la construction des usages ? Comment les différents construits qui définissent la construction de nouveaux usages à partir des données ouvertes sont-ils mis en relation dans la littérature scientifique ?

Pour mettre à jour la mise en relation des différents construits dans la littérature, mon raisonnement réutilisera le *pattern* de *framework development* développé par Vaishnavi et Kuelcher (2015, p. 115). Ce *pattern* invite le chercheur à effectuer une analyse morphologique (Zwicky, 1967) des informations contenues dans les champs de littérature scientifique. Plus concrètement, il s'agit d'utiliser un *framework* visuel permettant de classer et de créer des formes avec les contenus scientifiques sur lesquels ma revue de littérature s'appuie. Au-delà de dresser un panorama de la recherche existante sur mon sujet de thèse, ce *framework* me permettra d'identifier (morphologiquement) la *gap* de recherche que ma thèse visera à combler.

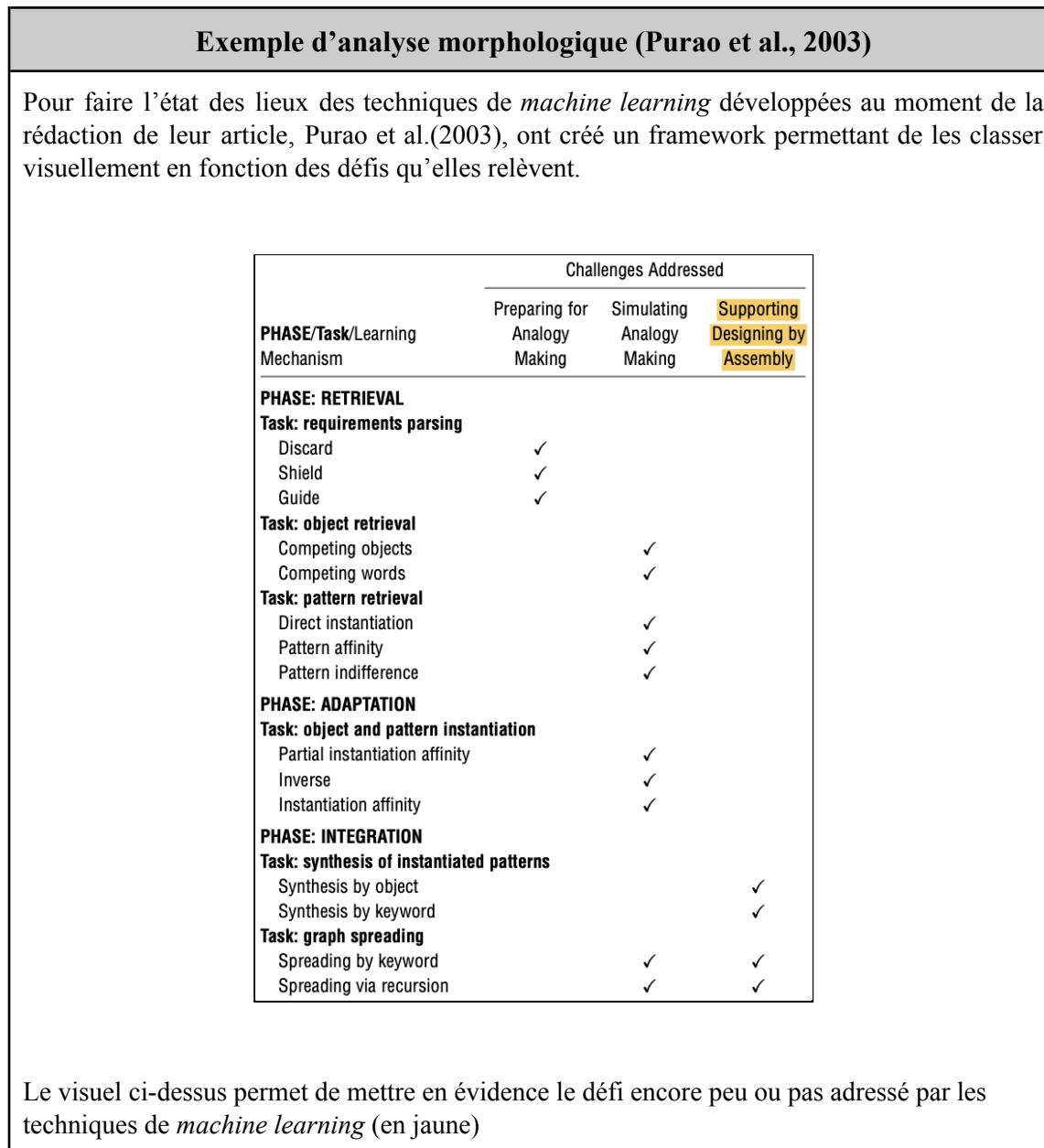


Figure 1.1 : Exemple d'analyse morphologique

J'utiliserai ici le framework multi-niveau développé par Chesbrough and Bogers (2014) et affiné par Bogers et al. (2017) pour classer la littérature sur le concept d'innovation ouverte¹(Chesbrough, 2003) selon le niveau d'analyse. Il distingue cinq unités d'analyse permettant de classer les déterminants, processus et résultats d'un phénomène :

¹ Termes originaux : open innovation

- le niveau intra-organisationnel
- le niveau organisationnel
- le niveau extra-organisationnel
- le niveau inter-organisationnel
- le niveau industriel, territorial et sociétal

À ces cinq niveaux, j’en ajouterai un sixième, qui permet de prendre en compte les travaux qui superposent plusieurs niveaux d’analyse, et un septième pour lister les travaux reconnaissant une tension paradoxale (K. Smith et al., 2017) entre les différents niveaux d’analyse.

Définition : Tensions paradoxales : “les différents niveaux d’analyse (qui) se confrontent les uns aux autres de manière contradictoire tout en étant nécessairement interdépendants” (K. Smith et al., 2017, p.305).

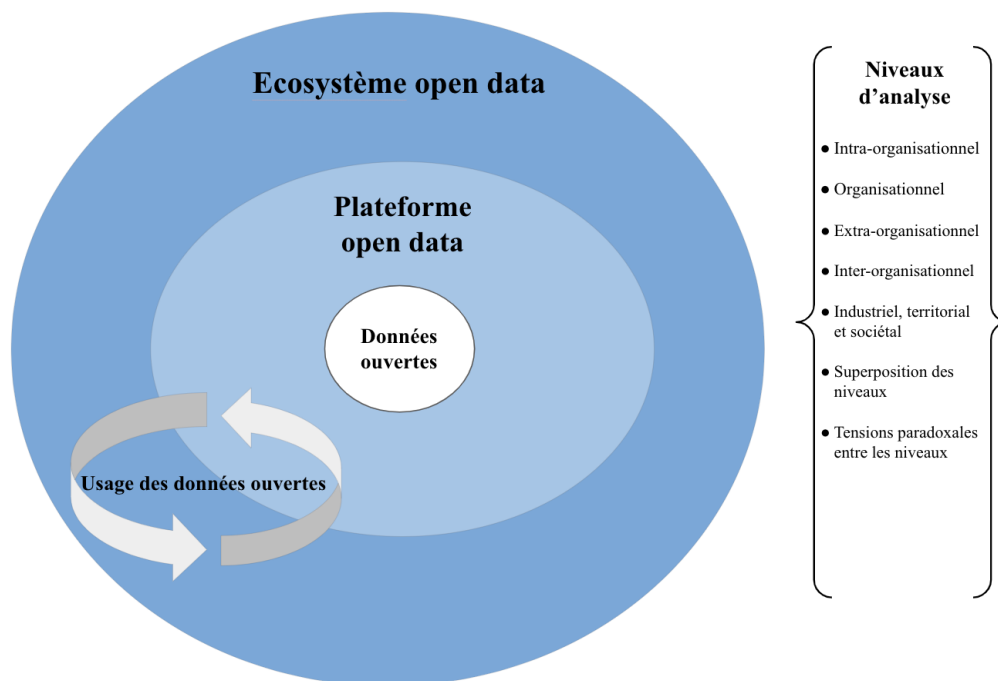


Figure 1.2 : Les relations entre les termes clés de la revue de littérature

1.1 Usages des données ouvertes

1.1.1 Les usages des données ouvertes au niveau intra-organisationnel

Les études ayant couvert le niveau d'analyse intra-organisationnel s'intéressent principalement au travail concret que les individus doivent effectuer pour effectivement faire usage des données. Si la majorité des auteurs note que les données ouvertes force à l'émergence de nouvelles pratiques intéressantes, (Henriksen, 2015), ils soulignent néanmoins le caractère difficile de ces nouvelles pratiques. Les auteurs pointent principalement deux types de difficultés, le premier relatif à l'incompatibilité de la substance même des données par rapport à l'usage, l'autre concernant les relations et collaborations entre usagers

1.1.1.1 Les difficultés d'usage relatives au contenu et au format des données

Le terme "données ouvertes" peut laisser penser qu'il s'agit d'un ensemble cohérent et défini de données dont le contenu et le format seraient uniformes. Or, cet *a priori* est vite démentie par la littérature qui estime que les données ouvertes constitue un volume de données hétérogènes en expansion et de qualité variable (Heise et Naumann, 2012 ; Böhm et al., 2010 ; Wirtz et al., 2016 ; Mainka et al., 2015).

D'un point de vue technique, l'hétérogénéité empêche tout traitement technique et de l'autre, le volume en expansion fait obstacle à toute prise en main manuelle. L'utilisateur se trouve dans une double contrainte à laquelle il ne peut apporter de solution, situation qui aboutit dans la plupart des cas à l'abandon des données ouvertes (Heise et Naumann, 2012).

Si le traitement technique des données est difficile, la substance même des données rend le travail d'usage des données difficile. Les auteurs pointent que les jeux de données sont rarement accompagnés de documents qui permettraient de donner du sens aux données (Böhm et al., 2010).

Sans ces éléments, l'utilisateur ne peut ni placer les données dans un contexte cohérent ni en extraire des informations utiles. Ironiquement, les auteurs parlent alors d'une ouverture en demi-teinte puisque les données qui sont ouvertes ne sont pas interprétables. La transparence que promet les données est en demi-teinte pour les usagers qui récupèrent des données qu'ils ne peuvent comprendre (Wirtz et al., 2016 ; Janssen, 2011).

Face à ces difficultés des auteurs ont proposé des artefacts pour évaluer le niveau d'utilisabilité des données ouvertes (Solar et al., 2012), des bonnes pratiques (Kuhn, 2011), certains allant jusqu'à proposer des méthodes d'ouverture allant de la recherche des données jusqu'à l'accompagnement des usagers des données (Böhm et al., 2010).

1.1.1.2 Les difficultés d'usage relatives aux relations et collaborations entre usagers

Le deuxième type de difficultés qui bloque l'usage au niveau intra-organisationnel relève des relations et collaborations entre usagers. En effet, faire usage des données suppose d'adopter une attitude collaborative de la part des acteurs (Ruijter et al., 2017), une attitude qui vise à attirer et engager de plus en plus d'acteurs (citoyens et développeurs) dans la démarche (Kassen., 2013).

Or, historiquement, les acteurs qui ont intérêt à engager d'autres acteurs dans l'usage ne possèdent ni les outils ni les compétences, ni la culture pour le faire. Ce manque de compétence résonne particulièrement chez les responsables des organismes publics en charge d'ouvrir les données, qui n'arrivent pas à solliciter les développeurs et citoyens externes à l'organisme (Ganapati et Reddick., 2012). L'usage des données ouvertes réveille, selon Tinati et al. (2012), des pratiques sociales et culturelles préexistantes qui sont en opposition avec une posture collaborative et d'incitation à l'engagement. Parmi elles, Relly and Sabharwal (2009) note que la tendance au silotage des informations vient freiner les pratiques compatibles avec les données ouvertes comme la transmission libre d'informations entre les départements d'une organisation.

La clé de l'usage au niveau intra-organisationnel réside dans la capacité des usagers à convaincre d'autres usagers du potentiel que recèlent les données ouvertes et de l'intérêt que d'autres acteurs pourraient porter au développement de nouveaux usages (Hartog et al., 2014), mais aussi du caractère attractif de cette activité (Juell-Skielse et al., 2014). On retrouve ici une des justifications de l'organisation de nombreux hackathons ou autres concours d'innovations tournés autour des données ouvertes et qui visent à attirer des usagers potentiels.

Enfin, la littérature fait état de la présence de nombreux usagers en sommeil, qui, avec un appui adapté de la part des producteurs des détenteurs des données, pourraient créer de nombreuses applications et services utiles (Mainka et al., 2015). Des citoyens anonymes se sont déjà saisis des données ouvertes pour

partager de l'information sur la conduite des organismes publics, charge aux détenteurs de données de mettre à disposition des moyens pour développer des solutions utiles (Desouza et Bhagwatwar., 2012).

1.1.2 Les usages des données ouvertes au niveau organisationnel

D'une certaine manière, le niveau d'analyse organisationnel fut le point de départ de la ruée vers l'or universitaire sur le potentiel d'usage des données ouvertes. En effet, je rappelle qu'à l'origine du mouvement d'ouverture des données se trouve l'idée selon laquelle une fois les données ouvertes, elles seraient réutilisées et transformées par les organisations, avec à la clé de la création de valeur économique (Chignard, 2013).

On trouve ce postulat pour la première fois dans un rapport de la commission européenne de 2006, intitulé *Mesure les ressources informationnelle du secteur européen*² (Dekkers et al., 2006). Celui-ci estime que l'ouverture des données et les informations publiques pourraient créer un marché estimé à 40 milliards d'euros par an. Je décrirai ici la logique construite par la commission européenne et qui a permis de supporter cette estimation (figure ci-dessous).

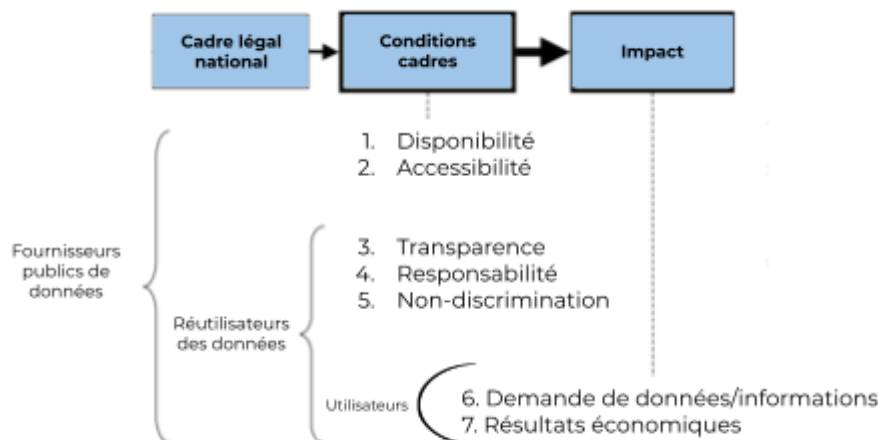


Figure 1.3 : Logigramme établissant le lien entre l'offre de données (conditions cadres), la demande de données et les résultats économiques de l'ouverture des données (impact)

² Termes originaux : *Measuring European Public Sector Information Resources*

En premier lieu, la commission européenne identifie trois rôles que les organisations impliquées dans l'ouverture des données peuvent jouer :

1. Fournisseur public de données : tout organisme public ou organisme régi par le droit public qui fournit des données à un marché.
2. Réutilisateur de données : les organismes (hors secteur public) qui fournissent de l'information au marché.
3. Utilisateur : toute personne ou organisation utilisant les informations fournies par les réutilisateurs.

Puis, elle articule les relations entre ces rôles en suivant une logique constituée de cinq conditions cadres, et de deux types d'impacts économiques. Les premières correspondent à des principes généraux s'appliquant à l'offre de données et d'informations tandis que les seconds ont trait aux conséquences économiques de l'application des principes généraux (description ci-dessous).

Conditions cadres	1. Si les données ne sont pas disponibles pour les réutilisateurs alors il n'y aura pas de marché.
	2. Si les données disponibles ne sont pas accessibles , alors il ne peut y avoir de réutilisations.
	3. Si les conditions de réutilisation des données ne sont pas transparentes , alors l'accessibilité perd tout son sens et il ne pourra y avoir de marché lié à la réutilisation des données.
	4. Les producteurs de données sont responsables et doivent rendre des comptes vis-à-vis des trois premiers points.
	5. Les quatre premiers points doivent être respectés et ce, quel que soit le réutilisateur.
Impacts économiques	6. L'égalité et l'équité face aux données/informations publiques stimulera la demande actuelle de données et d'informations .

	7. Cette demande se traduira éventuellement par des résultats économiques directs (plus de chiffre d'affaire pour les réutilisateurs) et indirectes (plus d'activités commerciales basées sur la réutilisation des informations publiques)
--	---

Table 1.1 : Les conditions et impact économique des données ouvertes

En substance, la commission européenne estime que si les organisations productrices de données ouvertes font en sorte que les données soient disponibles (point 1), accessibles (point 2), transparentes (point 3), que les fournisseurs de données s'engagent à rendre des comptes (point 4) et à fournir un accès égal et équitable à tous (point 5 et point 6), alors les données seront réutilisées. Autrement dit, la commission européenne soutient que les acteurs publics qui respecteront ces cinq critères verront leurs données réutilisées naturellement.

Si ce raisonnement a participé à l'avènement du mythe selon lequel les données ouvertes donnent lieu à un surcroît de croissance économique, il a été largement mis de côté par la littérature scientifique. Si le niveau organisationnel est effectivement considéré comme le plus pertinent pour comprendre les évolutions des usages mais aussi des démarches d'ouverture des données prises dans leur ensemble, la construction naturelle de nouveaux usages du fait de la simple ouverture a été fortement démentie. Au niveau organisationnel, la construction de nouveaux usages issus des données ouvertes rencontre de nombreux obstacles, qui s'expriment en termes économiques et structurels.

1.1.2.1 Les obstacles économiques à la construction de nouveaux usages issus des données ouvertes : la question du business model

Comme évoqué précédemment, l'usage des données ouvertes est une activité foncièrement collaborative et qui nécessite un engagement de la part de plusieurs acteurs dont les organisations productrices de données. Or, cet engagement se trouve grippé par l'absence de modèle économique qui permettrait de soutenir le travail nécessaire à l'ouverture. Selon Wang et Lo. (2016), sans modèle économique viable pour les producteurs, il est peu probable que les démarches d'ouverture des données ait un impact significatif en termes d'usage. Il est aussi peu probable dans cette configuration qu'ils s'investissent dans

les relations avec les développeurs ou citoyens qui pourraient mettre en place ces usages (Desouza et Bhagwatwar., 2012).

Il en est de même pour les organisations qui, de l'autre côté, pourraient faire usage des données. Sans modèle économique, elles ne peuvent soutenir les activités incluses dans l'usage des données. C'est avec cet enjeu en tête que de nombreux auteurs ont cherché à développer des modèles économiques soutenables pour des organisations (Magalhaes et al., 2014, Janssen et Zuiderwijk, 2011). Je note que ces modèles économiques partagent l'idée de voir les organisations intégrées dans un *écosystème open data*, ce qui suppose une structure organisationnelle spécifique.

1.1.2.2 Les obstacles structurels à la construction de nouveaux usages issus des données ouvertes

Comme le soulignent justement Léveillé et Timms (2015), pour qu'une organisation fasse usage des données et s'inscrive à son avantage dans un écosystème, il s'agit pour elle de développer un cadre structurel qui la mette en capacité de créer, contrôler, compléter des données fiables. Ce cadre permet non seulement d'assurer la qualité du management organisationnel de la donnée mais aussi de faciliter la construction de relations de confiance avec les autres acteurs de l'écosystème. Autrement dit, les auteurs considèrent que pour faire usage des données ouvertes, la seule ouverture des données n'est pas suffisante. Il faut également apporter des changements significatifs à la structure de l'organisation qui s'en charge (O'Hara, 2012).

On trouve dans la littérature des exemples d'organisations ayant réussi à faire de l'usage des données ouvertes un levier pour leur développement. Parmi elles, l'exemple des organisations non gouvernementales (ONG) en charge du développement économique international (Linders et al., 2013) ou des agrégateurs d'informations (Janssen et Zuiderwijk, 2014 ; Vert et Vasiliu, 2015).

1.1.3 Les usages des données ouvertes au niveau extra-organisationnel

Antoine Picon rappelait dans son ouvrage phare *Smart cities: a spatialised intelligence* (2015) que les données ne sont pas :

“des entités éthérées qui tels des esprits voleraient au-dessus des rues et des bâtiments. Elles ont nécessairement une dimension locale, inséparables de leur environnement physique et des habitants qui y vivent” (Picon, 2015, p. 105).

Par là, il émettait l'idée que les données et leurs usages étaient systématiquement ancrés dans des communautés données, qui dépasse le cadre strict d'une organisation. Il faisait alors appel aux travaux de Townsend (2013) qui étudient la communauté des *civics hackers*, pionniers dans l'usage des données ouvertes.

Avec les travaux d'Antoine Picon, on retrouve un champ entier de littérature qui aborde l'usage des données depuis un niveau d'analyse communautaire, qui fait partie, dans notre typologie, du niveau extra-organisationnel. Ces travaux se sont concentrés sur les freins à la prise en main de l'usage des données ouvertes à ce niveau et sur les moyens de l'améliorer.

1.1.3.1 Sur les freins à la prise en main de l'usage par les communautés

Pour Zuiderwijk et Janssen (2013), les activités des communautés *open data* ne sont pas coordonnées. Tant et si bien que les producteurs de données sont rarement au fait de ce qui est fait avec les données qu'ils ont ouvertes. Les producteurs et réutilisateurs agissent comme deux agents isolés au sein d'une communauté, ce qui aboutit à des usages peu fiables.

Dans un autre registre, Chan (2013) met en avant les stratégies gouvernementales qui ne prennent pas correctement appui sur ces agents extérieurs que sont les communautés. Il enjoint même les agences publiques à "réaliser que les initiatives de données ouvertes qui marchent sont celles basées sur l'usage communautaire des données ouvertes et non simplement sur la création d'un portail de données ouvertes." (p.1891). Les communautés devraient être au cœur des initiatives de données ouvertes, celles qui en spécifiant aux organisations les données pour lesquelles des opportunités d'usage se présentent (Power et al., 2015).

Du fait qu'elles détiennent une connaissance sur le quotidien de ses membres, les communautés sont considérées comme plus à même que les producteurs de données de connaître les besoins auxquels les usages peuvent répondre et donc demander les données appropriées (Radl et al., 2013). Cependant, les auteurs constatent qu'elles n'arrivent pas à se faire entendre (Bertot et al., 2014).

1.1.3.2 Sur les moyens d'améliorer la prise en main de l'usage par les communautés

Bertot et al. (2014) suggèrent aux communautés d'identifier les bonnes pratiques à appliquer dans l'usage des données mais aussi de collaborer avec d'autres communautés pour avoir plus de poids face aux organisations productrices de données ouvertes. C'est du rapprochement des communautés avec les organisations productrices de données que dépend la qualité des usages (Parycek et al., 2014).

D'un autre côté, les communautés ne pourront faire usage des données seules. Elles doivent s'inscrire dans des réseaux de valeurs établis (Lindman et al., 2014) qui leur apporteront les ressources dont elles ont besoin pour développer les usages. Ces réseaux de valeurs auxquelles elles participent doivent équilibrer le rapport de force entre l'offre et la demande de données mais aussi maximiser l'échanges d'informations entre communautés existantes (Ohemeng et Ofofu-Adarkwa., 2015). Ils sont également un vecteur de structuration qui profite aussi bien aux communautés qu'aux organisations productrices.

1.1.4 Les usages des données ouvertes au niveau inter-organisationnel

Les analyses qu'on peut trouver à ce niveau empruntent toutes la même position : la formation d'alliance, de réseau ou d'écosystème de différentes organisations n'est pas suffisante en l'état pour stimuler la construction des usages (Susha et al., 2015a ; Kalampokis et al., 2013 ; Yang et al., 2015 ; Janssen et Zuiderwijk, 2014 ; Shum et al., 2015 ; Cowan et al., 2015 ; Veeckman et Van Der Graaf., 2015 ; Susha et al., 2015b.). Aussi s'appuient-ils sur cette position pour faire des suggestions aux organisations intéressées. Ces suggestions sont de deux ordres : d'un côté on trouve les auteurs qui proposent des solutions sociales (i.e, qui regardent aux relations entre les agents des organisations) et de l'autre les auteurs qui proposent des solutions techniques.

1.1.4.1 Les solutions sociales proposées pour favoriser la formation d'alliance, de réseau ou d'écosystème

On retrouve à nouveau la distinction entre les organisations qui produisent les données et celles qui les réutilisent dans la mesure où les solutions proposées insistent sur l'implication sur l'un ou l'autre type d'organisations.

Des auteurs, à l'image de Susha et al. (2015a., 2015b) recommandent grandement aux organisations productrices de données ouvertes (organismes publics) d'améliorer le support formel qu'elles offrent aux organisations réutilisatrices (les entreprises). Ils constatent que le support est aujourd'hui mis en place de manière informelle par des agents du service particulièrement acquis à la cause des données ouvertes. Cette configuration a selon les auteurs atteint ses limites. De fait, ils proposent des grilles d'évaluation de support qui insisteraient non pas sur le nombre de jeux de données ouverts mais sur les nouveaux processus ou résultats issus des collaborations entre organismes publics et entreprises.

D'autres auteurs recommandent aux organisations réutilisatrices de s'impliquer plus largement dans les relations inter-organisationnelles et de passer d'une position passive où ils observent simplement les jeux de données ouvertes à une position active dans laquelle ils dressent la liste des jeux de données dont ils aimeraient faire usage (Yang et al., 2015). Ils participeraient alors à la création de chaînes de valeurs unissant les organisations publiques et privées mais aussi d'autres parties prenantes des données ouvertes comme les journalistes, les étudiants, les journalistes ou les militants politiques (Shum et al., 2015 ; Veeckman et Van Der Graaf (2015)).

Enfin, je note une troisième alternative proposée par Janssen et Zuiderwijk (2014) : l'instauration d'organisations info-médiatrices dont le but est d'être un support entre les organisations productrices et les organisations réutilisatrices de données ouvertes.

1.1.4.2 Les solutions techniques proposées pour favoriser la formation d'alliance, de réseau ou d'écosystème

On trouve également des auteurs qui insistent sur le rôle que les solutions techniques peuvent avoir pour favoriser la formation de continuum inter-organisationnels. Du fait que les données ouvertes sont dispersées sur une multitude de dispositifs techniques appartenant à des organisations, les solutions techniques sont pertinentes en la matière.

À l'image de Kalampokis et al. (2013), il est question d'adopter le paradigme des données liées³ qui se réfèrent aux données publiées sur le Web de telle manière à ce qu'elles soient lisibles par les machines, "que leur contenu soit explicitement défini, lié à d'autres jeux de données externes et pouvant être lié à

³ Termes originaux : *Linked data*

d'autres jeux de données externes" (Kalampokis et al., 2013, p.101). Ce paradigme se traduit par des solutions techniques qui permettent de récupérer les données dispersées, de les lier et de les agréger pour ensuite proposer des mécanismes préparant leur usage (en effectuant par exemple, des analyses statistiques préliminaires).

Cowan et al. (2015) précisent que ces solutions sont essentielles pour développer des processus de collaborations entre organisations. Le secteur privé serait selon eux le plus à même de développer et de maintenir de telles solutions techniques, dont ils précisent les contours : architecture technique, modes d'opérations et modèle de gouvernance envisageables.

1.1.5 Les usages des données ouvertes au niveau industriel, territorial et sociétal

Les nouveaux usages issus de l'ouverture des données sont corrélés dans la littérature avec l'émergence de nouvelles formes d'industries, de territoires et de sociétés. Les données ouvertes seraient devenues une forme de peau digitale⁴ à la surface de laquelle émergent de nouveaux services ou applications (Rabari et Storper, 2015).

Je note dès à présent que cette peau digitale est perçue de manière ambiguë. D'un côté, elle permettrait d'offrir des solutions permettant de résoudre les problèmes politiques, économiques, environnementaux ou sociaux à ce niveau d'analyse. Certains auteurs rappellent que l'utilisation et l'analyse des données peut donner des pistes pour améliorer la performance des services de transport, de santé et d'éducation (Cury et al., 2015). De l'autre, elle peut également aggraver des déséquilibres économiques, sociaux et environnementaux. Face à cette ambiguïté, Rabari et Storper (2015) invitent au développement de recherches en sciences sociales sur le sujet, invitation à laquelle je me rallie dans cette thèse.

Je classe les études correspondant à ce niveau d'analyse dans une des deux positions mentionnées ci-dessus.

⁴ Termes originaux : *digital skin*

1.1.5.1 Les usages des données ouvertes pour améliorer la performance des Etats et des services public

Selon Whitmore (2014) les données ouvertes sont un moyen pour mieux anticiper le déclenchement de nouveaux conflits internationaux armés. Les données relatives à la dépense militaire peuvent informer les gouvernements sur les zones dans lesquelles investir en préparation d'une guerre probable. Dans la même vague, Lofi et Krestel (2012) montrent comment l'usage des données ouvertes peut faciliter la transparence et l'engagement des citoyens dans la co-production des services publics. Schkabatur (2012) montre comment un usage massif des données ouvertes peut améliorer le fonctionnement des systèmes judiciaires tout en élargissant l'accès au droit.

1.1.5.2 Les usages des données ouvertes et le risque d'aggravation des déséquilibres sociétaux

Höchtel et Reichstädter (2011) sont les premiers qui à ma connaissance ont mentionné les risques que comportent les usages des données ouvertes en termes d'aggravation des déséquilibres sociétaux. Ils considèrent que si les données ouvertes ne sont pas de qualité suffisante ou si les réutilisateurs manquent d'informations sur leur contexte de production, les usages construits peuvent être néfastes en amenant par exemple à des revendications politiques non-justifiées. Aussi, pointent-ils que tous les organismes publics ne sont pas égaux face aux données qu'ils publient et que ce sont les plus munies financièrement qui bénéficieront plus des usages.

On retrouve alors dans la littérature l'idée selon laquelle il faut que les États prêtent une attention particulière aux bonnes pratiques en matière d'ouverture et de management des données (Lofi et Krestel., 2012), tout en ayant conscience que la non-ouverture des données peut avoir un coût d'opportunité assez élevé avec des innovations, des initiatives entrepreneuriales et des business plans jamais réalisés (Lakomaa et Kallberg., 2013).

1.1.6 Les usages des données ouvertes vus depuis la superposition de plusieurs niveaux d'analyse

Pour Courmont (2016), les usages des données ouvertes ne peuvent être abordés par un seul niveau d'analyse dans le mesure où ils opèrent "une médiation entre tous les collectifs qui s'y rattachent" (p.52). Plus précisément les usages créent un enchevêtrement d'acteurs, de dispositifs techniques et de conventions de mesures dans lesquels les niveaux d'analyse se superposent. Il prend pour exemple l'usage des données du référentiel géographique dans les métropoles. Ces données sont créées par le Système d'Information Géographique (SIG). Leur usage premier, à savoir donner une image fiable du territoire, se situe au niveau territorial. Or, il fait aussi usage de ses données à un niveau organisationnel puisque les services de la métropole (voiries, eau, cadastre par exemple) s'en saisissent dans la gestion de leurs affaires courantes (voir figure ci-dessous).

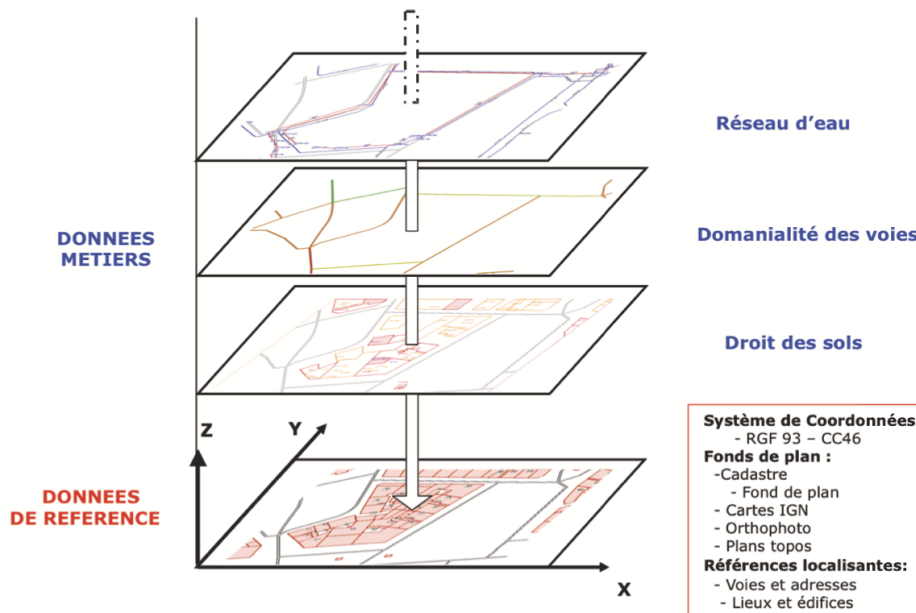


Figure 1.4 : Visualisation des différents niveaux d'analyse de l'usage des données du référentiel géographique (Courmont, 2016, p.81)

Cette approche multi-niveaux a été déclinée de plusieurs manières. Par exemple, pour Khayyat et Bannister (2017), l'usage des données ouvertes peut se comprendre selon un modèle tiré de la théorie de

l'activité (Engeström, 1999). Ce modèle intègre plusieurs facteurs positionnés sur plusieurs niveaux d'analyse. Au niveau intra-organisationnel, l'usage des données ouvertes est influencé par la motivation et le niveau de sensibilisation (*level of awareness*) des agents. Au niveau organisationnel, l'usage est déterminé en partie par la division du travail, les barrières et les règles en application dans les organisations. Au niveau inter-organisationnel, l'usage est déterminé par l'environnement direct des organisations. Enfin, au niveau industriel, territorial et sociétal, l'usage est déterminé par les plateformes *open data* disponibles.

Worthy (2015) s'est quant à lui servi de l'approche multi-niveaux pour analyser des pattern d'usage des données ouvertes. Il conclut que l'usage des données ouvertes implique nécessairement, pour l'acteur qui l'entreprend, d'osciller entre des niveaux d'analyse microscopiques (intra-organisationnel et organisationnel) et macroscopiques (extra-organisationnel, inter-organisationnel et sociétal). Il devra par exemple s'informer de faits microscopiques pour comprendre les besoins auxquels les usages répondront tout en les faisant correspondre avec des données ouvertes couvrant tout le périmètre national et produites par des organismes nationaux. L'idéal étant d'arriver à intégrer tous les outils utilisés par obtenir une vision d'ensemble.

Cette intégration que les auteurs appellent de leurs vœux leur permet de convoquer le concept de gouvernance, qui permet de faire fonctionner conjointement les intérêts des citoyens à l'échelle locale avec des infrastructures territoriales et des politique nationales (Balena et al., 2013).

Pour Yang et al. (2015), la gouvernance est un champ de recherche porteur dans la mesure où il permet d'envisager les usagers sous quatre perspectives simultanément : perspective technologique, organisationnelle, législative, politique et environnementale.

1.1.7 Les usages des données ouvertes vus depuis les tensions paradoxales entre plusieurs niveaux d'analyse

Relativement à ce positionnement méthodologique, je n'ai identifié qu'un seul auteur l'ayant emprunté. En effet, seul Dominique Boullier dans son célèbre ouvrage *Sociologie du numérique (2016)* a analysé les usages des données ouvertes depuis les tensions paradoxales entre plusieurs niveaux d'analyse. Il y démontre qu'on assiste à un conflit dans les formes d'usage des données, ces formes dépendant des

différentes déclinaisons du critère de circulation des données. Il en liste quatre qu'il met en scène dans une ville fictive :

- La forme d'usage inspirée de la "good old city" où les données restent dans les silos traditionnels, chaque producteur de donnée restant le propriétaire et l'utilisateur exclusif de celle-ci
- La forme d'usage inspirée de la "IBM city" où poste de centralisation capte tous les flux de données pour établir des usages sous forme de tableaux de bord qui permettront de prendre les décisions les plus rationnelles pour le fonctionnement de la ville
- La forme d'usage inspirée de la "Google city" où toutes les données produites dans l'espace urbain sont mis en relation sans présumer d'aucun usage au préalable. Cette forme renvoie à une utilisation accentuée du big data dans la gestion des villes.
- La forme d'usage inspirée de la "Wiki city" où la ville repose cette fois-ci sur la collecte des contributions et usages des citoyens. Une forme donc où la place du *crowdsourcing* est prépondérante pour prendre les décisions concernant le fonctionnement de la ville et mettre en place de nouveaux dispositifs.

Par rapport aux autres niveaux d'analyse, je note un faible volume d'études recensées. Ce petit volume d'articles et d'études ayant adopté ce niveau d'analyse indique un gap dans la littérature que ma thèse se propose de combler.

1.2 Plateformes *open data*

1.2.1 Les plateformes *open data* au niveau intra-organisationnel

Au niveau intra-organisationnel, je peux dégager chez les auteurs une certaine ligne de pensée partagée : les interactions entre les plateformes *open data* et les entités intra-organisationnelles (individu, groupe et business unit) constituent un angle mort pour les éditeurs de logiciels qui produisent ces plateformes et pour les chercheurs.

Ces interactions, qui intègrent toutes “les opérations concrètes et les compétences sur lesquelles reposent l’entretien et l’alimentation quotidien (commune aux) infrastructures informationnelles” (Denis et Pontille, 2012, p. 11), sont aujourd’hui encore invisibles. Le rôle de la littérature est alors de “faire remonter à la surface le travail plus ou moins visible qui participe à la codification des assemblages sociotechniques composant le monde” (Star, 1999, p.132).

Cette compréhension serait d’autant plus importante qu’elles permettrait de décrire ce que Gardey (2001) nomme les «conditions de félicité» de la performativité des données. Ce qui, dans leur existence la plus ordinaire, assure la capacité des données ouvertes à agir. En ce sens, les plateformes et les données ouvertes partagent une même capacité à agir sur leur environnement.

Malgré la relative invisibilité du phénomène, certains auteurs ont posé les premiers jalons conceptuels pour saisir le travail d’alimentation et d’entretien des plateformes *open data* mis en place à un niveau intra-organisationnel. Ce travail est nommé *open data work* ou travail des données ouvertes, et se décline en quatre catégories :

- Le travail de production des données ouvertes
- La travail de maintenance des données ouvertes
- Le travail de coordination autour de la plateforme *open data*
- Le travail de coopération autour de la plateforme *open data*

1.2.1.1 Le travail de production des données ouvertes et d'alimentation des plateformes open data

L'open data work ou travail des données est un concept qui trouve ses origines dans la sociologie des sciences et des techniques et qui renvoie à la réalité des coulisses des infrastructures informationnelles à l'image des plateformes *open data*. Ces infrastructures organisationnelles, à l'image des plateformes *open data* supporte les activités des administrations, des entreprises ou plus particulièrement des relations de services. Elles représentent des univers peuplés de documents écrits mais aussi de données.

Le concept du travail des données ouvertes s'oppose à la vision cybernétique d'une information qui serait naturelle et détachée de tout contexte humain. Il met en avant le fait que la pratique de la production et de l'utilisation de données ouvertes sont nécessairement coincées dans des tensions entre un cadre technologique et réglementaire qui impose un certain formalisme et le travail réel des agents qui peut s'écarter du cadre pour répondre à ses propres exigences. On notera ici des échos évidents avec la théorie de la structuration technologique d'Orlikowski. Par là, la sociologie des sciences et des techniques sous-entend que les informations et les données ouvertes sont plus le fruit d'un travail et d'une coordination entre des intérêts divergents que d'un phénomène naturel.

Ce travail de production se caractérise tout d'abord par l'importance de l'engagement corporel et cognitif des agents qui y participe. En effet, contrairement à ce que l'absence de reconnaissance de ce travail de petite main laisse à penser, d'alimentation des plateformes *open data* est véritablement un travail intellectuel. Denis et Pontille (2012) le comparent au travail de mise à jour d'une base de données de biomédecine. Ce travail suppose notamment que l'agent en charge arrive à identifier les instruments ayant produit les données et documents afin d'en évaluer la pertinence ou alors d'en affiner le traitement. De réelles connaissances des mécanismes de production de la donnée sont ici nécessaires. Ils citent également tous les travaux qui consistent à recomposer l'ordonnancement de certains documents ou de certains champs de jeux de données ou ceux qui à l'inverse suppose de rassembler des données de différentes sources pour ensuite faciliter leur traitement.

Ensuite, ce travail de production suppose de pouvoir effectuer des opérations cognitives complexes dans lesquelles l'agent doit prendre des décisions. Pour Denis et Pontille (2012, p.13) :

“la production d’information, la codification, la saisie de données reposent sur une série de décisions plus ou moins grandes, plus ou moins collectives qui assure au jour le jour la qualité des infrastructures informationnelles”.

Enfin, le travail de production accentue fortement sur une nouvelle définition de la notion de matérialité. La conception classique de matérialité, explique Ingold (2007) s’accroche à un découpage entre d’un côté la matière et de l’autre les informations et les données qui s’inscrivent dans cette matière. Or, en considérant pragmatiquement le travail d’alimentation des plateformes *open data*, cette distinction s’effondre. En effet, si nous prenons l’exemple des informations financières d’une entreprise, leur mise à disposition sous la forme d’un rapport ou sous la forme de données ouvertes reviendra à disposer de deux données différentes. Le concept d’open data work implique que les données dépendent de l’instrument matériel qui les produit et inversement, que la capacité des instruments matériels dépend fortement de la donnée dont ils disposent en entrée :

“L’utilisation de 2 dispositifs pour saisir une série de chiffres participe à la production de 2 informations distinctes qui n’ont ni la même forme, ni la même valeur, ni la même portée”

Denis et Pontille (2012, p.14)

1.2.1.2 Le travail de maintenance des données ouvertes sur les plateformes open data

La donnée ouverte est un objet dit scriptural, c’est-à-dire qu’il est composé de symbole étant de la matière vivante. Il peut être dégradé, abîmé et se détériorer au cours du temps. Pour contrer la déliquescence de cette matière, un travail de maintenance est nécessaire. Un travail qui peut être assuré par des gestionnaires de la donnée ouverte aussi nommée éditeurs (Boutet, 2001).

Edwards et al. (2011) note que cette vision dynamique de la maintenance des données contrevient à l’idée selon laquelle les plateformes et données ouvertes seraient figées une fois pour toutes. Elle prend le contre-pied de cet immobilisme et instaure un “processus éphémère de communication” entre l’agent et la plateforme. Le maintien laisse place aux échanges «informels», leur reconnaissent une utilité, et qui tiennent leur valeur dans une certaine forme de souplesse, et dans les mises à jour et autres ajustements dont elles font l’objet.

1.2.1.3 Le travail de coordination autour des plateformes open data

Le travail de coordination autour des plateformes *open data* renvoie au fait de fédérer une communauté autour d'une plateforme open data. Il ne s'agit pas de travailler spécifiquement sur le développement d'une application ou d'un logiciel à partir de données mais de s'occuper du management des compétences qui y participe.

Eckleberg et al. (2018) estiment que cette coordination autour des plateformes *open data* est un facteur critique dans la mesure où il assure l'éventuelle analyse et réutilisation des données. Sans la conjonction des compétences d'experts métiers, de designers de bases de données et de développeurs web, les données sur les plateformes open data seront de mauvaise qualité ou difficilement accessibles. Or, notent-ils, il y a là un gap théorique difficile à surmonter. Zhu et Freeman (2018) abondent dans ce sens lorsqu'ils analysent les plateformes *open data* américaines et montrent leur faible capacité à faciliter l'engagement, et la participation croisée de plusieurs acteurs.

1.2.1.4 Le travail de coopération autour des plateformes open data

Le travail de coopération autour des plateformes *open data* consiste à contribuer directement aux objets au cœur des plateformes : les données ouvertes. C'est un travail qui repose sur des formes de contribution que Licoppe, Proulx et Cudicio (2010) définissent comme "des transactions fondées sur l'échange de messages isolables et districts (information nom, numéro de version de la donnée) orientée vers le service, l'entraide interpersonnelle. Pour les plateformes, ce travail de coopération porte sur l'amélioration de la qualité des données (Couture, 2012).

Pour Edwards (2010), ce travail n'est pas sans causer des frictions. Il prend pour exemple les données climatiques mondiales qui nécessitent un grand nombre de capteurs de données répartis sur de nombreuses zones géographiques :

"Cette coordination implique une longue chaîne d'opérations, comprenant l'observation, l'enregistrement, la collecte, la transmission, le contrôle de la qualité, le rapprochement, le stockage, le catalogage et l'accès. Chaque maillon de cette chaîne représente une interface d'information soumise à des frictions de données. Chaque point où des données sont déplacées ou transformées représente une opportunité de perte ou de corruption de données" Edwards, (2010, p. 84).

1.2.2 Les plateformes *open data* au niveau organisationnel

En passant du niveau intra-organisationnel au niveau organisationnel, les auteurs font évoluer leurs centres d'intérêts pour s'intéresser à deux autres caractéristiques des plateformes *open data* : leur architecture technique et leur gouvernance.

1.2.2.1 L'architecture technique des plateformes *open data*

Définition : **L'architecture d'une plateforme** : se définit comme le schéma conceptuel décrivant les composants stables et homogènes de la plateforme, les modules hétérogènes qui gravitent autour de ces composants, et les règles de design qui régissent l'articulation entre ces deux éléments (Tiwana *et al.*, 2010).

L'architecture d'une plateforme possède des fonctions primaires qui constituent le coeur de la valeur d'une plateforme ainsi que des fonctions secondaires (dites "ilités") dont la valeur varie en fonction du cycle de vie de la plateforme : durabilité, maintenabilité, flexibilité, etc.

Dans son travail de doctorat, Courmont (2016) décrypte la mise en place d'une plateforme *open data* au sein de l'entreprise JC Decaux et note qu'elle a été pensée comme une réponse à des problématiques techniques.

*“Ne diffusant pas leurs données, ils étaient l'objet de tentatives de récupération de leurs données sur leur site internet par des robots informatiques. Ces opérations leur posaient deux problèmes : elles mettaient en péril leur infrastructure technique, et elles ne garantissaient pas la qualité des données récupérées. Pour éviter l'utilisation « sauvage » de leurs données et en conserver la maîtrise, l'entreprise JC Decaux a par exemple fait le choix de mettre à disposition ses données de VLS sur une plateforme *open data*”*

Une solution technique pour répondre à un problème technique dans un premier temps, puis, pour favoriser l'utilisation des données ouvertes par des acteurs tiers. Branschweig *et al.* (2012), dans son analyse plus de 50 plateformes s'intéressent au niveau de standardisation, de découvrabilité et de lisibilité permis par l'architecture technique des plateformes *open data*. Ils concluent qu'au moment d'écrire l'article, les solutions techniques des plateformes n'encouragent pas l'utilisation des données et dressent un certain nombre de difficultés.

1.2.2.2 La gouvernance des plateformes *open data*

Les mécanismes de gouvernance d'une plateforme correspondent aux processus mis en place pour administrer, au sein d'une organisation les opérations des deux côtés de la plateforme et maintenir son écosystème. En effet, comme l'indique le schéma ci-dessous tiré d'Eisenmann et al. (2008), une plateforme se positionne comme un intermédiaire entre deux types d'utilisateurs : les fournisseurs de données hétérogènes d'un côté (offre) et de l'autre les bénéficiaires de ces données ouvertes (demande).

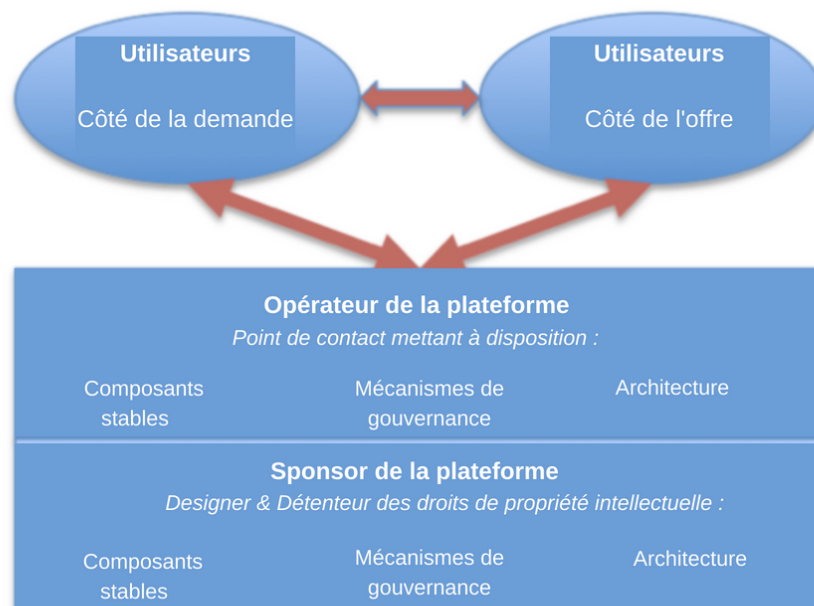


Figure 1.5 : Les éléments de l'écosystème d'une plateforme selon Eisenmann et al. (2008, p.164)

Le design de mécanismes de gouvernance permet à l'opérateur de la plateforme de stimuler l'adaptabilité et la résilience de l'écosystème en incorporant la connaissance qu'il possède des interdépendances et de la co-évolution entre les deux fournisseurs et les utilisateurs de données. (Allen and Varga, 2006 ; Benbya and McKelvey, 2006 ; Tanriverdi et al., 2010)

À terme, le design de ces mécanismes tend vers un équilibre permettant de garder suffisamment de contrôle sur le développement de l'écosystème pour assurer l'intégrité de la plateforme tout en laissant des zones de liberté aux acteurs et utilisateurs. Ceci afin qu'ils puissent innover au travers de la création de nouvelles données venant s'incorporer à la plateforme.

Pour la gouvernance des plateformes *open data*, les auteurs se concentrent sur un aspect sous-jacent : la gouvernance des données. Pour Briner et Rome (2018), la gouvernance des données est fondamentalement composée de quatre éléments dont les caractéristiques et les configurations diffèrent selon les organisations et contextes d'applications :

- **Des sponsors exécutifs** : ils ont pour fonction de faire pression, c'est-à-dire de faire usage de leur autorité pour donner du poids au programme de gouvernance de la donnée
- **Des gestionnaires de données** : ils travaillent avec le groupe de surveillance pour s'assurer que les ressources appropriées sont mises à disposition et qu'elles répondent aux besoins de gouvernance des données de l'organisation.
- **Un groupe de surveillance** : il développe la stratégie de données de l'organisation et s'assure de la conformité de la gouvernance vis-à-vis des standards, politiques et procédures de l'organisation.
- **Une équipe dédiée à la gouvernance des données** : elle facilite les démarches entreprises par les acteurs cités ci-dessus.

Cette gouvernance des données apporte une valeur ajoutée certaine aux organisations dans la mesure où elle permet (Al-Ruithe et al., 2017) :

- de clarifier les missions de chacun, d'accroître la confiance dans l'utilisation et la valeur des données ouvertes, d'établir des responsabilités, des axes stratégiques de développement et des indicateurs pour suivre le développement de la démarche.
- d'améliorer le processus de décision, de réduire les frictions opérationnelles, de protéger les besoins de parties prenantes des données ouvertes
- de former l'ensemble des parties prenantes aux approches centrées autour des données, de construire des standards, des processus fluides, de réduire les coûts et d'améliorer l'efficacité à travers la coordination et d'établir une certaine transparence dans les processus.

1.2.3 Les plateformes *open data* au niveau extra-organisationnel

On retrouve au niveau extra-organisationnel, la question de la gouvernance des plateformes. La différence réside dans le fait que les auteurs ne s'intéressent plus à la manière de manager la donnée ouverte en interne d'une organisation mais plutôt des moyens d'actions disponibles pour constituer une communauté autour d'une organisation.

Tiwana et al. (2010) identifient trois approches communautaires de la gouvernance : la gouvernance par le partage de l'autorité et des responsabilités (approche centrée autour de la répartition du pouvoir de décision), la gouvernance par l'alignement des intentions (approche centrée autour des démarches de contrôle) et la gouvernance par le partage de l'intérêt (approche centrée autour de la propriété de la plateforme).

Selon l'approche de la gouvernance par le partage de l'autorité et des responsabilités, la gouvernance des plateformes doit répondre à trois questions et par les réponses apportées diviser l'autorité et la responsabilité entre les propriétaires de la plateforme et les développeurs membres de la communauté de la plateforme :

- *Quelle est la finalité de chaque sous-ensemble de la plateforme ?*
- *Comment ce sous-ensemble devrait-il faire ce qu'il fait ?*
- *Qui contrôle les interfaces internes (i.e, les éléments qui définissent les frontières des modules) de l'écosystème ?*

Selon l'approche de la gouvernance par l'alignement des intentions, la gouvernance d'une plateforme vise à mettre en place des mécanismes formels et informels visant à encourager des comportements souhaitables chez les membres de la communauté. Le contrôle formel peut prendre deux formes :

- Le contrôle de résultat par lequel le propriétaire de la plateforme précise les critères d'évaluation des modules développés.
- Le contrôle du processus par lequel le propriétaire de la plateforme flèche des méthodes et procédures aux développeurs de module.

Enfin, selon l'approche de la gouvernance par le partage des intérêts, la gouvernance de la plateforme revient à distribuer les droits de la propriété de la plateforme et d'établir dans quelle mesure la propriété est partagée entre plusieurs membres de la communauté (Eisenmann et al., 2006).

Courmont (2016) propose une classification alternative au travers de trois schémas type en s'appuyant sur les écrits de Loveluck (2015) :

1. **Le modèle de la “captation”** dans lequel la plateforme est tenue par un acteur unique qui centralise l'ensemble des données avant de les redistribuer auprès de la communauté selon certaines modalités spécifiques.
2. **Le modèle de la “dissémination”** qui renvoie à une forte déconcentration des flux d'information, distribués en un grand nombre d'acteurs autonomes. Dans ce cas, la coordination des acteurs est assurée par des protocoles techniques. Par exemple, il existe des plateformes qui mettent en place des données de références qui permettent de relier toutes les données et tous les producteurs de données entre eux.
3. **Le modèle de “l'auto-institution”** qui repose sur des “dispositifs collaboratifs d'échange et des droits communautaires de propriété intellectuelle” (Courmont, 2016) amenant à la création d'une communauté dédiée à la gestion et préservation de la valeur unitaire de la plateforme (Parker et al., 2016). Cette valeur unitaire se définit comme la ressource centrale échangée sur la plateforme. Dans le cas des plateformes open data, il s'agit des jeux de données eux-mêmes.

L'instauration de l'un ou l'autre de ces types de gouvernance communautaire dépend pour les auteurs de deux facteurs : la licence d'utilisation de la plateforme et les protocoles techniques de la plateforme.

1.2.3.1 La licence d'utilisation de la plateforme comme facteur différenciant des communautés rassemblées autour des plateformes *open data*

Relativement à la licence d'utilisation, Pitt *et al.* (2017) distinguent principalement les plateformes sous licence propriétaire, détenues par un seul acteur dont l'accès est conditionné par le paiement d'une redevance ou par le consentement à voir ses données personnelles utilisées, et les plateformes sous licence ouverte, dont l'accès est gratuit et dont le code source peut être modifié par toutes les personnes y voyant un intérêt. Le choix d'une licence propriétaire ou ouverte est crucial dans la mesure où il influence la

manière dont les membres de la communauté échangent des ressources. Plus précisément, une plateforme sous licence propriétaire amène à une asymétrie dans la distribution des ressources qui peut inhiber l'action collective issue des échanges entre les membres de l'écosystème et donc la co-création de service. À l'inverse, une plateforme sous licence ouverte facilite la distribution de ressources, démocratise son utilisation au profit de mécanismes de co-création de services plus participatifs.

1.2.3.2 Les protocoles techniques comme facteur différenciant des communautés rassemblées autour des plateformes *open data*

En s'appuyant sur l'analyse d'une plateforme *open data*, Wikidata, que j'érigerai en modèle de gouvernance communautaire, Allison-Cassin et Scott (2018) mettent en avant que, bien que la plateforme soit sous licence ouverte, cela ne garantit aucunement la participation d'une large communauté dans l'alimentation et le maintien des données ouvertes. Il faut créer des protocoles techniques qui permettent à tout type de public de pouvoir modifier les jeux de données et proposer des améliorations techniques de la plateforme. Pfenniger et al. (2017) citent l'exemple des chercheurs dans le domaine de l'énergie qui, s'ils sont prêts à faire de leurs données des données ouvertes, sont bloqués par le manque de protocole technique pour ouvrir facilement leurs données et interagir avec les éventuels autres producteurs.

Neumaier et al. (2018) précisent cette approche en spécifiant qu'il faut instaurer des protocoles semi-automatiques permettant à une communauté d'acteurs d'améliorer la qualité des données présentes sur la plateforme. La plateforme ADEQUATE qu'ils proposent est un exemple de réponse à ce défi. Elle combine des évaluations automatiques de la qualité des données avec des dispositifs permettant à plusieurs acteurs de modifier et de republier les jeux de données avec des améliorations. Pour Lapi et al. (2012), il s'agit de sortir de la conception *stricto sensu* d'une plateforme *open data*, qui serait limitée à être un lieu de dépôt de données. Au contraire, pour développer ces communautés, les plateformes *open data* doivent permettre d'intégrer des composants et des outils propres aux communautés.

1.2.4 Les plateformes *open data* au niveau inter-organisationnel

Le niveau inter-organisationnel est un niveau très utilisé dans les recherches sur les plateformes *open data*. Elles trouvent leur point d'appui sur l'observation que les plateformes *open data* sont souvent à la base de l'émergence d'écosystèmes *open data*. Plus précisément, c'est la réutilisation de chaque

instrument et outil disponibles sur la plateforme *open data* qui amènent à la création d'un écosystème *open data* (Zuiderwijk, Janssen et Davies., 2014 ; Lusch et Nambisan, 2015).

La plateforme serait au croisement entre un écosystème et des mécanismes de co-crédation de valeur. Cette intermédiation est rendue possible par la centralisation de nombreuses ressources détenues par chaque membre de l'écosystème ou, pour reprendre les termes les termes de Lusch et Nambisan (2015, p. 157), par la maximisation de leur densité. Aussi peut-elle être cet intermédiaire dans la mesure où elle facilite la dynamisation des échanges de ressources entre les membres de l'écosystème. Autrement dit, elle augmente la liquidité des ressources. Enfin, forte d'une densité importante de ressources liquides, la plateforme constitue le lieu d'élection pour développer des mécanismes de co-crédation de services entre les membres de l'écosystème (Walravens et al., 2016).

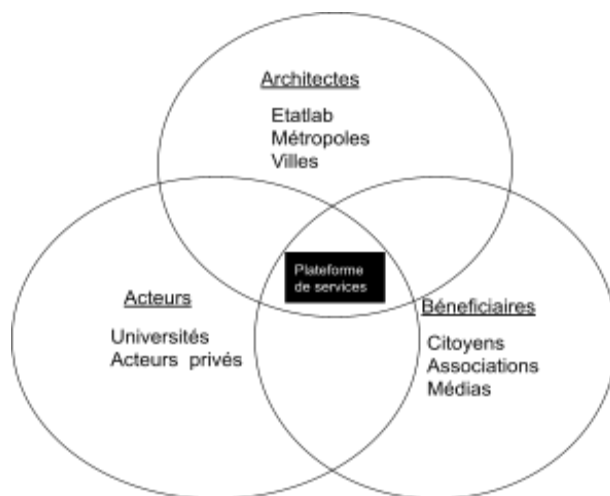


Figure 1.6 : La typologie des parties prenantes impliquées dans les plateformes *open data* (Lusch et Nambisan, 2015)

Selon Lusch et Nambisan (2015), Il existe trois types de parties prenantes investies dans les écosystèmes autour des plateformes *open data* : les architectes, les acteurs et les bénéficiaires.

Les architectes créent l'architecture de participation qui clarifie la façon dont la co-crédation de valeur se déroule et la façon dont cette valeur est répartie entre les acteurs de l'écosystème. Autrement dit, ils mettent en place une feuille de route⁵ stratégique, soit une direction à suivre par tous les acteurs impliqués dans l'écosystème de services. Aussi créent-ils les mécanismes de gouvernance permettant de piloter le

⁵ Termes originaux : road map

fonctionnement et l'évolution de l'écosystème de services vers les objectifs fixés par la feuille de route stratégique de la plateforme.

Les acteurs ont, de leur côté, les moyens techniques et les compétences pour exploiter les ressources présentes sur la plateforme dans le but de co-créeer des usages. Ils peuvent occuper trois rôles : idéateur, designer et intermédiaire. Les éditeurs apportent des connaissances sur l'écosystème, ses besoins et les usages existants. Les designers créent des services à partir de nouvelles configurations de ressources. Enfin, les intermédiaires créent des synergies entre les différentes parties de l'écosystème. Ces acteurs constituent les premiers utilisateurs des mécanismes de gouvernance et participent à leur construction dans le cas des plateformes ouvertes.

Entre ces acteurs, et avec la complexité de la participation d'autant d'acteurs hétérogènes, la question se pose du partage du contrôle et de la décision (Lee et al., 2018). Une question qui resterait sans réponse jusqu'à présent ou, selon Parker et al. (2016) dont la réponse dépasserait le simple cadre de la plateforme *open data* puisque les organisations qui participent à l'écosystème pourraient s'en trouver "inverser", en externalisant massivement tout le développement de leurs innovations.

Je trouve néanmoins des auteurs à l'image de Recupero et al. (2016) et Welle Donker et van Loenen (2017) pour proposer des modèles qui optimisent la co-création de valeurs entre les acteurs hétérogènes. Le premier propose de centrer les plateformes autour de l'engagement des usagers et de l'interopérabilité de manière à pouvoir relier un maximum de systèmes d'information autour de la plateforme. Le deuxième quant à lui préconise que les autorités publiques à l'origine des initiatives *open data* devraient laisser les futurs usagers des données ouvertes construire eux-même la plateforme. La raison en est que les usagers sont plus proches des besoins finaux auxquels la plateforme *open data* vient répondre, c'est-à-dire la construction de nouveaux usages utiles à l'écosystème. Cette proximité est selon les auteurs source d'une meilleure connaissance sur la plateforme et les composants qu'elle doit intégrer.

1.2.5 Les plateformes *open data* au niveau industriel, territorial et sociétal

Au niveau industriel, territorial et sociétal, j'ai pu observer que les auteurs centrent leurs contributions autour des moyens d'accroître l'utilisation citoyenne des plateformes *open data* et leur impact sociétal.

Force est de constater que les plateformes *open data* s'adressent en premier lieu à un public de spécialistes de la donnée (développeurs, *data scientists*) et qu'il est relativement difficile pour un citoyen de s'approprier leur contenu. Les plateformes sont conçues majoritairement pour utilisateurs "techniciens", délaissant les "non-techniciens". Elles incarnent un type de logiciel développé par les programmeurs et pour les programmeurs, limitant ainsi leur capacité à adresser de réels problèmes de société (Osagie et al., 2017).

Le faible impact sociétal des plateformes s'expliquent aussi selon certains auteurs par le fait qu'elles s'adressent en premier lieu aux fournisseurs de données et non aux utilisateurs. Par exemple, un utilisateur ne sera pas assisté dans la recherche, la modélisation ou l'analyse des données (Millette et Osein., 2016 ; Osagie et al., 2017) tandis que le fournisseur sera guidé dans chaque étape de la publication des données.

Plusieurs stratégies sont proposées dans la littérature pour adresser ce manque d'utilisation et d'impact citoyen des plateformes. Pour, Attour et Rallet (2014), ce sont les collectivités et organismes publics qui ont la charge d'organiser les flux entre citoyens et plateformes, en partant des données sur les équipements qu'ils possèdent. Ces équipements constituent une essentielle de données locales, dans lequel les citoyens peuvent facilement se reconnaître dont l'ouverture sur une plateforme va permettre aux différents acteurs d'intégrer leurs apports ou briques technologiques en vue de les valoriser. L'idée, reprise également par Aguilera et al. (2017) est d'arriver à décomplexifier l'utilisation des données ouvertes et de l'amener dans le quotidien des citoyens pour progressivement les faire passer d'une position de consommateur à une position producteur de données.

Pour d'autres, la réponse se trouve, à nouveau, dans l'utilisabilité des plateformes *open data*. Weerakkody et al. (2017) se proposent de concevoir cette utilisabilité en s'appuyant sur la théorie de la diffusion de la technologie de Rogers ; Viljoen et al. (2016) misent sur la maximisation de la capacité des plateformes à intégrer d'autres systèmes d'informations et faire rentrer les plateformes *open data* dans le paysage informationnel des citoyens. Cordasco et al. (2017) suggèrent à cet égal d'intégrer certaines formes de réseaux sociaux dans les plateformes, facilitant ainsi la formation de groupes citoyens, stimulant leur intérêt vis-à-vis des données ouvertes et les invitant à participer à la co-construction de connaissances autour des données.

1.2.6 Les plateformes *open data* vus depuis la superposition de plusieurs niveaux d'analyse

La superposition de plusieurs niveaux d'analyse n'a fait l'objet que de très peu de recherches. De ce constat émerge un paradoxe dans la mesure où une plateforme *open data* peut être considérée, dans son essence même, comme la superposition d'un système cyber-physique qui capte les données et un système socio-technique qui les utilise (Pitt et al., 2017).

Un système cyber-physique intègre des objets informatiques (ex : caméras, capteurs) et des artefacts matériels (ex : ordinateurs, serveurs) dans le but de capter, d'analyser et d'exploiter des données sur un espace physique (ex : une ville). Parmi les nombreux exemples de tels systèmes, il existe le système de localisation du Syndicat Mixte de Transports en Commun (SMTC) de la ville de Grenoble. Chaque bus et tram de la ville possède des capteurs qui permettent de les géolocaliser. Ces données sont collectées par des serveurs appartenant à la métropole grenobloise qui, grâce à elles, peut organiser et gérer les transports communs de la ville en temps réel.

Un système socio-technique se définit comme le point de rencontre entre d'un côté des individus, organisations, institutions, communautés et sociétés, et de l'autre, des machines et technologies. L'Open lab *La Péniche* à Grenoble constitue un système socio-technique à part entière dans la mesure où il rassemble les parties prenantes de l'Open data au sein de lieux physiques et virtuels et met à disposition des machines et technologies permettant de créer des services issus des données.

Parmi le peu de travaux recensés, il existe l'étude de Denis (2011) qui analyse le travail d'alimentation de plateformes par les agents d'une organisation hors de la relation de service au cours de laquelle ces agents se trouvent face aux demandes de client⁶. Il met en contraste la vision organisationnelle et intra-organisationnelle de ce travail en analysant parallèlement les structures organisationnelles conçues autour de ce travail et les activités réelles des opérateurs au quotidien.

⁶Termes originaux : back office

1.2.7 Les plateformes vues depuis les tensions paradoxales entre plusieurs niveaux d'analyse

Au sein de la littérature, je n'ai pu recenser que deux articles empruntant brièvement la position des tensions paradoxales entre plusieurs niveaux d'analyses : Villum (2014) à travers l'utilisation du concept d'*open washing* et Daneels et al. (2017) à travers de celui de plateforme autopoïétique.

Pour Villum (2014) les plateformes constituent un objet incarnant la tension entre les intérêts organisationnels et sociétaux autour des données ouvertes. D'un côté, la plateforme *open data* permet à une organisation de montrer sa transparence et de servir des intérêts communicationnels. Cependant, la mise à disposition de données organisationnelles pouvant lui être préjudiciable dans la mesure où la société pourra exercer un contrôle sur ses activités, beaucoup d'organisations limitent les conditions d'utilisation des données. De fait, ces plateformes sortent du périmètre des plateformes *open data*. Il écrit:

Nous constatons une tendance croissante à ce que l'on peut appeler un "open washing" (inspiré du "greenwashing") - c'est-à-dire des éditeurs de données qui prétendent que leurs données sont ouvertes, même si elles ne le sont pas - mais plutôt disponibles sous des conditions restrictives. Si nous - à ce moment critique de la période de formation de la société axée sur les données - ne sommes pas conscients de la différence, nous finirons par mettre nos flux de données vitales dans des infrastructures en silo construites et détenues par des sociétés internationales. Mais aussi pour faire l'éloge et soutenir le mauvais type de développement technologique non durable - Villum (2014, p.32)

Daneels et al. (2017) partent eux aussi du principe que les niveaux d'analyse doivent être combinés bien que s'opposant les uns aux autres. Ils créent le concept de plateforme autopoïétique pour incarner une plateforme qui garantirait un équilibre entre les différents niveaux d'analyse. Elle serait un lieu virtuel où chaque individu ou chaque organisation oscillerait entre des actions pour son intérêt personnel et des actions collectives pour le bien d'une communauté, d'un écosystème ou d'une société. L'équilibre dépendrait de mécanisme de coordination et de gouvernance positionné à tous les niveaux d'analyse.

1.3 L'ouverture des données

1.3.1 L'ouverture des données vue depuis le niveau intra-organisationnel

À ce niveau d'analyse, on retrouve de nombreux acteurs qui note une absence de compréhension des processus d'ouverture de données empruntés par les agents (Kuk et Davies, 2011 ; Bunakov et Jeffrey, 2013 ; Conradie et Choenni, 2012).

Si la littérature s'est peu intéressée à ces agents, c'est qu'ils sont, avec leur milieu, traditionnellement exclus des études classique des mécanismes de formation de données, d'informations ou de connaissances:

"on ne dit pas grand-chose sur la manière dont ces agents (pourraient) éventuellement trouver le soutien social, cognitif, organisationnel et institutionnel complémentaire nécessaire pour qu'ils (les acteurs non économiques) puissent apporter des données au marché"

Bunakov et Jeffrey, (2013, p.933)

Je retrouve chez Kuk et Davies (2011) le constat similaire au sein de leur étude dédiée aux mécanismes de complémentarité permettant de transformer les données ouvertes. Ils affirment que peu d'études se sont intéressées aux activités menées par ces agents dans le cadre de leur engagement pour la libre utilisation des données produites par les acteurs publiques. Un engagement qui plus motivé par les défis intellectuels qu'il sous-tend et le plaisir lié à la résolution de problèmes difficiles qu'à la création d'un logiciel ou d'un service avec l'espoir de gains économiques (Laklani et Wolf, 2003). Corollaire de cet angle mort, il n'existe dans la littérature que peu d'indicateurs pour évaluer la qualité de l'ouverture des données. Or, ces indicateurs sont cruciaux dans la mesure où les connaissances développées à partir des données ouvertes dépendent fortement de leur qualité (Behkamal et al., 2014). De plus, pour certains agents qui réutilisent les données de manière automatique, des indicateurs sur la qualité des données doivent être modélisés, d'où le besoin d'avoir des informations détaillées sur les mécanismes utilisés lors de l'ouverture des données (Bunakov et Jeffrey., 2013).

Pour les études ayant pu analyser les processus d'ouverture des données à un niveau intra-organisationnel, elles exposent la présence de nombreuses barrières chez les agents eux-mêmes, notamment des barrières d'ordre culturel. Les barrières culturelles sont inhérentes à la nature de la donnée qui nécessite d'avoir

acquis certaines compétences avant de pouvoir comprendre son intérêt et sa valeur. Contrairement au mythe des débuts de mouvement d'ouverture des données qui affirmait que tout un chacun pourrait directement utiliser les données, il faut maîtriser certaines techniques statistiques et d'autres connaissances pour pouvoir analyser les données, leur donner du sens et comprendre les implications que son utilisation peut avoir (Janssen, Charalabidis & Zuiderwijk, 2012).

Grâce à des entretiens menés auprès de développeurs, premiers utilisateurs des données ouvertes, Le Corf (2016) montre les implications de ces difficultés sur la dynamisme des projets d'ouverture et de réutilisation de données. Selon son analyse, les barrières culturelles sont à l'origine de l'absence de modèle économique autour des données ouvertes via la commercialisation d'applications développées à partir d'elles.

Pour les surmonter, Conradie et Choenni (2012) préconise une approche nuancée de la publication des données, et de mettre en œuvre des petites mesures progressives pour explorer avec les agents les différentes phases qu'ils devront mettre en œuvre.

1.3.2 L'ouverture des données vue depuis le niveau organisationnel

Au niveau de l'analyse organisationnelle, les études se concentrent sur les agences gouvernementales ou autres organisations publiques en charge de l'ouverture des données. Plus précisément, elles interrogent et proposent des stratégies d'ouverture optimale et efficace. Pour Dawes et Helbig (2010) l'efficacité réside dans les stratégies basées sur l'information visant à promouvoir l'ouverture des données auprès du public et à générer une valeur sociale et économique grâce à leur utilisation ou réutilisation éventuelle.

Or, ils notent une certaine complexité inhérente à l'utilisation publique de données publiques et donc à la promulgation de stratégies conséquentes. Meijer et Thaens (2009) posent la question plus directement : face à cette complexité, comment les organisations publiques devraient-elles s'y prendre ? Ils présentent alors un cadre conceptuel permettant d'analyser les options stratégiques qui s'offrent aux organisations. Eu égard aux options stratégiques, Dawes et al. (2004) en dégagent les principaux jalons par une étude minutieuse de celles arrivées à maturité.

Aussi l'étude de ces stratégies visent-elles à protéger les organisations publiques d'une fin malheureuse pour leurs données : la captation à des fins privées via des pratiques qui renvoient aux phénomènes des enclosures. Dans un de ses rapports, la Commission Nationale de l'Informatique et des Libertés (2017) alertait les pouvoirs publics de la captation massive de données par les opérateurs privés du numérique et des réseaux. Le projet *Quayside*, porté par une filiale de Google à Toronto, est un exemple phare de cette captation de données publiques à des fins privées. Cette expérimentation donne libre champ à Google pour mettre en place dans un quartier des capteurs, exploiter les données de ces capteurs et adapter les services du quartier en fonction. Les laboratoires pharmaceutiques adoptent déjà les mêmes pratiques avec les données biomédicales (Heller., 1998) renvoie à la même dynamique : alors que ces données relèvent pour certaines de problèmes de santé public, les laboratoires se sont appropriées ces données et la connaissances liée grâce à des droits de propriété intellectuels excessivement protecteurs. De ces captations est née une réelle inquiétude de l'administration que les données ainsi communiquées ne soient réutilisées pour des mauvaises fins (Trojette., 2013).

Subsidiairement, on trouve des études comme celle d'Akrab (2016) qui notent que les mécanismes d'ouverture des données impliquent une transformation profonde des institutions publiques au niveau organisationnel et culturel que les organisations n'ont pas encore menées à terme. La théorie des systèmes est ici souvent utilisée pour expliquer cette transformation : l'ouverture des données et la mise en place d'espaces de dialogues revient à faire passer les organisations publiques d'un mode de fonctionnement correspondant à celui d'un système fermé vers un autre correspondant à un système ouvert. Un système fermé est facile à diriger car il n'est pas influencé par des facteurs extérieurs, instables. Il est d'ailleurs possible d'utiliser des processus de contrôle rigides de ce fait. À l'inverse, un système ouvert ne peut être contrôlé de façon aussi précise du fait de l'intrusion de nouveaux éléments, dont l'intégration a des effets difficiles à prévoir. Dès lors, le passage d'un système ouvert à un système fermé nécessite de passer d'une logique de contrôle mécaniste à une perspective évolutionniste dominée par l'auto-organisation et son caractère imprévisible.

1.3.3 L'ouverture des données vue depuis le niveau extra-organisationnel

L'ouverture des données publiques a amené les organisations publiques à "mettre en scène l'existence et la vitalité de la communauté" de producteurs et réutilisateurs de données ouvertes entourant les

plateformes *open data* (Mabi, 2014, p.58), modifiant par là le cours de l'action publique. La littérature note un changement dans les interactions entre les partenaires participant à la production de services publics. D'un système de production basé sur la délégation de services publics où les pouvoirs publics de la ville exercent toute l'autorité qui leur est conférée par les élections et leur capacité à solliciter des moyens autour de l'intérêt général, on s'oriente vers un système de coproduction dans lequel les acteurs participent aussi bien à la conception du service qu'à la gouvernance de celle-ci (Bourcier, 2013).

Avec la mise en scène de communautés, les données ouvertes sont conçues comme le substrat d'une couche de partage de connaissances communes à une multitude d'acteurs qui se retrouvent dans des initiatives de co-production de services basées sur les jeux de données ouvertes. Elles impliquent les pouvoirs publics et certaines communautés numériques, définies comme "un ensemble d'acteurs divers ayant des intérêts communs et collaborant à un même but sans devoir créer de *liens forts* entre les membres" (Bourcier, 2013).

Ces initiatives peuvent être classées selon trois types de collaborations : les services d'agrégation de données par les communautés, les collaborations des communautés sur des projets publics et la production de données par les communautés (Bourcier, 2013).

1.3.3.1 Les services d'agrégation de données par les communautés

Dans ce cas, les communautés agrègent des données provenant de différentes sources pour créer de nouvelles informations. L'objectif de ces agrégations est d'encourager d'autres producteurs ou fournisseurs à ouvrir leurs données dans un format qui facilite leur interconnexion pour arriver à constituer une base de données suffisamment importante pour créer des applications.

À cet égard, Zhang, Dawes et Sarkis (2005) alertent des difficultés rencontrées dans cette configuration. Les attentes des parties prenantes étant contradictoires, de nombreux conflits émergent entre les communautés. Leur étude présente une enquête exploratoire sur les attentes divergentes et convergentes des diverses parties prenantes d'un projet d'ouverture des données et expose les avantages et les obstacles du partage de données entre organisations.

1.3.3.2 Les collaborations des communautés sur des projets publics

Les communautés sont dans ce cas sollicitées pour collecter des données et générer des informations qui sont ensuite diffusées dans différents formats de façon à permettre une amélioration continue des données et par là, les services inhérents à ces données.

Detlor et al. (2013) prennent l'exemple de certaines collectivités locales qui travaillent avec les organismes communautaires de leur juridiction pour créer un point d'accès unique en ligne afin d'offrir aux citoyens un endroit où ils peuvent accéder à un ensemble électronique complet de données, d'informations et de services municipaux et communautaires. Ils nomment ces points d'accès "portails municipaux communautaires" (PMC) et les définissent comme des "sites web publics destinés à la fourniture d'informations, de services et de ressources concernant une municipalité ou une région administrative locale particulière.

1.3.3.3 La production de données par les communautés

Dans ce dernier cas, les communautés collectent les données mais produisent aussi des données publiques les concernant. Grâce à ces données, les communautés viennent en aide au service public qui utilise ces données pour adapter les services publics au plus près des besoins des communautés.

Kassen (2013) montre comment les données ouvertes peuvent révolutionner les mécanismes administratifs traditionnels descendant⁷ où les contributions de la société civile sont nulles. Le concept de données ouvertes réalisé au niveau local peut permettre aux collectivités d'exploiter la sagesse collective des communautés locales, leurs connaissances et leurs visions des défis locaux pour qu'elles puissent réagir et répondre aux besoins des citoyens d'une manière plus productive et plus rentable. Srinivasan (2012) prend l'exemple de données spatiales que les communautés humaines locales peuvent collecter pour créer un atlas consolidé et dynamique des zones humides du bassin du Tso Kar. Un atlas qui peut aider les pouvoirs publics à prendre des décisions informées et précises concernant les activités de développement d'une zone géographique.

⁷ Termes originaux : top-down

In fine, la mise en scène et l'implication d'une communauté a poussé certains auteurs à définir les données ouvertes comme un bien commun, tels que décrits par Elinor Ostrom et Charlotte Hess (2007), dans la mesure où elles possèdent tous leurs éléments fondamentaux :

- **Le libre accès au bien en question** matérialisé dans les données ouvertes par les portails,
- **Les règles de gestion communautaires** pour assurer la permanence d'un travail coopératif et protéger le bien d'un risque *d'enclosure*, matérialisé par la licence ouverte qui délimite les conditions d'utilisation des jeux de données,
- **La participation active d'une communauté** dans le développement et la protection d'un bien, matérialisés par les communautés numériques (Rheingold, 2015) et les systèmes de co-production de services décrits ci-dessus.

1.3.4 L'ouverture des données vue depuis le niveau inter-organisationnel

Au niveau inter-organisationnel, la littérature s'accorde sur le fait que les données ouvertes ouvrent de nouvelles possibilités d'alliances entre organisations publiques et privées (Strömberg et Näslund., 2017 ; Immonen., 2014). L'émergence de ces alliances étant dépendante de l'interopérabilité des données ouvertes.

1.3.4.1 Les données ouverte comme possibilités de nouvelles alliances entre organisations

Je commencerai par exposer l'exemple des villes intelligentes proposé par Strömberg et Näslund (2017). Ils perçoivent les données ouvertes comme la clé de voûte permettant la collaboration entre les organisations publiques et privées au sein des projets *smart city*. Immonen (2014) voit également de nouvelles possibilités commerciales se dessiner autour des données ouvertes. Elles peuvent offrir de nouvelles opportunités commerciales aux acteurs qui fournissent des données, aux acteurs qui consomment des données et aux acteurs qui développent des services et des applications innovants autour de ces données.

Ils posent néanmoins des conditions au succès de cette collaboration. Pour Immonen (2014), des *business model* et un environnement collaboratif - appelé écosystème - sont essentiels pour soutenir les entreprises basées sur des données ouvertes. Strömberg et Näslund (2017) listent plusieurs éléments nécessaires : une stratégie claire et approuvée d'ouverture et d'utilisation des données, l'implication de l'opinion publique sur les sujets de collecte, d'analyse et de transformation de données, un partage des connaissances entre

les parties prenantes des projets, des précisions sur les modalités de partage de données entre parties prenantes.

À propos de la nécessaire mise en place d'un écosystème, [Zuiderwijk \(2015\)](#) dessine les trois types de membres et de rôles qui devraient s'y trouver : les fournisseurs de données ouvertes, les utilisateurs de données ouvertes et les décideurs. Les fournisseurs de données ouvertes comprennent les agences gouvernementales qui produisent ou collectent de grandes quantités de données afin de remplir leurs tâches quotidiennes et les organisations financées par des fonds publics qui produisent et collectent des données au nom d'agences gouvernementales. J'ajoute à cette catégorie l'ensemble des acteurs civils et privés qui possèdent des données relevant de l'intérêt général et qui les rendent accessibles librement et gratuitement. À propos des utilisateurs, [Gascò-Hernández et al. \(2018\)](#) en distingue quatre types :

- Les fonctionnaires qui utilisent les données publiques pour améliorer les services publics autant que les processus de prise de décision.
- Les innovateurs, parmi lesquels nous retrouvons les développeurs informatiques et les entreprises, utilisent les données dans une optique de créer des nouveaux produits ou services innovants pour les commercialiser.
- Les chercheurs, data-journalistes et activistes utilisent pour leur part les données ouvertes dans le but de créer de la connaissance sous différents formats.
- Les citoyens utilisent généralement les données indirectement via certains médiateurs qui traduisent les données.

1.3.4.2 Des nouvelles collaborations comme fonction de l'interopérabilité des données ouvertes

Les auteurs insistent sur le fait que ces nouvelles collaborations dépendent de l'interopérabilité des données ouvertes, c'est-à-dire de leur capacité à communiquer entre elles.

Il faut comprendre derrière le concept de "données liées", un ensemble de meilleures pratiques pour la publication et l'interconnexion de données structurées sur le Web. Pour [Ngomo et al. \(2014\)](#), il serait intéressant d'adapter aux caractéristiques spécifiques des données ouvertes les nombreuses normes, méthodes et technologies développées par la communauté du web sémantique.

Un grand nombre de ces normes concernent la maintenance des métadonnées (données lisibles par ordinateur sur un ensemble de données et qui contiennent des informations pour la découverte, le contexte et l'utilisation détaillée des données) (Argyzoudis, 2014). Elles faciliteraient la connexion de données contenues dans différents portails de données ouvertes National Information Standards Organization (2014). Aussi sont-elles considérées comme un outil clé pour l'utilisation collaborative et efficace des données ouvertes liées (Zuiderwijk et al., 2012).

1.3.5 L'ouverture des données vue depuis le niveau industriel, territorial et sociétal

C'est à ce niveau d'analyse que la littérature retrace l'histoire du mouvement de l'ouverture des données et les potentialités qu'il recèle pour la société.

1.3.5.1 Histoire du mouvement de l'ouverture des données

Il trouve ses origines au début des années 2000 lorsque de nouvelles revendications citoyennes émergent par rapport à l'ouverture et à l'accessibilité des données publiques. En 2006, au Royaume-Uni, Charles Arthur et Michael Cross, journalistes au *Guardian* lancent un appel demandant l'ouverture des données publiques. Cet article intitulé « Give us back our crown jewels »⁸ expose les difficultés, les incohérences, l'absurdité résultant de l'utilisation des données publiques. Cet article marque le début du mouvement d'ouverture des données au Royaume-Uni. De part et d'autre du monde, des démarches d'ouverture des données similaires voient le jour. Néanmoins ces démarches d'ouverture et d'accessibilité des données publiques remontent bien avant les apports technologiques des NTIC, il né dans les années 1960, de l'ouverture des droits d'accès aux informations administratives.

C'est avec les nouvelles opportunités amenées par les NTIC et internet que le mouvement a pris de l'ampleur. En décembre 2007, le mouvement d'ouverture des données tel que nous le connaissons aujourd'hui prend forme à travers un plaidoyer en faveur de l'ouverture des données publiques.

⁸ <https://www.theguardian.com/technology/2006/mar/09/education.epublic> consulté le 16/03/18 à 11h15

En janvier 2009, Barack Obama signe deux mémorandums, faisant acte du fort engagement de son administration dans les démarches d'ouverture des données. Cette étape marque l'impulsion des mouvements d'ouverture dans le monde, le Royaume-Uni mettant en place une politique d'ouverture des données. En France, la création de la mission Etalab en 2011, lance l'accompagnement de l'Etat pour les démarches OD auprès des collectivités territoriales.

1.3.5.2 Le mouvement d'ouverture des données : un progrès sociétal aux multiples potentialités

Larroche et al. (2018) considèrent qu'au-delà de faciliter la mise à disposition des données ouvertes, ce mouvement incarne une tentative de création et d'organisation d'un espace ouvert à tous les acteurs qui souhaitent partager leurs données. Ce qui leur permet de conclure :

“Nous pouvons rapprocher ces données de « communs de la connaissance » si l'on suit la définition de Serge Proulx (2015), pour qui ces communs sont des corpus de savoirs et d'informations que l'on veut largement ouvert et directement accessibles au plus grand nombre d'utilisateurs sur la base de la gratuité du service.”

Larroche et al. (2018, p.278)

D'après Gil-Garcia *et al.*, (2015), les données sont à la fois l'une des quatre dimensions et l'un des dix éléments clés des nouvelles villes intelligentes. Ils spécifient que cette dimension et ce composant peuvent se concevoir comme des facilitateurs de toutes les autres dimensions et composants d'une ville intelligente.

Dans la même veine, certains auteurs considèrent même les données et la technologie comme le véritable cœur de l'intelligence urbaine. Pour Harrison *et al.*, (2010), les villes intelligentes sont des « zones urbaines qui exploitent des données opérationnelles telles que celles provenant des embouteillages, des statistiques de consommation d'énergie et des événements de sécurité publique pour optimiser le fonctionnement des services de la ville ».

Pour Goëta et Denis (2013), l'ouverture des données brutes des administrations et les dispositifs qui incitent à leur utilisation apparaissent “comme des moteurs de la reconfiguration non seulement de l'Etat (...) mais aussi des citoyens, instaurés en public du gouvernement et de ses données”. Selon l'ancien Premier Ministre britannique David Cameron, en créant un nouveau lien avec les citoyens, les données

ouvertes reconfigurent le rôle de l'Etat qui perdra le monopole des données et de leurs interprétations. Les citoyens pourront désormais intervenir dans le processus.

1.3.6 L'ouverture des données vue depuis la superposition de plusieurs niveaux d'analyse

D'une certaine manière la législation française appelait de ses vœux la superposition des niveaux d'analyse. Crosnier et Vidal (2017) rappellent à ce titre que la loi pour la République Numérique du 7 octobre 2016 considère :

“l'accès aux données publiques (comme allant) au-delà de la gratuité d'accès et demande que les données soient fournies dans un format ouvert, normalisé qui les rendent réutilisables. L'objectif est à la fois économique (susciter de nouvelles applications) et politique (permettre aux citoyens et leurs associations d'exercer le contrôle avec les données)” Crosnier et Vidal (2017)

Il est possible de rapprocher l'objectif économique de l'exploitation qui peut être faite d'un bien commun (niveau organisationnel) et l'objectif politique de la nécessaire constitution d'une communauté pour défendre ce bien contre les enclosures (niveau extra-organisationnel et sociétal).

Cette superposition des niveaux d'analyse se retrouve également dans l'analyse des barrières à l'ouverture des données. Par exemple, si Ruppert (2013) considère que les principales barrières à l'utilisation des données ouvertes par les citoyens sont principalement de l'ordre technique (manque d'équipement) et de la connaissance (en statistiques ou data-science), Ruijter et al. (2017) distinguent quatre catégories de barrières qu'il regroupe en deux dimensions.

Dimensions	Catégories
Culturelle	Intérêt et connaissance
	Valeur et qualité d'usage des données
Structurelle	Accès aux données
	Technique

Table 1.2 : Les barrières face à l'utilisation des jeux de données ouvertes (Ruijter et al., 2017)

Les barrières structurelles relèvent principalement des imperfections des démarches d'ouverture des données. L'accès aux données est souvent rendu difficile par le manque de méta-données qui qualifient les jeux ouverts et permettent de les référencer au mieux sur les moteurs de recherche. Ainsi, toute personne intéressée par un jeu de données, passera un temps important à parcourir de nombreuses plateformes indépendantes à la recherche de ce dernier, au risque de se décourager. Ce travail de production des métadonnées revient dans la plupart des cas au diffuseur de la donnée, et dans le cas des données ouvertes, aux collectivités publiques. Or, ce travail est rarement effectué du fait du manque de moyen de ces dernières (Braunschweig et al., 2012).

1.3.7 L'ouverture des données vue depuis les tensions paradoxales entre plusieurs niveaux d'analyse

Relativement à l'étude de l'ouverture des données depuis les tensions paradoxales, je n'ai pu identifier qu'une seule étude abordant l'ouverture des données depuis les tensions paradoxales entre les niveaux d'analyse, il s'agit de Robinson et Yu (2011). Chercheurs en technologie de l'information et en droit, ils montrent en quoi les données ouvertes, si elles permettent aux organisations gouvernementales de montrer leur intégrité et faire l'éloge de la transparence en invitant la société civile à utiliser les données, elles ne permettent pas pour autant à la société d'exercer un contrôle sur leurs activités ni à des utilisateurs de faire un usage des données dans le sens de leurs intérêts privés. Bien au contraire, ces deux aspects seraient opposés tant et si bien que la seule manière d'aboutir à une ouverture des données efficaces en termes d'usage serait de faire travailler ensemble la société civile et les organisations publiques sur les

données ouvertes, pour concilier les deux pôles opposés. Par leur recherche, ils éclairent les moyens de manager ces tensions paradoxales.

1.4 Les écosystèmes *open data*

1.4.1 Les écosystèmes *open data* vus depuis le niveau intra-organisationnel

Alors qu'aucun pourrait imaginer que l'étude des écosystèmes interdit un niveau d'analyse intra-organisationnel, plusieurs études contredisent ce préjugé. Un phénomène macroscopique n'empêche pas l'étude microscopique. Au contraire, l'étude des écosystèmes *open data* a invité les chercheurs à adopter des niveaux d'analyse nouveaux en étudiant les rôles des acteurs pris individuellement dans les écosystèmes d'un côté, et des modèles de management pour les équipes chargées de stimuler le fonctionnement des écosystèmes.

1.4.1.1 Le rôle des acteurs dans les écosystèmes *open data*

Outre les producteurs et réutilisateurs, amplement étudiés, la littérature s'est aussi intéressée au rôle des intermédiaires entre ces deux acteurs. Van Schallwyck et al. (2016) constatent qu'ils jouent plusieurs rôles importants dans l'écosystème : ils augmentent l'accessibilité et l'utilité des données ; ils peuvent assumer le rôle d'une "espèce clé" dans un écosystème de données ; et ils ont le potentiel de démocratiser les impacts et l'utilisation des données ouvertes. En prenant l'exemple de l'écosystème de données ouvertes de la gouvernance des universités publiques, les auteurs montrent que les intermédiaires peuvent réduire la viscosité des données gouvernementales et par là améliorer progressivement la qualité des données et le fonctionnement des écosystèmes. Intarakumnerd et Chaoroenporn (2013) adoptent un autre langage en affirmant que les intermédiaires jouent un rôle crucial pour compenser le manque de capital social qui entrave la fonctionnalité des écosystèmes. Enfin, Turki et al. (2017) conçoivent ce rôle d'intermédiaire comme celui d'un facilitateur qui contourne les obstacles rencontrés par les réutilisateurs potentiels de données ouvertes.

1.4.1.2 Des modèles de management pour les équipes d'animation des écosystèmes *open data*.

Les modèles de management pour les équipes d'animations sont inspirés des études sur le management d'écosystèmes biologiques. Sous le terme de “gestion des écosystèmes”, ces modèles recouvrent une réalité multiple (Yaffee, 1999) et suppose une gestion basée sur la collaboration des parties prenantes, la coopération entre les agences, l'intégration des informations scientifiques, sociales et économiques, la préservation des processus écologiques et une gestion adaptative (Koontz et Badine., 2007).

En reconnaissant la pluralité des modèles de management des écosystèmes, Yaffee (1999) estime que les prescriptions politiques, telles que les modifications de la législation, les mesures d'incitation et la fourniture d'informations, pourraient être mieux adaptées aux réalités des différentes données.

1.4.2 Les écosystèmes *open data* vus depuis le niveau organisationnel

Sur le sujet des écosystèmes *open data* et à ce niveau d'analyse, la littérature s'intéresse à la manière dont une organisation influence l'écosystème et inversement, comment elle peut être influencée par l'écosystème.

1.4.2.1 L'influence d'une organisation sur un écosystème *open data*

L'influence est double : d'un côté les organisations peuvent stimuler le fonctionnement d'un écosystème, de l'autre elles peuvent le dégrader et le faire disparaître (Ponte., 2015 ; Holling et Meffe., 1996 ; Briscoe et De Wilde., 2006).

Pour stimuler un écosystème *open data*, Ponte (2015) suggère à une organisation de comprendre la configuration et les mécanismes des écosystèmes afin de définir une stratégie d'influence. Il est ensuite utile de s'inspirer des cadres théoriques créés par la discipline de la gestion stratégique en établissant puis en évaluant plusieurs options stratégiques. Pour Kampars et al. (2019), la formation d'écosystèmes *open data* peut être pensée et améliorée par les organisations grâce à l'approche dites de “développement des capacités”⁹ des organisations. Selon cette approche, les écosystèmes *open data* seraient créés grâce au

⁹ Termes originaux : *capability-driven development*

déploiement de nouvelles capacités basées sur les systèmes d'information des organisations dont l'identification et l'utilisation de ressources présentes en dehors de l'organisation ou la modélisation de collaborations avec d'autres organisations.

À l'inverse, Holling et Meffe (1996, p.330) notaient qu'une intervention trop intense peut déclencher une "pathologie de la gestion des ressources". Une approche de captation, de contrôle ou de commande excessive de l'écosystème entraîne une perte de résilience du système qui ne s'adapte plus aux faibles variations. Il est nécessaire qu'une organisation comprenne alors que les mesures affecteront, à des degrés divers, la santé de l'écosystème *open data*, qui à son tour affectera en fin de compte les performances de l'organisation. Parce qu'une entreprise, comme une espèce individuelle dans un écosystème biologique, partage en fin de compte son sort avec l'ensemble de son écosystème, les organisations intelligentes poursuivent des stratégies qui profiteront à tous. La question restant alors de savoir comment promouvoir la santé et la stabilité de son propre écosystème tout en élaborant une stratégie à la hauteur de son rôle, contribuant ainsi à assurer le bien-être de l'organisation (Iansiti et Levin, 2004).

1.4.2.2 L'influence d'un écosystème *open data* sur une organisation

À l'image des travaux de Briscoe et De Wilde (2006), les écosystèmes *open data* peuvent profondément influencer l'architecture des systèmes d'information organisationnels. Pour les organisations se sachant dans un écosystème, elles peuvent faire évoluer l'architecture de leurs systèmes d'informations pour qu'ils puissent se recombinaer et évoluer dans le temps, en cherchant constamment à améliorer leur efficacité pour la base d'utilisateurs internes et externes.

Dans cette architecture, les applications sont créées en fonction des demandes des utilisateurs en utilisant l'optimisation évolutive pour agréger les services. Elles sont libres de migrer à travers l'écosystème *open data* et s'adapteront pour trouver des niches où elles seront utiles pour répondre aux demandes d'applications d'autres utilisateurs. Aussi ces systèmes d'informations seront libres d'intégrer et de combiner des applications et services extérieurs à l'organisation (Vossen, G., 2014).

1.4.3 Les écosystèmes *open data* vus depuis le niveau extra-organisationnel

Je trouve à ce niveau d'analyse un grand nombre d'études s'intéressant aux parties prenantes externes des écosystèmes centrés sur une organisation qui se trouve être, la plupart du temps, la collectivité productrice des données. La place centrale des collectivités s'explique par la légitimité naturelle que leur procure leur rôle de législateur, qui instaure le cadre juridique sous-jacent à la circulation des données ouvertes (Thomas., 2013).

Zuiderwijk et al. (2014) et Heimstädt et al. (2014) distinguent bien les rôles dans ce type d'écosystème. D'un côté, on trouve le fournisseur de données ouvertes, situé au centre de l'écosystème ; de l'autre, les intermédiaires et utilisateurs qui manipulent les données produites. La question que se posent ces auteurs est de savoir dans quelle mesure le fournisseur peut assurer et la croissance de son écosystème.

Concernant la croissance, pour Zuiderwijk et al. (2014), il faut que le fournisseur garantisse par des infrastructures quatre activités clés, à savoir :

- la publication de données ouvertes sur l'internet ;
- la recherche, la découverte, l'évaluation et la visualisation des données et des licences correspondantes ;
- le nettoyage, l'analyse, l'enrichissement, la combinaison, la liaison et la validation des données ;
- l'interprétation et la discussion des données et le retour d'information au fournisseur de données et aux autres parties prenantes.

En outre, pour intégrer les éléments de l'écosystème et leur permettre d'agir comme un tout intégré, il faudrait trois éléments supplémentaires :

- des parcours d'utilisateurs indiquant des orientations sur la manière dont les données ouvertes peuvent être utilisées ;
- un système de gestion de la qualité des données ;
- différents types de métadonnées pour pouvoir relier les éléments.

Sur la question de l'équilibre Heimstädt et al. (2014) et Ding et al. (2011) concluent qu'il s'agit pour les fournisseurs d'assurer la diversité des parties prenantes de l'écosystème avec, par exemple, les fonctionnaires qui conservent les données brutes, les développeurs qui créent des applications avec des

données ouvertes et les citoyens qui visualisent et analysent les résultats de ces données. Aussi s'agit-il d'assurer l'adéquation entre l'offre et la demande de données, à l'image d'un marché financier.

1.4.4 Les écosystèmes *open data* vus depuis le niveau inter-organisationnel

Alors que les études extra-organisationnelles ont porté leur intérêt sur les interactions entre une organisation productrice de données et ses parties prenantes (bidirectionnelles), les études inter-organisationnelles s'intéressent pour leur part aux interactions entre toutes les organisations (multilatérales) membres d'un écosystème et à leur modélisation

1.4.4.1 Les différents types d'interactions entre organisations membres d'un écosystème *open data*

On trouve différentes analyses de ces interactions inter-organisationnelles. Pour Hall et al. (2012), les organisations d'un écosystème fonctionnent à l'image d'un réseau avec des asymétries de pouvoir et d'information, des différences entre les niveaux de formation et de capacités de chacun mais surtout avec des relations d'interdépendances. Par exemple, les utilisateurs de données ouvertes dépendent des données et des métadonnées qui sont publiées par les fournisseurs de données ouvertes pour pouvoir utiliser ces données. Les fournisseurs de données ouvertes dépendent du retour d'information sur leurs données qu'ils obtiennent des utilisateurs de données ouvertes. De cette façon, les organisations qui fournissent des données ouvertes et les organisations qui utilisent les données ouvertes sont impliquées dans des réseaux d'interaction. Pollock (2011) préfère évoquer des boucles d'interactions¹⁰ qui alimentent des cycles de données ainsi que la vie de l'écosystème. La donnée est ici un objet qui se partage au gré d'échanges entre plusieurs acteurs. Entre le producteur et l'utilisateur, le producteur et l'intermédiaire et l'intermédiaire et l'utilisateur, de véritables boucles d'interaction se créent, aboutissant à l'émergence d'un écosystème.

Zuiderwijk (2015) distingue, quant à elle, les boucles d'interactions de Pollock (2011) des discussions et collaborations qui constituent un autre mode d'interaction entre organisations. Tandis que les boucles

¹⁰ Termes originaux : *feedback loops*

d'interactions sont généralement faites à propos des jeux de données produits et libérés, les discussions et collaborations peuvent avoir lieu dans ou hors de la plateforme, à propos d'une donnée ou d'autre chose.

Ces collaborations et discussions peuvent prendre plusieurs formes :

1. Des environnements de travail collaboratifs comme des wiki qui hébergent des discussions ou documents contenant le fruit des apprentissages de certains acteurs ;
2. Des conversations sur des forums ouverts ;
3. Des messages publics ou privés.

En dehors des infrastructures, les collaborations peuvent avoir lieu sur les réseaux sociaux (Twitter, Facebook et LinkedIn) ou dans des événements de type hackathons.

1.4.4.2 La modélisation des interactions inter-organisationnelles sous forme d'une chaîne de valeur

Le fonctionnement empirique d'un écosystème *open data* peut se modéliser sous la forme d'une chaîne de valeur au sein de laquelle chaque acteur apporte ses capacités (Cosgrave, 2013). Le *Climate group* dans son rapport *Information marketplaces : the new economics of cities* propose un exemple de chaîne de valeur d'un écosystème de ville intelligente.

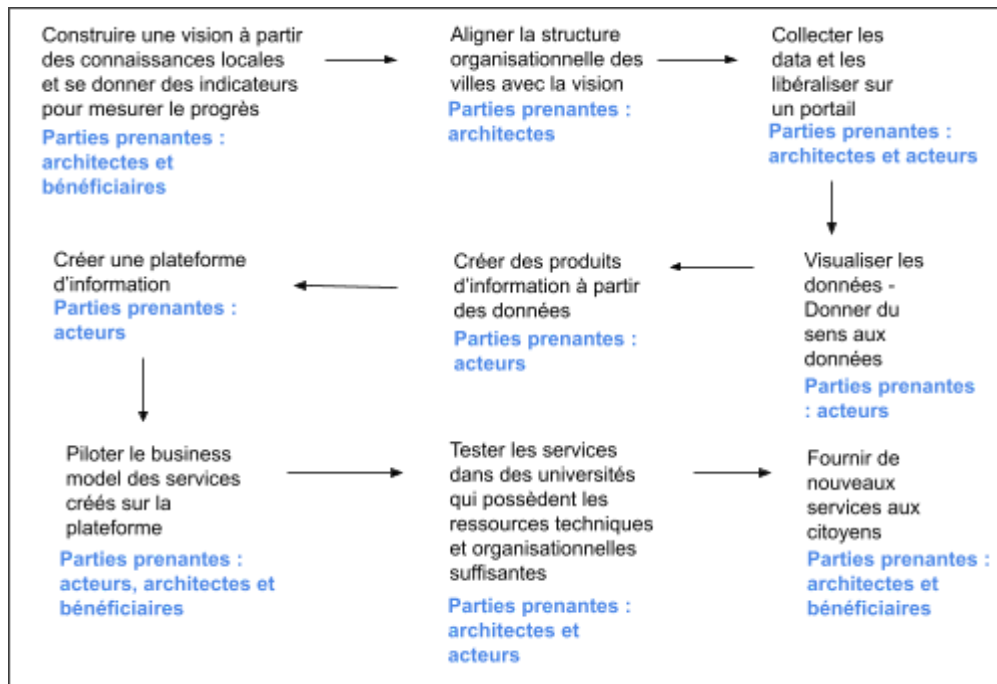


Figure 1.7 Chaîne de valeur d'un écosystème open data d'une ville intelligente

Du point de vue de la chaîne de valeur, l'écosystème *open data* adopte ce que Raymond (1999, p.120) appelle une approche "cathédrale", c'est à dire une approche centralisée «soigneusement élaborée par des sorciers individuels ou de petites bandes de mages travaillant dans un isolement splendide». On retrouve ce type de modélisation dans les rapports de la Commission européenne ou de l'OCDE (Ubaldi, 2013 ; Dekkers et al., 2006)

1.4.5 Les écosystèmes *open data* vus depuis le niveau industriel, territorial et sociétal

Quand on passe du niveau inter-organisationnel au niveau sociétal, l'intérêt des auteurs passe de l'étude des processus des écosystèmes à leurs résultats et impacts.

Pour Delacroix et Lawrence (2019), les écosystèmes *open data* sont un moyen de rétablir un rapport de force équitable entre les citoyens et les organisations qui collectent des données sur leurs agissements. Selon eux, l'écosystème devrait être juridiquement structuré sous la forme d'un ensemble défini comme

des organisations dans lesquelles les citoyens viendront déposer leurs données en échange de garanties sur leur protection¹¹. Reggi et Dawes (2016) prouvent quant à eux que les écosystèmes *open data* peuvent initier deux cycles d'influences sur la société. En premier lieu, ils encouragent la transparence des institutions, ensuite ils accroissent l'engagement du public et renforce la responsabilité des gouvernements.

Dans un point de vue plus technique, Boudreau (2007) envisage les écosystèmes comme les éléments préparatoires de véritables infrastructures d'innovation dans lesquelles les données circulent et interagissent les unes avec les autres. À terme, Heath et Bizer (2011) envisagent que ces infrastructures soient omniprésentes et qu'Internet devienne un écosystème *open data* global. Tout comme les hyperliens du Web classique relient les documents en un seul espace d'information global, les données utilisent des hyperliens pour relier des données disparates en un seul espace de données global. Ces liens, à leur tour, permettent aux applications de naviguer dans l'espace de données. Par exemple, un jeu de données décrivant une personne et mentionnant son lieu d'habitation pourra être relié avec d'autres données décrivant ce même lieu.

1.4.6 Les écosystèmes *open data* vus depuis la superposition de plusieurs niveaux d'analyse

Je commencerai cette partie par une citation Poikola et al. (2011, p.11) qui donne toute sa légitimité à la position du chercheur qui interroge plusieurs niveaux d'analyse pour comprendre les écosystèmes *open data*.

Définition : **Un écosystème** : se définit comme “une entité multi-niveaux et multidimensionnelle où la matière première, la donnée, en ce qui concerne sa distribution et son développement, est l'objet de la coopération”.

Sous l'égide de cette définition, j'ai recensé plusieurs études utilisant une approche multi-niveaux pour aborder les écosystèmes *open data*. Les auteurs empruntent plus particulièrement le cadre d'analyse de la théorie de l'activité pour comprendre la dynamique des écosystèmes *open data*. Grâce à cette théorie, il vise à atteindre deux objectifs : comprendre d'un côté les forces sociales et structures qui influencent les

¹¹ Termes originaux : *Data Trusts*

interactions entre utilisateurs et les relations entre les utilisateurs et les technologies de l'information présentes au sein d'un écosystème (Ojo, Janowski et Estevez., 2011).

La théorie de l'activité relie les activités dans lesquelles s'engagent des individus, la description de ces individus, leurs buts et intentions et les règles et normes qui délimitent les activités ainsi que la communauté impliquée dans ces activités (Kuuti, 1996). L'ensemble de ces éléments forme un système d'activité souvent illustré par une pyramide au sommet de laquelle se trouve la technologie de l'information qui centralise les échanges entre ses éléments : la plateforme open data.

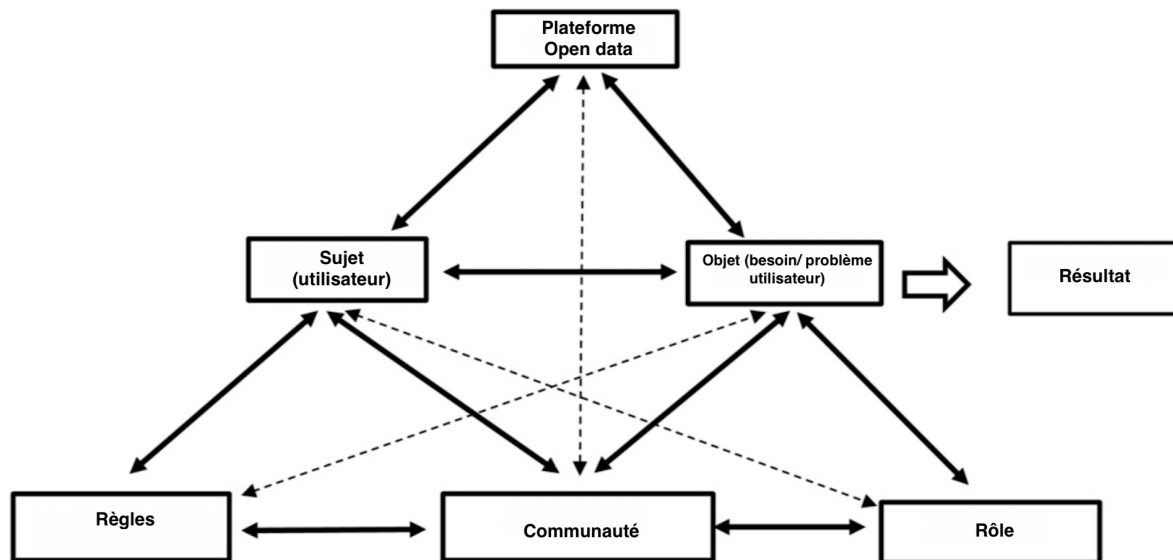


Figure 1.8. Le système d'activité open data côté utilisateur (d'après Engeström, 2001, p.140)

La plateforme open data revêt dans ce schéma deux fonctions :

- Faciliter la collaboration entre les membre du système d'activité pour qu'ils puissent identifier un besoin/problème (un objet) ou encore une tension dans le fonctionnement de l'écosystème qu'il s'agira de résoudre par une logique de co-création de service ;

- Jouer un rôle d'intermédiaire entre les deux systèmes d'activité qui forment l'écosystème (Lnenicka et Komarkova., 2019) : le système des fournisseurs de l'open data et celui des utilisateurs.

Sur la base de ces deux fonctions de la plateforme open data, Ruijer et al. (2017) aboutissent à une vision exhaustive de l'écosystème open data, de ses dynamiques ainsi que de sa finalité.

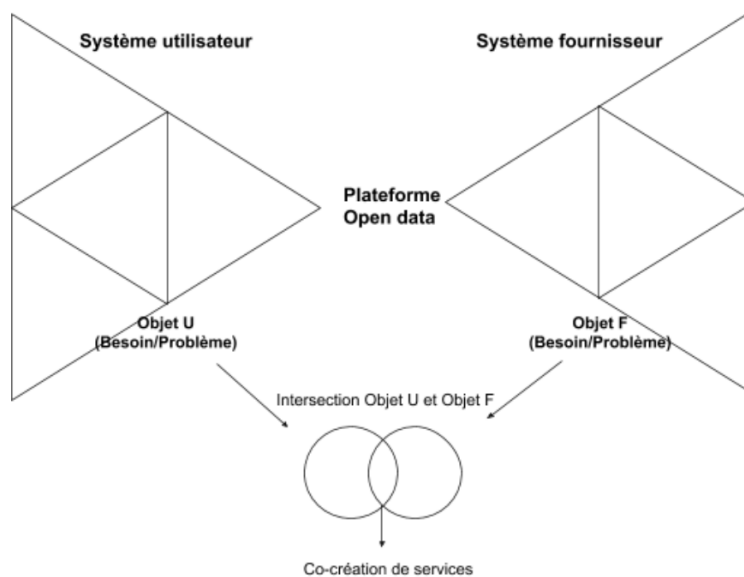


Figure 1.9 : Modèle des écosystèmes selon la théorie de l'activité (d'après Ruijer et al., 2017, p.48)

Davies (2011) se concentre sur les interactions entre utilisateurs des données et l'infrastructure technique, autrement dit entre la base et le sommet de la pyramide. Ce couplage permet selon lui de faire une distinction entre l'infrastructure, en tant que structures et installations physiques et organisationnelles de base nécessaires au fonctionnement d'une société ou d'une entreprise (souvent centralisée, normalisée et gérée par un petit groupe d'agents), et les composantes émergentes, autonomes et auto-organisées d'un écosystème, reliées entre elles dans des boucles de rétroaction.

1.4.7 Les écosystèmes *open data* vus depuis les tensions paradoxales entre plusieurs niveaux d'analyse

Enfin, sur l'approche des tensions paradoxales, je recense l'étude de Thomas et Autio (2003), dans laquelle ils insistent sur le fait que sans un modèle de compréhension clair et fiable des interdépendances contradictoires entre les membres, les responsables d'un l'écosystème *open data* ne peuvent identifier les capacités et les offres disponibles au sein de l'écosystème et manquer de perspicacité pour conduire des combinaisons mutuellement bénéfiques entre eux. De même, sans modèle et hypothèses fortes sur le fonctionnement d'un écosystème, les responsables ne pourraient empêcher les membres de se dissocier les uns des autres et l'écosystème pourrait se dissoudre.

La figure 1.10 ci-dessous présente l'analyse morphologique de la littérature.

Concept/Unité d'analyse	Intra-organisationnel	Organisationnel	Extra-organisationnel	Inter-organisationnel	Industriel, territorial et sociétal	Superposition des niveaux d'analyse	Tension paradoxale entre les niveaux d'analyse
Objets de recherche possibles	<ul style="list-style-type: none"> • Individuel • Groupe/EquipeProjet • Aire fonctionnelle • Business unit • Rôle 	<ul style="list-style-type: none"> • Entreprise • Autre organisation • Stratégie • Business model 	<ul style="list-style-type: none"> • Parties prenantes externes • Individus • Communauté • Organisation 	<ul style="list-style-type: none"> • Alliance • Réseau • Ecosystème 	<ul style="list-style-type: none"> • Développement d'une industrie • Différences inter-industrielles • Régions locales • Pays • Institutions supra-étatiques • Citoyens • Politique publique 	Combinaisons des niveaux précédents	Niveaux contradictoires et interdépendants
Usage des données ouvertes	<ul style="list-style-type: none"> • Heise et Naumann (2012) • Kuhn (2011) • Solar et al. (2012) • Henriksen (2015) • Wirtz et al. (2016) • Ganapati et Reddick (2012) • Tinati et al. (2012) • Böhm et al. (2010) • Mainka et al. (2015) • Desouza et Bhagwatwar (2012) 	<ul style="list-style-type: none"> • Dekkers et al. (2006) • Wang et Lo (2016) • Jaakola et al. (2015) • Desouza et Bhagwatwar (2012) • Vert et Vasu (2015) • Léveillée et Timms (2015) • Lindman et al. (2014) • Linders et al. (2013) • O'Hara (2012) • Hartog et al. (2014) 	<ul style="list-style-type: none"> • Picon (2015, p.27-28) • Zuidewijk et Janssen (2013) • Chan (2013) • Bertot et al., (2014) • Radl et al. (2013) • Parycek et al. (2014) • Power et al. (2015) • Kassen (2013) • Ohemeng et Ofosu-Adarkwa (2015) • Volpi et al. (2014) 	<ul style="list-style-type: none"> • Susha et al. (2015a) • Magalhaes et al. (2014) • Chan (2013) • Yang et al. (2015) • Janssen et Zuidewijk (2014) • Shum et al. (2015) • Cowan et al. (2015) • Veeckman et Van Der Graaf (2015) • Susha et al. (2015b.) • Kampars et al. (2019) 	<ul style="list-style-type: none"> • Rabari et Storper (2014) • Curry et al. (2015) • Whitmore (2014) • Nam (2015) • Maramieri (2014) • Höchtl et Reichstädter (2011) • Lakomaa et Kallberg (2013) • Lofi et Krestel (2012) • Shkabatur (2012) 	<ul style="list-style-type: none"> • Courmont (2016, p.78) • Khayyat et Bannister (2017) • Worthy (2015) • Balena et al. (2013) • Yang et al. (2015) 	<ul style="list-style-type: none"> • Boullier (2016)
Plateforme open data	<ul style="list-style-type: none"> • Eckelberg et al. (2018) • Zhu et Freeman (2018) • Denis et Pontille (2012) • Boutet (2001) • Couture (2012) • Star (1999) • Gardey (2001) • Gitelman (2014) • Edwards (2010) • Edwards et al. (2011) 	<ul style="list-style-type: none"> • Courmont (2016) • Eisenmann et al. (2008) • Allen and Varga (2006) • Benbya and McKelvey (2006) • Tanriverdi et al. (2010) • Briner et Rome (2018) • Al-Ruithe et al. (2017) • Braunschweig et al. (2012) • Liang et al. (2018) 	<ul style="list-style-type: none"> • Cathedral model (Raymond, 1999) • Cognitivist platform (Danneels, 2017) • Loveluck, 2015 • Schweik (2006) • Tiwana et al. (2010) • Schweik (2007) • Allison-Cassin et Scott (2018) • Pfenninger et al. (2017) • Neumaier et al. (2018) • Lapi et al. (2012) 	<ul style="list-style-type: none"> • Bourcier (2013) • Hogan (2017) • Lusch et Nambisan (2015) • Lee et al. (2018) • Parker et al. (2016) • Walravens et al. (2016) • Zuidewijk et al. (2014) • Welle Donker et van Loenen (2017) 	<ul style="list-style-type: none"> • IRM (2015) • Ojo et al. (2015) • Carrara et al. (2015) • Cordasco et al. (2017) • Villoen et al. (2016) • Aguilera et al. (2017) • Weerakkody et al. (2017) • Attour et Rallet (2014) • Osagie et al. (2017) • Millette et Osein (2016) 	<ul style="list-style-type: none"> • Plateforme connectionniste de Danneels (2017) • Denis (2011) • Pitt et al. (2017) 	<ul style="list-style-type: none"> • Villum (2014) • Plateforme autopoiétique de Danneels (2017)
Données ouvertes	<ul style="list-style-type: none"> • Ruijter et al. (2018) • Le Corf (2016) • Janssen, Charalabidis & Zuidewijk (2012) • Kuk et Davies (2011) • Laklani et Wolf (2003) • Behkamal et al. (2014) • Bunakov et Jeffrey (2013) • Conradie et Choenni (2012) 	<ul style="list-style-type: none"> • McLeod et Mc Naughton (2016) • Zuidewijk (2015) • CNIL (2017) • Heller (1998) • Akrab (2016) • Trojette (2013) • Dawes et al. (2004) • Dawes et Helbig (2010) • Meijer et Thaeens (2009) 	<ul style="list-style-type: none"> • Mabi (2014) • Bourcier (2013) • Hess et Ostrom (2007) • Levine (2007) • Gurstein (2011) • Kassen (2013) • Deltor et al. (2013) • Zhang, Dawes et Sarkis (2005) • Srinivasan (2012) 	<ul style="list-style-type: none"> • Strömberg et Näslund (2017) • Gascó-Hernández et al. (2018) • Zuidewijk (2015) • Ngomo et al. (2014) • Immonen (2014) • Zuidewijk et al. (2012) • National Information Standards Organization (2014) • Argyzoudis (2014) • Zuidewijk et al. (2012) 	<ul style="list-style-type: none"> • Ruppert (2012) • Denis et Goeta (2013) • Goeta (2016) • Flyverbom et Murray (2018) • Dawes et al. (2016) • Larroche et al. (2018) • Gil-Garcia et al. (2015), • Harrison et al. (2010) • Goeta et Denis (2013) 	<ul style="list-style-type: none"> • Ojo et al. (2015) • Khayyat et Bannister (2017) • Ruijter et al. (2017) • Ruppert (2013) 	<ul style="list-style-type: none"> • Yu (2016)
Ecosystème open data	<ul style="list-style-type: none"> • Harrison et al. (2012) • Koontz et Bodiine (2008) • Yaffee (1999) • Martin et al. (2017) • Turki et al. (2017) • van Schallwyck et al. (2016) • Intarakumnerd et Chaoroenporn (2013) 	<ul style="list-style-type: none"> • Kampars et al. (2019) • Koontz et Bodine (2007) • Briscoe et De Wilde (2006) • Holling et Meffe (1996) • Ponte (2015) • Vossen (2014) • Iansiti et Levin (2004) 	<ul style="list-style-type: none"> • Lee (2014) • Pollock (2011) • O'Reilly (2011) • Zuidewijk et al. (2014) • Ding et al. (2011) • Thomas (2013) • Mahmadi (2013) 	<ul style="list-style-type: none"> • McLeod et Mc Naughton (2016) • Pollock (2011) • Zuidewijk (2015) • Cosgrave, (2013) • Raymond (1999) • Hall et al. (2012) • Brodie (2010) 	<ul style="list-style-type: none"> • Delacroix et Lawrence (2019) • Boudreau (2007) • Fransman (2010) • Nardi et O'Day (1999) • Vander Sande et al. (2013) • Heath et Bizer (2011) • Reggi et Dawes (2016) 	<ul style="list-style-type: none"> • Welle Donker et van Loenen (2017) • Daves et al. (2016) • Ojo, Janowski et Estevez (2011) • Ruijter et al. (2017) • Poikola et al. (2011) • Ubaldi (2013) • Davies (2011) • Lnenicka et Komarkova (2019) 	<ul style="list-style-type: none"> • Autio et Thomas (2016)

Figure 1.10 - Analyse morphologique de la littérature

1.5 Identification Gap

J'ai pu identifier des gap dans la littérature en m'appuyant sur l'analyse morphologique du framework multi-niveau présenté ci-dessus (figure 1.10). Il a mis en exergue une zone obscure de la littérature. Une investigation de cette zone a finalement permis d'identifier précisément deux gaps qui seront comblés dans cette thèse.

Un faible volume de recherches étudiant la construction des usages depuis les tensions paradoxales entre les niveaux d'analyses.

L'analyse morphologique du framework multi-niveaux met en exergue le peu de recherches étudiant la construction des usages issus des données ouvertes depuis les tensions paradoxales entre les niveaux d'analyse. Le peu d'auteurs ayant pris cette position donne un indice quant à la présence probable d'un gap méthodologique dans la littérature.

En effet, en partant du point de vue qu'une étude à un certain niveau d'analyse est dépendante tout en s'opposant aux conclusions d'une étude à un autre niveau d'analyse, le chercheur se voit contraint d'adopter une posture méthodologique nouvelle et encore peu répandue en sciences de gestion. Il a face à lui un problème où l'effet de chaque solution prescrite à un certain niveau d'analyse sera annulée par une autre solution prescrite sur un autre niveau d'analyse. Autrement dit, chaque solution qu'il propose deviendra un problème sur un autre niveau d'analyse. Quelle grave incertitude, toutes les fois que l'esprit se sent dépassé par lui-même ; quand lui, le chercheur, est tout ensemble le pays obscur où il doit chercher et où tout son bagage ne lui sera de rien.

Alors comment doit-il chercher ? À ce stade, face à ce type de problème, il ne s'agit plus seulement de chercher mais de créer. Il faut que le chercheur se positionne en face de quelque chose qui n'est pas encore et que seul il peut réaliser, puis faire entrer dans sa lumière.

Avec ces tensions paradoxales, le chercheur fait place à des problèmes pernicioeux (Churchman, 1967, Rittel et Weber, 1973) caractérisés par :

- **Des supposés flous et des contraintes basées sur des contextes environnementaux aux contours mal définis.** Comme je l'ai montré ci-dessus, les écosystèmes open data répondent à ces caractéristiques.
- **Des interactions complexes entre les sous-composantes du problème et sa solution,** les sous-composantes étant ici les déclinaisons de la question de recherche aux différents niveaux d'analyse.
- **Une forte sensibilité au changement aussi bien concernant le processus de design de la solution que l'artefact qui y répond.** En effet, toute solution proposée sera rapidement modifiée du fait de la dynamique incessante des écosystèmes open data dans laquelle elle s'inscrira.
- **Une forte dépendance aux capacités cognitives humaines pour produire des solutions efficaces.** À nouveau, le chercheur se positionne en face de quelque chose qui n'est pas encore et que seul il peut réaliser, puis faire entrer dans sa lumière.
- **Une dépendance aux habiletés sociales des êtres humains pour produire des solutions efficaces.** En effet, dans la mesure où la solution implique la pluralité des acteurs d'un écosystème *open data*, l'efficacité de la solution dépendra de sa capacité à stimuler et orienter leurs habiletés sociales.

Or, d'après Hevner et Chatterjee (2010), les problèmes pernicieux impose au chercheur un type particulier de méthodologie : une Méthodologie de Recherche Design Science¹². Leur particularité provient du fait que ces recherches aboutissent, outre les éventuelles publications scientifiques, au design d'un artefact testé et évalué par les praticiens du secteur concerné. Dès lors, s'agit-il uniquement de chercher à comprendre la construction des usages issus des données ouvertes ? Non, pas seulement, c'est insuffisant, il faut créer.

¹² Termes originaux : *Design Science Research Methodology*

1.6 Une absence de recherches utilisant la méthodologie de recherche design science pour aborder l'intégralité du processus de construction des usages issus des données ouvertes

L'indice donné par l'analyse morphologique du framework multi-niveaux est confirmé dans la littérature : il n'existe, au moment de l'écriture de cette thèse, aucune recherche utilisant la méthodologie de recherche *design science* pour aborder l'intégralité de la construction des usages issus des données ouvertes. Aussi cette thèse exploite-t-elle ce gap. Elle se propose d'utiliser cette méthodologie en tant qu'activité scientifique innovante permettant de créer des connaissances pratiques et utiles aux praticiens de l'open data.

1.6.1 L'opportunité de solliciter la méthodologie de recherche design science pour combler un gap dans la littérature

Selon Hassan et Twinomurizi (2018), il n'y a pas eu d'études utilisant l'approche design science pour aborder la question des données ouvertes durant ces cinq dernières années. De plus, ils notent que l'opportunité d'exploiter ce gap est particulièrement intéressante dans la mesure il existe un grand nombre d'enquêtes et d'études identifiant des bonnes pratiques d'ouverture et d'usage des données. Ces connaissances pourraient être mises à contribution dans la conception d'un artefact via l'approche *design science*. Je nuancerai leur propos dans la mesure où j'ai pu identifier deux recherches ayant sollicité cette méthodologie : Alexopoulos et al. (2014) qui l'utilise pour concevoir une nouvelle génération de plateforme open data et Zuiderwijk (2015) pour concevoir une infrastructure visant à faciliter la coordination de l'usage de l'open data.

Si ces recherches appliquent bien la méthodologie de recherche *design science*, elles le font sur une conception restreinte de l'usage des données ouvertes qui ne permet pas de saisir l'ensemble du processus de construction des usages. Ils partent du principe, avec Davies (2010) que l'usage se subdivise en cinq étapes (Davies, 2010) et décident de ne retenir que les trois premières étapes pour concevoir l'usage des données ouvertes :

- L'étape de transformation de la donnée en fait
- L'étape de transformation de la donnée en information
- L'étape de transformation de la donnée en interface.

Ils décident d'exclure les deux dernières étapes, que je considère comme cruciales dans la mesure où elles représentent l'aboutissement de la construction des usages :

- L'étape de transformation de la donnée en données
- L'étape de transformation de la donnée en service

Zuiderwijk (2015) écrit :

"Les trois premières étapes sont fondamentales. Sans les trois premiers types d'utilisation des données, les deux derniers ne sont pas possibles. Par exemple, le développement d'un service nécessite que certaines informations soient extraites des ensembles de données. Les quatrième et cinquième types d'utilisation des données concernent le développement d'interfaces et de services innovants, qui font référence à l'utilisation par des développeurs et les entrepreneurs et qui exigent des compétences et une expérience considérables de la part de l'utilisateur des données. Comme cette étude se concentre sur l'utilisation par les chercheurs plutôt que par les développeurs ou les entrepreneurs, elle est orientée vers les trois premiers types d'utilisation des données de Davies (2010)" Zuiderwijk (2015, p.47)

Je conclus donc qu'il existe un gap théorique dans la mesure où les connaissances développées par Alexopoulos et al. (2014) et Zuiderwijk (2015) sur la construction des usages sont incomplètes. Elles couvrent de manière imparfaite le processus de construction des usages et ma thèse ambitionne de produire des connaissances qui viendront compléter leurs travaux. Plus précisément, ces connaissances seront de nature à informer tout chercheur ou praticien cherchant à concevoir des artefacts pour permettre, stimuler et faciliter la construction des usages dans un écosystème open data donné.

Aussi, par ricochet, suis-je face à un gap empirique dans la mesure où les travaux de Alexopoulos et al. (2014) et Zuiderwijk (2015) n'apportent que des solutions partielles aux usagers des données ouvertes. Leurs artefacts ne permettent que de solutionner les situations problématiques qui émergent lors des trois premières phases de l'usage de Davies (2010) au détriment des situations émergeant dans les deux dernières. Ma thèse complète ses travaux en proposant un artefact qui répond aux freins pouvant apparaître tout au long du processus de construction des usages issus des données ouvertes.

C'est dans le but de combler ce gap théorique que je pose ma deuxième sous-question de recherche, à savoir :

Quelles sont les situations problématiques qui freinent la construction des usages issus des données ouvertes ?

Pour combler ce gap théorique et empirique, je me vois dans l'obligation de mettre en place une méthodologie de recherche en *design science*. En effet, d'après Hevner et Chatterjee (2010), les problèmes pernicioeux, qui n'ont été confrontés que deux fois dans la littérature, et auxquels je choisis de faire face, imposent au chercheur de mettre en place ce type de méthodologie. Elle s'impose également dans la mesure où elle permet de prendre comme unité d'analyse les tensions paradoxales entre unités d'analyse.

Par conséquent, je présenterai dans la prochaine partie un état des lieux des méthodologies de recherche en *design science* dans le domaine des systèmes d'information.

Par la mise en place de cette méthodologie, je pourrai répondre à ma troisième sous-question de recherche, à savoir :

Quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ?

1.6.2 L'approche *design science* dans le domaine des systèmes d'information : une activité scientifique innovante pour créer des connaissances pratiques et utiles aux praticiens de l'open data

Dans un premier temps, je m'attacherai à retracer l'histoire de l'approche *design science* pour montrer comment elle est apparue au sein de la communauté scientifique et par quels biais elle a pu s'établir en tant qu'activité scientifique à part entière.

Une fois ce tour d'horizon historique établi, il s'agira dans un second temps de faire un tour d'horizon des controverses que cette approche suscite au sein de la communauté scientifique. Ainsi, je ferai un état des

lieux des interrogations qui font l'actualité de la littérature sur cette approche, avant d'identifier les gaps méthodologiques que ma thèse contribuera à combler.

Enfin, je montrerai en quoi cette approche est particulièrement adaptée pour la recherche et la pratique sur la construction des usages issus des données ouvertes. Pour cela, je montrerai les apports particuliers de l'approche *design science* par rapport à d'autres méthodologies et ferai état des travaux sur l'open data l'ayant sollicité.

In fine, cette troisième partie de ma revue de la littérature constitue une démonstration qui justifie le choix d'une méthodologie *design science* que je développerai dans le deuxième chapitre de ma thèse.

1.6.2.1 Les fondements et le développement de l'approche *design science* au sein de la communauté scientifique

En retraçant le développement de l'approche *design science* depuis son origine, il est question d'apporter un éclairage sur sa nature et sur les éléments qui la constituent. Tout d'abord, je montrerai que l'approche *design science* est une activité scientifique distincte qui utilise la conception d'artefact comme démarche d'étude scientifique. Ensuite, je montrerai que cette approche repose sur trois éléments principaux : des objets matériels dits artefacts, une démarche de recherche collaborative impliquant chercheur et praticien et un socle de connaissance particulier dit *design theory*.

1.6.2.1.1 Le développement historique des méthodologies de design science dans la communauté scientifique : de l'étude de la conception des systèmes d'information à la conception comme démarche d'étude scientifique

Résumé

Il s'agira dans cette partie de faire état du développement historique de l'approche *design science* en partant de son origine : l'étude de la conception de systèmes d'information. J'exposerai en premier lieu les différents instruments créés par ce courant de recherche visant l'étude de la conception de systèmes d'information. Ensuite il s'agira de mettre en lumière les points de rupture ayant amené les chercheurs

en systèmes d'information à basculer d'une approche analytique ayant pour objet la conception à une approche pragmatique utilisant la conception comme instrument de recherche.

a. L'étude de la conception de systèmes d'information

Définition : **Situation de conception** : issue des travaux de Midler (1996), cités par Pascal (2006), est définie comme l'ensemble des instants au cours desquels des acteurs sont amenés collectivement à produire un résultat qui sera soumis à un jugement externe (Girin, 1990).

La caractérisation de ces six situations constitue un bon point de départ pour comprendre le courant de recherche afférent à la conception dans le domaine des systèmes d'information. Ces six critères constituent le socle sur lequel est bâti le cadre d'analyse classique de la conception, que je présente ci-après :

1. Toute situation de conception est *“une heuristique tendue vers une situation globale”* (Midler, 1996). Autrement dit, la première étape d'une situation de conception consiste à imaginer une finalité à celle-ci. Matérielle ou immatérielle, cette finalité constitue un horizon vers lequel les activités des concepteurs tendent. Cela étant, cette finalité est définie de façon suffisamment large pour laisser les hasards de la démarche modeler le problème initial et le résultat final.
2. La conception est *“une affaire de communication et d'intégration de différentes logiques”* (Midler, 1996). Nécessairement collective, toute situation de conception suppose l'implication de divers acteurs et de diverses perspectives. De ce fait, la communication entre eux et l'intégration de leurs diverses logiques s'avèrent capitales pour espérer atteindre la finalité.
3. La conception est affaire de navigation dans l'incertitude. En effet, la conception faisant évoluer le problème initial et la forme du résultat final (critère 1) et impliquant des acteurs et logiques qui peuvent diverger et se confronter (critère 2), une situation de conception ne peut se planifier à l'avance.
4. La conception implique convergence et irréversibilité. Tous les choix faits par l'équipe de conception réduisent progressivement le nombre d'alternatives possibles concernant la forme du résultat final. Ainsi, plus la conception se poursuit, plus elle est irréversible.

5. Chaque projet de conception est singulier. Par conséquent, il n'existe aucune règle formelle pour conduire les négociations et compromis qui se nouent au cours de l'évolution de la conception.
6. L'espace de la conception est "*ouvert et fluctuant*". Autrement dit, il n'est pas possible de prévoir *a priori* l'ensemble des acteurs finalement impliqués dans la situation ni l'ensemble des formes possibles du résultat final.

b. Le cadre d'analyse classique de la conception : le cycle de vie d'un système d'information

Le cadre d'analyse classique de la conception se concentre autour de l'idée d'un cycle de vie du projet de conception. Un système d'information se construit et s'analyse à travers des stades successifs qui permettent d'aboutir à une solution répondant à la demande initiale (Reix et al., 2009). Ainsi, il est possible pour les praticiens et les chercheurs d'identifier très clairement les étapes à mettre en œuvre, leur résultat et les conditions de passage d'une étape à une autre.

Dans cette vision classique et cartésienne, il convient de retenir cinq étapes qui ensemble forment un cycle traduisant l'idée qu'un système d'information constitue bel et bien un système vivant et évolutif : l'analyse préalable (2.6.1.1.1), la conception (2.6.1.1.2), le développement (2.6.1.1.3), l'implantation (2.6.1.1.4) et la maintenance (2.6.1.1.5).

c. L'analyse préalable : définir le projet de conception

Par cette analyse préalable, il s'agit d'obtenir des réponses à deux questions :

- *Quel est le problème à résoudre ? Qui sont les acteurs et les activités concernées ?* Les concepteurs peuvent ainsi caractériser, définir les frontières du problème et localiser le niveau auquel il se situe dans l'organisation.
- *Quelles sont les solutions existantes, économiquement justifiées et techniquement viables pouvant répondre à ce problème ?* Par là, il s'agit de s'assurer que la solution à la problématique n'est pas trop coûteuse ou que les risques qu'elle comporte ne sont pas trop importants.

L'output de cette première étape prend généralement la forme d'un avant projet qui établit un diagnostic sur la faisabilité et la viabilité du projet de conception et peut s'étendre jusqu'à une esquisse de la solution à construire.

d. La conception générale

Cet étape repart de l'avant-projet formulé pour arriver à un projet détaillé décrivant toutes les actions à entreprendre au niveau conceptuel, organisationnel ou technique : “La conception établit un ‘pont’ entre la définition des besoins et la réalisation de la solution : elle explique par quels mécanismes seront réalisées les différentes fonctions requises pour la satisfaction des besoins” (Reix et al., 2009, p.400).

On distingue alors deux phases :

- Définition de la dimension informationnelle du SI. Il s’agit de construire un modèle représentant la nature et l'enchaînement des tâches du processus à concevoir
- Définition de la dimension organisationnelle du SI. On fixe alors quels sont les rôles à distribuer, la chronologie des tâches, ainsi que les méthodes utilisées pour mettre en place les nouvelles procédures au sein de l’organisation.

A la fin de cette deuxième étape, on aboutit à un cahier des charges qui fixe une image précise du système à concevoir ainsi que la façon d’organiser la conception dans l’organisation cible.

e. Le développement

Cette étape vise à traduire concrètement les choix technologiques effectués dans le cahier des charges. C’est à ce moment qu’on programme et qu’on teste les logiciels spécifiques. Encore, on distingue deux étapes :

- L’analyse détaillée des traitements d’une part et des données à solliciter d’autre part.
- La programmation et les tests des logiciels sur des échantillons de données

À l’issue de l’étape de développement, l’organisation dispose d’un prototype de système, prêt à l’implantation dans l’organisation.

f. L’implantation

On sollicite ici tous les niveaux du système d’information pour implanter concrètement le prototype de système. On comprend ici l’installation du matériel, le test de logiciels, la constitution de la base de

données et la formation des utilisateurs. Il existe ici plusieurs manières d'introduire progressivement le système dans l'ensemble de l'organisation : lancement en parallèle des anciens systèmes, lancement partiel concernant une seule partie du système conçu, etc.

g. La maintenance

Cette dernière étape exprime pleinement la dimension vivante de la conception d'un système d'information. Elle intègre toutes les étapes qui permettent de changer le système d'information conçu, que ce soit au niveau technique, organisationnel ou conceptuel.

Ce cadre d'analyse classique de la conception des systèmes d'information, qui a connu son apogée en France dans les années 1980-1990 avait pour avantage de bien clarifier les différentes dimensions sollicitées dans la conception et permettait de retracer très clairement les différentes stades permettant d'aboutir à un système d'information final. Son application concrète s'est traduite par l'utilisation de la méthode MERISE que je décris ci-après.

1.6.2.1.2 La méthode MERISE : une méthode pour modéliser les systèmes d'information et faciliter leur conception

Cette méthode reprend les fondements de la conception classique tout en instaurant une méthode particulière de représentation du système d'information. Au cycle de vie préexistant, la méthode MERISE ajoute un cycle d'abstraction permettant de représenter un système d'information dans sa forme réelle, c'est à dire perçue par l'utilisateur, et dans sa forme informatisée, c'est à dire technique et informatique. Précisons que cette méthode s'axe autour de la distinction entre les données gérées par un système d'information et les traitements qui lui sont appliqués. Le tableau ci-dessous retrace les différentes modélisations instaurées par la méthode MERISE.

	Données	Traitements	
<i>Quoi</i>	Quelles informations utiles ? (MCD)	Pour faire quoi ? (MCT)	Conceptuel
<i>Qui, Où, Quand</i>	Quelle structure de données ? (MLD)	Qui fait quoi et où, et quand ? (MOT)	Logique
<i>Comment</i>	Comment stocker les données ? (MPD)	Comment fait-on ? (MOPT)	Physique

Figure 1.11: Phases d'analyse des données et des traitements de la méthode Merise (Servigne, 2008, p.3)

Il est possible de retrouver ici trois couches de modélisation, qui s'applique chacune aux données et aux traitements :

La couche conceptuelle, qui a pour objectif de décrire le réel : la modélisation conceptuelle des données décrit le monde sous la forme d'entités étant en relation les unes avec les autres. Par exemple, on utilisera cette modélisation pour représenter les liens entre les données sur les clients d'une banque et celles relatives à son compte en banque. La modélisation conceptuelle des traitements a trait à la représentation des processus métiers exécutés avec les données. Il s'agira par exemple de représenter les mouvements de données entre le client d'une banque et son compte en banque (ouverture du compte, consultation, fermeture, etc.).

La couche logique, qui décrit la structuration des données et des traitements : la modélisation logique des données représente la structure des données. Par exemple, on pourra par cette modélisation se rendre compte que les données clients sont structurées sous formes de tables avec divers attributs (identifiants unique, adresse, etc.). La modélisation logique des traitements s'attache à décrire les opérations fonctionnelles effectuées sur les données.

La couche physique qui se concentre sur la dimension matérielle et informatique : la modélisation physique des données s'attache à décrire les supports physiques des données : langage de définition des données, produit logiciel choisi pour leur traitement, etc. Par exemple, on saura via cette modélisation que le numéro unique des clients correspond à une valeur entière. La modélisation physique des traitements détaille les programmes utilisés.

Le cadre d'analyse classique de la conception des systèmes d'information est ainsi doté d'un cycle de vie traduisant les différentes étapes à mettre en œuvre, et de modèles permettant d'avoir une vision détaillée et exhaustive du système d'information. Celui-ci a été adapté à la conception de systèmes d'information venant répondre à un besoin spécifique d'une organisation particulière et bien délimitée. Aussi venait-elle répondre aux besoins de directions informatiques qui opéraient principalement en silos et en consultant peu les utilisateurs finaux des systèmes d'information.

Or, avec l'émergence de nouvelles formes organisationnelles et d'une nouvelle tendance consistant à intégrer de plus en plus les utilisateurs dans la conception, de nouvelles formes d'analyse de la conception ont vu le jour.

1.6.2.1.3 Les dernières innovations en matière d'analyse de la conception : RAD, approches sociotechniques

Les démarches et les méthodes d'analyse de conception ont évolué pour s'adapter aux nouvelles exigences des organisations vis-à-vis de leurs systèmes d'information (Reix et al., 2009, p.397-409). Il est possible de distinguer le modèle en V (1) et le modèle W (2).

a. Le modèle en V face au besoin de systèmes d'information intégrés

Parmi les nouvelles exigences, on trouve le besoin d'étendre le champ d'action des systèmes d'informations, qui doivent pouvoir répondre à plusieurs besoins en parallèle. Il n'est plus question de concevoir un système d'information pour chaque besoin mais d'englober au sein d'une seule solution la réponse à plusieurs besoins métiers. On parlera de systèmes d'information intégrés. C'est dans ce cadre que le modèle en V a été créé (Figure ci-dessous). Celui-ci reprend le cadre d'analyse classique en introduisant deux particularités :

1. La décomposition du système global en sous-systèmes au moment des phases de test et de développement. Ainsi, les sous-systèmes sont développés et testés indépendamment.
2. L'intégration des sous-systèmes à l'issue des tests de chaque sous-système et avant de tester et d'implémenter le système global.

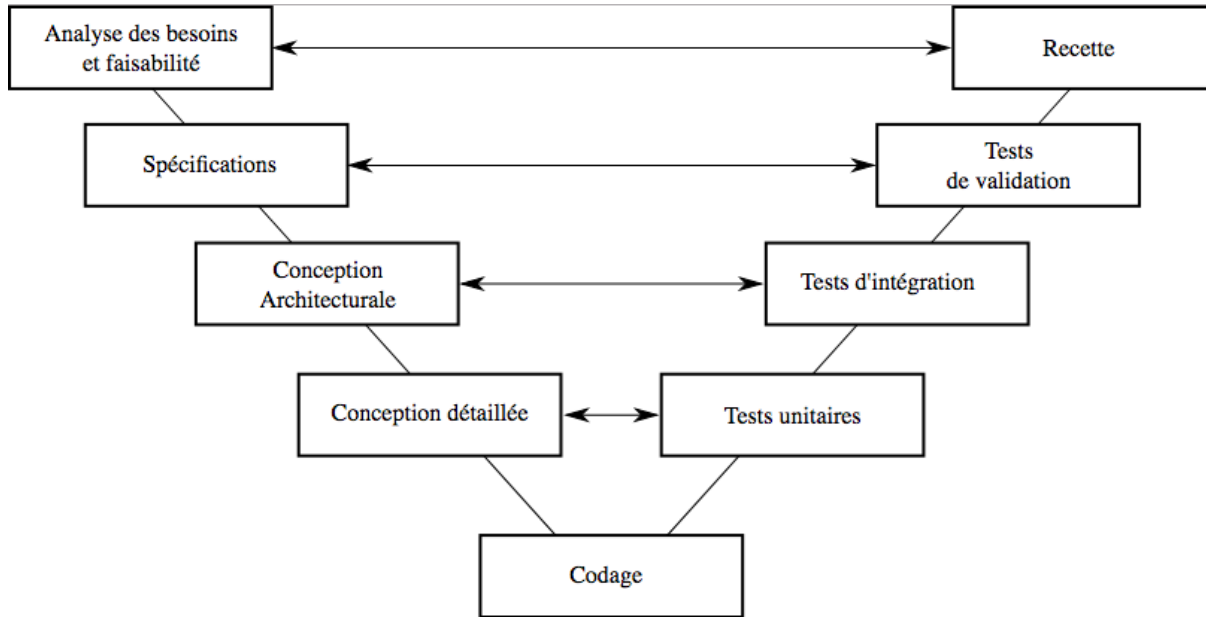


Figure 1.12: Exemple d'un modèle en V de conception d'un système d'information (Firesmith, 2014, p.14)

Avec ce modèle en V, l'analyse des systèmes d'information gagne en précision et permet de saisir la complexité grandissante des systèmes d'information.

b. Le modèle en W pour pallier à l'insuffisance de l'expression des besoins métiers

Autre problème grandissant dans la conception des systèmes d'information et ayant nécessité une adaptation des méthodes de conception et d'analyse : l'insuffisance des phases d'identification des besoins métiers. En effet, les résultats d'un recensement des besoins métiers à un temps donné s'avère souvent imprécis voire incorrects. Il est difficile pour les métiers d'identifier *a priori* leurs besoins. Ceux-ci se dévoilent généralement au fur et à mesure du développement du système. Le modèle en W intègre ce fait en introduisant dans la conception des maquettes ou des prototypes (figure ci-dessous).

Il s'agit de mettre à disposition des utilisateurs des représentations non-fonctionnelles (maquettes d'interfaces par exemple) ou fonctionnelles (prototype) du futur système et ce, le plus rapidement possible. Ces maquettes et prototypes servent alors "d'outils de dialogue entre le développeur et

l'utilisateur et permet à ce dernier, par une expérimentation limitée, de préciser ses besoins" (Reix et al., 2009, p. 405)

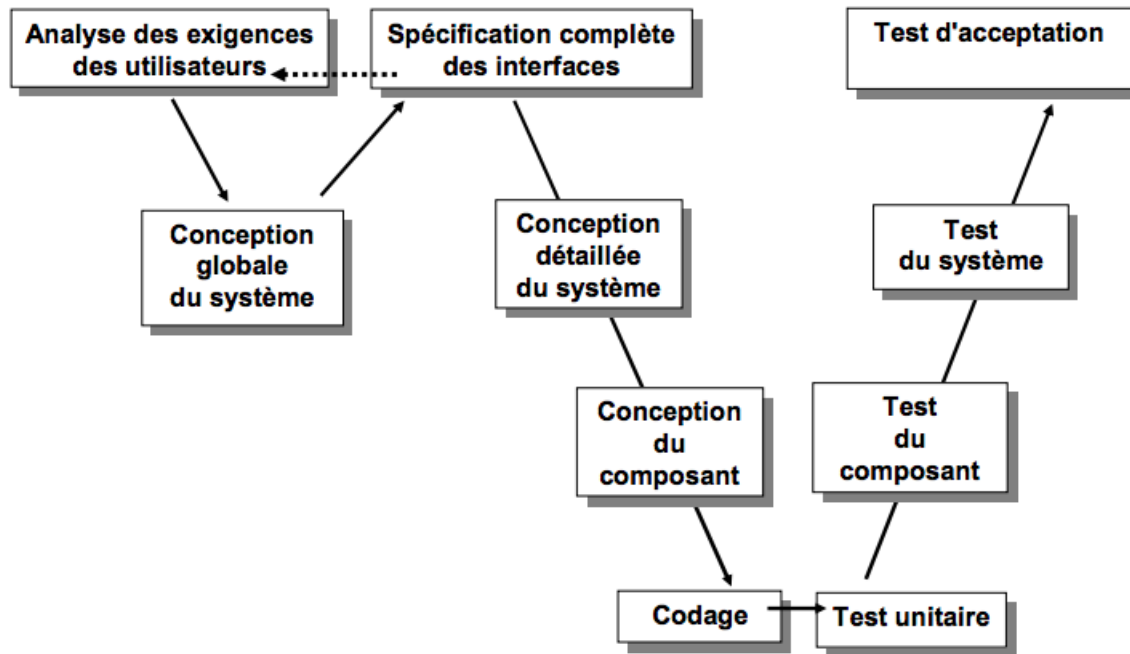


Figure 1.13 : Exemple d'un modèle en W de conception d'un système d'information (Petit, 2006,p.155)

c. La méthodologie RAD, cheffe de file des méthodes collaboratives : une réponse aux contraintes de temps et à l'évolution dynamique des besoins métiers

Un problème traditionnel rencontré dans les situations de conception consiste à perpétuellement perfectionner la solution et par la même occasion, à retarder régulièrement son achèvement et sa mise en œuvre. La méthodologie *Rapid Application Development* (RAD) prône face à ce biais la primauté du temps sur la qualité du système : on préfère un résultat imparfait rendu dans les temps à un résultat parfait rendu plus tard.

Cette méthodologie adjoint à chaque phase de conception une enveloppe de temps à l'issue de laquelle un prototype doit être fourni. Au fur et à mesure des étapes, les prototypes se complexifient progressivement pour former la solution finale. Plus précisément, on décompose la conception en 5 phases (ci-dessous) qui

doivent contenir les éléments suivants : un objectif précis, un acteur responsable, une durée maximale impérative, la fourniture obligatoire d'un résultat, des phases constitutives de son déroulement.

Phase	Description
Initialisation	Phase permettant d'organiser le démarche : définition du projet, ressources à organiser, etc.
Expression des besoins	Phase permettant d'identifier les besoins des futurs utilisateurs
Conception	Construction d'une première solution décrivant le futur système à deux niveaux : les flux et les données d'un côté, les maquettes d'écran et d'état de l'autre.
Construction	Développement du système par cycles successifs avec recours à des prototypes.
Mise en oeuvre	Installation du système

Figure 1.14 : phases de conception selon la méthode RAD (adapté de Reix et al., 2009, p.85)

Précisons que cette méthode inaugure une participation accrue des utilisateurs dans la démarche de conception. Lors des phases d'expression des besoins et de conception, il est question d'organiser des sessions de travail collaboratif auxquelles participent les utilisateurs et les concepteurs et qui permettent la co-construction de solutions mais surtout une meilleure adaptation à l'évolution dynamique des besoins des utilisateurs.

Cette méthode RAD fait office de cheffe de file d'un ensemble d'autres méthodes collaboratives donnant une place importante aux futurs utilisateurs et au temps de conception. À l'extrémité du spectre de ces méthodes, on trouve les méthodes *scrum* et l'*extreme programming* qui réduisent la conception à une série de périodes très courtes dites sprints auxquelles participent utilisateurs et concepteurs.

d. Les approches sociotechniques : la prise en compte de l'environnement extérieur dans le design et l'analyse de la conception

Si l'efficacité des méthodes agiles a assuré leur succès, elles possèdent néanmoins des limites pointées par la recherche académique. Baxter et Sommerville (2011) mettent en avant que, malgré l'instauration des sessions de dialogue avec les utilisateurs, l'utilisation des méthodes agiles n'est en aucun cas une garantie de solutions adaptées au contexte d'usage. Les moyens déployés n'assurent pas d'obtenir le résultat escompté. Cette incompatibilité est expliquée par le fait que les méthodes agiles se basent sur deux postulats simplistes, contraires à la réalité des organisations :

1. Les futurs utilisateurs sont disponibles pour interagir avec les concepteurs. Or, dans la réalité, les métiers n'ont souvent pas de temps à allouer au développement des systèmes d'informations.
2. Les besoins exprimés par les utilisateurs sont compatibles avec les contraintes exercées par l'organisation. Or, le niveau d'analyse individuel et organisationnel sont rarement compatibles.

Baxter et Sommerville (2011) résumant ces limites par une trop grande attention portée aux utilisateurs finaux par rapport au contexte organisationnel :

"Bien qu'il y ait certainement des idées intéressantes qui émergent des méthodes agiles, l'accent qu'elles mettent sur l'interaction avec les utilisateurs individuels ne répond pas au besoin d'une plus grande sensibilisation socio-technique dans l'ingénierie des systèmes"

Baxter et Sommerville (2011 p.7)

Les approches sociotechniques visent justement à rétablir un équilibre dans la prise en compte de ces deux éléments. Elles partent du principe que toute organisation est aujourd'hui constituée d'interactions entre des hommes, des machines et l'environnement extérieur et que, par conséquent, tous ces éléments sont à prendre en compte dans la conception de systèmes d'information.

Partant de ce principe d'ordre général, Baxter et Sommerville (2011) précisent les cinq critères qui guident la conception de systèmes d'information, considéré comme des systèmes socio-techniques :

1. Le système doit être constitué de parties interdépendantes
2. Le système doit s'adapter et participer à l'atteinte des objectifs de l'environnement extérieur

3. Le système doit comporter un environnement interne qui comprend un sous-système technique et un sous-système social interdépendants
4. Les systèmes sont caractérisés par l'équifinalité. Autrement dit, les objectifs des systèmes peuvent être atteints de plusieurs façons différentes. Ceci implique que des choix de design doivent être faits pendant les phases de développement du système.
5. La performance du système dépend de l'optimisation conjointe du sous-système social et du sous-système technique. En se concentrant sur un sous-système au détriment d'un autre, on risque de dégrader l'état du système global.

De nombreux facteurs ont modifié les méthodes de conception et, et par conséquent, le regard que portent les chercheurs à leur égard : extension du champ d'action des systèmes d'information, insuffisance des phases d'identification des besoins, réduction du temps de conception, prise en compte de l'environnement extérieur.

Sans prétendre à l'exhaustivité, cette brève revue des méthodes d'analyse dévoile la richesse d'une démarche de conception aussi bien pour les praticiens que pour les chercheurs : elle amène à identifier les différents acteurs et composantes impliqués, les interactions qui les lient ; elle pousse à interagir avec les futurs utilisateurs, à tester des prototypes, et à implanter des solutions dans une organisation. Cette richesse, comme nous le verrons par la suite, a attiré l'attention de la recherche académique qui l'a progressivement utilisée pour créer de nouvelles connaissances. Plus précisément, les chercheurs ont vu à travers les démarches de conception un moyen d'accroître la pertinence de leur recherche vis à vis des praticiens.

1.6.2.1.4 Le basculement des recherches en système d'information : passer du mode 1 au mode 2 de recherche pour gagner en pertinence

En 1999, Benbasat et Zmud constatent le déclin de la pertinence des recherches en systèmes d'information et mettaient en évidence cinq facteurs pour l'expliquer. Les deux premiers ont trait à la nature de la discipline des systèmes d'information, le troisième à l'objet même de la discipline, le quatrième au manque d'exposition des chercheurs aux contextes économiques et technologiques dans lesquels se trouvent les systèmes d'information étudiés, le dernier aux contraintes du monde académique. Le tableau ci-dessous synthétise les facteurs et leur provenance.

Facteur	Origine	Description
La rigueur plutôt que la pertinence	Histoire et nature de la discipline des SI	La discipline des systèmes d'informations faisant l'objet d'une remise en cause constante par les autres domaines quant à son assise scientifique, les chercheurs du domaine ont établi des critères de rigueur à respecter. Si la rigueur imposée a finalement permis à la discipline d'être consacrée en tant que science, le niveau de pertinence des travaux à quant à lui décrié.
Un manque de "tradition cumulative"	Histoire et nature de la discipline des SI	La discipline des SI s'appropriant de nombreux cadres théoriques en provenance d'autres disciplines, il est très difficile pour les chercheurs du domaine des SI de construire des modèles théoriques fédérateurs
Dynamisme des technologies de l'information	Objet de la discipline : les technologies de l'information	Le domaine des technologies de l'information se développe rapidement, dans une temporalité qui excède de loin celle de la recherche. Ainsi, la plupart des recherches sont rapidement obsolètes
Une faible exposition des chercheurs aux contextes économiques faisant usage des SI	Manque d'exposition des chercheurs aux contextes économiques	L'agenda des chercheurs ne leur permet pas de rester au contact des praticiens et des designers de technologies de l'information

Le facteur <i>publish or perish</i>	Contraintes du monde académique	Les chercheurs travaillent dans des institutions en compétition les unes avec les autres, avec pour critère de performance le nombre de publications. Les chercheurs sont logiquement plus poussés à publier qu'à s'intéresser aux événements dans les sphères économiques des praticiens.
-------------------------------------	---------------------------------	--

Figure 1.15 : tableau récapitulatif des facteurs expliquant le déclin de pertinence des recherches en systèmes d'information (Benbasat et Zmud, 1999, p.10)

a. Le mode 1 de recherche : accroître la pertinence des recherches en SI

Les premières recommandations pour accroître la pertinence des recherches en SI auprès des praticiens (Benbasat et Zmud, 1999).

Pour réorienter la recherche académique en systèmes d'information vers les intérêts des praticiens, Benbasat et Zmud, 1999 avancent un ensemble de propositions. Van de Ven et Johnson (2006) y ont vu une première piste pour réduire l'écart croissant entre la théorie et la pratique. On parle de mode 1 de la recherche.

Ce premier mode se résume à un ensemble de principes à suivre pour le chercheur :

- Choisir des thèmes de recherche directement en lien avec les problématiques rencontrés par les praticiens.
- Impliquer les éditeurs de revues de recherche et les chercheurs dans des discussions autour des retentissements pratiques de leurs travaux.
- Penser un projet de recherche à partir des résultats pratiques potentiels et non à partir des travaux de recherche existants.
- Produire des cadres de pensée compréhensibles et actionnables par les praticiens.
- Privilégier les formes claires et concises pour la présentation des résultats de recherche.

Ce mode 1 de recherche s'inscrit comme une première tentative de réduire l'écart entre la théorie et la pratique en incitant le chercheur à penser ses travaux du point de vue du praticien. Une démarche de recherche est ici toujours l'apanage de la communauté scientifique seule mais se conçoit en relation avec les praticiens.

Selon Pascal, 2012, cette façon d'accroître la pertinence des recherches en systèmes d'information présente des lacunes. Le mode 1 de recherche n'intègre pas pleinement les praticiens puisqu'il

“considère toujours la création de connaissances comme l'activité spécifique de la communauté académique” (Pascal, 2012, p.26)

Pour assurer une réelle pertinence, il s'agit plutôt de considérer que la théorie et la pratique constituent deux types de connaissances distinctes et complémentaires. Une approche que Van de Ven et Johnson (2006) nomment mode 2 de recherche.

b. Le mode 2 de recherche : faire participer activement les praticiens à la conduite de la recherche

Le mode 2 de recherche a été créé pour assurer la pertinence des recherches (Van de Ven et Johnson (2006)). Il propose une nouvelle solution au manque de pertinence des recherches en sciences de gestion. À l'inverse du mode 1 de recherche qui cible l'inefficacité du transfert de connaissance, le mode 2 de recherche s'intéresse à la problématique de production de connaissances :

"le fossé entre la théorie et la pratique peut être un problème de production de connaissances"

Van de Ven et Johnson (2006 p.808)

Dans ce mode, il ne s'agit plus de conduire une recherche pour les praticiens mais avec eux. La pertinence des recherches est ici assurée par une collaboration avec les praticiens et ce, lors des phases de design, de mise en œuvre et d'implémentation des recherches. Je préconise alors de former des équipes de recherches hétérogènes composées de chercheurs et de praticiens. L'effort de pertinence se porte alors plus sur les moyens de collaboration avec les praticiens que sur le choix et l'orientation des thèmes de recherches.

Avec l'intention de guider les chercheurs dans la mise en place de collaborations fructueuses avec les praticiens, Amabile et al. (2001) proposent cinq lignes directrices :

1. Choisir avec attention les praticiens à inclure et ce, en fonction de la complémentarité de leurs parcours, compétences, motivation intrinsèque avec celle du chercheur ainsi que de leur capacité à travailler avec des personnes au style cognitif et à la culture professionnelle différente.
2. Clarifier dès le début de la collaboration l'implication, le rôle, la responsabilité et les attentes de chacun, tout en les actualisant au fur et à mesure du processus.
3. Mettre en place des moyens et des habitudes de communication régulière.
4. Développer des modes de collaboration de façon à assurer une confiance mutuelle entre les acteurs.
5. Établir avec l'équipe des sessions de réflexion critique pour discuter des potentiels conflits.

Selon Pascal (2012), ce mode 2 de recherche incarne aujourd'hui l'avancée la plus significative dans la quête d'un gain de pertinence des travaux dans le domaine des systèmes d'information. La co-production de connaissances avec les praticiens permet une véritable prise en compte de leur réalité et facilite l'implémentation des résultats de la recherche auprès des sphères économiques et industrielles visées.

A mon sens, l'intégration croissante des praticiens constitue le point de basculement entre une approche scientifique ayant pour objet la conception à une approche scientifique tout autre utilisant la conception comme instrument de recherche.

Dans les travaux fondateurs de l'approche *design science* d'Hevner et al. (2004), on retrouve cette intégration dans le cycle d'activités appelé "boucle de pertinence". Il suppose un dialogue entre chercheurs et praticiens pour co-construire le problème initial de la recherche et évaluer le résultat de la recherche. Bien que, comme le note Pascal, 2012 la place des praticiens reste relativement restreinte, leur introduction dans le processus de recherche constitue une avancée critique qui a permis de construire le *design science* tel que nous le connaissons aujourd'hui.

1.6.2.1.5 La conception comme démarche scientifique

a. La reconsidération de la conception en sciences de gestion comme véritable champ scientifique

Il n'a pas fallu attendre l'approche *design science* pour émettre l'idée que la conception puisse être une démarche efficace en sciences de gestion. On trouve déjà dans l'ouvrage d'Herbert Simon, *Le nouveau management* (1981) une mention des démarches conception, qui selon lui revêtent une grande importance dans les sciences de gestion. En effet, il considère que les sciences de gestion sont avant tout des sciences de la conception et non des sciences de l'analyse. Autrement dit, la gestion se distingue par l'intérêt qu'elle porte aux façons de construire des instruments actionnables par la société. Selon Pascal (2006) :

“La spécificité de la gestion par rapport aux sciences institutionnelles telles l'économie ou la sociologie, tient à sa dimension ingénierique : la science de gestion peut, en effet, se définir au sens large comme l'ingénierie de l'organisation sociale ou de façon spécifique comme la coordination du pilotage finalisé de quelques fonctionnaires pré-identifiées”

Pascal (2006, p. 243)

Cette conception des sciences de gestion, que j'adopte dans cette thèse, fait de la conception un véritable outil permettant à la fois d'assurer la pertinence des recherches (voir partie précédente) mais aussi de produire des connaissances scientifiques d'un nouvel ordre. Celles-ci ont été décrites par la taxonomie établie par Gregor (2006) qui distingue les 5 types de théories développés en sciences de gestion :

1. Les théories pour analyser
2. Les théories pour expliquer
3. Les théories pour prédire
4. Les théories pour expliquer et prédire
5. Les théories pour concevoir et agir.

La particularité scientifique des sciences de gestion, et donc des systèmes d'information, réside précisément dans la combinaison de ces divers types de théories. Au-delà d'analyser, de comprendre et de prédire, les sciences de gestion s'attachent à produire des connaissances qui sont directement mises au service de la conception d'instruments et de l'action. Comme l'explique David (2000, p.210), le chercheur en science repose en équilibre entre les lois de la raison, et les possibilités de la création :

“Il lui fait accepter de contribuer scientifiquement à une pensée stratégique qui ne peut se réduire à la seule juxtaposition de résultats analytiques partiels, mais doit rechercher les articulations créatives, les combinaisons et les intégrations novatrices”

David (2000, p.122-123)

b. Les sciences de l’artificiel : un nouveau cadre scientifique pour développer des connaissances permettant la conception et l’intégration d’artefact dans la réalité (Avenier, 2009)

Pour asseoir définitivement la conception comme démarche scientifique, les auteurs en *design science* se rapportent tous aux travaux d’Herbert Simon (1969) qui instaure un nouveau paradigme scientifique : les sciences de l’artificiel. Ce paradigme vise à résoudre un paradoxe entre le monde qui entoure les hommes et les outils scientifique déployés pour le comprendre au moment de la publication. Alors que la plupart des sciences extraient les lois des objets qui nous entourent, considérant que ces objets en sont naturellement dotées, Herbert Simon observe que nous vivons principalement dans un monde façonné par l’homme. C’est à dire que les objets qui nous entourent sont bien plus définis par les fonctions et objectifs qui leur ont été assignés que par des lois qui proviendraient de la nature.

Pour résoudre ce paradoxe, Simon développe un cadre scientifique nouveau dont le but est de :

“Comprendre l’enchevêtrement de multiples projets humains évolutifs, leurs interrelations et leurs rapports avec des régulations perçues comme naturelles, en vue de la conception d’artefacts évolutifs destinés à fonctionner dans des environnement eux-mêmes susceptibles d’évoluer”

Avenier (2009, p. 56)

Ce cadre scientifique développe deux types de connaissances complémentaires, qui font écho avec la distinction entre connaissances théoriques et connaissances pratiques : d’un côté les connaissances de type conceptuel principalement produites selon le formalisme académique et dont le but est de décrire la réalité des artefacts qui nous entourent ; de l’autre les connaissances opératoires, plus orientées vers la pratique et dont le but est d’alimenter l’action de praticiens.

On retrouve ces deux types de connaissances décrits par Herbert Simon (1969) dans l'approche *design science*. Nous verrons par la suite ce qu'elles apportent à la fois aux praticiens des sciences de gestion et aux chercheurs.

1.6.2.1.6 Apports pratiques et théoriques de la conception comme démarche scientifique

L'apport principal des recherches utilisant l'approche design science réside dans la conception d'un artefact informatique (Hevner et al., 2004). Cette conception doit s'appuyer sur des théories existantes qui servent à la fois à épauler le chercheur dans les activités de conception mais aussi dans l'évaluation des artefacts créés. Ces connaissances correspondent aux connaissances opératoires décrites par Herbert Simon (1969). Ces aller-retours entre un stock de connaissance disponible et la conception d'un artefact forment selon Hevner et al. (2004) une boucle de rigueur qu'il schématise de la façon suivante (en rouge sur le schéma) :

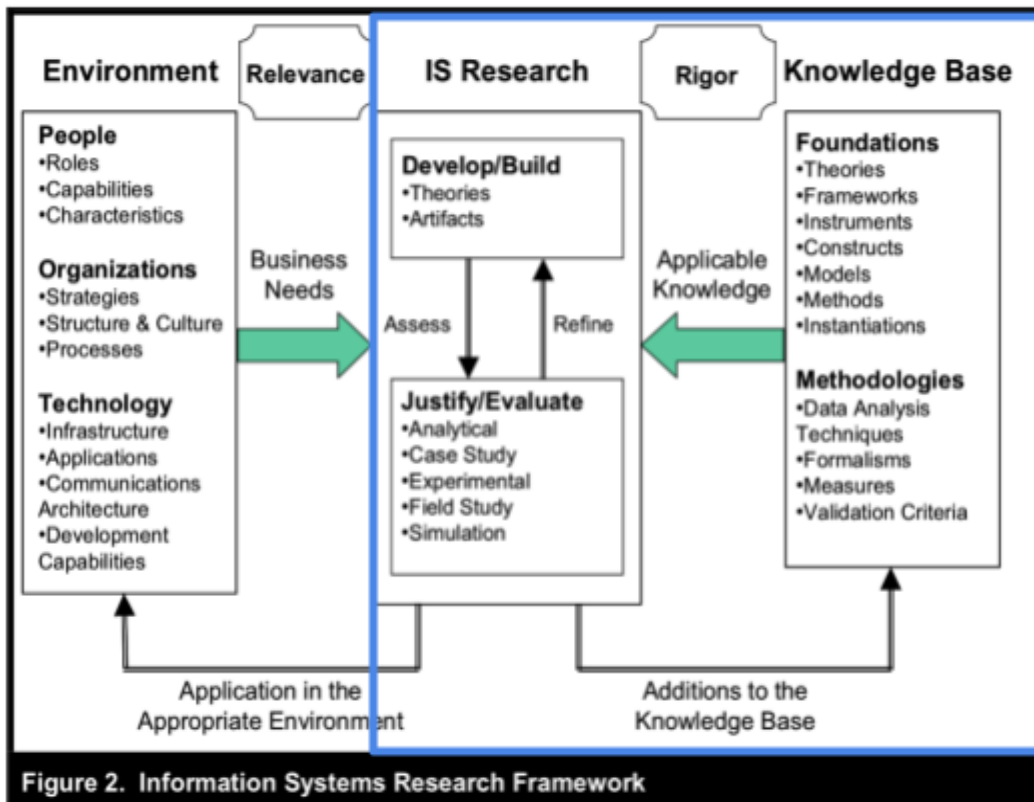


Figure 1.16 : la boucle de rigueur dans les Design Science Research Methodology (Hevner et al., 2004, p.80)

Cette rigueur scientifique dans les phases de conception et d'utilisation consacre un double apport empirique pour les praticiens du domaine étudié. Le premier apport réside dans l'acquisition d'un ou de plusieurs artefacts venant répondre aux problématiques rencontrées. En effet, un des critères de validité en design science research est d'aboutir à un artefact qui, à travers un certain nombre d'itérations, tend à remplir les conditions fixées par les praticiens au début de la recherche ou à s'insérer dans les pratiques de l'environnement adressé du fait de son utilité. Le deuxième apport réside dans la phase de recherche et de conception elle-même, au cours de laquelle le chercheur et ses outils vont occuper "à la fois un rôle de médiation qui supporte l'émergence de compromis socio-techniques et un rôle de facilitateur dans le développement d'un processus d'apprentissage collectif" (Pascal, 2012). En effet, les connaissances opératoires utilisées et créées par le chercheur peuvent être utilisées pour représenter avec un point de vue extérieur, une situation de conception. Dès lors, elles peuvent constituer la base sur laquelle des compromis peuvent émerger entre les différentes parties prenantes de la conception.

Au-delà des apports empiriques, la méthodologie de recherche design science comporte de nombreux bénéfices théoriques pour les sciences de gestion. Pour Aken (2004, p. 220), dans la mesure où les résultats de cette méthode sont prescriptifs et visent à décrire comment faire quelque chose, cette méthode complète les résultats de recherches descriptives qui s'attachent à détailler les problèmes. Ces mêmes problèmes étant justement les points de départ des travaux utilisant l'approche *design science*.

Aussi, les connaissances théoriques permettent-elles d'aborder, a posteriori, la complexité des processus de conception des systèmes d'information modernes :

"l'extrême complexité des systèmes d'information modernes nécessite, pour leur design et construction, un processus itératif dans lequel la documentation décrivant le fonctionnement du système et expliquant pourquoi il fonctionne, a de grande probabilité de voir le jour après la mise en œuvre du système. Les théories créées postérieurement au fait ne sont en aucun cas des formes inférieures de théorie du moment qu'elles répondent aux exigences d'abstraction et de généralisation". Gregor et Jones (2007).

Enfin, les connaissances théoriques permettent de réduire de deux façons différentes l'écart entre le théorie et la pratique en système d'information :

- Soit en répondant à des problématiques que se posent les praticiens, dans un langage qu'ils peuvent comprendre, et ce, en vertu de la première phase de la méthode qui consiste à reconnaître, à travers des discussions avec des praticiens, le problème que la recherche adressera.
- Soit en faisant participer activement et régulièrement les praticiens lors des activités de design, ce qui permet de co-produire les connaissances finales

Ainsi, nous avons appris que l'apport empirique de l'approche *design science* était double :

- Il permet de créer un artefact utile pour la résolution d'un problème pratique ;
- Le chercheur et ses outils jouent un rôle de médiation et de facilitation dont bénéficie les praticiens impliqués dans le recherche.

Aussi avons-nous appris que les 3 apports théorique de cette approche :

1. Les connaissances théoriques prescriptives créés viennent compléter les connaissance descriptives existantes ;
2. Elles permettent d'aborder des questions inaccessibles aux outils méthodologiques traditionnels : le conception de SI modernes innovants ;
3. Elles permettent de réduire l'écart entre le théorie et la pratique, ce qui permet d'améliorer la pertinence et la notoriété de la recherche en systèmes d'information.

1.6.2.2 La construction du cadre théorique propre à l'approche *design science*

Résumé

Après avoir décrit le développement historique ayant abouti à l'approche *design science*, je reviendrai sur les cadres théoriques sollicités pour conforter son assise scientifique. Dans la même logique que la partie précédente, je détaillerai ces cadres théoriques avant d'exposer le basculement ayant permis de passer d'une démarche scientifique empruntant des théories d'autres disciplines à une démarche théoriquement indépendante.

Cette partie est basé principalement sur une réinterprétation des travaux de thèse de Pascal (2006) qui à maints égards ont participé à construire et à faire connaître le *design science* en France. Plus particulièrement, je montrerai en quoi ses travaux retracent le développement théorique de l'approche

design science. En partant d'une méthodologie de recherche-intervention, elle a posé des jalons théoriques qui l'ont mené à bousculer cette méthodologie pour la transformer et la mener aux frontières de ce que nous connaissons aujourd'hui sous le terme de *design science*.

1.6.2.2.1 Attaches théoriques de l'approche *design science*

La majorité des auteurs ayant travaillé sur l'approche *design science* considère l'ouvrage de Herbert Simon (1969), *Les sciences de l'artificiel*, comme l'élément fondateur du corpus théorique qui la soutient. A titre de rappel, cet ouvrage part de l'observation que l'environnement moderne dans lequel nous évoluons est majoritairement constitué d'objets façonnés par les humains, chaque objet remplissant un but précis.

A partir de ce constat Simon établit la distinction entre les éléments artificiels, façonnés par l'homme et les éléments naturels qui émergent d'eux-mêmes. Ce monde artificiel appelle à un nouveau paradigme scientifique qui complète le paradigme des sciences de la nature. Ce paradigme n'a pas pour objectif de décrire les lois qui régissent les éléments naturels. Il s'attache plutôt à comprendre comment les éléments artificiels s'insèrent dans un environnement et à créer des connaissances qui supportent la création d'artefacts incorporant certaines fonctionnalités.

Or, selon Gregor et Jones (2007), si l'ouvrage de Herbert Simon (1969) est le premier à proposer une approche scientifique adaptée aux objets artificiels, les premières attaches théoriques de ce paradigme sont quant à elles bien plus anciennes. Elles remontent à Aristote (IV^e siècle av. J.-C) dans ses quatre explications de ce qu'est une "chose" (Hooker, 1993). Toute chose peut se définir selon Aristote comme :

- *Une causa finalis* : son but ou sa fonction ultime. Par exemple, une table a pour but de pouvoir manger dessus.
- *Une causa formalis* : son essence, ce qu'elle est fondamentalement. La table, par exemple, est une surface plane supportée par des pieds.
- *Une causa materialis* : sa constitution matérielle. La table est faite de bois.
- *Une causa efficiens* : la personne ou la chose qui a façonné cet objet. La table de notre exemple a été faite par un charpentier.

Je comprends dans ces quatre causes l'idée que toute chose est créée par autre chose et ce, de façon intentionnelle. Je peux retenir ici la définition d'un artefact (Heidegger, 1993).

Gregor et Jones (2007) précisent que si il est possible d'identifier une filiation d'auteurs ayant avancé des éléments théoriques qui tendent à supporter l'approche *design science*, aucun d'entre eux ne fait état d'une théorie du design à part entière. La raison en est que les philosophes ont une grande difficulté à établir la véritable nature de la connaissance issue de la technologie et de son lien avec la connaissance scientifique (Bunge, 1979).

Ainsi, les premières attaches théoriques décrites ici établissent le fait que l'approche *design science* s'attache bien plus à concevoir des objets concrets qu'à décrire des phénomènes. Pascal, (2012, [p.10](#)) résume brillamment le corpus théorique dont ont hérité les méthodes de design science et les implications pour les chercheurs souhaitant les utiliser :

“le travail des chercheurs en design science dans le domaine des Systèmes d'Information ne consiste pas (uniquement) dans la recherche de vérités ou de lois universelles mais en la transformation de pratiques existantes en situations désirables via des artefacts” Pascal (2012, p.10)

1.6.2.2.2 L'apport de la théorie de la traduction : l'émergence de supports matériels à la médiation entre parties prenantes d'une recherche scientifique

a. Les Objets Intermédiaires de Conception (OIC) : une représentation partagée de la situation de conception

Définition : Objet Intermédiaire de Conception (OIC) : se définit comme “un objet produit ou utilisé au cours du processus de conception, trace et support de l'action de concevoir en relation avec les outils, procédures et acteurs” (Jeantet, 1998, p.293). Ils permettent de fixer une représentation claire du processus en cours et des connaissances générées et permet aux acteurs impliqués de dialoguer autour.

Le concept d'Objet Intermédiaire de Conception (OIC) est né du constat que les méthodes d'ingénierie utilisées dans la conception n'étaient pas sans poser des problèmes aux entreprises qui devaient collaborer avec de nombreux partenaires (voir 2ème point d'une situation de conception selon Midler (1996), mais aussi les répartir dans les bons dispositifs tout en devant adapter leur propre structure à cette collaboration. C'est dans ce cadre que des objets permettant de créer une représentation partagée de la situation de conception ont été créés.

Pour Papadimitriou (2004), ces objets remplissent deux fonctions principales :

- Une fonction technique qui permet de calculer ou matérialiser les performances techniques du produit ;
- Une fonction sociale qui permet d'assurer la combinaison et la coordination entre les diverses logiques de conception présentes au sein de l'équipe de conception.

Issu des méthodologies de la recherche ingénierique, on retrouve derrière le concept d'OIC les prémisses de ce qui deviendra plus tard un *artefact* en *design science*. Les OIC, par leur rôle de médiateur entre divers acteurs, de facilitateur de relations sociales et de traducteur de logiques cognitives, possèdent les mêmes fonctions que les *artefacts* en *design science*. L'exemple de développement d'une plateforme web de cartographie de compétences décrite dans Pascal et al. (2013) est très intéressant. Elle y explique comment une représentation standardisée (*artefact*) de celle-ci a permis de fluidifier les échanges entre concepteurs et usagers (rôle de médiation) ainsi que de stimuler les apprentissages mutuels au sein de l'équipe (rôle de facilitateur)

b. Les Objets Frontières (OF)

Les objets frontières peuvent quant à eux être représentés comme les ancêtres d'un autre pan des *artefacts* en *design science* : leur capacité à traduire et intégrer les intérêts des diverses parties prenantes d'une recherche.

Le concept d'objet frontière a émergé du constat que tout travail de recherche se trouve dans une situation paradoxale du fait des intérêts divergents qui animent les parties prenantes. Il est rare, par exemple, que les intérêts des praticiens, des chercheurs et des financeurs soient identiques vis-à-vis d'un projet de recherche. Les praticiens chercheront par exemple à transférer les résultats d'une recherche en application industrielle tandis que les chercheurs seront motivés par l'identification de gaps exploitables sous forme de publication scientifique. Néanmoins, malgré ces divergences, ces acteurs doivent collaborer.

Or, "la coopération n'est possible que si un travail de 'traduction' des points de vue divergents permet un intéressement des différents acteurs" selon Pascal (2006, p.213). Elle prend pour illustrer ce travail de traduction l'exemple d'un chercheur en biologie, J.Grinnell, qui avait réussi à convaincre des acteurs

hétérogènes de participer à la constitution d'un muséum zoologique dont le mode de construction lui permet d'alimenter ses propres recherches théoriques. Les objets frontières donnent une forme matérielle à ce travail de traduction : ils sont conçus de façon à intéresser et faire participer des acteurs hétérogènes dans un projet de recherche et ce, en traduisant leurs intérêts divergents sous une forme qu'ils peuvent s'approprier. Pascal (2006) cite Vissac-Charles (1996) pour définir les objets frontières :

“les objets frontières schématisent le travail d’articulation, de coordination de l’action et de simplification du monde. À travers eux, différents acteurs qui n’ont pas obligatoirement les mêmes buts arrivent à se coordonner tout en conservant leur autonomie. Le concept d’objets frontières permet de se débarrasser de l’idée que la coopération passe nécessairement par le consensus et permet ainsi d’expliquer comment on peut gérer à la fois la diversité et la coopération. Les objets frontières sont des connecteurs de réseaux”

Vissac-Charles (1996)

Dans la conception d'Hevner et al. (2004) un artefact en *design science* présente cette même caractéristique. Les artefacts, suivant les objets frontières, doivent intégrer des logiques divergentes. Un artefact en design science doit adopter un formalisme et une rigueur qui permettront à la communauté scientifique d'évaluer son niveau de contribution à la recherche ; prouver sa capacité à répondre à une problématique concrète expérimenté par les praticiens ; utiliser des théories de la conception de façon à ce que d'autres concepteurs puissent s'inspirer du travail de recherche. Ici, les intérêts de 3 communautés d'acteurs sont à prendre en compte et à traduire : la communauté scientifique, les praticiens, les concepteurs d'artefact. Ces similitudes entre les objets frontières et les artefacts en *design science* nous amènent à établir un lien de filiation entre les deux.

La théorie de la traduction, à travers les concepts d'Objets Frontières et d'Objets Intermédiaires de Conception constitue un jalon central sur lequel s'est construit le *design science*. En justifiant l'apport pour la recherche des objets permettant d'intégrer les logiques d'acteurs hétérogènes et de fixer à un moment donné une représentation partagée de la situation de conception, la théorie de la traduction pose la structure théorique qui soutient l'utilisation des *artefacts* en design science dans le domaine des systèmes d'informations. Objets étant à la fois l'objet et le résultat de la recherche et qui, tant dans leur construction que dans leur implantation, produisent des résultats théoriques et pratiques, les artefacts sont le centre de gravité de l'approche *design science*. Un centre de gravité dont la place dans le corpus théorique est assuré par la théorie de la traduction.

1.6.2.2.3 L'apport de la théorie de la structuration : l'instauration des interactions entre concepteurs et utilisateurs

Il faut prendre appui sur les travaux d'Orlikowski (2000) qui utilise la théorie de la structuration de Giddens (1984) pour observer en pratique les interactions entre les individus et les technologies. J'évoque ici l'utilisation de cette théorie de la structuration dans la mesure où elle vient renforcer théoriquement une activité centrale des chercheurs en *design science* : les interactions avec les utilisateurs de l'artefact.

Sa lecture particulière des travaux de Giddens amène Orlikowski à établir une distinction fondamentale sur laquelle s'est appuyée la littérature en design science :

D'un côté, la technologie est un artefact technologique, c'est-à-dire "un ensemble de modules intégrés de logiciels et de matériel qui comporte des caractéristiques prédéfinies" (Pascal, 2006, p.214). Loin d'être neutre, cet ensemble est, selon Orlikowski (2000) l'incarnation de la vision du monde des concepteurs qu'elle nomme "esprit de la technologie". Je prendrai pour exemple le Web tel qu'il a été conçu à son origine par Tim Berners-Lee (1989). Suivant une philosophie propre, le Web devait être à l'origine "inscriptible" dans la mesure où toutes les pages HTML peuvent être modifiées par ceux qui y avaient accès.

De l'autre, la technologie est une technologie-en-pratique, autrement dit un ensemble de "propriétés structurelles énoncées de façon routinière dans l'usage de la technologie par les usagers". Pour reprendre l'exemple du Web, la dimension inscriptible du Web n'a pas perduré dans l'usage et aujourd'hui, les pages HTML ne peuvent être modifiées.

Ainsi, Orlikowski établit que la technologie doit être comprise comme une articulation permanente entre ces deux niveaux. Les concepteurs fixent une structure technologie issue de leur vision du monde ; les utilisateurs énoncent cette structure et révèlent une forme d'usage. A noter ici que l'énonciation de la technologie se distingue de l'appropriation de la technologie. Le premier concept renvoie au fait les utilisateurs des technologies font émerger les règles et ressources inscrites dans la technologie tandis que l'appropriation consiste à adopter des règles et ressources préexistantes et explicites.

Cette vision de la technologie entraîne deux corollaires qui justifient de favoriser les interactions avec les utilisateurs de l'artefact tout au long d'une démarche de *design science* :

Les utilisateurs énant la technologie, ils peuvent faire le choix ou non d'utiliser les fonctionnalités imaginées par les concepteurs et de les manipuler ou non selon les modalités prévues. De façon générale, ce corollaire implique que les fonctionnalités disponibles ne prédéterminent pas l'usage qui serait fait de l'artefact. Par conséquent, il est nécessaire pour un chercheur en *design science* de tester les fonctionnalités de son *artefact* en situation réelle pour aboutir à un ensemble de fonctionnalités réellement utiles.

L'usage de l'artefact est fortement influencé par les idéologies, images ou raisonnements sollicités au préalable par les utilisateurs. Ainsi, il est nécessaire pour le chercheur en design science de comprendre au préalable quels sont les éléments sollicités par les utilisateurs et d'établir dans quelle mesure il s'agit d'adapter l'artefact. L'objectif consiste ici à éviter les conflits avec ces éléments pour assurer l'utilité de l'artefact. On comprend alors pourquoi il est nécessaire d'interagir régulièrement avec les utilisateurs. Sans ce dialogue, le chercheur en *design science* prend le risque de voir son artefact rejeté du fait de sa dissonance avec les éléments cognitifs à disposition des futurs utilisateurs.

Les travaux d'Orlikowski sur la théorie de la structuration apportent un maillon essentiel dans la construction théorique de l'approche *design science*. L'établissement d'une vision duale de la technologie, constituée des articulations entre le monde des concepteurs et celui des utilisateurs, a entraîné toute une réflexion sur les interactions à mettre en œuvre dans le cadre du design d'un artefact. Sur la base de ces éléments, je conçois la théorie de la structuration comme un socle théorique sur lequel repose les boucles de pertinence et boucles de design imaginées par Hevner et al. (2004) et qui permet d'assurer respectivement la pertinence du problème adressé par la recherche et la valeur perçue par les praticiens de l'artefact conçu.

1.6.2.2.4 L'élaboration d'une design theory indépendante pour consolider les recherches de design science

Les travaux de Pascal (2006) que je réinterprète dans cette thèse comme une démonstration de la construction théorique des méthodologies de *design science* s'arrête en dressant quelques pistes de réflexion à propos d'une nouvelle "démarche scientifique de recherche-intervention". Sans mentionner le terme de *design science* et l'article d'Hevner et al. (2004), elle dresse un horizon qui, à mon sens,

préfigure l'élaboration d'une *design theory* qui s'affranchirait de l'héritage de la théorie de la structuration et de la traduction.

a. Les similitudes entre les nouvelles démarches scientifiques de recherche-intervention et la design theory soutenant les démarches de design science : preuve des liens de filiation entre les deux approches.

Je base cette affirmation sur les similitudes évidentes, présentées dans le tableau ci-dessous, entre les cinq composants clés de la nouvelle démarche scientifique évoquée par Pascal (2006) et certains des huit composants d'une *design theory* mis en lumière par l'article central de Gregor et Jones (2007).

Composants d'une nouvelle démarche scientifique (Pascal, 2006)	Composants d'une <i>design theory</i> (Gregor et Jones, 2007)
<p>Une science des organisations</p> <p>Elle opère comme une trame de fond théorique permettant d'expliquer les processus et résultats organisationnels impliqués dans la construction d'un système</p>	<p>Des connaissances justificatives</p> <p>Des théories ou connaissances sous-jacentes, provenant des sciences naturelles, sociales ou de design, qui donnent une base et une explication à la construction.</p>
<p>Des principes de construction</p> <p>« <i>Un ensemble cohérent de propositions, bâties sur l'état de l'art en sciences des organisations, pour produire de nouvelles conceptions et formes organisationnelles et redévelopper celles existantes</i> » (Romme et Endenburg, 2006, p.288 cité par Pascal, 2006, p.252)</p>	<p>Principes de forme et de fonction</p> <p>Un schéma directeur ou des principes d'architecture pour décrire et produire la forme et la fonction finale d'un artefact.</p>
<p>Règles de design</p> <p>Règles pratiques et heuristiques inspirées des</p>	<p>Principes d'implémentation de la théorie</p> <p>Une description des processus permettant</p>

principes de construction permettant d'assister la construction du système organisationnel	d'implémenter la <i>design theory</i> dans des contextes de conception particuliers.
Design organisationnel Développement d'un système organisationnel à partir des règles de design	Instanciation descriptive Une instanciation physique de l'artefact permettant de représenter la théorie et ce, de manière à pouvoir la tester ou l'expliquer.
Implémentation et expérimentation Implémentation du système et test des résultats engendrés	Hypothèses théoriques Affirmations relatives à la <i>design theory</i> pouvant être testées.

Figure 1.17 : représentation des similitude entre les composants de la nouvelle démarche scientifique évoquée par Pascal (2006, p. 241) et ceux de la design theory (Gregor et Jones, 2007, p.322)

Je détaille ci-après quelques unes de ces similitudes. Entre la science des organisations citée par Pascal (2006) et les connaissances justificatives évoquées par Gregor et Jones (2007), on retrouve en substance la même idée qui consiste à donner au design un cadre théorique initial. Ce dernier permet d'ancrer le design dans un univers de connaissance donnée, sans lequel le caractère scientifique de la démarche ne peut exister.

Encore, les règles de design et les principes d'implémentation de la théorie recèlent une démarche commune qui consiste à traduire des éléments théoriques (principes de construction pour Pascal (2006), *design theory* pour Gregor et Jones (2007) pour les incorporer dans une démarche de design d'un objet matériel.

Enfin, l'implémentation et l'expérimentation d'un côté et les hypothèses théoriques de l'autre sont reliées par le même principe, aujourd'hui essentiel aux démarches de *design science*, qui impose que les éléments théoriques qui ont été traduits dans une démarche de conception soient testés empiriquement.

Ces similitudes sont à mon sens la preuve que la *design theory* trouve ses racines dans les théories de la traduction et de la structuration dépeintes par Pascal (2006). Leur absence dans les travaux qui ont fait

naître l'école de la *design theory* en système d'information (Carlsson, 2010) doit à mon sens être corrigée dans la mesure où les liens de filiations se prouvent aisément. Aussi en réinvestissant ces deux théories dans leur rôle de socle théorique initial, cette thèse pourrait-elle apporter une contribution aux travaux sur la *design theory*.

Cela étant, je présenterai la *design theory* qui s'est détaché du développement historique que je dépeins dans cette partie avant d'établir quelques pistes qui permettrait de la rattacher aux théories de la traduction et de la structuration

b. L'anatomie d'une design theory indépendante pour appuyer un nouveau mode de recherche

Avant tout, précisons le périmètre de la *design theory* et les relations qu'elle entretient avec les phénomènes réels. La *design theory* a pour objectif et sous-tend le design aussi bien de méthodologies que de produits. Néanmoins, il s'agit de bien séparer ces outputs de la *design theory* qui existent dans le monde réel et la *design theory* en elle-même, qui est une pure entité abstraite. Le schéma ci-dessous extrait de Gregor et Jones (2007) représente plus précisément la nature des relations entre la *design theory* et les outputs matériels dit artefacts.

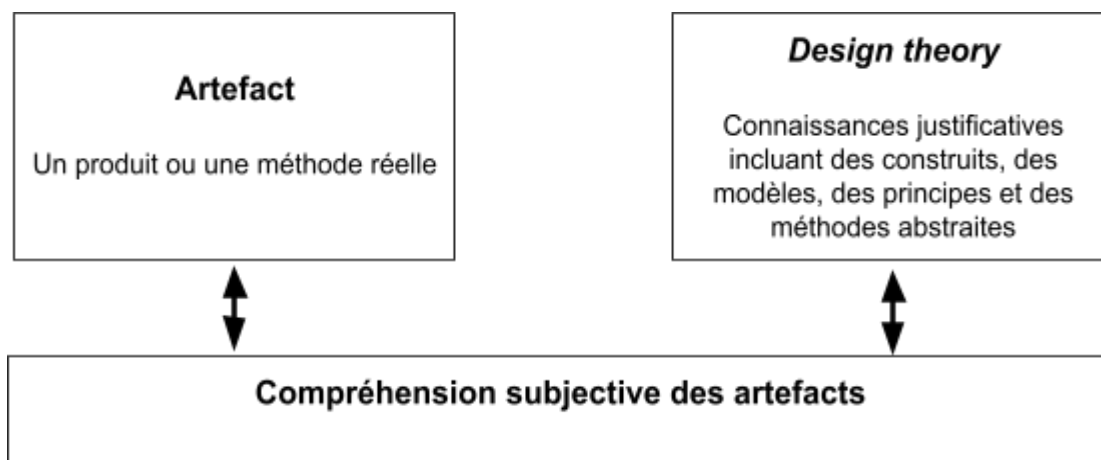


Figure 1.18 : Les relations entre la design theory et les artefact

Comme indiqué sur le schéma, la compréhension subjective constitue l'élément qui connecte la *design theory* aux artefacts. Les deux flèches symbolise le fait que les individus créent des *design theories* pour

conceptualiser et décrire des artefacts avec des termes abstraits et s'en servent pour la construction d'un artefact ; symétriquement, les êtres humains, par induction, peuvent tirer de l'observation du fonctionnement d'un artefact des construits, modèles et méthodes qui constitueront finalement une *design theory*.

Pour appuyer cette distinction, Gregor et Jones (2007) font un parallèle avec les trois mondes de Popper (1986) :

1. Le premier monde, correspondant aux artefacts, qui est le monde objectif des choses matérielles. Par exemple, une base de données relationnelles
2. Le deuxième monde, correspondant à la compréhension subjective des artefacts, qui est celui des états mentaux subjectifs. Par exemple, la compréhension qu'un chef de projet open data aura d'une base de données relationnelles, lui permettant d'importer ou d'exporter les données qu'elle contient.
3. Le troisième monde, correspondant à la *design theory*, qui est aussi un monde objectif mais abstrait cette fois-ci. Il est constitué de connaissances, de sciences et d'institutions. C'est le monde de règles qui ordonne le cours du premier et du deuxième monde.

Avec ces éléments, je peux conclure en citant Gregor et Jones (2007, p. 322) qu'une *design theory* :

“Expose les principes inhérents au design d'un artefact devant accomplir certains objectifs, et ce en se basant sur des connaissances à propos du comportement des Technologies de l'Information (artefacts) et des humains (compréhension subjective). Aussi cette théorie s'étend elle jusqu'à des prescriptions appliquées au design d'artefact” Gregor et Jones (2007, p.322)

Au sein du périmètre que nous venons de dresser, une *design theory* comporte huit éléments que nous décrivons ci-dessous :

1. **Un objectif et un périmètre** : les raisons d'être ou les pourquoi de l'artefact qui définissent également les frontières de la *design theory*.
2. **Des construits** : des représentation des entités réelle impliquées dans la *design theory*
3. **Principes de forme et de fonction** : un schéma directeur ou des principes d'architecture pour décrire et produire la forme et la fonction finale d'un artefact.

4. **La mutabilité de l'artefact** : les changements d'état de l'artefact tels qu'ils sont anticipés par la *design theory*. Autrement dit, il s'agit du degré de transformation que peut subir l'artefact au cours de son usage supposé.
5. **Hypothèses théoriques** : les affirmations relatives à la *design theory* pouvant être testées.
6. **Des connaissances justificatives** : des théories ou connaissances sous-jacentes, provenant des sciences naturelles, sociales ou de design, qui donnent une base et une **explication** à la construction.
7. **Principes d'implémentation de la théorie** : une description des processus permettant d'implémenter la *design theory* dans des contextes de conception particuliers.
8. **Instanciation descriptive** : une instanciation physique de l'artefact permettant de représenter la théorie et ce, de manière à pouvoir la tester ou l'expliquer.

Ces huit éléments montrent bien en quoi la *design theory* s'est émancipé des autres théories que l'on peut trouver en sciences de gestion dont les théories de la structuration et de la traduction. Il ne s'agit ni de décrire, ni de prédire, ni d'analyser, ni d'expliquer mais d'assister la conception d'artefact matériels. Il n'est donc pas simplement question de contribuer à alimenter une littérature mais aussi de s'impliquer dans le monde réel par l'intermédiaire de ces artefacts.

1.6.2.3 Les débats et controverses autour des méthodologies design science aujourd'hui

Résumé

La partie précédente m'a permis de mettre en lumière la nature et les composantes principales de l'approche *design science*. Ensemble, ils forment le coeur de l'approche. Cela étant, l'approche *design science* demeure jeune et de nombreuses interrogations doivent encore trouver des réponses avant qu'elle puisse acquérir ses lettres de noblesse auprès de la communauté scientifique :

- *Quelle(s) posture(s) épistémologique(s) adopter avec cette approche ? (1)*
- *En quoi cette approche se distingue-t-elle de la recherche-action qui, comme elle, a pour but de créer des connaissances avec et pour les praticiens ? (2)*
- *à quel(s) types de résultats scientifiques une telle démarche peut-elle aboutir ? (3)*
- *Quelles sont les limites de son périmètre d'étude, les systèmes d'information ? (4)*

Autant de questions qui font l'objet de controverses et qui alimentent aujourd'hui son développement dans le domaine des systèmes d'information. Cette partie vise justement à présenter ces controverses, les parties qui s'y opposent et la position de ma thèse en leur sein.

1.6.2.3.1 L'articulation entre méthodologie et épistémologie : la place controversée de l'approche design science

Résumé

Le premier objet de débat dans la littérature sur l'approche *design science* a trait à son articulation avec les postures épistémologiques. En effet, on trouve dans la littérature un panel de postures épistémologiques différentes sans que la question de leur compatibilité avec l'approche *design science* n'ait été posée. Cette partie est l'occasion de présenter ces diverses postures épistémologiques avant de prendre position en faveur de l'une d'entre elles, spécialement conçue pour s'imbriquer avec l'approche *design science*.

a. Les diverses acceptions de l'approche design science en fonction de la posture épistémologique choisie

À l'origine, les méthodologies de design science faisaient l'économie de se positionner au niveau de l'épistémologie. Les premiers travaux d'Hevner et al. (2004) et de Vaishnavi et Kuelcher (2004) sur méthodologies *design science* dans la revue *Management Information Systems Quarterly* ne font pas état d'un positionnement épistémologique particulier. Au mieux, les derniers évoquent une position ontologique et épistémologique mouvante chez le chercheur en design science :

"Les points de vue ontologiques et épistémologiques changent dans la recherche sur la conception au fur et à mesure que le projet suit des cycles de circonscriptions ... Cette itération est similaire mais plus radicale que les processus herméneutiques utilisés dans certaines recherches interprétatives". Vaishnavi et Kuelcher (2004)

Il faut attendre Carlsson (2010) pour rappeler qu'une démarche scientifique se doit de donner clairement et définitivement sa conception du monde et de la connaissance. Ce ne sont pas des options. Il en profite d'ailleurs pour pointer du doigt, reprenant les travaux d'Arnott et Pervan (2008), que les premiers auteurs en *design science* adoptent majoritairement une posture positiviste : 92.3 % des études empiriques analysées relèvent très clairement d'un positionnement positiviste.

S'opposant à cette vision monolithique du *design science*, il ouvre les perspectives et affirme que le *design science* peut découler d'autres sources épistémologiques et ontologiques, donnant alors des acceptions différentes des méthodologies *design science*. Pascal (2012) capitalise sur cette ouverture, énumère les paradigmes épistémologiques utilisés en *design science* et décrit "la nature des connaissances mobilisées, les méthodologies de recherche et les processus d'évaluation" de chacun d'entre eux :

- D'un positionnement épistémologique positiviste découlera une recherche de connaissances positives sur l'organisation et les artefacts. La méthodologie utilisée consistera alors à concevoir des artefacts qui, par leurs propriétés internes, s'intégreront dans l'organisation cible et produiront les effets escomptés. L'évaluation des résultats sera alors uniquement centré autour de l'artefact l'informatique.

Exemple d'une recherche positiviste en *design science*

Vaishnavi, V., Buchanan, G., and Kuechler, W. (1997). A Data/Knowledge Paradigm for the Modeling and Design of Operations Support Systems, *IEEE Transactions on Knowledge and Data Engineering*, 9(2), 275–291.

- Le positionnement épistémologique réaliste critique est à l'opposé du positivisme en ce qu'il considère que la science sociale est indéterminée et que les résultats d'une démarche dépendent majoritairement du contexte. Dès lors, un plus grand intérêt est porté à l'étude du contexte d'implémentation d'un artefact. L'artefact est construit autour de ce contexte et ses propriétés sont conçues dans l'optique d'une interaction avec un système socio-technique spécifique. *In fine*,

c'est au regard des nouvelles pratiques liées à l'implémentation de l'artefact que les résultats de la recherche sont évalués.

Exemple d'une recherche réaliste critique

Carlsson, S. A. and T. Kalling (2007) Implementation of a knowledge management initiative: turning kms use into profit, in *Proceedings of the 4th International Conference on Intellectual Capital, Knowledge Management and Organisational Learning (ICICKM)*.

- Le positionnement épistémologique constructiviste s'avère être le plus radical en ce qu'il considère qu'aucune connaissance indépendante de la relation entre le chercheur et son objet de recherche n'est possible. La connaissance est nécessairement construite, voire co-construite avec son terrain. Dans ce cadre, "le rôle de la connaissance n'est pas de construire des représentations proches de la vérité mais plutôt d'élaborer des constructions qui s'insèrent et fonctionnent dans le cadre de la relation chercheur-terrain" (Avenier, 2010, p.32). L'évaluation des artefacts n'est ici possible que dans un seul contexte quand le paradigme réaliste critique suppose une multitude de contextes similaires dans lesquels un artefact peut s'insérer avec les mêmes résultats.

b. L'idée d'un nouveau courant épistémologique indépendant dit design pour s'articuler avec les méthodologies de design science

Jusqu'à présent, nous avons dressé un panorama des postures épistémologiques existantes et utilisées dans la littérature sur le *design science*. Or, des auteurs comme Romme (2003, 2004) ou Romme et Edenburg (2006) considèrent que le *design science* n'est pas qu'une méthodologie dont les tenants et aboutissants varient au gré des postures épistémologiques. Ils affirment que le *design science* est un paradigme épistémologique à part entière. Autrement dit, le *design science* suppose une vision indépendante de la connaissance.

Pour distinguer cette posture épistémologique de toutes les autres citées dans la partie précédente, Romme (2003) établit une distinction entre trois courants épistémologiques utilisés dans les sciences de gestion :

- Le courant dit *science* qui renvoie au paradigme positiviste évoqué dans la partie précédente. Toute recherche qui s’inscrit dans ce sillage vise à extraire les lois qui animent des objets dotés de propriétés descriptives. Pour cela, le chercheur s’impose de forte contrainte pour s’assurer de construire des connaissances objectives et non déterminées par des biais personnels.
- Le courant dit *humanities* que l’on peut rattacher au paradigme constructiviste. Celui-ci ne vise plus à découvrir la vérité mais plutôt à comprendre et décrypter les pratiques organisées et ce, en se basant sur l’expérience des acteurs. Le chercheur en sciences de gestion utilisera alors des procédés narratifs pour rendre compte du fonctionnement de ces expériences et donner au lecteur la possibilité d’apprécier la complexité des phénomènes.
- Le courant dit *design science* qui opère un changement radical selon Pascal (2006) :
“L’idée sous-jacente au design consiste à investir des systèmes qui n’existent pas encore. Il y a un changement total de perspective puisque l’acteur ne cherche plus à répondre à la question « est-ce fondé ou vrai » mais plutôt « cela va-t-il marcher ». La connaissance est donc fondamentalement pragmatique c’est-à-dire au service de l’action”

À mon sens, ce courant explore un nouveau territoire pour les sciences de gestion dans lequel la conception de nouveaux artefacts et systèmes prédomine. La connaissance disponible, qu’elle émane des acteurs ou de la littérature est ici utilisée en *input* d’une démarche de conception de systèmes qui n’existent pas encore et se voit transformée au gré d’évaluations des systèmes sur le terrain. Que le système marche ou qu’il ne marche pas est ici un indicateur qui permet de construire de nouvelles connaissances venant s’intégrer au corpus existant.

Conclusion

Ma thèse s’inscrit dans ce troisième courant épistémologique dans la mesure où il vient parfaitement répondre la situation problématique dans laquelle se trouve aujourd’hui les praticiens et les chercheurs qui s’intéressent à l’ouverture des données : à savoir le manque d’usage des données, et donc le manque de services ou applications innovants qui crée de la valeur à partir de l’ouverture des données.

1.6.2.3.2 La complémentarité entre recherche-action et design science en débat

Résumé

En parallèle du débat sur l'articulation avec l'épistémologie, l'approche *design science* fait l'objet de controverses relatives à la nature de leurs relations avec la recherche-action. Du fait que ces méthodologies ont toutes les deux été développées pour créer de la connaissance avec les praticiens, une certaine confusion règne sur les traits qui les distinguent. Pour certains auteurs, les méthodologies *design science* sont une extension de la recherche-action tandis que d'autres prônent l'indépendance des premières. Je présenterai ces différentes conceptions avant de choisir et de justifier ma position dans ce débat.

a. La recherche-action comme guide des recherches sollicitant les méthodologies design science

Le camp qui prend position contre l'indépendance des méthodologies *design science*, représenté dans l'article de McKay et Marshall (2005), utilise un argument radical : telle que présentée par les auteurs américains comme Hevner, March ou Power, la méthodologie *design science* n'est pas une démarche scientifique. En effet, ces auteurs ont avant tout présenté le design science comme une méthode de résolution de problèmes identifiés par des praticiens. Or, McKay et Marshall expliquent avec éloquence (2005, p. 6) que le but de la recherche doit être avant tout de conduire des recherches sur un sujet et non de construire, développer et évaluer des artefacts.

Pour adresser cette confusion et préciser ce qui distingue la résolution de problème d'une démarche scientifique, McKay et Marshall proposent de s'appuyer sur les solutions proposées par la recherche-action, qui au cours de son développement a dû adresser la même controverse (McKay et Marshall, 2001). Ce courant a effectivement su développer au fil du temps une articulation équilibrée entre les deux pôles :

"la recherche-action doit être considérée comme comprenant deux cycles ou ensembles d'intérêts interdépendants, l'intérêt pour la résolution de problèmes et l'intérêt pour la recherche, qui sont quelque peu dépendants l'un de l'autre et doivent chacun progresser de manière satisfaisante grâce à l'intervention de recherche-action". McKay et Marshall (2005, p.6)

Il s'agit aussi de préciser que cette articulation n'est possible qu'au niveau analytique et conceptuel. Au niveau pragmatique, c'est-à-dire dans le quotidien d'un chercheur utilisant la recherche-action, la résolution de problème et la construction de connaissances sont entremêlées et indistinguables.

Sans emprunter à la recherche-action cette distinction analytique et conceptuelle, les méthodologies de *design science* seraient au mieux un moyen pour les chercheurs d'acquérir de l'expérience personnelles dans le conception d'artefact. Sans une volonté délibérée et sans la mise en place a priori de mécanismes permettant de générer et contrôler la création de connaissances, les méthodologies *design science* sortent du périmètre de la science. Alors et dans la mesure où le courant de la recherche-action a conceptualisé ces activités particulières de création de connaissances, il est tout à fait à propos de reprendre ses principes et d'en faire des guides de recherches. Telle est la première position, sensée et pragmatique, dans le débat sur l'indépendance des méthodologies en *design science*.

b. Concevoir les méthodologies design science indépendamment de la recherche-action

Les premières briques de l'édifice hétéroclite des méthodologies *design science* ont été apportées par Carlsson (2007) qui suggère de créer un nouvel horizon de recherche indépendant, attaché à consolider le développement de connaissance en design de système d'information. Il dépasse ainsi la vision réductrice de l'approche *design science* qui, avant ses travaux, restait cantonnée au développement d'artefact technologique ou à des théories de développement d'artefact spécifiques. Carlsson amorce ainsi une véritable réflexion sur la construction d'une méthodologie *design science* générique pour guider et contrôler la construction d'un nouveau type de connaissances.

Il établit ainsi qui sont les destinataires de ce nouveau type de connaissance, les besoins en connaissance de ces destinataires et les caractéristiques de ce type de connaissance. Les connaissances en design science s'adressent aux praticiens des systèmes d'information impliqués dans :

“la planification, le management, le design et la gouvernance, la construction, l'implémentation, l'opérationnalisation, la maintenance et l'évaluation d'initiatives en système ou en technologie de l'information” (Carlsson, 2007, p.77).

Ce groupe de praticiens cherche avant tout à obtenir deux types de connaissance :

1. Des connaissances sur le développement de technologie ou de système d'information pour répondre à des problématiques organisationnelles.
2. Des connaissances sur les moyens d'implémenter ces solutions dans un contexte organisationnel particulier.

Enfin, ces praticiens ont besoin d'une connaissance qui soit pratique et qui permette de mettre en place sur le terrain ce qu'il appelle des initiatives en système d'information. Il définit une initiative en système d'information comme "le design et l'implémentation d'une intervention dans un système socio-technique où les systèmes et technologies de l'information constituent des ressources critiques pour atteindre les résultats espérés". Par cette définition Carlsson sous-entend que les connaissances en *design science* ne sont pas centrées sur les artefacts technologiques mais aussi sur les organisations et individus qui conçoivent et implémentent les artefacts.

Bien qu'ils aient le mérite d'avoir initié le processus de construction d'une méthodologie indépendante, les travaux de Carlsson relèvent plus de l'ordre de l'intuition que de la science. De plus, ils ne donnent pas d'indications précises sur le processus à suivre pour construire des connaissances en *design science*. Or, une méthodologie est en premier lieu :

"un système de principes, de pratiques et de procédures appliquées à la construction d'une branche spécifique de la connaissance" (Archer, 1979, p. 57).

Les travaux de Peffers et al. (2007) comblent ce gap laissé par Carlsson, détaillent les composants d'une méthodologie en *design science* et prouvent en quoi elle se distingue de la recherche action ou de la recherche ingénierique. Il apporte cette contribution capitale en proposant trois éléments :

1. Un processus nominal qui, au-delà de donner des indications sur la façon de mener une recherche en *design science*, suggère la bonne manière de le faire. Par conséquent, ce processus nominal vient renforcer les pratiques d'évaluation de recherche en *design science*.
2. Un périmètre de littérature sur lequel les méthodologies doivent se baser et capitaliser.
3. Un modèle mental représentant les caractéristiques que doivent contenir les résultats d'une recherche en *design science*.

Le processus nominal propose une série d'activité normative pour le chercheur en *design science* (figure ci-dessous) :

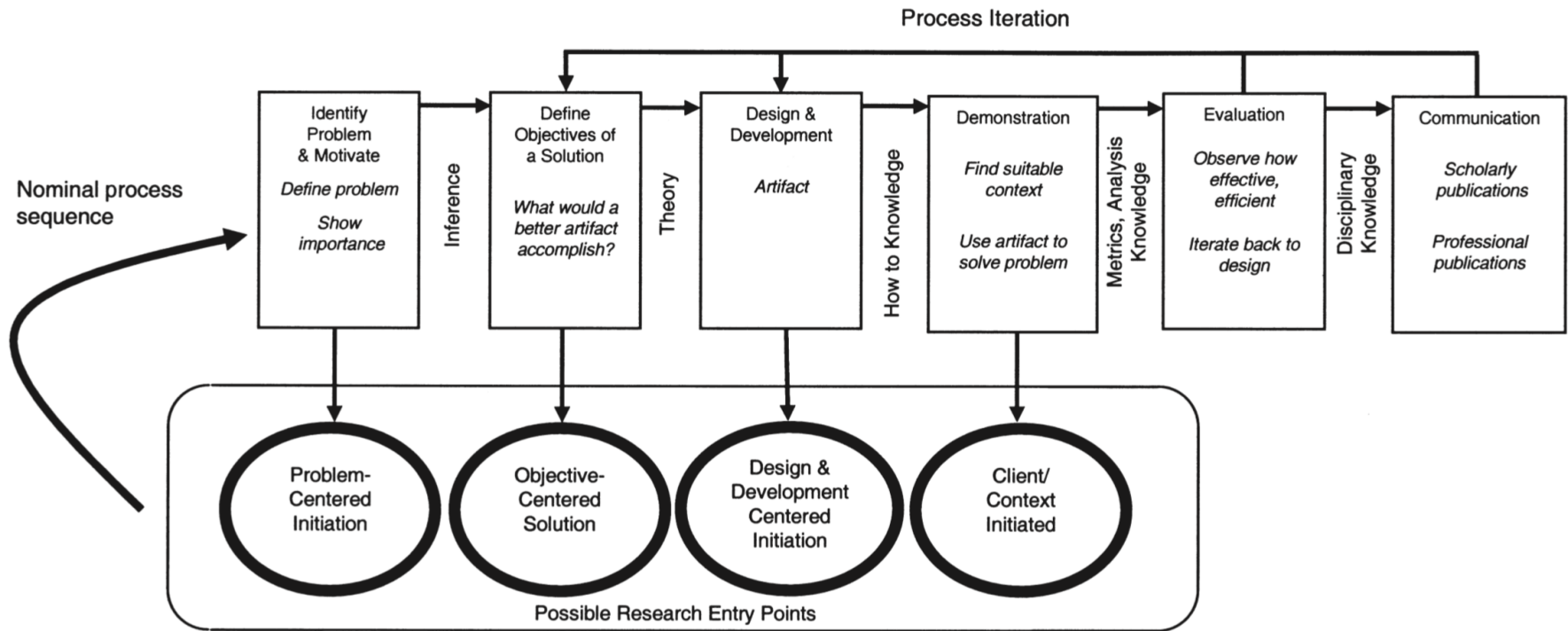


Figure 1.19 : Processus nominal comportant les différentes activités

À propos du périmètre de la littérature qui supporte les méthodologies *design science*, Peffers et al. (2007) établissent deux secteurs :

1. La littérature sur le *design research* qui contient les références expliquant les processus à mettre en place pour le design d'artefact dans des contextes particuliers. Cette littérature inclut des travaux de sciences ingénieriques qui se sont intéressés à l'articulation entre les différents types d'artefact et des processus de conception spécifiques ; des travaux en informatique sur le recensement des besoins métiers vis à vis d'un artefact informatique ; enfin et logiquement, des travaux en système d'information sur la construction d'une vision fonctionnelle d'un artefact par exemple.
2. La littérature sur la recherche en *design research* qui contient de nombreux éléments sur la façon de mener une recherche en *design science*. Ici, Peffers et al. (2007) expliquent que cette littérature ne fournit pas de méthodologies exhaustives mais plutôt des concepts que lesquels les chercheurs peuvent s'appuyer :
 - Des modèles abstraits décrivant la conduite d'une recherche (Nunamaker et al., 1990)
 - Des théories sur le design de système d'information (Walls et al., 1992)
 - Des lignes directrices pour conduire une recherche (Hevner et al., 2004)

Enfin, le modèle mental se définit de la façon suivante :

"modèle" de la réalité à petite échelle... qui] peut être construit à partir de la perception, de l'imagination ou de la compréhension du discours. Les [modèles mentaux] s'apparentent aux modèles des architectes ou aux diagrammes des physiciens en ce sens que leur structure est analogue à la structure de la situation qu'ils représentent, contrairement, par exemple, à la structure des formes logiques utilisées dans les théories des règles formelles" Johnson-Laird et Byrne (2000)

Ici, la réalité que représente le modèle mental sont les résultats issus d'une méthodologie *design science*. Ce modèle mental est représenté par les lignes directrices d'une recherche en *design science* mises en avant par Hevner et al. (2004) et que Peffers et al. (2007) interprète comme étant les composantes obligatoires des résultats d'une recherche en design science (figure ci-dessous).

Lignes directrices	Description
Ligne directrice 1 : La conception d'un artefact	La recherche en <i>design science</i> doit produire un artefact viable sous la forme d'une construction, d'un modèle, d'une méthode ou d'une instanciation.
Ligne directrice 2 : Pertinence du problème	L'objectif de la recherche en <i>design science</i> est de développer des solutions technologiques à des problèmes commerciaux importants et pertinents.
Ligne directrice 3 : Évaluation de la conception	L'utilité, la qualité et l'efficacité d'un artefact de conception doivent être rigoureusement démontrées par des méthodes d'évaluation bien exécutées.
Ligne directrice 4 : Contributions à la recherche	Pour être efficace, la recherche en <i>design science</i> doit apporter des contributions claires et vérifiables dans les domaines de l'artefact conçu, des fondements de la conception et/ou des méthodologies de conception.
Ligne directrice 5 : Rigueur de la recherche	La recherche en <i>design science</i> repose sur l'application de méthodes rigoureuses tant dans la construction que dans l'évaluation de l'artefact de design.
Ligne directrice 6 : Le design en tant que processus de recherche	La recherche d'un artefact efficace nécessite l'utilisation des moyens disponibles pour atteindre les objectifs souhaités tout en satisfaisant aux lois de l'espace-problème
Ligne directrice 7 : Communication de la recherche	La recherche <i>design science</i> doit être présentée de manière efficace à des publics aussi bien orientés vers la technologie que vers la gestion.

Figure 1.20 : Lignes directrices en design science (Hevner et al., 2004, p.83)

Selon Peffers et al. (2007), le trait qui valide définitivement l'indépendance de cette méthodologie par rapport à toutes les autres constitue le fait qu'elle repose sur une littérature spécifique qui se revendiquait jusque-là d'une approche *design science* en système d'information sans pour autant utiliser une méthodologie propre. Avec Peffers et al. (2007), les méthodologies *design science* sont bel et bien constitués d'un corpus scientifique, d'un processus de mise en place clair et d'un autre processus évaluant la qualité des connaissances qui en est issue.

Conclusion
<p>Cette thèse s'oppose aux auteurs qui préconisent de s'appuyer sur les canons méthodologiques de la recherche-action pour deux raisons principales :</p> <p>D'un côté, la recherche-action se préoccupe de créer des changements organisationnels. Or, à l'instar de</p>

Carlsson, je considère que la recherche en design science a pour objectif de développer des artefacts matériels et des connaissances pratiques dont la capacité à provoquer le changement a été prouvé scientifiquement avant de les mettre à disposition des praticiens. Mais il n'est pas question de s'impliquer dans le changement.

De l'autre, la littérature scientifique dispose aujourd'hui de méthodologies de *design science* suffisamment matures pour être utilisées seules, à l'image de celles proposées par Pascal (2012) et Carlsson (2006, 2007)

1.6.2.3.3 La finalité des recherches en design science

Résumé

La controverse sur la finalité des recherches en design science occupe une place particulière dans la littérature du fait qu'elle questionne directement la validité scientifique des méthodologies de *design science*. De fait, c'est en grande partie à l'aune de la forme de ses résultats qu'on détermine la robustesse d'une science.

Dans cette controverse s'oppose d'un côté l'école de la design theory en système d'information, portée par des auteurs comme Walls et al. (1992) ou Gregor et Jones (2007), pour lesquels la formulation de nouvelles théories constitue la finalité nécessaire des méthodologies en *design science*. Une conception que je qualifie d'étroite dans la mesure où la théorie constitue la seule contribution scientifique possible des méthodologies en *design science*. De l'autre, l'école anglo-saxonne de la recherche en design science qui adopte un point de vue plus large sur la question et qui considère "qu'est contribution tout ce qui accroît la base de connaissance du domaine : approfondissement théorique, méthodologique ou nouvelle expérience acquise par la recherche" (Pascal et al., 2012)

a. Une conception étroite de la finalité des recherches en design science : la construction de théories nouvelles comme contribution essentielle

Reprenant les travaux de Cross (2001), Gregor et Jones (2007) fondent un argumentaire solide en faveur de cette conception étroite de la finalité des recherches en *design science*.

Le premier élément de cet argumentaire consiste à soutenir que la construction de nouvelles théories est le seul élément qui peut venir distinguer d'un côté les activités quotidiennes de design conduites par l'industrie du design, et les actions de design menées par les chercheurs de l'autre.

Les activités de l'industrie du design consiste à concevoir de nouveaux produits en utilisant les connaissances que les produits précédents contiennent sur la forme idéale d'un produit. Une recherche en *design science* serait, selon Gregor et Jones (2007) en tout point similaire aux pratiques de l'industrie si elle faisait de l'amélioration d'artefacts matériels son centre névralgique. Pour s'en distinguer et acquérir sa véritable dimension scientifique, la recherche en *design science* doit se baser et participer à la création de connaissances plus formelles et rigoureuses, dépassant ainsi les connaissances contenues dans les produits matériels. Ces connaissances formelles et rigoureuses constituent la théorie que les travaux en *design science* doivent construire.

La deuxième élément de l'argumentaire en faveur de la conception étroite consiste à dire que seule la construction de nouvelles théories permettrait de renforcer la place et la légitimité du *design science* en systèmes d'information comme discipline scientifique. A côté des sciences naturelles et des sciences sociales et de leurs théories descriptives, explicatives ou prédictives, les systèmes d'information se positionneraient en proposant des théories prescriptives pour l'action et la conception d'artefact. Sans elles, il est impossible de garantir la pérennité du *design science* en systèmes d'information au sein du paysage scientifique.

b. Une conception large de la finalité des recherches en design science : les construits, modèles et méthodes comme contributions potentielles

Face aux arguments la conception étroite présentés ci-dessus, l'école anglo-saxonne de la recherche en *design science* oppose une conception plus large de la finalité des recherches en *design science*. Pour elle, la construction de nouvelles théories n'est en aucun cas une contribution obligatoire pour une recherche. S'il est nécessaire pour cette école d'utiliser des connaissances théoriques en préambule de la recherche pour la positionner dans un champ scientifique, et au cours de la recherche pour assister le design et le prototypage d'artefact , l'*output* de la recherche peut revêtir toutes les formes qui permettent d'accroître la base de connaissance du domaine : construit, modèle, méthode, artefact ou expérience.

Pour les auteurs qui prônent cette conception (March et Smith, 1995 ; Hevner et al., 2004), il s'agit de défendre l'idée selon laquelle le *design science* est avant tout une nouvelle activité de recherche qui s'affranchit des considérations théoriques pour mettre l'accent sur la construction de connaissances pertinentes pour les praticiens d'un domaine. Ces auteurs s'inscrivent dans la lignée du constat de Benbasat et Zmud, 1999 selon lequel une trop grande importance donnée à la rigueur théorique est à l'origine de la perte de pertinence des travaux en systèmes d'information et du manque d'intérêt de la part des sphères économiques. Il s'agit donc, à travers cette conception large, de promouvoir les travaux qui, sans aller jusqu'à une contribution théorique, proposent des construits, modèles, méthodes, artefact ou expérience qui pourront être utiles aussi bien aux chercheurs qu'aux praticiens.

Conclusion

Ma thèse s'inscrit dans la deuxième perspective du fait que le champ de recherche qui s'intéresse à l'open data est particulièrement jeune et immature. Aujourd'hui, selon une des dernières revues de littérature sur le sujet publiée en 2018 (Hassan et Twinomurinzi, 2018), il n'existe aucun cadre théorique établi dans la littérature comme base sur laquelle il est possible d'établir une autre théorie, propre au domaine. Il serait trop ambitieux de prétendre à la construction d'une théorie nouvelle sachant que les théories sous-jacentes qui pourrait la faire émerger n'ont pas encore été identifiées.

1.6.2.3.4 Les limites du concept de système d'information

Résumé

Si la controverse sur les finalités des méthodologies questionne leur validité scientifique, celle sur les limites du concept de système d'information interroge leur périmètre d'action. En effet, la géométrie du concept de système d'information varie en fonction des auteurs et avec elle l'étendue des phénomènes adressés par les méthodologies de *design science*. A nouveau, on retrouve une opposition entre deux positions : les tenants d'une conception technologique du concept de système d'information d'une part ; leurs détracteurs qui défendent une conception systémique du concept d'autre part.

a. Une conception technologique du concept de système d'information

A l'image des travaux de Walls et al. (1992), March et Smith (1995) et Hevner et al. (2004), une partie de la littérature sollicitant les méthodologies de *design science* emprunte une conception technologique du concept de système d'information. Celle-ci restreint la définition d'un système d'information aux artefacts technologiques qui se présentent comme leur matière première évidente. On retrouve ici la position de Benbasat et Zmud (2003) pour lesquels ces artefacts technologiques et leur réseau nomologique (cf figure ci-dessous), que je définis comme l'ensemble des phénomènes naturellement liés aux artefacts technologiques (conception, usage, impact, etc.), constitue le coeur du concept de système d'information.

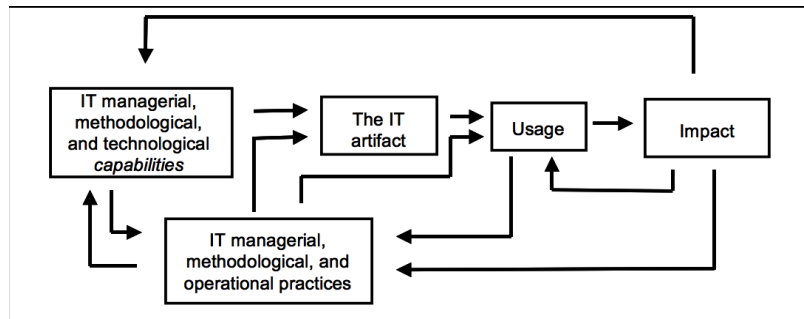


Figure 1.21 : Le réseau nomologique des artefacts technologiques (Benbasat et Zmud, 2003, p.130)

b. Une conception systémique du concept de système d'information

La conception technologique des systèmes d'information comporte certaines limites qui ont été justement soulignées par des auteurs comme Carlsson (2010) ou Alter (2003). En effet, en restreignant le concept de système d'information à ses aspects matériels et visibles (les artefacts), la conception technologique met de côté la dimension organisationnelle ou systémique inextricablement liée aux artefacts. Autrement dit, elle occulte les interdépendances entre les artefacts technologiques et les organisations qui les utilisent. D'ailleurs, si l'on se rappelle que la théorie de la structuration fait partie des jalons historiques du développement des méthodologies de *design science* et qu'elle prône la dualité de la technologie (technologie comme artefact technologique/technologie comme technologie-en-pratique), on penchera pour cette conception systémique, plus cohérente par rapport au développement historique des méthodologies de *design science*.

Cette conception systémique élargit donc le périmètre d'un système d'information pour inclure des problématiques 'non-technologiques'. Pour Agarwal et Lucas (2005), il s'agit d'adopter un point de vue macroscopique qui englobe l'ensemble des phénomènes relatifs au rôle de la technologie dans la transformation d'une organisation ou d'un réseau d'organisation. Le concept de système d'information englobe alors les phénomènes technologiques et les phénomènes organisationnels. Par conséquent, les méthodologies de *design science*, au delà de designer des artefacts, englobent dans leur démarche les dimensions organisationnelles qui interfèrent avec ces artefacts.

Partant de cette conception Carlsson (2010, p.213) avance les méthodologies de *design science* sont efficaces pour "développer ou améliorer les performances "d'initiatives en système d'information". Il les définit comme "le design et l'implémentation d'une solution dans un système socio-technique où les systèmes d'information constituent une ressource clé pour atteindre les objectifs de transformation à l'origine de l'intervention".

Conclusion

Bien que j'ai fait le choix de présenter cette controverse comme une arène opposant deux positions, je considère néanmoins que celles-ci sont compatibles et complémentaires et de fait, ma thèse emprunte des éléments aux deux conceptions.

La construction d'un artefact technologique innovant étant l'objectif central de ma thèse, j'emprunte à la conception technologique du système d'information. En parallèle, et du fait que la conception de l'artefact est directement dictée par le contexte organisationnel qui l'accueillera (les écosystèmes open data), ma thèse porte une grande attention à la dimension organisationnelle et par là, s'inspire de la conception systémique du concept de système d'information.

1.6.3 L'approche design science à l'appui de la recherche et de la pratique sur la construction des usages issus des données ouvertes.

Résumé

Après avoir présenté le cœur de l'approche *design science* (A) et ses développements récents (B), nous avons une vision précise de ce que cette approche est et de ce qu'elle devient. Désormais, il s'agit de rattacher cette approche à mon terrain et à ma problématique de recherche. Plus précisément, je démontrerai en quoi cette approche est plus pertinente que les autres étant donné la particularité de mon terrain et de ma problématique de recherche. Aussi, je prouverai que cette approche peut aboutir à la construction de connaissances nouvelles sur la transformation de l'open data en innovation de services.

1.6.3.1 La pertinence de l'approche design science pour aborder la transformation de l'open data en innovation de services

Résumé

Dans cette partie, nous mettrons en perspective la Design Science Research Methodology (DSRM) avec d'autres méthodologies orientées action et visant à résoudre des problématiques concrètes, avant d'expliquer en quoi la DSRM est la plus adaptée pour répondre à notre problématique de recherche

Je retrouve dans la littérature des rapprochements fréquents entre la DSRM et trois autres méthodologies visant à résoudre des problématiques concrètes : l'Action Design Research, la Soft Systems Methodology et la Recherche-Action (Pascal, 2012. Zuiderwijk, 2015).

1.6.3.1.1 L'Action Design Research (ADR)

L'Action Design Research (ADR) est une méthodologie qui, d'un côté, se rapproche de la DSRM en ce qu'elle vise à construire un artefact technologique d'information et de l'autre s'en distingue en ce qu'elle vient répondre à une des limites de la DSRM selon Sein et al. (2011) : l'absence de prise en compte de contextes organisationnels dans la conception et l'évaluation de l'artefact. Pour répondre à ce manque, l'ADR couple la conception d'un artefact dans un contexte donné avec de la recherche-action visant à résoudre une situation problématique dans ce même contexte.

1.6.3.1.2 La Soft Systems Methodology (SSM)

La Soft Systems Methodology (SSM) est une approche organisée permettant d'adresser des situations problématiques. Elle est orientée action. Elle offre un cadre de pensée de façon à ce que des actions pratiques puissent être envisagées pour améliorer la situation (Checkland et Poulter, 2010). Elle partage avec la DSRM son objectif qui est de résoudre une situation problématique mais par un intermédiaire différent. Quand la DSRM s'intéresse à la conception d'un artefact technologique tangible, la SSM vise à créer des modèles d'activités intangibles.

1.6.3.1.3 La Recherche-Action (RA)

Enfin, la Recherche Action (RA) implique pour "le chercheur de se préoccuper de créer des changements organisationnels et d'étudier simultanément le processus" (Pascal, 2012). La proximité entre la DSRM et la RA a amené des chercheurs comme Markus et al. (2002) ou McKay et Marshall (2005) à solliciter la première pour mener à bien la deuxième. Selon eux, la RA apporte à la DSRM une meilleure définition mise en pratique de l'évaluation de l'artefact lui-même mais aussi du protocole de recherche scientifique du fait des fondements solides et stables de la méthodologie de recherche-action. Or, comme l'ADR, la RA se déroule dans un contexte organisationnel défini alors que la DSRM s'adresse à des contextes trans-organisationnels.

Plusieurs raisons justifient mon choix de la DSRM parmi ces méthodologies. Tout d'abord, nous ne pouvons solliciter l'ADR et l'AR puisqu'elles s'appliquent dans des contextes organisationnels définis. Or, les réutilisateurs de l'open data et les écosystèmes open data qui constituent le cœur de notre recherche ne peuvent être circonscrits au sein d'une organisation. La DSRM permettant d'aborder des phénomènes trans-organisationnels, elle nous paraît être la plus adaptée. Ensuite, si la SSM s'applique bien à des contextes trans-organisationnels, elle n'aboutit pas à la création d'un artefact, qui est un de mes objectifs, d'autant plus important dans la mesure où cet objet permettra de réduire la complexité de la problématique de façon à ce que tous les profils impliqués puissent s'y rapporter et participer à la formulation des solutions.

Conclusion

Dans cette partie, nous avons montré les similitudes et différences entre la DSRM, l'AR, l'ADR et la SSM. Nous concluons que la DSRM reste la méthodologie la plus adaptée à notre problématique dans la mesure où elle s'applique à un contexte trans-organisationnel et vise à la conception d'un artefact.

1.6.3.1.4 Contextes d'application de l'approche design science

Les méthodologies design science sont pertinentes quand il s'agit de résoudre des problèmes non résolus ou mal résolus et qui sont identifiés comme tel par les praticiens du domaine étudié. En choisissant de mettre en place une démarche design science dans le cadre de cette thèse, je définirai par la suite un problème non résolu ou mal résolu avec les praticiens de l'open data (élus, chefs de projets, producteurs ou utilisateurs). L'identification du problème se fait au travers d'une "boucle de pertinence" qui correspond à des échanges itératifs entre l'environnement des praticiens et mes recherches académiques pour arriver à l'identification d'une problématique la plus précise possible (Pascal, 2012). Théoriquement parlant, l'environnement mentionné renvoie à un "espace-problème" et se compose de personnes, d'organisations et de technologies. Il comporte également des buts, activités et problèmes qui sculptent des besoins spécifiques. Ces besoins sont exprimés par des personnes au sein d'organisations, personnes dont la perception est affectée par les rôles et les responsabilités qu'ils y occupent (Hevner et al., 2004). La méthodologie de recherche en *design science* vise à répondre à ces besoins exprimés au sein de l'environnement des praticiens.

Ces besoins prennent la forme de wicked problems caractérisés par :

- Des supposés flous et des contraintes basées sur des contextes environnementaux aux contours mal définis ;
- Des interactions complexes entre les sous-composantes du problème et sa solution ;
- Une forte sensibilité au changement aussi bien concernant le processus de design que l'artefact ;
- Une forte dépendance aux capacités cognitives humaines pour produire des solutions efficaces ;
- Une dépendance aux habiletés sociales des êtres humains pour produire des solutions efficaces.

1.6.3.2 Les travaux sur la transformation de l'open data sollicitant l'approche design science

Résumé

Je présenterai dans cette partie les travaux ayant sollicité l'approche *design science* pour aborder le sujet de l'open data. Je m'attacherai particulièrement à montrer, à travers ces travaux, le potentiel de nouvelles connaissances que l'approche *design science* permet de créer sur la question de l'open data.

1.6.3.2.1 Zuiderwijk (2015) : la Design Science Research Methodology pour construire une infrastructure facilitant la coordination de l'usage de l'open data.

La thèse de Zuiderwijk (2015) constitue une des références en matière d'utilisation de la Design Science Research Methodology (DSRM) pour appuyer les praticiens de l'open data. Sa thèse trouve son origine dans l'identification d'un besoin des praticiens de l'open data, les chercheurs, et une absence de réponse concrète de la littérature à son égard.

En effet, dans les démarches d'utilisation par les chercheurs de l'open data, elle fait le constat d'une forte dispersion des outils nécessaires à une bonne utilisation de l'open data et d'une faible coordination entre les chercheurs. En parallèle, elle fait le constat que les études ayant abordé la question des facteurs influençant l'utilisation de l'open data ont adopté une posture abstraite ne permettant pas d'en déduire des recommandations pratiques que les chercheurs en prise avec l'open data pourraient utiliser dans leurs pratiques.

Pour répondre à la problématique des praticiens et combler le gap dans la littérature, elle adopte la DSRM. Elle justifie son choix en le mettant en perspective avec les autres méthodologies "orientées action", visant à résoudre des problématiques concrètes : la Recherche Action (AR), l'Action Design Research (ADS) et la Soft Systems Methodology (SSM).

La RA et l'ADS s'appliquant dans des contextes organisationnels définis, elles ne peuvent être pertinentes dans l'analyse de l'usage de l'open data. La raison en est que les utilisateurs de l'open data sont répartis dans des écosystèmes qui ne sont pas circonscrits au sein d'une seule organisation. Ce sont des écosystèmes ouverts et donc trans-organisationnels et trans-sectoriels (Crosnier, Vidal, 2017). Si la Soft System Methodology peut embrasser un périmètre d'étude trans-organisationnel, elle n'inclut pas le design d'un artefact innovant permettant d'améliorer l'utilisation de l'open data, ce qui est un des besoins clairement identifié par les chercheurs utilisant l'open data (Zuiderwijk, 2015).

Dès lors, l'approche design science apparaît comme la plus appropriée étant donné le terrain abordé (les chercheurs utilisateurs de l'open data) et les objectifs de recherche (construire un artefact permettant de faciliter l'utilisation de l'open data).

La méthodologie utilisée par Zuiderwijk (2015) comprend cinq étapes, répondant à cinq questions de recherche qui correspondent à cinq instruments de recherche. Ces étapes reprennent les étapes génériques propres à l'approche Design Science et les instruments de recherche sont définis comme "les différentes méthodes utilisées pour exécuter une stratégie de recherche particulière" (Gonzalez, 2010).

Étape de la Design Science Research Methodology	Question de recherche	Instrument de recherche
Identification du problème	Quels sont les facteurs qui influencent l'utilisation de l'open data ?	Revue de littérature
Définition des objectifs de la solution	Quelles sont les fonctionnalités qu'une infrastructure open data doit avoir pour faciliter la coordination de l'utilisation de l'open data ?	Etude de cas
Design d'un artefact	Quels sont les éléments fonctionnels pouvant traduire les fonctionnalités nécessaires citées ci-dessus ?	Analyse des besoins, design et revue de littérature
Développement d'un prototype	A quoi ressemble une infrastructure open data ayant ces éléments fonctionnels ?	Prototypage d'une infrastructure open data
Evaluation du prototype	Quels sont les effets de l'infrastructure développée sur la coordination de l'open data ?	Evaluation du prototype par des quasi-expérimentations

Figure 1.22 : Architecture de recherche des recherches de Zuiderwijk (2015,p.29)

Conclusion
<p>Par la problématique adressée et son application de la <i>Design Science Research Methodology</i>, la thèse de Zuiderwijk constitue une base sur laquelle cette thèse s'appuiera. En effet, l'artefact et les méthodes d'évaluation développées "illustrent dans quelle mesure une infrastructure open data peut-être utilisée pour faciliter et accélérer l'utilisation de l'open data et de faciliter la coordination de cette utilisation". Plus précisément, nous verrons dans la prochaine partie dans quelle mesure nous pouvons réutiliser les 22 principes de design de la coordination et les 15 principes de design d'interactions ? Et ce, étant donné les limites de cette thèse. Nous identifions 3 principales limites : cette recherche a exclu de son périmètre d'étude les interactions entre utilisateurs et fournisseurs de l'open data ; les utilisateurs étudiés sont très spécifiques puisqu'il s'agit des chercheurs en sciences humaines et sociales ; le type d'utilisation facilité par l'infrastructure créé est aussi très limité puisqu'il s'agit de l'utilisation opérationnelle de données de recherche structurées ouvertes.</p>

1.6.4 Les framework pour la construction d'usages issus des données ouvertes

1.6.4.1 Pertinence du choix d'un framework comme artefact central de la thèse

Grâce à la théorie de la structuration, nous avons pu identifier la nécessité d'avoir un cadre technologique orienté vers la construction des usages. En effet, si les travailleurs de l'open data emprunte leur comportement à une structure technologique isolée et contraignante, il est alors difficile de générer un dialogue avec un écosystème ainsi que la création de cadres cognitifs partagés chez les opérateurs de l'open data (Ruijter et al., 2018). Dès lors, je soutiens que les usages seront construits en partie grâce à un nouveau cadre (*frame*) permettant de faire advenir un travail de la donnée ouverte (*open data work*). On parle de *framework*.

Aussi, faut-il insister ici sur la spécificité de l'environnement que nous étudions (les écosystèmes open data) et qui pousse à un type d'artefact particulier. Harrison et al. (2012) notait que le produit principal d'un écosystème quel qu'il soit était le changement systémique et continu. Ils reprennent pour démontrer

l'importance de cette idée, les travaux de Nardi et O'Days (1999) ayant élaboré avec une précision particulièrement remarquable les liens et similitudes entre les écosystèmes biologiques et ce qu'ils nomment les *écologies informationnelles*. A la base de leur réflexion se trouve la définition d'un écosystème comme "un ensemble de composantes, fortement connectées par des interdépendances substantielles" (Harrison, 2012, p.908). De ce fait, Nardi et O'Days expliquent :

"Le produit de ces éléments interconnectés et en interaction est un changement systémique continu, car lorsqu'un élément est modifié, les effets peuvent être ressentis dans l'ensemble du système". Nardi et O'Days (1999, p.51)

Ils continuent en rappelant que ces changements continus et ces intégrations de nouvelles composantes sont non seulement le fruit des possibilités offertes par les nouvelles technologies de l'information mais surtout une nécessité pour la bonne santé des écosystèmes. En effet, les écosystèmes se nourrissent de la diversité des acteurs, des technologies et des outils accessibles pour s'adapter au changement permanent. Ils concluent en dressant les caractéristiques principales du comportement dynamique des écosystèmes :

"La dynamique des écosystèmes est une dynamique de flux et de mouvement - les personnes, les idées, les activités et les outils en mouvement, car chaque écosystème évolue continuellement sous la forme d'éléments qui s'ajustent et sont ajustés les uns par rapport aux autres, en essayant toujours de s'adapter parfaitement, sans jamais y parvenir. Cela fait partie de l'équilibre dynamique atteint dans les écologies saines - un équilibre trouvé dans le mouvement et non dans l'immobilité" Nardi et O'Days (1999, p. 53)

Conclusion

Avec ces caractéristiques, un artefact visant à faciliter l'innovation portée par des écosystèmes ne peut qu'encourager le flux de mouvement permanent et ces ajustements entre les différentes tâches de l'*open data work*. Aussi, cet artefact sera d'un type nouveau puisque dans un écosystème, les problèmes et solutions évoluent perpétuellement et que le temps du design ne peut en aucun cas saisir des besoins et problèmes évolutifs. Plutôt, il s'agit de créer un artefact offrant un espace permettant de générer de nouvelles combinaisons entre diverses composantes, donnant ainsi la possibilité aux acteurs clés de l'écosystème d'adapter en permanence les innovations de services ainsi que le travail de l'*open data*

1.6.4.2 Les caractéristiques d'un framework : un artefact de meta-design d'écosystèmes open data

Un artefact de meta-design peut se comprendre comme étant un artefact au service du design d'autres artefacts (Fischer et Giaccardi, 2006). Dans le cas de ma thèse, ces artefacts sont des usages issus de la transformation des données ouvertes.

Un artefact de meta-design vise à répondre à des situations, comme celles que l'on rencontre dans les écosystèmes open data, où les futurs usages et problèmes des utilisateurs ne peuvent être anticipés au moment du design (Fischer et Hermann, 2011). Des situations où on ne peut donc appliquer les méthodes de design visant à construire un système figé. La raison en est que les utilisateurs découvriront nécessairement des usages non-anticipés et inversement mettront à jour le manque de pertinence des fonctionnalités pourtant conçues pour répondre à leurs problématiques. Aussi, ces nouvelles opportunités et défauts seront-ils utilisés pour créer de nouvelles connaissances et de nouveaux besoins auxquels il faudra répondre par l'intermédiaire d'artefacts.

1.6.4.3 Le framework comme infrastructure d'écosystèmes

Le caractère infrastructurel d'un framework tient dans le fait que ce sont juste des briques qui permettent de construire un écosystème open data. Le framework est ici à percevoir comme un socle qui permet de faire fonctionner une dynamique open data à l'échelle d'une ville. On reprend ici la définition de l'infrastructure établie par Star (1999) pour laquelle une infrastructure est principalement définie par sa transparence ou son invisibilité et par son caractère relationnel. Son invisibilité tient au fait qu'à chaque nouvelle tâche effectuée sur la base de l'infrastructure, il n'est pas question de réassembler ses composantes. Son caractère relationnel tient au fait qu'une infrastructure sera considérée comme invisible par certains acteurs et comme un objet de travail pour d'autres acteurs. Nous pouvons prendre l'exemple du logiciel open source R, qui, du point de vue des utilisateurs est un outil qui permet de faire des calculs statistiques en utilisant des libraires de codes librement accessibles et pour lesquels l'infrastructure du logiciel est invisible ; et qui du point de vue des développeurs qui travaillent sur son architecture est un objet de travail quotidien.

Conclusion du chapitre 1

L'objectif de ce chapitre était de comprendre comment la littérature scientifique met en relation les quatre construits qui définissent la construction des usages des données ouvertes, à savoir les données ouvertes, l'usage, les plateformes et les écosystèmes de données ouvertes.

J'ai fait le choix d'utiliser un cadre d'analyse¹³ permettant de classer plus de deux cent articles scientifiques en fonction de leur niveau d'analyse. Il m'a permis d'établir dans les quatre premières sections une classification des articles selon sept niveaux : le niveau intra-organisationnel, organisationnel, extra-organisationnel, inter-organisationnel, territorial et sociétal, superposition des niveaux, tension paradoxale entre les niveaux.

Le choix de ce cadre d'analyse s'est avéré concluant dans la mesure où il m'a permis, dans la cinquième section de ce chapitre, d'identifier un gap dans la littérature. Si la littérature scientifique recense un grand nombre d'études qui permet de comprendre les dynamiques d'usage des données ouvertes sur une seule dimension (ou niveau d'analyse), rares sont les auteurs qui ont combiné les niveaux d'analyse et proposé une méthodologie pour attaquer l'usage depuis cet angle de vue.

Encore plus rare sont les auteurs étant partis du point de vue que les niveaux d'analyse sont en tension paradoxale quand il est question d'aborder l'usage des données ouvertes. Par exemple, peu d'auteurs ont mis en exergue le fait que l'analyse de l'usage des données ouvertes au niveau individuel devra être nécessairement complétée par une analyse au niveau organisationnel pour pouvoir en comprendre les tenants et aboutissants, et ce, bien que l'analyse au niveau organisationnel puisse invalider voire s'opposer à l'analyse au niveau individuel.

De fait, peu d'études ont proposé un appareil méthodologique qui permet d'aborder ce phénomène qui se comprend nécessairement selon différents niveaux d'analyse, chaque niveau ayant besoin de l'autre tout en s'y opposant. C'est pour répondre à ce gap que je présente une méthodologie *design science* dans la chapitre suivant.

¹³Termes originaux : *framework*

Chapitre 2

Méthodologie de recherche *design science*

Chapitre 2 : Méthodologie

Dans le chapitre précédent, j'ai identifié un gap dans la littérature scientifique sur lequel j'ai construit la question de recherche centrale de cette thèse.

Les recherches menées jusqu'à présent sur la construction des usages issues de l'ouverture des données se sont employées à décrire, comprendre et expliquer ce phénomène. Grâce à elles, j'ai une solide connaissance théorique des acteurs, processus, conditions et environnements qui participent à son émergence. A l'inverse, il existe peu de travaux qui explorent les moyens de favoriser cette innovation et qui permettraient de tester la pertinence empirique des connaissances théoriques. De fait, je propose la question de recherche suivante comme pilier central de cette thèse :

Comment résoudre les tensions paradoxales qui font obstacle à la construction de nouveaux usages à partir des données ouvertes ?

J'ai également qualifié la construction des usages issus de l'ouverture des données.

Je suis arrivé à la conclusion qu'il relevait d'une catégorie particulière de phénomène, connue sous le nom de problèmes pernicieux¹ et que sa compréhension nécessitait une méthodologie de recherche particulière : la méthodologie de recherche design science². J'ai exploré les différentes formes qu'elle pouvait prendre avant de justifier le choix d'une approche particulière.

Dans ce chapitre, je décris la méthodologie visant à répondre à notre question de recherche et qui permettra de saisir la nature particulière de l'objet qui l'intéresse de manière à construire des connaissances théoriques et pratiques. Dans un premier temps, je replacerai notre méthodologie dans le contexte plus global de l'architecture de recherche qui sous-tend l'ensemble de nos travaux. D'un côté, elle met en lumière le raisonnement qui me permet de mettre en relation notre posture scientifique et épistémologique avec notre posture méthodologique. De l'autre, elle permet de déduire de cette posture des principes opérationnels sur lesquels je me suis appuyé pour construire le design de recherche. Dans un second temps, je détaillerai chacune des cinq étapes de notre design de recherche ainsi que la nature des transitions qui permettent de les lier les unes aux autres. Plus précisément, je présenterai, pour

¹ Termes originaux : *wicked problems*

² Termes originaux : *design science research methodology*

chacune des étapes, la sous-question de recherche adressée, la méthode utilisée ainsi que les résultats espérés.

2.1 Présentation de l'architecture de recherche

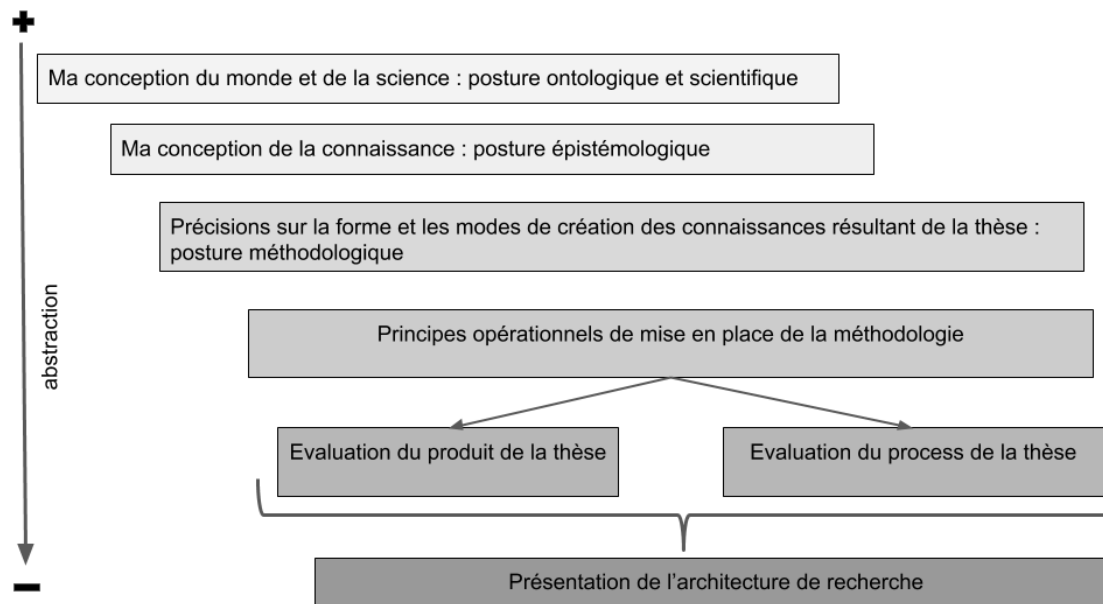


Figure 2.1 : Présentation des choix qui influencent l'architecture de recherche

2.1.1 De la posture scientifique et épistémologique à la posture méthodologique

2.1.1.1 Posture ontologique et scientifique : le choix du paradigme des sciences de l'artificiel

Je commencerai par préciser le paradigme scientifique que je solliciterai dans le cadre de cette thèse et par là, je donnerai ma vision de ce que “faire une thèse en sciences de gestion” veut dire.

Selon Avenier (2010), la plupart des recherches en sciences de gestion, bien qu'elles se revendiquent de paradigmes épistémologiques divers, ont en commun une conception de la science qui renvoie aux sciences naturelles comme la physique ou la biologie. Cette vision sous-entend que le but des sciences de

gestion est d'arriver à des connaissances objectives obtenues en découvrant les lois qui régissent des objets dotés de propriétés descriptives.

Or, Simon (1969) part du principe que tous les objets qui nous entourent ont été façonnés par l'homme, avec un but précis, c'est-à-dire avec des propriétés téléologiques plutôt que descriptives. En épousant cette vision de la science, ma thèse adopte le paradigme des sciences de l'artificiel, qui s'oppose au paradigme scientifique des sciences naturelles.

Par conséquent, la conduite d'une démarche scientifique et plus précisément la réalisation d'une thèse en science de gestion signifie pour moi être "capable de comprendre l'enchevêtrement des objets façonnés par l'homme dans l'optique de concevoir des artefacts qui pourront s'inscrire dans cet enchevêtrement et le faire évoluer" (Avenier, 2010). Cette thèse vise donc plus à répondre à la question "comment ça marche ?" qu'à la question "est-ce vrai ou faux ?"

Par ce choix, cette thèse se rapproche des postures épistémologiques fréquemment utilisées par les chercheurs en design science, selon lesquelles il existe de multiples états du monde alternatifs (alternative world states), contextuellement situés, qui sont rendus possibles par des artefacts socio-techniques (Kuelcher et Vaishnavi, 2015). Par exemple, dans cette thèse, les différentes étapes d'évolution d'un écosystème *open data* (Thomas et Autio, 2013) constituent les différents états du monde possibles. Par la construction d'artefacts, il s'agira non seulement de comprendre la construction des usages issus des données ouvertes mais aussi de créer de nombreuses connaissances sur le prochain état des écosystèmes *open data* qui est sur le point d'advenir et qui, en même temps, élargira encore les limites des capacités humaines et organisationnelles (Hevner et al., 2008).

Je précise que si Lyon et Grenoble seront les principaux écosystèmes à bénéficier des résultats, l'objectif de ce doctorat dépasse le cadre de ces trois terrains. En effet, ma thèse ne vise pas uniquement la création d'artefacts pour des écosystèmes donnés, ce qui ferait d'elle un produit de l'industrie du design, mais plutôt à combiner la production d'un artefact avec la production de nouvelles connaissances (Vaishnavi et Kuelcher, 2004). De cette manière, le paradigme que j'adopte, issu de la recherche en science du design, ne se limite pas à résoudre un problème de conception propre à quelques écosystèmes. Il contribue plutôt à une classe générale de problèmes et à une grande variété de contextes organisationnels et sociaux.

2.1.1.2 Posture épistémologique

2.1.1.2.1 La choix du paradigme épistémologique de design research établi par Romme (2003)

Dans le chapitre 1, nous avons passé en revue les différents paradigmes épistémologiques compatibles avec la posture scientifique et ontologique issus des *Sciences de l'artificiel* (Simon, 1969). J'ai choisi d'adopter un paradigme épistémologique dit design établi par Romme (2003) qui considère que l'objectif de la connaissance est de produire ou de servir la "production de systèmes qui n'existent pas encore et par là, induire des changements dans les systèmes existants pour obtenir ceux que l'on souhaite".

Ce paradigme épistémologique vient parfaitement répondre la situation problématique dans laquelle se trouve aujourd'hui les praticiens et les chercheurs qui s'intéressent à l'ouverture des données : à savoir le manque d'usage des données, et donc le manque de services ou applications innovants qui crée de la valeur à partir de l'ouverture des données.

En adoptant ce paradigme, ma thèse vise donc à produire un dispositif en Systèmes d'Information qui permettra la construction des usages issus des données ouvertes qui est l'objet de cette thèse et qui permettront, en parallèle, d'étudier ses processus.

2.1.1.2.2 Une posture épistémologique inspirée des travaux de Carlsson (2010)

Pour produire ce dispositif en Systèmes d'Information, j'emprunterai la posture épistémologique des travaux de Carlsson (2006, 2010) qui s'inspire du réalisme critique développé en premier lieu par Roy Bhaskar (1978). Elle définit trois domaines de réalités (le réel, l'actuel et l'empirique). Le réel est régi par des structures et des mécanismes ; l'actuel par des événements et des comportements ; l'empirique par des expériences. Selon cette posture, le but de la science est de comprendre les structures et mécanismes du réel et de les utiliser afin de modifier les expériences vécues dans le domaine empirique.

Ces modifications se font à travers des interventions en Systèmes d'Information (*Information system intervention*) qui sont constituées de cinq éléments (figure ci-dessous). Ces interventions sont déclenchées par une Situation Problématique (SP) située dans un Contexte (C) particulier. Les praticiens souhaitent que la SP la quitte pour atteindre une Situation Désirée (SD). Pour cela, le chercheur en système d'information met en place un dispositif en Systèmes d'Information (IS Initiative) qui vient utiliser les Mécanismes (M) pour assurer le passage de SP à SD.

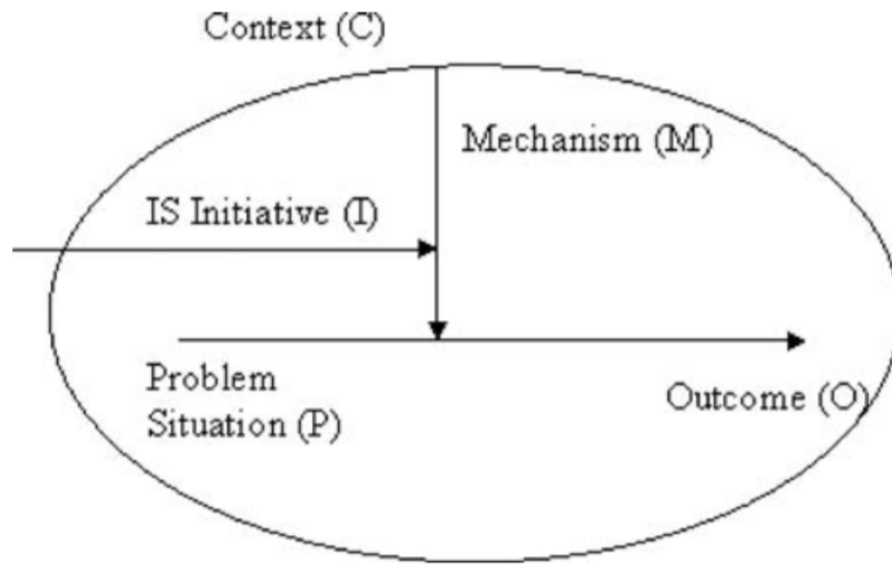


Figure 2.2 . Intervention en Systèmes d'Information (Carlsson, 2010, p.220)

2.1.1.3 Posture méthodologique

2.1.1.3.1 Une méthodologie pour contribuer aux efforts de construction d'une théorie sur l'usage et la transformation des données ouvertes

Le premier point que je souhaite éclaircir sur ma posture méthodologique concerne la nature des résultats scientifiques espérés par ma thèse

L'objectif final des méthodologies de design science fait l'objet de nombreux débats, dans lequel ma thèse se doit de se positionner. Pour résumer brièvement les opinions qui s'opposent, il existe:

- d'un côté les auteurs pour qui l'objectif de la recherche est la formulation de nouvelles théories.
- d'un autre côté des auteurs comme Hevner et al. (2004) qui élargissent les perspectives et considère "qu'est contribution tout ce qui accroît la base de connaissance du domaine : approfondissement théorique, méthodologique ou nouvelle expérience acquise par la recherche" (Pascal et al., 2012).

Ma thèse s'inscrit dans la deuxième perspective du fait que le champ de recherche qui s'intéresse à l'ouverture des données est particulièrement jeune et immature. Aujourd'hui, selon une des dernières revues de littérature sur le sujet publiée en 2018 (Hassan et Twinomurinzi, 2018), il n'existe aucun cadre théorique établi dans la littérature comme base sur laquelle il est possible d'établir une autre théorie,

propre au domaine. Il serait trop ambitieux de prétendre à la construction d'une théorie nouvelle sachant que les théories sous-jacentes qui pourrait la faire émerger n'ont pas encore été identifiées.

2.1.1.3.2 L'utilisation d'une méthodologie design science sans emprunt à la recherche-action

Le deuxième point que je souhaite éclaircir sur ma posture méthodologique concerne le mode d'interaction avec le terrain mis en place.

Comme la revue de littérature scientifique sur les méthodologies design science l'a démontré (voir Chapitre 1), on trouve des auteurs qui préconisent de s'appuyer sur les canons méthodologiques de la recherche-action pour combler le manque de références méthodologiques sur les recherches en *design science*. La recherche-action est ici considérée comme méthode comprenant deux cycles interdépendants, le cycle pour la résolution de problèmes et le cycle pour la recherche scientifique (McKay et Marshall, 2005). Elle peut servir de base pour déployer une méthodologie de recherche en *design science*.

Cette thèse s'oppose à cette vision pour deux raisons principales : d'un côté la recherche-action se préoccupe de créer des changements organisationnels. Or, à l'instar de Carlsson (2010), je considère que la recherche en design science a pour objectif de développer des dispositifs matériels et des connaissances pratiques pour les mettre à disposition des praticiens et dont la capacité à provoquer le changement a été prouvé scientifiquement. Mais il n'est pas question de s'impliquer dans le changement. D'un autre côté, la littérature scientifique dispose aujourd'hui de méthodologies de design science suffisamment matures pour être utilisées seules, à l'images de celles proposées par Pascal (2012) et Carlsson (2006, 2007)

2.1.1.3.3 Une méthodologie de thèse au croisement des deux écoles de recherche en design science

Enfin, et en guise de synthèse sur ma posture méthodologique, je souhaite établir mon positionnement par rapport aux deux écoles de méthodologies design science traditionnellement reconnues.

Deux écoles de recherches s'opposent traditionnellement à propos de la définition et des composants d'une méthodologie de design science (Carlson, 2012, p.229) :

- l'école anglo-saxonne qui place au centre l'ingénierie et la construction d'artefacts technologiques innovants.

- l'école européenne qui met l'accent sur le développement de connaissances théoriques design et l'intégration des artefacts dans des systèmes socio-techniques.

Cette thèse se positionne au croisement des deux écoles dans le mesure où :

- D'un côté, la construction d'un artefact technologique innovant constitue un objectif central de la thèse.
- De l'autre, la dimension socio-technique aura une place centrale puisqu'il sera question de l'usage réel de l'artefact technologique.

2.1.2 Principes opérationnels

2.1.2.1 Validité externe : les principes relatifs à l'évaluation du produit de la thèse

A l'instar des recommandations de Carlsson (2007) et Pascal (2012), l'évaluation du produit de la thèse suivra un double mouvement :

- d'un côté, c'est le dispositif ou artefact, c'est-à-dire la technologie informatique en elle-même qui sera évaluée. Le chapitre 6 qui traitera du prototypage du dispositif abordera ce point plus en détail.
- de l'autre, c'est l'usage de cette technologie par des membres d'écosystèmes *open data* qui sera évaluée. Le chapitre 7 relatif à l'évaluation du dispositif abordera ce point plus en détail.

2.1.2.2 Validité interne : les principes relatifs à l'évaluation du processus de la thèse

Le terme processus revêt ici un double sens, et, de fait, son évaluation suivra nécessairement un double mouvement :

- Le processus renvoie d'un côté à la méthodologie de développement du produit. Celle-ci sera évaluée par l'intermédiaire de présentations faites aux organisations qui développent de dispositifs pour le compte des opérateurs des données ouvertes. Je solliciterai en particulier l'entreprise Open Data Soft et la coopérative Dataactivist dont c'est le rôle sur le territoire français. Le chapitre 7 relatif à l'évaluation du dispositif abordera ce point plus en détail.

- D'un autre côté, le processus renvoie à la démarche de construction de connaissance mise en œuvre dans cette thèse. Cette démarche sera évaluée en participant à des consortium doctoraux comme celui de l'Association Information et Management ou des conférences scientifiques comme la conférence de Recherche en *Design Science* dans le domaine des systèmes d'informations³ dédiée au *design science* dans le domaine des systèmes d'information.

2.2 Présentation des cinq étapes du design de recherche

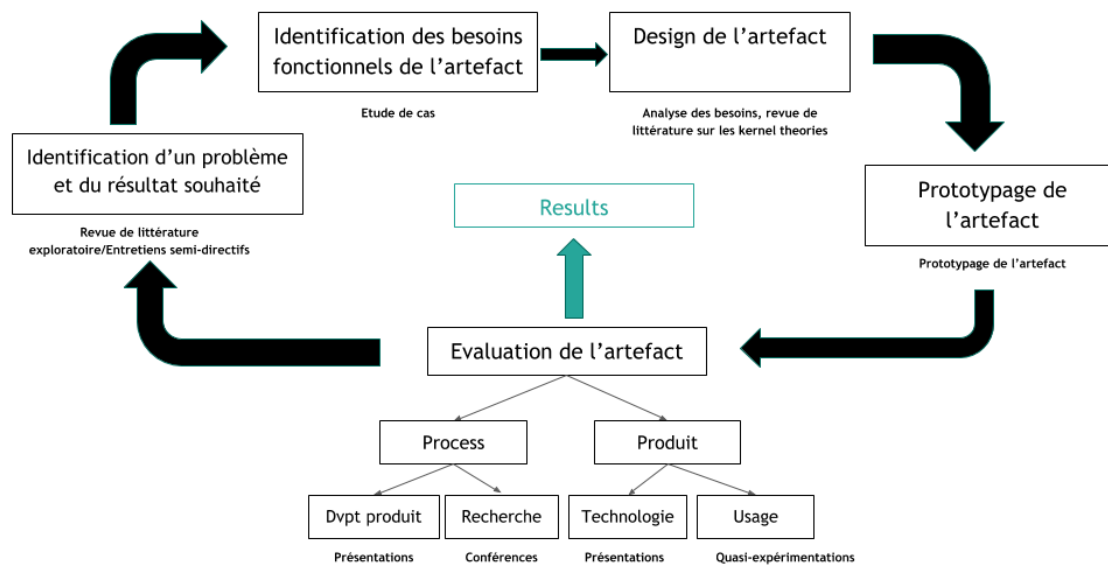


Figure 2.3. L'architecture de la recherche

Le design de la recherche est subdivisé en cinq parties itératives qui répondent aux trois sous-questions de recherche.

³ Termes originaux : *Design Science Research in Information Systems and Technology Recherche (DESRIST)*

Première partie : identification du problème et du résultat souhaité

Cette première partie répond à la première sous-question de recherche : quel est l'état des lieux concernant la construction des usages ? Pour la mener à bien, il s'agira, dans un écosystème donné, de comprendre les éléments contextuels qui freinent la construction des usages issus des données ouvertes. Pour cela nous utiliserons une méthode de revue de littérature ainsi qu'une méthode d'exploration hybride basée sur la méthode des cas.

Deuxième partie : Identification des besoins fonctionnels de l'artefact

Cette deuxième partie vise à répondre à la deuxième sous-question de recherche : quelles sont les situations problématiques qui freinent la construction des usages issus des données ouvertes ? Pour la mener à bien, il s'agit ici par l'intermédiaire d'études de cas d'établir la liste des fonctionnalités nécessaires d'un dispositif dont le but est d'accélérer la construction des usages issus des données ouvertes. Pour cela, nous utiliserons une méthode d'étude de cas inspirée des travaux sur la théorie ancrée.

Troisième partie : design de l'artefact

Cette troisième partie vise à répondre à la troisième sous-question de recherche : quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ? Pour la mener à bien, il s'agit d'identifier dans la littérature les éléments, puis les théories sur lesquelles la conception d'un dispositif sera basée (dites kernel theories) et enfin de construire les fonctions que ce dispositif devra incorporer. Pour cela, nous utiliserons une méthode de design particulière.

Quatrième partie : prototypage de l'artefact

Cette partie vise également à répondre à la troisième sous-question de recherche. Il s'agit de prototyper le dispositif selon les kernel theories identifiées mais aussi à décrire les composants du dispositif de façon à ce qu'il puisse être évalué. Pour cela, nous utiliserons une méthode de prototype particulière.

Cinquième partie : évaluation de l'artefact

Cette partie vise aussi à répondre à la troisième sous-question de recherche. Il s'agit, par divers outils, d'évaluer si le dispositif construit permet d'atteindre le résultat souhaité. Sinon, une nouvelle boucle de

conception se met en œuvre. Si oui, alors les boucles s'arrêtent et les résultats sont publiés par l'intermédiaire du document de la thèse ou de publications.

2.2.1 Les méthodes utilisées

Résumé

Je présenterai ci-après toutes les méthodes utilisées pour concevoir le dispositif, méthode qui découle de la posture épistémologique et méthodologique choisie à l'issue de la revue de littérature sur l'approche *design science*.

2.2.1.1 Méthode de revue de littérature

Dans le premier chapitre, je réaliserai une revue de littérature visant à fournir une image critique des différents courants qui irriguent la recherche sur l'innovation de services issue des données ouvertes. Mon raisonnement réutilisera le *pattern* de *framework development* développé par Vaishnavi et Kuelcher (2015, p. 115) et qui suggère au chercheur d'effectuer une analyse morphologique (Zwicky, 1967) des informations dans un domaine. Plus concrètement, il s'agit d'utiliser un framework visuel permettant de classer les contenus scientifiques sur lesquels je me suis appuyé. À la lecture de ce framework, il est possible d'identifier morphologiquement un gap de recherche dans les zones du domaine étudié.

J'utiliserai ici le framework multi-niveau développé par Chesbrough and Bogers (2014) et affiné par Bogers et al. (2016) pour classer la littérature sur le concept d'innovation ouverte⁴ (Chesbrough, 2003) selon le niveau d'analyse. Il distingue cinq unités d'analyse permettant de classer les déterminants, processus et résultats d'un phénomène :

- le niveau intra-organisationnel
- le niveau organisationnel
- le niveau extra-organisationnel
- le niveau inter-organisationnel

⁴ Termes originaux : open innovation

- le niveau industriel, territorial et sociétal

À ces cinq niveaux, j'en ajouterai un sixième, qui permet de prendre en compte les travaux qui superposent plusieurs niveaux d'analyse, et un septième pour lister les travaux reconnaissant une tension paradoxale (K. Smith et al., 2017) entre les différents niveaux d'analyse. Par tension paradoxale, j'entends "les différents niveaux d'analyse se confrontent les uns aux autres de manière contradictoire tout en étant nécessairement interdépendants" (K. Smith et al., 2017, p.305).

2.2.1.2 Méthode d'exploration hybride

Dans le deuxième chapitre, je présente une démarche d'exploration hybride (Thiétart et al., 2014, p. 93) en utilisant la méthode des cas (Hlady-Rispal, 2002). Cette méthode allie, de manière itérative, des observations participantes et des entretiens. La construction de cette démarche d'exploration met en perspective trois phases ou séquences (Giordano et Jolibert, 2012, p.39) :

2.2.1.2.1 Phase d'invention et de conceptualisation du phénomène

2.2.1.2.2 Choix des observations à effectuer

2.2.1.2.3 Confrontation des observations avec les praticiens et formulation d'un nouveau cadre conceptuel

2.2.1.2.1 Phase d'invention et de conceptualisation du phénomène

Cette première phase a été réalisée dans le cadre de la conclusion de la revue de littérature. Elle a pour objet de mobiliser et d'intégrer les concepts étudiés de manière à créer un modèle théorique ou canevas de recherche (Hlady-Rispal, 2002) permettant d'analyser le phénomène au cœur de la recherche. A savoir, les difficultés rencontrées dans l'innovation de services issue des données ouvertes. J'ai abouti à l'utilisation d'un modèle heuristique des écosystèmes *open data* (Dawes et al., 2016) qui me permet d'analyser le processus d'innovation de services issu des données ouvertes et ses difficultés.

Pour inventer un tel modèle préliminaire de l'usage des données ouvertes je m'inspire ici du pattern d'analyse de système complexe (Vaishnavi et Kuelcher, 2015, p. 107) qui permet de décomposer le fonctionnement d'un système complexe en quatre couches :

1. La structure statique du système complexe avec ses différents sous-systèmes interdépendants. Il s'agit ici de comprendre quels sont les éléments constitutifs d'un système complexe et leurs liens.

Pour illustrer ce que représente la structure statique d'un système complexe, prenons l'exemple de l'article de Vaishnavi et al.(1997) sur le fonctionnement des objets intelligents⁵. Il commence par montrer comment s'organise les objets intelligents entre eux (figure ci-dessous)

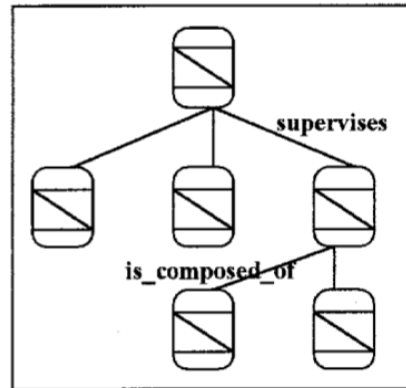


Figure 2.4 : exemple de structure statique d'un système d'objets intelligents (Vaishnavi et al., 1997, p.277)

2. La structure statique de chaque sous-système avec ses subdivisions. Il s'agit ici de comprendre les éléments constitutifs communs à chaque sous-système et leurs liens. Pour reprendre le même exemple, Vaishnavi et al. (1997) présente la structure statique de chaque objet intelligent.

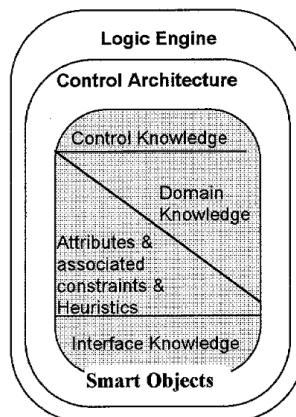


Figure 2.5 : Exemple de structure statique d'un objet intelligent (Vaishnavi et al., 1997, p.281)

⁵ Termes originaux : *Smart objects*

Dans le troisième chapitre, je décrirai plus précisément les éléments constitutifs des systèmes de production, d'intermédiation et de transformation de données. Je m'appuierai sur une convention de notation établie par Datta (1998) pour représenter les différents éléments qui forment les activités de production, d'intermédiation et de transformation de données. Selon cette convention, chaque processus est constitué d'une séquence d'états d'un objet (S1,S2,S3 dans la figure ci-dessous) sur lesquels les acteurs agissent avec des artefacts qui sont d'un certain type et qui peuvent atteindre certains objectifs. Par convention, on note cette conjonction de facteurs avec deux séquences de lettres ("R-SR" dans la figure ci-dessous)

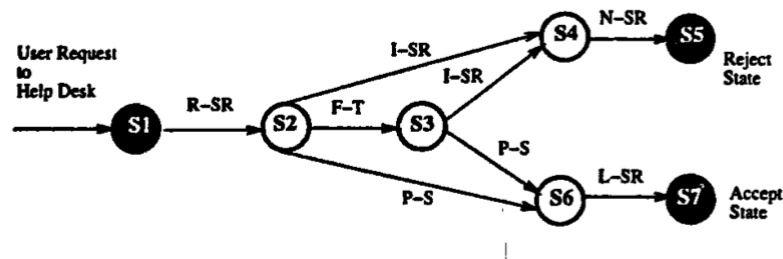


Figure 2.6 : Représentation d'un processus selon Datta (1998, p.279)

- Le comportement dynamique de chaque sous-système et la manière dont le comportement du système est produit. Dans notre exemple, Vaishnavi et al. (1997) présente comment les objets intelligents interagissent entre eux par l'intermédiaire d'un moteur d'inférence avec 3 étapes : lecture, inférence, écriture.

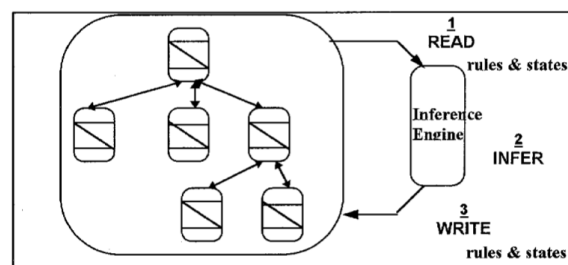


Figure 2.7 : Exemple de comportement dynamique d'un système d'objets intelligents (Vaishnavi et al., 1997, p.277)

Dans le troisième chapitre, je décrirai plus précisément comment les artefacts technologiques, sociaux et leurs capacités d'action s'imbriquent dans des séquences d'articulations pour faire passer les données d'un état à un autre.

4. L'évolution historique des systèmes et de leurs interactions. Traditionnellement, cette évolution historique permet d'analyser le cycle de vie d'un système complexe avec ses différentes phases.

2.2.1.2.2 Choix des observations participantes à effectuer et réalisation

Dans la première étape, il s'agit de définir les éléments et concepts constitutifs d'un modèle heuristique d'usage de la donnée. J'obtiens neuf concepts. Dans la deuxième étape, il est question de construire une grille d'observation à partir de ces neuf concepts du modèle. Cette grille constituera le principal outil de collecte de données. Je présente ci-dessous les concepts, les interrogations suscitées et la grille d'observation qui les intègrent toutes.

2.2.1.2.2.1 Une plateforme open data

Il résulte des développements précédents qu'une plateforme *open data* était constituée de trois éléments : une architecture, des composants stables et des mécanismes de gouvernance. Ces éléments permettent de comprendre le niveau de participation et de collaboration voulu par les concepteurs de la plateforme. En ce sens, les observer revient pour le chercheur à affiner sa compréhension du rôle des plateformes *open data* dans l'usage des données, tel qu'imaginé par les concepteurs des plateformes. Dès lors, il est primordial de les identifier et de les observer dans les investigations empiriques.

Question(s) empirique(s) : Quels sont les éléments tangibles et intangibles qui définissent les plateformes open data ? Quel est le rôle des plateformes open data dans l'usage des données ?

2.2.1.2.2.2 Les parties prenantes des plateformes open data

Les parties prenantes des plateformes jouent un ou plusieurs rôles dans l'offre, la demande ou la maintenance des données ouvertes. De fait, ils exercent une influence certaine sur la manière d'utiliser les données aujourd'hui et sur les évolutions de ces usages demain. Par conséquent, il sera important d'établir qui participe à l'évolution de la plateforme et comment.

Question(s) empirique(s) Qui sont les parties prenantes des plateformes ? Quel(s) rôle(s) jouent ces parties prenantes ?

2.2.1.2.2.3 Les données ouvertes en tant qu'espèce ingénieur

En tant qu'espèce ingénieur, les données ouvertes structurent tout un ensemble d'outils informatiques. Établir et comprendre quels sont ces outils permettra au chercheur de délimiter, voire de cartographier "l'espace d'influence" des données (nommé écosystème *open data*) et ainsi d'évaluer la magnitude du phénomène d'usage des données ouvertes. Par conséquent, il sera important d'établir et de comprendre les outils qui constituent l'écosystème *open data*.

Question(s) empirique(s) Quels sont les outils informatiques structurés par les données ouvertes ?

2.2.1.2.2.4 L'écosystème *open data*

L'écosystème *open data* est constitué de trois systèmes interdépendants qui assurent la circulation des données ouvertes et le fonctionnement d'une chaîne qui produit, ouvre et transforme les données ouvertes en services. Sans aucune information sur les données en entrée et en sortie de chacun des systèmes, et sur l'identité des entités qui les forment, il est difficile de détailler les étapes de l'usage des données ouvertes. Par conséquent, il sera nécessaire dans ces investigations empiriques d'étudier l'influence des acteurs de chaque système sur l'état des données.

Question(s) empirique(s) Qui sont les acteurs qui constituent les systèmes de production, d'intermédiation et de transformation des données ouvertes ? Quel est l'état initial des données en entrée du système et l'état final des données en sortie du système ?

L'écosystème de chaque cas pourra être représenté de la façon suivante :

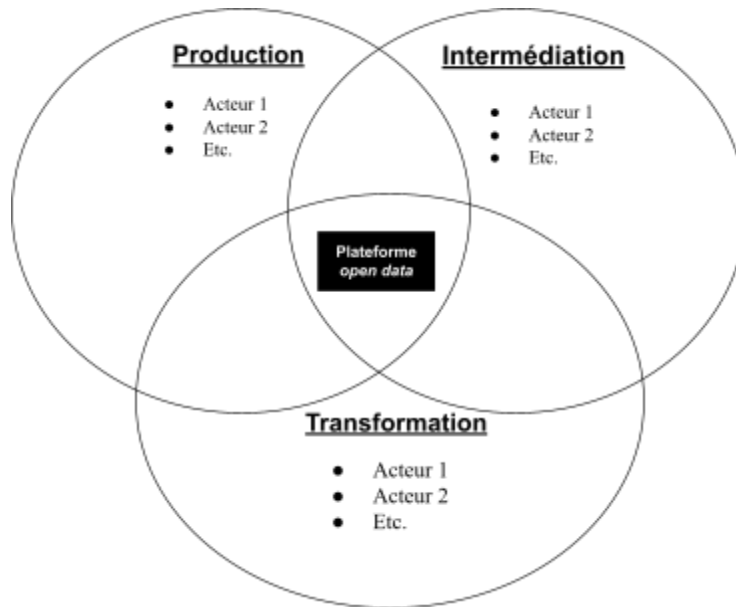


Figure 2.8 : La représentation d'un écosystème open data (d'après Harrison et al., 2012, p.908)

2.2.1.2.2.5 Les artefacts technologiques et sociaux

Les artefacts sont d'une certaine manière les maillons les plus importants de la chaîne formée par chacun des systèmes. Le choix de concevoir et d'utiliser tels ou tels artefacts permet de comprendre les différentes trajectoires d'usage réelles et possibles des données ouvertes. De ce fait, il sera crucial de comprendre la nature et les objectifs des artefacts à l'œuvre dans l'usage des données ouvertes.

Question(s) empirique(s) Quels sont les types et les objectifs des artefacts à l'œuvre dans l'usage de l'open data ?

2.2.1.2.2.6 Leurs capacités d'action (material et human agency)

La capacité d'action constitue le mécanisme qui explique directement les différents états des données ouvertes. Par conséquent, il est évident que les investigations empiriques viseront à les identifier et à étudier leur influence sur l'état des données ouvertes.

Question(s) empirique(s) Quelles sont les capacités d'action des artefacts ? Quel est leur influence sur l'état des données ouvertes ?

2.2.1.2.2.7 Les états des jeux de données ouvertes

Les états des jeux de données ouverts sont témoins de l'usage qui en sont fait. Dans ce cadre, les investigations empiriques viseront à comprendre les différentes configurations prises par les données ouvertes au cours de leur usage.

Question(s) empirique(s) Quelles sont les différentes configurations dans lesquelles se trouvent les jeux de données ouvertes ?

2.2.1.2.2.8 Les séquences d'articulation

Avec des données sur les différentes séquences qui constituent l'usage des données ouvertes, le chercheur s'ouvre la possibilité d'une compréhension nouvelle du comportement dynamique non seulement de chaque système mais aussi de l'écosystème global. Dans ce cadre, les investigations empiriques s'attacheront à comprendre les différentes séquences qui définissent l'usage des données ouvertes.

Question(s) empirique(s) Quelles sont les différentes séquences d'articulation identifiables dans l'usage des données ouvertes ?

2.2.1.2.2.9 Les complémentarités formées

Enfin, les complémentarités sont le résultat de chacune des séquences d'articulations. Elles incarnent la valeur ajoutée à chaque séquence d'articulation et, finalement, permettent de comprendre la valeur de service innovant issu de l'usage des données ouvertes. De fait, une grande attention sera portée sur le résultat des séquences d'articulation et la valeur ajoutée générée par chacune d'entre elles.

Question(s) empirique(s) Quel est le résultat de chacune des séquences d'articulation ? Quelle est la valeur ajoutée émergente de chacun de ces résultats ?

Toutes ces questions forment une grille d'observation des cas exploratoires présentée dans la figure ci-dessous.

Strate	Element	Question empirique	Influence sur l'observation
Statique/globale	Plateforme open data	<ul style="list-style-type: none"> • Quels sont les éléments tangibles et intangibles qui définissent les plateformes open data ? • Quel est le rôle des plateformes open data dans la construction de nouveaux usages à partir des données ouvertes ? 	Etablir et observer la composition des plateformes open data et leur rôle dans la construction de nouveaux usages à partir des données ouvertes
Statique/globale	Parties prenantes	<ul style="list-style-type: none"> • Qui sont les parties prenantes des plateformes ? • Quel(s) rôle(s) jouent ces parties prenantes ? 	Etablir qui participe à l'évolution des plateformes et comment
Statique/globale	Données ouvertes	<ul style="list-style-type: none"> • Quels sont les outils informatiques structurés par les données ouvertes ? 	Etablir et comprendre les outils qui constituent l'écosystème open data
Statique/globale	Ecosystème open data	<ul style="list-style-type: none"> • Qui sont les acteurs qui constituent les systèmes de production, d'intermédiation et de transformation des données ouvertes ? • Quel est l'état initial des données en entrée du système et l'état final des données en sortie du système ? 	Etudier l'influence des acteurs d'un système sur l'état des données
Statique/systémique	Artefacts	<ul style="list-style-type: none"> • Quels sont les types et les objectifs des artefacts technologiques et sociaux à l'oeuvre dans l'usage de l'open data ? 	Comprendre la nature et les objectifs des artefacts à l'oeuvre dans l'usage des données ouvertes
Statique/systémique	Capacité d'action	<ul style="list-style-type: none"> • Quelles sont les capacités d'action des artefacts ? • Quel est leur influence sur l'état des données ouvertes ? 	Etudier l'influence des capacités d'action sur l'état des données ouvertes
Statique/systémique	Etats	<ul style="list-style-type: none"> • Quelles sont les différentes configurations dans lesquelles se trouvent les jeux de données ouvertes ? 	Comprendre les différentes configurations prises par les données ouvertes au cours de leur usage
Dynamique/systémique	Séquence d'articulation	<ul style="list-style-type: none"> • Quelles sont les différentes séquences d'articulation identifiables dans l'usage des données ouvertes ? 	Comprendre les différentes séquences qui définissent l'usage des données ouvertes
Dynamique/systémique	Complémentarités	<ul style="list-style-type: none"> • Quel est le résultat de chacune des séquences d'articulation ? • Quelle est la valeur ajoutée émergente de chacun de ces résultats ? 	Etudier le résultat des séquences d'articulation et la valeur ajoutée générée par chacune d'entre elles

Table 2.1 : La grille d'observation des cas exploratoires

Ensuite, grâce à cette grille d'observation, j'ai effectué des observations dans quatre cas de projets d'usage des données : "Quartier", "Transport", "Logement" et "Collaboration". Pour des raisons de confidentialité, j'ai anonymisé le nom des projets, des acteurs et des technologies utilisées.

Ces observatoires ont été réalisées en adoptant la posture du balancier, suggérée par Hlady-Rispal (2002, p.109) qui consiste à osciller entre l'impératif de participation et l'impératif de décentration. D'un côté, pour pouvoir être adopté par le milieu observé et récupérer des informations riches et pertinentes, j'ai dû adopter les codes de communication et comportements qui lui sont propres (impératif de participation). Cette Immersion a cependant dû être contrebalancé par un mouvement de prise de recul et ce afin d'éviter l'absorption, synonyme d'arrêt du projet de recherche. Pour cela, j'ai mis en place des dispositifs de collecte et élaboré des protocoles visant à assurer la collecte des données (impératif de décentralisation).

Aussi, pour constituer notre échantillon d'observation ai-je utilisé le principe d'opportunisme méthodologique. Ceci m'a amené à explorer l'ensemble des opportunités se présentant à moi et de fait de multiplier les observations.

2.2.1.2.3 Confrontation des observations avec les praticiens et formulation d'un nouveau cadre conceptuel

Cette étape vise à confronter les praticiens⁶ avec les conjectures tirées de nos observations. Pour constituer l'échantillon de praticiens à interroger, j'ai utilisé la technique de la boule de neige. Cette procédure est conseillée quand les populations à interroger sont difficiles à identifier (Thiétart et al., p.234) ou réticentes à l'exercice des entretiens. C'est le cas des parties prenantes des plateformes *open data*. J'ai donc trouvé un premier interlocuteur avant de lui demander d'en désigner d'autres qui présentent des caractéristiques similaires, et ainsi de suite.

Le protocole des entretiens semi-directifs avec ces praticiens consistait en premier lieu à évoquer nos conjectures dans le but de savoir si elles correspondaient à leur expérience. Cette confrontation me permettait alors d'affiner ou de réfuter notre canevas de recherche et d'identifier d'autres opportunités d'observation (étape 2).

In fine, c'est le canevas de recherche affiné qui servira de modèle de compréhension initial du phénomène et qui clôturera les premières activités du cycle de pertinence. Il sera exporté vers d'autres contextes

⁶ (i.e les parties prenantes des plateformes)

“parents” (Thiéart et al., 2014, p. 93 ; Passeron, 2006) et permettra alors d’initier les cycles de design et de rigueur (Hevner et Chatterjee, 2010, p.16) qui, avec le cycle de pertinence, forment l’ensemble du processus de recherche de la méthodologie de recherche design science.⁷

2.2.1.3 Méthode d’étude de cas

Dans le troisième chapitre, je présente une étude de cas comparative des écosystèmes *open data* de la métropole Grenoble-Alpes et du Grand Lyon. Comme l’indique l’architecture générale de cette thèse, elle a pour but premier, dans chaque itération, de construire les besoins fonctionnels auxquels un dispositif en système d’information doit répondre pour passer de la situation problématique à la situation désirée. Ces besoins fonctionnels comportent trois éléments : une description du contexte (C) social et technique, des différents éléments de situation problématique (ESP) ainsi que des solutions (S) entrevues par les opérateurs des données ouvertes pour atteindre la situation désirée (SD). *In fine*, cette étude de cas permettra d’établir une liste de besoins fonctionnels qui seront exprimés de la façon suivante :

Pour atteindre la SD dans C, il s’agit d’adresser ESP en mettant en place S - **Exemple de besoin fonctionnel, inspirée de la logique CIMO créée par Denyer et al. (2008) pour exprimer des propositions de design**

Pour aboutir à ces besoins fonctionnels, il s’agira, pour chaque itération, de mettre en place une méthodologie d’étude de cas qui s’inspire principalement des travaux de Eisenhardt (1989). La particularité de ses travaux réside dans la combinaison des de Miles et Huberman sur l’analyse des données qualitatives (1984), de ceux de Yin (1981, 1984) sur le design d’études de cas et les recherches de Glaser et Strauss (1967) sur le développement de théories ancrées⁸. Elle mobilise tour à tour ces travaux pour proposer une feuille de route qui précise les étapes à suivre pour construire des connaissances théoriques à partir d’une étude de cas. Dans notre étude de cas, les connaissances théoriques prennent la forme des besoins fonctionnels mentionnés ci-dessus.

La théorie ancrée est une méthode systématique qui permet de construire des connaissances théoriques via la collecte et l’analyse méthodique de données et d’informations (Glaser et Strauss, 1967). Je reprendrai les neuf étapes proposées par Eisenhardt (1989) qui permettent de développer des connaissances théoriques ancrées, qui se trouvent être dans mon cas des besoins fonctionnels et ce, à partir d’études de cas :

1. Définition de la sous-question de recherche adressée
2. Sélection des cas

⁷ Termes originaux : *Design Science Research Methodology*

⁸ Termes originaux : *grounded theories*

3. Construction du protocole et des instruments de l'étude de cas
4. Immersion dans le terrain
5. Analyses intra-cas
6. Recherche de pattern inter-cas
7. Ébauche d'hypothèses
8. Confrontation avec la littérature
9. Atteinte du point de clôture

2.2.1.3.1 Définition de la sous-question de recherche adressée

La première étape consiste à rappeler la sous-question de recherche adressée par cette étude de cas. Elle se définit eu égard à notre compréhension de la situation problématique, de son contexte, et évolue avec elle. En effet, pour rappel, cette méthodologie de recherche en *design science* est intrinsèquement itérative. À la fin de chaque itération (chapitre 7), notre compréhension de la situation problématique se déplace et se précise. Elle évolue pour renforcer graduellement sa **pertinence** vis à vis de la réalité des praticiens auxquels cette thèse s'adresse. De ce fait, la sous-question de recherche sera redéfinie en fonction des évolutions de notre compréhension de la situation problématique.

Grâce aux études de cas exploratoires, j'ai posé comme première hypothèse (à ne pas prendre dans un sens positiviste) que le faible usage des données ouvertes était dû à la présence de barrières socio-techniques dans les écosystèmes *open data*. Par conséquent, la première itération de l'étude de cas vise à répondre à la sous-question de recherche suivante :

Quels sont les besoins fonctionnels auxquels un dispositif en Systèmes d'Information doit répondre pour lever les barrières socio-techniques qui se dressent face à la construction de nouveaux usages issus des données ouvertes ?

À nouveau, il est important de rappeler que cette première question sera précisée ou complètement redéfinie dans les prochaines itérations en fonction de la reconceptualisation de la situation problématique, elle-même issue des phases d'évaluation (chapitre 7) du dispositif que je designerai (chapitre 5) et prototyperai (chapitre 6).

2.2.1.3.2 Sélection des cas

Contrairement à la première étape dont le contenu évoluera au gré des itérations, cette deuxième étape est statique. En effet, elle sera consacrée à l'étude des cas des écosystèmes *open data* des métropoles Grenoble-Alpes et du Grand Lyon. Cette sélection de cas se justifie, au-delà de leur facilité d'accès pour l'équipe de recherche, par le fait que ces deux écosystèmes possèdent un double avantage. D'un côté, ils sont suffisamment mûrs pour que je puisse y observer des phénomènes d'usage des données. De l'autre, ils sont suffisamment hétérogènes en termes de géométrie, de mode de gestion et d'orientation, pour que les résultats communs aux deux cas puissent prétendre à une certaine validité externe.

2.2.1.3.3 Construction du protocole et des instruments de l'étude de cas

Je mobiliserai plusieurs modes de collecte de données hétérogènes de façon à bénéficier d'une triangulation des données. Aussi cette pluralité d'instruments de renforcer la validité des construits et d'établir les différentes dimensions de l'usage de la donnée étudiées ici.

À nouveau, cette partie verra son contenu varier en fonction des évolutions de la compréhension de la situation problématique et donc de la sous-question de recherche adressée par l'étude de cas. Quoi qu'il en soit, pour lancer la première itération, j'ai conduit deux activités :

- la réalisation d'entretiens semi-directifs en suivant un protocole de collecte et d'analyse de contenu thématique (Miles et Huberman, 2004) cadré par l'heuristique d'écosystème *open data* et centré autour des deux unités d'analyses choisies dans le chapitre précédent.
- l'analyse de documents internes produits directement par des opérateurs des données ouvertes à Lyon et à Grenoble. J'ai également suivi un protocole d'analyse de contenu thématique (Miles et Huberman, 2004)

2.2.1.3.4 Immersion dans le terrain

À l'image de cette thèse, la phase de collecte de données et d'immersion dans le terrain a évolué entre chaque itération. Les recherches ont été orientées en fonction des opportunités de collecte suggérées par l'itération précédente. À ce propos, Eisenhardt rappelait que pour les recherches dans lesquelles il est question de construire un cadre théorique, cette flexibilité dans la collecte des données était parfaitement légitime :

Pour la recherche théorique, (la modification des méthodes de collecte de données) est légitime dans la mesure où les enquêteurs essaient de comprendre chaque cas individuel, aussi profondément que possible. Ainsi, si une nouvelle opportunité de collecte de données se présente et si une nouvelle ligne de pensée émerge pendant la recherche, il est logique d'en profiter en modifiant la collecte de données, si une telle modification est susceptible de mieux ancrer la théorie ou de fournir de nouvelles perspectives théoriques » - Eisenhardt (1984, p. 539)

2.2.1.3.5 Analyse intra-cas

Dans cette cinquième phase, si les objectifs de l'analyse évolueront au gré des itérations, le protocole sera le même. J'effectuerai une longue description des cas qui aidera à traiter les larges volumes de données récoltées. Ensuite, je reprendrai la méthode utilisée par Leonard-Barton (1988) qui utilise des tableaux et des graphiques d'information pour représenter les données.

2.2.1.3.6 Recherche de pattern inter-cas

À l'image de la cinquième étape, si les objectifs de l'analyse évolueront au gré des itérations, le protocole sera le même. Il s'agira de mettre en œuvre le procédé qui consiste à sélectionner des catégories ou des dimensions particulières puis de confronter les similitudes intra-cas avec les différences entre les cas (Eisenhardt, 1984, p.540). Cette méthode permettra d'identifier aussi bien des traits spécifiques que des traits communs aux deux cas.

2.2.1.3.7 Ébauche d'hypothèses

À nouveau, si le contenu des hypothèses évoluera au gré des itérations, le protocole d'ébauche sera le même. Cette étape hautement itérative consiste à comparer systématiquement les résultats émergents, avec les preuves issues de chaque cas, de façon à évaluer dans quelle mesure les résultats épousent les données récoltées. Il s'agit ici de comparer continuellement la littérature sur les besoins fonctionnels avec les données de façon à renforcer la validité empirique des premiers. Pour cela, je raisonnerai en deux étapes. La première étape, elle-même subdivisée en deux sous-parties, consiste à affiner les besoins fonctionnels. La première sous-partie consiste à affiner la définition des besoins fonctionnels avec la littérature, la deuxième sous-partie à apporter la preuve que les données évoquent des problèmes qui

peuvent être levées par ces besoins fonctionnels. La deuxième étape consiste pour sa part à s'assurer que les relations entre les besoins fonctionnels épousent bien les relations entre les problèmes évoqués.

2.2.1.3.8 Confrontation avec la littérature

Dans la continuité, si le contenu de la littérature à laquelle les hypothèses seront confrontées évoluera au gré des itérations, le protocole de confrontation sera le même. Cette étape consiste à confronter les besoins fonctionnels émergents avec la littérature existante sur les artefacts facilitant l'usage de la donnée. Il s'agit plus particulièrement d'extraire d'articles de recherche les besoins fonctionnels auxquels les artefacts se sont adaptés.

Cette phase de confrontation permettra d'identifier les similitudes ou au contraire les contradictions avec cette littérature. En juxtaposant des résultats contredisant la littérature ou inversement, le chercheur sera amené à penser différemment, de manière plus créative et ce, de façon à préciser l'expression des besoins fonctionnels identifiés.

2.2.1.3.9 Atteinte du point de clôture

Enfin, cette dernière étape consiste à décider quand il s'agit d'arrêter les itérations entre la littérature et les données. Idéalement, cette étape s'arrêtera quand la saturation théorique sera atteinte, c'est-à-dire quand l'apprentissage marginal de chaque nouvelle donnée sera minime voire nulle.

2.2.1.4 Méthode de design

Dans le cinquième chapitre, j'utiliserai une méthode qui permet de passer des besoins fonctionnels au design d'un dispositif qui permet de répondre à la situation problématique. Cette méthode se subdivise en trois étapes :

1. L'identification d'éléments fonctionnels identifiés à partir des besoins fonctionnels et qui assurent la transformation de la situation problématique en situation désirée. Chacun de ces éléments est décrit sous le forme d'une proposition de design
2. La construction de mécanismes issus de *kernel theories* qui expliquent en quoi les éléments fonctionnels permettant le passage de la situation problématique à la situation désirée. Chaque mécanisme est décrit sous la forme d'un principe de design

3. La construction d'un modèle conceptuel qui dessine l'architecture de l'artefact, intègre les éléments fonctionnels et explique les fonctions qu'il assurera dans un contexte réel. Pour parler du modèle, on parle du design du système (que constitue l'artefact).

Pour appliquer cette méthode, j'emprunterai les prescriptions du pattern de design élégant (Vaishnavi et Kuelcher, p.132-133). Ce pattern établit qu'il faut simplifier au maximum la description de l'artefact conçu. Pour cela, il s'agit d'utiliser un vocabulaire fonctionnel qui décrit les propriétés de l'artefact en termes de ce qu'il fait plutôt que les détails de sa construction et de son organisation.

2.2.1.4.1 Identification d'éléments fonctionnels et construction des propositions de design

Pour pouvoir construire les propositions de design, j'ai réalisé une nouvelle revue de la littérature visant à identifier des éléments fonctionnels qui puissent répondre aux 12 besoins fonctionnels identifiés dans le chapitre précédent. Cette revue de littérature vient compléter celle présentée au chapitre 2 en ce qu'elle vise à identifier des articles et atteindre des résultats complémentaires.

Il ne sera pas question de revenir sur l'approche complète de la revue de littérature détaillée dans le chapitre 3 mais de préciser les aspects méthodologiques qui diffèrent entre la première revue de littérature du chapitre 2.

Ces différences ont pour origine des inspirations différentes. Alors que la méthodologie de la première s'inspirait des travaux de Bogers et al.(2016), fixant des règles pour mener une littérature classée selon des niveaux d'analyse, celle-ci se base sur les travaux de Tranfield, Denyer et Smart (2003) et de Denyer, Tranfield et Van Aken (2008). Ces auteurs proposent une méthodologie de revue de littérature visant à contrecarrer le manque de pertinence empirique des revues de littératures dites traditionnelles. L'objectif de cette méthodologie est de permettre au chercheur d'extraire de la littérature des éléments qui puissent informer et guider les praticiens dans la réponse à apporter à leurs besoins opérationnels.

En premier lieu, et contrairement à la première revue de littérature, celle-ci n'a pas été structurée sur la base d'une liste de concepts visant à décrire un phénomène. Cette revue de littérature a été réalisée en suivant les cinq étapes d'usage des données ouvertes construites au chapitre 4. Ainsi, j'ai sélectionné tous les articles évoquant au moins une des cinq étapes de construction de nouveaux usages issus des données ouvertes

J'obtiens ainsi un corpus d'articles correspondant à l'ensemble cinq étapes d'usage. Ce corpus a ensuite été classé. Ce classement s'est fait en attribuant une étiquette à chaque article correspondant à l'élément fonctionnel qu'il décrit. Les articles ont ensuite été classés par groupe d'éléments fonctionnels similaires. Les groupes d'éléments fonctionnels ont été ensuite regroupés dans des catégories plus grandes correspondant à des éléments fonctionnels intégrant plusieurs groupes et ainsi de suite jusqu'à obtenir un minimum d'éléments fonctionnels couvrant un maximum d'étapes d'usage (figure ci-dessous). Par cette démarche de classification des articles, il est question d'arriver à un nombre d'éléments fonctionnel qui soit en cohérence avec mes capacités de conception. J'ai estimé que je pouvais concevoir au maximum trois éléments fonctionnels.

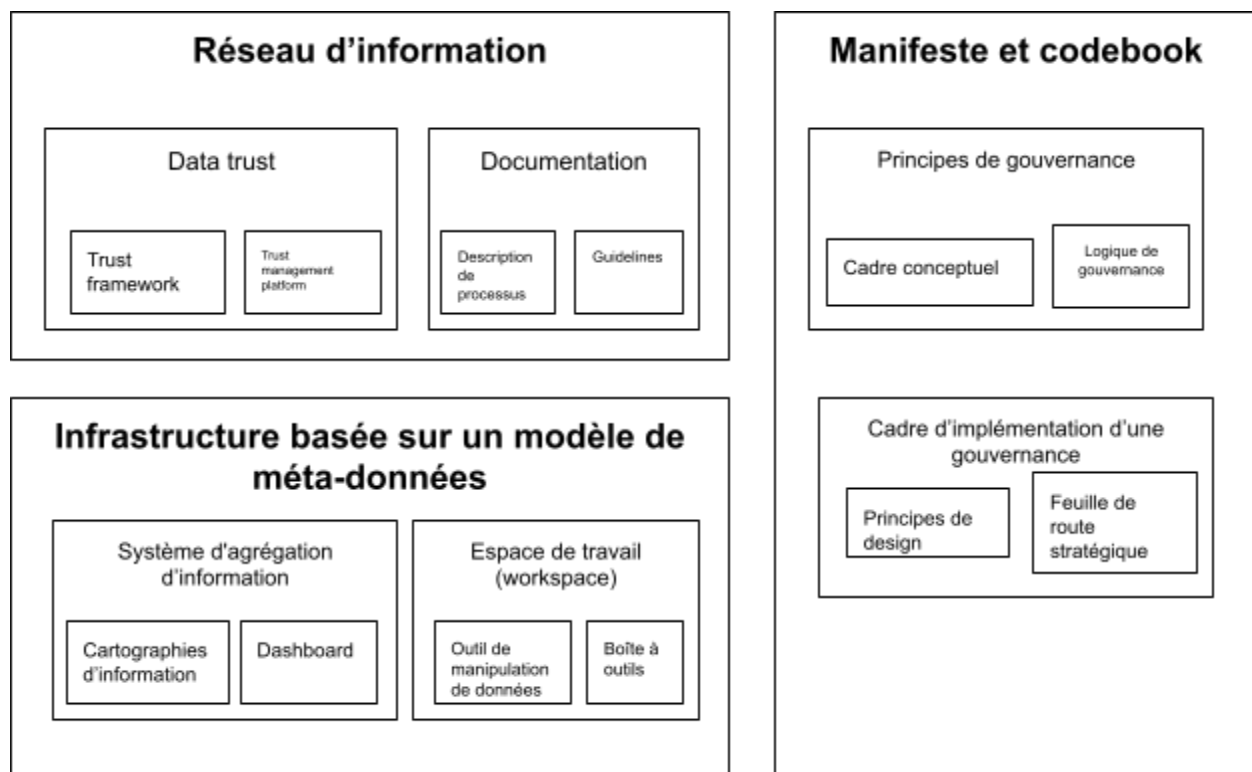


Figure 2.9 : catégorisation récursive des articles par groupes d'éléments fonctionnels

En second lieu, et contrairement à la première revue de littérature, celle-ci n'adopte pas le niveau d'analyse comme critère de sélection des articles. Plutôt, cette revue de littérature sélectionne les articles sur la base des trois critères suivants :

- cohérence entre la problématique de recherche adressée par l'article d'une part et un ou plusieurs des douze besoins fonctionnels d'autre part.

- cohérence et similitude entre le contexte empirique auquel répond l'article d'une part et le contexte propre à l'usage des données ouvertes tel que décrit au chapitre 4 d'autre part.
- l'article contient une description suffisante de l'élément fonctionnel qu'il propose pour répondre à sa problématique et fait référence à des travaux de recherche qui se sont attachés à le décrire et comprendre ses mécanismes sous-jacents.

Enfin, cette revue de littérature a été menée sur la base d'une liste de mots-clés (et leurs combinaisons) venant compléter celle utilisée pour réaliser la première revue de littérature. Cette liste, présentée dans ce chapitre, a permis d'orienter clairement la recherche d'articles de manière à identifier des articles en lien avec les douze besoins fonctionnels. Ces mots-clés ont permis d'effectuer des recherches sur plusieurs bases de données (Crossref, Google Scholar, Microsoft Academic et Scopus) et m'ont permis d'obtenir de riches corpus d'articles.

	Mots-clés
Revue de littérature - Etat des lieux (chapitre 1)	<i>open data, platform, ecosystem, use</i>
Revue de littérature - Propositions de design (chapitre 5)	<i>open data, platform, ecosystem, use</i> <i>trust, user interface, usability</i> <i>data visualisation, evaluation, domain knowledge</i> <i>data literacy, information overload, data expertise, coordination</i> <i>shared cognitive framework, Open Data Work</i> <i>data institutions, data use, formal organization, institutional distance, cultural distance</i>
Revue de littérature - Principes de design (chapitre 5)	<i>meta-design artefact, paradox theory, paradoxical tension</i> <i>information network, social construction of technology</i>

	<i>infrastructure, meta-data model, distributed cognition theory manifesto, codebook, middleground theory</i>
--	---

Figure 2.10: Mots clés choisis dans les différentes revues de littérature

2.2.1.4.2 Construction des principes de design

Définition Kernel theories : Théorie choisie parmi les quatre premiers type de théorie de la classification de Jones (les théories pour analyser, les théories pour expliquer, les théories pour prédire, et les théories pour expliquer et prédire) et qui informe le cinquième type, à savoir les théories pour agir et designer.

Ces kernel theories jouent le rôle de “noyau”⁹ théorique autour duquel on peut construire le design de l’artefact, résultat final de cette thèse. Hanseth et Lyytinen (2004, p.208) les définissent comme : “les théories issues des sciences naturelles et sociales qui gouvernent les exigences de conception ou le processus qui permet de les construire”

Les principes de design précisent les propositions mentionnées ci-dessus.

Pour construire les principes de design, j’ai orienté la revue de littérature dans une nouvelle direction : la recherche de *kernel theories*. Ces kernel theories jouent le rôle de “noyau” théorique autour duquel je construirai le design de l’artefact, résultat final de cette thèse. Hanseth et Lyytinen (2004) les définissent comme : “les théories issues des sciences naturelles et sociales qui gouvernent les exigences de conception ou le processus qui permet de les construire”.

Bharosa (2011) précise que ces théories acquièrent un caractère opérationnel par leur traduction en “lignes directrices et normatives, formulées de façon à permettre aux architectes de système d’information d’agir” (Bharosa, 2011, p.118). À ce propos, il convient de préciser dès à présent que cette thèse part , avec Gregor (2006), du postulat que des théories peuvent parfaitement acquérir un caractère normatif et directif. Dans son article majeur sur la théorie du design, Gregor (2006) établit qu’il existe cinq types de théories interdépendantes : les théories pour analyser (1), les théories pour expliquer (2), les théories pour prédire (3), les théories pour expliquer et prédire (4), et les théorie pour agir et designer dites (5). Ces théories ne s’opposent pas. Elles se complètent. En l’espèce, les *kernel theories*, choisies parmi les quatre

⁹ Traduction : kernel (en anglais)

premiers types de théories, “informer” le cinquième type : les théories pour agir et designer (Carlsson, 2006, p.222). Pour faciliter sa compréhension de cette relation entre théories, le lecteur pourra considérer que les *kernel theories* sont un noyau théorique duquel une lecture “orientée pratique” permet d’extraire un design et/ou une théorie du design.

Si les *kernel theories* sont issues d’une lecture “orientée pratique” des quatre autres types de théories, il s’agit pour les identifier d’approfondir la revue de littérature mentionnée dans la section précédente. Pour cela, il a fallu reprendre les mêmes articles, en cherchant les références théoriques qui pouvaient m’orienter vers des *kernel theories*. Une fois ces *kernel theories* identifiées, j’ai ajouté leur nom à la liste des mots-clés. utilisés pour effectuer nos recherches dans les bases de données (Crossref, Google Scholar, Microsoft Academic et Scopus). De ces recherches, j’ai extrait des articles délimitant le champ des courants de littérature associées aux *kernel theories*.

Définition : Intervention : Design et implémentation d’un dispositif dans un système socio-technique où les systèmes d’information sont cruciaux pour atteindre les objectifs visés par l’artefact (Carlsson, 2010).

Finalement, j’ai analysé ces articles en cherchant à extraire les configurations PIMCO¹⁰ (*Problem, IS Intervention, Mechanisms, Context, Outcome*) décrites par Carlsson (2006). Il s’agit d’extraire la situation problématique adressée par la théorie (*Problem*) ; l’intervention en Systèmes d’Information évoquée dans la théorie pour le résoudre (*IS Initiative*) ; les mécanismes sociaux, naturels ou techniques qui justifient le changement opéré par l’intervention et son dispositif (*Mechanisms*) ; le contexte de mise en oeuvre de cette intervention (*Context*) ainsi les résultats escomptés (*Outcome*).

2.2.1.4.3 Construction du design de l’artefact

Sur la base des douze principes de design, j’ai élaboré un dispositif en système d’information visant à faciliter l’usage des données ouvertes. Tous les principes ont été incorporés dans le design.

Pour rappel, au regard du modèle théorique développé au chapitre 3, le dispositif en système d’information que j’ai décidé de designer est un environnement de conception des usages des données ouvertes dit artefact de meta-design (Fischer et Giaccardi, 2006).

Définition : Artefact de meta-design : Infrastructure technique et sociale dans laquelle de nouvelles formes de design collaboratifs peuvent prendre forme. Il dépasse la notion traditionnelle d’artefact en ce qu’il intègre

¹⁰ (*Problem, IS Initiative, Mechanisms, Context, Outcome*)

un processus co-adaptatif entre l'artefact et l'utilisateur, élevant ainsi l'utilisateur au rang de co-développeur ou de co-designer de l'artefact (Fischer et Giaccardi, 2006, p.1).

Un artefact de meta-design se définit comme “une infrastructure technique et sociale dans laquelle de nouvelles formes de design collaboratifs peuvent prendre forme. Il dépasse la notion traditionnelle d'artefact en ce qu'il intègre un processus co-adaptatif entre l'artefact et l'utilisateur, élevant ainsi l'utilisateur au rang de co-développeur ou de co-designer de l'artefact” (Fischer et Giaccardi, 2006, p.1).

On distingue quatre caractéristiques chez les artefacts de meta-design (Giaccardi, 2003, p.156) : son habitabilité, traduite par un espace électronique dans lequel l'utilisateur peut vivre et avec lequel il peut interagir ; un paramètre relationnel, traduit par le fait que la forme de l'espace électronique peut être négocié entre individus, en relation avec la technologie. ; sa co-évolutivité mise à disposition par des interactions réciproques et récursives entre l'artefact et l'utilisateur, laissant ainsi transparaître le caractère “ouvert” de l'artefact ; et enfin sa convivialité mise en oeuvre par la possibilité offerte par l'artefact de faciliter sa gouvernance par une communauté.

Dans cette partie, il s'agit de construire le design du système que constitue l'artefact de meta-design. Le design d'un système décrit la structure et le comportement de ce système (Offermann et al., 2010). Il convient de décrire la structure et le comportement de l'artefact de meta-design dans son ensemble avant de se tourner vers les trois éléments fonctionnels : le réseau d'information, l'infrastructure basée sur un modèle de métadonnées et la manifeste associée à son codebook (voir figure ci-dessous).

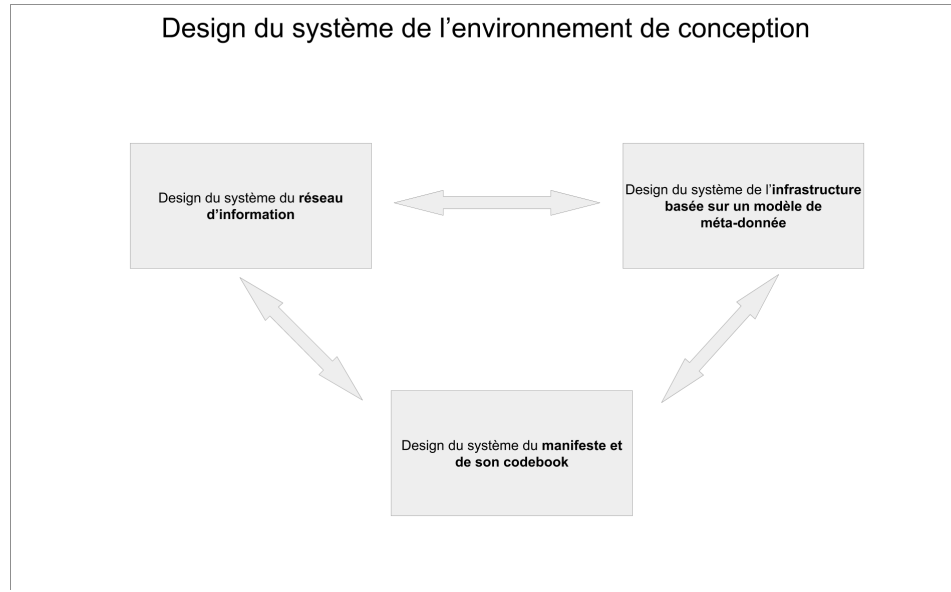


Figure 2.11 : Design du système de l'environnement de conception des usages et des trois éléments fonctionnels

Ensuite, il s'agira de définir les pattern de coordination. Ces pattern définissent les parties réutilisables d'un design, leurs apports et comment ils peuvent être appliqués dans un contexte réel, ainsi que les relations qu'entretiennent les différentes parties du design entre elles (Offermann et al., 2010). La figure ci-dessous (Figure 2.12) représente ces pattern de coordination sous forme de flèches colorées. J'ai fait le choix de présenter ces pattern de coordination sous la forme de scénario d'usage, c'est à dire sous une forme narrative "qui décrit la situation dans laquelle l'usage de l'outil en cours de conception, va se construire et prendre sens"(Pascal et Rouby, 2006). Pour construire ce scénario, je mobiliserai le modèle de scénario d'usage. proposé par Pascal et Rouby (2006).

Cette approche possède une double utilité pour le chercheur-concepteur que je suis. La première utilité tient au fait qu'elle offre une aide substantielle pour m'aider à concevoir un environnement de conception des usages des données qui soit compatible avec l'environnement et le quotidien de ses futurs utilisateurs, que l'on désignera comme "système fonctionnel d'accueil" (Pascal et Rouby, 2006, p.3). En d'autres termes, cette approche prête au scénario d'usage la qualité très utile de renforcer la coordination de mon artefact avec le contexte réel de ses futurs utilisateurs. Pour cela, le scénario d'usage viendra peupler les items suivants tirés des données récoltées pour l'étude de cas présentée au chapitre 4 :

- Caractéristiques de l'utilisateur de l'artefact : identification, statut, rôle, caractéristiques personnelles susceptibles d'impacter ses attentes concernant l'artefact (formation, expérience ...).
- Tâches réalisées à l'aide de l'artefact et type de déroulement (séquentiel, itératif, en parallèle) : voir section précédente
- Les partenaires qui prendront part à la réalisation des tâches
- Les ressources qui seront mobilisées pour réaliser les tâches
- Les informations et flux d'informations nécessaires pour réaliser les tâches
- Les outputs des tâches réalisées

La deuxième utilité tient au fait qu'elle propose une aide pour m'aider à intégrer l'artefact dans l'environnement structurel de référence des usages (ensemble de règles et ressources). En d'autres termes, l'approche de cognitivo-structurationniste prête au scénario d'usage la qualité de renforcer la coordination de mon artefact avec le contexte structurel, (évidemment plus large que le contexte quotidien des futurs utilisateurs) des grandes logiques d'action sollicitées par les acteurs et organisations impliquées dans le développement des usages des données ouvertes. Pour cela, le scénario d'usage viendra peupler les items suivants tirés des données récoltées pour l'étude de cas présentée au chapitre 4 :

- Les schèmes interprétatifs utilisés pour guider les actes des individus (modes de raisonnement, cadres de référence ou représentation collective). Ces schèmes sont compris ici comme les enjeux, contours et contenus du processus entourant les tâches réalisées par les acteurs
- les normes qui justifient et légitiment les actes, comprises comme moyen d'assignation de rôles et de devoirs aux utilisateurs de l'artefact et à leurs partenaires
- les facilités, comprises comme les ressources d'allocation et d'autorité que les auteurs mobilisent dans leur action, et qui vont subséquemment permettre de solliciter des ressources et des informations spécifiques.

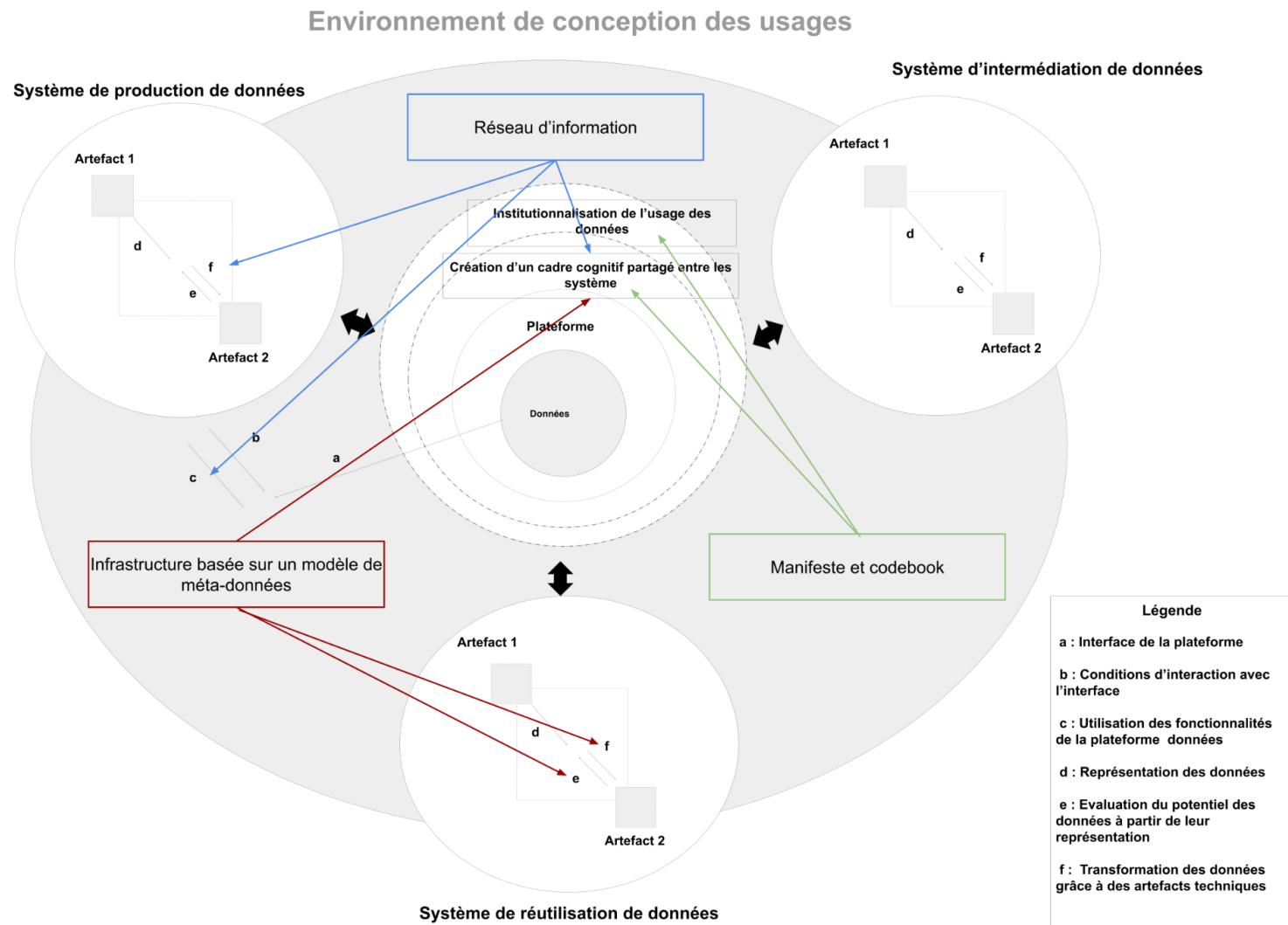


Figure 2.12 : Pattern de coordination de l'environnement de conception des usages.

Enfin, je m’efforcerai de décrire les fonctions du design, c’est-à-dire “les choses que l’objet désigné doit faire pour être une réussite” (Dym et Little, 2004, p.79). Pour cela je m’appuie sur les principes de design construits dans la partie précédente, qui m’a permis d’établir une liste de fonctions, avant de faire des liens avec le design du système et les pattern de coordination. *In fine*, ces fonctions peuvent être perçues comme une traduction pratique et concrète de ces deux éléments.

2.2.1.5 Méthode de prototypage

L’objet du sixième chapitre est d’assurer le passage de la dimension conceptuelle à la dimension matérielle. Autrement dit, il s’agit de mettre en oeuvre un processus d’instanciation qui permet :

“d’implémenter des fonctionnalités dans un objet à partir d’une structure prédéfinie et pré-optimisée”
(Cofer et Harding, 2006, p.120).

Pour chacune des itérations de cette thèse, il sera pertinent de suivre une méthode d’instanciation basée sur une actualisation des travaux de Ince et Hekmatpour (1987). Ces derniers ont défini un processus en trois étapes :

1. Définir les objectifs du prototype relativement au projet de recherche
2. Sélectionner les fonctions du prototype à mettre en oeuvre
3. Construire le prototype

Or, ce processus a été développé pour le prototypage de logiciels informatiques, des artefacts produits, tandis que cette thèse vise à prototyper un dispositif qui combine des produits et des processus. Pour cette raison, j’ai ajouté une quatrième étape au processus qui permet de satisfaire cette particularité des dispositifs en système d’information. De plus, les seuls moyens de mise en place du processus mentionnés par Ince et Hekmatpour (1997) sont des langages et des méthodes formelles propres aux sciences informatiques et à la programmation. Or, les compétences et ressources à disposition de ce travail de thèse ne me permettent pas d’utiliser de telles méthodes. Par conséquent, j’ai dû contourner ce manque de compétences et de ressources. En lieu et place des langages de programmation, j’ai mobilisé des outils logiciels (Ince et Hekmatpour, 1987, p.11). Ces outils sont des programmes qui ne nécessitent pas d’avoir des compétences de programmation mais à partir desquels il est possible de construire un logiciel. En lieu et place des méthodes formelles faisant appel aux mathématiques, j’ai utilisé des *pattern* de prototypage développés par Vaishnavi et Kuelcher (2015, p.119-156).

Finalement, pour chaque itération, il s'agira de suivre un processus en quatre étapes et d'utiliser comme moyen de mise en place des étapes des *pattern* de prototypage et des outils logiciels :

1. Définir les objectifs du prototype relativement au projet de recherche
2. Sélectionner les fonctions du prototype à mettre en oeuvre
3. Construire le prototype
4. Établir les mécanismes permettant de venir implémenter le prototype dans un contexte d'usage réel

2.2.1.5.1 Définir les objectifs du prototype vis à vis de la recherche et le type de prototypage

Dans la première étape, il s'agit de déterminer ce que le chercheur cherche à accomplir avec le prototype vis-à-vis du projet de recherche. Trois alternatives se dessinent (Carey et Mason, 1983). Un prototype peut avoir pour objectif de préciser la compréhension des besoins fonctionnels des futurs utilisateurs du prototype (*improved functional requirements*). Dans ce cas, le prototype devient un intermédiaire, un moyen d'apprentissage, une loupe d'observation au service de la compréhension des besoins des utilisateurs. Aussi, un prototype peut avoir pour objectif d'améliorer les interactions entre ce dernier et les utilisateurs (*improved interaction requirements*). Il est fréquent qu'un prototype ne soit pas utilisé du fait du design de l'interface qui décourage les utilisateurs. Dans ce cas, on intègre de nouvelles interfaces de façon à améliorer les interactions entre l'utilisateur et l'artefact. Enfin, un chercheur peut mettre en place un prototype pour s'assurer de sa capacité à s'adapter aux évolutions des besoins des utilisateurs (*easier evolution of requirements*). Il est pertinent de concevoir un système indéfini, qui peut prendre plusieurs formes en fonction des contextes d'usage et de leur évolution. Les systèmes d'aide à la décision sont de bons exemples de systèmes indéfinis qui peuvent être façonnés au gré des situations rencontrées par les utilisateurs.

2.2.1.5.2 Le choix d'un type de prototypage (throw-it-away, incremental, evolutionary)

Le choix de ces objectifs va avoir une incidence sur le type de prototypage mis en place. De ce fait, dans la première étape, il faudra également définir, pour chaque itération, le type de prototypage qui sera mis en place.

Il est pertinent de distinguer le prototypage *throw-it-away*, le prototypage incrémental et le prototypage évolutif. Le prototypage *throw-it-away* est mis en place pendant la période de compréhension des besoins

fonctionnels des utilisateurs. Il permet d'obtenir un objet très rapidement que les utilisateurs peuvent évaluer. Dans la plupart des cas, l'objet est rapidement écarté (*thrown away*) car de nouveaux besoins fonctionnels ont été identifiés (ce qui donne son nom à ce type de prototypage). Comme indiqué dans le tableau 6.1, ce type de prototypage est adapté quand les objectifs sont de mieux comprendre les besoins des utilisateurs et d'améliorer les interactions avec l'artefact. Le prototypage incrémental est quant à lui utilisé quand les concepteurs ont décidé d'un modèle d'artefact et figé le liste des besoins fonctionnels des utilisateurs. Il s'agit alors de subdiviser le modèle et de construire chaque partie progressivement. Ce type de prototypage permet de se rendre compte rapidement des erreurs et des problèmes d'implémentation qui peuvent apparaître pour les résoudre avant de créer l'artefact final. Comme indiqué dans le tableau 6.1, ce type de prototypage est uniquement utilisé quand le but est d'améliorer les interactions entre les utilisateurs et l'artefact. En effet, en mettant en place successivement chaque partie, il est possible de régler au cas par cas les frictions qui peuvent naître entre l'artefact et les utilisateurs. Enfin, le prototypage évolutif est le type le plus flexible. Il consiste à développer un système graduellement qui puisse s'adapter à des changements substantiels et souvent inévitables : modification des besoins fonctionnels, changement du cadre juridique, des organisations ou encore précision des *kernel theories* utilisées pour guider le design. Comme indiqué dans le tableau 6.1, ce type de prototypage est adapté quelque soit les objectifs.

Figure 2.13. Tableau de correspondance entre les objectifs et les types de prototypage.

	Prototypage throw-it-away	Prototypage incrémental	Prototypage évolutif
Improved functional requirements	X	X	
Improved interaction requirements	X	X	X
Easier evolution of requirements			X

2.2.1.5.3 Sélectionner les fonctions du prototype (classées par élément fonctionnel)

La deuxième étape consiste à sélectionner, par élément fonctionnel, les **fonctions** qui seront prototypées. Pour effectuer ce choix, je mobiliserai les critères du pattern *Easy solution first*¹¹ (Vaishnavi et Kuelcher, 2015, p.131). Ce pattern vise à toujours privilégier les fonctions faciles à mettre en œuvre en période de prototypage. Ainsi, je sélectionnerai les fonctions pour lesquelles il existe une solution de prototypage facile à mettre en œuvre et à évaluer. Autrement dit, je privilégierai les fonctions qui peuvent être facilement mises en place par des outils logiciels connus et ne nécessitant pas de connaissances techniques. Cette façon de sélectionner les fonctions a notamment été choisie par Berners-Lee and Cailliau (1990) quand ils ont présenté pour la première fois le concept de *World Wide Web*¹². Internet était dépeint comme une combinaison de deux éléments fonctionnels (un navigateur et un serveur) avec des fonctions simples : le navigateur a pour fonction de lire des liens hypertextes qui correspondent à des noeuds d'un réseau et de garder un historique des liens lus ; le serveur, quant à lui, stocke tous les réseaux de noeuds et négocie avec le navigateur le format sous lequel les liens peuvent être lus.

2.2.1.5.4 Construction du prototype

La troisième étape consiste à construire concrètement le prototype, c'est-à-dire à mettre en place les fonctions sélectionnées selon le type de prototypage choisi. Comme mentionné précédemment, je mobiliserai des outils logiciels pour prototyper chaque fonction. Plus précisément, j'utiliserai des outils configurables à partir d'un tableur excel ou Google sheet du fait que ces outils sont connus et utilisés de tout le monde. Cette diffusion massive facilitera la prise en main des prototypes par l'ensemble des concepteurs impliqués dans cette thèse mais aussi la future implémentation dans des situations réelles. Il sera en effet d'autant plus facile de mettre en place ces artefacts si les utilisateurs peuvent être rapidement en capacité de configurer ses fonctions au gré de leurs envies.

2.2.1.5.4.1 Architecture (modèle simple) - rappel phase design

Dans cette construction, je reprendrai le modèle de design créé et les éléments fonctionnels identifiés dans le chapitre 5 pour dessiner à partir de là une architecture générale du prototype. Je diviserai ensuite cette

¹¹ Citation originale : Easy solution first

¹² Citation originale : *World Wide Web*

architecture en différentes aires fonctionnelles dans lesquelles je viendrai placer les fonctions sélectionnées. Alors, je suivrai le pattern des *blocs de construction*¹³ (Vaishnavi et Kuelcher, 2015, p.138). Ils préconisent de décomposer chaque fonction en fonctions plus petites jusqu'à pouvoir faire correspondre à chaque fonction un outil logiciel dédié. Ensuite, il s'agit de construire la fonction puis, de les assembler récursivement jusqu'à que toutes les fonctions soient assurées et que les aires fonctionnelles soient opérationnelles.

2.2.1.5.4.1 Opérations/Prototypage des interfaces

Une fois le prototype construit, il s'agit de rédiger des scénarios d'utilisabilité, indépendant du contexte d'usage, qui permettent de dessiner le fonctionnement opérationnel du prototype et montre sa facilité d'utilisation (Pascal et Roubey, 2006). Il permet de simuler la façon dont les diverses fonctions opèrent en interactions avec l'utilisateur. Il est possible de trouver dans ces scénarios des prototypes d'interfaces utilisateurs qui permettent de se rendre compte de l'aspect visuel des fonctions et du point de vue de l'utilisateur lorsqu'il les utilise.

2.2.1.6 Méthode d'évaluation

J'ai mené des évaluations dans le but d'estimer successivement la capacité des différentes versions d'Innodata (i.e les artefacts) à conduire l'innovation de services issue des données ouvertes dans les villes intelligentes vers l'idéal qu'elles se sont fixées. En effet, le choix de la méthodologie de recherche *design science* suppose d'adopter une démarche intrinsèquement itérative dans laquelle une première réponse artefactuelle est apportée avant d'être évaluée. Si la réponse est suffisamment pertinente empiriquement et rigoureuse scientifiquement, alors la recherche s'arrête. Dans le cas contraire, la question de recherche est précisée et un nouveau cycle de recherche démarre. Cette nouvelle itération apporte une deuxième réponse artefactuelle qui est à son tour évaluée, et ainsi de suite.

La stratégie globale d'évaluation des artefacts que j'ai mis en place (et qui s'est déclinée en fonction des spécificités de chaque itération) reprend les travaux de Venable et al. (2012). Ils proposent un processus de décision éprouvé qui permet d'établir une stratégie d'évaluation d'artefact. Je l'ai appliqué à chaque itération. Il était également possible de se baser sur les travaux de Pries-Heje (2008). Or, ces derniers proposent un processus de décision moins exhaustif et des stratégies d'évaluation pour un nombre

¹³ Citation originale : *Building Blocks*

restreint de types de projets de recherche en *design science*. Le processus de Venable et al. (2012) est dès lors plus adapté aux spécificités de cette thèse.

2.2.1.6.1 Construction de la stratégie d'évaluation : analyse de l'évaluation

La première étape du processus de décision consiste à évaluer le contexte et les objectifs de la phase d'évaluation. Venable et al. (2012) préconise de déterminer les sept critères suivants et leur importance relative :

1. **La forme de l'artefact à évaluer** : s'agit-il d'un concept ? d'un modèle ? d'une méthode ? d'une instanciation ? d'une *design theory* ?
2. **La nature de l'artefact à évaluer** : on distingue généralement les artefacts qui prennent la forme d'un produit et ceux qui prennent la forme d'un processus. Les processus se distinguent des produits en ce qu'ils ajoutent aux aspects techniques d'un artefact (technologie utilisée, architecture matérielle ou logicielle) une dimension socio-technique dans la mesure où une interaction avec les humains est nécessaire pour donner à l'artefact toute son utilité.
3. **Les propriétés de l'artefact à évaluer** : s'agit-il d'évaluer l'utilité ? l'efficience ? l'efficacité ? la qualité ?
4. **Les objectifs de l'évaluation** : s'agit-il d'évaluer si l'artefact permet d'atteindre des objectifs fixés au préalable ? de le comparer à d'autres artefacts ? d'évaluer les effets collatéraux ou les conséquences inattendues de l'implantation de l'artefact ?
5. **Les contraintes de l'environnement de recherche** : quelles sont les ressources à disposition : temps ? compétences ? budget ? accès aux terrains ?
6. **La rigueur à appliquer** : s'agit-il d'une évaluation préliminaire ou d'une évaluation rigoureusement détaillée ?
7. **Pondération des critères précédents** : établir une pondération des six critères précédents en fonction de leur pertinence vis-à-vis du projet de recherche. J'ai attribué à chacun des critères un coefficient allant de 1 (=critère non-pertinent) à 5 (=critère essentiel).

2.2.1.6.2 Construction de la stratégie d'évaluation : choix d'une stratégie

Une fois ces éléments identifiés, il s'agit de se reporter au framework de sélection des stratégies d'évaluation des recherche en *design science* construit par Venable et al. (2012, p.11). Il fait le lien entre ces critères et différentes stratégies possibles pour l'évaluation d'un artefact : évaluation *ex ante* ou *ex post*, évaluation naturaliste ou artificielle. Chaque itération se voit alors attribuer un quadrant avec les caractéristiques de la(les) stratégie(s) choisie(s).

Framework de Selection des Stratégies d'Evaluation		Ex ante (avant le prototypage)	Ex post (après le prototypage)
		<ul style="list-style-type: none"> - Evaluation formative - Cout moins élevé - Plus rapide - Evaluation du design, d'un prototype partiel ou complet - Moins de risque pour les participants durant l'évaluation - Risque plus élevé de faux positifs 	<ul style="list-style-type: none"> - Evaluation sommative - Cout plus élevé - Moins rapide - Evaluation des instances - Plus de risques pour les participants durant l'évaluation - Risque moins élevé de faux positifs
Naturaliste	<ul style="list-style-type: none"> - Plusieurs parties prenantes hétérogènes - Conflits substantiels - Artefacts socio-technique - Cout plus élevé - Evaluation longue - Accès au terrain requis - Evaluation de l'utilité de l'artefact - Niveau de rigueur : "On juge l'arbre à ses fruits" - Risque plus élevé pour les participants - Risque de faux positif plus faible 	<ul style="list-style-type: none"> ● Utilisateurs réels, problèmes réels et système irréel ● Cout faible-moyen ● Vitesse correcte ● Risque faible pour les participants ● Risque élevé de faux positifs 	<ul style="list-style-type: none"> ● Utilisateurs réels, problèmes réels et systèmes réels ● Cout le plus élevé ● Risque le plus élevé pour les participant ● Meilleure évaluation de l'utilité ● Identification des effets collatéraux ● Risque le moins élevé de faux positifs
Artificielle	<ul style="list-style-type: none"> - Peu de parties prenantes similaires - Peu ou pas de conflits - Artefacts purement techniques - Cout moins élevé - Evaluation courte - Niveau de rigueur : contrôle des variables - Evaluation de l'efficacité - Risque moins élevé pendant l'évaluation - Risque de faux positif plus élevé 	<ul style="list-style-type: none"> ● Utilisateurs irréels, problèmes irréels et système irréels ● Cout le moins élevé ● Vitesse la plus élevée ● Risque le moins élevé pour les participants ● Risque le plus élevé de faux positifs 	<ul style="list-style-type: none"> ● Système réel, problème irréel et utilisateurs irréels ● Cout moyen/élevé ● Vitesse correcte ● Risque faible-moyen pour les participant

Table 2.2 : Framework de Sélection des Stratégies d'Évaluation (Venable et al., 2012)

2.2.1.6.3 Construction de la stratégie d'évaluation : choix des méthodes

À ce premier framework correspond un second, le framework de sélection des méthodes d'évaluation (Venable et al., 2012, p.14) qui fait correspondre à la stratégie choisie une liste de méthodes d'évaluations appropriées. Il s'agit ensuite de choisir la (ou les) méthode(s) à appliquer à l'artefact.

Framework de Sélection des Méthodes d'Évaluation	Ex ante	Ex Post
Naturaliste	Recherche-action Focus Group	Recherche-action Étude de cas Focus Group Observation Participante Ethnographie Phénoménologie Enquête
Artificiel	Preuve logique ou Mathématique Évaluation basée sur des critères Expérimentation Simulation informatique	Preuve logique ou mathématique Expérimentation Serious Game Simulation informatique

Table 2.3 : Framework de Sélection des Méthodes d'Évaluation (d'après Venable et al., 2012, p.14)

2.2.1.6.4 Construction de la stratégie d'évaluation : design du détail de l'évaluation

Dans cette étape, il s'agit d'articuler les différentes stratégies et méthodes choisies de façon à former un design cohérent. Il s'agit également d'apprendre le fonctionnement des méthodes choisies et de les désigner individuellement.

J'ai, dans cette thèse, conçu des *focus group* permettant d'évaluer l'utilité de la version 0 du prototype (voir section . Ceux-ci ont trois objectifs:

1. Récupérer des données relatives à l'utilité de l'environnement de conception des usages sous la forme de feedbacks qualitatifs
2. Confirmer ou infirmer l'utilité du design de l'environnement de conception des usages, i.e sa capacité à répondre aux 12 besoins fonctionnels
3. Récupérer des feedbacks permettant d'informer la construction d'une nouvelle version du prototype.

Pour atteindre ces trois buts, j'ai mis en place une méthode en sept étapes. Elle suit les travaux de Tremblay, Hevner et Berndt (2010) qui adaptent ;

“les méthodes traditionnelles de mise en place, de conduite et d'analyse de focus group aux spécificités de l'évaluation de prototypes dans les recherches utilisant les méthodologies de recherche en design science”
Hevner et Berndt (2010, p.601).

Ainsi, il s'agira :

1. De formuler la question de recherche à laquelle les *focus group* répondront de manière à pouvoir orienter les contenus présentés aux participants ;
2. D'identifier un échantillon de participants en termes de profils et de volume de participants nécessaires ainsi que le nombre de focus group à mener ;
3. D'identifier un modérateur pour l'animation des *focus group* ;
4. De développer et pre-tester le conducteur, c'est à dire le déroulé des *focus group* de manière à s'assurer que les *focus group* sont bien alignés avec les objectifs de recherche ;
5. De recruter les participants
6. De conduire et animer le *focus group* ;
7. D'analyser et interpréter les données en s'inspirant des méthodes d'analyse de données qualitatives ;
8. De présenter les résultats en fonction de la structure de l'interprétation et avec des verbatim des participants aux *focus group* ;

Formulation de la question de recherche

En lien avec les deux premiers objectifs, la question de recherche à laquelle les *focus group* répondent est la suivante : dans quelle mesure l'environnement de conception des usages permet-il de répondre aux

douze besoins fonctionnels des acteurs ? Autrement dit, dans quelle mesure l'Open Data Canvas est-il en capacité de lever les barrières socio-techniques qui freinent :

1. l'utilisation des fonctionnalités d'une plateforme *open data* (Etape 1 de l'usage des données) ?
2. l'évaluation du potentiel des données à partir de leur représentation (Etape 2 de l'usage des données) ?
3. la transformation des données grâce à des artefacts techniques ?
4. la création d'un cadre cognitif partagé entre les systèmes ?
5. l'institutionnalisation de l'usage des données ?

Identification d'un échantillon de participants

J'ai identifié *a priori* un échantillon de profils de participants en fonction de la cartographie de l'architecture métier de l'Open Data Canvas. Il s'agissait de construire des *focus group* rassemblant des acteurs jouant au moins deux des quatre rôles précisés dans l'architecture, à savoir : un meta-designer (par exemple une collectivité supra-communale ou des directions d'un grand groupe), un producteur de données, un intermédiaire de données et un réutilisateur de données (ces trois derniers rôles faisant partie du groupe des designers d'usage). Le choix d'instaurer un minimum de rôles présents pour pouvoir organiser un *focus group* tient du fait que je souhaitais faire interagir ces rôles de manière à pouvoir analyser des données croisant les points de vue des acteurs sur la version 0 du prototype de l'Open Data Canvas.

Quant au choix du volume de participants et au nombre de *focus group* organisés (5 *focus group* organisés avec un total de 19 participants), il a été guidé principalement par les opportunités et les contraintes auxquelles j'ai fait face : - capacité à trouver des participants intéressés par l'idée de l'Open Data Canvas - temps disponible - absence de budget limitant les possibilités de rencontre des acteurs disséminés sur une grande partie du territoire français.

Identifier un modérateur

J'ai joué le rôle de modérateur des *focus group*. Pour cela, j'ai pris appui sur mon expérience d'animation de *focus group* accumulée en amont de la thèse. J'ai également pris le rôle de conducteur développé ci-après.

Développement et pre-test du conducteur

Pour pouvoir animer les *focus group*, un conducteur a été développé (Annexe 14). La première version du conducteur a été co-construite avec la coopérative Dataactivist et la Fondation Internet Nouvelle Génération (FING). Il a été rapidement mis à l'épreuve avec un premier *focus group* pilote mené avec la métropole de Lille. Sur la base des résultats, nous avons révisé le conducteur. Ce conducteur révisé a été utilisé dans cette forme pour les cinq *focus group*.

Plusieurs éléments composent le conducteur :

- une présentation PowerPoint du modèle théorique de l'usage des données reprenant les schémas des cinq étapes de l'usage des données et des descriptions des types de barrières rencontrées dans chacune d'entre elles
- un guide d'entretien (uniquement à disposition du modérateur) basé sur l'expérience de Dataactivist et de la FING avec une liste exhaustive des tâches cognitives et comportementales fréquemment réalisées dans les projets de développement des usages.
- le diagramme de l'architecture métier de la version 0 du prototype de l'Open Data Canvas
- un template permettant de classer les notes prises lors des *focus group* et plus particulièrement de hiérarchiser les feedback sur la version 0 du prototype de l'Open Data Canvas en fonction de l'importance des besoins auxquels il répond.

Ma stratégie a été d'amener les participants à dérouler, dans l'ordre chronologique, un projet de développement des usages qu'ils ont piloté. À chaque étape de leur récit, il s'agissait pour moi de les inviter à décrire le plus précisément possible les tâches qu'ils ont dû réaliser et les barrières qu'ils ont rencontrées. Lorsque ces points d'achoppements émergeaient dans la conversation, je les amenais à me décrire les solutions mises en place pour les résoudre ou les dépasser. Une fois ces éléments notés, je leur demandais de les classer par ordre d'importance : de la barrière la plus pénalisante pour la suite du projet, à la moins pénalisante.

L'espace problème des participants ayant été clairement saisi, je pouvais présenter l'Open Data Canvas en tant que solution potentielle. Je présentais alors la cartographie de l'architecture métier de la version 0 du prototype. Cette présentation me permettait à la fois d'éclairer les participants sur la refonte de la conduite d'un projet de développement des usages proposé sur le Canvas mais aussi sur le support technique qu'il

offre. Pour utiliser la langue universitaire, je les amenais à percevoir que le Canvas était de la même nature que les barrières qu'ils ont rencontrées : de nature socio-technique.

Sans leur indiquer au préalable les besoins exprimés auquel le Canvas pourrait répondre, je laissais les participants s'exprimer librement sur le Canvas et leur ressenti sur son utilité. Cette phase d'expression libre m'a permis de récupérer des données qualitatives sur l'utilité du Canvas dans sa globalité. Enfin, je reprenais leur hiérarchie en leur posant une série de question visant à comprendre dans quelle mesure il pensait que le Canvas pourrait répondre à leurs besoins et se substituer aux outils aux outils qu'ils utiliseraient, à défaut de pouvoir bénéficier du Canvas.

Recrutement des participants

Les participants ont été recrutés au sein du réseau Open Data France. J'ai bénéficié de l'appui de la présidente du réseau qui a relayé auprès des 75 membres (principalement des collectivités territoriales) une brève description du projet de recherche, du déroulement et des objectifs du *focus group*. L'invitation au *focus group* précisait que les participants devaient (de préférence) avoir déjà mené un projet de développement des usages des données ouvertes, quel que soit le contexte sous-jacent (mise en place de l'ouverture des données publiques, projet ville intelligente, demande des élus, etc.). Le lecteur trouvera dans la prochaine partie une description des profils des participants.

Conduite et animation des focus group

En premier lieu, les participants étaient informés, en amont du *focus group*, du déroulé, du fait que les échanges seraient enregistrés et qu'ils pouvaient communiquer au modérateur les éventuels points qu'ils souhaitaient aborder. Ensuite, au lancement d'un *focus group*, le modérateur commençait par demander une brève présentation du profil des participants ainsi qu'une description du projet d'usage qui fera l'objet de la session.

Le modérateur présentait ensuite à l'écran une présentation Powerpoint avec les schémas du cadre théorique de l'usage des données ouvertes. Il présentait chacune des cinq étapes de l'usage, les barrières identifiées avec des exemples concrets rencontrés à Grenoble et Lyon.

Les participants étaient ensuite invités à décrire leur projet d'usage à la manière d'un déroulé narratif, en insistant sur les barrières rencontrées. Les cinq étapes étaient toujours à l'écran pour que les participants

aient en tête l'ensemble des éléments qu'ils pouvaient amener dans le projet. Au cours de la discussion, le modérateur s'assurait que le récit des participants correspondait avec la chronologie de la conduite du projet de développement des usages. Aussi, sans en informer les participants, le modérateur avait à disposition un guide d'entretien avec une liste comportant l'ensemble des tâches couramment effectués lors des cinq étapes de l'usage. Cette liste était utilisée de manière à pouvoir scanner avec les participants un maximum de tâches et faire émerger un maximum de barrières.

Le modérateur pouvait ensuite présenter le diagramme cartographie de l'architecture métier de l'Open Data Canvas (sur l'outil de diagramme Lucidchart) en précisant qu'il s'agissait d'une version intermédiaire d'un artefact qui vise, dans l'avenir, à répondre à un maximum des besoins fonctionnels qu'ils venaient d'évoquer dans récit. La discussion qui s'en suivait laissait les participants donner leur idée sur l'utilité de l'architecture métier et la viabilité de sa mise en place dans leur organisation.

Passant du général pour aboutir au particulier, le modérateur résumait les barrières évoquées lors de leur récit et demandait aux participants de les hiérarchiser par ordre d'importance. Cette hiérarchie établie, le modérateur posait une liste de question permettant de savoir s'il percevait le Canvas comme une solution à chacune des barrières et dans quelle mesure il pouvait se substituer aux solutions que les participants auraient avaient utilisé.

Analyse et interprétation des données

L'analyse s'est basée sur la retranscription intégrale des enregistrements des *focus group*, le template ayant permis de hiérarchiser les notes prises (Annexe 14, slide 4), ainsi que sur les retours par mail que les participants m'ont envoyé suite à la transmission du résumé des échanges.

La première étape de l'analyse a consisté à coder les retranscriptions avec Nvivo. J'ai créé un système de code (Annexe 15) avec pour premier niveau la liste des cinq étapes de l'usage des données ouvertes, comme deuxième niveau les barrières socio-techniques relatives à chacune d'entre elles, et comme troisième niveau les preuves ou contre-preuve de l'utilité du Canvas. Ces codes ont permis de classer toutes les données relatives aux tâches que les participants ont effectué et les barrières concrètes qu'ils ont rencontrées dans leur projet. J'ai ensuite créé une table (un extrait se trouve au Chapitre 7, Table 7.6, p.469) qui classe par étape d'usage les barrières rencontrées. Ce codage a été effectué suivant les préceptes de la une théorie ancrée¹⁴. J'ai commencé par utiliser la premier niveau de code pour rattacher

¹⁴ Termes originaux : *grounded theory*

les entretiens aux cinq étapes de l'usage des données. Dans un deuxième temps, au sein de chaque unité de sens de premier niveau, j'ai analysé à nouveau les données pour les rattacher à un besoin fonctionnel (deuxième niveau de code). Enfin, dans un troisième temps, j'ai analysé les données affiliées à chaque besoin fonctionnel pour extraire les preuves et contre-preuves de l'utilité ou de l'inutilité du Canvas (troisième niveau de code)

La deuxième étape a consisté à compléter la table de classement et de synthèse des *focus group* (Chapitre 7, Table 7.6, p.469). J'ai rattaché à chaque barrière rencontrée par les participants, leurs réponses quant à l'utilité du Canvas pour les lever. Au final, cette table synthétise les données suivantes :

- étapes d'usage
- tâches effectuées
- barrières rencontrées
- utilité du Canvas ? (Oui Non)
- Justification de l'utilité
- Outils utilisés auparavant
- Substitution possible ? (Oui Non)
- Suggestions annexes.

Présentation des résultats

Les résultats finaux de l'analyse consistent en un tableau qui établit si oui ou non, le Canvas permet de répondre à chacun des 12 besoins fonctionnels, les verbatim qui m'ont permis de tirer cette conclusions, ainsi que les verbatim qui montrent les limites du Canvas pour répondre aux besoins fonctionnels.

2.2.1.6.5 Construction de la stratégie d'évaluation : réalisation des évaluations

Les méthodes choisies sont mises en place selon le design choisi. Il faut collecter les données avant de les analyser et de tirer des conclusions qui visent à savoir dans quelle mesure l'artefact a répondu à la sous-question de recherche qui structure l'itération.

2.2.1.6.6 Construction de la stratégie d'évaluation : implication sur la prochaine itération

Dans cette dernière étape, j'évalue la pertinence empirique et la rigueur scientifique de l'artefact avant de noter spécifiquement les manquements relatifs à ces deux aspects qui sont à améliorer dans la prochaine itération.

Chapitre 3

Les barrières à la construction des usages issus des données ouvertes

Chapitre 3. Les barrières à la construction des usages issus des données ouvertes

Eu égard à la méthodologie *design science* présentée dans le chapitre précédent, ce troisième chapitre vise à mettre en place sa première étape, c'est à dire établir l'identification des situations problématiques auxquelles les usagers font face et qui les bloquent dans leurs démarches de construction. Ce chapitre sera guidé par la deuxième sous-question de recherche :

Quelles sont les situations problématiques qui freinent la construction de nouveaux usages issus des données ouvertes ?

Je rappelle que dans le premier chapitre, je suis arrivé à la conclusion que la construction des usages issus des données ouvertes constituait un problème pernicieux¹ (Rittel et Weber, 1973 ; Churchman, 1967). Aussi ai-je fait le choix de circonscrire le périmètre d'étude de cette thèse aux écosystèmes *open data*.

De ce fait, dans ce troisième chapitre, je me suis rapproché des opérateurs des données ouvertes (Denis, 2018) qui animent ces écosystèmes à savoir les producteurs, éditeurs et réutilisateurs de données ouvertes. Ce sont les praticiens auxquels cette thèse s'adresse et que j'ambitionne de pouvoir assister grâce à nos résultats.

Pour identifier les situations problématiques qu'ils rencontrent et répondre à la question de recherche, j'ai mis en place une démarche d'exploration hybride (Thiétart et al., 2014, p. 93) en utilisant la méthode des cas (Hlady-Rispal, 2002) qui allie, de manière itérative des observations participantes et des entretiens (voir Chapitre 2). Le lecteur pourra se référer à la figure 0.4 de l'introduction générale pour suivre les différentes étapes que j'ai empruntées.

¹ Termes originaux : *wicked problem*

3.1 Invention et conceptualisation des difficultés rencontrées dans l'innovation de services issue des données ouvertes.

Résumé

Pour inventer un modèle préliminaire de l'usage des données ouvertes dans un écosystème, qui puisse mettre en lumière les difficultés rencontrées dans l'innovation de services, je m'inspire ici du pattern d'analyse de système complexe (Vaishnavi et Kuelcher, 2015, p. 107). Il est utile au chercheur en science de gestion lorsqu'il cherche à décomposer le fonctionnement d'un système complexe. Il se décompose en quatre phases (voir chapitre 2 pour un exemple d'analyse de système complexe) :

1. Invention de la structure statique de l'écosystème en la décomposant en différents systèmes interdépendants ;
2. Invention de la structure statique de chaque système en les décomposant en subdivisions plus petites ;
3. Invention du comportement dynamique de chaque système et de la manière dont ce comportement est produit ;
4. Invention de l'évolution historique des systèmes et de leurs interactions.

Précisons que les concepts constitutifs de ce modèle, et leur applicabilité aux situations réelles d'innovation de services issue des données ouvertes, sont présentés de manière exhaustive dans le glossaire. Dans cette partie, il est simplement question de les présenter brièvement ainsi que de poser quelques postulats quant à notre méthodologie de modélisation de ces concepts. Pour illustrer nos propos, j'ai décidé de prendre un exemple simple d'usage des données ouvertes mais qui a l'avantage de couvrir la totalité du modèle conceptuel proposé. J'utiliserai cet exemple tout au long du chapitre.

Je prendrai dans cette partie l'exemple d'une métropole qui, au titre de la Loi pour une République numérique, décide d'ouvrir l'annuaire de tous les services publics actifs sur son territoire dit guichets publics locaux. On nommera ce jeu de données "Annuaire". Ces données fournissent principalement les coordonnées (adresses, téléphones, site internet, horaires d'ouverture, coordonnées et géolocalisation) et le type de service rendu dans chacun de ses guichets. Ici, la métropole dispose de trois types de services :

des services de police répartis dans différents commissariats, des services fiscaux situés dans le centre des impôts et des services judiciaires accessibles depuis le Tribunal.

Pour pouvoir mettre ces données à disposition, la métropole choisit de créer une plateforme dédiée “data.metro.fr” sur laquelle seront déposées ces données et mises en libre accès. Cette plateforme est mise à disposition de tous les acteurs privés ou civils qui souhaiteraient diffuser leurs données en *open data*. La responsabilité de diffusion de ces données échoit à la Direction des Systèmes d'Information de la métropole (DSI Métro). Cette direction a pour mission de coordonner les actions des départements des Systèmes d'Information de la Police (DSI Police, des Impôts (DSI Impôts) et de la Justice (DSI Justice) pour pouvoir créer et diffuser l'Annuaire, d'assurer la maintenance de la plateforme “data.metro.fr”, avant de communiquer et faire connaître auprès de l'écosystème numérique l'Annuaire pour que des acteurs privés ou des acteurs issus de la société civile les utilisent.

En l'occurrence, après avoir mis en place cette plateforme, l'Annuaire est utilisé par la commune ABA, rattachée à la métropole, qui l'injecte dans l'application de gestion de projet (AppG) de la mairie pour que tous les agents puissent trouver rapidement le numéro des services de Police, de Justice et des Impôts en cas de besoin. L'entreprise CBC, quant à elle, profite de ces données pour compléter l'application “ServicesPublic&Moi”(SP&M) qu'elle commercialise et qui permet d'obtenir l'itinéraire le plus rapide pour se rendre dans le commissariat de Police, le Tribunal ou le centre des impôts le plus proche. Enfin, Paul B. s'empare également de ces données pour créer une carte interactive (Cartel) sur laquelle sont disposés tous les services de la métropole.

Conclusion et transition

Cet exemple établi, je peux commencer à inventer les différentes strates qui composent l'usage des données ouvertes. Je commence par inventer la structure globale statique, à l'œuvre dans l'usage des données ouvertes.

Cette structure globale se comprend grâce aux concepts de plateforme *open data*, de parties prenantes, d'espèces ingénieurs, l'écosystème *open data*, de système de production, d'intermédiation et de transformation de données. Je définirai ces termes. Ces termes sont largement utilisés dans la littérature sur les données ouvertes et que cette partie ne fait que les adapter. Une fois que ces termes seront définis, je les illustrerai grâce à l'exemple type.

3.1.1 Invention de la structure globale statique à l'œuvre dans l'usage des données ouvertes.

Définition : Plateforme open data : structure modulaire composée d'éléments tangibles et intangibles qui facilitent l'interaction des acteurs et des données ouvertes. Selon Lusch et Nambisan (2015), une plateforme permet de maximiser la densité des données ouvertes à un endroit donné et d'augmenter leur liquidité soit leur circulation entre parties prenantes.

Une plateforme *open data* se définit par trois éléments tangibles et intangibles : une architecture qui peut être ouverte ou fermée, un ensemble de composants stables et des mécanismes de gouvernance.

Concernant l'architecture, elle a une incidence sur la possibilité pour des acteurs externes de contribuer à son développement. Dans le cas où elle serait ouverte, des acteurs externes pourraient librement ajouter des modules à la plateforme pour permettre, par exemple, de visualiser les données ouvertes.

Les composants stables sont le catalogue de jeux de données ouverts ainsi que les modalités d'accès à ce catalogue qui peuvent être de deux ordres : par téléchargement ou par une interface de programmation applicative². Les mécanismes de gouvernance constituent les règles instaurées par le propriétaire de la plateforme pour assurer la gestion à la fois de l'approvisionnement en données (offre) et l'utilisation des données (demande).

Ces plateformes sont conçues ici comme des infrastructures qui permettent à l'usage des données ouvertes de naître et de se développer. Outre le fait qu'elles sont le moyen technique de diffusion des données ouvertes, elles constituent de véritables infrastructures d'innovation (Boudreau, 2007). Dans notre exemple, pour pouvoir utiliser l'Annuaire, il faut en premier lieu y avoir accès. Pour cela, il faut se rendre sur la plateforme "data.metro.fr". Celle-ci constitue en premier lieu un moyen de diffusion de l'Annuaire accessible selon différentes modalités. On peut télécharger les données grâce à des liens ou bien y accéder de manière continue grâce à des interfaces de programmation applicative APIs. C'est grâce à cette plateforme que la commune ABA, l'entreprise CBC et Paul B. peuvent injecter l'Annuaire dans une application, un logiciel ou dans tout système d'information qui pourrait bénéficier de cette nouvelle ressource. La figure ci-dessous illustre notre conception de la place de la plateforme *open data* dans l'usage des données ouvertes à partir de l'exemple de l'Annuaire diffusé sur la plateforme "data.metro.fr".

² Termes originaux : Application Programming Interface

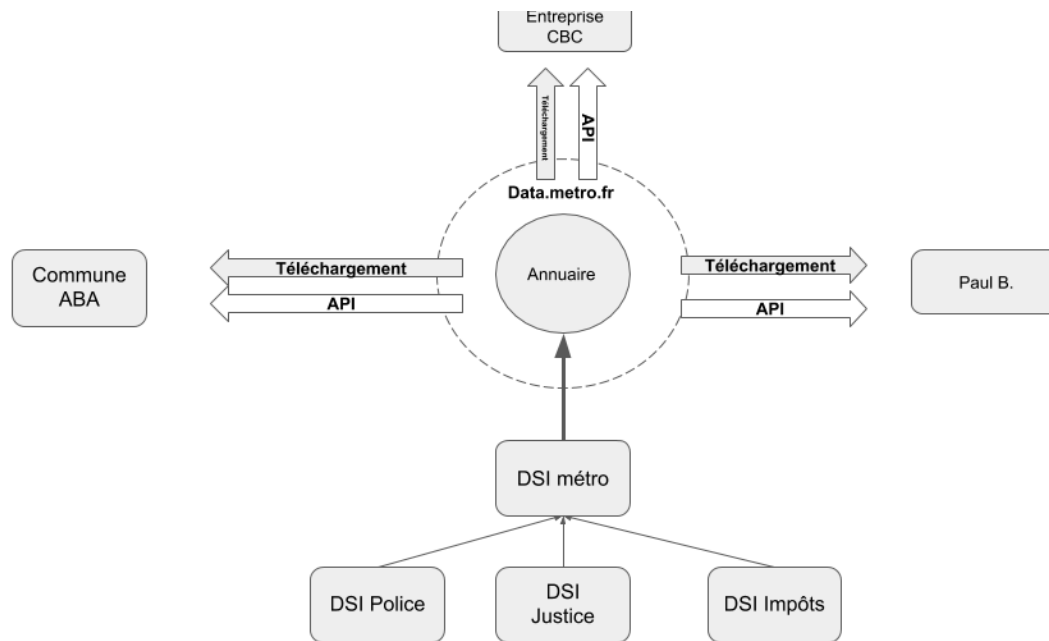


Figure 3.1 : Place de la plateforme open data “data.metro.fr” dans la structure globale de l’usage des données ouvertes.

Conclusion et transition

Le rôle de la plateforme étant défini, il s’agit désormais de définir le rôle de toutes les organisations qui gravitent autour de la plateforme comme le représente la figure ci-dessus. Ce rôle est conçu grâce au concept de **partie prenante** de plateforme.

Définition : Partie prenante de plateforme : Organisation ou individu intéressé par au moins un des trois éléments d’une plateforme et pouvant jouer un ou plusieurs rôles dans l’offre, la demande ou la maintenance des données ouvertes. Aussi, ils supportent proactivement le processus d’innovation de services qui peut en découler en établissant de nouveaux mécanismes organisationnels ou en influençant les mécanismes de gouvernance internes à la plateforme.

On identifie les parties prenantes des plateformes par leur relation directe avec la plateforme, soit en tant que fournisseur (côté offre) ou en tant qu'utilisateur (côté demande) de modules et de données. Ces parties prenantes se caractérisent par la pluralité de rôles qu'elles peuvent jouer dans le fonctionnement de la plateforme. On distingue trois rôles différents : le rôle d'architecte, d'acteur et de bénéficiaire. L'architecte participe à l'architecture technique de la plateforme (apport de modules ou de données) ainsi qu'au développement de mécanismes de gouvernance qui clarifient la façon dont l'innovation de services se déroule. L'acteur met à disposition ses moyens techniques et ses compétences pour exploiter les modules et données présentes sur la plateforme et participe ainsi à l'innovation de services. Le bénéficiaire fait, quant à lui, l'expérience de la valeur créée par la plateforme. On distingue quatre types d'expériences : l'expérience pragmatique, l'expérience socialisante, l'expérience d'usage et l'expérience hédonique (Nambisan & Nambisan, 2008). Cette valeur se matérialise finalement pour les bénéficiaires par une amélioration de leurs activités propres.

Dans notre exemple, les DSI Métro, DSI Police, DSI Justice et DSI Impôts sont des parties prenantes "fournisseur" puisqu'ils fournissent les données présentes sur la plateforme. La Commune ABA, l'entreprise CBC et Paul B. sont quant à eux des parties prenantes "utilisateur". Du côté des parties prenantes fournisseurs, la DSI Métro est la seule à assurer le rôle d'architecte puisqu'elle maintient la plateforme "data.metro.fr" et qu'elle coordonne les actions des autres DSI via des mécanismes de gouvernance. Du côté des parties prenantes utilisateurs, toutes les organisations et Paul B. jouent ici le rôle de bénéficiaire faisant plus particulièrement l'expérience pragmatique et l'expérience d'usage de la plateforme "data.metro.fr".

Conclusion et transition

Maintenant que la place et le rôle des différentes organisations et individus gravitant autour des plateformes ont été définis, il s'agit de s'intéresser à l'effet de l'usage des données ouvertes à la fois sur les utilisateurs mais aussi sur les producteurs. Cet effet sera défini grâce au concept d'espèce ingénieur. Notons que ce concept est le seul à être importé d'une autre discipline que les systèmes d'information, à savoir l'écologie. J'exposerai plus en détail les raisons de cet emprunt transdisciplinaire.

Définition : Espèce ingénieur : aussi appelée “espèce clé de voûte”, ce concept issu de l’écologie décrit les espèces qui structurent et assurent l’équilibre d’un écosystème (Byers et al., 2006).

Le concept d’espèce ingénieur est utilisé en écologie pour décrire les espèces qui, à l’image des loups gris:

“interagissent avec de nombreux éléments d’un environnement et dont l’impact sur la diversité et la compétition entre ces espèces est disproportionné par rapport au volume de biomasse qu’elle représente”
(Davic, 2003)

Je pose le postulat selon lequel les données ouvertes constituent l’espèce ingénieur de l’écosystème qui se forme autour des plateformes. En effet, les données ouvertes agissent à la manière d’un patrimoine informationnel sur lequel est basé tout un ensemble d’outils informatiques. Autrement dit, elles sont des ressources qui irriguent et permettent le fonctionnement d’outils informatiques. Sans elles, ces outils se verraient obligés d’utiliser d’autres données équivalentes. Des données qui seraient moins fiables car ne provenant pas des organisations qui les ont produites. Avec elles, ces outils donneraient de fausses informations qui pourraient induire en erreur les personnes ou organisations qui consomment ces informations. D’une certaine manière, les données ouvertes structurent un écosystème informatique comme les loups gris structurent certains écosystèmes naturels. La présence de loups gris assure le fonctionnement d’une chaîne écologique stable : ils éloignent les herbivores de zones boisées, lesquelles sont le grenier des castors, qui peuvent construire des barrages qui retiennent le cours de l’eau, permettant ainsi aux sols de se renouveler et aux espèces qui en dépendent de perdurer. De manière analogue, les données ouvertes assurent également le fonctionnement stable d’une chaîne mais qui est dans ce cas informationnelle.

Conclusion et transition

Pour désigner la chaîne informationnelle par laquelle passent les données, j’utiliserai le concept d’écosystème *open data*, lui même basé sur celui, plus générique, d’écologie informationnelle. J’adapterai ici les définitions instaurées par Harrison et al. (2012), Ruijer et al. (2017) et Nardi and O’Day (1999).

Définition : Écosystème open data : Écologie informationnelle dont l’équilibre repose sur la circulation de jeux de données ouvertes qui agissent à l’image d’une “espèce ingénieur“. Cette écologie est constituée de

trois systèmes : le système de production de données, le système de transformation de données en services informationnels et le système d'intermédiation qui coordonne les interactions entre les deux.

Dans mon exemple, illustré par la figure ci-dessous, l'application "ServicesPublic&Moi" de l'entreprise CBC, la carte interactive de Paul B. et l'application de gestion de projet de la commune ABA se structurent autour de l'Annuaire (1). Ces trois acteurs forment ensemble le système de transformation de données en services informationnels. Ce système a pour rôle de transformer les données ouvertes pour créer des services informationnels permettant de transformer les comportements dans le sens des intérêts collectifs. Si ce système est indépendant, il sera néanmoins en interaction directe avec le reste de l'écosystème. Plus particulièrement et du fait que ce système base des outils sur l'Annuaire, ce système va demander à ce que la qualité de l'Annuaire soit garantie, c'est-à-dire que les coordonnées soient mises à jour et que les informations qui s'y trouvent soient certifiées comme étant exactes. Alors, le système de transformation de données entrera en interaction avec la DSI Police, la DSI Justice et la DSI Impôts pour obtenir cette garantie (2).

Ces trois DSI constituent alors un système de production de données ouvertes. Ce système a un rôle de fournisseur de données ouvertes. Il récupère les données produites par les systèmes d'information pour les transformer en données ouvertes. Dans cet exemple, du fait de la demande du système de transformation, les trois DSI vont interroger régulièrement leur système d'information pour trouver les coordonnées des guichets qu'ils mettent à disposition et créer des jeux de données avec elles. Comme le système de transformation, le système de production est indépendant mais il est également en interaction directe avec le reste de l'écosystème. Plus particulièrement, le système de production va transmettre les données à la DSI de la métropole qui va se charger de combiner ces données et de les intégrer à la plateforme "data.metro.fr"(3)

La DSI de la métro constitue alors le système d'intermédiation de l'écosystème. Il a un rôle d'intermédiaire qui consiste à coordonner la production (offre) et l'utilisation (demande) de données ouvertes. Dans ce même exemple, ce système d'intermédiation accompagnera le système de production dans la réponse qu'il apportera à la demande du système de transformation, à savoir la garantie sur la qualité et la mise à jour de l'Annuaire. Il pourra se charger notamment de nettoyer les données transmises par le système de production ou d'automatiser leur mise à jour (4).

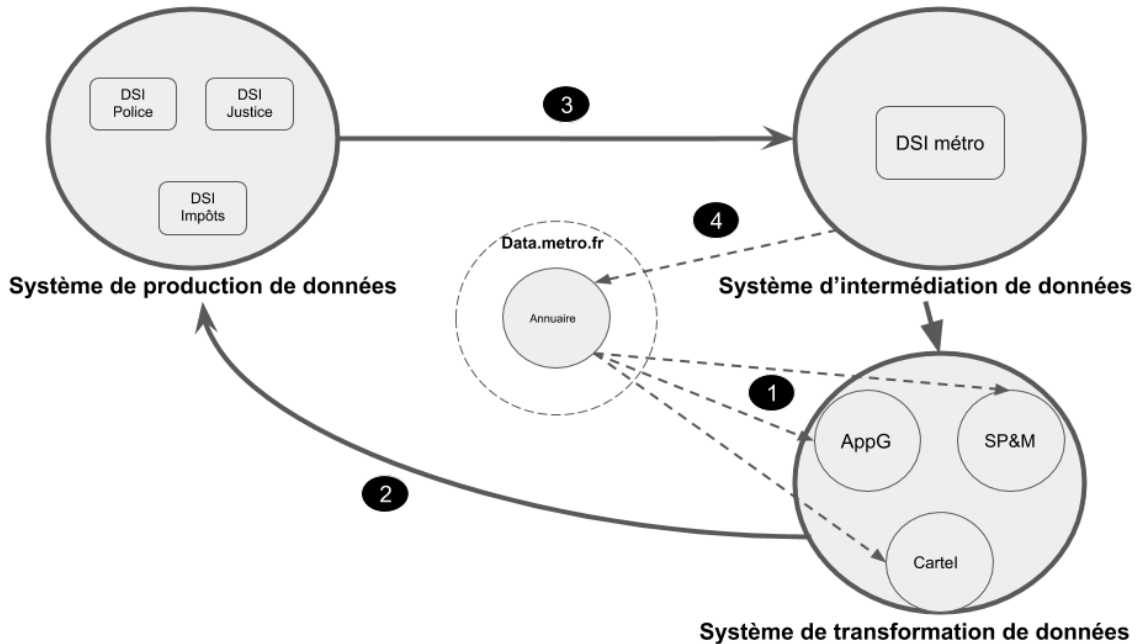


Figure 3.2 : Fonctionnement de l'écosystème open data autour de l'Annuaire

Conclusion et transition

De cette partie, je conclus que les concepts de plateforme *open data*, de parties prenantes, d'espèce ingénieur, d'écosystème *open data*, de système de production, système d'intermédiation et système de transformation constituent la structure globale à l'œuvre dans l'usage des données ouvertes. Après avoir défini la structure globale à l'œuvre dans l'usage des données ouvertes, il s'agira d'aborder la deuxième étape. Elle consiste à appliquer le même raisonnement qu'à la première étape mais à l'échelle des systèmes. En l'occurrence, il s'agit de proposer une liste d'éléments qui forment la structure de chaque système. Cette structure se comprend grâce aux concepts d'artefact technologique, d'artefact social, de capacité d'action (agency) et d'état de la donnée. Je décrirai ces concepts en adaptant les définitions proposées par Kuk et Davies (2011) puis je les aborderai en adaptant la représentation graphique proposée également par Kuk and Davies (2011). Une fois ces termes définis et représentés, je les illustrerai grâce à l'exemple type.

3.1.2 Invention de la structure statique des systèmes à l'oeuvre dans l'usage des données ouvertes

Dans chacun des trois systèmes d'un écosystème *open data*, les parties prenantes créent et utilisent des artefacts technologiques pour jouer leur rôle et utiliser les données.

Définition : Artefact technologique : Objet ou solution technique conçu dans le but d'accomplir une certaine tâche ou d'induire un changement comportemental chez les utilisateurs. L'utilité d'un artefact technologique est mesurée à l'aune de sa capacité à atteindre les objectifs imaginés par ses concepteurs.

J'ai présenté dans le chapitre 2 la convention que j'utilise pour représenter les éléments qui forment les activités de chacun des trois systèmes. Pour rappel, je mobilise une convention de représentation où un artefact technologique est défini par l'ensemble fini [Tat, Oat] où Tat correspond à un type d'artefact et Oat correspond à l'objectif qu'il doit atteindre.

Dans le système de transformation de l'exemple, Paul B. utilise une solution de cartographie de données, comme *geojson.io*, pour pouvoir afficher les données de l'Annuaire sur une carte et la publier. Cette solution, dans la mesure où elle a été conçue pour cartographier des données brutes comme l'Annuaire, est un artefact technologique. Aux côtés de ces artefacts technologiques, il existe des artefacts sociaux.

Définition : Artefact social : Instance, méthode ou processus visant à organiser les agissements d'individus ou de groupes d'individus dans le but d'accomplir une certaine tâche ou d'induire un changement comportemental. L'utilité d'un artefact social est mesurée à l'aune de sa capacité à atteindre les objectifs imaginés par ses concepteurs.

Comme présenté dans le chapitre 2, j'utilise une convention de représentation où un artefact social est défini par l'ensemble fini [Tas, Oas] où Tas correspond au type de l'artefact (instance, méthode ou processus) et où Oas correspond à l'objectif qu'il doit atteindre.

Dans le système d'intermédiation de notre exemple, la DSI de la métropole a créé un groupe de travail auquel participe la DSI de la Police, de la Justice et des Impôts et qui suit un processus pour créer l'Annuaire à partir des "jeux de données métiers" fournis par chacune d'entre elles. Cette instance et le processus qu'elles utilisent sont des artefacts sociaux, dans la mesure où ils ont été conçus pour produire l'Annuaire.

Ces deux types d'artefacts étant définis, il s'agit désormais de poser un postulat quant à leur mode d'action dans chacun des systèmes. Je pose l'hypothèse qu'ils agissent grâce à leur capacité d'action (agency).

Définition : Capacité d'action ou agency : Capacité d'une entité sociale d'agir indépendamment de toute structure ou capacité d'une entité artificielle d'agir indépendamment de toute intervention humaine. À noter que la capacité d'action d'une entité s'exprime à travers sa performativité, c'est à dire sa capacité à faire advenir ce pour quoi il a été conçu (Leonardi, 2011).

Dans le cas de l'usage des données ouvertes, si la capacité d'action d'un artefact technologique ou social est indépendante, elle influence néanmoins l'état des données ouvertes. Dans cet exemple, le groupe de travail mis en place par la métropole va permettre à trois jeux de données à l'état de "données métiers", de devenir un seul jeu de données à l'état de "donnée ouverte", à savoir l'Annuaire. Avant de continuer, il convient de définir le concept "d'état d'un jeu de données".

Définition : État d'un jeu de données : configuration dans laquelle se trouve un jeu de données. Un état peut être conçu comme une image instantanée d'un jeu de données.

Il est possible de distinguer l'état initial (donnée métier), l'état final (nouvel usage) et les états intermédiaires d'un jeu de données. Le rôle de chacun des trois systèmes consiste à partir d'un jeu de donnée dans un état initial (Ei) et de l'amener à un état final (Ef) en passant par un état intermédiaire (Ei1, Ei2, Ei3, Ein)

Les concepts d'artefact technologique, d'artefact social, de capacité d'action et d'état d'un jeu de données étant définis, il est possible désormais de présenter la méthode de représentation graphique de la structure statique des systèmes à l'oeuvre dans l'usage des données ouvertes et de l'appliquer à l'exemple.

Comme l'indique la figure ci-dessous, il s'agit dans chaque système de faire passer un jeu de données dans un état initial (Ei) à un état final (Ef) en passant par des états intermédiaires. Pour cela, le système crée et utilise des artefacts technologiques (Tat, Oat) et des artefacts sociaux (Tas, Oas). En fonction de l'état de la donnée en entrée de l'artefact, il met en oeuvre une capacité d'action qui fait passer la donnée à l'état suivant. Ce schéma ou *pattern* se répète jusqu'à ce que le jeu de données soit dans l'état final (Ef).

Dans le système de production de l'exemple représenté dans la figure, la donnée se présente dans l'état Ei qui correspond aux trois jeux de données métiers de la DSI Police, Justice et Impôts. L'objectif est de passer à l'état de données ouvertes (Ef). La DSI de la métropole crée et utilise alors un groupe de travail. Ce groupe de travail met en oeuvre sa capacité d'action qui consiste à fusionner les trois fichiers. Ainsi la

donnée passe de l'état E_i à l'état E_{i1} . À ce moment, la DSI de la métropole va télécharger le jeu de données sur la plateforme "data.metro.fr" qui constitue un artefact technologique. Cette plateforme va mettre en œuvre sa capacité d'action qui consiste à lier un jeu de données à une licence d'utilisation qui lui donnera le statut et l'état de données ouvertes. Ainsi la donnée passe de l'état E_{i1} à l'état E_f et le rôle du système de production est assuré.

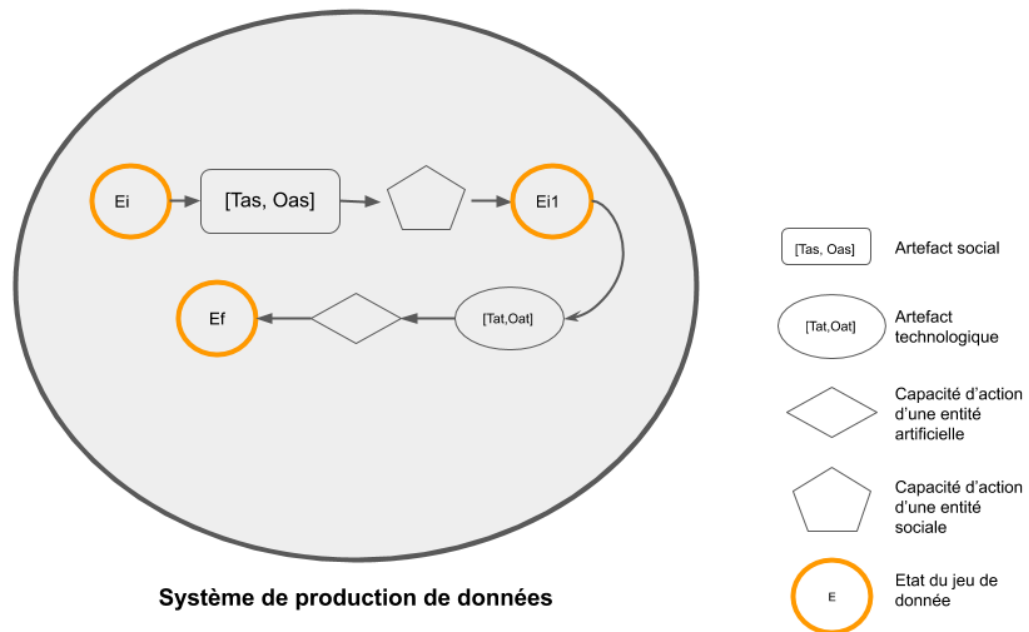


Figure 3.3: Représentation graphique de la structure statique des systèmes à l'œuvre (adapté de Kuk et Davies, 2011)

Conclusion et transition

Grâce aux deux premières étapes, je dispose désormais d'une représentation de la structure statique à l'œuvre dans l'usage des données ouvertes au niveau écosystémique et systémique. Désormais, il s'agit de poser des postulats quant à la manière dont cette structure se met en mouvement, et produit le comportement qui aboutit à des innovations de services. J'ai choisi de reprendre ici les concepts de séquence d'articulation et de complémentarité que j'adapterai à partir de la conception de Kuk et Davies (2011).

3.1.3 Invention du comportement dynamique des systèmes et de la manière dont ce comportement est produit

Définition : Séquence d'articulation : Unité logique de l'usage des données ouvertes, une séquence d'articulation est une suite particulière de capacités d'action sociale (human agency) et technologique (material agency) dans laquelle l'une et l'autre interagissent, se renforcent mutuellement et mettent en place des conditions propices au changement d'état d'une entité ou d'une organisation.

Le passage d'un état initial du jeu de donnée à un état final se produit grâce à une ou plusieurs séquences d'articulation entre des capacités dites d'action sociale et des capacités d'action technologique. Si bien que le passage d'un état E_i à un état E_f résulte toujours des interactions et du renforcement mutuel des capacités de plusieurs artefacts.

Dans le système de production de données de l'exemple, c'est l'interaction entre la capacité d'action sociale du groupe de travail et la capacité d'action technologique de la plateforme *open data* qui a permis à l'Annuaire de passer de l'état E_i "données métiers" à l'état E_f "données ouvertes".

Je présume également que l'état final du système de production constitue l'état initial du système d'intermédiation ; que l'état final du système d'intermédiation constitue l'état initial du système de transformation ; finalement, que l'état final du système de transformation s'avère être l'innovation de service.

Aussi, je présume que le comportement dynamique de l'ensemble de l'écosystème se conçoit également grâce au concept de séquence d'articulation à la différence près que ce ne sont pas les capacités d'action qui sont articulées d'une certaine manière mais les rôles de chaque système. Ainsi, il existe des suites particulières de rôles dans lesquelles ils interagissent, se renforcent mutuellement et que de cette interaction émergent les conditions propices au changement d'état de l'écosystème.

Pour préciser encore ce que produisent ces séquences d'articulation aussi bien au sein des systèmes que de l'écosystème, je mobiliserai la notion de complémentarité.

Définition : Complémentarité : Concept issu de la théorie économique éponyme (théorie des complémentarités) et qui décrit le résultat d'activités coordonnées et se renforçant mutuellement. Ce résultat possède la caractéristique particulière d'être supérieur à la somme de ses parts.

Ce concept admet deux postulats implicites. Le premier est que seules les activités coordonnées offrent les conditions optimales pour voir émerger des complémentarités. Le deuxième est que les activités qui

sont coordonnées sont indépendantes et peuvent perdurer l'une sans l'autre. Je nuancerai ces postulats dans la dernière étape de cette partie.

Ainsi, les états intermédiaires et finaux des jeux de données constituent des complémentarités, ce qui inclut les usages en tant qu'ils constituent l'état final du système de transformation. De la même manière, les changements d'état de l'écosystème sont des complémentarités. J'illustrerai ces concepts dans la figure ci-dessous.

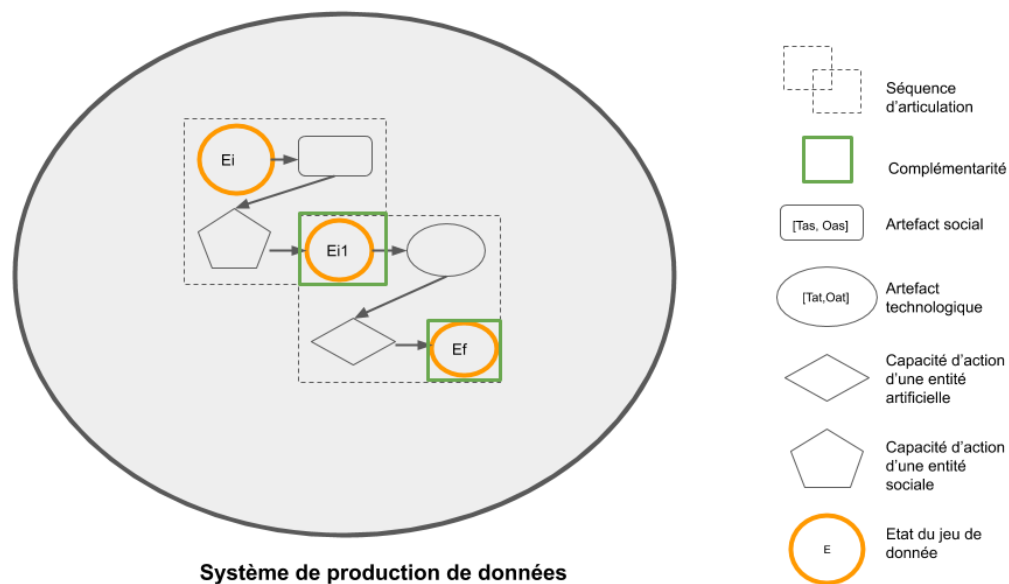


Figure 3.4 : Représentation graphique du comportement dynamique au sein des systèmes

Conclusion et transition

Maintenant que le cadre conceptuel permettant de comprendre le comportement dynamique du système et de l'écosystème est établi, il est opportun de poser quelques postulats concernant l'évolution historique de la structure globale. Pour pouvoir comprendre cette évolution historique j'utiliserai le concept de cycle de vie d'un écosystème et préciserai le concept d'espèce ingénieur déjà utilisé auparavant. J'adapterai, une nouvelle fois, les définitions issues de la littérature sur les écosystèmes informationnels et biologiques.

3.1.4 Invention de l'évolution historique de la structure globale

Définition : Cycle de vie d'un écosystème : Le cycle de vie d'un écosystème comprend quatre phases cycliques : la genèse, l'expansion, le leadership et l'auto-renouvellement (adapté de Moore, 1993, 1996).

Pour les écosystèmes *open data*, la phase de genèse définit la période pendant laquelle les parties prenantes de la plateforme *open data* construisent son architecture, réfléchissent à des mécanismes de gouvernance et mettent à disposition leur premier jeu de données. La phase d'expansion correspond au moment où les jeux de données sont de plus en plus utilisés par les acteurs et les artefacts technologiques de plus en plus hétérogènes. Commence alors la structuration d'un certain écosystème. Durant la phase de leadership, les parties prenantes de la plateforme identifient et formalisent les systèmes qui constituent l'écosystème, mettent en place des processus relatifs à l'usage des données ainsi qu'une gouvernance pour l'évolution de la plateforme et de ses composants stables (données et modules). La phase finale de renouvellement est une phase où l'écosystème doit s'auto-renouveler pour faire face à une menace extérieure ou bien pour trouver de nouvelles ressources sans lesquelles il disparaît.

Au cours de ces différentes phases, le statut des données ouvertes évolue. Pendant la phase de genèse et d'expansion, les données ouvertes occupent la place d'une véritable espèce fondatrice tandis que dans les phases de leadership et d'auto-renouvellement, elles occupent la place et le rôle d'une espèce ingénieur. Il convient de définir le concept d'espèce fondatrice par rapport à l'espèce ingénieur avant d'illustrer l'évolution historique d'un écosystème.

Définition : Espèce fondatrice et Espèce ingénieur : Selon les écologistes, les espèces fondatrices possèdent trois caractéristiques distinctives : elles sont numériquement abondantes et constituent la majorité du volume de biomasse ; ensuite, elles sont situées à la base des réseaux qui caractérisent les écosystèmes ; enfin, les nombreuses interactions qu'elles entretiennent avec les autres espèces sont de nature non-trophiques et mutuelles en ce sens qu'elles consistent principalement à leur apporter un soutien structurel ou à moduler les flux au sein de l'écosystème (adapté de Ellison, 2019). Une fois que ces espèces ont fondé un écosystème, d'autres espèces se développent et les espèces fondatrices voient leur prédominance déclinées. Cela étant, leur rôle dans l'écosystème reste tout aussi important puisqu'elles assurent son équilibre. Elles deviennent alors des espèces ingénieurs.

J'illustrerai l'évolution historique dans la figure ci-dessous :

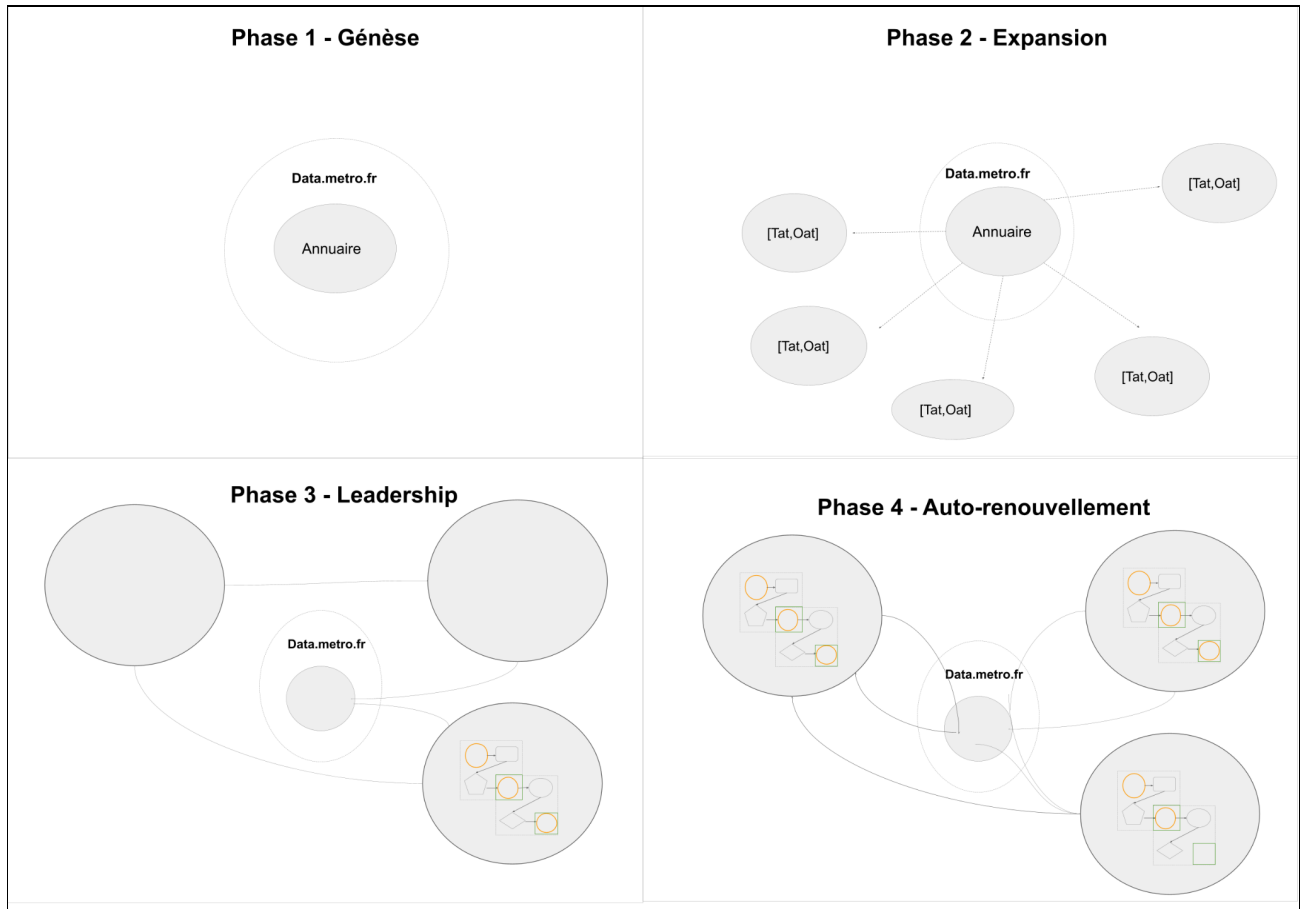


Figure 3.5 : Évolution de l'écosystème open data

3.2 Choix des observations à effectuer et réalisation

Résumé

Dans la partie précédente, j'ai posé le fait qu'un modèle préliminaire d'usage de la donnée était constitué de neuf éléments : une plateforme *open data* (1) et ses parties prenantes (2), des données ouvertes prenant la place d'espèce ingénieur (3) dans un écosystème *open data* (4), des capacités d'action (5) d'artefacts technologiques et sociaux (6) agissant l'état de jeux de données (7), celles-ci étant articulées en séquences d'articulation (8) produisant des complémentarités (9). Ces concepts m'ont permis de créer une grille d'observation, principal instrument de collecte de données d'études de quatre cas exploratoires. Dans cette partie, je décrirai le champ empirique constitué de quatre cas d'usage des données ouvertes. Pour chaque cas, je présenterai d'abord le contexte dans lequel il a émergé, son évolution dans le temps, et la relation entre les différents acteurs et technologies utilisées d'un point de vue juridique et opérationnel (3.2.1, 3.2.2, 3.2.3, 3.2.4) Enfin, je réaliserai une analyse transversale des cas et plus particulièrement des difficultés rencontrées dans l'usage des données ouvertes (3.2.5).

3.2.1 Présentation et analyse du cas d'usage 1 “Quartier” : Utiliser les données ouvertes *smart living* et *smart mobility* d'une ville pour faciliter les interactions entre des salariés et leur quartier

3.2.1.1 Genèse et évolution du cas d'usage “Quartier”

Le cas d'usage “Quartier” a été initié à la demande d'un grand groupe immobilier dit “Batiments”. Ce dernier souhaitait mettre à disposition de ses clients professionnels (les entreprises locataires des bâtiments) une nouvelle gamme de services digitaux. Cette gamme devait permettre de faciliter les déplacements et d'améliorer le cadre de vie des salariés occupant les bâtiments construits par le groupe. Par exemple, ce groupe imaginait mettre à disposition des occupants une application *smartphone* donnant tous les modes de transports à leur disposition pour rejoindre le bâtiment ainsi que le temps de trajet. Aussi souhaitait-il créer une plateforme de conciergerie mettant en relation les commerçants des alentours

du bâtiment et ses occupants. Cette plateforme aurait par exemple permis aux occupants de pouvoir disposer d'un service de pressing ou d'épicerie fine directement au sein du bâtiment.

Pour parvenir à créer ces services, ce groupe s'est mis en relation avec la ville "Smart" qui accueille ses bâtiments. En effet, cette ville produit et met à disposition en *open data*, sur une plateforme, un grand nombre de jeux de données pouvant l'aider dans sa démarche. Parmi eux se trouve un jeu de données recensant tous les commerces de la ville, leur adresse, leurs horaires d'ouverture et les modalités de livraison ou encore un jeu de données affichant en temps réel la disponibilité de places de parking.

Après avoir établi le schéma directeur de l'infrastructure digitale cible à mettre en oeuvre pour importer et utiliser ces données ouvertes, ce groupe s'est rapproché d'un éditeur de logiciel, "Plateforme", pour pouvoir créer une plateforme de données centralisant les données ouvertes de la ville, les données privées de chaque bâtiment et les données des prestataires du groupe immobilier. Cette plateforme ayant pour vocation de devenir la base de données centrale sur laquelle seront construits les services. En parallèle, le groupe a fait appel à un cabinet de conseil, "Assist", pour être assisté dans l'usage des données et la conception des services souhaités. Ce cabinet de conseil a été choisi pour son réseau qui comprend un grand nombre d'organisations et d'individus habitués à utiliser les données ouvertes. D'ailleurs, pendant la période d'observation, de nombreuses start-up et entreprises dites "Users" ont été sollicitées pour proposer des prototypes de services utilisant les données du groupe et les données ouvertes.

Progressivement, le groupe "Bâtiment", l'éditeur "Plateforme", la ville "Smart", le cabinet "Assist" et les "Users" se sont organisés pour utiliser les données ouvertes de *smart living* et de *smart mobility* et créer des services à partir d'elles. Au cours de ce projet, le cabinet "Assist" a servi de coordinateurs entre ces différents acteurs et que l'éditeur "Plateforme" a mis en place l'interface technique permettant d'échanger et de construire les usages.

3.2.1.2 Plateforme open data

La plateforme *open data* est constituée d'outils de collecte de données brutes et plus particulièrement d'une interface de programmation applicative (API)³

³ Termes originaux : *Application Program Interface (API)*

Définition : **An Application Programming Interface** (API) Une API permet à une application de communiquer avec une autre application, ou un système d'exploitation, une base de données, un réseau, etc. Par exemple, l'API Google Maps permet à une application d'intégrer du contenu tiers, tel que des restaurants superposés sur un Google Carte (Conrad, 2016).

Elle est également constituée d'outils de transfert des données collectées vers d'autres applications. Ces outils de transfert peuvent aussi être dans notre cas des API.

Son rôle dans l'innovation est en premier lieu de centraliser et d'agréger les données qui permettront de créer de nouveaux services. De manière plus annexe, la plateforme *open data* assure le rôle de partage d'informations à la fois avec les membres de l'écosystème mais aussi avec les organisations chargées de concevoir de nouveaux services.

3.2.1.3 Parties prenantes

Les observations participantes réalisées n'ont révélé aucune partie prenante en charge de l'évolution de la plateforme.

3.2.1.4 Écosystème

Acteurs	Système	Etat initial	Etat final
Prestataires de services	Production	Données métier issues de capteurs fermés	Données métier issues de capteurs partagés
Gestionnaire du bâtiment	Transformation	Service conçu	Service utilisé
Mairie	Intermédiation	Données demandées	Données partagées et données ouvertes
Départements du groupe "Bâtiment"	Production	Données métiers dispersées	Donnée métiers centralisées
Cabinet d'avocat	Production	Donnée métiers fermés	Données fermés partageables

Cabinet Assist	Production	Données ouvertes produites manuellement et de manière discontinue	Données ouvertes produites automatiquement et en continu
Start-up IoT	Intermédiation	Données métiers dispersées	Données métiers connectées (Linked Data)
Cabinet U	Transformation	Données ouvertes	Concept de services
Comité RSE	Intermédiation	Données métier dont la qualité n'est pas garantie	Donnée métier dont la qualité est garantie
Cabinet Assist	Production	Données métier dispersées et non-documentées	Données métier centralisées et documentées
Plateforme	Intermédiation	Données métier dispersées non-exploitées	Donnée métier centralisées et exploitées

Table 3.1 : Présentation de l'écosystème du cas "Logement"

3.2.1.5 États des données

Données indépendantes et dispersées / Données indexées et centralisées

Cette distinction met en relief les données prises indépendamment de toute classification et dispersées sur plusieurs systèmes d'information des données rattachées à une ontologie, un référentiel ou un catalogue se trouvant sur un système d'information unique.

Données brutes / Données connectées (Linked Data) (MAJIC, BAN)

Cette distinction met en contraste les données construites à partir d'une source unique (données brutes) de données dites connectées, construite à partir du croisement de plusieurs sources de données.

Données privées / Données partageables et utilisables

Cette distinction permet de mettre en relief les données qui ne sont pas accessibles de fait de clauses juridiques qui en interdisent la diffusion, le partage ou l'usage, de données qui, à l'inverse, peuvent être partagées et utilisées et ce, en vertu de clauses juridiques spécifiques.

Données calculées / Données interprétées

Il s'agit ici de distinguer les données comme les ratios ou les fréquences qui sont susceptibles de donner une information des données qui, assimilées par un artefact technique, social ou d'un agent, déclenche une réaction.

Données statiques / Données dynamiques

Cette distinction met en contraste les données qui représentent des occurrences à un moment donné et celles qui s'actualisent en suivant la production des occurrences. À l'extrême, les données dynamiques peuvent se mettre à jour en temps réel.

Données indexées / Données normalisées

Cette distinction permet de mettre en relief les données faisant partie d'une classification des données qui répondent à un modèle spécifique. Ce modèle permet d'aider l'intégration des données entre elles.

Données demandées / Données partagées et données ouvertes

Cette distinction sépare les données dont une personne connaît l'existence et qu'elle demande à son producteur, des données tels qu'il les obtient de la part de ce même producteur (via un partage ou via l'ouverture des données)

3.2.1.6 Données ouvertes et outils

Les jeux de données ouvertes de la ville "Smart" relatifs aux transports et au cadre de vie alimentent en premier lieu des visualisations. Ceux-ci cartographient l'ensemble des travaux en cours sur la voie publique ou encore l'emplacement des espaces verts.

Toutefois, les données ouvertes n'alimentent pas que des outils informatiques. En effet, j'ai pu observer qu'elles structuraient également des études sur le changement des habitudes de déplacement des citoyens.

3.2.1.7 Artefacts

Nom	Type	Objectif
Comité data	Social	Arbitre et prend les décisions relatives à l'exploitation, la sécurité et la qualité des données
Cellule du comité	Social	Instances en charge de l'exploitation, la sécurité et la qualité des données
Annexe aux contrats de prestation de service	Social	Créer de lisibilité, de la confiance et un équilibre dans l'échange de données entre le groupe "Bâtiment" et ses prestataires de services
Plateforme open data	Technique	Centraliser les APIs qui récupèrent les données sur un système d'exploitation, normaliser les données récupérées
Système d'exploitation de données	Technique	Interfacer les données produites par les capteurs des prestataires de services ; Exploiter les données produites par les prestataires de service
Logiciel de modélisation	Technique	Modéliser des espaces physiques et leurs surfaces de manière dynamique
Tableur collaboratif "Airtable"	Technique aux finalités sociales	Recenser, documenter et échanger autour des données à ouvrir et/ou à utiliser

Table 3.2 : Présentation des artefacts du cas "Logement"

3.2.1.8 Capacité d'action

Artefact	Capacité d'action	Influence sur état
Comité data	Saisir les cellules du comité pour réaliser des exploitations de données pilotes, sécuriser la collecte par des moyens juridiques, renforcer la sécurité et la qualité des données	Données privées → Données partageables ; Données calculées → Données interprétées

Cellule du comité	"	"
Annexe aux contrats	Sécuriser la collecte des données	Données privées → Données partageables
Système d'exploitation des données	Interfacer les données relatives à la Gestion Technique du Bâtiment et fixer des seuils d'alertes pour identifier les incidents	Données brutes → données connectées ; Données calculées → Données interprétées
Plateforme open data	Centraliser les données relatives par bâtiment, normaliser les données, ouvrir les données	Données indépendantes → données annexées
Logiciel de modélisation	Transformer le plan d'un espace physique en un modèle numérique	Données statiques → données dynamiques
Tableur collaboratif "Airtable"	Etablir un recensement des données à collecter par catégorie de producteur	Données indépendantes / Données indexées

Table 3.3 : La capacité d'action des artefacts du cas "Logement"

3.2.1.9 Séquence d'articulation

L'ensemble de ces observations m'ont permis de mettre à jour X séquences d'articulations. Elles représentent et définissent l'usage des données ouvertes du cas "Quartier". Afin de ménager notre lecteur, je m'efforcerai de ne présenter en détail qu'une seule des séquences d'application mise à jour.

Séquence d'articulation 1

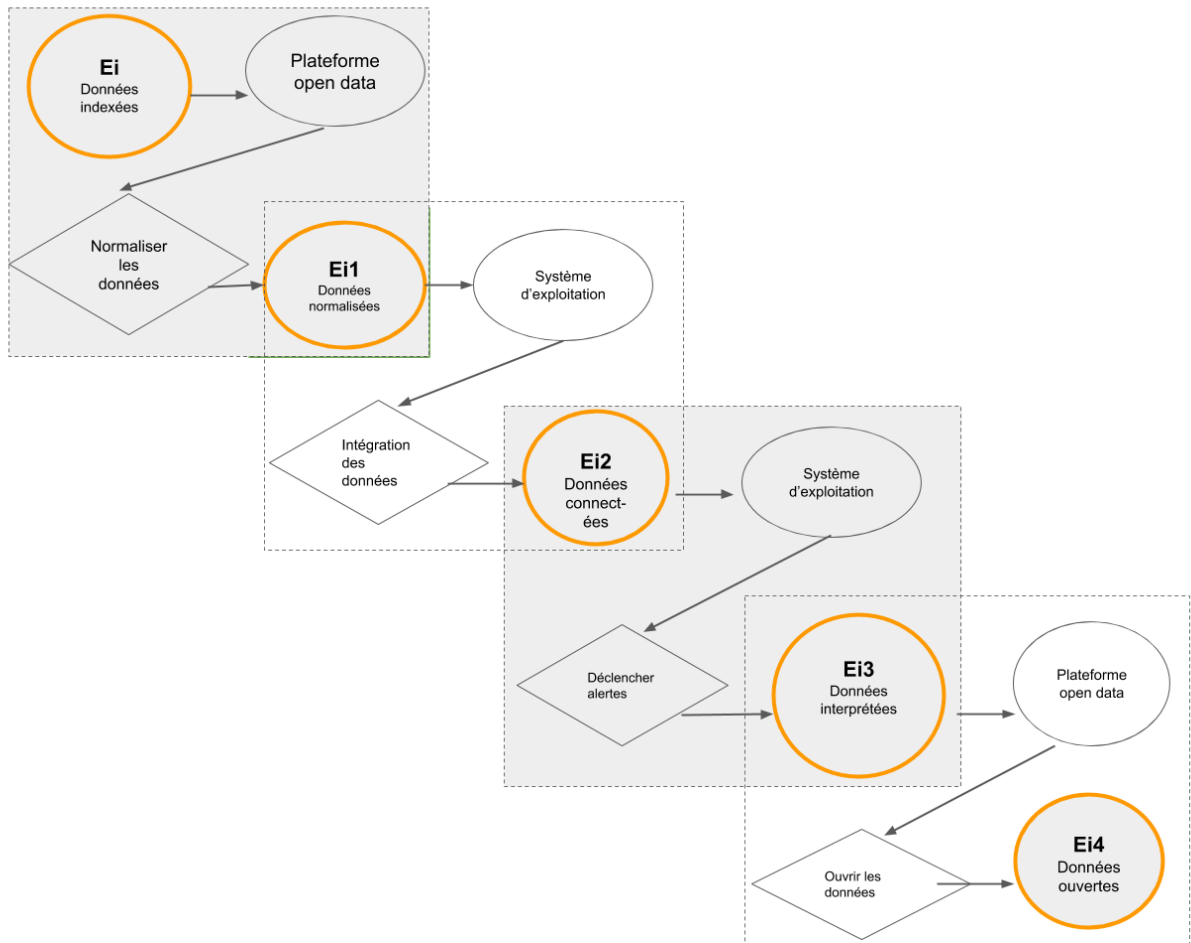


Figure 3.6 : Illustration de la première séquence d'articulation du cas d'usage "Quartier"

Cette première séquence se déroule de la façon suivante :

La première articulation débute avec les données recensées grâce au tableur collaboratif "Airtable", des données à l'état "indexées" (Ei). Elle fait appel à la plateforme *open data* pour pouvoir normaliser les données récoltées. Autrement dit, elle va faire en sorte que les données se conforment à un modèle spécifique, dit modèle normal. Cette normalisation va permettre de faciliter les interactions entre ces données. Ainsi, à l'issue de la première articulation, les données sont passées de l'état "indexées" à l'état "normalisées" (Ei1).

La deuxième articulation implique le système d'exploitation des données. Ce dernier va permettre d'interfacer ou d'intégrer les données recensées avec, par exemple, le plan des bâtiments construits par "Tourgroup". Cette intégration était essentielle pour envisager de créer des services qui facilitent les interactions entre la vie au sein du bâtiment et la vie en dehors du bâtiment. Ainsi, à la fin de la deuxième articulation, les données sont passées de l'état "normalisées" à l'état "connectées" (Ei2).

La troisième articulation fait également appel au système d'exploitation qui va être mis à contribution pour déclencher des alertes dès qu'apparaît une anomalie dans les données. En déclenchant ces alertes, le système d'exploitation permettra ainsi de fiabiliser progressivement les données. Pour pouvoir déclencher ces alertes, le système d'exploitation interprète les données. Dès lors, à la fin de la troisième articulation, les données passent de l'état "connectées" à l'état "interprétées" (Ei3).

Enfin, la quatrième articulation fait appel à nouveau à la plateforme *open data*. Elle va permettre d'ouvrir les données qui, une fois fiabilisées, permettrait aux partenaires de Tourgroup de créer des services à partir d'elles. Ainsi, à la fin de la quatrième articulation, les données passent de l'état "interprétées" à l'état "ouvertes" (Ei4).

Toutefois, cette séquence d'articulation s'est très vite arrêtée. En effet, au cours de notre période d'observation, la source des données alimentant la première articulation s'est très vite tarie. Au début du projet, la séquence était alimentée par les données déjà détenues par "Tourgroup". Ensuite, il a fallu l'alimenter avec des données appartenant aux partenaires et prestataires de services de "Tourgroup". C'est à partir de ce moment que les données ont arrêté d'affluer vers la première articulation.

L'encadré ci-dessous décrit une explication de ce tarissement de données ayant amené à l'arrêt de cette séquence d'articulation.

Exemple de blocage : tarissement de la source de données

(source : extraits des notes d'observation)

Les rencontres avec les partenaires et prestataires, fournisseurs de données révèlent qu'ils expriment tous le besoin de connaître davantage la finalité de leur utilisation afin de transmettre des données plus pertinentes. Aussi s'interrogent-ils sur le cadre juridique et éthique de l'utilisation des données. "TourGroup" s'interroge alors sur la nécessité de créer un comité qui donnera ce type d'informations aux organisations requérantes.

3.2.1.10 Complémentarité

La première articulation s'achève avec des données normalisées. Les données normalisées ont une valeur intrinsèquement liée à leur interopérabilité puisqu'elles réduisent le risque d'erreur dans l'analyse dont elles peuvent faire l'objet. D'une certaine manière, la valeur des données normalisées se révèle à travers la fiabilité de leur usage futur.

Le deuxième articulation s'achève par la connexion des données entre elles.. La valeur ajoutée de données connectées s'incarne dans la création de ponts entre applications et systèmes d'informations. Par exemple, le système d'exploitation a permis de connecter les données de capteurs de taux de COE installées dans les bâtiments avec les données relatives aux équipements installés dans les bâtiments (ventilation, éclairage automatique). Ces données connectées permettront de lier la détection d'un taux élevé de CO2 dans le bâtiment avec l'activation automatique de la ventilation.

La troisième articulation aboutit à l'obtention de données interprétées, prenant la forme d'alertes en cas d'anomalie dans les données. La valeur de ce résultat est évidente dans la mesure où les occupants du bâtiment pourront en faire l'expérience. Par exemple, en cas d'un taux de CO2 trop élevé dans le bâtiment, la ventilation se mettra en marche pour renouveler l'air.

Enfin, la quatrième articulation produit des données ouvertes, dont la valeur se traduit par les potentiels nouveaux services qui pourront être créés grâce à elles. Alors qu'elles auront déjà permis de créer des bâtiments plus agréables, elles pourront également servir à d'autres organisations et individus.

3.2.3 Présentation et analyse du cas d'usage 2 “Transport” - Utiliser les données *smart mobility* pour créer une plateforme de gestion et d'innovation autour des mobilités

3.2.3.1 Génèse et évolution du cas d'usage

Le cas d'usage “Transport” est né suite au lancement d'un projet par la métropole “SmartMetro”. Ce projet consistait à fédérer l'ensemble des outils d'information des services de mobilités de son territoire, et ce sur une seule et unique plateforme. Le but final du projet consiste à créer, grâce à cette plateforme, des services communs à tous les modes de transport (bus, tram, co-voiturage, etc.) et pouvant se déployer sur l'ensemble du territoire.

Pour y parvenir, “SmartMetro” a lancé un partenariat d'innovation avec un groupement d'entreprise mené par un grand groupe dit “BTP”. Pour faciliter ses futures relations avec les acteurs locaux de la donnée, le groupe “BTP” a fait appel à une coopérative “Coop”. “Coop” est spécialisée dans la médiation de données (activité consistant à traduire les enjeux techniques de la donnée auprès de publics non-spécialistes)". Coop a intégré le partenariat d'innovation en qualité de coordinatrice des acteurs locaux de la donnée, tous producteurs de données relative à la *smart mobility*, dont des données ouvertes. Parmi ces données ouvertes, on peut citer les jeux de données contenant l'emplacement des parkings, la disponibilité des parkings ou encore les itinéraires exacts des pistes cyclables.

Dans le but de centraliser toutes les données *smart mobility* ouvertes et privées recensées sur le territoire, “BTP” a également fait appel à un éditeur de logiciel “OpSoft”, spécialisé dans les plateformes de gestion de données ouvertes. Cette plateforme devait jouer 2 rôles : permettre de créer des prototypes de services de mobilité pouvant être testé auprès des futurs utilisateurs (les citoyens de SmartMetro) et alimenter les solutions logicielles qui développeront des services spécifiques sur la base de ces données. “BTP” a fait par exemple appel à “ParkingSoft” pour développer un système de guidage des automobilistes vers les places de parking disponibles sur le territoire de SmartMetro.

Progressivement, la métropole “SmartMetro”, le groupe “BTP”, l'éditeur “OpSoft”, la coopérative “Coop”, et l'éditeur “ParkingSoft” se sont organisés selon le modèle du partenariat d'innovation et ont mis à jour deux services. Le premier est une application mobile permettant d'obtenir tous les titres de transports possibles sur l'ensemble du territoire. Le deuxième est un système de guidage des automobilistes vers les places disponibles de la métropole.

3.2.3.2 Plateforme open data

La plateforme est dans ce cas constituée d'un hub de données.

Définition : Hub de données : « système permettant d'améliorer la cohérence des données qui y transitent offrant une vue 360° fiable et compréhensible. (Il s'agit d'un) endroit unique, où se retrouvent l'ensemble des données d'un domaine, sans discrimination aucune. Elles y sont intégrées quelles que soient leurs sources, structures, formats, granularité, niveaux d'agrégations, type de contenus, et surtout quelle que soient leurs destinations. En effet, le mot Hub connote bien l'idée de transit d'un point de collecte vers un ou plusieurs points cibles » (DataGalaxy, 2018)

Elle est également constituée de différents tableaux de bord (dashboard).

Définition : Dashboard : « Les tableaux de bord remplissent généralement trois fonctions importantes pour les utilisateurs: ils surveillent et suivent les indicateurs importants, ils fournissent des analyses pour déterminer les tendances et les conditions d'exception, ils fournissent des informations pour faciliter l'étude et le diagnostic et identifier les actions correctives si nécessaire » - Vora (2009, p.120)

La plateforme dans l'innovation de services a un rôle d'intermédiaire entre les objets et technologies qui récupèrent la donnée, les agents et outils chargés de concevoir de nouveaux services, et les agents en charge de contrôler les effets de la mise en œuvre des nouveaux services.

3.2.3.3 Parties prenantes

Les observations participantes réalisées n'ont révélé aucune partie prenante en charge de l'évolution de la plateforme.

3.2.3.4 Écosystème

Acteurs	Système	Etat initial	Etat final
Réseau de transport	Transformation	Informations intégrées dans un service d'information et de	Service utilisé

		paiement (application)	
Centrale régionale de la billettique transport	Production	Données métiers	Données partagées
Plateforme publique de covoiturage	Production	Données métiers	Données partagées
Smart Metro	Intermédiation	Données métiers dispersées	Donnée métiers centralisées
BTP	Transformation	Donnée partagées	Informations intégrées dans un service d'information et de paiement (application)
BTP	Transformation	Service utilisé	Adaptation des infrastructures au nouveaux usages des citoyens de la métropole
Parking Soft	Transformation	Données ouvertes	Données cartographiées et données intégrées à un service de guidage
Op Soft	Intermédiation	Données métiers disséminées	Données métiers et données ouvertes centralisées
Mairies	Production	Données fermées	Données ouvertes

Table 3.4 : Présentation de l'écosystème du cas "Transport"

3.2.3.5 États des données

Données indépendantes / Données indexées

Cette distinction met en relief les données prises indépendamment de toute classification des données rattachées à une ontologie, un référentiel ou un catalogue.

Données / Informations

J'applique ici la distinction classique d'Ackoff entre les données qui représentent des occurrences sans aucune précision contextuelle et les informations qui peuvent être interprétées en fonction du contexte.

Données statiques / Données dynamiques

La première distinction met en contraste les données qui représentent des occurrences à un moment donné et celles qui s'actualisent en suivant la production des occurrences. À l'extrême, les données dynamiques peuvent se mettre à jour en temps réel. Les données des temps de passage des bus et trams aux arrêts sont un exemple type de données dynamiques en temps réel.

Données découvrables / Données non découvrables

Cette distinction évidente oppose les données dont il est possible de vérifier l'existence, des données pour lesquelles il faut questionner son potentiel producteur.

Données en création

L'état originel d'une donnée représente le moment où une occurrence particulière est capturée puis transcrite dans un système d'information en suivant des procédures particulières.

Besoins des utilisateurs / Informations nécessaires

Cette distinction met en relief les besoins des futurs utilisateurs du services, et les informations qui sont nécessaires pour répondre à ces besoins. Une fois ces informations identifiées, il est possible d'induire les données à exploiter.

Données simples / données calculées

Il s'agit ici de distinguer les données comme ratios ou fréquences basées sur d'autres données, aux données qui représentent une occurrence unique.

3.2.3.6 Données ouvertes et outils

Les outils structurés par les données ouvertes relatives à la mobilité sont les mêmes que ceux qui résultent de l'usage des données relatives à la mobilité. C'est-à-dire l'application mobile permettant d'obtenir des titres de transports et le système de guidage des automobilistes vers les places disponibles dans la métropole.

3.2.3.7 Artefacts

Nom	Type	Objectif
Plateforme de management de la mobilité urbaine	Technique	Fédérer l'ensemble des outils d'information et de distribution des services de mobilités du territoire, sur une seule et unique plateforme.
Dispositif mobilité	Technique aux finalités sociales	Permettre aux citoyens de gérer un compte de mobilité métropolitain, d'accéder à l'information multimodale (données théoriques et dynamiques) et de calculer un itinéraire.
Référentiel de données	Technique	Classer les données et assurer leur interactions avec les systèmes d'information
Coordination des systèmes billettiques	Technique	Centraliser tous les dispositifs de billettique de la région
Solution de mutualisation des parkings	Technique aux finalités sociales	Permettre aux organisations de mutualiser leurs parking
Capteurs sur les places de parking	Technique	Détecter les places disponibles pour en informer les automobilistes
Plateforme de services	Technique aux finalités sociales	Obtenir des informations, réserver et payer des places de parking
Algorithme prédictif	Techniques	Prédire la disponibilité des places de parking et guider les automobilistes vers leur place de parking
Espace de co-création	Social	Offrir un lieu pour co-créer des services à partir des données
Tableur collaboratif "Airtable"	Technique à vocation sociale	Recenser les données ouvertes, non-ouvertes, à transformer, à générer
Plateforme Opsoft	Technique à vocation sociale	Permettre l'intégration des données à des applications tierces

Table 3.5 : Présentation des artefacts du cas d'usage "Transport".

3.2.3.8 Capacité d'action

Artefact	Capacité d'action	Influence sur état
Plateforme de management de la mobilité urbaine	Intégrer les données aux outils d'information ; Fédérer les outils de billettique pour créer un point unique de souscription aux services de mobilité	Données indépendantes → données indexées
Dispositif mobilité	Non-observé	Non-observé
Référentiel de données	Non-observé	Non-observé
Coordination des systèmes billettiques	Non-observé	Non-observé
Solution de mutualisation des parkings	Tester la mutualisation sur les parkings en ouvrage	Données → Informations
Capteurs sur les places de parking	Identifier les places disponibles et leur emplacement	Données en création → Données créées
Plateforme de services	Traduire les données de façon à indiquer à un automobiliste l'emplacement de places disponibles	Données → Information
Algorithme prédictif	Non-observé	Non-observé
Espace de co-création	Identifier les besoins des citoyens relatifs à leur mobilité	Besoins → Informations
Tableur collaboratif "Airtable"	Technique à vocation sociale	Recenser les données ouvertes, non-ouvertes, à transformer, à générer

Plateforme Opsoft	Permettre l'intégration des données à des applications tierces	Données → Information
-------------------	--	-----------------------

Table 3.6 : Les capacités d'action des artefacts du cas d'usage "Transport"

3.2.3.9 Séquence d'articulation

Ces observations ont permis de mettre à jour deux séquences d'articulations. Elles représentent et définissent l'usage des données ouvertes du cas "Transport". Afin de ménager notre lecteur, je m'efforcerai de ne présenter en détail qu'une seule des séquences d'application mise à jour (Figure 3.7).

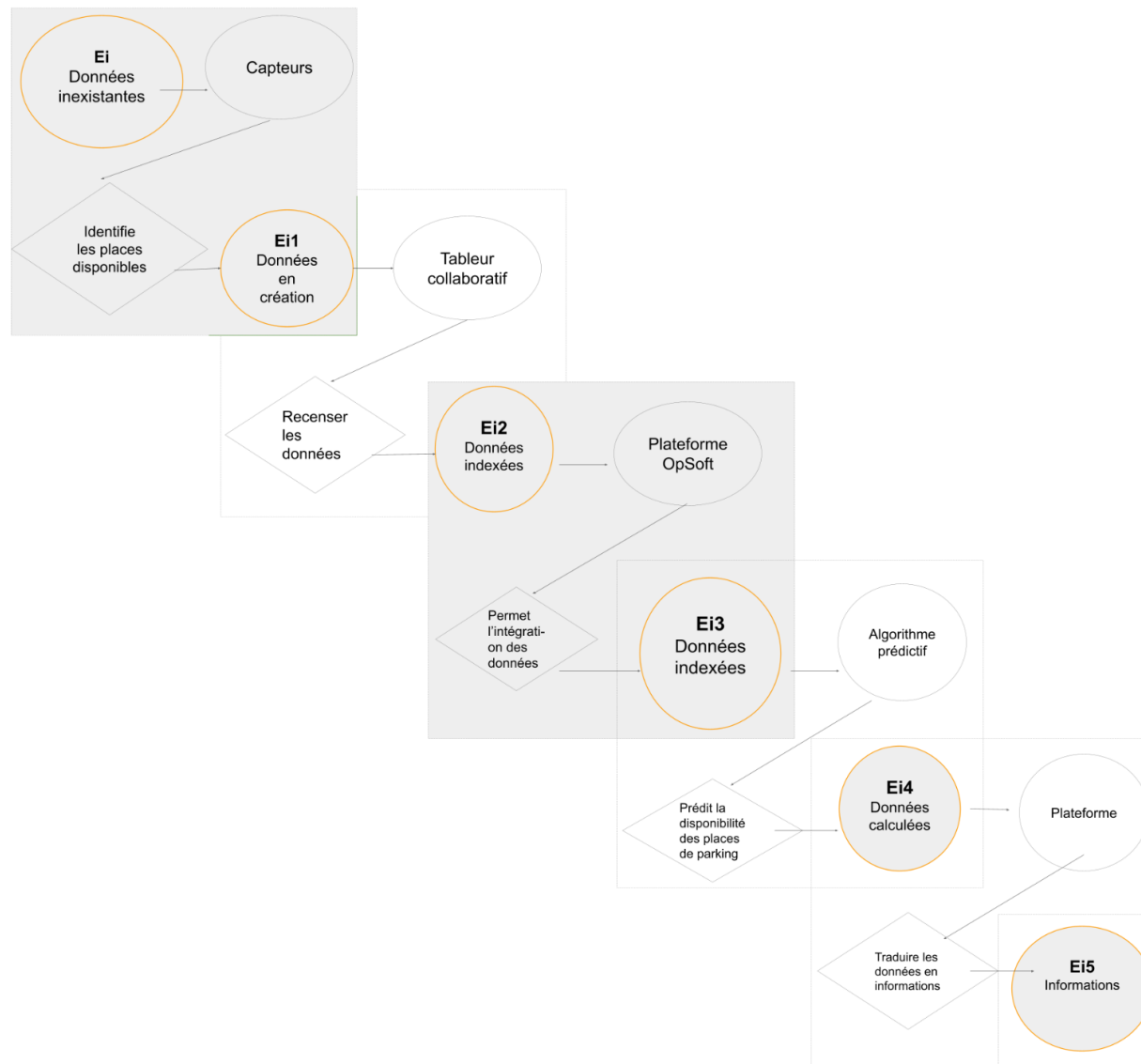


Figure 3.7 : La représentation de la première séquence d'articulation du cas d'usage "Transport"

Cette première séquence se déroule de la façon suivante :

La métropole “SmartMetro” souhaitait mettre à disposition des automobilistes une application pour faciliter la recherche de places de parking. Elle avait besoin pour cela de données qui n'existaient pas (Ei). Dès lors, elle a sollicité ParkingSoft qui a installé des capteurs dans trois parkings pilotes. Ces capteurs collectent des données en temps réel sur la disponibilité de chaque place de parking. À l'issue de cette première articulation, les données sont passées de l'état “inexistantes”(Ei) à l'état “en création” (Ei1).

La deuxième articulation fait appel à un tableur collaboratif permettant de recenser des données. Dans cette séquence, il est utilisé pour recenser les données qui pourraient venir compléter les données issues des capteurs. Ce tableur a par exemple permis de recenser les données relatives au volume du trafic. Ces données, une fois croisées avec celles des capteurs, pourraient par exemple permettre de créer un service qui redirige les automobilistes coincés dans une circulation dense vers les places de parking disponibles autour d'eux. À la fin de cette articulation, les données sont passées de l'état “en création” (Ei1) à l'état “indexées” (Ei2).

La troisième articulation sollicite la plateforme OpSoft, qui, outre le fait de centraliser les données, permet d'intégrer les données à des applications tierces. En ce sens, elle permet à des données indexées d'être exploitables par des applications tierces. Par conséquent, à la fin de la troisième articulation, les données passent de l'état “indexées” (Ei2) à un autre état “indexées” (Ei3).

La quatrième articulation implique l'algorithme prédictif de la société ParkingSoft. Celui-ci permet de prédire les places de parking disponibles mais aussi d'orienter les automobilistes vers celles-ci. Pour cela, il exploite les données. Ainsi, cette quatrième articulation permet aux données de passer de l'état “indexées” (Ei3) à l'état “calculées” (Ei4).

Enfin, la cinquième articulation fait appel à la plateforme de services souhaitée par la métropole SmartMetro et livrée par le groupe “BTP”. Celle-ci permet de traduire les données calculées par l'algorithme en informations qui vont effectivement permettre à l'automobiliste de se rendre vers les places disponibles. Ainsi, à la fin de la cinquième articulation, les données sont passées de l'état de données calculées (Ei4) à l'état d'information (Ei5).

Malgré le fait que cette séquence d'articulation soit arrivée à son terme et ait permis la création d'un service de guidage opérationnel, celle-ci n'a pas exploité tout son potentiel. En effet, il était question de créer une multitude d'autres services sur la base de croisements de données de parking avec d'autres

données relatives à la mobilité. Or, les acteurs du projet ont rencontré des difficultés dans la troisième séquence. En l'occurrence, ils n'ont pas pu récupérer les données qu'ils avaient recensées.

L'encadré ci-dessous décrit une explication de cette difficulté à récupérer les données.

Exemple de blocage : le producteur de données ne donne pas suite

(source : extraits des notes d'observation)

“Les acteurs sont toujours dans la phase où ils cherchent beaucoup de données. Pour les données de covoiturage sur la métropole, cela n'a pas marché (au moins à présent) et pour le trafic ils étaient obligés de payer (ce qui aurait réduit leurs budgets...). Les raisons de cet arrêt sont difficiles à saisir. Concernant le covoiturage, la startup contactée a apparemment un autre projet en cours et n'a pas le temps. Pour les données parking, les capteurs sont un succès, mais pas pour faire de l'ouverture de donnée, plutôt pour faire des visualisations intelligentes. À mon avis, il est difficile d'obtenir les données car les entreprises ont peur d'ouvrir leurs données. Dans la mesure où elles ont une valeur commerciale et, surtout, dans la mesure où elles ont un coût pour eux (mise en forme, anonymisation...), elles veulent une contribution financière”.

3.2.3.10 Complémentarité

Le résultat de la première articulation⁴ se traduit par un nouvel actif que les acteurs du cas vont pouvoir exploiter. La valeur de cette première articulation sera révélée par les prochaines articulations.

La deuxième articulation qui se conclut par des données indexées dans le tableur collaboratif, produit un résultat inattendu : les acteurs du projet vont pouvoir identifier les synergies entre fournisseurs de données qu'ils seraient intéressant de mettre en place dans le cadre du projet. Par exemple, le tableur collaboratif a indexé les données relatives au trafic routier, produites par l'entreprise en charge de la maintenance des routes et les données relatives à l'occupation des places de parking produites par Parking Soft. Ce tableur a donnée l'idée aux acteurs de créer une synergie entre l'entreprise et Parking Soft qui serait

⁴ (i.e de nouvelles données en création)

mutuellement bénéfique : d'un côté Parking Soft pourrait guider les automobilistes vers leurs places en fonction du trafic ; de l'autre, l'entreprise pourrait anticiper les aménagements des routes qu'il faudrait envisager, forte de la connaissance que Parking Soft lui apporterait sur les trajectoires des automobilistes sur le point de se garer.

La troisième articulation aboutit à des données indexées sur la plateforme OpSoft mais aussi partagées avec tous les acteurs du projet. Cette nouvelle indexation va jouer le rôle de référentiel commun à tous les acteurs du projet. La valeur de ce référentiel réside dans le fait qu'il va permettre aux acteurs hétérogènes de s'entendre sur une compréhension commune de la disponibilité des parkings et des actions à entreprendre.

Le résultat de la quatrième articulation (i.e prédictions sur la disponibilité des places de parking) se traduit par les nouvelles possibilités d'anticipation offertes aux acteurs. Dans la mesure où ils sont capables de prédire le taux de remplissage des parkings, ils sont également en capacité de prévenir leur saturation et les difficultés de circulation liées.

Enfin, la cinquième articulation se termine sur les informations données aux automobilistes quant au trajet à suivre pour se garer au plus vite. Sa valeur est évidente. Pour les automobilistes, elle se traduit par une diminution du temps d'attente pour trouver une place de parking ; pour la collectivité par une meilleure gestion du trafic routier.

3.2.4 Présentation et analyse du cas d'usage 3 “Logement” - Utiliser les données *smart economy* pour maximiser l'activité commerciale d'une zone touristique

3.2.4.1 Genèse et évolution du cas d'usage

Le cas d'usage “Logement” est né d'un projet d'un grand groupe dit “TourGroup”, gestionnaire d'infrastructures touristiques. Ce projet visait à créer un service d'accompagnement de propriétaires fonciers dans la location de leurs biens immobiliers situés dans les zones touristiques où “TourGroup” est implanté. “TourGroup” espérait développer ce service en exploitant les informations sur l'occupation des logements contenues dans les données ouvertes et les données fiscales. Par ce projet, “Tourgroup” espérait augmenter le nombre de personnes présentes dans certaines zones et par là, son activité commerciale.

Pour y parvenir, “Tourgroup” s’est mis en relation avec la commune dite “ComFisc” pour explorer les données fiscales auxquelles cette dernière a accès et qui donne des informations précieuses sur le taux d’occupation des logements. En effet, par l’intermédiaire des registres des taxes de séjour notamment, les communes et l’administration fiscale ont des informations très précises sur les logements occupés et inoccupés tout au long de l’année. En utilisant ces données, “TourGroup” pensait pouvoir cibler les propriétaires des logements sous-occupés, les accompagner pour louer plus régulièrement leur logement, et indirectement, maximiser son activité commerciale. Autre piste : “TourGroup” envisageait d’utiliser toutes les données ouvertes sur les logements touristiques du territoire de “ComFisc” et, d’en induire des informations sur le taux d’occupation de tel ou tel parc de logement. Parmi elles, les données des enquêtes de fréquentation des zones touristiques réalisées par l’INSEE, librement disponibles.

Pour centraliser ces données ouvertes et, éventuellement les données fiscales auxquelles “ComFisc” a accès, “TourGroup” a souhaité lancer une expérimentation avec l’éditeur de logiciel “InfraData”, fournissant une infrastructure logicielle pour stocker des données ouvertes et confidentielles. Aussi, “TourGroup” a fait appel à une entreprise de conseil “DataAdvisor” pour l’accompagner dans le pilotage du projet et plus particulièrement, dans les interactions avec “ComFisc” et dans l’utilisation de données sensibles comme les données fiscales.

Au fur et à mesure du projet, “ComFisc”, “TourGroup”, “InfraData” et “DataAdvisor” ont exploré plusieurs pistes d’utilisation des données *smart economy* et avec elles, plusieurs modes d’organisation pour développer et pérenniser ce service d’accompagnement de propriétaires de logement.

3.2.4.2 Plateforme open data

La plateforme en question est constituée uniquement d’un catalogue de données structurées et consolidées au préalable.

Son rôle dans l’innovation est de constituer un socle technique sur lequel seront basés tous les artefacts technologiques chargés d’exploiter les données.

3.2.4.3 Parties prenantes

Les observations participantes réalisées n’ont révélé aucune partie prenante en charge de l’évolution de la plateforme.

3.2.4.4 Écosystème

Acteurs	Système	Etat initial	Etat final
Collectivités territoriales	Production	Données fermées	Données partagées
Tourgroup	Production	Données partagées	Service commercial à titre gracieux (cadeaux)
Agence du développement touristique de la France	Production	Données fermées	Données partagées
DataAdvisor	Intermédiation	Données partagées	Donnée ouvertes
Institut National de l'Information Géographique et Forestière	Intermédiation	Donnée ouvertes non-exploitable (complexité)	Données ouvertes exploitables
Union européenne	Intermédiation	Données métiers non-anonymisées	Données métiers anonymisées

Table 3.7 : L'écosystème du cas d'usage "Logement"

3.2.4.5 États des données

Données simples / données calculées

Il s'agit ici de distinguer les données, comme ratios ou fréquences basées sur d'autres données, et les données qui représentent une occurrence unique.

Données non-certifiées / données certifiées

Cette distinction met en relief les données dont la qualité et la fiabilité sont certifiées par d'autres données. En l'espèce, il s'agit des données comportant un indice de fiabilité.

Données brutes / données contextualisée (taxes de séjour)

Cette distinction oppose les données produites sans qu'aucune utilisation future n'ait été imaginée et des données produites dans un but particulier.

Données découvrables / Données non-découvrables

Cette distinction évidente oppose les données dont on peut vérifier l'existence des données pour lesquelles il faut questionner son potentiel producteur. En l'espèce, "DataAdvisor" a dû constituer une liste de données existantes et elle a pu pour cela s'appuyer sur les bases de données et informations disponibles en ligne (données découvrables) mais a dû interroger les producteurs éventuels de données (données non-découvrables) sur lesquelles aucune information n'était disponible.

Données brutes / Données connectées (Linked Data) (MAJIC, BAN)

Cette distinction met en contraste les données construites à partir d'une source unique (données brutes) de données dites connectées, car construite à partir du croisement de plusieurs sources de données.

Données privées / Données partageables et utilisables

Cette distinction permet de mettre en relief les données qui ne sont pas accessibles de fait de clauses juridiques qui en interdisent la diffusion, le partage ou l'usage, de données qui, à l'inverse, peuvent être partagées et utilisées et ce, en vertu de clauses juridiques spécifiques.

3.2.4.6 Données ouvertes et outils

J'ai pu constater que les données ouvertes sur les logements touristiques structurent un outil au nom de Module d'Évaluation du Poids Touristique des Territoires (METT) et qui produit, par territoire, quatre indicateurs relatifs à la capacité d'accueil de touristes, le volume de fréquentation du territoire, la dépense moyenne des touristes et les effectifs salariés.

3.2.4.7 Artefacts

Nom	Type	Objectif
Tableur collaboratif "Airtable"	Technique aux finalités sociales	Recenser, documenter et échanger autour des données à ouvrir et/ou à utiliser

Argumentaire de lobbying en faveur de l'ouverture des données	Social	Influencer les décideurs politiques en faveur de l'ouverture des données
Portail gouvernemental "demarches-simplifiées.fr"	Technique	Demander à l'administration des données
Procédure de choix de la structure juridique portant l'utilisation des données	Social	Guider les producteurs et utilisateurs des données dans le choix d'une structure juridique adaptée aux usages envisagés de la donnée
Convention producteur-intermédiaire-utilisateur	Social	Fixer un cadre juridique de collaboration entre les trois acteurs

Table 3.8 : Les artefacts du cas d'usage "Logement"

3.2.4.8 Capacité d'action

Artefacts	Capacité d'action	Influence sur état
Tableur collaboratif "Airtable"	Recenser les données ouvertes, non-ouvertes, à transformer, à générer	Données non-découvrables → Données découvrables
Argumentaire de lobbying	Convaincre les décideurs d'ouvrir leurs données	Données non-ouvrables → données ouvrables
Portail gouvernemental "demarches-simplifiées.fr"	Non-observé	Non-observé
Procédure de choix de la structure juridique portant l'utilisation des données	Inspirer la rédaction de la convention producteur-intermédiaire-utilisateur	Non-observé
Convention producteur-intermédiaire-utilisateur	Non-observé	Non-observé

--	--	--

Table 3.9 : Les capacités d'action des artefacts du cas d'usage "Logement"

3.2.4.9 Séquence d'articulation

L'ensemble de ses observations m'ont permis de mettre à jour deux séquences d'articulations, qui représentent et définissent l'usage des données ouvertes du cas "Logement". Afin de ménager le lecteur, je m'efforcerai de ne présenter en détail qu'une seule des séquences d'articulation mises à jour. Les schémas descriptifs de l'autre séquence se trouvent en Annexe.

Séquence d'articulation 1

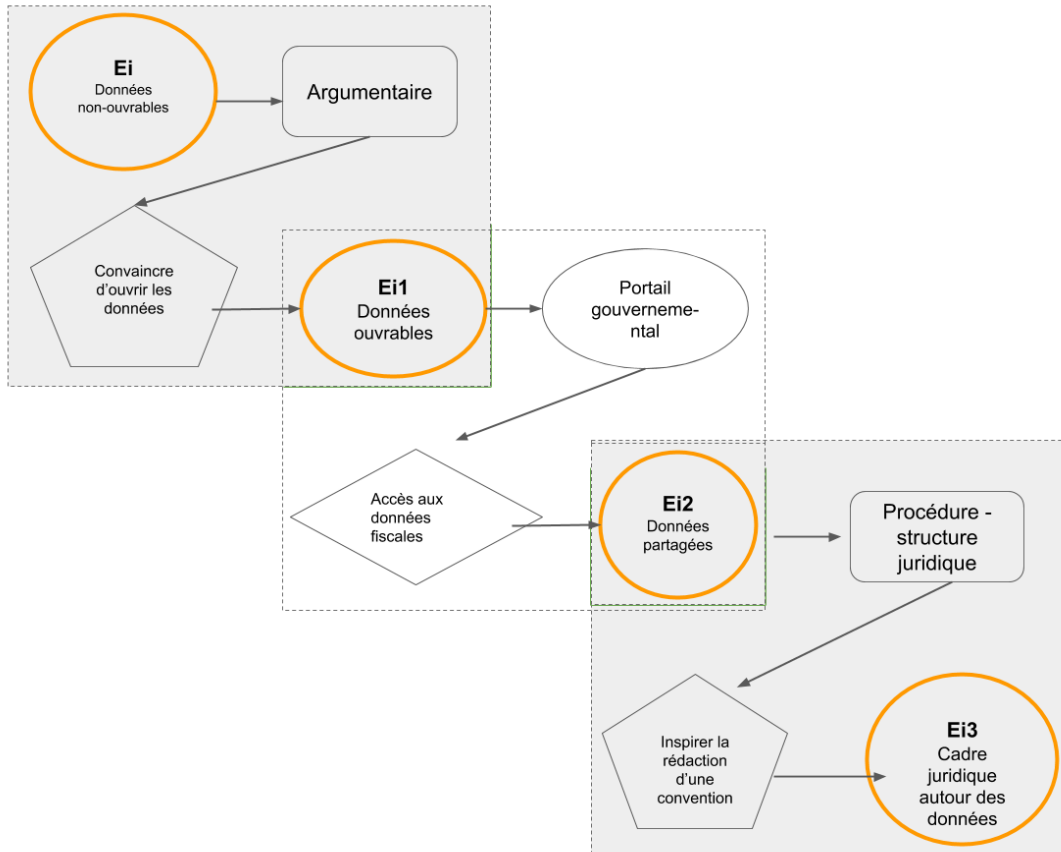


Figure 3.8 : représentation de la première séquence d'articulation

Cette première séquence se déroule de la façon suivante :

Les données fiscales dont a besoin “Tourgroup” pour créer son service commercial ne sont pas ouvrables. La raison en est que la commune qui peut avoir accès aux données n’est engagée dans aucune démarche d’ouverture des données. Pour convaincre la commune d’ouvrir ces données, “Tourgroup” fait rédiger par

le cabinet "DataAdvisor" un argumentaire de lobbying. Ainsi à la fin de cette première articulation les données sont passées d'un état non-ouvrable (Ei) à un état ouvrable (Ei1).

La seconde articulation implique le portail gouvernemental sur lequel la commune peut avoir accès aux données. Elle effectue cette demande et obtient de manière discrétionnaire les données. À la fin de cette deuxième articulation, les données sont passées d'un état ouvrable (Ei1) à un état partagé (Ei2).

Enfin, la troisième articulation fait appel à la procédure de choix d'une structure juridique devant accueillir l'utilisation des données fiscales. En prenant connaissance de cette procédure, la commune "ComFisc" et le groupe "Tourgroup" se mettent d'accord pour construire et signer une convention juridique qui encadrera les futurs utilisations des données, dont le service commercial souhaité par "Tourgroup". Ainsi, à l'issue de la troisième articulation, les données sont passées d'un état partagé (Ei2) à un état utilisable (Ei3).

Néanmoins, cette séquence d'articulation s'est arrêtée très vite. En effet, le projet s'est arrêté soudainement au début de la deuxième séquence. Bien que la commune "ComFisc" ait été convaincue par l'argumentaire de la pertinence de l'ouverture des données, elle n'a pas donné suite et n'a jamais fait les démarches de demande des données sur la portail gouvernemental. De ce fait, les données n'ont jamais été partagées et aucune convention n'a été rédigée. La séquence est restée à l'état d'idée mais n'a jamais vu le jour.

L'encadré ci-dessous décrit les raisons possibles de l'arrêt du projet à ce stade de la séquence d'articulation.

Exemple de blocage : le producteur de données ne donne pas suite

(source : extraits des notes d'observation)

"Les représentants de TourGroup supputent que le maire n'a pas compris réellement ce que nous pourrions faire avec ces données. Leur hypothèse est que le travail de demande de la donnée est perçu au sein de la mairie comme un boulot "en plus" et "lointain" de leurs enjeux à court terme (assurer l'ouverture des écoles, la sécurité des lieux touristiques, etc.)".

Aussi, j'ai observé des blocages dans la seconde séquence identifiée. Cette séquence s'est elle aussi bloquée dès les premières articulations. Or, le blocage est cette fois-ci d'ordre technique. En l'occurrence,

les données à utiliser contenaient un trop grand nombre de valeurs fausses pour pouvoir s'y fier et baser dessus le comptage des logements vides.

3.2.4.10 Complémentarité

La première articulation se traduit par le fait que les données fiscales sont désormais ouvrables, la commune étant convaincue de l'intérêt de les ouvrir. En tant que tel, ces données ouvrables n'ont aucune valeur dont les acteurs peuvent faire l'expérience. Ce sont les prochaines articulations qui donneront les premiers résultats avec une valeur directe.

Le résultat de la deuxième articulation (i.e données partagée) se traduit par le fait que "Tourgroup" va pouvoir évaluer la compatibilité réelle de ces données avec la création de son service commercial. Bien qu'il ne pourra pas encore exploiter les données, le fait qu'elles soient déjà partagées lui permettra de faire une analyse descriptive de leur contenu (granularité, dates de mises à jour, etc.). Alors, il pourra évaluer dans quelle mesure ces données lui permettront de mettre en place le service commercial voulu.

Enfin, la troisième articulation, qui se conclut par des données utilisables, produit un résultat évident : Tourgroup et Comfisc bénéficient d'un nouvel actif qui va leur permettre à tous les deux d'améliorer leur offre de service. Tourgroup va pouvoir créer son service commercial tandis que Comfisc pourra également optimiser sa collecte d'impôt. En effet, ces données fiscales lui donneront certainement des indications sur des anomalies quant à la collecte de taxes liées à l'occupation de logements. Des anomalies qu'elle pourra rectifier à l'avenir.

3.2.5 Présentation et analyse du cas d'usage 4 "Collaboration" - Utiliser les données de *smart governance* pour créer un service public de la donnée

3.2.5.1 Genèse et évolution du cas d'usage

Le cas d'usage "Collaboration" est né d'une délibération de la métropole "SoliData" actant la mise en œuvre d'un projet de décloisonnement des systèmes d'informations des communes de son territoire par l'échange et l'usage croisé des données ouvertes. À travers ce projet, la métropole "SoliData" entendait mettre en place un service de mise à disposition des données publiques dit "service métropolitain de la donnée" à destination des communes et collectivités du territoire puis progressivement à destination des associations, ou des communautés d'utilisateurs.

Pour aboutir à la mise en place de ce service, "SoliData" a mis en place un groupe de travail rassemblant toutes les communes de son territoire et partagé un jeu de données dit "Wanted Data List" listant toutes les données devant être produites, partagées et ouvertes. L'utilisation de ce jeu de données devait permettre de créer un tableau bord permettant de suivre et de piloter la construction du service public de la donnée. "SoliData" a mutualisé ce groupe de travail avec le syndicat de transport de son territoire "Syndic" et la ville principale de son territoire "VilleD" de façon à pouvoir offrir un service de la donnée unique à toutes les collectivités intéressées. En parallèle, "SoliData" a conçu une plateforme dite "Data.Metro" permettant de stocker et de mettre à disposition les données.

Au fur et à mesure du projet, "SoliData", "Syndic" et "VilleD" se sont organisés autour de "Wanted Data List" pour recenser, partager et ouvrir des données à destination des collectivités dans un premier temps puis des communautés d'utilisateurs dans un second temps.

3.2.5.2 Plateforme open data

La plateforme *open data* dans ce cas est constituée d'un catalogue de données.

Définition : Catalogue de données : «Un catalogue de données maintient un inventaire de données grâce à la découverte, la description et l'organisation des ensembles de données. Le catalogue fournit un contexte permettant aux analystes de données, aux scientifiques des données, aux gestionnaires de données et à d'autres consommateurs de données de trouver et de comprendre un ensemble de données pertinent dans le but d'extraire de la valeur commerciale. » - Gartner (2017)

Il est également constitué d'un logiciel de gestion de contenu ou CMS.

Définition : Logiciel de gestion de contenu : « Les systèmes de gestion de contenu (CMS) modernes, dont WordPress est un exemple bien connu, utilisent des bases de données pour gérer des informations. Ce qui peut sembler être un simple détail technique a eu des conséquences si importantes qu'il a marqué un avant et un après dans la nature des sites Web. Dans un CMS, les informations sont stockées dans une base de données, ce qui signifie que les pages sont créées de manière dynamique chaque fois qu'une demande d'informations est adressée au serveur, plutôt que d'être stockées de manière statique comme dans les systèmes de publication Web d'origine. ». Montoro et Codina (2017)

En backoffice, on trouve un serveur alimenté par des scripts Extraction, Transformation et chargement⁵ (ETL) qui récupèrent les données dans les applications métiers.

Définition : ETL : « ETL pour extraction, transformation et chargement (loading) et un ensemble de trois bases de données qui sont consolidées dans un système pour extraire les données d'une base de données et les repérer dans une autre base de données. » - Tiwari (2016)

En termes d'éléments intangibles, la plateforme *open data* évolue en coordination avec un acteur national clé dans le domaine de l'*open data* : Open Data France. Par exemple, les thématiques du catalogue de données sont celles choisies au niveau d'Open Data France pour uniformiser au maximum l'organisation des plateformes à l'échelle du pays. On note également que l'évolution des éléments qui constituent la plateforme est fortement inspirée des évolutions observées sur les autres plateformes au niveau national.

Le rôle de la plateforme dans l'innovation de services est double : valoriser les données publiées et animer les interactions avec les acteurs concernés par les données ouvertes. Le rôle de la valorisation est notamment assuré par les outils de visualisation dynamique de données intégrés à la plateforme, celui d'animation par le CMS qui permet de transmettre des informations sur les événements autour des données ouvertes mais aussi qui permet à tout un chacun de demander des données à "Solidata".

3.2.5.3 Parties prenantes

Les parties prenantes impliquées dans l'évolution de la plateforme sont ici la ville principale du territoire de la métropole, le syndicat de transport de cette même ville ainsi qu'un éditeur informatique mandaté par

⁵ Termes originaux : *Extract-transform-load (ETL)*

la métropole. Il convient de préciser que les agents sollicités sont tous des développeurs informatiques. Par conséquent, leur rôle est strictement cantonné au développement des modules technologiques de la plateforme.

3.2.5.4 Ecosystème

La figure ci-dessous résume les acteurs qui constituent les systèmes de production, d'intermédiation et de transformation ainsi que les états d'entrée et de sortie des données qui passent par eux.

Acteurs	Système	Etat initial	Etat final
Premier ministre	Production	informations/connaissances fermées	informations/connaissances partagées
Prestataire (éditeur de logiciel)	Intermédiation/Transformation	données ouvertes	Informations intégrée dans un service de renseignement (recherche d'itinéraires de montagne) matérialisé par une tablette
Metro "Solidata"	intermédiation	données non-documentée ouverte	donnée ouverte documentée
Communes du territoire	production	donnée métier	-> donnée ouvertes
Tiers-lieux/Coopérative	Transformation	Donnée brute	Donnée recontextualisée
Service du Système d'Information Territorial (SIT)	Production	Données métiers disséminées	Données métiers regroupées
Etalab et Openlaw	Intermédiation(license)	Données ouvertes (open data)	Données partagées
Metro "Solidata"	Production	Données fermées	données ouvertes
Service géomatique	intermédiation	données ouvertes sous un format unique	Données ouvertes sous plusieurs formats
Etalab	Intermédiation	données ouvertes sur une plateforme locale	données ouvertes sur une plateforme nationale
SIT	intermédiation	données métier regroupées	données ouvertes

Table 3.10 : L'écosystème du cas d'usage "Collaboration"

3.2.5.5 États des données

Données statiques / Données dynamiques

La première distinction met en contraste les données qui représentent des occurrences à un moment donné et celles qui s'actualisent en suivant la production des occurrences. A l'extrême, les données dynamiques peuvent se mettre à jour en temps réel. Les données des temps de passage des bus et trams aux arrêts sont un exemple type de données dynamiques en temps réel.

Données consultables / Données réutilisables (opérationnel)

La deuxième distinction observée sépare les données produites et ouvertes dans le seul but d'être consultée et de faire acte de transparence des données qu'il est possible de réutiliser à d'autres fins. Les données sur les délibérations des conseils municipaux sont un exemple de données consultables quand les données des transports sont un exemple type de données réutilisables.

Données découvrables / Données non-découvrables

Cette distinction évidente oppose les données dont il est possible de vérifier l'existence des données pour lesquelles il faut questionner son potentiel producteur.

Données à jour / Données non mises à jour

Cette distinction majeure met en contraste les données à jour des dernières occurrences des données dont la dernière mise à jour remonte dans le temps.

Données brutes / Données visualisées

Il s'agit ici de distinguer les données qui prennent une forme brute que l'on peut imaginer comme des listes de valeur nues, des données qui prennent la forme visuelle de carte, de graphique, ou de tableau de bord par exemple.

Données non-documentées / Données documentées

J'observe une distinction entre les données sur lesquelles aucune information n'existe quant à leurs modalités de création, leur mise à jour ou leur contenu, des données auxquelles sont adjoints ce type d'informations aussi appelées métadonnées.

Données sous un format unique / Données sous plusieurs formats

Je distingue ici les données qu'il est possible d'obtenir sous un seul format des données disponibles sous plusieurs formats.

Données en création

L'état originel d'une donnée représente le moment où une occurrence particulière est capturée puis transcrite dans un système d'information en suivant des procédures particulières.

Données métiers / Données manuelles

Enfin, cette distinction met en relief les données qui sont produites par des logiciels (métiers), des données qui sont renseignées manuellement par des agents.

3.2.5.6 Données ouvertes et outils

Parmi les outils informatiques structurés par les données ouvertes, on trouve une application utilisée par l'observatoire fiscal de la métropole. Elle permet de faire des simulations sur le potentiel fiscal des communes du territoire de la métropole "Solidata". Cette application se base sur l'exploitation de données ouvertes comme le budget des communes mis en ligne sur la plateforme *open data* data.gouv.fr

Les données ouvertes structurent également une application créée par une entreprise tierce. Cette application permet de dresser le portrait socio-économique de territoires en présentant des indicateurs (revenu moyen, accès aux services publics, etc.). Elle se base sur l'exploitation des données présentes sur la plateforme *open data* de la métropole.

Enfin, les données ouvertes alimentent et structurent un concours de création d'application annuel. Cet évènement rassemble des citoyens, développeurs et designers. Leur objectif est de créer des applications pour améliorer la vie des citoyens de la métropole. Les participants ont pour consigne d'utiliser les données ouvertes et ont un accès direct au groupe de travail pour demander des données complémentaires.

3.2.5.7 Artefacts

Nom	Type	Objectif
-----	------	----------

Comité de pilotage	Social	Définir les objectifs du projet et le calendrier d'ouverture des données publiques ; valider le type de données à mettre à disposition ; déterminer les modes de concertation avec les habitants et usagers du territoire ; fixer les moyens à mettre en oeuvre pour animer la plateforme open data
Réunion des Directeurs des Systèmes d'Information du territoire	Social	Informier et sensibiliser à l'ouverture des données
Ateliers Open Data France	Social	Former aux outils créés par OpenDataFrance pour faciliter l'ouverture des données
Laboratoire d'innovation	Social	Faire émerger des applications utilisant les données ouvertes
Prototypes d'application	Technique	Transformer les données ouvertes en informations utiles au citoyen
Métrosphère	Technique aux finalités sociales	Mettre à disposition des agents de la métropole un espace pour poser leurs questions sur l'open data
Plan de charge des données à libérer	Technique	Planifier le travail de production des données
Convention juridique	Social	Acter un engagement mutuel des métropoles et des communes dans la production de données ouvertes
Serious game "Explorateur des données territoriales"	Social	Simuler le travail de production et d'intermédiation des données ouvertes pour former les agents
Script ETL Talend	Technique	Extraire régulièrement et automatiquement les données des applications métiers vers le serveur qui alimente la plateforme open data
Réseau social "teamopendata.org"	Technique aux finalités sociales	Échanger et obtenir des réponses de la part d'une communauté de spécialistes de l'open data
Outil de validation de données Validata	Technique	Valider la qualité de futures données ouvertes
Comité de Pilotage "Données" des communes	Social	Valider l'ouverture des données sur la plateforme de la métropole

Table 3.11 : Les artefacts du cas d'usage "Collaboration"

3.2.5.7 Capacité d'action

Artefact	Capacité d'action	Influence sur état
Comité de pilotage	renseignement de la fréquence de mise à jour des données, publication de jeux de données, rédaction d'une Wanted Data List, réponse aux demandes du laboratoire d'innovation, essaimage de la démarche	Données fermées -> Données ouvertes
Réunion des Directeurs des Systèmes d'Information du territoire	non-observé car non-sollicité	non-observé car non-sollicité
Ateliers Open Data France	non-observé	non-observé
Laboratoire d'innovation	Demande de jeux de données complémentaires, Emergence de deux prototype d'application, demande de jeux de données plu fiables pour renforcer l'application	Données ouvertes ->Prototype de service
Prototypes d'application	Transformer les données ouvertes en informations utiles au citoyen	Données ouvertes -> Informations
Métrosphère	non-observé car non-sollicité	non-observé car non-sollicité
Plan de charge des données à libérer	non-observé car non-sollicité	non-observé car non-sollicité
Convention juridique	non-observé car non-sollicité	non-observé car non-sollicité
Serious game "Explorateur des données territoriales"	non-observé	non-observé
Script ETL Talend	Extraction des données budgétaires nationales et	Données brutes -> données contextualisées

	transformation pour créer des fichiers budgétaires par collectivité	
Réseau social “teamopendata.org”	non-observé car non-sollicité	non-observé car non-sollicité
Outil de validation de données Validata	non-observé car non-sollicité	non-observé car non-sollicité
Comité de Pilotage “Données” des communes	non-observé car non-sollicité	

Table 3.12 : Les capacités d'action du cas d'usage “Collaboration”

3.2.5.8 Séquence d'articulation

Les précédentes observations permettent de mettre à jour les quatre séquences d'articulations. Elles représentent et définissent l'usage des données ouvertes du cas “Collaboration”. Afin de ménager le lecteur, je m'efforcerai de ne présenter en détail qu'une seule des séquences d'application mises à jour.

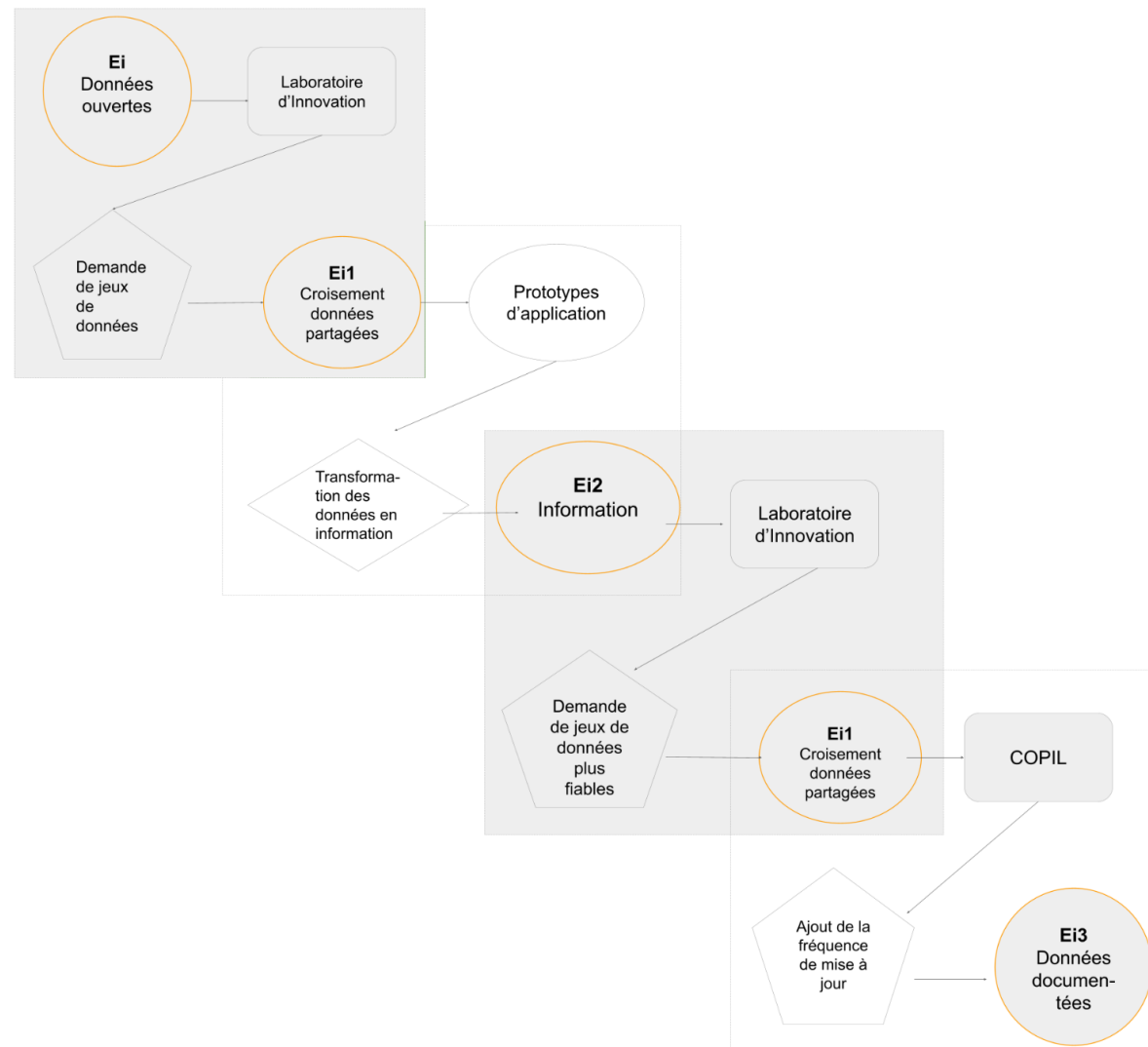


Figure 3.9 : La représentation de la première séquence d'articulation

Cette première séquence d'articulation se déroule de la façon suivante.

Le laboratoire d'innovation prend connaissance des données ouvertes à disposition sur la plateforme *open data* de la métropole "Smart data". Les membres de ce laboratoire émettent le besoin d'avoir accès à des jeux de données complémentaires. Leur but est de créer des applications basées sur plusieurs jeux de données complémentaires. Alors, le comité de pilotage récupère les données demandées et les partage avec eux. À la fin de la première articulation, les données dans un état de données ouvertes (Ei) sont maintenant croisées et complétées avec des données partagées. Les données ouvertes sont devenues un croisement de données (Ei1).

Le deuxième articulation commence par l'utilisation du croisement pour créer des prototypes d'applications. Ces prototypes ont la capacité de transformer le croisement de données en informations utiles pour le citoyen. Par exemple, les données contenant les coordonnées GPS des points de collecte de déchets deviennent un itinéraire à suivre pour déposer ses déchets. À l'issue de cette deuxième articulation, le croisement de données (Ei1) emprunte un nouvel état, celui d'informations utiles au citoyen (Ei2).

La troisième articulation met en œuvre le laboratoire d'innovation qui, après avoir créé des prototypes d'application, souhaiterait maintenant développer des applications fonctionnelles. Or, pour cela, il a besoin de jeux de données fiables, c'est-à-dire contenant des valeurs exactes et mises à jour. Alors, il revient vers le comité de pilotage et demande à ce que les données qui constituent le croisement (Ei1) puissent être fiabilisées. À la fin de cette troisième articulation, les données sont revenues à l'état précédent : elles sont passées de l'état d'informations (Ei2) à celui de croisement (Ei1).

Enfin, la quatrième articulation fait appel au Comité de pilotage qui, pour répondre à la demande du laboratoire d'innovation, va venir documenter l'ensemble des données ouvertes présentes sur la plateforme. Plus précisément, il va venir indiquer la fréquence de mise à jour de tous les jeux de données ouvertes.

In fine, je constate que cette séquence d'articulation ne s'est déroulée qu'une seule fois sans encombre. Dans la majorité des cas, cette séquence était stoppée dès la première articulation. Autrement dit, j'ai pu observer que les demandes de données du Laboratoire d'Innovation qui aboutissent étaient rares et que, par conséquent, le Laboratoire d'Innovation ne pouvait que très rarement baser ses prototypes d'applications sur les données ouvertes.

Par ailleurs, j'ai observé des blocages similaires dans les trois autres séquences d'articulation identifiées.

L'encadré ci-dessous illustre la difficulté que le Comité de Pilotage rencontre dans le partage des données.

Exemple de blocage : les difficultés à partager les données

(source : extraits des notes d'observation)

“L'implication sur le Laboratoire d'Innovation est difficile à comprendre pour les services de la métropole. Ils ne comprennent pas pourquoi ils devraient passer du temps à trouver et partager des données pour alimenter des projets peu ou mal définis. De plus, cela pose un problème de définition de priorités par rapport aux demandes pré-existantes du Comité de Pilotage (feuille de route, wanted data list à tenir à jour, etc.)”

3.2.5.9 Complémentarité

Le résultat de la première articulation relative au croisement de données ouvertes et de données partagées, se traduit par une connaissance accrue du territoire pour les membres du laboratoire d'Innovation. En effet, ces croisements vont permettre aux participants de coupler des informations sur le territoire de façon à en tirer de nouvelles connaissances. Prenons l'exemple d'un projet de ce laboratoire visant à améliorer le tri des déchets dans la métropole. En croisant les données sur la localisation des zones de dépôt de déchets et sur les consignes de tri relatives à ces zones, elle a pu créer de nouvelles connaissances sur l'offre actuelle des services publics relatifs au traitement des déchets.

La deuxième articulation a permis de créer des informations utiles aux citoyens. Prenons l'exemple d'un projet du laboratoire visant à créer des tableaux de bords pour les agences immobilières. Ce projet a permis de générer des informations sur la qualité de vie dans les différentes zones de la métropole. Ces informations permettront aux agences immobilières de mieux conseiller leurs clients sur l'emplacement d'un bien à louer ou à acheter. Sur la base de cet exemple, je peux affirmer que la valeur de ces informations réside dans le fait qu'elles éclairent les décisions des citoyens du territoire de la métropole.

Le résultat de la troisième itération (i.e demande de données plus fiables) se traduit par une amélioration indirecte du “service métropolitain de la donnée” à l'origine de ce cas. Cette demande permet en effet au

groupe de travail de lancer une réflexion autour de la fiabilisation des données et de mettre en place les processus nécessaires. La quatrième itération en témoigne.

Finalement, la quatrième itération se termine par la publication de “données documentées”. La valeur de ce résultat se traduit par une meilleure appréhension de la qualité des données ouvertes par les utilisateurs. Avec les indications sur la fréquence de mise à jour des données, un data scientist peut décider plus facilement d'utiliser ou non ces données, ou de prendre certaines précautions pour ne pas fausser les analyses.

3.2.3 Analyse transversale des quatre cas exploratoires d'usage des données ouvertes

Il est maintenant possible d'identifier les différents éléments constitutifs de l'usage de la donnée dans nos quatre cas et d'émettre des conjectures sur le phénomène qu'est la construction des usages issus des données ouvertes. Par conjecture, il est question d'hypothèses dont il faudra confirmer le bien-fondé dans la suite de la thèse.

3.2.3.1 Une plateforme *open data*

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Plateforme open data	Quels sont les éléments tangibles et intangibles qui définissent les plateformes open data ?	<ul style="list-style-type: none"> • API entrantes • API sortantes 	<ul style="list-style-type: none"> • Hub de données • Dashboard 	Catalogue de données	<ul style="list-style-type: none"> • Catalogue de données • Logiciel de gestion de contenu • Scripts ETL • Coordination avec Open Data France
	Quel est le rôle de la plateforme open data dans la construction de nouveaux usages à partir des données ouvertes ?	<ul style="list-style-type: none"> • Centraliser et agréger les données qui permettront de construire de nouveaux usages • Partage d'informations entre les membres de l'écosystème 	Intermédiaire entre les artefacts technologiques chargés de la production de données, les artefacts sociaux chargés de la construction de nouveaux usages et les artefacts sociaux chargés du suivi des effets de ces nouveaux usages	Socle technique des artefacts chargés d'exploiter les données	<ul style="list-style-type: none"> • Valoriser les données • Animer les interactions avec les membres de l'écosystème

Table 3.13 : Les composants des plateformes *open data*

De cette analyse transversale, je conclus que les plateformes *open data* sont constituées de plusieurs ensembles hétérogènes d'éléments tangibles et intangibles qui lui confèrent une large palette de fonctionnalités et de rôles au sein de la construction des usages.

3.2.3.2 Les parties prenantes des plateformes *open data*

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Parties prenantes	Qui sont les parties prenantes des plateformes ?	Aucune partie prenante identifiée	Aucun partie prenante identifiée	Aucune partie prenante identifiée	<ul style="list-style-type: none"> ● Ville ● Syndicat de transport ● Editeur informatique
	Quel(s) rôle(s) jouent ces parties prenantes ?	Non-observé	Non-observé	Non-observé	Développement des modules technologiques de la plateforme

Table 3.14 : Les parties prenantes des plateformes *open data*

À l'issue de cette analyse transversale, je constate une absence de mécanisme de gouvernance au sein des plateformes *open data* (exception faite de la coordination avec l'association Open Data France observée dans le cas "Collaboration"). Aussi n'ai-je pas identifié de parties prenantes participant à l'évolution de la plateforme, exception faite des agents en charge du développement technique de la plateforme dans le cas "Collaboration". Ces deux constats m'amènent à émettre la conjecture selon laquelle il existe un(des) blocage(s) social qui empêche(ent) la mise en place d'une organisation sociale dotée d'une gouvernance propre autour des plateformes *open data*.

3.2.3.3 Les données ouvertes en tant qu'espèce ingénieur

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Données ouvertes	Quels sont les outils informatiques structurés par les données ouvertes ?	8 visualisations de données	<ul style="list-style-type: none"> • Application mobile de paiement de titres de transport • Système de guidage des automobilistes vers les places disponibles 	Outil d'évaluation des capacités d'accueil touristique des territoires	<ul style="list-style-type: none"> • Application de simulation du potentiel fiscal du territoire • Application permettant de dresser le portrait socio-économique du territoire

Table 3.15 : Les outils structurés par les données ouvertes

Si les plateformes sont bien constituées d'éléments techniques qui permettent aux données de devenir l'espèce ingénieur des écosystèmes, elles structurent un nombre négligeable d'outils informatiques (1 à 2 outils dans chaque cas). Je peux dès lors émettre la conjecture selon laquelle il existe un(des) blocage(s) qui empêche(nt) d'activer les fonctionnalités techniques des plateformes *open data*.

3.2.3.4 Les écosystèmes open data

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Ecosystème open data	Qui sont les acteurs qui constituent les systèmes de production, d'intermédiation et de transformation des données ouvertes ?	<ul style="list-style-type: none"> • Système de production : concepteurs de capteurs, cabinet de conseil, business unit de grand groupe privé • Système d'intermédiation : collectivités territoriales, éditeur de logiciel, business unit de grand groupe privé • Système de transformation : cabinet de conseil, managers d'entreprise 	<ul style="list-style-type: none"> • Système de production : collectivités territoriales, plateforme de services publics • Système d'intermédiation : collectivités territoriales, éditeur de logiciel • Système de transformation : grand groupe privé, réseau de transport 	<ul style="list-style-type: none"> • Système de production : collectivités territoriales, grand groupe privé • Système d'intermédiation : cabinet de conseil, représentants politiques internationaux • Système de transformation : grand groupe privé 	<ul style="list-style-type: none"> • Système de production : collectivités territoriales et représentants politiques nationaux • Système d'intermédiation : éditeur de logiciel, administration nationale et collectivités territoriales • Système de transformation : éditeur de logiciel et tiers-lieux
	Quel est l'état initial (Ei) des données en entrée du système et l'état final (Ef) des données en sortie du système ?	<ul style="list-style-type: none"> • Système de production : données métiers fermées(Ei) --> données ouvertes(Ef) • Système d'intermédiation : données métiers partagées et dispersées (Ei1) données ouvertes demandées (Ei2) --> données métiers centralisées, documentées, connectées, exploitées et garanties (Ef1) données ouvertes téléchargées (Ef2) • Système de transformation : données ouvertes (Ei) --> nouvel usage (Ef) 	<ul style="list-style-type: none"> • Système de production : données métiers fermées(Ei) --> données ouvertes(Ef) • Système d'intermédiation : données métiers dispersées (Ei1) données ouvertes dispersées (Ei2) --> données métiers centralisées (Ef1) et données ouvertes centralisées (Ef2) • Système de transformation : données ouvertes (Ei) --> nouvel usage (Ef1) adaptation des infrastructures au nouvel usage (Ef2) 	<ul style="list-style-type: none"> • Système de production : données métiers fermées(Ei) --> données métier partagées (Ef) • Système d'intermédiation : données métiers partagées (Ei) --> données ouvertes exploitables(Ef) • Système de transformation : données ouvertes (Ei1) données métiers partagées (Ei2) --> nouvel usage (Ef) 	<ul style="list-style-type: none"> • Système de production : connaissances fermées (Ei) --> données ouvertes(Ef) • Système d'intermédiation : données ouvertes (Ei1) données métiers regroupées (Ei2) --> données ouvertes documentées (Ef1) données ouvertes (Ef2) • Système de transformation : données ouvertes (Ei) --> service de renseignement (Ef)

Table 3.16 : Les éléments des écosystèmes open data

Cette analyse transversale confirme le bien-fondé de la subdivision en trois sous-systèmes des écosystèmes open data tirés de la littérature : le système de production de données, le système de transformation des données et le système d'intermédiation.

Aussi, en analysant les états des données en entrée et les états des données en sortie des systèmes, je remarque un chevauchement dans les activités des différents systèmes. Ce constat m'amène à émettre la conjecture selon laquelle il existe un(des) blocage(s) social qui empêche(ent) la mise en place d'une coordination entre les systèmes lors de la construction des usages. Ce système de coordination fluidifierait l'usage des données ouvertes.

3.2.3.5 Les artefacts techniques et sociaux et leur capacité d'action

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Artefacts	Quels sont les types et les objectifs des artefacts technologiques et sociaux à l'oeuvre dans l'usage de l'open data ?	<ul style="list-style-type: none"> • 3 artefacts sociaux en charge de prendre les décisions relatives à la sécurité, la qualité et l'exploitation des données • 3 artefacts techniques ayant pour but de récupérer, normaliser, connecter et exploiter les données • 1 artefact technique aux finalités sociales ayant pour but de recenser, documenter et échanger autour des données 	<ul style="list-style-type: none"> • 1 artefact social ayant pour but de mettre à disposition un lieu pour co-crée des nouveaux usages à partir des données • 5 artefacts techniques ayant pour but de créer, classer, centraliser les données ainsi que de prédire les valeurs des données futures • 5 artefact technique aux finalités sociales ayant pour but de recenser et d'intégrer les données à des applications tierces ainsi que de fournir un service informationnel au citoyen 	<ul style="list-style-type: none"> • 3 artefacts sociaux en charge d'influencer, de guider et d'assurer la collaboration entre acteurs • 1 artefact technique permettant de demander les données • 1 artefact technique aux finalités sociales ayant pour but de recenser, documenter et échanger autour des données à ouvrir et/ou à utiliser 	<ul style="list-style-type: none"> • 7 artefacts sociaux en charge de prendre les décisions relatives à l'usage des données ouvertes • 4 artefacts techniques ayant pour but de récupérer, d'éditer et d'utiliser les données • 2 artefacts techniques aux finalités sociales ayant pour but d'échanger autour de l'usage des données ouvertes

Table 3.17 : Les artefacts sociaux et techniques impliqués dans l'usage de l'open data

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Capacité d'action	Quelles sont les capacités d'action des artefacts ?	7 capacités d'action exprimées sur les 7 artefacts	7 capacités d'action exprimées sur les 11 artefacts dont 4 techniques non-exprimés	capacités d'action exprimées sur les 5 artefacts dont 1 technique et 1 social non-exprimé	4 capacités d'action exprimées sur les 13 artefacts dont 4 technique et 5 social non-exprimé
	Quel est leur influence sur l'état des données ouvertes ?				

Table 3.18 : Les capacités d'action des artefacts et leur influence sur l'état des jeux de données ouverts

De cette analyse transversale, je conclus qu'il existe une différence entre le nombre d'artefacts évoqués pour permettre l'usage des données ouvertes et le nombre d'artefacts ayant effectivement exprimé leurs capacités d'action durant la durée de l'observation et donc modifier l'état des jeux de données. Ce constat m'amène à émettre la conjecture selon laquelle il existe un(des) blocage(s) qui empêche(ent) aux artefacts d'exprimer leurs capacités d'action et de modifier l'état des données.

3.2.3.6 Les états des jeux de données ouvertes

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Etats	Quelles sont les différentes configurations dans lesquelles se trouvent les jeux de données ouvertes ?	<ul style="list-style-type: none"> • données indexées • données normalisées • données statiques • données dynamiques • données calculées • données interprétées • données privées • données partageables et utilisables • données brutes • données connectées • données indépendantes • données indexées 	<ul style="list-style-type: none"> • données simples • données calculées • besoin des utilisateurs • informations nécessaires • données en création • données découvrables • données non-découvrables • données statiques • données dynamiques • informations 	<ul style="list-style-type: none"> • privées • données partageables et utilisables • données brutes • données connectées • données découvrables • données non-découvrables • données brutes • données contextualisées • données non-certifiées • données certifiées • données simples • données calculées 	<ul style="list-style-type: none"> • données en création • données métiers • données manuelles • données format unique • données plusieurs formats • données non-documentées • données documentées • données brutes • données visualisées • données non mises à jour • données à jour • données non-découvrables • données consultables • données réutilisables • données statiques • données dynamiques

Table 3.19 : Les différentes configurations des jeux de données ouverts

L'analyse transversale des états des jeux de données ouvertes confirme les arguments de la revue de littérature selon lesquels les données ouvertes passent par différents états lors de la construction de nouveaux usages.

3.2.3.7 Les séquences d'articulation

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Séquence d'articulation	Quelles sont les différentes séquences identifiables dans l'usage des données ouvertes ?	2 séquences d'articulation dont 1 séquence dédiée à la collecte manuelle des données et l'autre à leur traitement automatique.	2 séquences d'articulation dont 1 séquence dédiée à la production et l'exploitation autour des données relatives au parking et l'autre à la production et à l'exploitation autour des données relatives à la souscription aux services de mobilité.	2 séquences d'articulation dont 1 séquence dédiée à rendre utilisables les données et l'autre à connecter les données fiscales avec d'autres données dans le but de construire un nouvel usage.	4 séquences d'articulation dont la première dédiée à la mise à jour des données ouvertes, la deuxième à l'extraction automatique des données métiers, la troisième à mise à jour du plan de charge d'ouverture et la quatrième également dédiée à l'extraction automatique de données.

Table 3.20 : Les séquences d'articulation identifiables dans l'usage de l'open data

Enfin, j'ai pu me rendre compte de différents blocages qui empêchent le déroulement des séquences d'articulation. Ce constat m'amène à émettre la conjecture selon laquelle ces blocages arrêtent les séquences d'articulation permettant la construction des usages issus des données ouvertes.

3.2.3.8 Les complémentarités formées

Element	Question empirique	Cas 1 - Quartier	Cas 2 - Transport	Cas 3 - Logement	Cas 4 - Collaboration
Complémentarités	Quel est le résultat de chacune des séquences d'articulation ?	<ul style="list-style-type: none"> Donnée normalisées Connexion des données entre elles Données interprétées, prenant la forme d'alertes en cas d'anomalie dans les données. 	<ul style="list-style-type: none"> Données indexées dans le tableur collaboratif Données indexées sur la plateforme OpSoft Prédictions sur la disponibilité des places de parking Informations données aux automobilistes quant au trajet à suivre pour se garer au plus vite 	<ul style="list-style-type: none"> Données ouvrables Evaluation de la compatibilité réelle de ces données avec la création d'un service commercial. Nouvel actif pour améliorer leur offre de service 	<ul style="list-style-type: none"> Croisement de données Création d'informations utiles aux citoyens Demandes de données plus fiables Données documentées
	Quelle est la valeur ajoutée émergente de chacun de ces résultats ?	<ul style="list-style-type: none"> Amélioration de la qualité des données. Réduction du risque d'erreur dans l'analyse dont elles peuvent faire l'objet. Création de ponts entre applications et systèmes d'informations. Alerte en cas d'anomalie dans fonctionnement des équipements 	<ul style="list-style-type: none"> Identification de synergies pertinentes à mettre en place entre les différents fournisseurs de données Compréhension commune de la disponibilités des parkings et des actions à entreprendre. Prévenir saturation des parkings et les difficultés de circulation liées. diminution du temps d'attente pour trouver une place de parking Meilleure gestion du trafic routier 	<ul style="list-style-type: none"> Amélioration de la prospection commerciale Amélioration de l'offre de service 	<ul style="list-style-type: none"> Connaissance accrue du territoire pour les membres du laboratoire d'Innovation Des informations qui permettront aux agences immobilières de mieux conseiller leurs clients sur l'emplacement d'un bien à louer ou à acheter. Amélioration du service métropolitain de la donnée Meilleure appréhension de la qualité des données ouvertes par les utilisateurs.

Table 3.21: Les complémentarités formées lors des séquences d'articulation

De cette analyse transversale, je tire la conclusion que les complémentarités formées sont dans la majorité des cas relatives aux données mêmes et que peu d'usages réels sont construits à l'issue des projets. Ce constat m'amène à émettre la conjecture qu'il existe certains blocages qui empêchent la construction des usages issus des données ouvertes.

3.2.3.9 Synthèse des résultat de l'analyse transversale

Ces analyses transversales des données m'invitent à émettre une conjecture quant aux blocages les plus récurrents dans l'usage des données. Ils peuvent être classés en deux catégories : les blocages d'ordres

sociaux et les blocages d'ordre technique. Ces analyses ont été rendues possibles par l'utilisation du logiciel Nvivo 2. Je détaillerai ci-après ces deux catégories de blocage.

3.2.3.9.1 Les blocages d'ordre technique

Si les plateformes sont bien constituées d'éléments techniques qui permettent aux données de devenir l'espèce ingénieur des écosystèmes, elles structurent un nombre négligeable d'outils informatiques (1 à 2 outils dans chaque cas). Je peux dès lors émettre la conjecture selon laquelle il existe un(des) blocage(s) qui empêche(nt) d'activer les fonctionnalités techniques des plateformes *open data*.

J'observe une différence entre le nombre d'artefacts évoqués pour permettre l'usage des données ouvertes et le nombre d'artefacts ayant effectivement exprimé leurs capacités d'action durant la durée de l'observation. Ce constat m'amène à émettre la conjecture selon laquelle il existe un(des) blocage(s) qui empêche(ent) aux artefacts techniques d'exprimer leurs capacités d'action.

Enfin, j'ai pu me rendre compte de différents blocages d'ordre technique empêchant le déroulement des séquences d'articulation. Parmi eux, des données de mauvaise qualité (valeurs manquantes ou erronées) ou un volume de données insuffisant pour envisager une exploitation ultérieure. Ce constat m'amène à émettre la conjecture selon laquelle ces blocages techniques bloquent les séquences d'articulation permettant l'usage des données ouvertes.

3.2.3.9.1 Les blocages d'ordre social

Dans les quatre cas, j'ai constaté l'absence de mécanisme de gouvernance au sein des plateformes *open data* (exception faite de la coordination avec l'association Open Data France observée dans le cas "Collaboration"). Aussi n'ai-je pas identifié de parties prenantes participant à l'évolution de la plateforme, exception faite des agents en charge du développement technique de la plateforme dans le cas "Collaboration". Ces deux constats m'amènent à émettre la conjecture selon laquelle il existe un(des) blocage(s) social qui empêche(ent) la mise en place d'une organisation sociale dotée d'une gouvernance propre autour des plateformes *open data*.

Dans les quatre cas, je remarque un chevauchement dans les activités des différents systèmes. Ce chevauchement se constate en analysant les états des données en entrée et les états des données en sortie des systèmes.

Ce constat m'amène à émettre la conjecture selon laquelle il existe un(des) blocage(s) social qui empêche(ent) la mise en place d'une coordination entre les systèmes. Ce système de coordination fluidifierait l'usage des données ouvertes.

J'ai pu apprécier une différence entre le nombre d'artefacts évoqués pour permettre l'usage des données ouvertes et le nombre d'artefacts ayant effectivement exprimé leurs capacités d'action durant la durée de l'observation. Ce constat m'amène à émettre la conjecture selon laquelle il existe un(des) blocage(s) socia(ux)l qui empêche(ent) aux artefacts sociaux d'exprimer leurs capacités d'action.

Enfin, je constate de nombreux blocages d'ordre social empêchant le déroulement des séquences d'articulation. Parmi eux, le manque d'informations et de compréhension des acteurs quant aux finalités de l'usage des données ; les peurs que véhiculent l'ouverture des données ; ou encore le manque de contreparties financières en échange de la mise à disposition de données. Ce constat m'amène à émettre la conjecture selon laquelle ces blocages sociaux bloquent les séquences d'articulation permettant usage des données ouvertes.

Ce classement a donné lieu à de nouvelles récoltes de données. En l'occurrence, il a permis de guider la réalisation d'entretiens semi-directifs.

Conclusion et transition

L'analyse transversale des cas m'a permis de mettre en lumière les difficultés dans l'usage des données ouvertes. J'ai émis des conjectures afférentes à ces difficultés et les ai classées en deux catégories. La première catégorie regroupe trois conjectures relatives aux blocages d'ordre technique tandis que la deuxième englobe quatre conjectures relatives aux blocages d'ordre social. Ces conjectures ont guidé la réalisation d'entretiens semi-directifs visant à les confronter avec l'expérience de praticiens impliqués dans l'usage des données ouvertes. La prochaine partie présente les résultats de ces confrontations ainsi que notre cadre conceptuel affiné grâce à ces dernières.

3.3 Confrontation des observations avec les praticiens et formulation d'un nouveau cadre conceptuel

Résumé

Dans cette partie, l'objectif est d'affiner mon modèle conceptuel. Pour cela, il est question de confronter le modèle empirique construit dans la partie précédente aux expériences vécues de parties prenantes de plateformes *open data*. Dans un premier temps, j'exposerai comment je suis parvenue à constituer notre échantillon de parties prenantes. Dans un deuxième temps, je présenterai la méthode de collecte et d'analyse des entretiens des parties prenantes. Dans un troisième temps, je révélerai les résultats de ces analyses avant de conclure en présentant le modèle théorique affiné.

3.3.1 Constitution de l'échantillon de parties prenantes

L'échantillon est constitué de parties prenantes de l'ouverture des données. J'ai, au préalable établi que cette population était constituée des organisations ou individus jouant les rôles suivants :

- architecte : il participe à l'architecture technique de la plateforme (apport de modules ou de données) ainsi qu'au développement de mécanismes de gouvernance;
- acteur : il met à disposition ses moyens techniques et ses compétences pour exploiter les modules et données présentes sur la plateforme;
- bénéficiaire : il fait l'expérience de la valeur créée par la plateforme;

Les investigations ont été orientées de manière à obtenir un échantillon représentatif de cette distribution de rôles. De fait, notre échantillon est constitué selon des critères théoriques du *choix raisonné* (Royer et Zarlowski, 2014 in Thiétart et al., p.233). Or, très vite, des difficultés sont apparues pour identifier les parties prenantes selon ces critères et ce, pour plusieurs raisons.

La première raison réside dans le fait que l'usage de l'open data est un phénomène émergent et que les organisations ou individus qui y contribuent n'ont pas forcément conscience du rôle qu'elles y jouent. Dès lors, il existe très peu d'informations disponibles permettant *a priori* d'identifier des architectes, des

acteurs et des bénéficiaires. La deuxième raison est que la très grande majorité des collectivités qui tombent sous le coup de l'obligation d'ouverture des données, ne la respectaient pas au moment où j'ai mené cette recherche exploratoire. Elles se trouvaient dans l'illégalité. On comprend aisément les difficultés que peut rencontrer un chercheur qui cherche à identifier des individus ou organisations qui n'existent pas encore, ou qui existent mais qui ne sont pas encore en conformité dans la loi.

Du fait de ces difficultés, j'ai appliqué la technique de la boule de neige. Celle-ci consiste à :

“trouver un premier répondant qui vérifie les critères de sélection définis par le chercheur. On demande à ce premier interlocuteur d'en désigner d'autres, qui seront, eux aussi, susceptibles de présenter les caractéristiques requises, et ainsi de suite” Royer et Zarlowski (2014 in Thiéart et al., p.234).

In fine, avec cette technique, j'ai constitué l'échantillon représenté dans la figure ci-dessous :

Côté	Rôle	Poste	Durée de l'entretien
Offre	Architecte	Responsable des Systèmes d'Information - Communauté de Commune du Massif du Vercors	41 min
Offre, Demande	Architecte, Acteur, Bénéficiaire	Directeur d'études Offre de Mobilité - CEREMA	46 min
Maintenance	Architecte, Acteur	Co-fondateur - Dataactivist	35 min
Offre	Architecte	Conseiller stratégique “valorisation des données” - Etalab	43 min
Demande	Acteur, Bénéficiaire	Co-fondateur - Tiers lieux sur les nouvelles mobilités	28 min
Demande	Acteur, Bénéficiaire	Concepteur d'une application utilisant les données des vélos en libre-service	35 min
Demande, Maintenance	Acteur, Bénéficiaire	Chargé d'expérimentation data - Laboratoire d'innovation	59 min
Demande	Bénéficiaire	Chargé de projet innovation ouverte	63 min

Offre	Architecte	Responsable de l'ouverture des données transports - Transports.data.gouv.fr	63 min
Offre, Maintenance	Architecte	Chargé de mission Open Data - Grand Lyon	43 min
Demande	Bénéficiaire	Responsable du pôle ressource et données touristiques - Atout France	35 min
Demande	Acteur	Chef de projet Data - Laboratoire d'innovation	55 min
Demande	Bénéficiaire	Consultant Open Data - Indépendant	35 min
Offre, Maintenance	Architecte	Directeur des Systèmes d'Information - Métropole	41 min

Table 3.22 : L'échantillon de parties prenantes

3.3.2 Méthode de collecte et d'analyse des entretiens

3.3.2.1 Collecte

J'ai réalisé des entretiens semi-directifs auprès de ces quatorze parties prenantes. Pour les mener à bien, j'ai élaboré un guide d'entretien. Ce guide d'entretien est divisé en une liste de thèmes qui reprend tous les éléments constitutifs du modèle empirique : plateforme *open data*, parties prenantes, données ouvertes, écosystème open data, artefacts, capacité d'action, états, séquence d'articulation, complémentarités.

L'interviewé découvre chaque thème au fur et à mesure et peut donner son point de vue sur celui-ci. Au cours de la discussion sur chaque thème, je lui présente les conjectures tirées du modèle empirique. En réaction à cette présentation, l'interviewé a la possibilité de donner son ressenti vis-à-vis d'elles.

Par la suite, j'ai retranscrit l'ensemble des entretiens.

3.3.2.2 L'analyse

Pour analyser les retranscriptions, je me suis doté du logiciel Nvivo. Dans un premier temps, j'ai découpé le texte en autant de parties que d'éléments du modèle évoqués au cours de l'entretien. J'ai ensuite créé un système de code permettant d'interpréter le texte. Ce système de code est construit selon les différentes subdivisions de chaque élément. Par exemple, pour le texte relatif aux plateformes *open data*, le système de code reprend les subdivisions suivantes :

- Architecture de la plateforme
- Composants stables de la plateforme
- Mécanismes de gouvernance

À la suite du codage du texte, j'obtiens une liste de verbatim liés à chaque subdivision. À ce moment, il m'est possible de commencer l'interprétation.

L'interprétation vise à pouvoir confronter les verbatim avec les conjectures tirées du modèle empirique. Au regard de ces dernières, deux voies d'interprétation sont possibles :

1. La première voie consiste à extraire des verbatim une liste d'éléments pour pouvoir la comparer avec la liste tirée du modèle empirique ;
2. La seconde voie consiste à extraire des verbatim le sens que les parties prenantes interviewées donne à certains éléments pour pouvoir le comparer avec le sens déterminé par le modèle empirique.

In fine, l'interprétation permet de réaliser un tableau de comparaison avec d'un côté les éléments du modèle empirique et de l'autre, les éléments issus des verbatim.

Élément du modèle	Conjecture	Verbatim
Plateforme open data - Architecture	Les fonctionnalités techniques des plateformes sont inhibées par des blocages	<p>Entretien 2 : “Je trouve l’analogie (de l’open data) avec les logiciels libres très intéressantes. Hier, j’ai voulu télécharger des CRM en logiciels libres sauf qu’il faut installer 1000 trucs, j’ai rien compris. L’open data c’est pareil. Quand j’arrive sur une plateforme, un truc de data, j’y comprends rien. Et c’est pour ça que ça reste dans un entre-soi.”</p>
Artefacts technologiques	Des blocages empêchent aux artefacts techniques d’exprimer leurs capacités d’action.	<p>Entretien 5 : “Si on a pas de données sur des événements avec combien de personnes, si on a pas des données sur des flux non statiques... ce n’est pas avec des données pour savoir où sont les silos à verre qu’on va générer des millions d’euros. Avec des données statiques, on ne peut rien faire”</p>
Plateforme open data - Gouvernance	Il existe un(des) blocage(s) social qui empêche(ent) la mise en place d’une organisation sociale dotée d’une gouvernance propre autour des plateformes open data.	<p>Entretien 8 : “Il y a tout un tas de directions réticentes, qui ne veulent pas une centralisation des données et qui ne veulent pas les mettre à disposition. Et ce, pour des raisons de pouvoirs mais aussi du fait qu’ils n’ont pas les mêmes stratégies derrière la donnée, ils n’ont pas le même rapport à la diffusion”.</p>
Séquence d’articulation	Des blocages sociaux bloquent les séquences d’articulation permettant l’usage des données ouvertes.	<p>Entretien 4 : “(Animer des réseaux, c’est pas leur cœur de métier. En 2015, on a publié un ouvrage (.?.) où on parle des freins à l’ouverture, côté collectivité où on parle des questions d’animation. Les grands freins n’ont pas trop changé. Pas beaucoup mis l’accent sur les ré-utilisateurs, pas d’animation pour faire suite aux initiatives ponctuelles ”</p>

Table 3.23 : Extrait de la confrontation des éléments empiriques et des verbatims des entretiens

3.3.3 Résultats et présentation du modèle théorique affiné

Dans cette section, il est question de confronter les résultats de nos interprétations avec la littérature et ce, dans le but d'affiner notre modèle théorique. Grâce à la littérature, je vais caractériser plus précisément la nature des barrières socio-techniques mentionnés dans les entretiens analysés ci-dessus et qui freinent l'usage des données ouvertes. Je mobiliserai pour cela les catégories construites précédemment. Les détails apportés par la littérature me permettront d'affiner le modèle théorique. Enfin, je présenterai la représentation de ce modèle théorique affiné.

3.3.3.1 Les barrières techniques (human-computer interactions)

3.3.3.1.1 Les barrières techniques inhibant les fonctions des plateformes open data

- Une première barrière technique apparaît dans le fait que les plateformes *open data* ne sont pas multilingues, ce qui peut empêcher un utilisateur de comprendre les données présentes sur la plateforme ainsi que le moyen de les obtenir (Zuiderwijk et Janssen, 2014 ; Petychakis et al., 2014)?
- Aussi, une barrière inhibant les fonctions des plateformes réside dans le fait qu'elles ne sont pas toujours constituées d'interfaces claires et compréhensibles montrant les différentes modalités de recherche de données disponibles. De ce fait, les utilisateurs de la plateforme ont des difficultés à trouver les données qu'ils recherchent et les chances qu'ils reviennent sur la plateforme sont faibles (Zuiderwijk et al., 2012 ; Petychakis et al., 2014).
- Aussi, le fait que certaines plateformes soient d'un accès technique difficile (faible temps de réaction des interfaces, temps de chargement des pages élevées) n'inspire pas confiance à l'utilisateur. Dès lors, il a peu de chances de baser une application, un service ou tout autre forme de réutilisation sur les données ouvertes offertes par la plateforme (Petychakis et al., 2014).

3.3.3.1.2 Les barrières techniques inhibant les capacités d'action des artefacts techniques

- Le fait que les données publiées sur certaines plateformes ne soient pas sous un format lisible par machine peut participer à l'inhibition des capacités d'action des artefacts techniques. Le fait que les formats ne soient pas lisibles par machine implique que toute réutilisation demandera un travail

important pour le réutilisateur avant que d'autres artefacts techniques puissent exploiter les données. Cette charge de travail peut décourager l'utilisateur qui ne pourra pas mettre en œuvre les artefacts qu'il avait imaginés pour utiliser les données (Petychakis et al., 2014).

- Des données qui sont obsolètes, non-valides ou qui ne sont pas mises à jour régulièrement bloquent rapidement les capacités d'action des artefacts techniques. J'ai présenté un exemple de ce type de barrière dans le cas "Logement" de la section précédente. Les données MAJIC contenaient des adresses inexactes, rendant impossible tout usage de ces données par des data-scientist ou autres techniciens de la donnée, sous peine de faire des interprétations erronées (Lee et Kwak, 2012).

3.3.3.1.3 Les barrières techniques forçant l'arrêt des séquences d'articulation entre artefact technique et artefact social

- Les séquences d'articulation impliquant le passage d'un artefact technique à un artefact social peuvent être bloquées par le faible niveau d'informations sur les métadonnées. Sans ces informations, il est très difficile d'interpréter le sens des données. De fait, tout artefact social, comme un groupe de travail, un comité ou une instance exécutive pourra difficilement prendre une décision ou poursuivre tout simplement le travail d'exploitation des données (Conradie et Choenni, 2012).
- Aussi, un artefact social pourra-t-il être découragé de poursuivre toute utilisation des données en constatant que les procédures et conditions de réutilisation des données sont différentes et hétérogènes. Cela impliquerait pour lui de négocier avec chaque producteur de données pour pouvoir utiliser ses données et donc, d'investir énormément de temps sans même savoir si les données pourraient lui permettre de développer les services informationnels souhaités (Janssen, 2011).
- Cette même incertitude sur les bénéfices potentiels qu'un utilisateur pourra tirer des données implique un arrêt des séquences d'articulation dans le cas où les données sont de mauvaise qualité. Un utilisateur sera rarement enclin à investir du temps pour améliorer les données d'un côté, et prendre le risque qu'elle ne lui apporte aucun bénéfice de l'autre (Zuiderwijk et Janssen, 2015).

3.3.3.2 Les barrières sociales

3.3.3.2.1 *Les barrières sociales empêchant la mise en place d'une organisation sociale dotée d'une gouvernance propre autour des plateformes*

- La première de ces barrières évoquées dans la littérature est le manque d'un cadre juridique et politique clair en matière d'ouverture des données. Celui-ci empêche la mise en œuvre d'initiatives durables et résilientes et donc d'une organisation stable. Sans cadre juridique et politique stable, le statut des données peut changer à tout moment, et les modèles de financement de l'usage peuvent être fragilisés. Il est très difficile dans ce contexte de mettre en place toute forme d'organisation sociale (Barry et Bannister, 2014).
- Le fait que l'open data ne fait pas partie du travail courant des organisations, constitue une deuxième barrière empêchant la mise en place d'une organisation sociale dotée d'une gouvernance propre autour des plateformes. En effet, cela a pour conséquence qu'elles ne mettent en place aucun processus d'ouverture ou d'utilisation des données et par conséquent aucune organisation sociale (Conradie et Choenni, 2012).
- Enfin, le manque de vision cohérente sur le financement et de la structuration de l'utilisation des données est aussi pointé du doigt dans la littérature comme un motif entraînant l'absence d'organisation sociale. Sans cette vision, il est difficile d'unir les ressources et compétences d'acteurs hétérogènes autour d'une organisation (Martin, 2014).

3.3.3.2.2 *Les barrières sociales empêchant la mise en place d'une coordination entre les systèmes*

- Une des principales barrières empêchant la mise en place d'une coordination entre les systèmes réside dans les conflits autour de la définition des données. Par exemple, plusieurs départements d'une entreprise peuvent avoir des définitions différentes et divergentes de ce qu'est un client : certains départements le définissent comme l'entité effectuant le paiement d'une facture, d'autres comme la personne bénéficiaire d'un produit ou d'un service, etc. Dès lors, il est très difficile de créer des informations homogènes entre les participants d'un projet dans la mesure où chacun aura

sa propre conception de différents concepts ou entités. *In fine*, il sera très difficile de coordonner les actions de ces acteurs (Zhang et al., 2005).

- L'autre barrière plus directe concerne le manque d'information des producteurs sur l'identité des utilisateurs externes de leurs données. La coordination des acteurs apparaît alors difficile. Les producteurs ont rarement cette information en raison d'un principe inhérent au mouvement d'ouverture des données : la mise à disposition des données sans demande d'aucune contrepartie, ce qui inclut la déclaration de son identité en échange de l'utilisation des données (Archer et al., 2003).
- Enfin, une ultime barrière peut être identifiée du côté de la recherche, qui s'est peu intéressée aux relations entre l'offre et la demande de données. Par conséquent, il existe peu d'éclairage sur les façons de mettre en place une coordination entre systèmes.

3.3.3.2.3 *Les barrières sociales inhibant les capacités d'action des artefacts sociaux*

- La littérature a relevé aussi que certains artefacts sociaux, bien qu'en capacité de faire usage des données se trouvaient inhibés par certains facteurs. Parmi eux, on trouve la peur que de mauvaises interprétations soient faites par des utilisateurs extérieurs à l'organisation, ce qui empêche certains producteurs de décider de l'ouverture de leurs données (Conradie et Choenni, 2012).
- Un autre facteur peut expliquer ces barrières : la difficulté de résoudre le dilemme moral que pose l'ouverture des données entre l'intérêt public de l'ouverture d'un côté et la potentielle menace sur les intérêts privés qu'elle génère. (Martin, 2014). Ce dilemme se pose aussi bien pour les producteurs, les intermédiaires que les utilisateurs.
- Enfin, les artefacts sociaux peuvent aussi être bloqués par une culture organisationnelle fondée sur l'aversion au risque. De fait, ils ne pourront mettre en œuvre une stratégie d'ouverture des données dans laquelle les réutilisations futures sont inconnues (Martin, 2014).

3.3.3.2.3 Les barrières sociales forçant l'arrêt des séquences d'articulation entre artefact social et un artefact technique

- Les séquences d'articulation impliquant le passage d'un artefact social à un artefact technique peuvent être bloquées par un manque de connaissance d'un domaine, ce qui fait qu'un individu, un groupe ou toute autre instance sera dans l'incapacité de connecter, comparer et croiser des données de manière à leur donner du sens (King et al., 2011).
- Aussi le manque de compétences et d'expertise en management, visualisation et analyse de données empêche les groupes d'utilisateurs de tirer quelque valeur des données ouvertes (Graves et Hendler, 2014).
- Enfin, ces séquences peuvent être également bloquées du fait des profils difficiles à intégrer dans le processus d'utilisation et de transformation des données. Selon Klerk (2011) et Eckerberg et al. (2018), il est très difficile d'intégrer des experts dans un domaine de connaissance et des experts techniques de la données au sein d'un même processus et de faire converger leur efforts

La figure ci-dessous résume l'ensemble des barrières socio-techniques relevées dans la littérature :

Type de barrières	Groupe/cluster	Barrières à l'usage de l'open data	Source
Technique human-computer interaction	Inhibition des fonctionnalités techniques des plateformes	Les plateformes ne sont pas multilingues	Zuiderwijk et Janssen (2014) ; Petychakis et al. (2014)
		Les plateformes open data ne sont pas toujours constituées d'interfaces claires et compréhensibles montrant les différentes modalités de recherche de données disponibles	Zuiderwijk et al. (2012) ; Petychakis et al. (2014)
		L'accès difficile temps de réponse, chargement des pages aux plateformes n'inspire pas confiance aux utilisateurs	Zuiderwijk et Janssen (2014)
	Inhibition des capacités d'action des artefacts techniques	Les données publiées sur certaines plateformes ne soient pas sous un format lisible par machine	Petychakis et al., (2014)
		Les données ne sont pas mises à jour régulièrement ou sont obsolètes et non-valides	Lee et Kwak (2012)
	Arrêt des séquences d'articulation	Le faible niveau d'information sur les métadonnées rend l'interprétation du sens des données difficile	Conradie et Choenni (2012)
		Les procédures et conditions de réutilisation des données sont différentes et hétérogènes, ce qui rend difficile le développement de services informationnels	Janssen (2011)
		Dans la mesure où les utilisateurs ne peuvent anticiper les bénéfices potentiels tirés des données, ils ne les utiliseront pas si les données sont de mauvaise qualité	Zuiderwijk et Janssen (2015)

Table 3.24 : Extrait du tableau récapitulatif des barrières socio-techniques freinant l'usage des données

3.3.3.3 Présentation du modèle théorique affiné

Grâce à la littérature, j'ai pu identifier un ensemble de barrières socio-techniques freinant l'usage des données ouvertes. Elles ont été classées en utilisant les catégories construites à partir d'études de cas empiriques décrites précédemment dans ce chapitre. Ce classement permet d'aboutir à un framework qui peut être utilisé pour identifier les barrières socio-techniques que les acteurs souhaiteraient voir lever pour faire usage des données. Aussi me permet-il d'affiner le modèle conceptuel que je mobiliserai pour comprendre l'usage des données ouvertes. Je présenterai ci-dessous le framework et le modèle conceptuel représentant l'usage des données ouvertes.

3.3.3.3.1 Un framework pour identifier les barrières à lever pour mettre en oeuvre l'usage des données ouvertes

À partir du classement effectué ci-dessus, il est possible de conclure que l'usage des données ouvertes peut se décomposer en cinq étapes. à chacune de ces étapes un certain nombre de barrières peuvent apparaître, freinant ainsi l'usage des données. La figure ci-dessous résume ces étapes et les barrières afférentes.

Etape d'usage	Barrières socio-techniques
Utilisation des fonctionnalités de la plateformes <i>téléchargement/flux de données</i>	Manque de confiance dans la plateforme
	Plateforme disponible dans une seule langue
	Manque de lisibilité des interfaces
Evaluation du potentiel des données à partir de leur représentation	Manque de connaissance du domaine relatif aux données
	Manque de compétences et d'expertise technique
	Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation
Transformation des données grâce à des artefacts techniques	Format de données illisible par machine
	Mauvaise qualité des données
	Hétérogénéité des condition d'utilisation des données
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	Manque d'un cadre juridique et politique clair
	Manque d'une vision cohérente autour de la production et de l'utilisation des données
	Le classement de l'open data en bas de la liste des priorités
Institutionnalisation de l'usage des données	L'incompatibilité des structures organisationnelles avec l'ouverture des données
	Divergence culturelle de l'ouverture des données par rapport aux cultures organisationnelle

Table 3.25 : framework d'identification des barrières face à l'usage de l'open data

3.3.3.3.2 Un modèle théorique affiné pour comprendre l'usage des données ouvertes

Le framework présenté ci-dessus me permet d'affiner le modèle théorique présenté dans la première section de ce chapitre. Plus particulièrement, il vient modifier notre compréhension de trois éléments :

- la nature des relations entre la plateforme *open data* et les système qu'elle alimente,
- la nature des enchaînements entre les différentes articulations qui forment une séquence,
- les interactions entre les systèmes de production, d'intermédiation et de transformation.

La figure ci-dessous représente le modèle théorique affiné.

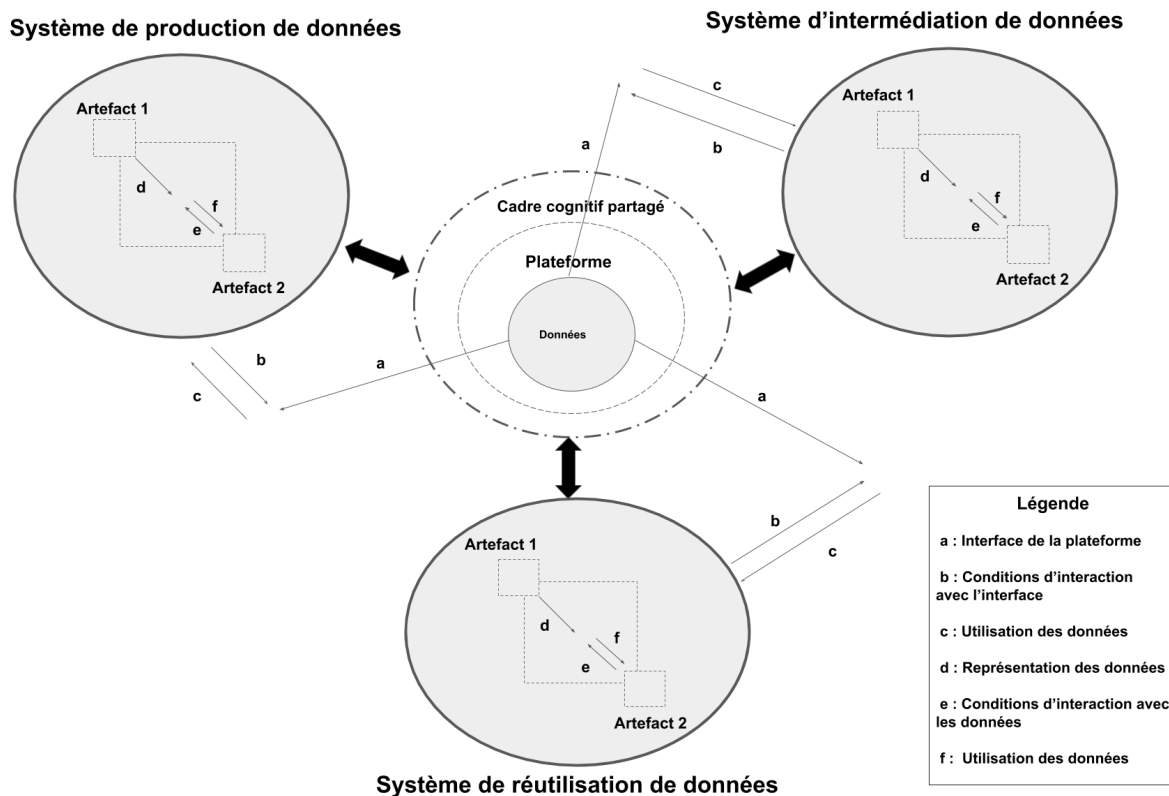


Figure 3.10 : Un modèle conceptuel affiné de l'usage de l'open data

Conclusion du chapitre 3

L'objectif de ce chapitre était d'identifier les situations problématiques auxquels les usagers des données ouvertes font face et qui bloquent leurs activités de construction.

La première section de ce chapitre m'a permis, à partir des construits définis dans la revue de littérature, d'inventer un modèle conceptuel permettant de comprendre la construction des usages des données ouvertes.

En premier lieu, ce modèle établit un premier socle de compréhension de la structure statique de l'écosystème dans lequel la construction des usages a lieu. Je conclus que la construction des usages a lieu dans un écosystème alimenté par une plateforme *open data* qui diffuse des données agissant comme une espèce ingénier en structurant les interactions entre trois systèmes (système de production, d'intermédiation et de réutilisation de données).

Ensuite, passant du général au particulier, ce modèle établit un premier socle de compréhension de la structure statique de la construction des usages, c'est-à-dire des acteurs et artefacts impliqués. Je conclus que la construction des usages relève de transformations de jeux de données dans un état initial en jeux de données dans un état final et ce, par la mise en oeuvre de capacités d'action spécifiques d'artefacts sociaux et techniques dont le rôle est de faire passer les jeux de données d'un état à un autre.

Le modèle ajoute une couche de compréhension dynamique au-dessus de la couche statique, visant à comprendre comment les artefacts et acteurs impliqués se mettent en mouvement pour construire les usages. A ce niveau, je conclus que la construction résulte de capacités d'action qui s'articulent selon des séquences d'articulation définies comme une suite particulière dans lesquelles ces dernières interagissent, se renforcent mutuellement et mettent en place les conditions propices au changement d'état des jeux de données. En ce sens, les usages construits sont définis comme des complémentarités. Autrement dit, les usages sont le résultat de capacités d'action coordonnées et se renforçant mutuellement, chaque usage possédant la caractéristique particulière d'être supérieur à la somme des capacités d'action qui le compose.

Enfin, ce modèle établit que l'émergence de complémentarités que sont les usages vient alimenter l'évolution des écosystèmes dans lesquels la construction des usages a lieu. Les écosystèmes passeraient

ainsi par quatre phases : une phase de genèse, une phase d'expansion, une phase de leadership et une phase d'auto-renouvellement.

Ce modèle m'a permis, dans la deuxième section du chapitre de construire un instrument de collecte de données sous la forme d'une grille d'observations applicables au projet de construction des usages des données ouvertes (Table 3.8). J'ai ensuite mis en place une démarche d'exploration hybride en utilisant la méthode des cas qui m'a amené à effectuer des observations participantes dans quatre démarches de construction d'usage des données ouvertes. Ces observations participantes m'ont permis d'établir 7 conjectures sur les situations problématiques auxquels les usagers des données faisaient face et de les classer en deux catégories : les situations problématiques d'ordre social et les situations problématiques d'ordre technique.

Dans la troisième section de ce chapitre, j'ai confronté les conjectures et leur classification avec l'expérience vécue par les usagers des données ouvertes. Ces confrontations m'ont permis d'affiner ma compréhension des situations problématiques. Je conclus que les usagers des données ouvertes font face à quatorze situations problématiques prenant la forme de barrières socio-techniques et qu'on peut les classer selon deux catégories : les barrières socio-techniques relatives aux interaction homme-machine⁶ et les barrières socio-techniques relatives aux interactions sociales entre les opérateurs des données ouvertes (Table 3.37).

Je termine ce chapitre en proposant un modèle théorique selon lequel la construction des usages se subdivise en cinq étapes :

1. Utilisation des fonctionnalités de la plateformes
2. Evaluation du potentiel des données à partir de leur représentation
3. Transformation des données grâce à des artefacts techniques
4. Création d'un cadre cognitif partagé entre les systèmes pour les coordonner
5. Institutionnalisation de l'usage des données

Le modèle théorique et les barrières socio-techniques identifiées me serviront d'instrument de recherche dans le chapitre suivant pour réaliser une étude de cas comparative de deux écosystèmes open data. Cette étude visera à comprendre ce dont les usagers des données ouvertes ont besoin pour lever les barrières et socio-techniques et ainsi pouvoir construire de nouveaux usages.

⁶ Termes originaux : *Human-Computer Interaction*

Chapitre 4

**Les attentes des opérateurs de la donnée vis-à-vis
d'un dispositif en Systèmes d'Information**

Chapitre 4. Les attentes des opérateurs de la donnée vis à vis d'un dispositif en Systèmes d'Information

Ce quatrième chapitre a pour objet de comprendre les attentes des opérateurs de la donnée vis à vis d'un dispositif en Systèmes d'Informations permettant de faciliter l'usage de la donnée. Il s'inscrit dans le cycle de pertinence de cette recherche dans la mesure où il vise à interpréter les besoins du terrain de manière à pouvoir apporter une réponse scientifique pertinente. Aussi sera-t-il guidé par la deuxième sous-question de recherche :

Quelles sont les situations problématiques qui freinent la construction de nouveaux usages issus des données ouvertes ?

Je rappelle que dans le chapitre précédent, j'ai mené une étude de cas exploratoire à l'échelle du territoire français. Elle m'a permis de représenter le contexte (C) dans lequel se situe la situation problématique (Carlsson, 2010) adressée par cette thèse. L'usage de la donnée est réalisé par des opérateurs de la donnée qui s'inscrit dans le cadre d'écosystèmes *open data*. À travers des entretiens semi-directifs, des observations participantes et des analyses de documents, j'ai établi une heuristique de ces écosystèmes. Cette représentation décrit les différents sous-systèmes qui composent un écosystème, les différentes phases de son évolution et les mécanismes qui les régissent. Aussi, cette étude de cas exploratoire a permis d'établir les deux unités d'analyse pertinentes pour avoir une compréhension à la fois globale et détaillée de la situation problématique :

- les interactions homme-machine (human computer interaction) entre les opérateurs de la donnée et les technologies qu'ils utilisent et ce lors des différentes phases de l'usage des données
- les interactions entre les agents des différents sous-systèmes nécessaires lors de l'usage des données

Métaphoriquement, ces deux éléments constituent la carte de la situation problématique. L'heuristique est à la situation problématique ce que la topographie est à une carte : une représentation normée de ces différentes aires et reliefs. Aussi, les deux unités d'analyse sont à la situation problématique ce qu'une échelle est à une carte : un moyen de mesure permettant de décrire des phénomènes en deux dimensions.

Or, la carte n'est pas le territoire (Korzybski, 1998) et ce chapitre vise à construire le second sur la base de la première. Pour cela, il est question de réaliser une étude de cas comparative des écosystèmes *open data* de la métropole Grenoble-Alpes et du Grand Lyon. Comme l'indique l'architecture générale de cette thèse, elle a pour but premier, dans chaque itération, de construire les besoins fonctionnels auxquels un dispositif en Systèmes d'information doit répondre pour passer de la situation problématique à la situation désirée.

4.1 Sélection et présentation des cas

4.1.1 Sélection des écosystèmes de la métropole Grenoble Alpes et de la métropole du Grand Lyon

Dans cette étude de cas, nous avons sélectionné les écosystèmes open data des métropoles Grenoble-Alpes et du Grand Lyon. Cette sélection de cas se justifie, au-delà de leur facilité d'accès pour l'équipe de recherche, par le fait que ces deux écosystèmes possèdent un double avantage.

D'un côté, ils sont suffisamment mûrs pour que je puisse y observer des phénomènes d'usage des données. De l'autre, ils sont suffisamment hétérogènes en termes de géométrie, de mode de gestion et d'orientation, pour que les résultats communs aux deux cas puissent prétendre à une certaine validité externe.

4.1.2 Présentation des cas

Avant de présenter les cas de Lyon et de Grenoble, il faudra commencer par préciser les unités d'analyse choisies pour les analyser.

4.1.2.1 Cas 1 - L'usage des données ouvertes dans la métropole grenobloise

La première étude de cas concerne l'usage des données ouvertes dans la métropole grenobloise. La figure ci-dessous précise les frontières de cette étude de cas et clarifie les unités d'analyse choisies.

En ce qui concerne les unités d'analyse, une approche holistique qui permet d'étudier les interactions complexes entre différents niveaux d'analyse, sera mobilisée (Zuiderwijk, 2015). Ce choix se justifie par le fait que les besoins fonctionnels qu'il s'agit d'identifier dans cette étude de cas viennent de ces interactions complexes. La figure illustre les deux interactions complexes sur lesquelles il s'agira de se concentrer : 1. Les interactions homme-machine (*human computer interaction*) au sein de chaque système de l'écosystème *open data* grenoblois. Dans le chapitre précédent, il a été établi que des barrières sociaux-techniques apparaissent à ce niveau. En effet, ces dernières bloquent les séquences d'articulation qui combinent des artefacts techniques et des artefacts sociaux et, de fait, l'usage des données. 2. Les interactions sociales entre les différents systèmes de l'écosystème grenoblois. De la

même manière, le chapitre précédent conclut à l'apparition de barrières socio-techniques à ce niveau d'analyse.

Les frontières de cette étude de cas sont fixées par :

- la production des données ouvertes par la métropole grenobloise
- leur publication et leur édition sur la plateforme `data.metropolegrenoble.fr` par le groupe de travail *Open data*. Précisons que ce groupe est constitué d'une multitude d'acteurs publics et privés constituant l'écosystème *open data*.
- leur réutilisation par des acteurs au sein et en dehors de la métropole grenobloise.

Ainsi, l'usage des données est étudié à travers une approche holistique qui englobe la production, leur publication et leur édition ainsi que leur réutilisation au sein et en dehors de l'organisation à l'origine de la démarche d'ouverture des données.

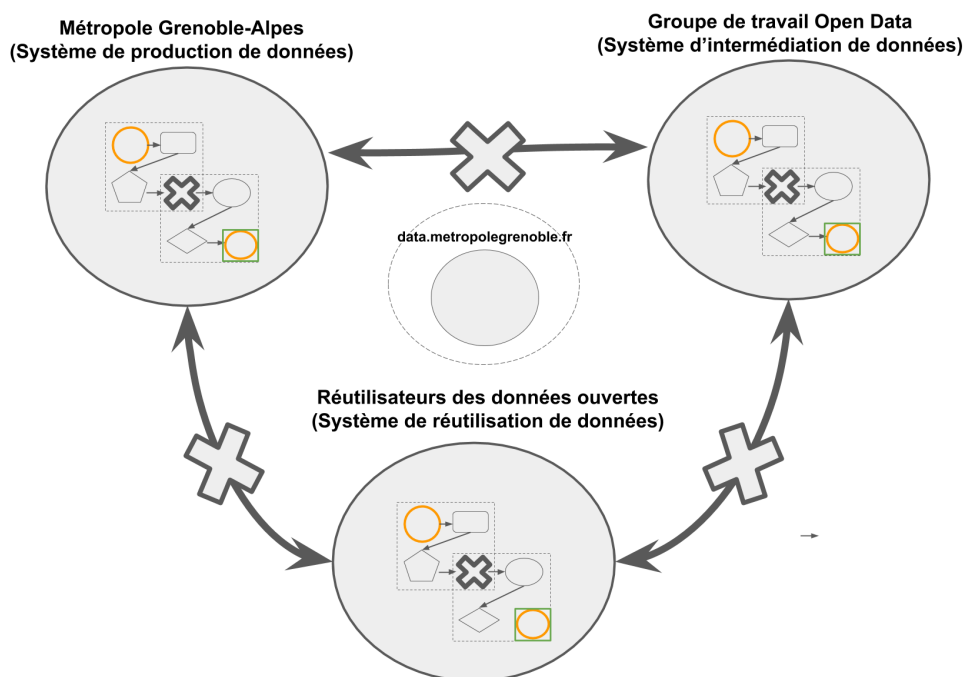


Figure 4.1 : Illustration des frontières et unités d'analyse choisie pour le cas de Grenoble

Le producteur de données choisi dans cette étude de cas est la métropole Grenoble-Alpes. C'est une métropole rassemblant 49 communes et plus de 600 000 habitants, connue comme étant un centre majeur de l'innovation et de l'économie de la connaissance et figurant régulièrement aux premières places dans les classements internationaux en la matière. Elle a été créée officiellement suite à la loi

MAPTAM le 1er janvier 2015 et possède de nombreuses compétences dont le développement économique, l'aménagement de l'espace métropolitain, l'énergie ou la culture. Elle emploie fin 2016 plus de 1700 agents (Grenoble-Alpes Métropole, 2016) répartis en 6 directions générales administratives : finances et gestion, cohérence territoriale, cohésion sociale et urbaine, services techniques métropolitains, cohésion sociale et urbaine et développement et attractivité.

La démarche d'ouverture des données de la métropole remonte à une délibération de la ville de Grenoble du 27 octobre 2014 dans laquelle elle s'engage à : * Mener une démarche d'ouverture des données commune avec la ville et le Syndicat Mixte de Transports en Commun (SMTC) * Mettre en oeuvre un portail *open data* à partir duquel toutes les données et toutes les applications et tous les services disponibles développés sont disponibles * S'appuyer sur une stratégie d'opportunité en publiant les données déjà fiabilisées et en lien avec les projets en cours * Constituer et participer à un comité de pilotage * Créer un lien fort avec les acteurs de l'écosystème, associatifs, entreprises, chercheurs et citoyens.

La métropole Grenoble-Alpes se lance dans une démarche d'ouverture des données pour décroiser l'information entre les services, entre les collectivités et entre les collectivités et la société et ce dans le but d'augmenter l'efficacité administrative et améliorer la transparence de l'action publique. Aussi vise-t-elle à favoriser l'émergence de nouveaux usages et l'amélioration de services existants grâce aux applications nouvelles qui pourraient être imaginées grâce à la réutilisation des données.

L'augmentation de l'efficacité administrative et de la transparence peut apparaître de la façon suivante. L'ouverture des données peut permettre d'économiser le temps passé par les services et les collectivités à s'échanger des informations et des données entre elles ou avec les citoyens. En effet, en publiant les données sur une seule plateforme, il n'y a plus besoin de s'échanger les données. Quant à l'émergence de nouveaux usages, elle peut apparaître grâce aux initiatives spontanées d'acteurs privés ou de citoyens qui décident de s'emparer des données pour créer de nouveaux services d'intérêt général.

La métropole Grenoble-Alpes lance la plateforme data.metropolegrenoble.fr en mars 2016. En 2020, 93 jeux de données sont publiés, classés en 6 groupes : aménagement du territoire, transport et mobilité, solidarité et citoyenneté, éducation en enfance, culture, environnement. Ces données proviennent de 7 organisations incluant les membres fondateurs de la démarche et des villes de la métropole. Cette plateforme permet également de demander des données à la métropole, de suivre les actualités de la démarches et d'accéder aux applications et services créés avec elles. Toutes les

communes et tous les services de la métropole peuvent publier leurs données sur ce portail. Les données de ce portail sont moissonnées par le portail national data.gouv.fr.

La plateforme data.metropolegrenoble.fr est principalement utilisée en interne, pour permettre aux services de la métropole de communiquer entre eux. Aussi, est-elle utilisée tous les ans dans le cadre d'un laboratoire d'innovation publique dit "CivicLab", visant à créer des services pour la métropole à partir des données. En 2020, la plateforme recense plus de quinze réutilisations émanant de développeurs indépendants, de collectifs citoyens ou d'éditeurs de logiciel. Par elles, une application permettant de calculer des itinéraires de déplacement au sein de Grenoble ou une cartographie des arbres de Grenoble.

4.1.2.2 Cas 2 - L'usage des données ouvertes dans la métropole du Grand Lyon

La deuxième étude de cas se concentre sur l'usage des données ouvertes dans la métropole du Grand Lyon. La figure ci-dessous précise les frontières de cette étude de cas et clarifie les unités d'analyse choisies. Comme pour la première étude de cas, une approche holistique qui permet d'étudier les interactions complexes entre différents niveaux d'analyse, sera mobilisée (Zuiderwijk, 2015). Les unités d'analyse choisies sont les mêmes que celles de la première étude de cas.

Les frontières de cette étude de cas sont fixées par :

- la production des données ouvertes par la métropole du Grand Lyon
- leur publication et leur édition sur la plateforme data.grandlyon.com par la Chief Data Officer (CDO), la Direction des Systèmes d'Information (DSI) et le laboratoire d'innovation Tubà.
- leur réutilisation par des acteurs au sein et en dehors de la métropole du Grand Lyon

Comme pour la première étude de cas, l'usage des données est étudié à travers une approche holistique qui englobe la production, leur publication et leur édition ainsi que leur réutilisation au sein et en dehors de l'organisation à l'origine de la démarche d'ouverture des données.

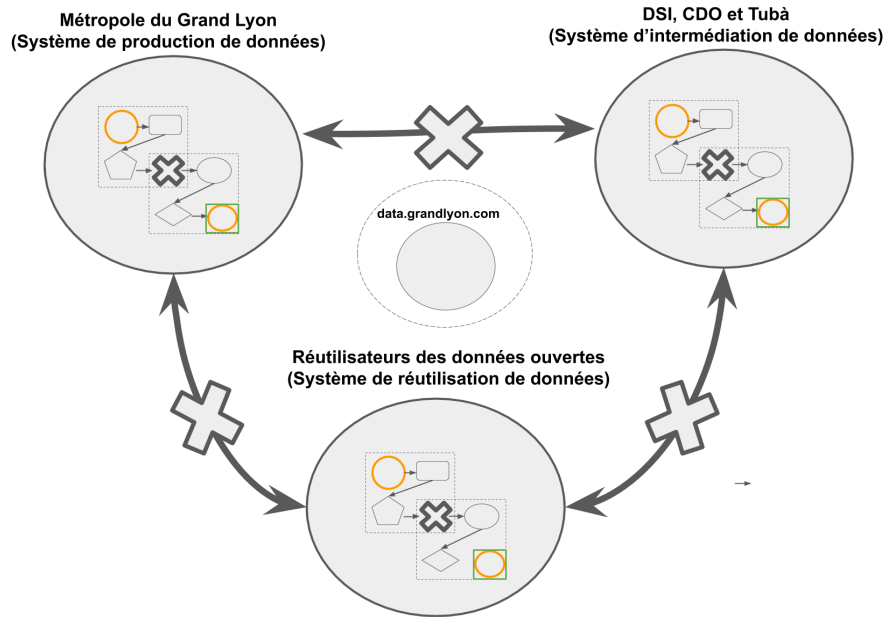


Figure 4.2 : Illustration des frontières et unités d'analyse choisie pour le cas du Grand Lyon

Le producteur de données choisi dans cette étude de cas est la métropole du Grand Lyon. C'est une métropole qui rassemble 59 communes et plus de 1,4 millions d'habitants. Elle a pour particularité administrative d'exercer à la fois les compétences d'un département (celui du Rhône) et d'une métropole. Le Grand Lyon est connu pour avoir été une métropole pionnière dans l'avènement de la *smart city*. Une avance qu'elle confirme chaque année à travers son classement au sein du Smart City Index : 23ème place en 2019 et 1ère ville française (IMD et SUTD, 2019). Elle a été créée officiellement le 1er janvier 2015 en tant que collectivité à statut particulier. Elle exerce les mêmes compétences que les métropoles auxquelles s'ajoutent celles du département du Rhône à savoir l'insertion, les personnes âgées, la mobilité, la voirie ou encore l'agriculture. Elle emploie en 2020 plus de 9200 agents (Grand Lyon, 2020) répartis en 5 directions générales : le développement solidaire, l'habitat et l'éducation ; le développement économique, emploi et savoirs ; le développement urbain et cadre de vie ; territoires et partenariats ; et les ressources.

La démarche d'ouverture des données de la métropole du Grand Lyon remonte à 2012 et s'inscrit comme une réponse aux nouvelles contraintes à laquelle les aires urbaines doivent faire face. Elle est intrinsèquement liée à la création de la *smart city* qui doit trouver de nouveaux modes de gestion de

l'énergie, de l'environnement, du logement ou de l'emploi dans un contexte d'intensification de la concentration urbaine (Larroche et Vila, 2015).

À travers cette démarche, elle s'engage à :

- Ouvrir sa plateforme à l'ensemble des acteurs du territoire. Ainsi, il s'agit d'une plateforme mutualisée
- Mettre à disposition les données dans des formats ouverts et normalisés afin de faciliter leur interopérabilité, leur croisement et leur enrichissement
- Géoréférencer les données et donner un accès via des webservices qui garantissent leur mise à jour automatique et ce, afin de développer des applications
- Appliquer un contrôle sur l'utilisation des données en appliquant plusieurs modalités de diffusion : une diffusion sans contrôle de l'utilisation des données, une diffusion avec authentification de l'utilisateur et contrôle de l'usage, une diffusion avec authentification de l'utilisateur, contrôle de l'usage et paiement des redevances

La métropole du Grand Lyon se lance dans l'ouverture des données en premier lieu pour faciliter l'échange des données entre les acteurs du territoire. Aussi vise-t-elle à valoriser le potentiel économique des données publiques. Enfin, son objectif est d'encourager la participation des citoyens.

Les échanges entre les acteurs d'un même territoire peuvent être facilités grâce à la plateforme et ses gestionnaires qui s'assurent de publier des données fiables et utiles aux acteurs souhaitant mettre en place des actions sur leur territoire. La valorisation des données peut advenir grâce aux expérimentations et aux innovations pilotées par le laboratoire d'innovation Tuba. Ce dernier est chargé de créer de nouveaux services à partir des données. Enfin, la Grand Lyon espère favoriser la participation des citoyens en publiant des données sur l'action publique. Elles peuvent permettre aux citoyens de s'informer et de questionner en retour les collectivités sur leur gestion des fonds publics.

La plateforme data.grandlyon.com est lancée en mai 2013. En 2020, 770 jeux de données sont publiés, classés en 13 groupes : localisation, occupation du sol, imagerie, services, environnement, urbanisme, transport, équipements, citoyenneté, limites administratives, culture, économie et accessibilité. Ces données proviennent de 31 organisations incluant les membres fondateurs de la démarche, les villes de la métropole mais aussi des entreprises comme JCDecaux, Indigo ou BlaBlaBus, ou des chambres de métiers et de l'artisanat. Cette plateforme met également à disposition des services qui facilitent l'exploitation des données ouvertes comme un géocodeur ou un fond de carte permettant de

cartographier les données. Enfin, elle donne accès à toutes les applications qui réutilisent les jeux de données ouvertes.

La plateforme data.grandlyon.com est utilisée à trois égards (Larroche et Vila, 2015). En premier lieu, elle agrège les données hétérogènes provenant des 31 acteurs hétérogènes, les consolide pour les rendre plus facilement manipulables. Ensuite, elle permet aux données d'être utilisées par des acteurs extérieurs dans le but de fournir un service potentiellement marchand. La laboratoire d'innovation Tuba est officiellement chargée de stimuler cette utilisation de la plateforme *open data*. La plateforme recense 7 applications utilisant les données ouvertes sans compter les nombreux projets menés par le Tuba tout au long de l'année. Parmi elles, on peut citer Handisco, une canne blanche connectée qui utilise les données de géolocalisation des arrêts de bus pour guider les personnes malvoyantes ou aveugles.

4.2 Construction du protocole et instruments de l'étude de cas

Résumé

Après avoir présenté les cas des métropoles grenobloise et lyonnaise, il s'agira de construire dans cette section le protocole que j'utiliserai pour les analyser. Par là, il est question, dans un premier temps, j'exposerai la manière dont j'ai abordé le terrain, les guides d'entretien utilisés ainsi que le cadre théorique. Dans un second temps, je présenterai les instruments et les méthodes utilisées pour collecter les données.

4.2.1 Approche du terrain et guides d'entretien

Dans ces études de cas, la collecte et l'analyse des données se succèdent et se chevauchent. Ces aller-retours sont particulièrement adaptés pour les recherches qui, comme le présent chapitre, visent à partir du terrain pour aller vers la théorie. Il est nécessaire de rappeler au lecteur qu'il n'est pas question dans cette thèse de tester des cadres théoriques existants puisqu'aucun d'entre eux n'émergent dans la littérature sur l'*open data* (Hassan et Twinomurinzi, 2017). Plutôt, cette thèse participe à la construction d'un cadre théorique visant à comprendre l'usage des données ouvertes.

Or, alterner collecte et analyse de données offre au chercheur une certaine flexibilité. Elle lui permet de faire rapidement des ajustements dans la collecte des données sur la base d'analyses préliminaires. Par là, le chercheur s'offre la possibilité de réduire progressivement ses champs d'investigation et de se concentrer de plus en plus profondément sur des thèmes spécifiques. Cette investigation profonde, comme l'écrit Eisenhardt, est propice à la construction d'un cadre théorique :

"Pour la recherche d'élaboration de théories, (modification des méthodes de collecte de données) est légitime parce que les enquêteurs essaient de comprendre chaque cas individuellement et de manière aussi approfondie que possible. Ainsi, si une nouvelle opportunité de collecte de données se présente et si une nouvelle ligne de pensée émerge au cours de la recherche, il est logique d'en tirer profit en modifiant la collecte de données, si une telle modification est susceptible de mieux fonder la théorie ou de fournir un nouvel éclairage théorique". - Eisenhardt (1984, p. 539)

Suivant cette logique, j'ai commencé par examiner les cas de Grenoble et de Lyon en établissant une liste de thèmes à aborder avec chacun des trois systèmes impliqués dans l'usage des données ouvertes. Ils seront synthétisés dans la table ci-dessous. Ces listes ont également servi de socle pour construire les guides d'entretien semi-directifs. Les données collectées suivant cette vaste liste ont été analysées durant la première partie de nos études de cas. Sur la base de ces analyses, il était opportun de restreindre la liste des thèmes à aborder.

Cette restriction s'est faite en suivant le cadre théorique des barrières socio-techniques freinant l'usage des données ouvertes, développé à partir de la littérature dans le chapitre 3. Ce framework a délimité les frontières de la deuxième partie de nos études de cas et prescrit les aspects des cas à investiguer plus en profondeur.

Le résultat de cette analyse préliminaire a été communiqué aux participants de l'étude. Ces communications ont été l'occasion de récolter des précisions et des retours sur la pertinence de ces analyses. Ces précisions et retours ont également été le moyen d'affiner la collecte et l'analyse des données. Finalement, les besoins fonctionnels qui résultent de ce chapitre ont été construits grâce à un processus itératif dans lequel l'ensemble des phases présentées ci-dessus se sont succédé plusieurs fois.

	Producteurs de données	Intermédiaire de données	Réutilisateur de données
Première liste étendue	<p>Contexte :</p> <ul style="list-style-type: none"> - lois et règlements entourant et influençant l'accès et l'utilisation des données dans la ville - type d'organisation impliqués dans la collecte et la gestion des données ouvertes - écosystème autour des données ouvertes - volonté politique/Leadership dans la démarche de production des données - ressources et compétences techniques mises à disposition de la production de données ouvertes - contexte économique de production des données <p>Données produites :</p> <ul style="list-style-type: none"> - définition d'une donnée ouverte - dimensions techniques, légales et sociales des données ouvertes - jeux disponibles et jeux à ouvrir - procédure d'analyse de qualité des données <p>Utilisation des données produits</p> <ul style="list-style-type: none"> - nom et caractéristiques des utilisateurs - but des utilisations <p>Impact des données produites</p> <ul style="list-style-type: none"> - bénéfices sociaux - bénéfices environnementaux - optimisation de la gouvernance 	<p>Objectifs de la démarche open data</p> <ul style="list-style-type: none"> - ambition - dispositifs mis en place - indicateurs, méthodes et suivi de performances <p>Situation actuelle</p> <ul style="list-style-type: none"> - nombre de jeux de données - nombre de réutilisations de jeux de données <p>Actions pour réduire l'écart entre les objectifs et la situation actuelle</p> <ul style="list-style-type: none"> - animation de la communauté - appel à projet <p>Caractéristiques de la plateforme open data</p> <ul style="list-style-type: none"> - type de plateforme - documentation mise à disposition - fonctionnalités à disposition de l'utilisateur <p>Guide d'utilisation de la plateforme</p> <ul style="list-style-type: none"> - Prise en main de la plateforme par des acteurs externes - Phases de prise en main de la plateforme - Standard de la méthode de prise en main <p>Evaluation de la méthode de déploiement de la démarche open data</p> <ul style="list-style-type: none"> - Précision de la méthode - Evolutions de la méthode 	<p>Présentation de la réutilisation</p> <ul style="list-style-type: none"> - Origine de l'idée - Objectif de la réutilisation - Créateur de la réutilisation - Fonctionnalités <p>Vision et stratégie de la réutilisation</p> <ul style="list-style-type: none"> - Stratégie commerciale ? sociale ? - Articulation entre différents parties prenantes <p>Gouvernance de la donnée</p> <ul style="list-style-type: none"> - cadre d'usage de la donnée - actions autour de la donnée - collaborations autour de la donnée <p>Impact de la réutilisation</p> <ul style="list-style-type: none"> - Indicateurs de suivi de la réutilisation

Table 4.1 : thèmes abordés avec les personnes interrogées

Les démarches de collecte de données ont été initiées à partir des organisations produisant les jeux de données ouvertes, à savoir les métropoles de Grenoble et de Lyon. Ensuite, j'ai ensuite appliqué la procédure de la boule de neige en leur demandant de nous désigner les individus et organisations jouant un rôle dans le système d'intermédiation des données. De la même manière, il leur a été demandé de désigner d'autres individus et organisations jouant un rôle dans le système de transformation des données.

Cette procédure de collecte se justifie par le fait que l'usage des données ouvertes découle de la coordination des actions des producteurs de données, des intermédiaires qui éditent et publient la donnée et des réutilisateurs qui transforment la donnée en services. Par conséquent, en délimitant notre étude de cas en suivant ces trois systèmes, j'ai pu analyser un maximum d'aspects de l'usage des données ouvertes. Aussi ce choix se justifie par le fait qu'il permet d'étudier de manière systématique les barrières socio-techniques qui se dressent face à l'usage des données ouvertes. En basant cette étude sur les systèmes de production, d'intermédiation et de production, il est possible de délimiter clairement les frontières de l'étude de cas et de comparer les cas de Grenoble et de Lyon.

4.2.2 Guide pour la réalisation de rapports d'études de cas

Les rapports d'étude de cas ont été réalisés en suivant le framework des barrières socio-techniques freinent l'usage des données ouvertes tel que décrit dans le chapitre 3. Ce framework décrit les séquences d'articulation qu'il s'agit d'analyser dans les trois systèmes, ainsi que l'emplacement des barrières socio-techniques qu'il s'agit d'examiner. À savoir, les interactions entre les agents d'un système et les artefacts technologiques qu'ils utilisent (*human-computer interactions*) et les interactions sociales entre les agents des trois systèmes. Les besoins fonctionnels ont été construits sur la base des données récoltées sur les barrières qui apparaissent au niveau de ces deux types d'interactions.

4.2.3 Sources et méthodes de collecte de données

Il s'agit de combiner plusieurs sources et méthodes de collecte de données et ce afin de maximiser la validité et la fiabilité du construit (Yin, 2003). Ces données sont souvent intégrées en utilisant la triangulation. En l'espèce, j'ai basé l'étude de cas sur trois sources et méthodes de collecte de données.

Étude de cas	Documents étudiés	Entretiens semi-directifs	Observations participantes
Metropole Lyon	13	10	Pendant 10 rencontre
Metropole Grenoble	19	11	Pendant 13 rencontres

Figure 4.3 : La liste des sources de données

Je présenterai ci-dessous chacune des sources et méthodes de collecte :

- L'étude de documents : j'ai étudié des comptes rendus officiels des réunions de travail et d'évènements des métropoles de Grenoble et de Lyon, des rapports de recherche, des appels d'offre, des délibérations administratives et directives ministérielles, des présentations PowerPoint internes aux métropoles et des documents de travail interne à certains réutilisateurs. Une liste des documents étudiés est disponible en Annexe.
- Les entretiens semi-directifs : j'ai mené des entretiens auprès d'un échantillon équitablement constitué de personnes responsables de la production de données ouvertes, de personnes en charge de la publication et de l'édition de données ouvertes et enfin avec des réutilisateurs de données ouvertes. Les thèmes abordés au cours des entretiens dérivent directement de la liste présentée dans la figure ci-dessus (approche du terrain et guides d'entretien). Je me suis particulièrement attaché à centrer les entretiens autour des barrières opérationnelles rencontrées au quotidien. Pour éviter les biais d'interprétations basées sur les propos d'une seule personne, plusieurs personnes ont été interrogées dans chacun des systèmes, ces personnes occupant des positions différentes dans les organisations : élus locaux, responsables opérationnels, agents administratifs, porte-parole de communautés, étudiants, dirigeants d'entreprise. La figure ci-dessous résume le profil et la durée des entretiens semi-directifs menés.

#	Fonction	Ville	Rôle	Durée
1	Dirigeant d'entreprise	Lyon	Intermédiaire/Editeur de données	47 min
2	Agent de collectivité	Lyon	Intermédiaire/Editeur de données	82 min
3	Agent de collectivité	Lyon	Fournisseur & Intermédiaire/Editeur de données	59 min
4	Dirigeant d'entreprise	Grenoble	Réutilisateur d'open data	40 min
5	Agent de collectivité	Grenoble	Intermédiaire open data	67 min
6	Agent de collectivité	Grenoble	Fournisseur & Intermédiaire/Editeur de données	48 min
7	Agent de collectivité	Grenoble	Fournisseur & Intermédiaire/Editeur de données	61 min
8	Agent de collectivité	Grenoble	Intermédiaire open data	45 min
9	Dirigeant d'entreprise	Grenoble	Réutilisateur d'open data	78 min
10	Agent de collectivité	Grenoble	Intermédiaire/Editeur de données	100 min
11	Etudiant	Lyon	Réutilisateur d'open data	38 min
12	Dirigeant d'entreprise	Grenoble	Utilisateur d'open data	46 min
13	Agent de collectivité	Lyon	Open data provider	100 min
14	Designer	Lyon	Intermédiaire/Editeur de données	40 min
15	Designer	Lyon	Intermédiaire/Editeur de données	37 min
16	Dirigeant d'entreprise	Grenoble	Réutilisateur d'open data	59 min
17	Elue	Grenoble	Intermédiaire/Editeur de données	68 min

18	Agent de collectivité	Grenoble	Fournisseur d'open data	76 min
19	Dirigeant d'entreprise	Lyon	Réutilisateur d'open data	56 min
20	Dirigeant d'entreprise	Lyon	Réutilisateur d'open data	39 min
21	Dirigeant d'entreprise	Lyon	Réutilisateur d'open data	70 min
			Total	1256 min

Figure 4.4 : liste des entretiens semi-directifs conduits

- Les observations participantes : j'ai observé en tant qu'observateur et en tant que participant direct à certaines activités relatives à l'usage des données ouvertes dans la métropole Grenoble-Alpes et la métropole du Grand Lyon. J'ai plus particulièrement assisté à des réunions de travail autour de l'ouverture des données, des ateliers de réutilisation des données ouvertes, des discussions avec des réutilisateurs autour de la conception d'un service, ainsi qu'avec des producteurs autour de la sensibilisation aux enjeux des données ouvertes. Les notes prises durant ces observations ont été étudiées.

4.3 Analyse des données : codage des entretiens en utilisant un système permettant d'identifier les barrières socio-techniques et les moyens imaginés pour les lever

Résumé

Dans cette section, les cas des métropoles Grenoble-Alpes et du Grand Lyon ont été longuement décrites. Je décris dans chacun de ces cas les différentes étapes de l'usage de données, telles que décrites dans le framework construit dans le chapitre précédent, ainsi que les barrières que les barrières rencontrées par les praticiens. Sur la base de ces descriptions je m'efforcerai d'identifier dans la prochaine section les besoins fonctionnels qu'un dispositif en système d'information doit combler pour lever ces barrières.

4.3.1 Utilisation des fonctionnalités de la plateformes (téléchargement/flux de données)

Il s'agit de définir comme fonctionnalités de la plateforme *open data* toutes les modalités permettant à un utilisateur de la plateforme (qu'il soit producteur, intermédiaire ou réutilisateur) d'y ajouter des données ou d'accéder aux données qu'elle contient déjà.

Pour les producteurs de données, les plateformes des deux cas leur offrent des espaces qu'ils peuvent utiliser de manière autonome pour ajouter des données. Ces espaces ont été conçus de manière à faciliter au maximum le travail d'ajout des données sur la plateforme et sont perçus comme simples d'utilisation.

“(Pour ajouter des données) c'est du suivi de procédure et ça se fait sur le Web donc c'est pas compliqué. Il suffit de briefer les agents et, normalement, tous les agents peuvent le faire, dont les agents de la communication qui n'ont pas de formation technique” (Entretien 7)

Ainsi, lorsqu'un agent producteur de données ajoute des données sur la plateforme, il ne doit pas modifier ses habitudes ni acquérir de nouvelles compétences. Il doit simplement suivre les instructions. À ce propos, les concepteurs des espaces précisent :

“L'objectif de cet espace, c'est d'offrir (au producteur) un processus qui soit le plus simple possible (...) et qui désilote les données mais qui ne modifie pas le mode de gestion de l'expert qui se trouve en face de l'outil” (Entretien 13)

Toutefois, une difficulté freine l'ajout de données : *“renseigner correctement la donnée si ce n'est pas le cas dès l'origine” (Entretien 7)*

Le fait de renseigner correctement la donnée implique de fournir des informations sur les modalités de création de la donnée comme les outils utilisés, la date de production ou la fréquence de mise à jour. Ces informations sont appelées *métadonnées*. Or, ces informations sont rarement considérées comme importante par le producteur qui doit donc investir du temps pour les obtenir :

“il y avait des informations mais diffuses, par exemple sur les données des bureaux de vote et les accès PMR. Sait-on quand le recensement des accès PMR a été effectué ? Il se trouve que oui. Mais la réponse au début n'était pas intuitive. Enfin, à mon niveau je ne m'étais pas posé la question et mes collègues, au niveau des services techniques ou au niveau des services

population non plus. Ils n'avaient pas recensé l'information en se disant qu'elle était utile à l'extérieur" (Entretien 3)

Conclusion et transition

Finally, from the point of view of producers, the interfaces fulfill their conditions of interaction (simplicity, familiarity and didactics) and the functionalities of the platform are effectively used. If difficulties there are, they come rather from the difficulty to provide mandatory information. On the other hand, from the point of view of reusers, the use of the platform proves to be more difficult. In both cases, I observe a certain mistrust vis-à-vis the data present on the platform and the modalities of access and use of these.

La méfiance vis-à-vis des données provient du fait que les réutilisateurs ont conscience que les données affichées sur la plateforme n'ont pas été produites à des fins d'ouverture et de réutilisation :

“Les données sur la plateforme n'ont pas été produites pour la plateforme. Elles ont été produites en vue de faire quelque chose. Par exemple, les données de capteurs de trafic, c'est en vue de faire de la régulation. Les données emplacement de marchés, les travaux c'est en vue de faire de l'information voyageurs. Toute donnée produite qui est mise en ligne sur la plateforme n'est pas produite pour la plateforme mais dans un autre but. Pour bien interpréter les forces et les limites de cette donnée, il faut comprendre le point de vue du producteur original” (Document 1L).

Ainsi, les réutilisateurs ont conscience que les plateformes n'offrent pas de données brutes et donc qu'il est assez dangereux de se baser entièrement sur elles.

La méfiance vis-à-vis des modalités d'accès provient d'un certain flou sur les licences d'utilisation des données. Pour la plateforme du Grand Lyon, la méfiance vient du fait que la plateforme présente deux licences atypiques en France : la licence “engagée” et la licence “associée”, la première donne un droit de regard de la métropole sur la réutilisation des données et la deuxième dernière prévoyant le paiement d'une redevance pour l'accès à certaines données par l'intermédiaires d'une API (Application Programming Interface). Un développeur explique :

« Une API est une surcouche qui donne des services en plus qui sont les bienvenus, mais du coup il y a des restrictions... Une API, c'est bien pour un développeur qui veut tester des usages, mais si l'on veut

faire une entreprise derrière, il faut faire attention parce qu'il peut y avoir des restrictions »
(Document 1L)

Pour la plateforme de la métropole grenobloise, la méfiance porte sur la licence associée aux données transports, à savoir la licence ODBL. Son article 4.4 spécifie que les réutilisateurs avaient :

“l'obligation de repartager l'ensemble de la base dans les mêmes conditions” (Document 1G).

De fait, tout réutilisateur se trouvait dans l'obligation d'ouvrir sa base de données s'il souhaitait utiliser les données transports. Cette clause a été perçue comme inconfortable par certains réutilisateurs qui, par conséquent, préféraient ne pas utiliser la plateforme et ses données.

Conclusion et transition

Enfin, du point de vue des producteurs, les plateformes ne remplissent pas certaines conditions d'interaction comme l'absence de restriction d'usage ou la garantie du caractère brut des données. Dans ce cadre, l'utilisation des fonctionnalités de la plateforme sont freinées.

Pour certains, ces manquements contribuent à faire des plateformes des “vitrine de données” (Entretien 20) et non des points d'accès direct à ces dernières. A Lyon, le représentant d'une start-up nous a expliqué que :

“dans (ses) missions, on se sert des plateformes open data pour pointer à nos clients qu'il existe des données détenues par des acteurs publics et qu'il serait intéressant pour eux et pour nous de les démarcher pour obtenir des données. Les plateformes open data, pour nous, c'est la vitrine des données, surement plus complexes et techniques, qui sont dans l'arrière cuisine et qui pourrait avoir une réelle valeur dans un cadre industriel” (Entretien 20).

À Grenoble, un intermédiaire en charge de la qualité des données géographiques précise que les plateformes *open data* sont des outils de :

“marketing territorial, très important pour afficher une image puis ensuite tu grattes derrière et tu vois si il y a vraiment de la donnée” (Entretien 7).

Le tableau ci-dessous résume les barrières rencontrées dans cette première étape d’usage des données ouvertes.

Etape d’usage	Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon
Utilisation des fonctionnalités de la plateforme (téléchargement/flux de données)	Manque de confiance dans la plateforme	Absence de garantie sur la caractère brut des données
		Manque d’informations sur la production des données (métadonnées)
		Doute sur les restrictions d’usage aposées aux données par les licences (restriction économique, conditions de réutilisation)
	Plateforme disponible dans une seule langue	
	Manque de lisibilité des interfaces	Plateforme perçue comme des vitrines et non comme un point d’accès aux données

Table 4.2 : barrières freinant l’utilisation des fonctionnalités de la plateforme open data

4.3.2 Évaluation du potentiel des données à partir de leur représentation

Pour les producteurs de données, évaluer le potentiel d’une donnée s’avère très difficile. Lorsque ces acteurs sont chargés de produire des jeux de données qui seront ouverts, ils se trouvent en grande difficulté d’évaluer, *a priori*, le potentiel de ces données en termes de réutilisations futures. Pour Laurence Comparat, élue à la mairie de Grenoble en charge des questions d’ouverture des données et présidente d’Open Data France, cette difficulté à évaluer le potentiel est consubstantielle de l’essence de l’ouverture des données et touche aussi bien Lyon que Grenoble:

“Il y a un problème intrinsèque à l’ouverture des données, c’est qu’elles sont en libre accès. Donc vous n’avez pas l’obligation que les gens qui l’utilisent vous en parle. Il y a un côté bouteille à la mer qui fait que j’ouvre les données mais”so what ?“n qu’est-ce qu’elles produisent ? Donc la question pour toutes les collectivités c’est : est-ce qu’il y a moyen de court-circuiter cet effet bouteille à la mer ?” (Entretien 17).

Conclusion et transition

À cause de cet effet “bouteille à la mer”, les producteurs et intermédiaires de données peinent à identifier les acteurs qui réutilisent les données et donc, ne peuvent mesurer la valeur des données qu’elles libèrent.

Du point de vue des intermédiaires, ce potentiel est également difficile à évaluer pour plusieurs raisons. La première raison relève du format dans lequel ils reçoivent les données de la part des producteurs. Ces formats ne sont pas forcément lisibles par machine et donc ne sont pas intégrables dans des outils qui permettraient de mieux comprendre le contenu des données. À Grenoble, un agent explique :

“Parfois on reçoit des excels avec des couleurs ou avec d’autres choses qui ne sont pas du tout adaptées à une base de données. Donc il y a toujours un long travail de lissage voire de nettoyage avant de pouvoir l’intégrer dans une base de données”.

À Lyon, un agent dresse le même constat :

“souvent, les services m’envoient un fichier, et pas un tableur mais un tableau en présentation texte avec des listes de numéros. Derrière, il faut transformer les données pour les ouvrir” (Entretien 3)

Quand les formats sont lisibles par machine, les intermédiaires font face à une autre difficulté qui s’aggrave avec le temps : leur hétérogénéité et leur volumes croissants. Les intermédiaires se retrouvent avec des données de plus en plus nombreuses qui proviennent d’instruments de plus en plus divers et complexes : capteurs, compteurs communicant, *smartphones*, applications, logiciels.

Cette complexification entraîne un déséquilibre entre les données disponibles et la capacité des agents à les évaluer :

“On a un très haut potentiel de la production des données et une très faible capacité d'évaluation, c'est compliqué par le fait que les facteurs d'évaluation sont extrêmement nombreux et complexes” (Document 1L).

À Grenoble, un agent prédit que :

“ (les) données produites demain seront des agrégats de données personnelles, de données provenant d'objets connectés ou de capteurs. Pour les comprendre et les évaluer, il faudra alors expliciter tous les traitements de toutes ces données. L'essentiel du travail sera là” (Entretien 5)

Pour illustrer cette complexification, l'exemple des données transports est évocateur. Les données de transports sont désormais gérées par des capteurs présents sur chacun des véhicules et en temps réel :

“ce sont des données difficiles à gérer et à interpréter; ça implique de gros volumes de données, de connexions entre machine en temps réel, des contraintes de la chaîne technique” (Document 1L).

Pour pouvoir absorber cette complexification, les intermédiaires doivent perpétuellement se former :

“On aura beau avoir un référent données, il nous faut un référent donnée capable de monter en compétences sur la donnée. Quand on voit tous les volets “IoT”, “Big Data”, “Intelligence Artificielle”. Pour tout ce qui a trait au numérique, on ne peut pas dire qu'on est compétent à un instant t, c'est une approche itérative et nécessairement continue parce qu'on a des standards qui évoluent, des types de données nouveaux”(Entretien 13).

Or, les collectivités ne sont pas forcément armées de profils en mesure de monter perpétuellement en compétences :

“ Là, on a un problème de compétences dans nos collectivités. L'open data c'est nouveau, les gens ne savent pas ce que c'est et ils n'ont pas les réflexes de montée en compétences, et ils ne savent pas ce qui est de l'ordre du normal ou pas dans ce que racontent les éditeurs de logiciel. À quel point est-on armé quand on est une collectivité avec des informaticiens qui sont surtout des gestionnaires de parc et pas des spécialistes des bases de données. On n'est pas armé. Il y a un enjeu de montée en compétence des agents et des élus. Si on ne forme pas nos agents et agentes, ça pose problème.” (Entretien 17)

Conclusion et transition

Finalement, pour les intermédiaires, l'évaluation s'avère difficile du fait de la complexification constante des données ou du fait qu'elles ne se présentent sous un format lisible par machine. En revanche, pour les réutilisateurs, cette évaluation s'avère facile à réaliser, et sans appel : les données ne sont pas de qualité suffisante pour être réutilisées.

Du point de vue des réutilisateurs, et du système de transformation, l'évaluation des données apparaît comme une formalité et leur avis est souvent tranché : les données ne sont pas utilisables car de trop mauvaise qualité. À Lyon comme à Grenoble, de nombreuses données sont perçues comme inutiles dès le premier coup d'oeil :

“Quand tu consultes les données, tu t'imagines que tu vas pouvoir faire des choses super avec ça... ils t'annoncent qu'il y a 1043 jeux de données mais il n'y a rien de qualitatif là-dedans. Pas de données venant d'acteurs privés, pas de données sur la consommation énergétique à l'échelle d'un quartier, pas des données sur les poubelles ramassées. On a pas de données sur des événements précis avec des mesures précises et non-statiques. Ce n'est pas ce n'est pas avec des données pour savoir où sont les silos à verre qu'on va générer des millions d'euros.” (Entretien 14).

Les réutilisateurs évaluent également les données en fonction de leur compatibilité avec les données avec lesquelles ils souhaitent les croiser ou avec les systèmes dans lesquels ils souhaitent injecter les données. Or, il est courant que les données soient incompatibles pour plusieurs raisons. Un industriel lyonnais explique :

“nous ne croisons pas de l'open data avec de la donnée de systèmes industriels. La typologie, les systèmes de production de données, la fréquence de mise à jour, rien n'est pareil (...) Dans ce contexte, la donnée ouverte n'est pas encore prête à être valorisée” (Entretien 20).

Le tableau ci-dessous résume les barrières rencontrées dans cette deuxième étape d'usage des données ouvertes.

Étape d'usage	Barrières socio-techniques	Exemple - Cas de Lyon et Grenoble
Evaluation du potentiel des données à partir de leur représentation	Manque de connaissance du domaine relatif aux données	

	Manque de compétences et d'expertise technique	Format illisible par machine
		Complexification constante des données et accroissement du volume
		Manque d'expertise technique spécifique aux données
	Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation	Méconnaissance des réutilisateurs de données du fait de la libre utilisation des données (<i>Effet bouteille à la mer</i>)

Table 4.3 : barrières freinant l'évaluation du potentiel des données

4.3.3 Transformation des données grâce à des artefacts techniques

Du point de vue des intermédiaires, le choix des capacités d'action à mettre en œuvre pour transformer les données revient à choisir quel outil permettra, le plus efficacement possible, de mettre un maximum de données dans un format facilitant l'ouverture des données. Il peut s'agir d'un format répondant à un standard particulier ou d'un format qui permet de représenter les données sous une forme compréhensible par tous (carte ou visualisation).

Cette tâche s'avère difficile tant les données sont hétérogènes et non standardisées. Elles proviennent d'une myriade d'applications et de logiciels qui possèdent leur propre format de données, difficile à homogénéiser :

“Le problème peut être au niveau des standards pour interconnecter. Alors, ce que fait Open Data France est un premier pas, ou les standards que nous imposent l'Etat en termes de marchés publics est aussi autre chose comme les marchés” (entretien 13).

À Grenoble, le responsable de l'ouverture des données rappelle que :

“toute la difficulté, c'est que le spectre d'intervention, qui est tellement large. A Gières par exemple, il y avait 20-30 applications métiers et tous les excels à droite à gauche. Et si tu dis OK, je vais prendre en main un ETL et mettre les données au bon format et que tu t'en sers une fois tous les ans, ça va être compliqué parce que tu dois faire une remise en main complète de ton ETL” (Entretien 8).

Du point de vue des réutilisateurs, la difficulté réside dans la maintenance des artefacts techniques. Soit du fait de la non-mise à jour des données ou de l'évolution constante des formats et volumes, les

réutilisateurs doivent en permanence maintenir les artefacts qu'ils utilisent. Le responsable de l'ouverture des données de la métropole grenobloise a notamment évoqué un échange avec un groupement d'industriels qui affirmait que :

“les données (ouvertes) ne les intéressaient pas, tant elles ne sont pas de bonne qualité. Et puis sur certaines données, en consultant d'autres sources que les métropoles, ils se sont rendus compte que les données ne sont pas les mêmes d'une année sur l'autre. Selon eux il n'y a pas de fiabilité, ni de pérennité. A la limite, on est prêt à payer pour que vous puissiez fournir de la donnée de qualité” (Entretien 8).

Les producteurs et intermédiaires sont conscients de l'enjeu de la maintenance des données pour faciliter le travail des réutilisateurs. À Lyon comme à Grenoble, les responsables cherchent à mettre en oeuvre des processus à cet égard :

“à Grenoble, je préfère mettre peu de jeux de données en ligne mais des données dont on est sûr que le processus de libération est bon, qu'on peut mettre la donnée à jour parce qu'on a le processus et qu'on a une demande. On se concentre là-dessus” (Entretien 17).

À Lyon, cette nouvelle exigence est présente dès la production des données :

“Notre priorité est de garantir la qualité des données, qualité en continue, la structuration en continue. On va aller voir la direction de l'eau, de la voirie, des déchets. Et là, on veut être sûr de la qualité et de la pérennité du jeu de données. Et quand on a un expert devant nous, on veut qu'il nous assure de la qualité et de la pérennité du jeu de données. Et du coup, on peut aussi valoriser, en plus de la donnée, l'expertise métier qu'il y a derrière.”(Entretien 8)

Enfin, une autre difficulté réside dans l'application des réglementations relatives aux données, dont la plus importante, le Règlement pour la Protection des Données Personnelles (RGPD). Il est difficile pour les réutilisateurs de savoir quelles sont les dispositions réglementaires qu'ils doivent prendre en compte pour transformer les données dans la mesure où ils ne savent pas dans quelle catégorie elles se trouvent. Qu'est-ce qui donne le caractère personnel à des données ? le caractère public ou privé ? Ces questions s'avèrent complexe et peuvent bloquer l'utilisation des données :

“il y a des cadres mais c'est dans l'évolution que ça pose problème. On avait open data, maintenant RGPD, demain l'Internet des Objets avec les données personnelles...Et une grosse problématique liées à cette évolution, c'est que la frontière entre public et privé est floue. Aujourd'hui, on a des données de caractère d'intérêt général, qui sont captées par des acteurs privés. Mais on peut aussi avoir des données d'intérêt général, captées par les collectivités mais qui intéresse les acteurs privés.”(Entretien 13)

Le tableau ci-dessous résume les barrières rencontrées dans cette troisième étape d’usage des données ouvertes.

Étape d’usage	Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon
Transformation des données grâce à des artefacts techniques	Format de données illisible par machine	
	Mauvaise qualité des données	Données non-maintenues dans le temps
	Hétérogénéité des conditions d’utilisation des données	Complexité des dispositifs réglementaires
		Spectre des données à traiter trop large par rapport aux ressources disponibles

Table 4.4 : Barrières freinant la transformation des données grâce à des artefacts techniques

4.3.4 Création d’un cadre cognitif partagé entre les systèmes pour les coordonner

Un cadre cognitif partagé se définit comme un ensemble d’éléments permettant aux agents d’acquérir une compréhension commune d’un phénomène et des pratiques qui le supportent, à savoir dans notre étude : l’usage des données ouvertes.

Les métropoles de Grenoble et de Lyon apparaissent dans les données récoltées comme fortement impliquées dans la construction d’éléments de compréhension et de pratique commune. Dans la délibération à l’origine de la démarche d’ouverture de la métropole Grenoble-Alpes, on trouve mentionné l’importance de développer des “interactions avec le réseau Open Data France et les associations nationales et internationales sur l’Open data pour la définition de référentiels communs” (Document 2G). Par “référentiel commun”, il est ici fait référence aussi bien à des standards de données qu’à l’instauration de pratiques communes pour la production, l’édition et la transformation des données.

On trouve ce même engagement au sein de la métropole de Lyon qui souhaite faire de son territoire :

“ un archipel, où nous serions au centre pour créer des ponts entre les acteurs de la donnée, entre réutilisateurs, entre producteurs, entre producteurs et réutilisateurs et avec le reste de la sphère qui sont les associations, les chercheurs, le monde académique au niveau régional et national ” (Entretien 13).

Pour créer ce cadre cognitif, les deux métropoles se sont engagées dans un travail dit “d'évangélisation”, une connotation religieuse qui trouve son sens quand il est question de vouloir rassembler des acteurs autour d'une même conception des choses. À Grenoble, un acteur nous a livré que :

“l'open data, c'est en premier lieu des tâches de fonds, aller voir chaque commune pour que ça prenne doucement. D'où l'importance (des chefs de projets open data) pour soutenir, l'animer, aider les gens. Ols peuvent s'appuyer aussi sur Open Data France mais ils font déjà pas mal de travail d'évangélisation par eux-mêmes. Et après, c'est l'effet boule de neige qui va devenir intéressant” (Entretien 7).

À Lyon, cette évangélisation est portée par un laboratoire d'innovation urbaine, le Tuba. Au lieu d'évangélisation les membres de ce laboratoire préfère parler de “médiation” Un chargé d'expérimentation explique :

“c'est très important de faire de la médiation car on est sur des problématiques très complexes quand tu parles d'open data à un représentant d'un quartier de la ville de Lyon il est un peu perdu et il a besoin de médiation sur le sujet. On fait de la médiation pour se faire comprendre auprès des gens avec lesquels on va mener nos expérimentations sur les nouveaux services urbains à partir des données. Sans cette médiation on arrivera jamais à transformer la ville comme on veut le faire” (Entretien 14)

Conclusion et transition

À la suite de cette évangélisation, il est question pour les métropoles de fédérer les acteurs de son territoire autour de pratiques communes. C'est par la pratique qu'elles comptent favoriser le plus l'usage des données ouvertes.

Du côté des producteurs, les métropoles accompagnent toutes les communes de leur territoire pour les rendre autonomes. Pour cela, elles mettent en œuvre des process et bonnes pratiques. À Lyon sont organisées des réunions :

“sur la structuration des données, sur comment on définit une structure commune de données. C'est là qu'on interagit en direct avec les communes, avec la métropole. C'est là qu'on a pu échanger sur nos différents points de vue autour des données qu'on a tous. Sachant qu'on a certaines communes qui

veulent aller plus loin que la structuration. Ça dépend des moyens qu'on a pour arriver ensuite à produire des données".

À Grenoble, il est question d'organiser des ateliers :

"dans une commune pour leur demander quels sont les jeux de données qu'ils souhaitent libérer. On a listé les jeux de données à libérer puis on les accompagne derrière pour sortir les données en open data"
(Entretien 8)

Du côté des réutilisateurs, les métropoles financent l'animation de groupes de réutilisateurs. Cette animation passe notamment par l'organisation d'évènements au cours desquels elles accompagnent et aident à structurer des projets de réutilisation de données. Ces évènements sont tous les deux portés par des laboratoires d'innovation : la Péniche à Grenoble et le Tuba à Lyon. A Lyon, leur objectif est d'arriver à :

"écrire des protocoles pour mener des projets avec l'open data, ça peut être une méthodologie de projet mais ça peut être aussi un guide d'entretien ou des choses plus loufoques avec des méthodologies un peu plus innovantes" (Entretien 14)

Dans l'évangélisation et la mise en place de pratiques communes, tous les acteurs impliqués dans l'usage des données ouvertes font face à une difficulté : le manque d'un cadre stratégique clair. Les objectifs et les moyens alloués à l'ouverture, l'édition et la transformation ne sont pas clairs, pour ne pas dire inexistant :

"Oui on a des difficultés à articuler une stratégie mais on se met dans l'optique d'en créer une stratégie. Ce qu'il faut rappeler c'est qu'entre 2015 et 2017, ils ont fait comme ils ont pu. Et ce n'est que depuis mon retour, on s'est dit qu'on allait cadrer les choses, se donner une ligne de conduite pour savoir ce qu'on fait"
(Entretien 8).

À Lyon, si ce cadre existe par l'intermédiaire d'un document cadre sur la stratégie Ville intelligente de la métropole lyonnaise, c'est la mise en œuvre au-delà des déclarations d'intention stratégique qui se trouve bloqué, à l'image du Tuba. Alors que celui-ci a été créé en 2012 à la base pour faciliter la réutilisation et l'agrégation des données, ils n'ont :

"jamais vu un jeu de données transiter par chez (eux)" ; "Les partenaires ne voulaient pas partager des données ou alors sur des domaines très précis donc c'était du partage et non de l'ouverture. Finalement sur la démarche Open Data on est un peu frustré parce que les partenaires qui ont rejoint le programme ne l'ont pas fait pour ouvrir leur données mais ils l'ont plus fait dans une dynamique politique" (Entretien 14)

Le tableau ci-dessous résume les barrières rencontrées dans cette quatrième étape d'usage des données ouvertes.

Étape d'usage	Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	Manque d'un cadre juridique et politique clair	Manque d'un cadre stratégique clair <i>objectifs moyens alloués</i>
	Manque d'une vision cohérente autour de la production et de l'utilisation des données	Coût trop élevé du travail d'évangélisation par rapport aux ressources humaines disponibles
	Le classement des données ouvertes en bas de la liste des priorités	

Table 4.5 : barrières freinant la création d'un cadre cognitif partagé entre les systèmes

4.3.5 Institutionnalisation de l'usage des données

L'institutionnalisation est ici entendue comme " l'intégration de l'enjeu de gouvernance de la donnée au sein des organisations " (Courmont, 2015, p.328).

Les métropoles de Lyon et Grenoble ont toutes deux initié la mise en oeuvre, rencontrant de nombreuses difficultés :

"Au Grand Lyon, c'est assez compliqué. Il y a avait une forme de gouvernance des données sur les données géographiques, avec des instances de décisions et une certaine forme d'organisation. C'était en place depuis les années 80, 90. (...) plus récemment, c'est compliqué du fait des grandes reconfigurations au sein de l'institution : en 2015, la communauté urbaine de Lyon est devenue métropole et a fusionné avec le département du Rhône, ce qui a conduit à pas mal de réorganisations de l'administration. Tout le process autour de l'action et du numérique s'est retrouvé au milieu de ce chamboulement, avec les données" (Entretien 2).

À Grenoble, la mise en place d'une gouvernance des données au sein des organisations est en encore au stade de l'idée :

“Pour nous, une vraie barrière, c’est de savoir quelle organisation on met en oeuvre pour faire ça. Aujourd’hui, on fait juste comme on peut. Pour cela, il faut qu’on bénéficie d’un minimum d’accompagnement pour avoir apporter quelque chose en plus à l’écosystème” (Entretien 8).

Pour les producteurs, la première chose sur laquelle ils trébuchent est l’architecture des système d’information, souvent sectorialisée, qui rend difficile la production de données ouvertes :

“La mise en place d’une organisation dépend de la structuration du système d’information qui peut être très sectorialisée par métier ou beaucoup plus transverse, avec un avantage pour la deuxième. Ça dépend des organisations et ça joue beaucoup sur ce que tu peux faire avec les données” (Entretien 2).

En termes d’architecture, les producteurs et intermédiaires se heurtent également à la “démarche applicative” qui prévaut dans la façon de l’envisager. Ce mode de gestion des systèmes d’information veut qu’à chaque besoin de logiciel, il y ait la création d’une nouvelle base de donnée au lieu de récupérer les données déjà présentes :

“on a vraiment cette dimension qui est du au fait qu’on adjoint à chaque besoin une application. Et pour que les services se mettent en commun et créent une base commune, il faudra un certain temps. Et du coup, on se retrouve avec des agents et des prestataires qui ne gèrent que le logiciel et qui ne font pas le lien avec la donnée. Il faut comprendre que quand j’identifie un besoin, il est beaucoup plus facile de faire un cahier des charges en isolant tout ce qui a trait à ce logiciel parce que derrière les éditeurs vont répondre avec leurs outils propres sur chacun des points listés. Alors que si on créé une base de donnée communes et qu’on demande au prestataire de s’y connecter, ça ne va pas plaire et le prestataire va dire que c’est au risque et péril du commanditaire parce qu’il n’aura pas la main sur les données.” (Entretien 8)

Parmi les autres freins à l’institutionnalisation, les acteurs pointent l’incompatibilité des cultures organisationnelles avec l’usage des données ouvertes. Pour les organisations publiques, l’incompatibilité vient du fait que l’ouverture des données porte l’accent sur la transversalité, la mise en lumière des métiers et de leur expertise, ce qui :

“bouleverse l’organisation parce que vous désilotez. L’accès à la donnée est l’occasion d’encourager le partage et le savoir. On va décrire ensemble les méta-données on va décrire les données et le métier qu’elle rend visible. Il faut essayer de minimiser ce bouleversement sous peine de voir les projets rejetés” (Entretien 13).

Pour d'autres organisations, la difficulté réside dans le fait que l'usage des données ouvertes n'a pas encore fait ses preuves en termes de résultat concrets et à court-terme :

“De mon point de vue les entreprises freinent des quatre fers pour ne pas avoir à libérer leur données. Ils n'ont pas entendu parler de projet concret avec les données ouvertes donc ils ne les utiliseront et n'ouvriront pas les leurs non plus. A la limite, ils peuvent faire de l'Open Data washing en produisant quelques graphes ridicules” (Entretien 14).

Pour les petites collectivités, un constat similaire est établi :

“Et l'open data ne fait pas partie du mode de fonctionnement de la commune. Ça paraît logique que ça ne rentre pas dans les priorités. Le but d'une collectivité, c'est déjà de rendre des services à ses administrés au quotidien et pour les communes, la data n'est pas encore un enjeu crucial à ce niveau là. Alors que pour le travail des métropoles, c'est essentiel” (Entretien 6)

En termes de culture organisationnelle, la difficulté vient aussi du fait que l'ouverture des données, autant dans sa production que dans sa réutilisation est perçue comme comportant de nombreux risques. Or, certaines cultures organisationnelles présentent une aversion au risque.

En libérant des données, certaines organisations considèrent qu'elles prennent le risque de voir leurs dysfonctionnements mis à jour :

“prenons l'exemple des opérateurs dans le cadre d'une délégation du service public comme Vinci, Eiffage ou Véolia, etc. Ces acteurs-là n'ont pas du tout envie de publier ces données en OD parce qu'on peut observer ce qu'ils font par exemple si le tram a du retard tout le temps, ils n'aiment pas du tout donner de l'information qui risquerait de leur retomber sur la tête, parce que quelque part c'est donner à observer la qualité de service qu'ils devraient assurer” (Document 1L)

Encore, l'ouverture des données est-il perçu comme un facteur pouvant fragiliser les avantages concurrentiels des organisations et la maîtrise de leur stratégie :

“Pour des raisons de pouvoirs, certaines directions au sein de la métropole n'ont pas le même rapport à la diffusion. Et j'imagine que c'est pareil dans les autres organisations. Par exemple on peut avoir la direction de l'eau qui a sa propre stratégie en termes de délégation de services publics et qui ne voit pas l'intérêt de mettre à disposition ces données, car ils ne veulent pas perdre la maîtrise de leurs stratégies” (Entretien 2)

Le tableau ci-dessous résume les barrières rencontrées dans cette cinquième et dernière étape d'usage des données ouvertes.

Étape d'usage	Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon
Institutionnalisati on de l'usage des données	L'incompatibilité des structures organisationnelles avec l'ouverture des données	Incompatibilité de l'architecture des systèmes d'information avec l'institutionnalisation de la production
		Incompatibilité des données ouvertes avec la démarche de développement des système d'information dites "applicatives" (<i>un besoin, une application, une base de données</i>)
		Opposition entre la silotage organisationnel et la transversalité portée par l'ouverture des données
	Divergence culturelle de l'ouverture des données par rapport aux organisations	Opposition entre exigences de résultats à court-terme des organisations et les changements structurels portés par l'ouverture des données.
		Opposition entre culture organisationnelle averses aux risques et la culture opportuniste portée par l'ouverture des données.
		Opposition entre culture organisationnelle portée sur le secret des affaires et la culture de la transparence portée par l'ouverture des données.

Table 4.6 : barrières freinant l'institutionnalisation de l'usage des données

4.4 Résultats et discussion : un besoin d'avoir des connaissances partagées entre opérateurs pour lever plus facilement les barrières

Dans cette section, il est question de définir comment un dispositif en Systèmes d'Information devrait s'insérer pour faire évoluer les deux cas présentés dans la section précédente et lever les obstacles qui font face à l'usage des données ouvertes. En effet, il s'agira de se baser dans cette section sur les

barrières identifiées dans la section précédente pour pouvoir identifier les besoins fonctionnels adressés par un dispositif visant à lever ces barrières. Autrement dit, alors que dans la section précédente, il s'agissait de définir les barrières qui freinent les différentes phases de l'usage des données ouvertes, il sera question dans la section présente de les traduire en besoins fonctionnels concrets. Ces besoins seront adressés subséquentement par le design d'un dispositif en Systèmes d'Information

Les besoins fonctionnels sont les éléments qui définissent les difficultés qu'un système vient résoudre et qui sculptent comment les fonctionnalités d'un système peuvent être utilisées (Stellman et Greene, 2005). En identifiant ces besoins fonctionnels, nous répondons à un des objectifs principaux de cette étude, à savoir de comprendre quel type de dispositif en Systèmes d'Information sera perçu comme utilisable et fonctionnel pour lever les barrières qui font face à l'usage des données ouvertes. Plus précisément, nous répondons à la nécessité de découvrir ce que les acteurs attendent d'un dispositif en Systèmes d'Information

Cette section sera organisée en suivant le framework des barrières liées à l'usage des données ouvertes présenté dans le chapitre 3. Des besoins fonctionnels seront identifiés dans chacune des cinq phases de l'usage des données ouvertes. Ce choix a été fait pour permettre l'émergence d'un cadre directeur pour la formulation des besoins fonctionnels. Par exemple, la quatrième phase de l'usage des données ouvertes (à savoir la création d'un cadre cognitif partagé entre les systèmes pour les coordonner) suggère indirectement que le dispositif en Systèmes d'Information doit mettre à disposition des éléments de compréhension des données ouvertes qui puissent être partagés et utilisés par les différents systèmes.

4.4.1 Les besoins fonctionnels pour lever les obstacles dans l'utilisation des fonctionnalités d'une plateforme *open data*

Pour lever les obstacles dans l'utilisation des fonctionnalités d'une plateforme, il est principalement question pour un dispositif en systèmes d'information d'améliorer le niveau d'information fourni sur les modalités de production et de réutilisation de données présentes sur la plateforme *open data*. Je présenterai ci-dessous les besoins fonctionnels adressant le manque de confiance dans la plateforme et le manque de lisibilité des interfaces.

4.4.1.1 Manque de confiance dans la plateforme

Du point de vue des réutilisateurs, il a été dit précédemment que le manque de confiance trouvait son origine dans le manque d'information sur le contexte de production des données affichées sur la plateforme. Des informations que les producteurs ont du mal à faire remonter jusqu'à la plateforme. Ceci m'amène à formuler **le premier besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit supporter la documentation du processus de production de donnée**, de manière à ce que les producteurs et intermédiaires puissent afficher sur la plateforme des informations sur le contexte de production des données et les différents traitements qui lui ont été appliqués.

Les acteurs interrogés dans les deux cas ont noté que cette documentation était nécessaire pour améliorer la confiance que les réutilisateurs peuvent avoir sur la plateforme. Selon eux, il est essentiel que les réutilisateurs soient tenus au courant, via la plateforme, du "qui fait quoi" (Entretien 3). Autrement dit, il s'agit de leur fournir sur la plateforme des informations sur les instruments utilisés pour collecter la plateforme, les méthodes de traitement des données ou les expertises sollicitées, les instances ayant validé la diffusion de la donnée, ou tout autre information susceptible d'intéresser les futurs réutilisateurs.

De ce premier besoin fonctionnel découle le deuxième besoin fonctionnel dispositif en Systèmes d'Information, à savoir qu'il doit supporter la documentation de tous les mécanismes techniques (à l'image des API) ou juridique (à l'image des licences) qui fixent des restrictions sur les conditions d'utilisation des données. Ceci afin que les réutilisateurs soient fixés sur ce qu'il est possible de faire avec les données et sous quelles conditions. Bien que ce deuxième besoin fonctionnel ne soit qu'une déclinaison du premier, il paraissait important dans la mesure où il insiste sur la nécessité du dispositif en Systèmes d'Information d'améliorer le niveau d'information sur les deux côtés d'une plateforme *open data* : la production et la réutilisation des données.

4.4.1.2 Manque de lisibilité des interfaces

Les cas de Grenoble et de Lyon ont tous les deux démontré que la plateforme *open data* était avant tout perçue comme un outil marketing avant d'être approché comme un point d'accès direct aux données. Cette perception constitue un obstacle à l'utilisation des fonctionnalités principales de la plateforme, à savoir le téléchargement ou l'accès à des flux de données. Cette analyse nous amène à

formuler le troisième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit faciliter le design de plateformes *open data* à double entrée, la première étant réservée aux développeurs et experts de la données et la deuxième réservée au grand public. La première entrée offrirait des interfaces mettant en avant les modalités techniques d'accès et de réutilisation des données (avec leur documentation) et la deuxième des interfaces avec des informations plus générales sur la démarche d'ouverture des données.

Les réutilisateurs des deux cas ont mentionné l'utilité qu'aurait cette double entrée. Elle permettrait de susciter l'intérêt de profils d'experts de la donnée autour des plateformes mais également celui d'organisations qui se positionneraient dans le système des intermédiaires de la donnée. Par exemple, des agrégateurs de données pourraient s'intéresser aux plateformes *open data* comme nouvelle source de données, les nettoyer et les intégrer à leur patrimoine de données. Par là, l'usage des données serait maximisé.

Le tableau ci-dessous résume les besoins fonctionnels relatifs à l'utilisation des fonctionnalités d'une plateforme *open data*.

Barrières socio-techniques	Exemples de barrières dans les cas de Grenoble et Lyon	Besoins fonctionnels
Manque de confiance dans la plateforme	Absence de garantie sur la caractère brut des données	BF1. Le dispositif en système d'information doit supporter la documentation du processus de production de donnée
	Manque d'informations sur la production des données <i>métadonnées</i>	BF2. Le dispositif en système d'information doit supporter la documentation de tous les mécanismes techniques ou juridiques qui fixent des restrictions sur les conditions d'utilisation des données
	Doute sur les restrictions d'usage apposées aux données par les licenses (restriction économique, conditions de réutilisation)	
Manque de lisibilité des interfaces	Plateforme perçue comme des vitrines et non comme un point d'accès aux données	BF3. Le dispositif en système d'information doit faciliter le

		design de plateformes <i>open data</i> à double entrée, la première étant réservée aux développeurs et experts de la données et la deuxième réservée au grand public.
--	--	---

Figure 4.5: Les besoins fonctionnels relatifs à l'utilisation des fonctionnalités d'une plateforme *open data*.

4.4.2 Les besoins fonctionnels pour lever les obstacles dans l'évaluation du potentiel des données à partir de leur représentation

Pour lever les obstacles dans l'évaluation du potentiel des données à partir de leur représentation, il est principalement question pour un dispositif en systèmes d'information de centraliser les flux d'informations et de connaissances sur les possibilités de transformation des données. Je présente ci-dessous ses besoins fonctionnels adressant la manque de compétences et d'expertises techniques ainsi que la difficulté d'intégration de profils experts et techniques dans les processus d'utilisation.

4.4.2.1 Manque de compétences et d'expertise technique

À plusieurs reprises, les personnes interrogées dans la métropole de Grenoble et de Lyon ont fait part de la difficulté à évaluer des données transmises dans des formats n'étant pas lisibles par machine. Par exemple, les acteurs nous ont fait part de données transmises dans des fichiers Word ou des fichiers excels rendant impossible leur intégration dans des outils d'analyse de données. Ceci m'amène à formuler le quatrième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit faciliter la sensibilisation les acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre. Cette sensibilisation de base faciliterait la circulation des données et permettrait d'identifier rapidement les données présentant un fort potentiel de réutilisation.

De ce premier quatrième besoin fonctionnel découle le cinquième. En effet, dans les cas de Grenoble et de Lyon, les acteurs ont noté, qu'au-delà de la sensibilisation, que la montée en compétence perpétuelle était une nécessité pour être en capacité d'évaluer le potentiel des données. Ne pas avoir ce

réflexe constitue un réel blocage dans l'usage des données ouvertes. Ceci m'amène à formuler le cinquième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel. Ces flux de connaissances doivent être conçus de manière à atteindre les acteurs dans leur quotidien. Ils permettent ainsi d'apporter des réponses à tout individu qui souhaite évaluer le potentiel de nouvelles données et ainsi de faciliter sa prise de décision concernant l'utilisation des données ouvertes.

Les personnes interrogées à Grenoble et Lyon ont insisté sur l'opportunité de venir centraliser de nombreuses ressources déjà disponibles librement et fournissant des connaissances clés pour évaluer le potentiel des données. Ils nous ont précisé que depuis l'avènement de l'ouverture des données, un très grand nombre d'associations, de collectifs ou d'instituts de formation avaient produit des contenus de formation sur l'ouverture et les données. Cependant, personne ne s'était encore appliqué à les centraliser et à fournir des guides permettant aux acteurs de s'orienter dans un réel océan de connaissance.

4.4.2.2 Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation

Du fait du principe de libre usage des données, les producteurs et les intermédiaires ne peuvent pas identifier directement les réutilisateurs des données. Ceci pose une réelle difficulté à l'évaluation du potentiel des données dans la mesure où les producteurs et intermédiaires ne peuvent pas solliciter directement les feedbacks des réutilisateurs sur le potentiel des données ouvertes. Ceci m'amène à formuler le sixième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre. Ainsi, les producteurs pourraient demander l'avis des réutilisateurs sur le potentiel de certains jeux de données et inversement, les réutilisateurs pourraient faire part de leurs besoins en données.

Si les producteurs et intermédiaires lyonnais et grenoblois rencontrent de manière sporadique les réutilisateurs au cours d'événements, ils ont exprimé le besoin d'avoir des contacts plus réguliers avec eux. À cet égard, ils emploient la métaphore de l'écosystème perçu comme un organisme vivant dans lequel les différentes parties s'adaptent les unes aux autres. Cette métaphore nous a été présentée comme une représentation de l'idéal à atteindre en matière d'usage des données ouvertes.

Dans le tableau ci-dessous, sont résumés les besoins fonctionnels relatifs à l'évaluation du potentiel des données à partir de leur représentation.

Barrières socio-techniques	Exemples rencontrés dans les cas de Lyon et Grenoble	Besoins fonctionnels
Manque de compétences et d'expertise technique	Format illisible par machine	BF4. Le dispositif en système d'information doit faciliter la sensibilisation des acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre
	Complexification constante des données et accroissement du volume	BF5. Le dispositif en système d'information doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel.
	Manque d'expertise technique spécifique aux données	
Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation	Méconnaissance des réutilisateurs de données du fait de la libre utilisation des données\Effet bouteille à la mer	BF6. Le dispositif en système d'information doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre

Figure 4.6 : Les besoins fonctionnels relatifs à l'évaluation du potentiel des données à partir de leur représentation

4.4.3 Les besoins fonctionnels pour lever les obstacles dans la transformation des données grâce à des artefacts techniques

Pour lever les obstacles dans la transformation des données grâce à des artefacts techniques, il est principalement question pour un dispositif en systèmes d'information d'offrir des outils permettant de filtrer, trier, structurer et classer les données ouvertes disponibles selon des critères de transformabilité comme la maintenance dans le temps ou les dispositifs réglementaires encadrant la réutilisation.

Il s'agira de présenter ci-dessous les besoins fonctionnels adressant le manque de qualité des données et l'hétérogénéité des conditions d'utilisation de celles-ci.

4.4.3.1 Manque de qualité des données

Les réutilisateurs de données rencontrés se déclarent unanimement gênés dans leurs actions par le manque de fiabilité et de pérennité des données dans le temps. Ceci m'amène à formuler le septième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit offrir des outils permettant de comparer la qualité des données entre elles et surtout dans le temps. De fait, les producteurs et intermédiaires pourront prendre conscience de la dégradation des données dans le temps si tel est le cas ou inversement, valoriser l'expertise métier qui permet de maintenir la qualité dans le temps. Aussi, les réutilisateurs pourront adapter plus facilement les capacités d'action à mettre en œuvre pour transformer les données. Par exemple, ils pourront mettre en œuvre des dispositifs permettant de maintenir la qualité des données dans le temps comme des indicateurs de fiabilité à prendre en compte dans les transformations.

À Lyon comme à Grenoble, tous les acteurs souhaitent voir ce type d'outils apparaître. Pour les producteurs et intermédiaires, ce serait un moyen de valoriser à l'extérieur de l'organisation l'expertise métier qu'il y a derrière la donnée. Ces outils traçant la qualité permettraient de mettre à jour les compétences de la chaîne qui permet d'avoir des données mises à jour. Côté réutilisateur, l'enjeu de la maintenabilité des données est si important que certains seraient prêts à payer pour voir la qualité des données tracées et maintenues.

4.4.3.2 Hétérogénéité des conditions d'utilisation

Par ailleurs, les réutilisateurs comme les producteurs ont noté la difficulté de choisir les capacités d'action à mettre en œuvre tant le spectre de données disponibles est large et tant les dispositifs réglementaires qui s'appliquent peuvent s'avérer complexes. De cette analyse émerge le huitième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité. Ainsi, ces outils permettraient de réduire le volume d'information à traiter par les acteurs chargés de transformer les données. Également, il faciliterait la structuration de la transformation en fixant des points de repères pour les réutilisateurs.

À Lyon comme à Grenoble, les acteurs ont fait part de ce besoin de structurer le large volume d'information disponible sur la transformation des données ouvertes. Ils m'ont précisé que cette structuration doit être dynamique pour s'adapter au profil de la personne qui cherche ses informations mais aussi au rythme soutenu de la croissance du volume de données disponible et à la complexification de données disponibles. Par exemple, les acteurs anticipent que les données ouvertes provenant de l'Internet des Objets vont soulever une nouvelle vague de transformation de données. Avec cette vague, de nouvelles questions vont être soulevées par les producteurs, les intermédiaires et les réutilisateurs. Autant de nouvelles questions qu'il faudra prendre en compte pour adapter les outils de filtrage. Le tableau ci-dessous résume les besoins fonctionnels relatifs à la transformation des données grâce à des artefacts techniques.

Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon	Besoins fonctionnels
Mauvaise qualité des données	Données non-maintenues dans le temps	BF7. Le dispositif en système d'information, à savoir qu'il doit offrir des outils permettant de comparer la qualité des données entre elles et surtout dans le temps
Hétérogénéité des conditions d'utilisation des données	Complexité des dispositifs réglementaires	BF8. Le dispositif en système d'information doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité
	Spectre des données à traiter trop large par rapport aux ressources disponibles	

Figure 4.7 : besoins fonctionnels relatifs à la transformation des données grâce à des artefacts techniques.

4.4.4 Les besoins fonctionnels pour lever les obstacles dans la création d'un cadre cognitif partagé entre les systèmes

Pour lever les obstacles dans la création d'un cadre cognitif partagé entre les systèmes, il est en grande partie question pour un dispositif en systèmes d'information de permettre aux acteurs d'intégrer l'ouverture des données dans leur quotidien. Dans cette partie, je présenterai les besoins fonctionnels relatifs au manque d'un cadre juridique et politique clair ainsi qu'au manque d'une vision cohérente autour de la production et de l'utilisation des données.

4.4.4.1 Manque d'un cadre juridique et politique clair

À Lyon comme à Grenoble, certains acteurs interrogés m'ont fait part de la difficulté d'unifier la conception de l'ouverture des données sans cadre stratégique clair qui précise les objectifs de l'usage des données ouvertes et les moyens alloués pour les atteindre. Dans ce contexte, la citation de Sénèque s'applique : "Il n'y a point de vent favorable pour qui ne sait où il va". À maints égards, la production ou la transformation des données ouvertes reste circonscrite à des actions isolées, utilisant très peu de moyens, sous couvert d'être des expérimentations et dont la continuité et l'intégration n'est pas assurée. Ceci m'amène à formuler le neuvième besoin fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit fournir un parcours de tâches concrètes et collectives à mettre en œuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles. Ce parcours permettrait de donner aux acteurs le moyens d'ouvrir les données et d'intégrer au sein d'un cadre unique toutes leurs tâches. Aussi, les producteurs, intermédiaires et réutilisateurs seront amenés naturellement à échanger de plus en plus ensemble ce qui favorisera la création d'éléments de compréhension partagé entre les systèmes.

À Lyon comme à Grenoble, j'ai retrouvé ce besoin de passer de l'idée d'ouverture des données à la pratique concrète. Maintenant que les dispositifs juridiques sont éprouvés et que l'ouverture des données commence à se répandre dans un grand nombre de collectivités, les producteurs, intermédiaires et réutilisateurs demandent à pouvoir engager des actions concrètes. À Grenoble, j'ai observé que les petites communes étaient en attente du soutien des agents de la métropole pour pouvoir effectivement déposer leurs premiers jeux de données sur la plateforme. À Lyon, j'ai observé des chefs de projets qui ne demandaient que de pouvoir utiliser des données ouvertes dans leurs projets, notamment par l'intermédiaire du Tuba. Dans les deux cas, le manque de moyen bloquait la mise en œuvre de ces actions.

4.4.4.2 Manque d'une vision cohérente autour de la production et de l'utilisation des données

Plus développée à Lyon qu'à Grenoble, la mise en place d'une vision cohérente par l'évangélisation se heurte au temps limité à disposition des producteurs et intermédiaires. Selon les acteurs interrogés, il est très difficile de concilier le travail de production ou d'intermédiation avec le travail d'évangélisation. Pour Laurence Comparat, présidente d'Open Data France, cette incompatibilité peut d'ailleurs se généraliser à toutes les métropoles de France (Entretien 17). De ce constat, je peux formuler le dixième besoin fonctionnel du dispositif en Systèmes d'Information selon lequel il doit offrir des outils permettant d'automatiser et d'autonomiser une partie du travail d'évangélisation. De cette manière, les chargés de projet d'ouverture des données se verront libérer du temps pour mener effectivement le travail de production ou d'intermédiation.

Dans les deux cas, les acteurs interrogés ont fait part de leur besoin de renforcer le travail d'évangélisation d'Open Data France et de ses campagnes d'action nommée "Open Data Locale". Celle-ci supporte une grande partie du travail d'évangélisation en France et commence à autonomiser ce travail en nommant successivement des référents Open Data dans les territoires qui se forment par transmission successive. En répondant à ce dixième besoin fonctionnel, le dispositif en Systèmes d'Information viendrait renforcer ce dispositif en proposant des outils librement mis à disposition comme des cours en ligne ou des supports de communication.

Le tableau ci-dessous résume les besoins fonctionnels relatifs à la création d'un cadre cognitif partagé entre les systèmes.

Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon	Besoins fonctionnels
Manque d'un cadre juridique et politique clair	Manque d'un cadre stratégique clair <i>objectifs moyens alloués</i>	BF9. Le dispositif en système d'information doit fournir un parcours de tâches concrètes et collectives à mettre en oeuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles.
Manque d'une vision cohérente autour de la production et de l'utilisation des données	Coût trop élevé du travail d'évangélisation par rapport aux ressources humaines disponibles	BF10. Le dispositif en système d'information doit offrir des outils permettant d'automatiser et d'autonomiser une partie du travail d'évangélisation.

Figure 4.8 : Les besoins fonctionnels relatifs à la création d'un cadre cognitif partagé entre les systèmes

4.4.5 Les besoins fonctionnels pour lever les obstacles dans l'institutionnalisation de l'usage des données

Quant aux moyens à mettre en oeuvre pour lever les obstacles dans l'institutionnalisation de l'usage des données, il est en grande partie question pour un dispositif en système d'information de proposer un socle sur lequel les organisations pourront développer des documents cadres fixant des règles et principes à suivre en ce qui concerne le développement d'une organisation autour de l'usage des données. Dans cette partie, je présenterai les besoins fonctionnels relatifs à l'incompatibilité des structures organisationnelles avec les données ouvertes et aux diverses cultures de l'*open data* par rapport aux cultures organisationnelles.

4.4.5.1 Incompatibilité des structures organisationnelles avec l'*open data*

Il ne fait aucun doute pour l'ensemble des acteurs lyonnais et grenoblois que les structures des organisations qui produisent les données sont un frein à l'institutionnalisation de l'usage des données. Avec des systèmes d'information en silo ou des démarches de développement applicative basée sur le principe "un besoin, une application, une base de données", il est difficile de mettre en place une organisation autour de l'ouverture des données qui porte les principes contraires (désilotage et démarche intégrée autour d'une base de donnée unique). Ceci m'amène à formuler le onzième besoin

fonctionnel du dispositif en Systèmes d'Information, à savoir qu'il doit fournir un exemple de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes. Ces règles agissent à la manière d'une constitution organisationnelle dont la mise en œuvre assurerait la mise en œuvre d'une organisation favorisant l'usage des données ouvertes.

Les acteurs lyonnais et grenoblois ont exprimé ce besoin en faisant souvent référence à des ensembles de règles développées dans d'autres métropoles françaises ou étrangères. Ainsi, Lyon faisait souvent référence à la "charte des données métropolitaines" de la métropole de Nantes. Elle fixe par exemple pour règles que lorsque des données sont considérées comme "d'intérêt métropolitain" soit mis systématiquement en œuvre un cadre de dialogue avec les réutilisateurs potentiels qui débattrait des modalités d'accès et de réutilisation des données. Ce mécanisme est un exemple de règles clés qui permet le développement de structures organisationnelles adaptées à l'usage des données ouvertes.

4.4.5.2 Divergence culturelle de l'ouverture des données par rapport aux organisations

Enfin, les acteurs interrogés m'ont fait part de la distance culturelle qui sépare l'ouverture des données des organisations qui produisent et ouvrent les données. Ces divergences empêchent d'inscrire durablement l'ouverture des données dans les organisations. De cette analyse, je décrirai le douzième besoin fonctionnel du dispositif en Systèmes d'Information selon lequel il doit fournir des exemples de principes qui cadrent la manière de concevoir l'ouverture des données. Par exemple, les acteurs évoquent régulièrement le principe de "l'open data à la demande" qui veut que la production soit dirigée par les demandes des réutilisateurs et non par les producteurs eux-mêmes. On trouve dans les textes de l'Union Européenne des références au principe "une fois pour toutes" qui veut privilégier l'échange de données entre organismes publics comme source de données principale pour éviter aux contribuables de transmettre leurs données plusieurs fois. Ces principes permettent de cadrer le développement d'une nouvelle culture organisationnelle qui réduit les divergences de l'ouverture des données par rapport aux organisations.

Dans nos entretiens, les acteurs ont insisté sur l'importance d'avoir des grands principes, souvent présentés par une métaphore pour encourager l'institutionnalisation de l'usage des données. Ainsi, les agents de la métropole lyonnaise aiment à dire que les données ouvertes sont l'ADN ou le génotype de la métropole et que pour activer son potentiel elles ont besoin d'un environnement favorable. Le

produit de l'interaction entre un génotype et son environnement étant un phénotype, qui en l'espèce prend la forme de la métropole intelligente.

Le tableau ci-dessous résume les besoins fonctionnels relatifs à l'institutionnalisation de l'usage des données.

Barrières socio-techniques	Exemples rencontrés dans les cas de Grenoble et Lyon	Besoins fonctionnels
Manque d'un cadre juridique et politique clair	Manque d'un cadre stratégique clair <i>objectifs moyens alloués</i>	BF9. Le dispositif en système d'information doit fournir un parcours de tâches concrètes et collectives à mettre en œuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles.
Manque d'une vision cohérente autour de la production et de l'utilisation des données	Coût trop élevé du travail d'évangélisation par rapport aux ressources humaines disponibles	BF10. Le dispositif en système d'information doit offrir des outils permettant d'automatiser et d'économiser une partie du travail d'évangélisation.
L'incompatibilité des structures organisationnelles avec l'open data	Incompatibilité de l'architecture des systèmes d'information avec l'institutionnalisation de la production	BF11. Le dispositif en système d'information doit fournir un exemple de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes.
	Incompatibilité de l'open data avec la démarche de développement des système d'information dites "applicative" (un besoin, une application, une base de données)	
	Opposition entre la silotage organisationnel et la transversalité portée par l'open data	
Divergence culturelle de l'open data par rapport aux organisations	Opposition entre exigences de résultats à court\-terme des organisations et les changements structurels portés par l'open data	BF12. Le dispositif en système d'information doit fournir des exemples de principes qui cadrent la manière de concevoir l'open data

	Opposition entre culture organisationnelle averse aux risques et la culture opportuniste portée par l'open data\.	
	Opposition entre culture organisationnelle portée sur le secret des affaires et la culture de la transparence portée par l'open data	

Figure 4.9 : Les besoins fonctionnels relatifs à l'institutionnalisation de l'usage des données

Conclusion du chapitre 4

L'objectif de ce chapitre était de comprendre quels sont les besoins fonctionnels auxquels un dispositif en Systèmes d'Information doit répondre pour lever les barrières socio-techniques qui freinent la construction de nouveaux usages issus des données ouvertes.

La première, la deuxième et la troisième section de ce chapitre m'ont permis de construire un protocole et des instruments de récolte de données pour établir une étude de cas de l'usage des données ouvertes dans la métropole Grenoble-Alpes et dans la métropole du Grand Lyon. Ce protocole s'est conclu par l'étude de trente documents, la réalisation de vingt et un entretiens directifs et des observations participantes réalisées au cours de 23 rencontres avec les usagers des données ouvertes.

Au cours de la quatrième et dernière section de ce chapitre, j'ai analysé les données récoltées. De cette analyse je conclus qu'un dispositif en systèmes d'information visant à lever les barrières socio-techniques qui freinent la construction de nouveaux usages doit répondre à 12 besoins fonctionnels qui se répartissent le long des cinq étapes de la construction des usages :

1. Le dispositif en système d'information doit supporter la documentation du processus de production de donnée
2. Le dispositif en système d'information doit supporter la documentation de tous les mécanismes techniques ou juridiques qui fixent des restrictions sur les conditions d'utilisation des données
3. Le dispositif en système d'information doit faciliter le design de plateformes *open data* à double entrée, la première étant réservée aux développeurs et experts de la données et la deuxième réservée au grand public.
4. Le dispositif en système d'information doit faciliter la sensibilisation des acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre
5. Le dispositif en système d'information doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel.
6. Le dispositif en système d'information doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre

7. Le dispositif en système d'information, à savoir qu'il doit offrir des outils permettant de comparer la qualité des données entre elles et surtout dans le temps.
8. Le dispositif en système d'information doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité
9. Le dispositif en système d'information doit fournir un parcours de tâches concrètes et collectives à mettre en œuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles.
10. Le dispositif en système d'information doit offrir des outils permettant d'automatiser et d'autonomiser une partie du travail d'évangélisation.
11. Le dispositif en système d'information doit fournir un exemple de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes.
12. Le dispositif en système d'information doit fournir des exemples de principes qui cadrent la manière de concevoir l'ouverture des données.

Évidemment, eu égard à la méthodologie *design science* choisie, il est question dans la suite de cette thèse de créer un artefact qui vise à combler ces douze besoins fonctionnels. Le chapitre suivant aura justement pour objet d'identifier et de combiner des éléments fonctionnels connus de la littérature scientifique dans le domaine des Systèmes d'Information et dont la capacité à répondre à ces besoins a déjà été prouvée.

Chapitre 5

Design d'un dispositif en Systèmes d'Information

Chapitre 5. Design d'un dispositif en Systèmes d'Information permettant de résoudre les situations problématiques qui freinent l'usage des données ouvertes

Ce cinquième chapitre a pour objet de concevoir un dispositif en Systèmes d'Information permettant de résoudre les situations problématiques qui freinent la construction des usages issus des données ouvertes. Il s'inscrit à la fois dans le cycle de design et de rigueur de cette recherche dans la mesure où il vise à développer un artefact qui répond aux besoins exprimés par le terrain (cycle de design) en prenant des théories scientifiques, dites *kernel theories*, pour appui (cycle de rigueur). Aussi sera-t-il guidé par la troisième sous-question de recherche :

Quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les situations problématiques au sein des écosystèmes open data ?

Je rappelle que dans le chapitre précédent, nous avons identifié douze besoins fonctionnels et qu'il est désormais question de passer des besoins fonctionnels au design d'un dispositif qui permet d'adresser les barrières qui freinent la construction des usages issus des données ouvertes. Le design d'un prototype est constitué de trois éléments :

1. Des éléments fonctionnels identifiés à partir des besoins fonctionnels et qui vont permettre de lever les barrières face à la construction des usages. Il s'agit d'une proposition de design (voir définition dans le glossaire ou dans le chapitre 2). La figure ci-dessous illustre le rôle des propositions de design dans cette thèse.

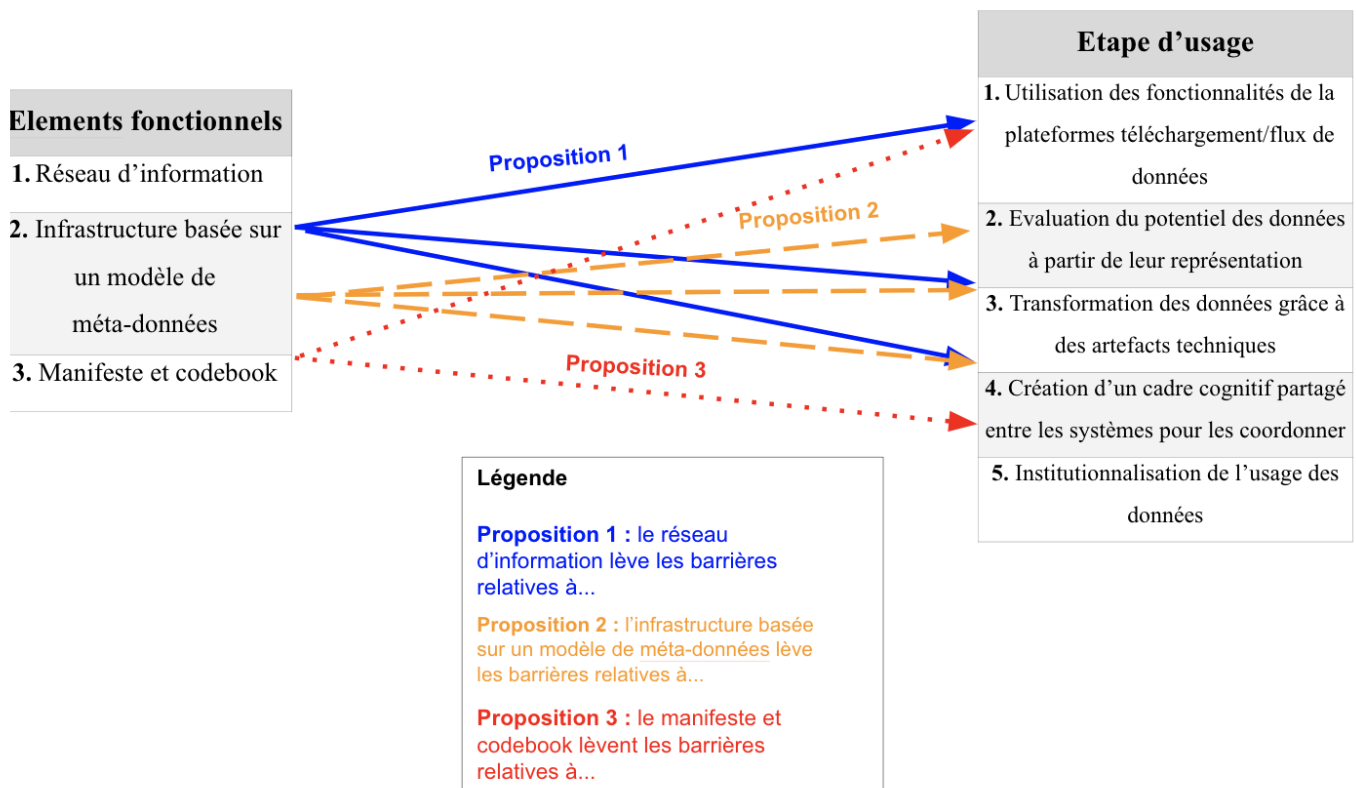


Figure 5.1 : Le rôle des propositions dans le design du dispositif

2. Des mécanismes issus de *kernel theories* qui expliquent comment les éléments fonctionnels vont permettre de lever les barrières . Il s'agit de principes de design (voir définition dans le glossaire ou dans le chapitre 2).
3. Un modèle conceptuel qui dessine l'architecture du dispositif, intègre les éléments fonctionnels et explique les fonctions qu'il assurera dans un contexte réel. On parle de design du système (voir définition dans le glossaire ou dans le chapitre 2).

La figure ci-dessous illustre le rôle des principes, mécanismes et fonctions dans le design du dispositif.

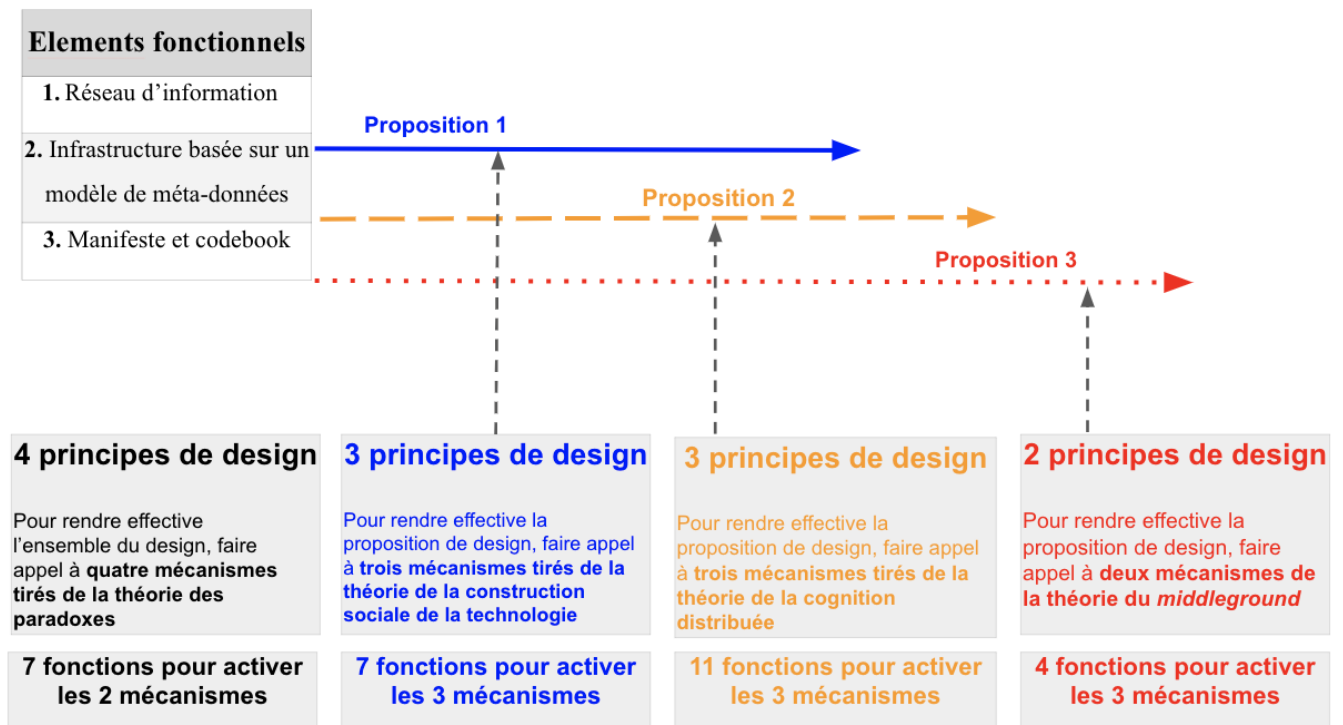


Figure 5.2 : Le rôle des mécanismes, principes et fonctions dans le design du dispositif

Aussi, je rappelle que le design du dispositif s'inscrit dans un cycle d'activité de recherche dédié. Par conséquent, les résultats de ce chapitre seront amenés à évoluer dans le temps. Au moment de terminer le travail de thèse (novembre 2020), le design arbore les caractéristiques suivantes :

- Il intègre trois éléments fonctionnels : un réseau d'information, un système de gestion des tâches relatives à construction des usages ainsi qu'un manifeste et un codebook pour faciliter la compréhension de la démarche.
- Le design est basé sur quatre théories noyau dites *kernel theories* (le lecteur trouvera le détail de l'articulation entre ces quatre cadres théoriques dans la section 8.1.2.2, p.503-504) :
 - les théorie des paradoxes (Smith et al., 2017)
 - la théorie de la construction sociale de la technologie (Pinch et Bijker, 1984)
 - la théorie de la cognition distribuée (Hutchins, 1995)
 - la théorie du *middleground* (Cohendet et al., 1994)

- Trois scénarios d'usage du dispositif ont été construits :
 - Un dispositif pour rallier les diverses communautés d'utilisateurs des données ouvertes autour de piliers définissant l'usage des données ouvertes.
 - Un dispositif pour mutualiser et classer les ressources permettant de développer les usages des données ouvertes
 - Un dispositif pour mettre à disposition du plus grand nombre une forge des usages des données ouvertes.

- Le design comporte 28 fonctions relatives aux trois éléments fonctionnels et à son fonctionnement global.

- Deux versions du prototype ont été créées :
 - Une version 0 qui comporte les cartographies des architectures métier, fonctionnelle et applicative du dispositif.
 - Une version alpha qui est un prototype fonctionnel mais non-opérationnel construit à l'aide d'outils logiciels accompagné d'une méthode d'implémentation dans des contextes d'usages réels.

- La version 0 du prototype a été évaluée eu égard à son utilité, c'est-à-dire à sa capacité à lever les barrières qui freinent la construction des usages issus des données ouvertes. Autrement dit, il a été évalué dans sa capacité à répondre aux douze besoins fonctionnels que le chapitre 2 a permis d'identifier. Il répond aux besoins fonctionnels suivants qui représentent 50 % de la liste initiale :
 - BF4 : Le dispositif en système d'information doit faciliter la sensibilisation des acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre
 - BF5 : Le dispositif en système d'information doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel.
 - BF6 : Le dispositif en système d'information doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre

- BF8 : Le dispositif en système d'information doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité
- BF9 : Le dispositif en système d'information doit fournir un parcours de tâches concrètes et collectives à mettre en oeuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles.
- BF10 : Le dispositif en système d'information doit offrir des outils permettant d'automatiser et d'autonomiser une partie du travail d'évangélisation.

5.1. Construction des propositions de design

Résumé

Le but de ma thèse est de trouver un mode de résolution des tensions paradoxales qui font aujourd'hui obstacle à l'usage des données et ce, dans le but de favoriser la construction des usages qui en sont issus. À cette fin, cette partie vise à extraire de la littérature des éléments qui puissent informer et guider les praticiens dans la réponse à apporter à ces problèmes, formulés sous la forme de besoins fonctionnels au chapitre 4. J'ai présenté dans le chapitre 2 le processus adopté pour sélectionner et classer les articles de la littérature de manière à obtenir un minimum d'éléments fonctionnels couvrant un maximum de besoins fonctionnels.

Ces éléments sont présentés sous la forme de propositions de design. Elles se présentent sous la forme suivante : étant donné le contexte empirique $C(e)$ pour résoudre le problème P et dans l'objectif d'atteindre l'objectif O , alors utiliser le dispositif en Systèmes d'Information I . Dans toutes les propositions de design, les contextes $C(e)$ correspondent au contexte de l'usage des données ouvertes dépeint au chapitre 4 les `problèmes P ` aux barrières qui freinent l'usage des données ouvertes. L'objectif O de chaque proposition est évidemment de lever une ou plusieurs barrières freinant aujourd'hui l'usage des données ouvertes. Enfin, l'identification du dispositif en Systèmes d'Information I est le résultat de la revue de littérature et sera annoncé à la fin de cette première section.

5.1.1 Première proposition : Un réseau d'information autour des données ouvertes pour lever les barrières face à la construction des usages

Traditionnellement, un réseau d'information se définit comme :

Définition : Réseau d'information : “un ensemble de voies relationnelles aptes à faire circuler des flux d'informations entre des acteurs” (Proulx, 1995,p.150 ; Planque, 1983).

Selon la classification de Proulx (1995), les réseaux d'information font partie de la catégorie des réseaux utilitaires, qui constituent :

“Voies relationnelles privilégiées, construites sur la base de mobiles particuliers, de cibles tacites, de tâches non officielles, afin de compléter le travail ou d'améliorer l'efficacité du tissu de réseaux fonctionnels souvent trop rigides ou trop conservateur dans leur topologie ou dans leurs substances transitées” - Proulx (1995, p.159).

Ces réseaux d'information se distinguent par le fait que leur structuration est supportée par des utilités ou des besoins implicites et qu'ils véhiculent des informations qui vont permettre aux acteurs de renforcer leurs connaissances quant à leur rôle et les fonctions qu'ils peuvent exercer au sein du réseau.

Dans ce cadre, un réseau d'information autour des données ouvertes se définit comme un ensemble de relations privilégiées, aptes à faire circuler des flux d'informations entre les acteurs autour de l'usage des données ouvertes et ce afin d'améliorer l'efficacité des organisations impliqués dans la production, l'édition ou la transformation de données. Ce réseau d'information, supporté par un cadre commun adaptable et reproductible, entraîne ses acteurs à partager et ouvrir des données de manière équitable, efficace, juste et sécurisée.

Si la définition abstraite du réseau d'information fait encore aujourd'hui consensus dans le secteur académique, on trouve dans la littérature des visions différentes quant aux réalités qu'elle recouvre. Ainsi, un réseau d'information pourra définir le réseau Usenet, crée en 1979, qui permettait à des universités non liées à l'ancêtre d'Internet (Arpanet) d'échanger des informations sur des sujets informatiques à travers des listes de discussion et des *Newsgroup* (Schafer, 2012).

On trouvera aussi attachée à cette définition des groupements d'établissements de santé qui, à travers un réseau informatique, partagent et s'échangent des rapports sur la qualité des données de santé qu'ils produisent et les moyens de l'améliorer (Tuti et al., 2016). Enfin, il est courant de voir les réseaux d'informations rattachés aux réseaux d'innovation et aux communautés de pratiques qui se développent

notamment dans les secteurs high-techs (Depret et Hamdouch, 2009). Ces réseaux sont supportés par des technologies comme des Intranet ou des Systèmes de Management des Connaissances qui permettent à ses membres d'accéder aux meilleures pratiques mondiales pour les incorporer dans leur contexte spécifique (Tessier, Bourdon et Kimble, 2014).

Définition Data trust : ensemble de relations, supporté par un cadre commun adaptable et reproductible, entraînant ses parties à partager les données de manière équitable, juste et sécurisé (Hall-Prenti report, 2017).

Dans la courte histoire du phénomène d'ouverture des données, les réseaux d'informations ont (re)fait leur apparition très récemment à travers l'émergence de nouvelles formes organisationnelles que l'on retrouve communément sous le terme de *data trust*. On trouve sa première définition dans le rapport Hall-Prenti selon lequel "un data trust consiste en un ensemble de relations, supporté par un cadre commun adaptable et reproductible, entraînant ses parties à partager les données de manière équitable, juste et sécurisé" (Hall-Prenti report, 2017).

Bien que de multiples définitions et conceptions contradictoires aient vu le jour depuis (Open Data Institute, 2018), j'ai choisi de rester fidèle à cette conception première. Je conçois ici un *data trust* comme un réseau d'informations qui se développe en trois étapes :

1. Les producteurs et utilisateurs de données, sur la base de cas concrets d'usage, se rassemblent et transmettent régulièrement, grâce à des supports technologiques, des informations sur les difficultés pratiques et concrètes de l'usage des données dans toutes ses dimensions, ainsi que sur les bonnes pratiques à adopter.
2. Ensemble, ils développent un cadre commun adaptable et reproductible dans lequel ils spécifient un ensemble de pratiques à appliquer par chacun des membres du réseau.
3. Enfin, ils exportent ce cadre pour l'appliquer dans des cas d'usage parents de celui à l'origine de la création du *data trust*.

Un réseau d'information permet de lever les barrières qui freinent l'usage des données de plusieurs manières.

En premier lieu, il vient répondre au flou dans lequel évoluent les organisations qui lancent une démarche d'ouverture des données. Il vient combler le besoin des organisations qui se demandent comment, concrètement, partager et ouvrir leurs données de manière à ce qu'elles produisent de la valeur, tout en évitant qu'elles soient utilisées à des fins frauduleuses (Paprica et al., 2020). A cet égard, le réseau

d'information fournit des processus rigoureusement définis pour permettre la collecte, le stockage, l'utilisation et la publication de données. Un processus qui, documenté et affiché sur les plateformes, fournit des informations précieuses aux utilisateurs. Ces processus peuvent être issus de document cadre utilisés à l'échelle internationale. Par exemple, les *Five Safe* de Ritchie (2017), les recommandations de l'OCDE pour la gouvernance des données de santé (2017) ou encore le cadre de partage responsable des données de la Global Alliance for Genomics and Health (2014).

Définition : Substances : matériau (bien, service, information, données) transitant à travers un réseau

Ensuite, le caractère pragmatique et activable des “substances” (Proulx, 1995) qui circulent à travers le réseau d'information permet d'améliorer la création de valeur initiée par les développeurs tiers qui utilisent les plateformes *open data* (Söderman, 2016). Ces réseaux peuvent par exemple véhiculer des parties de code informatique montrant ce que les producteurs et éditeurs considèrent comme un bon usage de la plateforme et des données. Inversement, ils peuvent faire circuler des parties codes qui montrent un mauvais usage de la plateforme et des données. Pour Monino et Sedkaoui (2016), ces substances pragmatiques sont une solution à la “loi malthusienne des données” :

Définition : Loi malthusienne des données : loi selon laquelle il existe un décalage entre l'augmentation géométrique du volume de données disponibles et l'augmentation au mieux linéaire du volume de données utilisées.

Aussi, le réseau d'information apparaît-il comme un élément fonctionnel d'autant plus intéressant du fait qu'il s'avère peu coûteux en temps et ressources. Il contourne le problème du coût matériel élevé qu'une organisation devrait supporter pour améliorer la documentation et la qualité des données. Ceci s'explique tout d'abord par le fait que le réseau fournit des *feedback* réguliers et gratuits aux producteurs de données qui peuvent ainsi améliorer en continu leurs données. L'autre explication provient du fait que les règles comportementales appliquées au sein du réseau permettent, par capillarité, d'améliorer la culture de la donnée au sein des organisations membres (Tuti et al., 2016). Les règles pratiques du réseau se répandent au sein des organisations membres. Par exemple, le fait que le réseau fixe comme règle d'interaction que toute nouvelle information proposée soit documentée inspirera les membres à faire de même dans le travail quotidien de production de données (Lee et al., 2020).

Ces constats tirés de la littérature me permettent de construire notre première proposition de design selon laquelle un réseau d'information autour des données ouvertes peut permettre de lever des barrières qui freinent plusieurs étapes de l'usage des données.

Exposition de la première proposition de design

Étant donné le contexte de développement des usages des données ouvertes (C), et dans l'objectif (O) de lever les barrières relatives à l'utilisation des fonctionnalités d'une plateforme *open data* (E1), la transformation des données grâce à des artefacts techniques (E3) et la création d'un cadre cognitif partagé entre les systèmes (E4), il est pertinent de mettre en oeuvre un réseau d'information (I)

Plus précisément, nous formulons l'hypothèse selon laquelle le réseau d'information est un élément fonctionnel répondant aux besoins fonctionnels suivants :

- Le dispositif en Systèmes d'Information doit supporter la documentation du processus de production de donnée (BF1)
- Le dispositif en Systèmes d'Information doit supporter la documentation de tous les mécanismes techniques ou juridiques qui fixent des restrictions sur l'utilisation des données (BF2)
- Le dispositif en Systèmes d'Information doit offrir des outils permettant de comparer la qualité des données entre elles et surtout dans le temps (BF7)
- Le dispositif en Systèmes d'Information doit fournir des parcours de tâches concrètes et collectives à mettre en oeuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles (BF9)

5.1.2 Proposition 2 : Une infrastructure open data basée sur un modèle de métadonnées pour lever les barrières face à la construction des usages

Selon Zuiderwijk (2015), une infrastructure de données ouvertes se définit de la façon suivante :

Définition Infrastructure de données ouvertes : "Système partagé, (quasi-)public et évolutif, composé d'un ensemble d'éléments sociaux interconnectés (par exemple les opérations des utilisateurs) et d'éléments techniques (par exemple les outils et technologies d'analyse de données ouverts, les services de données ouverts) qui permettent conjointement l'utilisation d'open data" - Zuiderwijk (2015, p.45)

L'émergence de ces infrastructures de données a pour condition nécessaire le partage d'informations par un nombre très large d'utilisateurs (Janssen et al., 2009). Elles évoluent grâce aux interactions entre des éléments sociaux et techniques (Henfridsson et Bygstad, 2013) et finissent par devenir une ressource publique, partagée, utilisée comme un socle pour une communauté (Hanseth, 2004). Sur ce socle, cette communauté peut développer de nouveaux composants, de nouvelles couches ou des sous-infrastructures (Hanseth, 2004).

Pour Zuiderwijk (2015), une infrastructure *open data* repose sur trois éléments : un modèle de métadonnées, des mécanismes d'interactions et des indicateurs de qualité des données. Dans cette partie, le modèle de métadonnées apparaît comme la colonne vertébrale d'une infrastructure *open data* qui permet de lever plusieurs barrières qui freinent l'usage des données ouvertes.

Les métadonnées se caractérisent généralement comme des "données sur les données". Plus précisément :

Définition Les métadonnées : «Informations structurées qui décrivent, expliquent, localisent ou facilitent d'une autre manière la récupération, l'utilisation ou la gestion d'une ressource d'information» - Organisation nationale de normalisation de l'information (2004, p. 1)

Dans ce cadre, un modèle de métadonnées est un modèle abstrait qui permet d'organiser les "données sur les données", définit les relations entre elles mais aussi leurs relations par rapport aux propriétés d'entités réelles. Le modèle proposé par Zuiderwijk (2015) suggère d'organiser les métadonnées en trois niveaux : le niveau découverte, le niveau contextuel et le niveau "domaine".

Le niveau découverte couvre toutes les métadonnées qui supportent la recherche de données ouvertes et représente un pan majeur de l'usage des données ouvertes que j'ai défini comme la production des données ouvertes. Le niveau contextuel contient toutes les informations qui facilitent le renseignement d'information concernant la production des données. Il représente un autre pan de l'usage des données ouvertes que j'ai défini comme l'intermédiation de données ouvertes. Enfin le niveau domaine concentre les informations sur l'usage possible de certaines données standardisées et propres à des domaines

spécifiques, à l'image des données géographiques. Il représente le dernier pan de l'usage des données ouvertes que j'ai défini comme la transformation des données ouvertes.

Dans la littérature, j'ai identifié de nombreux articles affirmant qu'une infrastructure *open data* basée sur un modèle de métadonnées permet de lever les barrières qui freinent l'usage des données de plusieurs manières.

En premier lieu, et par ses propriétés structurantes, une infrastructure de données ouvertes vient répondre aux problèmes de surcharge informationnelle auxquels font face les producteurs, intermédiaires et utilisateurs. Grâce au modèle de métadonnées, l'infrastructure permet de trier, classer et d'établir des relations intelligentes entre les nombreuses informations que les acteurs doivent traiter pour faire bon usage des données (Curts et Campbell, 2001). Elle leur permet de naviguer et d'identifier rapidement les informations dont ils ont besoin. Par là, l'infrastructure réduit la charge informationnelle (Gavshon et Gorur, 2019).

L'infrastructure basée sur un modèle de métadonnées permet d'autant plus de résoudre les difficultés de surcharge informationnelle qu'elle peut supporter le développement de composants évolutifs. Par conséquent, elle peut absorber la complexification future des données et des informations qui s'y rattachent en supportant le développement de nouveaux composants. Pour Katz et al. (2011), cette capacité à absorber en continu des informations de plus en plus diverses s'explique par la capacité d'une infrastructure à normaliser des informations hétérogènes de manière à ce qu'elles puissent être traitées et que les nouveaux composants puissent être intégrés. Le modèle de métadonnées est ici le principal mécanisme à l'œuvre.

La combinaison des propriétés structurantes et évolutives d'une infrastructure de données ouvertes basée sur un modèle de métadonnées lui permettent de répondre au manque d'un cadre stratégique évolutif exprimé par les producteurs, intermédiaires et utilisateurs. Elle devient le socle sur lequel les acteurs peuvent construire des espaces de travail personnalisés (Denis, 2018) avec des parcours de tâches concrètes à mettre en œuvre pour répondre à leurs problématiques particulières. Ces espaces peuvent être adaptés dans le temps et en fonction des contextes (McQueary et al., 2016). D'une certaine manière, l'infrastructure devient une source de ressources structurantes pour les producteurs, intermédiaires et utilisateurs.

En parallèle, l'infrastructure de données ouvertes basée sur un modèle de métadonnées facilite les interactions entre les différents acteurs au-delà des frontières sectorielles (public, privé, société civile).

Janssen, Chun, & Gil-Garcia (2009) détaille ce mécanisme et avance qu'en facilitant ces interactions, elles font émerger un ensemble dynamique et cohérent, capable de s'adapter aux changements de son environnement. En cela, les auteurs comparent les infrastructures de données aux Systèmes Complexes Adaptatifs systèmes basés sur des agents individuels capables de s'adapter au gré de leurs interactions et des interactions avec leur environnement (Auyang, 1998). Ils utilisent une métaphore qui résume les mécanismes sous-jacents :

"Un exemple typique de CAS est une volée d'oiseaux dans laquelle les oiseaux individuels s'adaptent continuellement aux changements de leur environnement, par exemple la distance entre les oiseaux et la direction du vent. Chaque oiseau suit des règles simples et interagit avec les autres pour former un ensemble cohérent et dynamique." - Janssen, Chun, & Gil-Garcia (2009, p.4)

Enfin, l'infrastructure basée sur un modèle de métadonnées peut apporter une solution à deux problèmes connexes exprimés par les praticiens : le coût trop élevé du travail d'évangélisation et le manque de sensibilisation des acteurs sur la question de l'exploitabilité des données. La solution provient du fait que l'infrastructure de données est une ressource publique et utilisée au quotidien par les acteurs. En effet, cette position lui permet d'introduire indirectement une culture de la donnée au sein des organisations (Tuti et al., 2016), de sensibiliser les agents sur la question de l'exploitabilité et ainsi de les "évangéliser". Pour Denis (2018), ce phénomène s'explique par le fait que ces infrastructures sont des "artefacts scripturaux qui forment (un) vaste processus d'inscription dans lequel nos vies infusent et qui participent à la construction du monde que nous partageons". Il fait appel aux travaux de Smith (1984) qui écrivait à cet égard :

"Nos vies, à un degré plus étendu que nous ne le pensons, sont imprégnées d'un processus d'inscription, produisant des traces imprimées ou écrites ou travaillant à partir d'elles". - Smith (1984, p. 59)

Ces constats tirés de la littérature me permettent de construire notre deuxième proposition de design selon laquelle une infrastructure *open data* basée sur un modèle de métadonnées peut permettre de lever les barrières qui freinent plusieurs étapes de l'usage des données.

Exposition de la deuxième proposition de design

Étant donné le contexte de développement des usages des données ouvertes (C), et dans l'objectif (O) de lever les barrières relatives relatives à l'évaluation du potentiel des données à partir de leur

représentation (E2), la transformation des données grâce à des artefact techniques (E3) et la création d'un cadre cognitif partagé entre les systèmes (E4), il est pertinent de mettre en oeuvre une infrastructure *open data* basée sur un modèle de métadonnées.

Plus précisément, je formulerai l'hypothèse selon laquelle une infrastructure *open data* basée sur un modèle de métadonnées est un élément fonctionnel répondant aux besoins fonctionnels suivants :

- Le dispositif en Systèmes d'Information doit faciliter la sensibilisation des acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre (BF4)
- Le dispositif en Systèmes d'Information doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel (BF5)
- Le dispositif en Systèmes d'Information doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre (BF6)
- Le dispositif en Systèmes d'Information doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité (BF8)
- Le dispositif en Systèmes d'Information doit fournir des parcours de tâches concrètes et collectives à mettre en oeuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles (BF9)
- Le dispositif en Systèmes d'Information doit offrir des outils permettant d'automatiser et d'autonomiser une partie du travail d'évangélisation (BF10)

5.1.3 Troisième proposition : Un manifeste et sa déclinaison opérationnelle, le codebook pour lever les barrières face à la construction des usages

Un manifeste se définit comme :

Définition Un manifeste : “un artefact, implicite ou explicite, qui fait valoir la positionnement stratégique déterminé d'une nouvelle idée, exprimé en termes de différenciation et de valeurs (...) et qui assure l'expression d'une orientation partagée dans une dimension symbolique” (Cohendet et Simon, 2015, p.2).

Il existe des exemples de manifestes dans l'art, avec notamment le manifeste du cubisme (Metzinger et Gleizes, 1913) dans lequel deux peintres défendent leur vision d'une nouvelle peinture, ou dans la cuisine avec le manifeste de la "nouvelle cuisine" des chefs Adria, Blumenthal et McGee (2006) dans lequel ils réinstaurent l'innovation, comme concept central de la cuisine et délaissent le culte de la nouveauté.

Pour Cohendet et Simon (2015), un manifeste joue le rôle suivant :

“Cela permet de comprendre la (nouvelle) idée comme une vision convergente qui ne nécessite pas nécessairement une coordination constante de la part d'un leader. Il fournit au collectif créatif un accord sur l'orientation des efforts, en se concentrant sur un sens partagé et sur un objectif commun bien compris et accepté”. Cohendet et Simon (2015, p.2)

Ce manifeste est souvent décliné dans le domaine opérationnel sous la forme d'un livre de code¹, qui vient préciser la façon dont la nouvelle idée va venir guider la pratique. Littéralement, un “codebook” désigne :

Définition Un codebook : pour une activité cognitive donnée, un codebook est un dictionnaire et un ensemble de règles de grammaire permettant l'usage efficace du dictionnaire (Cowan et al., 2000).

Pour reprendre l'exemple de la peinture et du cubisme, le codebook a pris la forme d'essais écrits par Kahnweiler, un galeriste, qui y transcrit les développements artistiques de Braque et Picasso, sans oublier certains détails techniques. Ces essais sont devenus par la suite l'équivalent de manuels dont le contenu était enseigné dans les écoles d'art parisiennes (Cohendet et al., 2014)

Toujours selon Cohendet et Simon (2015), un codebook remplit la fonction suivante :

« Le codebook émerge généralement de la projection de l'intention créative dans le domaine des utilisateurs: ce qu'ils doivent savoir et faire pour profiter pleinement de la nouvelle idée, une fois qu'elle a été concrétisée en un nouveau produit, service ou processus » Cohendet et Simon (2015, p.3)

Dans la littérature, j'ai identifié de nombreux articles affirmant qu'un manifeste et son codebook permettent de lever les barrières qui freinent l'usage des données de plusieurs manières.

En premier lieu, un manifeste et son codebook viennent apporter une solution au problème de la coordination de nombreuses parties prenantes hétérogènes et qui possèdent des visions différentes, qui doit être assurée quand il s'agit de créer une organisation autour de l'usage des données. Le manifeste permet d'établir des postulats philosophiques qui dépeignent une vision de l'usage des données et le livre

¹Termes originaux : *code book*

de code liste un certain nombre de lois explicatives et un code de conduite pour réaliser le potentiel de cette vision et créer un surplus de valeurs (Ladley, 2020). La coordination par l'intermédiaire de ces deux artefacts n'est possible que s'ils émergent d'un processus collaboratif au cours duquel des compromis ont été trouvés entre les différentes visions de l'usage (Cascio et al., 2019).

Dans la mesure où le manifeste dépeint une réalité donnée avec des lois fixes, il permet également de réduire la complexité de la prise de décision inhérente à l'institutionnalisation de l'usage des données. En effet, ces lois permettent de déduire facilement les décisions à prendre pour institutionnaliser efficacement l'usage. Lammers et Heideweg (2016) en font la démonstration en appliquant les lois de la théorie institutionnelle légale (ILT) et du cadre d'Analyse et de Développement Institutionnel d'Ostrom (2006) à l'implémentation d'une gouvernance des données pour le compte d'un réseau intelligent. Ces lois permettent de déduire des aires décisionnelles dont doivent se saisir les acteurs qui souhaitent organiser l'implémentation et accélère ainsi le processus de décision. De la même manière, Were et Moturi (2017) propose une vision de la gouvernance qui permet d'accélérer la mise en place d'une structure formelle en distinguant trois aires décisionnelles prioritaires : la création de principes de données² *data principles*, les règles de maintenance de la qualité de données, et les mécanismes d'accès aux données.

Aussi, ces deux artefacts permettent-ils de résoudre un dilemme majeur dans l'usage des données ouvertes qui bloque de nombreuses organisations dans l'usage des données : comment libérer, éditer et transformer des données de manière à maximiser la création de valeurs tout en s'assurant qu'aucune norme réglementaire (RGPD par exemple) ou principe (secret des affaires, éthique) n'ait été violé ? Par exemple, Dasgupta et al. (2019) montrent qu'en empruntant une vision de la construction de la gouvernance des données basée sur le cadre des 4I³, les acteurs peuvent progressivement institutionnaliser l'usage des données en s'assurant que l'ensemble des réglementations en matière de données a été respecté. De la même manière, Aisyah et Ruldeviyani (2018) suggèrent aux organisations - et plus particulièrement celles qui possèdent des données sensibles comme les banques et assurances, d'emprunter la vision du "Data Management Body of Knowledge" (DMBOK) créée par l'Association de Management des Données (DAMA International, 2009). Cette vision, lorsqu'elle est appliquée par le respect de son codebook, permet à la gouvernance des données d'intégrer une organisation tout en sécurisant légalement l'usage des données.

Le manifeste et le livre de code offrent aussi des solutions sur la question de l'incompatibilité des structures organisationnelles avec l'ouverture des données. Pour certains auteurs, ils peuvent être porteurs

² Termes originaux : data principles

³ Termes originaux : (Identify, Insulate, Inspect and Improve)

d'une vision quant à la transformation à apporter aux structures organisationnelles pour qu'elles deviennent compatibles avec l'usage des données. Par exemple, Benedict et Schielter (2015) proposent d'appliquer l'heuristique des écosystèmes aux structures organisationnelles (manifeste) en suivant une liste de treize principes formant un codebook. Cette heuristique suggère aux organisations de réfléchir aux moyens de s'ouvrir à son environnement par l'intermédiaire du design d'une plateforme ouverte. Cette démarche de design étant cadré par les 13 principes. Au final, la plateforme ouverte créée permet selon les auteurs de rendre progressivement la structure organisationnelle compatible avec l'usage des données. Reis et al. (2018) propose également de transformer les structures organisationnelles en mettant en place une plateforme, mais cette fois-ci en se basant sur la vision portée par l'Université Notre Dame (Barata et Prado, 2010). Cette vision subdivise en cinq piliers la mise en place d'une organisation compatible avec l'usage des donnée :

1. Définition de politiques en matière de gestion de la qualité des données
2. Choisir des standards de données
3. Définir les données sensibles
4. Définir le type de license à appliquer aux jeux de données
5. Définir une politique d'archivage des données

Notons que dans leurs travaux, Benedict et Schielter (2015) et Reis et al. (2018) affirment que cette transformation des structures organisationnelles a pour corollaire direct l'amélioration de l'utilisabilité et de la lisibilité des plateformes de données ouvertes, du fait d'une meilleure information quant à la production des données (informations utiles à tout réutilisateur potentiel). Par conséquent, le manifeste et le codebook offrent aussi des solutions sur la question du manque de lisibilité des interfaces exprimés par les praticiens rencontrés.

Ces constats tirés de la littérature me permettent de construire la troisième proposition de design selon laquelle un manifeste et son codebook peuvent permettre de lever les barrières qui freinent plusieurs étapes de l'usage des données.

Exposition de la troisième proposition de design

Étant donné le contexte de développement des usages des données ouvertes (C), et dans l'objectif (O) de lever les barrières relatives à l'utilisation des fonctionnalités de la plateforme *open data* (E1), et à l'institutionnalisation de l'usage des données (E5), il est pertinent de mettre en oeuvre un manifeste et son codebook.

Plus précisément, nous formulons l'hypothèse selon laquelle un manifeste et son codebook sont un élément fonctionnel répondant aux besoins fonctionnels suivants :

- Le dispositif en Systèmes d'Information doit faciliter le design de plateformes *open data* à double entrée, la première étant réservée aux développeurs et experts de la donnée et la deuxième réservée au grand public (BF3).
- Le dispositif en Systèmes d'Information doit fournir un exemple de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes (BF11).
- Le dispositif en Systèmes d'Information doit fournir des exemples de principes qui cadrent la manière de concevoir l'ouverture des données (BF12).

Conclusion et transition

Dans cette section, il s'agissait d'exposer les trois propositions de design. Elles établissent pour lever les barrières qui freinent les cinq étapes de l'usage des données, il serait pertinent de mettre en oeuvre trois éléments fonctionnels : un réseau d'information, une infrastructure basée sur un modèle de métadonnées ainsi qu'un manifeste et son codebook. Plus précisément et relativement aux cinq étapes d'usage :

- Etape 1 : il est pertinent de mettre en oeuvre un dispositif en Systèmes d'Information (I) combinant un réseau d'information avec un manifeste et son codebook pour lever les barrières relatives à l'utilisation des fonctionnalités des plateformes *open data*
- Etape 2 : il est pertinent de mettre en oeuvre une infrastructure basée sur un modèle de métadonnées pour lever les barrières relatives à l'évaluation du potentiel des données à partir de leur représentation

- Etape 3 : il est pertinent de mettre en oeuvre un dispositif en Systèmes d'Information (I) combinant une infrastructure avec un réseau d'information pour lever les barrières relatives à la transformation des données ouvertes grâce à des artefacts techniques
- Etape 4 : il est pertinent de mettre en oeuvre un dispositif en Systèmes d'Information (I) combinant une infrastructure avec un réseau d'information pour lever les barrières relatives à la création d'un cadre cognitif partagé.
- Etape 5 : il est pertinent de mettre en oeuvre un manifeste et son codebook pour lever les barrières relatives à la création d'un cadre cognitif partagé

5.2 Construction des principes de design

Résumé

Dans la partie précédente, j'ai établi trois propositions de design. Bien qu'elles donnent au chercheur-concepteur des premiers *inputs* pour guider la conception d'un dispositif en Systèmes d'Information, elles ne permettent pas encore de guider mon action. En effet, elles sont positionnées à un niveau d'abstraction trop élevé pour pouvoir guider la conception d'un artefact, qui sera l'objet de la partie suivante. Je propose donc dans cette partie de construire des éléments permettant aux propositions d'acquérir une dimension concrète et capable de guider mon action de chercheur-concepteur.

Ces éléments sont présentés sous la forme de principes de design. Ils se présentent de la forme suivante : en prenant en considération le contexte théorique $C(t)$ pour résoudre la classe de problème P , il est pertinent de mettre en oeuvre le dispositif en Systèmes d'Information I dans la mesure où elle déclenchera les mécanismes naturels, sociaux ou techniques M , moteurs pour atteindre l'objectif O .

Dans tous les principes de design, le contexte théorique ' $C(t)$ ' correspond à la vision d'un phénomène dictée par la *kernel theory* sous-jacente à chacun des trois éléments fonctionnels (le réseau d'information, l'infrastructure, le manifeste et son codebook). Pour leur part, la classe de problème ' P ' correspond à un type de problème traité par la *kernel theory*, le dispositif en Systèmes d'Information ' I ' à une action que le chercheur-concepteur devra instantier dans son artefact. Enfin, les mécanismes naturels, sociaux ou techniques ' M ' correspondent aux éléments médiateurs, tirés des *kernel theories* et

qui assurent le lien de cause à effet entre le dispositif en Systèmes d'Information `I` et l'atteinte des objectifs `O`, à savoir lever les barrières relatives aux cinq étapes.

Pour identifier les principes de design, j'ai approfondi la revue de littérature établie dans la partie précédente sur la construction des propositions de design. Pour cela, j'ai identifié les références théoriques mentionnées dans les articles pour dresser la liste de l'ensemble des théories susceptibles d'être considérées comme *kernel theories*. Une fois ces *kernel theories* identifiées, j'ai ajouté leur nom à la liste des mots-clés utilisés pour effectuer nos recherches dans les bases de données (Crossref, Google Scholar, Microsoft Academic et Scopus).

5.2.1 Principes relatifs à l'ensemble du design construits à partir de la littérature sur le concept d'usage

Dans la mesure où ma thèse vise à favoriser l'usage des données ouvertes, j'ai en premier lieu ciblé la littérature sur la notion d'usage pour pouvoir identifier des principes de design. Ceux-ci donneront l'orientation générale du design et pourront s'appliquer aux trois éléments fonctionnels du design : le réseau d'information, le modèle de métadonnées ainsi que le manifeste et le codebook.

Définition : **Un usage** : construction socio-techniques émergeant dans un cadre de compréhension partagée et en conséquence des interactions entre trois sphères sociales : - la sphère de production des données ouvertes la sphère de réutilisation des données ouvertes - la sphère d'édition ou intermédiation des données ouvertes. (Ruijter et al., 2018)

D'après les travaux de Paquiénéguy (2017), la définition première de l'usage est issue des études menées dans les années 1980 dont l'objectif était "d'éprouver et de comprendre la manière dont le public rencontre les nouveaux objets qu'étaient alors la télévision et le téléphone" (p. 468). Dans ce contexte, l'usage se définit comme "l'action finalisée d'un usager utilisant un service et les machines qui s'y rattachent, lesquelles ne lui appartiennent pas : il en a la simple jouissance (...) et non la nue-propriété" (p.467).

Toujours selon Paquiénéguy (2017), l'usage et la manière dont il a été étudié révèle que c'est un phénomène :

“à travers lequel se joue un ensemble de tensions. Ces tensions qui sont à l’oeuvre (révèlent) des aspects particuliers de la vie quotidienne, et particulièrement la confrontation des logiques sociales et de la place des technologies de l’information et de la communication (Tic) dans un quotidien forcément social” - Paquienséguy (2017,p. 466)

La prégnance du concept de tension dans l’étude de l’usage m’amène à convoquer les théories des paradoxes, qui conceptualise le phénomène de tension paradoxale, pour identifier des principes de design relatifs à l’ensemble du design. En effet, je choisis de solliciter principalement les théories des paradoxes parmi le florilège de théories et de paradigmes sollicités pour comprendre l’usage (on pourra se référer à l’anthologie des usages de Paquienséguy pour saisir l’étendue du domaine).

Outre le fait que les tensions occupent une place centrale dans l’usage, mon choix se porte sur les théories des paradoxes car elle me semble particulièrement féconde pour identifier des principes de design. Elle a le double avantage d’être suffisamment appareillée conceptuellement pour faire face “aux relations parfois irrationnelles entre des pôles opposés” (Smith et al., 2017, p.313) ainsi que de “fournir une approche prescriptive et managériale qui s’attache à comprendre les pratiques, cognitions et arrangements organisationnels que les managers emploient individuellement et collectivement pour manager les contradictions auxquelles ils font face sans pour autant les résoudre” (Van de Ven et Hargrave, 2017, p.328).

Du fait de l’angle d’attaque prescriptif de les théories des paradoxes, j’ai pu, en scannant les articles qui s’en inspirent, identifier plusieurs mécanismes permettant de favoriser l’usage des données ouvertes, et ainsi donner une orientation générale au design de l’artefact.

D’ailleurs, on se rappellera que plusieurs besoins fonctionnels identifiés au [chapitre 4](#) faisaient directement référence à des tensions paradoxales qui freinent l’usage des données. Ainsi du troisième besoin fonctionnel qui met à jour une tension entre les attentes des développeurs et du grand public vis à vis des plateformes *open data* ; du sixième besoin fonctionnel qui fait état du besoin de faire interagir des acteurs hétérogènes dont les profils sont difficiles à intégrer en temps normal ; du dixième besoin fonctionnel qui répond au manque d’une vision unique et cohérente intégrant la production et l’utilisation des données ; et enfin des onzième et douzième besoin fonctionnel qui répondent aux oppositions structurelles et culturelles engendrées par l’usage des données ouvertes.

D’après les travaux de Hargrave et Van de Ven (2017) qui se basent sur une large revue des différentes approches managériales permettant d’aborder les tensions paradoxales, il est possible isoler quatre mécanismes permettant de favoriser l’usage. Je considère ces quatre mécanismes, et la configuration

PIMCO (Carlsson, 2006) dans lesquelles ils s'inscrivent, comme quatre principes de design donnant l'orientation générale du design de mon artefact.

5.2.1.1 Premier principe

Le premier principe permettant de favoriser l'usage se réfère à une habilitation individuelle : l'acceptation cognitive. Elle se définit comme :

Définition L'acceptation cognitive : “habileté à reconnaître l'interdépendance et la coexistence d'éléments contradictoires et à accepter le paradoxe avant d'adopter un point de vue critique sur les présupposés de la contradiction leur permettant de percevoir les opposés sous un angle plus utile et propice à l'action” (Hargrave et Van de ven, 2017, p.322).

Dans des situations où de nouvelles contradictions émergent et que les individus n'arrivent pas à réconcilier leurs termes contradictoires (*Contexte*), ils font régulièrement face à de la confusion, de l'anxiété et à une forme d'ignorance qui les amènent à adopter une posture défensive (Jarzabkowski, Lê, et Van de Ven, 2013 ; Smith et Berg, 1987). Avec elle, une inertie se forme et les usages ne peuvent se développer (*Problème*). Toutefois, lorsque le management encourage la poursuite simultanée de plusieurs objectifs contradictoires (*Dispositif en Systèmes d'Information*), les individus développent naturellement des formes d'acceptation cognitive et expérimentent alors un état psychologique dit état synergique ou état de flow (Quinn et Cameron, 1988). Celui-ci est une source d'énergie créatrice (Csikzentmihalyi, 1976) (*Mechanisms*) qui permet alors aux individus de s'appuyer sur les tensions pour expérimenter de nouvelles approches, et d'alimenter alors le développement des usages (*Outcome*).

5.2.1.2 Deuxième principe

Dans le même contexte d'émergence de nouvelles contradictions, un deuxième principe, cette fois-ci organisationnel, permet de favoriser l'usage : la synergie. Elle se définit comme :

Définition : La synergie : manière de coordonner des éléments contradictoires visant à satisfaire chacun d'entre eux” (Hargrave et van de Ven, 2017, p.324).

Adopter une forme de management synergique permet de répondre aux situations dans lesquelles différents éléments d'une organisation (comme des directions, ou départements) s'affrontent (bien

qu'elles soient interdépendantes), suite à une restructuration ou à la promulgation de nouvelles réglementations ou directives (*Context*). En mettant en place des actions conjointes aux éléments qui s'affrontent (coopération, collaboration, questionnement stratégique dialectique, arrangements structurels, etc.)(*dispositif en Systèmes d'Information*), les organisations évitent la paralysie de l'affrontement, la renversent (*Problem*), pour créer des cycles vertueux d'apprentissage et d'amélioration de la performance organisationnelle (Smith et Lewis, 2011) (*Outcome*). Ce renversement est rendu possible par le fait que la synergie est un processus continu et désorganisé qui permet aux organisations de traverser (*work through*, Luscher et Lewis, 2008) l'incertitude et les conflits pour construire des "certitudes activables" qui permettent l'action (Luscher et Lewis, 2008)(*Mechanisms*).

5.2.1.3 Troisième principe

Le troisième principe peut être considéré comme parent du deuxième. Il concerne la capacité d'action des acteurs sous-jacente à la mise en place de synergie ou *agency*. Elle se définit comme :

Définition L'agencement : "approche téléologique, empruntant ses codes à la rationalité limitée pour prendre des décisions complexes qui satisfont des objectifs contradictoires" (Hargrave et Van de ven, 2017).

Cette capacité d'action est considérée comme particulièrement pertinente dans un contexte complexe, conflictuel et qui appelle à un changement rapide (*Context*). Elle peut être encouragée par la mise en oeuvre de stratégies connues et théorisées visant à "faire face" (*cope with*, Vlaaar, Van den Bosch et Volberda, 2007) ou à "traverser" les paradoxes (*work through*, Lüscher et Lewis, 2008)(*Dispositif en Systèmes d'Information*). Cette capacité d'action spécifique favorise la mise en place de synergies, et favorise donc indirectement l'usage (*Outcome*), en ce qu'elle invite les acteurs à prendre des décisions de manière incrémentale et inductive, c'est à dire de manière dynamique (*Mechanisms*). Ceci alimente des cycles vertueux et débloquent des confrontations néfastes. (*Problem*) (Hargrave et van de Ven, 2017, p. 328)

5.2.1.4 Quatrième principe

Enfin, le quatrième principe qui favorise l'usage fait écho à un phénomène bien connu dans l'étude des changements à l'échelle supra-organisationnelle : le feedback et la transformation. Il se définit comme :

Définition : Le feedback : “la reproduction, la révision ou la transformation de tensions dynamiques suite à une action de management délibérée (*feedback loop*) donnant lieu à de l'innovation” (Hargrave et van de Ven, 2017, p.323).

Ce feedback et cette transformation permettent de répondre à un besoin de réconciliation ou de dissolution de contradiction entre éléments contradictoires (*Problem*). Ils aboutissent à de l'innovation (*Outcome*) uniquement dans le cas où les éléments contradictoires se trouvent dans un rapport de force équilibré, c'est-à-dire quand l'un et l'autre sont en mesure de s'affronter à armes égales. (Poole et van de Ven, 1989)(*Context*). Alors, dans ce cadre, l'énergie émotionnelle générée par la tension se trouve libérée (Mason, 1996) et vient sortir les acteurs de leur inertie cognitive et comportementale (Coser, 1957 ; Dewey, 1930)(*Mechanisms*). Ce mécanisme présuppose qu'une forme d'antagonisme entre les éléments soit volontairement attisée, sans laquelle il n'est pas possible de voir apparaître une transformation (Coser, 1957)(*Dispositif en Systèmes d'Information*).

Exposition des principes relatifs à l'ensemble du design

Premier principe : lorsque les individus n'arrivent pas à réconcilier les termes contradictoires d'une nouvelle contradiction (C(t)), pour éviter l'inertie cognitive qui se forme alors (P), il est pertinent d'encourager la poursuite simultanée de plusieurs objectifs contradictoires (I) dans la mesure où elle libérera une source d'énergie créative (M), moteurs pour développer de usages (O).

Deuxième principe: lorsque différentes organisations ou parties d'une organisation interdépendantes s'affrontent suite à un changement organisationnel (C(t)), pour éviter la paralysie de l'affrontement (P), il est pertinent de mettre en oeuvre des actions conjointes (coopération, collaboration, etc.)(I) dans la mesure où elles permettront de traverser l'incertitude et les conflits pour construire des “certitudes

activables” qui permettent l’action (M). Ce processus permet de créer des cycles vertueux d’apprentissage et d’amélioration de la performance organisationnelle (O).

Troisième principe: dans un contexte de tensions paradoxales intra-organisationnelles ou inter-organisationnelles qui demandent un changement rapide (C(t)), pour éviter les confrontations néfastes (P), il est pertinent de mettre en oeuvre des stratégies visant à “faire face” (cope with, Vlaar, Van den Bosch et Volberda, 2007) ou à “traverser” les paradoxes (I) dans la mesure où elles favorisent des prises de décisions incrémentales et inductives, moteurs pour la mise en place de synergie (O)(voir deuxième principe).

Quatrième principe: lorsque les éléments contradictoires se trouvent dans un rapport de force équilibré (C(t)), pour répondre à un besoin de réconciliation ou de dissolution de contradiction entre éléments contradictoires (P), il est pertinent d’attiser l’antagonisme entre les éléments (I) dans la mesure où il libérera l’énergie émotionnelle générée par tension et sortira les acteurs de leur inertie cognitive et comportementale (M), moteurs pour générer de l’innovation (O)

5.2.2 Principes relatifs au design d’un réseau d’information construits partir de la littérature sur la théorie de la construction sociale de la technologie (Pinch et Bijker, 1984)

Pour identifier les principes relatifs au design du réseau d’information, je me suis tourné vers la littérature sur la théorie de la construction sociale de la technologie.

Définition La théorie de la construction sociale de la technologie : “théorie qui considère la technologie comme une boîte noire dans laquelle il faut entrer afin de saisir comment le social a influencé son élaboration ou son modelage. (Autrement dit), elle se conçoit comme un assemblage produit à la suite d’un processus de sélection entre différentes options possibles (proposées par différents groupes sociaux)” (Doray, 2017, p.57).

À l’origine, dans leur article séminal, Pinch et Bijker (1984) distinguaient trois moments clé dans la construction sociale de la technologie. Le premier voit des “groupes significatifs” (*relevant groups*),

porteurs de représentations et interprétations, s'affronter pour attribuer un sens à un artefact en élaboration:

Définition Les groupes significatifs : groupe porteurs de représentations et d'interprétations influençant la conception d'un ensemble d'individus

“Se jouent dans cette controverse des visions différentes de l'usage de l'artefact, des correspondances diverses entre l'objet matériel et la signification que des groupes sociaux attribuent à l'objet, et la constitution matérielle des objets.” - Doray (2017, p.57)

Il existe des controverses similaires sur l'ouverture des données : tandis que certains groupes (dont je fais partie) considèrent que la donnée ouverte n'a de valeur que dans ses usages, c'est à dire quand elle est transformée, d'autres tendent plutôt à voir la valeur de l'ouverture des données dans le rôle de courtier informationnel qu'il fait porter aux administrations et organismes publics (Bonacorsi, 2016). Le deuxième moment clé voit la fixation du design de l'artefact par un processus de fermeture des débats. Ce stade est atteint via un consensus qui se forme soit par la disparition des problèmes (les groupes ne voient plus les oppositions qui les divisaient) soit par la redéfinition des problèmes par un groupe qui emporte l'adhésion de tous les autres. Le troisième moment clé voit l'ouverture des groupes et la réintroduction de contextes plus large (sociopolitiques, culturels, etc.) qui viennent façonner la signification attribuée à l'objet.

J'ai choisi ce cadre théorique car les articles sélectionnés dans la partie précédente pour construire les propositions de design, se réfèrent à plusieurs courants théoriques qui me semblent converger vers la théorie de la construction sociale.

Ainsi de la théorie de la diffusion de l'innovation (Rogers, 1983) sur laquelle s'appuient Lee et al. (2020) et qui, dans ses derniers développements, avance qu'avec les nouvelles TIC, la diffusion se confond désormais avec la production et l'appropriation des innovations. De là, je déduis que la diffusion se confond désormais avec la manière dont des groupes d'utilisateurs (les groupes significatifs) vont construire ou co-construire l'innovation technologique et se l'approprier. En somme, les théories de la diffusion de l'innovation et de la construction sociale de la technologie sont devenues interdépendantes.

Définition La théorie de l'apprentissage contextualisé : Théorie sur l'apprentissage dont la focale se porte sur la relation entre l'apprentissage (considérée comme une co-construction de connaissance) et le contexte social dans lequel il a lieu. Cette théorie affirme la place centrale de l'environnement de travail dans les situations (Lave et Wenger, 1991)

Aussi de la théorie de “l’apprentissage contextualisé”⁴ (Lave et Wenger, 1991) sollicitée par Tuti et al.(2016) dont la focale se porte sur l’étude de la relation entre l’apprentissage (considérée comme une co-construction de connaissance) et le contexte social dans lequel il a lieu. Cette théorie partage avec la théorie de la construction sociale la place centrale attribuée à l’environnement de travail. Dans l’une comme dans l’autre, “l’environnement dans lequel se déroulent les pratiques peut être considéré comme le prolongement des capacités cognitives des êtres humains qui le constituent.” (Proulx, 1999).

Ainsi, la co-construction de la connaissance et la co-construction de la technologie sont tous deux grandement influencées par cet environnement cognitif que l’on retrouve dans la théorie de la construction sociale sous le concept de “cadre technologique”⁵. En somme, sur ce point précis, les théories de l’apprentissage contextualisé et de la construction sociale de la technologie sont parentes.

5.2.2.1 Premier principe

Définition La surdétermination de l’usage : stabilisation temporairement des modes d’usage des données dans une direction donnée (Proulx, 2000)

Le premier principe de design du réseau d’information se réfère à une faculté du design l’objet technique qui supporte la circulation des flux d’information entre les acteurs : celle de surdéterminer l’usage qui sera fait du réseau d’information (Woolgar, 1991). En ce qu’il développe un cadre commun adaptable et reproductible dans lequel est spécifié un ensemble de pratiques à appliquer par chacun des membres du réseau (voir description d’un réseau d’information), il peut surdéterminer l’usage (*Mechanism*), c’est à dire stabiliser temporairement des modes d’usage des données dans une direction donnée (Proulx, 2000)(*Outcome*).

Cette stabilisation temporaire s’explique par le fait que ce cadre commun agit alors comme un “cadre technologique”⁶ qui vient fermer un débat latent entre différents groupes constitutifs (*Context & Problem*). Le cadre peut fermer le débat latent car il emporte l’adhésion de tous les autres groupes significatifs. Pour Doray (2017), il est alors question dans ce contexte de comparer ce cadre technologique à un objet-frontière, à savoir une entité qui sert d’interface entre des mondes sociaux et des acteurs ayant des

⁴ Termes originaux : *situated learning theory*

⁵ Termes originaux : *technological frame*

⁶ Termes originaux : *technological frame*

perspectives différentes. Pour déclencher ce mécanisme, le réseau d'information peut recourir à trois moyens directement intégrés dans le design du réseau (Thévenot, 1993) : des prescriptions d'interdictions, c'est à dire des injonctions à ne pas utiliser l'objet technique de telle ou telle manière, et ce pour mettre de côté toute action ou pratique qui irait à l'encontre du maintien du consensus ; des dispositifs contraignants pour orienter les pratiques dans le sens du cadre technologique fixée ; des impositions de normes de "bon usage" en faisant constamment références aux bonnes manières de faire usage des données (*Dispositif en Systèmes d'Information*).

5.2.2.2 Deuxième principe

Le deuxième principe de design du réseau tient de la même capacité du design d'un objet technique à surdéterminer l'usage qui en sera fait, mais permet cette fois-ci aux membres du réseau d'acquérir une certaine culture de la donnée (*Outcome*). L'acquisition de cette culture est rendue possible en attribuant au support technique du réseau une composante éminemment *interactive*, c'est à dire en utilisant des logiciels ou interfaces qui permettent un dialogue constant entre l'utilisateur et le support technique comparable au dialogue qu'aurait un utilisateur avec des données ouvertes (*Dispositif en Systèmes d'Information*). En effet, Proulx écrivait déjà à propos du rôle des dispositifs dans l'acquisition d'une culture numérique cette fois-ci :

"L'attribution de la caractéristique « interactivité » à un dispositif induit une posture spécifique à l'usager potentiel de ce dispositif (Jouët, 1993). L'utilisateur est invité à être « actif » dans le processus i.e. que cette injonction à l'interactivité oblige l'humain à acquérir et maîtriser un minimum de savoir-faire techniques (protocoles, procédures pour manipuler « correctement » l'objet technique) pour pouvoir dialoguer avec le système informatique (Mechanism). L'on pourrait parler ici de l'injonction à acquérir les rudiments d'une culture numérique entendue au sens d'un ensemble d'habiletés liées à la maîtrise de l'intelligence informatique et des protocoles" - Proulx (2001, p.64)

Par conséquent, ce principe est particulièrement pertinent dans un contexte où les acteurs d'un écosystème *open data* doivent acquérir manquant du savoir-faire (*Problem*) dont ils ont besoin pour mener à bien les tâches relatives à l'usage des données ouvertes (*Context*).

5.2.2.3 Troisième principe

Le troisième principe de design du réseau d'information a trait à la place qu'il peut occuper dans la cognition des acteurs, à savoir "devenir le prolongement des capacités cognitives des êtres humains qui le constituent" (Proulx, 2001, p.62). Autrement dit, le réseau d'information peut devenir un stock duquel les agents tirent des ressources cognitives (connaissances formelles, informelles, procédures, buts, etc.) afin d'effectuer des tâches nécessairement collectives et qu'un agent solitaire ne pourrait accomplir. Il se trouvera alors particulièrement efficace dans des situations où des agents isolés doivent coopérer pour effectuer certaines tâches (*Context & Problem*). Ainsi, le réseau d'information permettrait de supporter l'accomplissement de tâches collectives comme la documentation du processus de production des données ou la documentation des restrictions concernant l'utilisation des données (*Outcome*). Un réseau d'information peut produire cet effet en vertu d'un mécanisme ou processus mis en lumière par Hutchins (1995) : la cognition distribuée.

Définition La cognition distribuée/la distribution de la cognition : Théorie selon laquelle le processus mental qui permet d'effectuer des tâches concrètes ne se situe pas exclusivement dans le corps et l'esprit de l'utilisateur : il est distribué entre plusieurs agents humains et non-humains qui coopèrent (Hutchins, 1995).

Celui-ci stipule que le processus mental qui permet d'effectuer des tâches concrètes ne se situe pas exclusivement dans le corps et l'esprit de l'utilisateur : il est distribué entre plusieurs agents humains et non-humains qui coopèrent (*Mechanism*).

Il est possible de stimuler la distribution cognitive induite par le réseau en travaillant sur le couplage des virtualités (Bardini, Horvath, 1995) de chacun des membres du réseau, c'est à dire sur la recherche de complémentarité entre les différentes représentations que les acteurs se font de l'usage du réseau considéré comme environnement cognitif commun (Proulx, 2001)(*Dispositif en Systèmes d'Information*).

Exposition des principes relatifs au design d'un réseau d'information

Premier principe : lorsque plusieurs groupes constitutifs sont en phase de débat pour stabiliser l'usage d'un objet technologique dans une direction donnée (C(t)), pour éviter de cristalliser le débat (P), il est pertinent d'incorporer dans le design du réseau d'information un "cadre technologique" de l'objet sous la forme de prescription d'interdiction, des dispositifs contraignants et des impositions de normes de "bon

usage" du réseau (I). Ces trois moyens vont permettre de surdéterminer l'usage du réseau et, par extension, de l'objet dans une direction donnée (M) pour finalement stabiliser le mode d'usage de l'objet dans une direction donnée.

Deuxième principe : lorsqu'il est question pour des acteurs de mener à bien des tâches relatives à l'usage des données C(t), bien qu'ils ne possèdent pas le savoir-faire nécessaire (P), il est pertinent de rendre le réseau d'information interactif (I) dans la mesure où cette interactivité va obliger les acteurs à acquérir un minimum de savoir-faire technique (M), leur permettant d'acquérir alors un minimum de culture de la donnée (O) moteur pour la réalisation des tâches.

Troisième principe : lorsque des acteurs isolés (P) doivent coopérer pour effectuer certaines tâches (C(t)), il est pertinent de présenter le réseau de manière complémentaire à la représentation que les acteurs ont de l'usage du réseau (I), dans la mesure où cela facilitera la distribution de la cognition (M), motrice pour permettre l'accomplissement de tâches collectives.

5.2.3 Principes relatifs au design d'un modèle de métadonnées construit à partir de la littérature sur la théorie de la cognition distribuée (Hutchins, 1995)

Pour construire les principes relatifs au design d'un modèle de métadonnées, je me suis appuyé sur la littérature autour de la cognition distribuée (Hutchins, 1995).

Cette théorie part du postulat que la cognition humaine s'opère au-delà des limites du corps et du cerveau d'un individu. La cognition est perçue comme désinternalisée et désindividualisée (Proulx, 2001). De fait, les mécanismes qui permettent d'effectuer des tâches sont partagés par plusieurs individus. On trouve alors derrière cette théorie la notion de coopération en réseau orientée vers la réalisation de tâches complexes (Hutchins, 1995).

À partir de ce postulat, il est pertinent de tirer deux grandes conséquences. La première étant que lorsqu'un individu réfléchit à effectuer une tâche, comme faire usage de données ouvertes, cette réflexion va alors s'inscrire dans un contexte plus large que lui-même. Sa représentation personnelle de la tâche à effectuer va être influencée par d'autres représentations individuelles présentes dans son environnement

mais aussi par des représentations plus larges “présentes dans l’esprit du temps” (Proulx, 2001). La deuxième conséquence regarde elle, le fait que les pratiques d’un individus vont être structurées par les ressources auxquelles la distribution de la cognition lui donne accès. Pour Serge Proulx (2001, p.63) : “(la cognition distribuée) induit des possibilités de communication pratique” entre l’individu et son environnement cognitif qui tend vers une structuration mutuelle.

Reposer la construction des principes de design relatifs au modèle de méta-données sur cette théorie m’a semblé pertinent dans la mesure où tous les articles sélectionnés dans la partie précédente font appel à des théories que je considère comme des affluents qui trouvent tous leurs sources dans ce grand fleuve qu’est la théorie de la cognition distribuée. Ainsi de la théorie du deuxième mode de production des connaissances scientifiques (Gibbons et al., 1994) sollicitée par Lipton (2020). Cette théorie explique que la connaissance est nécessairement co-construite et pratique, c’est à dire qu’elle est le fruit d’interaction entre plusieurs processus cognitifs et qu’elle aboutit à des applications concrètes. Ici, je considère qu’il existe un lien de parenté évident entre ces deux théories. Aussi de la théorie des infrastructures, sollicitée et développée par Denis (2018) qui cite explicitement cette théorie comme un de ses piliers historiques (p.107).

5.2.3.1 Premier principe

Définition Une structure médiatrice : Institution ou artefact qui se tient à la frontière entre la sphère privée de l’individu et la sphère publique (Berger, 1976)

Le premier principe de design traite de la capacité du modèle de métadonnées à jouer le rôle d’une “structure médiatrice” (*mediating structure*, Hutchins, 1995) qui organise la réalisation d’une tâche. En tant que structure médiatrice, le modèle de méta-données facilite la propagation d’une forme d’organisation et de savoir-faire présents dans une partie de l’environnement cognitif vers toutes les autres parties (*Outcome*). Faciliter cette propagation peut répondre aux problèmes rencontrés lorsque les acteurs peu nombreux souhaitent transmettre des savoir-faire pratiques à tout un environnement constitué d’un grand nombre d’acteurs. À cet égard, dans son ouvrage *Cognition in the Wild*, Hutchin prend l’exemple de la transmission d’une procédure que tous les membres d’équipage doivent apprendre, bien qu’elles ne soient connues initialement que du capitaine (*Problem & Context*). On peut encourager cette propagation en introduisant dans un système, comme dans une infrastructure, une constitution flexible (*flexible constitution*, Hutchins, 1995).

Par là, j’entends que le système peut être alors organisé de plusieurs manières différentes en gardant toujours l’objectif de la réalisation de tâches particulières (*Dispositif en Systèmes d’Information*). Ce qui fait le lien entre la constitution flexible et la propagation d’une organisation tient au fait que la cognition humaine possède elle-même une architecture extrêmement flexible capable d’effectuer de nombreux modes de raisonnement différents. Dès lors, en assurant à un système une certaine flexibilité, il est possible aux individus de lui donner la forme qui correspond à leur architecture cognitive spécifique (*Mechanism*).

5.2.3.2 Deuxième principe

Le deuxième principe de design est un cas particulier tiré du premier et tient d’un aspect spécifique d’une “structure médiatrice” : la disposition spatiale de ses éléments. D’après Hutchins (1995), la disposition spatiale des éléments d’une structure médiatrice, comme une procédure écrite ou un modèle de métadonnées, détermine l’ordre dans lequel un individu effectue une séquence d’action et le nombre de fois que chaque action doit être menée (p.295)(*Outcome*). Dès lors un modèle de métadonnées peut venir répondre discrètement au flou dans lequel évolue des acteurs qui ne savent pas ce qu’ils doivent faire et dans quel ordre (*Problem*). Pour Hutchins (1995), cette organisation spatiale est un aspect particulièrement important à prendre en compte dans le design d’un artefact quand celui-ci s’adresse à des individus non-qualifiés qui ne pourraient réaliser certaines tâches sans l’aide d’une structure médiatrice (*Context*). Ainsi, en fonction de la séquence d’action qu’un artefact doit susciter, le design doit réfléchir à une représentation spatiale qui traduira au mieux l’ordre des tâches et la manière de les mener (*Dispositif en Systèmes d’Information*). Ce qui fait le lien entre la disposition spatiale des éléments et l’exécution d’une séquence d’action est ce que Hutchins nomme “un processus complexe de coordination” (p.296) entre la procédure suggérée par un artefact médiateur et la procédure mentale invoquée par les acteurs pour utiliser à bon escient cet artefact (*Mechanism*).

5.2.3.3 Troisième principe

La troisième principe du design du modèle de métadonnées permet de s’assurer que l’artefact met toutes les chances de son côté pour susciter une action réelle (*Outcome*) chez ses utilisateurs : il doit effectuer un lien direct entre la tâche suggérée et une tâche du monde réel (*task world*, Hutchins, 1995, p.299). Ce lien

évite aux utilisateurs qui ne sont pas familiers avec le domaine d'application de l'artefact (*Context*) de ne pas comprendre le sens, au sens de finalité réelle, des tâches suggérées (*Problem*). Sans pouvoir accéder au sens, les acteurs délaissent l'artefact. Pour éviter ce délaissement, les désigner peuvent faire implicitement et/ou explicitement faire mention du monde réel auquel est confronté l'utilisateur à chacune des tâches suggérées (*le dispositif en Systèmes d'Information*). Ceci maximisera les chances que la tâche soit réalisée puisque la "coordination entre l'artefact, son sens, et le monde" (Hutchins, 1995, p.300) est assurée (*Mechanism*), donnant ainsi à l'acteur tous les ingrédients nécessaires à son action.

Exposition des principes relatifs au design d'un système basé sur un modèle de méta-données

Premier principe : lorsqu'un groupe d'acteurs peu nombreux souhaite transmettre des savoir-faire pratiques à tout un environnement constitué d'un grand nombre d'acteurs (C(t), P), il est pertinent de permettre au système basé sur un modèle de métadonnées d'être organisé de manière différente (I) dans la mesure où cela lui permettra de s'adapter aux différentes architectures cognitives des futurs utilisateurs (M), et par là, de faciliter la propagation d'une forme d'organisation et de savoir-faire d'une partie de l'environnement cognitif vers toutes les autres parties (O).

Deuxième principe : lorsque des acteurs se trouvent en situation de devoir réaliser certaines tâches pour lesquelles ils ne sont pas qualifiés (C(t)), et donc qu'ils ne savent que faire et dans quel ordre (P), il est pertinent de permettre au système basé sur un modèle de méta-données d'emprunter différentes représentations spatiales traduisant l'ordre des tâches à effectuer et la manière de les mener (I), dans la mesure où cela facilitera la coordination entre la procédure suggérée par un artefact médiateur et la procédure mentale invoquée par les acteurs pour l'utiliser (M). Cela permettra aux acteurs de comprendre l'ordre dans lequel il s'agit d'effectuer une séquence d'action et le nombre de fois que chaque action doit être menée (O).

Troisième principe : lorsque des acteurs ne sont pas familiers avec le domaine d'application d'un l'artefact (C(t)), et qu'ils ne peuvent donc avoir accès au sens des tâches suggérées par l'artefact, ce qui comporte un risque de délaissement de l'artefact par ces acteurs (P), il est pertinent de mentionner le monde réel auquel est confronté l'utilisateur à chacune des tâches suggérées par le système basée sur le modèle de métadonnées (I) dans la mesure où cela assurera la coordination entre l'artefact, son sens et

le monde (M), coordination importante pour maximiser les chances que l'artefact suscite une action réelle chez ses utilisateurs (O).

5.2.4 Principes relatifs au design d'un manifeste et d'un codebook construits à partir de la littérature sur la théorie du *middleground* (Cohendet et al., 2014)

Enfin, pour construire les principes relatifs au design d'un manifeste et d'un codebook, je me suis appuyé sur la littérature autour de la théorie du *middleground* (Cohendet et al., 2014).

Définition La théorie du *middleground* : théorie expliquant comment certaines connaissances novatrices (*creative ideas*) scientifiques, technologiques ou artistiques peuvent s'extraire de leur milieu d'origine pour rayonner à une échelle géographique plus importante.

Elle considère que ces connaissances novatrices sont portées à l'origine par des “communautés épistémiques” (*epistemic communities*), “groupes d'individus mus par un désir de connaissance et soudés par l'intermédiaire d'un but commun, d'un cadre cognitif commun et d'une compréhension partagée de leur travail” (Cohendet et al., 2014, p.930). Pour arriver à rayonner à une échelle globale, ces connaissances traversent et se transforment au gré de conflits entre différents “cadres de référence” (*frames of reference*, Cohendet et al., 2014). Ils sont au nombre de trois : *l'underground*, *le middleground*, et *l'upperground*. Pour qu'une nouvelle connaissance atteigne son apogée, elle suit un processus conflictuel qui l'emmène de *l'underground* vers *l'upperground* en passant par le *middleground*.

L'underground constitue le socle de l'émergence de connaissances novatrices, c'est un cadre de référence centré autour de la génération d'idées nouvelles sans aucun lien avec le monde industriel ou commercial. A l'inverse, *l'upperground* englobe la vision des institutions ou des organisations dont les objectifs sont intéressés et liés à l'intégration de nouvelles idées sur le marché. Entre les deux, la cadre du *middleground* regroupe “un ensemble de plateformes intermédiaires animées par des groupes de communautés reliant la culture informelle de *l'underground* (individus, groupes et communautés) avec les institutions de *l'upperground* (collectivités publiques, industriels)” (Cohendet et al., 2014, p.930).

Il allait de soi d'utiliser la littérature autour de cette théorie pour construire les principes relatifs au design d'un manifeste et d'un codebook dans la mesure où les concepts même de manifeste et de codebook sont tirés de cette récente théorie. Aussi, cette théorie paraissait d'autant plus pertinente qu'elle offre de

nombreux éclairages théoriques sur les mécanismes qui permettent d'institutionnaliser des pratiques nouvelles comme l'usage des données ouvertes peut l'être.

5.2.4.1 Premier principe

Le premier principe de design du manifeste et du codebook permet de garantir qu'ils auront une influence en dehors de la communauté qui les a créés. Il a trait à leur capacité d'intégrer en continu "des répertoire de pratiques et de processus de plus en plus nombreux" (Cohendet et al., 2014, p. 936).

Dans de nombreux cas, les membres de la communauté épistémique à l'origine de la connaissance novatrice sont confrontés au risque de voir leur manifeste mal interprété, de sorte que les procédures, méthodes et expériences promues dans le codebook ne sont pas adoptées en dehors de la communauté (*Context & Problem*). Pour contrecarrer cette tendance, Cohendet et al. (2014) indiquent que les membres de la communauté épistémique doivent communiquer leur manifeste aux autres acteurs et communautés et le confronter à leurs conceptions et pratiques. Aussi suggèrent-ils que le manifeste et le codebook doivent être conçus de manière à pouvoir intégrer les conceptions et pratiques qu'ils rencontreront au cours de ces confrontations (*Dispositif en Systèmes d'Information*). En anticipant cette intégration, le manifeste et le codebook vont pouvoir attirer et convaincre de plus en plus d'acteurs et de communautés qui feront de ces deux artefacts leur référence cognitive et (*Outcome*). Cohendet et al. (2014) explique le lien de cause à effet entre l'évolutivité du manifeste/codebook et la croissance de leur audience en faisant appel au mécanisme de traduction/inscription tel que décrit par Callon et Latour (1991) (*Mechanism*).

5.2.4.1 Deuxième principe

Dans la continuité du premier principe, le deuxième principe oriente le design du manifeste et du codebook de manière à rendre les connaissances novatrices qu'ils véhiculent transmissibles et traduisibles à des structures pré-existantes. Il incite le designer à traduire le manifeste et le codebook en prototypes fonctionnels (un business model, une offre de service, une application par exemple) Ces prototypes vont venir répondre à la problématique à laquelle font face les communautés épistémiques, à savoir convaincre la hiérarchie d'une structure que la connaissance novatrice peut donner lieu à de l'innovation et de la création de valeur, et qu'en sus, cette innovation peut s'insérer dans la structure (*Problem & Context*). Pour que cette connaissance "atterrisse" (Cohendet et Simon, 2015, p.7) dans la structure (*Outcome*), la

théorie du middleground suggère d'alimenter et d'actualiser la connaissance novatrice en l'intégrant, sous le forme de prototype, dans des actions concrètes de l'organisation (*Dispositif en Systèmes d'Information*). Aussi incite-t-elle à manager avec précaution les idées "secondaires" (Hargadon et Sutton, 1999) qui émergeraient en cours de route. Ce sont ces idées et l'expérience concrète accumulée au cours du processus qui contribueront à la traduction finale de la connaissance dans la structure. Pour Cummings, Bilton, and Ogilvie (2015), on active alors le mécanisme de "creativitying" qui complète la production d'idée, souvent restreinte à un petit nombre d'acteurs et théorique, par sa dimension collective et ancrée dans l'action (*Mechanism*).

Exposition des principes relatifs au design d'un manifeste et d'un codebook

Premier principe : lorsque les membres d'une communauté épistémique sont confrontés au risque de voir leur manifeste mal interprété (C(t)), de sorte que les procédures, méthodes et expériences promues dans le codebook ne sont pas adoptées en dehors de la communauté (P), il est pertinent de concevoir le manifeste et codebook de manière à ce qu'ils puissent intégrer de nouvelles conceptions et pratiques (I) dans la mesure où cela leur permettra de faire appel au mécanisme de traduction/inscription (M), moteurs pour attirer de plus en plus d'acteurs et de communautés et d'organisation dans leur giron (O).

Deuxième principe : lorsqu'une communauté épistémique cherche à convaincre la hiérarchie d'une structure que la connaissance novatrice peut donner lieu à de l'innovation et de la création de valeur, et qu'en sus, cette innovation peut s'insérer dans la structure (C(t), P), il est pertinent d'intégrer les connaissances du manifeste et du codebook, sous formes de prototypes, dans des actions concrètes (I) dans la mesure où cela active le mécanisme de "creativitying" qui permet que la connaissance atterrisse dans une structure pré-existante (O).

Conclusion et transition

Sur la base de quatre *kernel theories* (la théorie des paradoxes, la théorie de la construction sociale de la technologie, la théorie de la cognition distribuée et la théorie du middleground), j'ai pu construire douze principes de design. Ils donnent des précisions quant à la manière de concevoir le design du dispositif en Systèmes d'Information.

5.3 Construction du design de l'artefact

Dans la partie précédente, douze principes de design ont été identifiés. Ils permettent de guider l'action du chercheur-concepteur. Dans cette partie, je m'appuie sur ces principes pour construire le design de notre artefact : un environnement de conception des usages des données. Le design se présente sous la forme de combinaisons de mécanismes qui expliquent le comportement des quatre "aires fonctionnelles" de l'artefact : l'environnement de conception pris dans son ensemble, son réseau d'information, son système de gestion des tâches basées sur un modèle de métadonnées et le manifeste et son codebook. La figure ci-dessous (Figure 5.3) schématise la place de l'artefact et de ses différentes aires fonctionnelles dans le modèle théorique de l'usage des données ouvertes construit au chapitre 3. À la fin de la première section de cette partie, le lecteur trouvera la liste des mécanismes décrivant le comportement de l'environnement de conception des usages des données ouvertes.

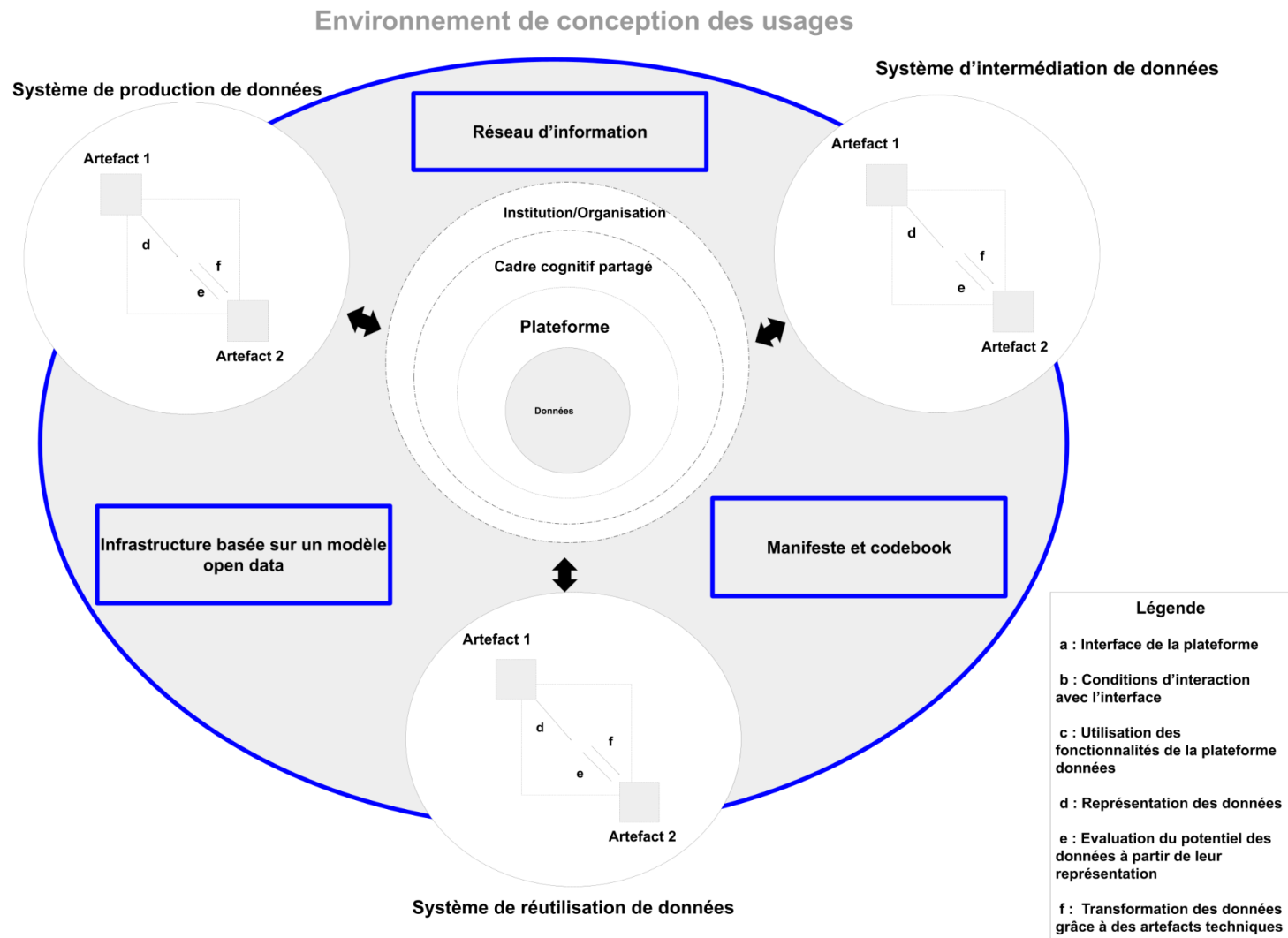


Figure 5.3 : La place du dispositif et de ses différentes aires fonctionnelles dans le modèle théorique de l'usage des données ouvertes

Le comportement défini, il est question dans la deuxième section de décrire comment l'artefact s'insère dans le contexte réel et quotidien de ses futurs utilisateurs. Dans la méthodologie *design science*, répondre à cette question équivaut à construire des *pattern* de coordination, autrement dit des schémas archétypaux décrivant comment l'utilisateur se saisit de l'artefact et l'utilise à son profit. Par soucis d'esthétisme, de lisibilité et pour rester en harmonie avec la rigueur méthodologique, j'ai fait le choix de présenter ces *pattern* sous une forme narrative. En effet, je présenterai trois scénarios d'usage. Ils correspondent aux trois grandes logiques d'action qui ont été synthétisées à partir des douze besoins fonctionnels issus du chapitre 4. Une logique d'action correspond à :

Définition : Logique d'action : tirée d'une vision rationaliste des organisation, une logique d'action correspond à un schéma récurrent et stabilisé qui définit le comportement des organisations et dont les éléments forment un certain équilibre (Hutchel et Molet, 1986, p.183)

En effet, les douze besoins fonctionnels du chapitre 4 ont permis de dégager trois grandes logiques d'action relatives au développement des usages des données ouvertes, appelées de leurs vœux par les acteurs interrogés :

- Scénario 1 : Rallier les diverses communautés autour de piliers définissant l'usage des données ouvertes ;
- Scénario 2 : Mutualiser et classer les ressources permettant de développer les usages ;
- Scénario 3 : Mettre à disposition du plus grand nombre une forge des usages des données ouvertes.

Le lecteur trouvera à la fin de la deuxième section de cette partie les différents éléments constitutifs de ces trois scénarios. Enfin, la section qui conclut cette partie liste les vingt-huit fonctionnalités que j'incorporerai dans l'environnement de conception des usages lors de son prototypage qui fait l'objet de mon septième chapitre.

5.3.1 Construction du design du système de l'environnement de conception des usages dans son ensemble

Dans la section précédente, j'ai identifié quatre principes qui guident le design d'ensemble de l'environnement de conception des usages. Ces principes contiennent quatre mécanismes permettant de favoriser l'usage : 1. l'acceptation cognitive et l'énergie créative qu'elle libère. 2. la synergie qui permet

de passer de l'incertitude et des conflits à des "certitudes activables" qui permettent l'action 3. la capacité d'action des acteurs sous-jacente à la mise en place de synergie qui alimente des cycles vertueux et débloquent des confrontations néfastes. 4. le feedback et la transformation qui libère l'énergie émotionnelle générée par la tension et vient sortir les acteurs de leur inertie cognitive et comportementale pour permettre l'innovation.

A partir de ces 4 principes, nous avons construits 2 mécanismes qui définissent le comportement de l'environnement de conception des l'usage, à savoir le sparring (Lüscher et Lewis, 2008) des acteurs souhaitant développer des usages, et la boucle de rétroaction⁷ qui alimente la transformation des tensions dynamiques pour donner lieu à de l'innovation (voir figure ci-dessous).

Définition Le sparring : processus collaboratif sous forme de dialogue visant à identifier et explorer les tensions paradoxales sous-jacentes au développement des usages pour finalement identifier des pistes d'action visant à les résoudre (à partir de Lüscher et Lewis (2008)).

Définition La boucle de rétroaction : mécanisme qui permet au producteur, à l'intermédiaire ou au réutilisateur d'avoir un retour immédiat de la part des autres acteurs sur les actions, tâches, stratégies telles que formulées à la fin du processus de sparring.

Il est important de préciser, dès à présent que ces deux mécanismes sont nécessairement itératifs et activés ponctuellement tout au long et entre les cinq étapes du développement de l'usage des données ouvertes. Ils constituent à la fois le cœur du comportement de l'artefact de meta-design et le fil rouge des projets de développement des usages.

⁷ Termes originaux : *feedback loop*

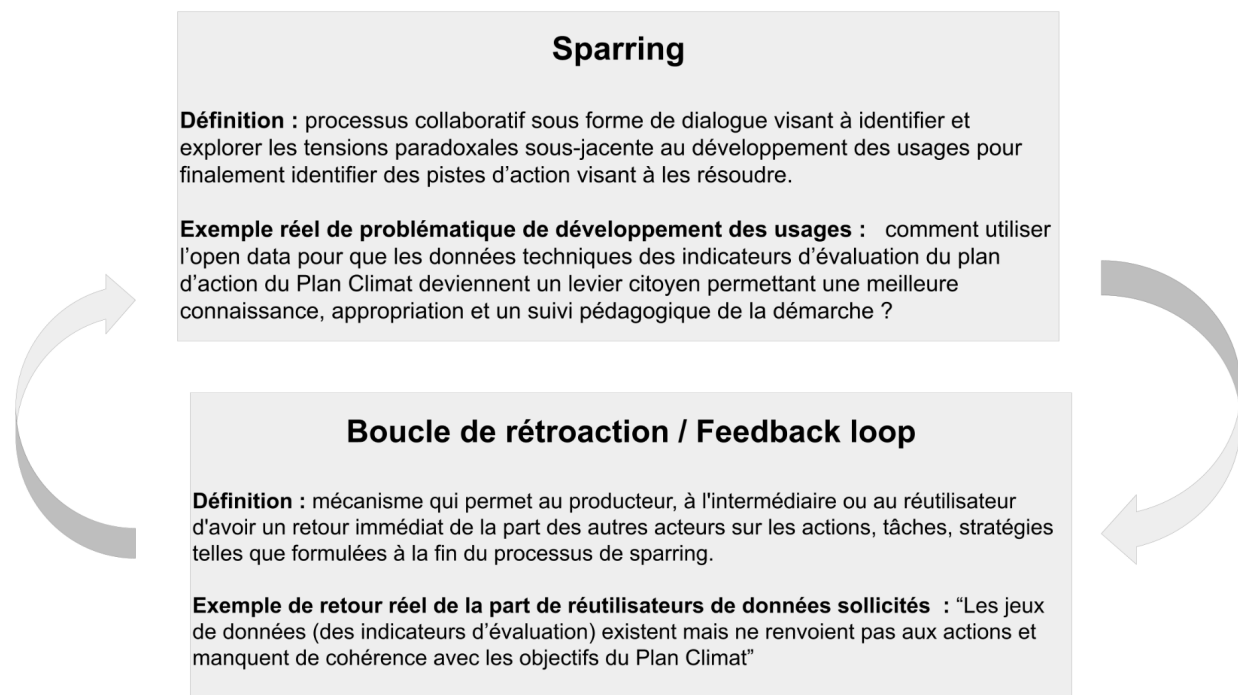


Figure 5.4 : Design des deux mécanismes principaux définissant le comportement du système de l'artefact de meta-design.

5.3.1.1 Le sparring : un processus collaboratif permettant d'aller au bout/de traverser (work through) les paradoxes et tendre vers l'innovation (Luscher et Lewis, 2008)

Le sparring va permettre aux acteurs de décrire les usages qu'ils souhaitent développer, les problématiques qu'ils pensent devoir résoudre et de mettre à l'épreuve leurs conceptions de manière à pouvoir construire des "certitudes activables" qui permettent d'agir pour développer les usages (Luscher et Lewis, 2008).

Dans le sens commun, le *sparring* renvoie au *sparring partner*, défini comme l'individu qui participe à l'entraînement d'un sportif avant une confrontation. Tout en gardant l'idée originelle du terme, il est pertinent d'emprunter ici la conception de *sparring* élaborée par Lüscher et Lewis (2008) pour lesquelles ce terme renvoie à un :

“dialogue visant à mettre à l'épreuve, par des questions exigeantes et stimulantes, les problématiques que rencontrent les acteurs qui se trouvent face à des tensions paradoxales” (p.225)

Ce dialogue a lieu au sein de l'environnement de conception et plus spécifiquement dans des espaces de discussions privés auxquels n'ont accès que l'acteur qui souhaite développer les usages (partie intégrante du système de production, d'intermédiation ou de transformation) et l'intervenant en Systèmes d'Information qui accompagne la mise en oeuvre de l'environnement de co-conception. Le processus suit les cinq étapes décrites par Lüscher et Lewis (2008) et qui portent le nom du résultat ou *output* que chacune d'entre elles doit atteindre : le bazar, le problème, le dilemme, le paradoxe et la certitude activable.

1. **Bazar** Dans un premier temps, le producteur, l'intermédiaire ou le réutilisateur soumet à l'intervenant en Systèmes d'Information une description libre des usages qu'ils souhaitent mettre en place et des difficultés qu'ils rencontrent. L'environnement lui offre la possibilité de décrire librement l'usage et de lister autant de difficultés qu'ils souhaitent sans aucune restriction ni formalisme. Encouragent cette expression libre, il est question pour l'intervenant en Systèmes d'Information d'obtenir le plus d'informations possibles pour la deuxième étape.
2. **Problème.** Sur la base de la description fournie par l'acteur, l'intervenant en Systèmes d'Information pose des questions linéaires (Lane, 1994), qui visent la description des usages visés et des problématiques rencontrées. La réponse à ces questions est censée “viser en priorité un changement à opérer chez l'intervenant, c'est-à-dire dans sa propre compréhension du système-client” (Lane, 1994, p.66). À la suite de ces questions, l'acteur et l'intervenant co-formulent un/des problèmes clairs qu'il s'agit de résoudre. L'environnement permet de basculer automatiquement les réponses de l'acteur et les problématiques formulées dans une base de données relationnelle en ligne.
3. **Dilemme.** Dans cette troisième phase, il s'agit de complexifier les problématiques. Pour cela, l'intervenant en Systèmes d'Information pose des questions circulaires qui demandent à l'acteur de se mettre de se mettre la peau d'autres acteurs appartenant aux systèmes différents du sien. Par ses

réponses, l'acteur est amené à acquérir une perception différente des interactions dans lequel il est imbriqué (Lane, 1994). L'intervenant en Systèmes d'Information peut alors faire émerger des éléments ou pôles contradictoires et montrer qu'un dilemme se forme. Cette phase a pour but "d'intensifier la perception de la paralysie chez les acteurs en exposant les éléments opposés d'une problématique" (Lüscher et Lewis, 2008, p.229). À nouveau, l'environnement transfère les réponses de l'acteur et les dilemmes formulés dans une base de données relationnelle en ligne.

4. **Paradoxe.** Sur la base des dilemmes, l'intervenant en Systèmes d'Information choisit de poser des questions réflexives. Ces questions font appel à la créativité de l'acteur qui, par les questions, est amené à se projeter dans le futur et/ou à se placer dans une position d'observateur. Par exemple, l'intervenant en Systèmes d'Information peut l'inviter dans un premier temps à imaginer des solutions pour résoudre le dilemme, puis les effets de ses solutions sur les autres systèmes, avant de l'inviter à se placer comme observateur externe analysant objectivement ces effets. L'acteur se rendra alors compte qu'aucun choix entre l'un ou l'autre élément du dilemme ne résout le problème, puisqu'ils sont interdépendants. Il voit alors apparaître des paradoxes. Ces paradoxes sont transférées dans une base de données relationnelle en ligne.
5. **Certitude Activable.** Enfin, lors de la dernière étape, l'intervenant en Systèmes d'Information va venir attiser la tension inhérente au paradoxe et confronter plus directement les solutions simplistes amenées par l'acteur. Par exemple, il peut lui demander si les solutions sont vraiment réalistes ou si ils pensent sincèrement qu'elles auront l'effet souhaité puisqu'elles vont nécessairement attiser la tensions entre les éléments du paradoxe. L'acteur sera alors contraint d'imaginer de nouveaux modes d'action, de nouvelles tâches ou de nouvelles stratégies qu'il devra formuler au sein de l'environnement. Encore une fois, les actions, tâches et stratégies formulées sont transférées dans une base de données relationnelle en ligne.

5.3.1.2 La boucle de rétroaction pour alimenter la transformation

La boucle de rétroaction permet au producteur, à l'intermédiaire ou au réutilisateur d'avoir un retour immédiat de la part des autres acteurs sur les actions, tâches, stratégies telles que formulées à la fin du processus de *sparring*. L'acteur a besoin de ces retours pour s'assurer qu'elles sont perçues comme mutuellement bénéfiques par les autres acteurs, et alors générer la synergie nécessaire au développement des usages.

Pour avoir ces retours, l'intervenant en Systèmes d'Information et l'acteur à l'initiative du développement des usages vont identifier les acteurs appartenant aux autres systèmes de l'écosystème et les inviter dans un nouveau groupe de discussion fermé de l'environnement de conception (ou bien dans tout autre environnement qui conviendrait à tous les acteurs). Dans ce groupe, les acteurs seront notifiés de l'existence de la liste d'action, en les invitant à donner leurs retours.

Les retours peuvent être effectués directement sur un tableur rattaché à la base de données relationnelle qui liste toutes les actions, tâches, stratégies envisagées ou par l'intermédiaire d'un questionnaire généré automatiquement à partir du tableur. Les retours peuvent concerner la description des actions, les méthodes choisies pour les mettre en place, leurs objectifs, les rôles que devront jouer les différents acteurs, les données concernées ou encore les domaines de travail (*workspace*) choisis pour mettre en place ces actions.

Les retours sont soumis à l'acteur à l'initiative du développement des usages qui, avec l'intervenant en Systèmes d'Information va alors modifier la liste des tâches envisagées, notifier les acteurs consultés des modifications pour qu'ils puissent approuver ou proposer leur propre liste d'action.

Exposition des mécanismes décrivant le comportement de l'environnement de conception pris dans son ensemble

Mécanisme 1 : le sparring

Dialogue entre l'utilisateur et l'intervenant en Systèmes d'Information ayant lieu dans des espaces de discussion privés au sein de l'environnement de conception. Ce mécanisme ponctue l'utilisation de l'environnement de conception des usages. À des intervalles de temps donné, ce mécanisme permet à l'utilisateur de faire le point sur l'état de développement de l'usage, et de questionner les paradoxes. Ce mécanisme assure également le stockage des éléments du dialogue dans une base de données relationnelle collaborative.

Mécanisme 2 : le feedback et la transformation

Mécanisme permettant à l'utilisateur d'avoir à la fois des *feedbacks* et de transformer la liste des actions, tâches, stratégies telles que formulées à la fin du processus de *sparring*. Ces retours se trouvent sur la même base de données collaborative.

5.3.2 Construction du design du système du réseau d'information

Dans la section précédente, trois principes relatifs au design du réseau d'information ont été identifiés. Ces principes contiennent trois mécanismes permettant de favoriser trois étapes de l'usage des données à savoir l'utilisation des fonctionnalités des plateformes *open data*, la transformation des données grâce à des artefacts techniques et la création d'un cadre cognitif partagé : 1. la surdétermination de l'usage du réseau incorporé dans son design (voir page 298) qui permet de stabiliser les modes d'usage des données dans une direction donnée 2. l'interactivité du réseau qui permet de passer d'acquérir une certaine culture de la donnée 3. la capacité du réseau à devenir le prolongement des capacités cognitives des êtres humains qui permet de supporter l'accomplissement de tâches collectives comme la documentation du processus de production des données ou la documentation des restrictions concernant l'utilisation des données

A partir de ces trois principes, en ressort trois mécanismes qui définissent le comportement du réseau d'information, à savoir la surdétermination de l'usage par la mise en oeuvre de normes techniques et sociales implémentées dans l'objet technologique qui supporte le réseau d'information ; un mécanisme de gestion collaboratif des composants du réseau à travers une base de données collaborative qui rend interactif l'usage du réseau et permet alors l'acquisition d'une certaine culture de la donnée ; et des interfaces de gestion des composants du réseau sous forme de tableur, qui établissent une continuité directe entre les interfaces les plus courantes utilisées individuellement par les acteurs pour travailler avec des données et celles avec lesquelles ils interagissent pour effectuer des tâches collectives avec les membres du réseau autour des données ouvertes (voir figure ci-dessous).

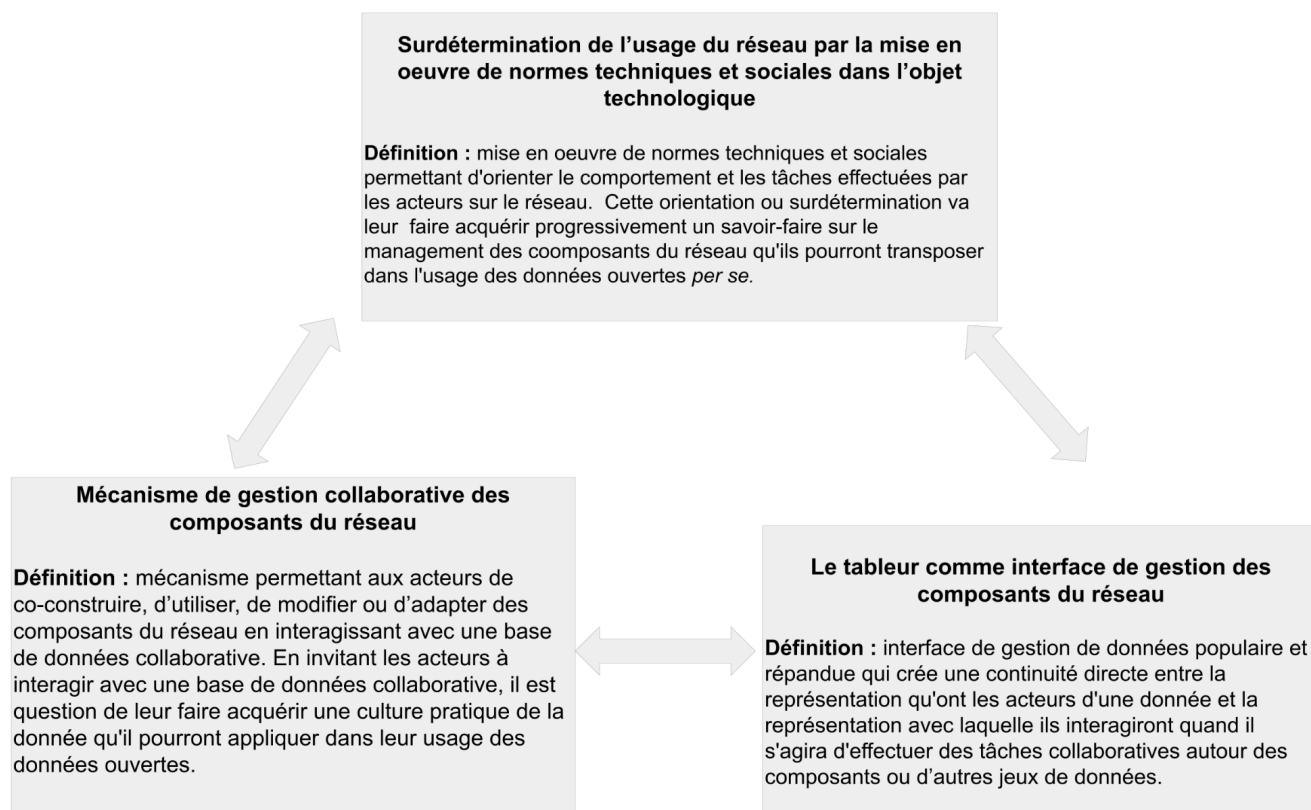


Figure 5.5 : Design des 3 mécanismes principaux définissant le comportement du système du réseau d'information

5.3.2.1 La surdétermination de l'usage du réseau par la mise en oeuvre de normes techniques et sociales implémentées dans l'objet technologique (Lessig, 1999)

La mise en oeuvre de normes techniques et sociales est un moyen d'orienter le comportement et les tâches effectuées par les acteurs sur le réseau (qu'ils soient producteurs, intermédiaires ou réutilisateurs). Cette orientation ou surdétermination va leur faire acquérir progressivement un savoir-faire spécifique sur le management des composants qui constituent le cadre du réseau (voir définition du réseau d'information)

et sur l'accomplissement de tâches collectives. Deux savoir-faire qu'ils pourront transposer dans l'usage des données ouvertes *per se*.

Nous avons conçu deux niveaux de normes interdépendants : un niveau technique et un niveau social présentés ci-après.

5.3.2.1.1 Normes techniques guidant le management des composants du réseau

Les normes techniques s'appliquent au management des composants du réseau. Elles sont portées à la connaissance de tout acteur qui souhaiterait ajouter au cadre du réseau une pratique qui lui semble pertinente pour développer les usages des données (connaissance qui deviendrait alors un composant du réseau). Ces normes s'appliquent également lorsqu'un acteur souhaite réutiliser ces composants, c'est-à-dire les transformer pour son propre usage. Précisées dans le formulaire de sourcing de composants, ces normes techniques donnent des indications sur le fond et la forme que la connaissance doit revêtir pour devenir un composant du réseau ainsi que sur le fond et la forme de toute connaissance créée à partir d'un composant doit revêtir.

Ces normes techniques ont été créées en transposant les principes de Sébastopol, originellement créés pour spécifier la forme et le fond d'un jeu de données ouvert (voir revue de littérature du Chapitre 2). Ainsi, j'ai conçu six normes techniques : la complétude de la connaissance, son caractère générique, son caractère génératif, son caractère collaboratif, son caractère récursif et son respect de l'esprit du manifeste.

Lorsqu'un acteur soumet une connaissance nouvelle ou créé à partir d'un composant existant, l'intervenant en Systèmes d'Information s'appuie sur une grille d'évaluation transparente construite à partir des six critères. Si la pratique ou connaissance respecte les six principes, alors l'auteur de la proposition est notifié de l'acceptation de la connaissance. Dans le cas contraire, l'intervenant en Systèmes d'Information motive le rejet du composant et notifie l'auteur des modifications qu'il faut apporter pour que la connaissance ou pratique intègre le cadre du réseau.

5.3.2.1.2 Normes sociales guidant les interactions entre membre du réseau et l'accomplissement de tâches collectives

Pour que le réseau d'information se développe en volume et se structure autour d'un cadre qui intègre de plus en plus d'aspects de l'usage des données, j'ai élaboré une hiérarchie de règles sociales et d'utilisateurs. Elle est directement inspirée de celle mise en place par la communauté fondatrice de Wikipédia. Elle vise tout comme sa source d'inspiration à instaurer une forme de gouvernance des

interactions entre les membres du réseau et permet de résoudre le dilemme de la création collective de connaissance tel qu'exposé par Earl (2005) :

« D'une part, il y a un réel avantage à rassembler la sagesse collective (augmentation du volume de connaissances); d'autre part, il y a des préoccupations concernant la validité et les incitations ou, plus particulièrement, les dissuasions pour les contributeurs de contenu » - Earl (2005)

Ces règles sont notifiées aux acteurs à leur entrée au sein de réseau et sont mises en avant dès qu'ils souhaitent échanger sur les espaces de discussion du réseau. Quelques exemples de règles sociales qui réglementent les interactions peuvent être cités ici :

- La règle NPOV (neutral point of view) invite les acteurs, avant de poster un message dans les espaces de discussion, à présenter conventionnellement leur pratique et les faits qui s'y rattachent, de manière la moins biaisée et partielle possible.
- La règle "No personal attacks" stipule que si on peut commenter le contenu, on ne doit en aucun cas juger son auteur.
- La règle du "work-in-progress" rappelle que toutes les informations partagées ne sont que partielles, et que la perfection n'est pas requise pour soumettre ses idées au réseau. Aussi précise-t-elle que le réseau est un espace d'entraide, sans pour autant être limité à un rôle de support de discussion.

Enfin, la promulgation et le respect de ces règles est garanti par une hiérarchie d'utilisateurs. Outre les concepteurs du réseau, les utilisateurs qui contribuent le plus au réseau sont investis de rôles et de responsabilités visant à s'assurer que les règles sont respectées et à sanctionner les éventuels écarts. Les utilisateurs grimpent les échelons de la hiérarchie sur la base d'un suivi de leurs interactions sur le réseau (réponses, messages privés, etc.) paramétré par l'intervenant en Systèmes d'Information.

5.3.2.2 Un mécanisme de gestion collaboratif des composants du réseau à travers une base de données collaborative

Lorsqu'un acteur souhaitera participer à la construction des composants du réseau, les commenter, les utiliser ou les adapter pour son propre usage, il devra interagir avec une base de données collaborative qui sert de support de management aux composants. En invitant les acteurs à interagir avec une base de données collaborative, il est question de leur faire acquérir une culture pratique de la donnée qu'ils pourront appliquer dans leur usage des données ouvertes. En un sens, le réseau favorise l'usage des

données ouvertes aussi bien par les informations et composants qu'ils véhiculent, que par l'apprentissage de son utilisation.

Illustration possible du mécanisme

Par exemple, si un acteur ne comprend pas comment utiliser dans sa pratique un composant, il pourra interpeller son auteur directement sur le jeu de données dans lequel le composant est une entité. En retour, l'auteur pourra renseigner des informations plus spécifiques (toujours dans le jeu de données) sur les moyens concrets d'utiliser le composant, et ce, en renseignant directement les métadonnées correspondant au jeu de données en question. Autre exemple : dans le cas où un acteur souhaite exporter une liste de composants qui l'intéresse sur son ordinateur, il serait amené à télécharger un fichier CSV au format non-propriétaire, sa familiarisant ainsi avec le format et la représentation de données brutes. Enfin, si un acteur souhaite inviter des personnes à ajouter des informations à propos de composants, il sera invité à créer un questionnaire dont les champs viendront automatiquement remplir un jeu de données.

5.3.2.2 Le tableur comme interface de gestion des composants du réseau

Pour que l'acteur ne soit pas déstabilisé au moment d'utiliser le réseau et ses composants (càd pour que le réseau s'inscrive facilement comme le prolongement des capacités cognitives des individus), j'ai choisi de présenter la base de données collaborative qui sert de support de gestion des composants, sous la forme de tableur. Dans la mesure où le tableur est l'interface la plus populaire et répandue quand il s'agit de faire usage de données brutes, cette forme d'interface ne créera aucune rupture entre la représentation qu'ont les acteurs d'une donnée et la représentation avec laquelle ils interagiront quand il s'agira d'effectuer des tâches collaboratives autour des jeux de données.

Dans la continuité du mécanisme de gestion collaboratif des composants, ces interfaces vont permettre d'effectuer toute sorte de tâches collaboratives autour de données ouvertes. À l'image du management des composants du réseau, chaque membre pourra également se servir des interfaces pour documenter le

processus de production des données ou les mécanismes de restriction juridique et technique relatifs à l'usage des données, grâce à des zones de texte librement renseignables ou en liant plusieurs tableurs via des associations entre entités. Il est aussi possible d'imaginer que les acteurs se saisissent de ces interfaces pour tout autre pratique collaborative en lien avec l'usage des données ouvertes.

Illustration possible du mécanisme

Par exemple, si le producteur de donnée souhaite renseigner toutes les informations qui ont permis de recenser les immeubles d'une ville, il pourra alors renseigner du texte sur la colonne "ville" du tableur ou bien associer l'entité "ville" présente dans une table de la base de données à une autre table qui contient la description de toutes les entités.

Exposition des mécanismes décrivant le comportement du réseau d'information

Mécanisme 3 : surdétermination de l'usage du réseau par la mise en oeuvre de normes techniques et sociales

Ce mécanisme prend la forme de normes dont l'utilisateur est informé lors de son accès à l'environnement de conception et dont le respect est assuré par l'intervenant en système d'information et la communauté des fondateurs l'environnement. Ce mécanisme permet d'orienter le comportement des acteurs humains au sein du réseau de manière à leur faire acquérir un savoir-faire technique relatif au management des composants du réseau et un savoir-faire relationnel relatif aux interactions au sein du réseau. Il faut préciser que ces savoir-faire s'appliquent également au management des données ouvertes et aux interactions autour de l'usage des données ouvertes.

Mécanisme 4 : gestion collaborative des composants du réseau à travers une base de données collaborative

Ce mécanisme permet à l'utilisateur de participer avec les autres utilisateurs de l'environnement à la gestion des composants du réseau, en interagissant avec une base de données relationnelles collaborative. Par là, il est question de leur faire acquérir une culture pratique de la donnée qu'ils pourront appliquer dans leur usage des données ouvertes.

Mécanisme 5 : un tableur comme interface de gestion des composants du réseau

Ce mécanisme assure la représentation de la base de données relationnelle dans une interface sous forme de tableur. Il permet de faciliter les interactions entre la base de données et l'utilisateur.

5.3.3 Construction du design du système de gestion des tâches à partir d'un modèle de méta-données

Dans la section précédente, trois principes relatifs au design d'un modèle de métadonnées ont été identifiés. Ces principes contiennent trois mécanismes (ci-dessous) permettant de favoriser trois étapes de l'usage des données à savoir l'évaluation du potentiel des données à partir de leur représentation, la transformation des données grâce à des artefacts techniques (E3) et la création d'un cadre cognitif partagé entre les systèmes (E4) :

1. la structuration de systèmes flexibles d'organisation des tâches (Hutchins, 1995), ce qui permet d'organiser de plusieurs manières différentes la réalisation de tâches relatives à l'usage des données ouvertes, et ce tout en gardant la même structure logique (voir concept de structure médiatrice).
2. la plasticité des disposition spatiales de la représentation des tâches qui permet l'exploration et la découverte de la disposition spatiale qui traduira le mieux l'ordre et la manière de mener les tâches relatives au développement des usages des données ouvertes.
3. la création de liens directs entre les tâches suggérées par l'environnement et le monde réel dans lequel elles s'inscrivent qui permet aux acteurs d'accéder au sens, dans le sens de finalité directe et réelle, des tâches.

À partir de ces trois principes, j'ai construit trois mécanismes qui définissent le comportement du modèle de métadonnées, à savoir la flexibilité organisationnelle du système d'organisation des tâches permettant ainsi de maximiser le spectre d'architectures cognitives pouvant se coordonner avec le modèle et la structure logique sous-jacente ; la flexibilité spatiale des représentations des tâches qui permet de trouver la disposition la plus parlante en fonction des acteurs utilisant l'artefact, tout en respectant la structure logique du modèle ; et le rattachement systématique des entités du modèle et des tâches subséquentes à des exemples réels et concrets, qui facilitera l'appropriation du modèle par les acteurs et maximisera les chances que les tâches soient réalisées effectivement (voir figure ci-dessous). Les flèches bidirectionnelles indiquent que les trois mécanismes sont reliés entre eux.

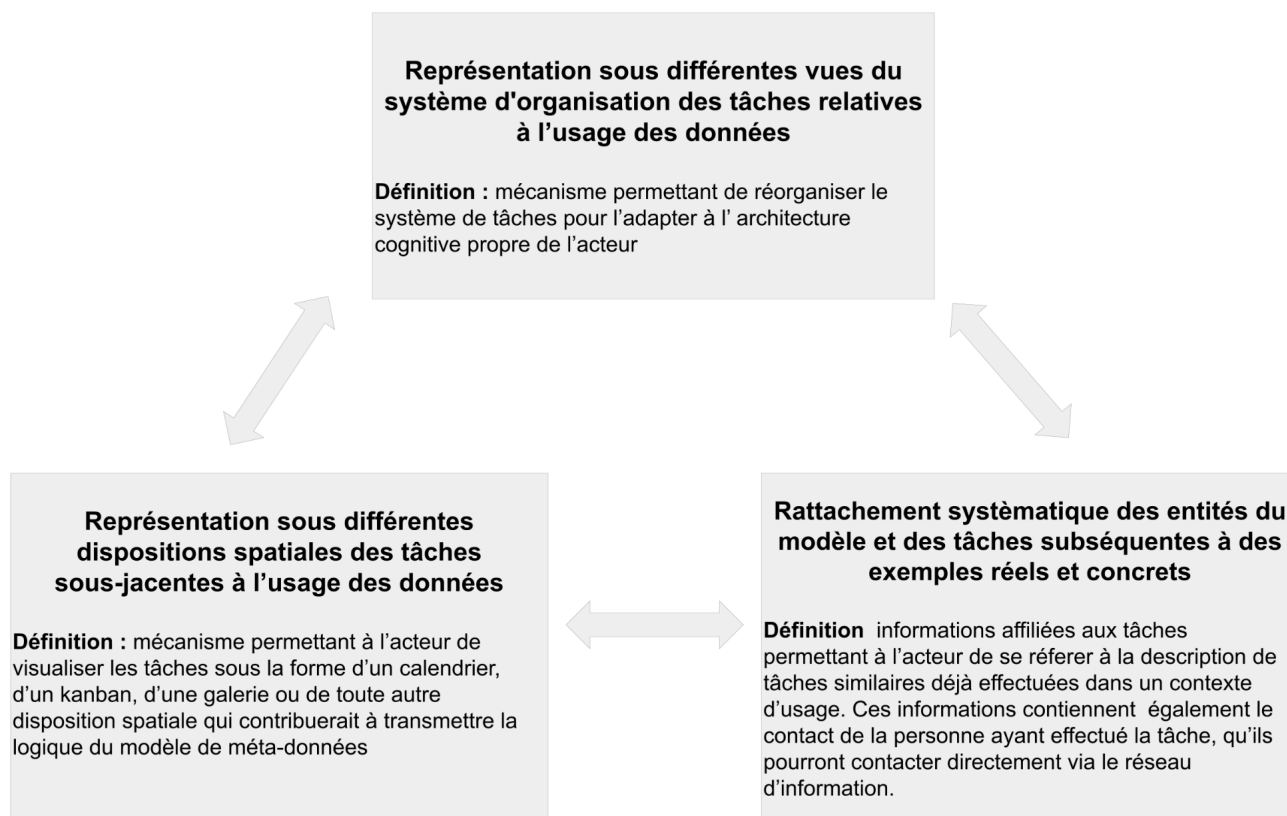


Figure 5.6 : Design des 3 mécanismes principaux définissant le comportement du système de modèle de méta-données

Avant de décrire en détail le design de cette flexibilité organisationnelle, il semble pertinent de revenir sur le design du modèle de métadonnées en lui-même..

5.3.3.1 Le modèle de métadonnées

Le modèle va permettre d'infrastructurer (Star et Bowker, 2002) la manière de concevoir et de réaliser les différentes tâches relatives à l'usage des données. Le modèle de méta-données représente les différentes étapes qui constituent l'usage des données. Elles sont les entités du modèle. À chaque étape est lié un

ensemble de tâches. Ces tâches sont des tâches cognitives complexes que Van Merriënboër (1997) décompose quatre éléments, liste à laquelle nous ajoutons un cinquième élément qui va permettre de rattacher des entités du modèle à des exemples réels et concrets :

- les sous-objectifs de la tâche, qui participent à la réalisation de l'étape lié à la tâche
- des méthodes qui facilitent la réalisation de la tâche ;
- des opérations qui peuvent être entièrement cognitives, donc abstraites et invisibles, ou comportementales, donc visibles et concrètes ;
- des règles de sélection permettant de choisir la meilleure méthode en fonction du contexte.
- un exemple réel et concret de tâche similaire déjà effectuée dans un contexte d'usage réel de données ouvertes.

Les tâches et les éléments constituent les attributs des entités du modèle.

Le modèle établit ensuite les relations que possède chaque entité avec les autres entités. L'ensemble constitue la cardinalité du modèle.

Ensemble, les entités, les attributs et la cardinalité constituent la structure logique inaliénable sur la base de laquelle le système d'organisation des tâches est construit.

5.3.3.2 La flexibilité organisationnelle du système d'organisation des tâches basé sur le modèle

Lorsque l'utilisateur souhaitera réaliser les tâches relatives à l'usage des données, il aura à sa disposition un système qu'il pourra réorganiser comme il l'entend. Ce système intègre la structure logique du modèle de métadonnées. Il se présente sous la forme de tables, constituées de colonnes (représentant les attributs) et de lignes, ces tables étant liées entre elles (ces liaisons représentent la cardinalité du modèle). Il ne peut changer ni le lien entre les tables ni modifier les colonnes présentes dans les tables.

Cependant, il pourra effectuer tout un ensemble de réorganisations pour adapter le système à son architecture cognitive propre. L'intervenant en Systèmes d'Information assistera l'acteur pour adapter au mieux le système. Il mettra à disposition de l'utilisateur une organisation des tâches initiales et lui indiquera les réorganisations possibles. Pour l'utilisateur, cela suppose un coût initial à supporter, à savoir

la rémunération du temps passé par l'intervenant pour configurer les tâches initiales et pour former l'utilisateur de manière à ce qu'il puisse réorganiser lui-même les tâches.

Illustration possible du mécanisme

L'utilisateur pourra regrouper les tâches entre elles selon les critères de son choix, les filtrer ou encore ajouter des colonnes pour désigner des responsables de tâches. *In fine*, il aura à sa disposition des tables de données organisées selon son architecture cognitive spécifique.

5.3.3.3 La plasticité des dispositions spatiales de la représentation des tâches basé sur le modèle

Une fois que l'utilisateur aura organisé les tables selon son architecture cognitive, il pourra également les visualiser de plusieurs manières différentes. Avec l'assistance de l'intervenant en Systèmes d'Information, l'acteur pourra visualiser les tables sous la forme d'un calendrier, d'un kanban, d'une galerie ou de toute autre disposition spatiale qui contribuerait à transmettre la logique du modèle de métadonnées. Pour cela, il pourra exporter le tableur sous un format CSV, et paramétrer avec le fichier d'autres systèmes offrant des représentations spatiales spécifiques (un arbre de décision ou un modèle de *workflow* par exemple).

5.3.3.4 Le rattachement des entités du modèle à des exemples réels et concrets

Enfin, avant de se lancer dans la réalisation d'une tâche, l'utilisateur pourra se référer à la description d'une tâche similaire déjà effectuée dans un contexte d'usage réel des données. Cette description lui donne également le contact de la personne ayant effectué la tâche, qu'il pourra contacter directement via le réseau d'information pour plus d'information ou pour un simple retour d'expérience.

Exposition des mécanismes décrivant le comportement du système de gestion des tâches basé sur un modèle de métadonnées

Mécanisme 6 : un modèle de métadonnées

Le modèle de métadonnées représente la réalité du processus d'usage qui s'applique aux données ouvertes. Il est considéré comme mécanisme en tant qu'il permet de structurer le système de gestion des tâches décrit.

Le modèle se décompose en entités qui correspondent aux cinq étapes de l'usage des données. À chaque entité correspondent des tâches, elles-mêmes subdivisées en cinq éléments. Les tâches et éléments constituent les attributs des entités du modèle. Enfin, il établit des liens entre les différentes entités liées que l'on nomme cardinalité du modèle.

Mécanisme 7 : flexibilité organisationnelle du système d'organisation des tâches

Mécanisme permettant à l'utilisateur de réorganiser le système de gestion des tâches de manière à ce qu'il corresponde avec ses priorités dans le développement d'un usage.

Mécanisme 8 : plasticité des dispositions spatiales de la représentation des tâches

Mécanisme de représentation de la base de données relationnelle dans une interface sous forme de tableur. Il permet de faciliter les interactions entre la base de données et l'utilisateur. Cette réorganisation se trouve limitée par la structure logique du modèle de métadonnées

Mécanisme 9 : rattachement des tâches du système à des exemples réels et concrets

Mécanisme permettant de faire le lien entre les tâches présentes dans le système et des témoignages d'acteurs ayant effectué ces tâches.

5.3.4 Construction du design du système du manifeste et de son codebook

Dans la section précédente, j'ai identifié deux principes relatifs au design du manifeste et du codebook. Ces principes contiennent deux mécanismes permettant de favoriser l'usage : 1. l'incomplétude (*incompleteness*, Garud et Jain, 2008) du manifeste et du codebook, qui va permettre d'attirer et de convaincre de plus en plus d'individus et de communautés de les utiliser. 2. le prototypage du manifeste et du codebook qui va encourager leur intégration au sein de structures pré-existantes.

À partir de ces deux principes, j'ai construit deux mécanismes qui définissent le comportement du manifeste et du codebook, à savoir le sourcing et l'intégration constante de nouvelles pratiques qui va

permettre d'élargir le périmètre des individus et communautés dont les pratiques sont représentées dans le manifeste et le codebook, et le développement et l'expérimentation de produits et services à partir d'eux (voir figure ci-dessous).

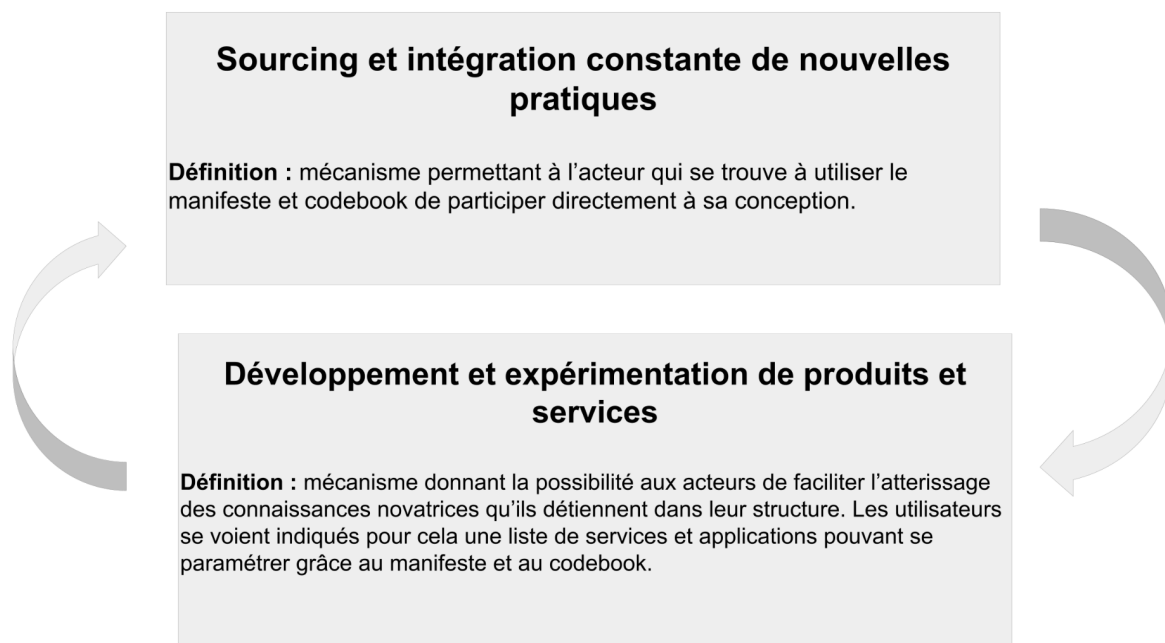


Figure 5.7 : Design des 2 mécanismes principaux définissant le comportement du manifeste et du codebook

5.3.4.1 Le sourcing et l'intégration constante de nouvelles pratiques dans le manifeste et le codebook

Le sourcing et l'intégration constante de nouvelles pratiques permettent à l'acteur qui utilise le manifeste et codebook de participer directement à sa conception. Utiliser et co-concevoir le manifeste et le codebook sont synonymes de son point de vue.

Ainsi, il peut venir directement enrichir le manifeste et son codebook avec des connaissances que son organisation ou lui-même détiendraient sur la compréhension de l'usage des données ouvertes et les pratiques qui peuvent supporter cette compréhension. Par exemple, il peut proposer des déclinaisons du manifeste et du codebook en fonction du type de données. En effet, les conceptions et pratiques de l'usage peuvent varier en fonction du type de données ouvertes.

Illustration possible du mécanisme

L'utilisateur pourra proposer de nouvelles versions du manifeste et du codebook sur un outil de gestion de version. L'intervenant en Systèmes d'Information prendra connaissance de ces nouvelles versions et décidera s'il y a lieu de considérer les propositions comme des manifestes et codebook alternatifs ou comme des développements s'inscrivant dans la continuité des premières versions.

5.3.4.2 Le développement et l'expérimentation de produits et services à partir du manifeste et du codebook

En permettant le développement et l'expérimentation de produits et services à partir du manifeste et du codebook, l'environnement de conception donne la possibilité aux acteurs de faciliter l'intégration des connaissances novatrices qu'ils détiennent dans leur structure.

Après avoir pris connaissance de leur contenu, les utilisateurs se voient indiquer les différentes applications et logiciels ou modèles de service dans lesquels il peut intégrer la conception du manifeste et les règles du codebook. Au cours de cette thèse, j'ai par exemple fait usage du *codebook* au cours d'une conférence réunissant les institutions leaders sur l'ouverture des données, comme d'un "service d'assistance" auprès des producteurs, intermédiaires et utilisateurs de données ouvertes. Je récolte les besoins des acteurs, avant d'effectuer une recherche dans le codebook et de recontacter ces derniers pour leur indiquer la règle qui serait la plus à même de répondre à leur problématique. Encore, j'ai créé un site web à partir du codebook, pouvant servir de support de communication pour les acteurs qui souhaiteraient communiquer autour des données ouvertes.

Avec ces produits et services, l'utilisateur peut plus facilement interpellier sa hiérarchie sur la valeur potentielle du manifeste et du codebook, et maximiser les chances qu'ils soient intégrés au sein de

structures pré-existantes. *In fine*, c'est l'institutionnalisation de l'usage des données porté par l'utilisateur qui est favorisé.

Exposition des mécanismes décrivant le comportement du manifeste et de son codebook

Mécanisme 10 : le sourcing et l'intégration constante de nouvelles pratiques dans le manifeste et le codebook

Via un outil de gestion de version contenu, ce mécanisme permet à tout utilisateur de l'environnement de conception d'amender le manifeste et son codebook en proposant des extensions, de nouvelles versions ou de simples modifications.

Mécanisme 11 : développement de produits et services à partir du manifeste et du codebook

Mécanisme permettant à l'utilisateur de développer des produits (sites web par exemple) ou des services (assistance à réutilisateurs de données) à partir du manifeste et son codebook.

5.4 Construction de scénarios d'usage compilant les pattern de coordination de l'artefact avec un contexte réel de développement des usages des données ouvertes

Dans cette section, il s'agira de mobiliser la méthodologie de construction des scénarios d'usage de Pascal et Rouby (2006). Elle utilise une approche qui combine les apports de la théorie de la cognition distribuée et la théorie de la structuration, dite approche cognitivo-structurationniste.

Le lecteur trouvera ci-dessous la description des trois scénarios d'usage distincts de l'artefact, correspondant aux trois grandes logiques d'action qui ont été synthétisées à partir des douze besoins fonctionnels issus du chapitre 4. Une logique d'action correspond à :

Définition : Logique d'action : tirée d'une vision rationaliste des organisation, une logique d'action correspond à un schéma récurrent et stabilisé qui définit le comportement des organisations et dont les éléments forment un certain équilibre (Hutchel et Molet, 1986, p.183)

Dans chacun de ces scénarios, il sera précisé l'ensemble des items mentionnés ci-dessus.

En effet, les douze besoins fonctionnels du chapitre 4 ont permis de dégager trois scénarios correspondant aux trois grandes logiques d'action relatives au développement des usages des données ouvertes, appelées de leurs vœux par les acteurs interrogés :

- Scénario 1 : Rallier les diverses communautés autour de piliers définissant l'usage des données ouvertes ;
- Scénario 2 : Mutualiser et classer les ressources permettant de développer les usages ;
- Scénario 3 : Mettre à disposition du plus grand nombre une forge des usages des données ouvertes.

5.4.1 Rallier les diverses communautés autour de piliers définissant l'usage des données ouvertes : construction du scénario 1

Dans ce premier scénario, on attend de l'artefact qu'il permette de rallier les diverses communautés qui, en France, favorisent le leadership du développement des usages sur les données ouvertes. Par là, il s'agit de les regrouper sous un étendard commun qui prend la forme de définitions communes et partagées de l'usage des données ouvertes. Les utilisateurs de l'environnement de conception sont a priori les acteurs institutionnels ou associatifs qui œuvrent au niveau territorial et/ou national pour favoriser l'usage des données ouvertes.

Illustration de communautés : transport.data.gouv.fr (Annexe 16)

Transport.data.gouv.fr est une plateforme publique destinée à rassembler toutes les données de mobilité utiles aux développeurs et éditeurs d'application qui facilitent le calcul d'itinéraire, au bénéfice des voyageurs. Autour de la plateforme s'est constitué la communauté des producteurs et utilisateurs des données de mobilités.

Pour les acteurs interrogés, l'artefact est un moyen de rallier un grand nombre de communautés promouvant l'usage des données ouvertes en France. Parmi elles la communauté de l'association Open Data France qui regroupe et soutient les collectivités dans leur démarche d'ouverture des données ; Etalab qui coordonne les actions des administrations de l'Etat et leur apporte son appui pour faciliter la diffusion et la réutilisation de leurs informations publiques ; la communauté Open Street Map spécialisée dans l'ouverture et la réutilisation de données géographiques ainsi que les nombreuses communautés de spécialistes de ce type de données (les géomaticiens) ; les groupements autour de données thématiques à l'image de *transport.data.gouv.fr* et du Point d'Accès National aux données de transport ; enfin, les nombreuses communautés territoriales intéressées dans l'ouverture et l'usage des données à un niveau local.

Entre toutes ces communautés, les acteurs perçoivent une dispersion qui déstabilise les acteurs, qui ne comprennent pas la référence ou l'autorité à suivre :

"Dans l'open data, combien de communauté comme celle-là ? Quel est le rapport avec Open Data France ? Il faut un positionnement visible. Des trucs sont faits à droite à gauche mais on ne sait pas qui est le chef de file. On se dit la même chose au niveau national : qui est chef de file. Mais c'est le problème il y a 5 ans, un foisonnement de truc de droite à gauche mais tu ne sais pas où est le fil". - Fonctionnaire de la métropole Grenoble-Alpes

Dans cette perspective, pour "rallier les diverses communautés autour de piliers définissant l'usage des données ouvertes", l'utilisateur pourra utiliser le manifeste et le codebook de l'environnement de conception des usages. L'acteur institutionnel ou associatif pourra consulter et co-concevoir le manifeste et le codebook de manière à pouvoir le diffuser de manière à ce qu'ils s'adressent à un grand nombre de communautés hétérogènes (voir section "Construction du design du système du manifeste et de son codebook"). Aussi pourra-t-il après les avoir co-conçu, développer des produits et services qui viendront renforcer l'adhésion des communautés visées à la connaissance détenue par le manifeste et le codebook.

Selon les acteurs, ce ralliement des communautés peut espérer voir le jour si l'effort est porté à l'échelle nationale *a minima* et dans le meilleur des cas au niveau européen :

"Ce que fait Open Data France est un premier pas, tout comme les standards que nous imposent l'Etat en termes de marchés publics. C'est intéressant mais un standard européen c'est le minimum qu'on doit viser car si on veut des services avec des modèles économiques qui sont solides. Le standard il faut le chercher au niveau européen. Mais si on en a niveau français, c'est déjà pas mal." - Fonctionnaire de la métropole du Grand Lyon

Cette coordination à l'échelle nationale ou européenne implique que les utilisateurs de l'artefact soient positionnés à différents échelons géographiques, certains agissant à l'échelle nationale et d'autres à l'échelle locale, et qu'ils coordonnent leurs actions. Ainsi il est préférable en termes de coordination que le codebook et le manifeste soient modifiés et co-conçus de manière à avoir à ce qu'il n'y ait qu'une voix à l'échelle nationale.

L'artefact devra alors solliciter et s'implanter dans des réseaux de dimension nationale déjà fortement structurés et qui possèdent la capacité de coordination à l'échelle nationale. Le réseau Open Data France a été évoqué à de nombreuses reprises du fait de l'étendue de son réseau. Sa présidente témoigne de l'importance de s'appuyer sur ces facilités :

“Si on reste dans l'approche des happy few, on s'éloigne de la philosophie de la loi pour une République numérique. C'est pour ça qu'on a lancé Open Data Locale 2 avec cette ambition de passage à l'échelle (2000 collectivités en 2020) sachant qu'on est 400 collectivités aujourd'hui. On s'inscrit dans cette ambition. Mais dans quelle mesure les collectivités vont être en mesure d'allouer les moyens nécessaires, c'est autre chose” - Élu(e) en charge de l'ouverture des données publique pour la ville de Grenoble

La communauté en ligne #teamopendata très active (80000 pages vues, 240 usagers inscrits, 206 sujets, 1,3k messages en 2017) pourra aussi être une ressource à solliciter.

Conclusion

En définitive, ce scénario précise la dimension géographique que doit viser l'artefact ainsi que les caractéristiques des utilisateurs centraux qui peuvent assurer l'implantation de l'artefact dans la réalité sous-jacente à l'usage des données ouvertes. Surtout, ce scénario précise que les effets visés par le manifeste et le codebook, à avoir favoriser la création d'un cadre cognitif commun et l'institutionnalisation de l'usage des données, seront sûrement les plus importants s'ils font l'objet d'une utilisation coordonnée entre différents échelons géographiques.

5.4.2 Mutualiser et classer les ressources permettant de développer les usages : construction du scénario 2

Dans ce deuxième scénario, il est attendu de l'artefact qu'il mutualise et classe les très nombreuses ressources qui ont été créées pour faciliter l'usage des données ouvertes par diverses communautés, organisations et individus en France. Par la mutualisation, il est question de mettre en commun, de centraliser toutes les ressources créées jusqu'à présent ; par classement, il est question de créer un système qui permet d'organiser logiquement toutes ces ressources. Elles recouvrent un large spectre de contenus : recommandations, tutoriels, applications Web, études de cas d'usage, outils de traitement de données, présentations PowerPoint, etc.

Les utilisateurs de l'environnement de conception sont *a priori* tous les acteurs individuels, quelque soit leur position hiérarchique, qui cherchent à développer l'usage des données ouvertes dans les trois systèmes d'un écosystème *open data*. Il peut s'agir des élus, des collectivités qui produisent les données, des responsables opérationnels en charge de développer des projets autour des données ouvertes, les agents et opérateurs qui sont en charge de manipuler les données ouvertes et de développer ses usages.

Par les acteurs interrogés, l'artefact doit constituer un objet qui lie les différentes sphères de l'usage des données entre elles, ainsi que leurs ressources :

“On voit qu'il y a un besoin de mutualisation, de capitalisation, entre collectivités, on le voit au niveau du territoire et entre acteurs publics et privés sur la donnée (...) Il s'agit de se positionner non pas comme un concentrateur mais comme un fédérateur, un acteur qui va créer des ponts entre les acteurs de la donnée, entre réutilisateurs, entre producteurs, entre producteurs et réutilisateurs et avec le reste de la sphère qui sont les associations, les chercheurs, le monde académique” - Fonctionnaire de la métropole du Grand Lyon

Dans cette perspective, pour “mutualiser et classer les ressources permettant de développer les usages”, l'utilisateur pourra utiliser le modèle de métadonnées et le système de gestion des tâches et ressources qui en est issu. Ils pourront y déposer facilement leurs ressources et expliquer les tâches effectuées grâce à elles. Ils placeront également leurs ressources dans le modèle de méta-données conçu.

En parallèle, l'acteur individuel pourra également utiliser et prendre connaissance de l'ensemble des tâches et ressources ayant déjà permis de développer les usages, ainsi que de les classer selon les critères de son choix. Le responsable du service Open Data de la métropole Grenoble-Alpes imaginait à cet égard un large ensemble de modes de tri :

“On imaginait plusieurs entrées : thématique, par acteur. Et à chaque ressource ou tâche, tu dis qui (elle) devrait intéresser. Si tu as une ressource préconisant un Chief Data Officer, tu vas la diriger au Directeur Général, mais pas aux autres. Il faut vraiment que tu sois orienté dans plein d’axes et tu peux faire des choses qui sont chouettes. Plus loin, on pourrait imaginer que les acteurs parlent à un bot qui sait décomposer une requête de façon intelligente pour proposer les bonnes choses.” - Responsable du chargé du service Open Data de la métropole Grenoble-Alpes

Ces ressources et tâches organisées sous la forme d’un tableur, chaque acteur pourra partager son plan d’action avec tout un chacun au sein de son organisation, sans les brusquer cognitivement. Cela sera également renforcé par la possibilité qu’aura l’acteur individuel de visualiser de différentes façons la liste de tâches et de ressources qu’il aura créée. À la suite de cette organisation, il pourra également tout long du développement des usages obtenir des retours d’expérience de la part des acteurs ayant mis à disposition les ressources et tâches, et ce par l’intermédiaire du réseau d’information de l’artefact.

Enfin, il faut préciser que tous les acteurs individuels, de par les normes techniques et sociales incorporées dans le design au réseau et au système de gestion des tâches, acquièrent un savoir-faire pratique unifié. De cette manière, à la mutualisation de ressources sera ajoutée une mutualisation et une convergence des savoir-faire autour de l’usage des données.

D’après les acteurs, la mutualisation se mettra en place à partir du moment où l’acte de partage des ressources sera associé à un projet et designé comme une mission à part entière du développement d’un projet d’usage des données. En associant la mutualisation à un projet spécifique, celle-ci devient un moyen de répondre aux objectifs de délais et de coût d’un projet, contournant alors l’effet “bouteille à la mer” du partage d’une ressource auprès d’un large réseau qui motive peu. On utilisera les ressources et tâches dans un but précis et visible :

“Faire un inventaire de ressources et données pour faire un inventaire, ça n’a aucun sens et ça ne marche pas. En revanche, avoir un projet qui mobilise les acteurs et qui vont les encourager à rendre accessibles leurs ressources et données, là ça marche. Ça c’est mobilisateur. Et le ROI (de la mutualisation) est tout de suite là parce que vous allez répondre aux objectifs de délais et de coût d’un projet. On est dans une dynamique et les acteurs vont être mobilisés naturellement” - Fonctionnaire de la métropole du Grand Lyon

Ceci implique que l’artefact puisse être utilisé “en mode projet”, au-delà d’être une infrastructure. Les objectifs du projet, les délais et les coûts doivent être visibles de tous.

En ce qui concerne le classement des ressources, les acteurs évoquent régulièrement l'importance de créer une classification modelable des données, rôle qu'occupe notre modèle de méta-données.

L'artefact devra alors s'appuyer sur la notoriété d'un acteur réputé pour être un "acteur de confiance", et un tiers éloigné d'intérêts économiques particuliers, auprès duquel il est possible de partager et mutualiser les connaissances sans aucun doute. Aussi cet acteur devra-t-il pouvoir bénéficier d'un accès direct au Systèmes d'information des projets pour récupérer les tâches et ressources directement à la source. À Lyon comme à Grenoble, cette position est unanimement considérée comme un *must-have* pour pouvoir considérer une mutualisation à grande échelle. Lors de l'évènement annuel de valorisation de l'ouverture des données lyonnais auquel j'ai assisté, l'élue du Grand Lyon en charge de l'innovation affirmait à cet égard :

"Nous voulons constituer un patrimoine informationnel commun et pour cela créer un cadre de confiance entre vous et la Métropole dans le but de créer des ponts entre les acteurs de la chaîne de valeur de la donnée" - Élué du Grand Lyon en charge de l'innovation

Conclusion

En définitive, ce scénario précise la dimension temporelle que doit viser l'artefact (le projet), la place que doit avoir cet artefact pour assurer l'implantation de l'artefact dans la réalité sous-jacente à l'usage des données ouvertes (en parallèle et dans la continuité direct d'un projet d'usage des données), ainsi que la notoriété de l'acteur qui le porte (un tiers de confiance). Principalement, ce scénario précise que les effets visés par le modèle de métadonnées et le système de gestion des tâches, à savoir favoriser la création d'un cadre cognitif commun, l'évaluation du potentiel des données ouvertes et leurs transformation, seront sûrement réalisés si l'artefact est consubstantiel de projets de développement des usages des données.

5.4.3 Mettre à disposition du plus grand nombre une forge des usages des données ouvertes : construction du scénario 3

Dans ce troisième scénario, il est attendu de l'artefact qu'il mette à disposition un espace de travail dans lequel on peut concevoir des usages, une forge des usages des données ouvertes.

Définition : Une forge : En informatique, le mot forge est utilisé pour décrire les systèmes de gestion de développement collaboratif de logiciel. Par analogie, je définis une forge comme un système de gestion de développement collaboratif des usages des données ouvertes.

Les utilisateurs sont *a priori* des groupes de réutilisateurs de données répartis par territoire, par organisation ou par thématique. Il peut s'agir d'une seule organisation qui déploie en interne l'artefact, d'une collectivité qui l'utilise comme outil d'accompagnement de développement des usages aux collectivités de niveau infra-territorial ou encore d'un acteur qui déploie l'artefact dans un domaine spécifique comme la santé ou l'énergie.

Par les acteurs interrogés, l'artefact doit constituer un objet qui permet à de nombreuses organisations de se lancer dans la dynamique d'usage des données ouvertes ou à de mettre d'autres organisations, déjà lancées dans l'usage des données, en capacité d'accompagner ces premières. À ce titre, des réutilisateurs de données grenoblois m'ont précisé le besoin auquel cet espace de travail viendrait répondre :

*“La principale barrière au développement de notre application ? Tout d'abord les ressources, qui est le nerf de la guerre, et puis aussi l'accompagnement (...) Aujourd'hui c'est un problème qu'on ait pas de ressource technique pour pouvoir développer le projet. Notre idéal, c'est d'avoir un développeur qui soit impliqué avec nous à plein temps dans le projet. Or, nous ne sommes pas du tout développeur, donc c'est assez difficile.” -
Porteuse d'un projet de réutilisation des données déchets de la métropole Grenoble-Alpes*

Dans cette perspective, pour “mettre à disposition du plus grand nombre une forge des usages des données ouvertes”, l'utilisateur pourra utiliser et combiner le réseau d'information, le système de gestion des tâches et les mécanismes de sparring et de feedback propres à l'environnement de conception pris dans son ensemble. À propos de ces deux derniers mécanismes, les groupes pourront les utiliser pour décrire les usages qu'ils souhaitent développer, formuler le paradoxe qu'ils contiennent, puis les modes d'action qu'ils envisagent pour l'alimenter avant de solliciter des *feedbacks* de la part d'autres acteurs. Par là, l'environnement viendra forger un plan de développement des usages, et, pour me permettre une formule

imaginée, solliciter et recruter d'autres forgerons. Voici pour une tâche que l'artefact peut aider à effectuer alors que l'usage est en encore à l'état d'idée ou d'inspiration.

Une fois un premier plan forgé, les utilisateurs pourront alors utiliser l'environnement de conception comme environnement de travail dans lequel ils pourront inventorier les données qu'ils souhaitent utiliser, les documenter, interagir de manière guidée grâce à l'intervenant en Systèmes d'Information et aux normes techniques et sociales intégrée dans le design du réseau. Encore pourront-ils utiliser l'ensemble des composants du réseau d'information pour guider la réalisation des tâches nécessaires au développement des usages.

D'après les acteurs interrogés, la mise à disposition de cette forge serait d'autant plus valorisée qu'elle permettrait aux acteurs ne s'étant jamais lancés dans l'usage des données ouvertes de mener une démarche de bout en bout, avec un accompagnement intensif et permanent, tout en étant frugal en termes de ressources. Le témoignage ci-dessous précise l'équilibre à trouver grâce à l'artefact :

“On a proposé une réunion aux collectivités pour évoquer la possibilité qu'ils ont d'être accompagné par la métropole. Alors, les réactions : le gars de XXX qui a dit : “Ils sont gentils quand ils présentent ce qu'ils veulent faire mais je tiens juste à dire qu'ils sont venus nous voir au mois de mai et qu'il ne s'est rien passé depuis. On attend qu'ils nous accompagnent pour ouvrir. Et oui, ils ont raison, on a fait un atelier de 3h pour leur demander quels sont les jeux de données qu'ils souhaitent libérer. Il se trouve que depuis on a rien fait. On a listé une dizaine de jeux de données à libérer mais nous, on n'a pas eu les moyens de les accompagner derrière pour sortir les données.” - Fonctionnaire de la métropole Grenoble-Alpes

Ceci implique que l'environnement de conception couvre l'ensemble des étapes propres à l'usage des données ouvertes, définies dans les chapitres précédents, et autonomise les groupes de réutilisateurs dans la réalisation de certaines tâches tout en minimisant les ressources humaines d'accompagnement qui devront être nécessairement sollicitées pour forger les usages.

La mise à disposition de cette forge suscitera l'intérêt des groupes d'utilisateurs si cette forge détient une ressource d'autorité particulière le sceau ou le “tampon officiel” d'une entité reconnue par les réutilisateurs :

“Je vois un exemple tout bête : une entreprise fournit une base de données. Nous on veut aller chercher les données qui concernent les communes de notre territoire. Et là, je veux casser le jeu de données par collectivité. Et là, j'ai trouvé deux sources de code qui peuvent m'aider à faire ça, autour de R (langage de programmation). Celle de #teamopendata, le mec s'appelle XXX et il appelle sa bibliothèque arielle je sais pas pourquoi. J'ai trouvé une partie de ce que je voulais mais j'irai jamais me connecter à un truc qui

s'appelle XXX.XXX. J'irai peut être si ça s'appelle data.gouv ou si il y a un tampon un minimum officiel. C'est pas un mec dans son coin, et tu sais pas ce que ça va devenir. Si il y a un tampon Jailbreak et Dataactivist, alors c'est du code officiel et ça donne un caractère officiel." - Responsable du chargé du service Open Data de la métropole Grenoble-Alpes

Conclusion

En définitive, ce scénario précise les caractéristiques des cibles finales de l'environnement de conception (les organisations qui ne se sont pas lancées dans l'open data ou celles qui les accompagnent) ainsi que le volume de charge supportable par les utilisateurs pour forger les usages. Aussi réaffirme-t-il que l'artefact doit couvrir les cinq étapes de développement des usages de bout en bout. Principalement, ce scénario précise que les effets visés par l'environnement de conception pris dans son ensemble et par l'interaction entre ces trois éléments fonctionnels, à savoir favoriser les cinq étapes de l'usage des données (Utilisation de la plateforme, Evaluation du potentiel, Transformation, Création d'un cadre cognitif, Institutionnalisation), seront sûrement réalisés si l'artefact permet de baisser les investissements nécessaires qu'une organisation doit entreprendre pour développer les usages.

5.5 Description des fonctions du design de l'environnement de conception des usages des données ouvertes.

5.5.1 Description des fonctions de l'environnement de conception pris dans son ensemble

La table ci-dessous décrit les fonctions de l'environnement de conception pris dans son ensemble. Ces fonctions se réfèrent plus particulièrement aux deux mécanismes (sparring, feedback et transformation) identifiés dans la première section de ce chapitre.

Étapes de l'usage des données	Fonction	Description de la fonction
Toutes les étapes	1) Création d'un chambre de conception dédiée à un territoire/une organisation	Les utilisateurs peuvent créer une chambre de conception dédiée à leur territoire dans laquelle les projets de développement des usages peuvent être débattus
	2) Création d'espaces de discussions dédiés à chacun des projets	Les utilisateurs peuvent créer des espaces dédiés à un projet qui accueillera quatre fils de discussion correspondant aux 4 étapes du processus de sparring
	3) Description de l'usage imaginé	Les utilisateurs peuvent librement décrire des usages qu'il souhaite mettre en place et des difficultés qu'ils rencontrent
	4) Dialogue en ligne	L'utilisateur-intervenant

		questionne l'acteur sur la base de template de questionnement linéaire qu'il adapte. L'utilisateur répond aux questions
	5) Transfert automatique des réponses	Les réponses de l'utilisateur sont transférées dans une base de données relationnelles
	6) Création de questionnaires	Les utilisateurs peuvent créer des questionnaires qui vont venir peupler la base de données relationnelles dans le but d'obtenir des feedback de la part des acteurs des autres systèmes
	7) Création de vues de la base de données relationnelles	Les utilisateurs vont pouvoir prendre acte des feedback en créant de nouvelles vues de la base de données pour montrer les évolutions du plan d'action suite aux retours

Figure 5.8 : Description des fonctions de l'environnement de conception des usages pris dans son ensemble

5.5.2 Description des fonctions du réseau d'information

La table ci-dessous décrit les fonctions de l'environnement du réseau d'information. Ces fonctions se réfèrent plus particulièrement aux trois mécanismes (surdétermination, gestion collaborative des composants du réseau et tableur de gestion des composants du réseau) identifiés dans la première section de ce chapitre.

Étapes de l'usage des données	Fonction	Description de la fonction
Utilisation des fonctionnalités de la plateformes (téléchargement/flux de données)	8) Création de chambre de discussion pour chaque plateforme open data française	Les utilisateurs-concepteurs peuvent créer des chambres de discussion pour accueillir toutes les demandes d'informations relatives aux données présentes sur les plateformes open data françaises
	9) Demande de précisions sur les données des plateformes	Les utilisateurs peuvent poser leurs questions sur des données spécifiques présentes sur les plateformes open data
	10) Mettre à jour la base de données de gestion des composants du réseau en collaboration avec d'autres membres	Les utilisateurs peuvent modifier à plusieurs la base de données de gestion des composants.
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	11) Proposition de nouvelles connaissances pour la liste des composants	Les utilisateurs peuvent proposer de nouvelles connaissances comme candidates à la liste des composants du réseau d'information
	12) Évaluation du respect des normes techniques des connaissances	Les utilisateurs-concepteurs peuvent évaluer les connaissances proposées. Ces évaluations sont transparentes et disponibles à l'ensemble du réseau
Transformation des données grâce à des artefacts techniques	13) Post de questions spécifiques à l'accomplissement des tâches spécifiques à l'usage	Les utilisateurs peuvent poser des questions pratiques quant à la réalisation de certaines tâches

		relatives à l'usage de la donnée et faire appel aux membres du réseau
	14) Evaluation du respect des normes sociales	L'utilisateur-administrateur surveillent le réseau et s'assure du bon respect des règles sociales.

Figure 5.9 : Description des fonctions du réseau d'information

5.5.3 Description des fonctions du système de gestion des tâches basé sur le modèle de méta-données

La table ci-dessous décrit les fonctions du système de gestion des tâches basé sur le modèle de métadonnées. Ces fonctions se réfèrent plus particulièrement aux trois mécanismes (structuration flexible, plasticité des dispositions spatiales et création de liens directs entre les tâches et le monde réel) identifiés dans la première section de ce chapitre.

Étapes de l'usage des données	Fonction	Description de la fonction
Utiliser les fonctionnalités de la plateforme open data	15) Télécharger le catalogue des données ouvertes d'une plateforme	L'utilisateur peut télécharger le catalogue de données ouvertes dans le système de gestion des tâches
	16) Documentation des données du catalogue	L'utilisateur peut renseigner les informations qu'il considère utile sur les données du catalogue
Evaluation du potentiel des données à partir de leur représentation	17) Télécharger les jeux de données à utiliser	L'utilisateur peut télécharger les données ouvertes dont il souhaite faire l'usage sur le système de gestion des tâches
	18) Visualiser les données ouvertes	L'utilisateur peut visualiser les données pour évaluer son potentiel

	19) Réorganisation des tables listant les tâches relatives à l'usage des données ouvertes	L'utilisateur peut réorganiser les tâches à réaliser en fonction des critères de son choix et ainsi créer un plan de développement de l'usage.
	20) Visualiser la séquence de tâches à réaliser sur les données	L'utilisateur peut visualiser de différentes manière l'ordre des tâches à effectuer sur le jeu de données
	21) Mise à disposition de témoignages de réalisation de tâches en situation réelle	L'utilisateur peut accéder à des témoignages concernant la mise en oeuvre réelle d'une tâche
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	22) Adaptation du modèle de méta-données	L'utilisateur-concepteur peut modifier le modèle de métadonnées qui structure le système de tâches en fonction d'ajouts de nouvelles pratiques dans le codebook ou de composants dans le réseau d'information
Transformation des données grâce à des artefacts techniques	21) Mise à disposition de témoignages de réalisation de tâches en situation réelle	L'utilisateur peut accéder à des témoignages concernant la mise en oeuvre réelle d'une tâche
	17) Télécharger les jeux de données à utiliser ou la liste de ces jeux	L'utilisateur peut télécharger les données ouvertes dont il souhaite faire l'usage sur le système de gestion des tâches
	24) Associer à la liste de données à utiliser une séquence de tâches	L'utilisateur peut relier les données qu'il compte utiliser à un ensemble de tâche qu'il effectue sur ces données

Figure 5.10 : Description des fonctions du système de gestion des tâches

5.5.4 Description des fonctions du manifeste et du codebook

La table ci-dessous décrit les fonctions du manifeste et du codebook. Ces fonctions se réfèrent plus particulièrement aux deux mécanismes (sourcing et intégration de nouvelles pratiques dans le codebook, développement et l'expérimentation de produits et services à partir du manifeste et du codebook) identifiés dans la première section de ce chapitre.

Étapes de l'usage des données	Fonction	Description de la fonction
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	25) Proposition de nouvelles pratiques dans le codebook	Les utilisateurs peuvent proposer l'ajout de nouvelles pratiques au sein du codebook
	26) Evaluation du respect de l'esprit du manifeste par les nouvelles pratiques proposées	Les utilisateurs-concepteurs évaluent si les nouvelles pratiques proposées correspondent à l'esprit du manifeste et communique leur choix d'intégration ou de rejet des nouvelles pratiques
	27) Déclinaison du manifeste dans un domaine de données spécifiques	Les utilisateurs peuvent proposer des déclinaisons du manifeste pour l'adapter à des domaines de données spécifiques
Institutionnalisation de l'usage des données ouvertes	28) Propositions de service et produits compatibles avec le codebook	Les utilisateurs se voient proposer et redirigés vers les services ou produits dans lequel ils peuvent injecter le codebook

Figure 5.11 : Description des fonctions du manifeste et du codebook

Conclusion du chapitre 5

L'objectif de ce chapitre était de comprendre quels sont les éléments qui constituent un dispositif en Systèmes d'Information répondant aux douze besoins fonctionnels identifiés dans le chapitre 5.

La première section m'a permis d'identifier trois éléments fonctionnels. Je conclus que pour lever les barrières qui freinent les cinq étapes de l'usage des données, il serait pertinent de mettre en œuvre trois éléments fonctionnels : un réseau d'information, une infrastructure basée sur un modèle de métadonnées ainsi qu'un manifeste et son codebook.

La deuxième section m'a permis de préciser les principes de design à respecter pour que les trois éléments fonctionnels puissent déclencher les mécanismes naturels, sociaux ou techniques qui leveront les barrières socio-techniques. J'ai établi une liste de douze principes de design classés en fonction de l'élément fonctionnel adressé.

Enfin, la troisième section m'a permis de construire le design de l'artefact qui intègre les trois éléments fonctionnels. Je conclus que le dispositif en Systèmes d'information permettant de répondre aux douze besoins fonctionnels prend la forme d'un environnement de conception des usages ou artefact de meta-design, nommé Open Data Canvas. Aussi, je conclus ce sixième chapitre par les vingt-huit fonctions techniques que l'Open Data Canvas doit contenir. Ces 28 fonctions se décomposent en quatre aires fonctionnelles.

La première aire de fonctionnelle englobe le fonctionnement de l'environnement de conception pris dans son ensemble :

Fonctions relatives à l'ensemble de l'environnement
1) Création d'un chambre de conception dédiée à un territoire/une organisation
2) Création d'espaces de discussions dédiés à chacun des projets
3) Description de l'usage imaginé
4) Dialogue en ligne
5) Transfert automatique des réponses
6) Création de questionnaires
7) Création de vues de la base de données relationnelles

Figure 5.12 : fonctions relatives à l'ensemble de l'environnement

La deuxième aire fonctionnelle regroupe toutes les fonctions du réseau d'information :

Fonctions relatives au réseau d'information
8) Création de chambre de discussion pour chaque plateforme open data française
9) Demande de précisions sur les données des plateformes
10) Mettre à jour la base de données de gestion des composants du réseau en collaboration avec d'autres membres
11) Proposition de nouvelles connaissances pour la liste des composants
12) Évaluation du respect des normes techniques des connaissances
13) Post de questions spécifiques à l'accomplissement des tâches spécifiques à l'usage

Figure 5.13 : fonctions relatives au réseau d'information

La troisième aire fonctionnelle regroupe toutes les fonctions de l'infrastructure basée sur un modèle de méta-données.

Fonctions relatives au système de gestion des tâches basée sur un modèle open data
15) Télécharger le catalogue des données ouvertes d'une plateforme
16) Documentation des données du catalogue
17) Télécharger les jeux de données à utiliser
18) Visualiser les données ouvertes
19) Réorganisation des tables listant les tâches relatives à l'usage des données ouvertes
20) Visualiser la séquence de tâches à réaliser sur les données
21) Mise à disposition de témoignages de réalisation de tâches en situation réelle
22) Adaptation du modèle de méta-données
21) Mise à disposition de témoignages de réalisation de tâches en situation réelle
17) Télécharger les jeux de données à utiliser ou la liste de ces jeux
24) Associer à la liste de données à utiliser une séquence de tâches

Figure 5.14 : fonctions relatives à un système de gestion des tâches basée sur un modèle open data

La quatrième et dernière aire fonctionnelle regroupe toutes les fonctions du manifeste et de son codebook.

Fonctions relatives au système de gestion des tâches basée sur un modèle open data
25) Proposition de nouvelles pratiques dans le codebook
26) Evaluation du respect de l'esprit du manifeste par les nouvelles pratiques proposées
27) Déclinaison du manifeste dans un domaine de données spécifiques
28) Propositions de service et produits compatibles avec le codebook

Figure 5.15 : fonctions relatives au manifeste et son codebook

Il est question d'intégrer ces vingt-huit fonctions dans trois scénarios qui précisent comment il s'insère dans le contexte réel et le quotidien des usagers des données ouvertes :

- Le premier scénario établit que l'Open Data Canvas peut être utilisé pour rallier les diverses communautés d'usagers des données ouvertes autour de piliers définissant l'usage des données ouvertes.
- Le deuxième scénario établit que l'Open Data Canvas peut être utilisé pour mutualiser et classer les ressources permettant de développer les usages des données ouvertes
- Le troisième scénario établit quant à lui que l'Open Data Canvas peut être utilisé pour mettre à disposition du plus grand nombre une forge des usages des données ouvertes.

Ces trois éléments, ces douze principes de design et ces trois scénarios d'usage n'appellent qu'une chose : la création d'un prototype de l'environnement de conception dit Open Data Canvas. Ils sont tels les plans d'un architecte qui n'attend que d'être transmis à un maître d'ouvrage. Ce sera l'objet du chapitre suivant qui donnera une forme matérielle et tangible à l'Open Data Canvas.

Chapitre 6

Prototypage d'un dispositif en Systèmes d'Information

Chapitre 6 : prototypage d'un dispositif en systèmes d'information à partir des kernel theories et permettant de faciliter les interactions homme-données

Ce septième chapitre a pour objet de donner une forme matérielle et tangible à l'environnement de conception des usages des données ouvertes dont nous avons présenté le design dans le chapitre précédent. Aussi, il sera guidé par la troisième sous-question de recherche :

Quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les situations problématiques au sein des écosystèmes open data ?

Pour rappel, j'ai détaillé dans le chapitre précédent le processus de design qui, à partir des besoins fonctionnels des opérateurs des données, m'a permis d'identifier des *kernel theories* pour guider le design d'un dispositif en systèmes d'information permettant de répondre à la question de recherche. J'ai également pu identifier dans la littérature un certain nombre d'éléments fonctionnels que le dispositif devra nécessairement intégrer pour assurer le passage de la situation problématique à la situation désirée (Carlsson, 2010, p.220). Ensemble, ces deux éléments m'ont permis de déterminer un modèle de design et de le rattacher à une classe (Offerman et al., 2010).

L'objet de ce chapitre est d'assurer le passage de la dimension conceptuelle à la dimension matérielle. Autrement dit, il s'agit de mettre en oeuvre un processus d'instanciation qui permet :

“d'implémenter des fonctionnalités dans un objet à partir d'une structure prédéfinie et pré-optimisée” (Cofer et Harding, 2006, p.120).

6.1 Objectifs du prototype

Le but de ma thèse est de trouver un mode de résolution des problèmes pernicieux qui font aujourd'hui obstacle à l'usage des données et ce dans le but de favoriser leur construction. J'ai fait le choix d'atteindre cet objectif par une méthodologie innovante issue de l'approche *design science*. Elle suppose de créer un artefact, conçu comme une instanciation de la réponse apportée à la question de recherche.

Dans le chapitre précédent, j'ai établi que cet artefact appartient à la classe des artefacts de *meta-design* et conçu le modèle de design sur la base duquel il sera prototypé. Il est question dans le chapitre présent d'éclairer le lecteur sur les différentes modalités ayant amené à prototyper cet artefact, intitulé Open Data Canvas. Aussi est-il question d'expliquer le place l'Open Data Canvas, eu égard à la résolution de ma problématique de recherche.

Cette partie vise plus précisément à expliquer en quoi le prototypage contribue à résoudre ma problématique de recherche, et d'en déduire le type de prototypage à mettre en place.

6.1.1 Les trois apports du prototype vis à vis de la résolution de ma problématique de recherche

Le prototype contribue à résoudre ma problématique de recherche de trois manières différentes :

1. Premièrement, il vise à préciser la compréhension des besoins fonctionnels auxquels un dispositif en Systèmes d'Information doit répondre pour lever les obstacles qui font face à l'usage des données ouvertes (*outputs* du chapitre 4)). En cela, il permet au chercheur-concepteur d'ajuster la loupe d'observation qu'il utilise pour concevoir et problématiser les barrières qui font face à l'usage des données ouvertes.
2. Deuxièmement, le prototype va permettre d'améliorer les interactions entre l'artefact et les futurs utilisateurs. Autrement dit, il va permettre de récupérer des *feedback* sur les facteurs qui peuvent décourager les futurs utilisateurs d'utiliser l'artefact en contexte réel. Par là, il permet au

chercheur-concepteur de maximiser ses chances de pouvoir observer si l'artefact génère les effets souhaités dans un contexte réel.

- Troisièmement, il vise à évaluer la capacité évolutive de l'artefact. Il est question alors pour le chercheur-concepteur d'établir dans quelle mesure l'artefact pourra s'adapter aux évolutions futures des besoins des utilisateurs dans le temps. En filigrane, cette étape vient renforcer la pérennité de l'artefact qui va de pair avec la validité externe des contributions finales de ma thèse.

6.1.2 Le choix d'un type de prototypage évolutif

En se référant au tableau de correspondance ci-dessous (Table 6.1), le lecteur constatera que les objectifs définis impliquent de choisir un type de prototypage particulier : le prototypage évolutif (Carey et Mason, 1983).

	Prototypage throw-it-away	Prototypage incrémental	Prototypage évolutif
Besoins fonctionnels à préciser	X	X	X
Interactions avec l'utilisateur à améliorer		X	X
Adaptabilité aux évolutions des besoins à renforcer			X

Table 6.1 : Table de correspondance entre les apports souhaités du prototype et le type de prototypage à effectuer (adapté d'Ince et Hekmatpour, 1987, p.12)

Ce type de prototypage amène à développer graduellement un prototype, de manière à ce qu'il puisse d'adapter aux changements substantiels et inévitables auxquels le chercheur-concepteur va faire face au cours de sa recherche. Parmi lesquels : modification des besoins fonctionnels ; changements ou échéances politiques, juridiques ; précision des *kernel theories* qui définissent le modèle de design ; augmentation ou diminution des moyens alloués au projet de recherche, etc.

Au cours des trois années du projet de recherche, j'ai été amené à co-construire avec la coopérative Dataactivist deux versions du prototype, correspondant à deux cycles de recherche en *design science* ainsi qu'à deux classes de design de la typologie établie par Carlsson (2006) :

1. Version 0 (2018-2019) : cartographie de l'architecture métier, fonctionnelle et applicative de l'Open Data Canvas, ce que Carlsson nomme *object design*. Ce prototype avait pour objectif spécifique de tester auprès des futurs utilisateurs le pertinence du design vis à vis de leurs besoins fonctionnels, et par là de préciser la compréhension des besoins fonctionnels.
2. Version *alpha* (2019-2020) : un prototype fonctionnel mais non-opérationnel construit à l'aide d'outils logiciels accompagné d'une méthode d'implémentation dans des contextes d'usages réels. Ce prototype avait pour objectif de tester la fiabilité technique des outils logiciels choisis et d'identifier les principaux défauts de la méthode d'implémentation.

6.2 Sélection des fonctions

Dans le chapitre précédent, j'ai dressé la liste des vingt-huit fonctions du design de l'environnement de conception des usages des données ouvertes. Sur la base de cette liste, ce sixième chapitre a pour objet de sélectionner et d'intégrer certaines d'entre elles dans des prototypes. Or, j'ai établi dans la partie précédente mon choix de développer graduellement ces prototypes. En collaboration avec la coopérative Dataactivist, j'ai co-construit trois versions du prototype. Dès lors, il sera question dans cette partie d'exposer les fonctions que j'ai décidé d'intégrer progressivement dans le prototype et qui donne leur forme matérielle à la version 0 et la version alpha du prototype (Table 6.2)

	Version 0	Version alpha
Ensemble - F1	X →	
Ensemble - F2	X →	
Ensemble - F3	X →	
Ensemble - F4	X →	
Ensemble - F5		
Ensemble - F6		
Ensemble - F7		
Réseau d'information - F8		
Réseau d'information - F9		
Réseau d'information - F10		
Réseau d'information - F11	X →	
Réseau d'information - F12	X →	
Système de gestion des tâches - F13		
Système de gestion des tâches - F14		
Système de gestion des tâches - F15		
Système de gestion des tâches - F16		
Système de gestion des tâches - F17		
Système de gestion des tâches - F18		
Système de gestion des tâches - F19		
Système de gestion des tâches - F20		
Système de gestion des tâches - F21		
Système de gestion des tâches - F22	X →	
Système de gestion des tâches - F23		
Système de gestion des tâches - F24	X →	
Manifeste et codebook - F25		
Manifeste et codebook - F26		
Manifeste et codebook - F27		
Manifeste et codebook - F28		

Table 6.2 : intégration et amélioration des fonctions dans les différentes versions du prototype

Rappelons que j'ai fait le choix méthodologique de sélectionner les fonctions selon les trois critères du pattern *Easy solution first* développés par Vaishnavi et Kuelcher (2015, p.131), décrit plus en détail dans la section 2.2.1.5.3 de cette thèse (p.190). Aussi, les différentes fonctions sélectionnées se réalisent de manière itérative, permettant ainsi à l'utilisateur d'étendre et d'améliorer son environnement de conception des usages.

6.2.1 Sélection des fonctions à intégrer dans la Version 0 du prototype

Comme expliqué dans la section précédente, la version 0 du prototype avait pour objectif spécifique de tester auprès des futurs utilisateurs le pertinence du design vis à vis de leurs besoins fonctionnels. Dans ce contexte, j'ai choisi d'intégrer dans ce premier prototype les fonctions les plus centrales de chaque élément fonctionnel.

6.2.1.1 Fonctions de l'environnement de conception des usages pris dans son ensemble

Avec la coopérative Dataactivist, j'ai intégré dans ce premier prototype les fonctions centrales de l'environnement de conception pris dans son ensemble. Ces fonctions sont les plus à même de permettre aux futurs utilisateurs de décrire et questionner les usages de données ouvertes qu'ils souhaitent développer (cf mécanisme de sparring) ainsi que leur permettre d'obtenir des *feedback* de la part d'autres systèmes de l'écosystème en question (boucle de rétroaction). Il s'agit des fonctions suivantes ci-dessous.

Étapes de l'usage des données	Fonction sélectionnée	Description de la fonction
Toutes les étapes	1) Création d'un chambre de conception dédiée à un territoire/une organisation	Les utilisateurs peuvent créer une chambre de conception dédiée à leur territoire dans laquelle les projets de développement des usages peuvent être débattus
	2) Création d'espaces de discussions dédiés à chacun des projets	Les utilisateurs peuvent créer des espaces dédiés à un projet qui accueillera quatre fils de discussion correspondant aux 4 étapes du processus de sparring
	3) Description de l'usage imaginé	Les utilisateurs peuvent librement décrire des usages qu'il souhaite mettre en place et des difficultés qu'ils rencontrent
	4) Dialogue en ligne	L'utilisateur-intervenant questionne l'acteur sur la base de template de questionnement linéaire qu'il adapte. L'utilisateur répond aux questions

Table 6.3 : Fonctions relatives à l'ensemble de l'environnement de conception sélectionnées pour intégrer la Version 0

Outre le fait qu'elles soient centrales à l'environnement de conception, j'ai choisi de les intégrer à la Version 0 du prototype dans la mesure où il existait des solutions de prototypage faciles à mettre en œuvre. En effet, pour prototyper les fonctions 1, 2, 3, 4 je pouvais déjà m'appuyer sur le forum de

discussion “#teamopendata” créée par Dataactivist, qui rassemblait en avril 2018 plus de 240 utilisateurs autour des questions relatives à l’ouverture des données ouvertes et le développement de ses usages.

Ce forum, basé sur un logiciel libre (et donc entièrement paramétrable), possède des fonctionnalités permettant de créer des chambres dédiées à un territoire ou une organisation ainsi que des espaces de discussions privés ou publics (voir prochaine partie).

6.2.1.2 Fonctions du réseau d’information

J’ai choisi de présenter les fonctions les plus centrales du réseau d’information, dans la mesure où elles assurent l’essentiel du comportement du réseau d’information, à savoir faire circuler des informations permettant aux acteurs de renforcer leurs connaissances sur l’accomplissement de tâches spécifiques et fixer un cadre commun adaptable et reproductible qui spécifie un ensemble de bonnes pratiques explicites et implicites à appliquer au sein et hors du réseau. Il s’agit des fonctions ci-dessous.

Étapes de l’usage des données	Fonction sélectionnée	Description de la fonction
Création d’un cadre cognitif partagé entre les systèmes pour les coordonner	11) Proposition de nouvelles connaissances pour la liste des composants	Les utilisateurs peuvent proposer de nouvelles connaissances comme candidates à la liste des composants du réseau d’information
	12) Évaluation du respect des normes techniques des connaissances	Les utilisateurs-concepteurs peuvent évaluer les connaissances proposées. Ces évaluations sont transparentes et disponibles à l’ensemble du réseau

Table 6.4 : Fonctions relatives au réseau d’information sélectionnées pour intégrer la Version 0

Outre le fait qu’elles soient centrales au réseau d’information, j’ai choisi de les intégrer à la Version 0 du prototype dans la mesure où il existait des solutions de prototypage faciles à mettre en œuvre. En effet, pour prototyper les fonctions 11 et 12, je pouvais déjà m’appuyer sur une base de données collaborative

créée par Dataactivist (et du questionnaire permettant directement, qui rassemblait déjà plus de 50 connaissances pratiques (les composants du cadre commun) provenant d'un grand nombre de communautés et organisations engagées dans le développement des usages des données ouvertes (voir prochaine partie).

6.2.1.3 Fonctions du système de gestion des tâches basé sur un modèle de méta-données

À nouveau, j'ai choisi les fonctions les plus centrales du système de gestion des tâches, dans la mesure où elles assurent l'essentiel son comportement, à savoir organiser les métadonnées et permettre aux utilisateurs d'obtenir des séquences de tâches à appliquer pour transformer les données ouvertes. Il s'agit des fonctions ci-dessous.

Étapes de l'usage des données	Fonction	Description de la fonction
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	22) Adaptation du modèle de méta-données	L'utilisateur-concepteur peut modifier le modèle de méta-données qui structure le système de tâches en fonction d'ajouts de nouvelles pratiques dans le codebook ou de composants dans le réseau d'information
Transformation des données grâce à des artefacts techniques	24) Associer à la liste de données à utiliser une séquence de tâches	L'utilisateur peut relier les données qu'il compte utiliser à un ensemble de tâche qu'il effectuera sur ces données

Table 6.5 : Fonctions relatives au système de gestion des tâches sélectionnées pour intégrer la Version 0

6.2.1.4 Fonctions du manifeste et de son codebook

Dans la Version 0, aucune des fonctions définies dans le chapitre précédent n'a été sélectionnée. Il a simplement été question de prototyper une première version du manifeste et de son codebook (voir prochaine partie).

6.2.2 Sélection des fonctions à intégrer dans la Version alpha du prototype

Comme expliqué dans la section précédente, la version *alpha* du prototype avait pour objectif spécifique de tester la fiabilité technique des outils logiciels choisis et d'identifier les principaux défauts de la méthode d'implémentation. Dans ce contexte, et au regard des moyens limités dont je disposais, j'ai choisi d'intégrer dans ce deuxième prototype les mêmes fonctions que celles déjà présentes dans le premier. La différence du deuxième prototype par rapport au premier se situe dans sa forme, du fait qu'il s'agit d'un dispositif en Systèmes d'Information tangible et non d'un *objet design*, intangible par définition.

Enfin, à la suite de l'évaluation du premier prototype, le prototypage du manifeste et de son codebook a été abandonné dans la mesure où leurs capacités à répondre aux besoins fonctionnels n'ont pas pu être prouvées.

6.3 Construction du prototype

Dans la partie précédente, j'ai établi la liste des fonctions à intégrer progressivement dans le prototype, qui comporte trois versions : la version 0 et la version alpha. Ces trois versions ont été co-construites tout au long des trois années du projet de recherche par le chercheur-concepteur avec l'aide de la coopérative Dataactivit.

Dans cette partie, je fournis dans un premier temps une description de l'architecture des trois versions construites en suivant le pattern des *building blocks* (Vaishnavi et Kuelcher, 2015, p.138), décrit plus en détail dans la section 2.2.1.1 de cette thèse (p.191). Dans un second temps, je présente des scénarios d'utilisabilité détaillant le fonctionnement du prototype. Je présenterai dans ces scénarios les interfaces qui permettent de se rendre compte de l'aspect visuel des fonctions et du point de vue de l'utilisateur lorsqu'il les utilise.

6.3.1 Construction de la version 0 du prototype

La version 0 du prototype est une cartographie de l'architecture métier, fonctionnelle et applicative de l'Open Data Canvas.

6.3.1.1 Architecture de la version 0

6.3.1.1.1 Cartographie de l'architecture métier de l'Open Data Canvas

La cartographie de l'architecture métier décrit, sous forme de diagramme, les processus que l'Open Data Canvas vient supporter. Chaque processus se subdivise en un enchaînement d'activités représenté par des diagrammes de *workflow*. Pour construire ces diagrammes et l'ensemble de l'architecture, j'ai utilisé la méthode de modélisation de processus *Processus d'entreprise et notation*¹. J'ai choisi cette méthode de modélisation pour sa capacité à simplifier la représentation des interactions complexes entre les acteurs parties prenantes d'un processus (Dijkman et al., 2008) . En effet, elle a pour avantage d'être compréhensible par tous les lecteurs de cette thèse mais aussi par tous les futurs utilisateurs du Canvas, quel que soit leur profil (analyste métier, développeur, consultant, etc.)

En collaboration avec Dataactivist, j'ai choisi de représenter les trois groupes d'utilisateurs du Canvas représenté par les trois rectangles (voir figure ci-dessous). En notation BPMN, on parle de *piscine* :

- **l'organisation meta-designer** : organisation cherchant à lancer le développement des usages des données ouvertes dans son écosystème designer en mettant à disposition l'Open Data Canvas, conçu ici comme objet facilitant le design des usages à partir des données ouvertes. Les membres de l'organisation meta-designer en charge de déployer le Canvas sont les utilisateurs-concepteurs du Canvas. Parmi les exemples d'organisation meta-designer : les collectivités supra-communales (comme la métropole de Grenoble ou de Lyon), les associations à l'échelle nationale ou les grands groupes.
- **l'écosystème designer** : ensemble de producteurs, d'intermédiaires et de réutilisateurs des données ouvertes coordonnés par l'organisation meta-designer et qui seront les utilisateurs finaux du Canvas. Ce triptyque d'acteurs est représenté en notation BPMN par les trois *lignes d'eau* de la *piscine* de l'écosystème designer. Parmi les exemples de membres d'un écosystème designer, sont présentes des communes, des divisions opérationnelles de grands groupes (en tant que productrice ou réutilisatrices de données ouvertes) ou des Directions des Systèmes d'Information de toute organisation (en tant qu'intermédiaire). Il peut s'agir par exemple des communes rattachées à la métropole de Lyon ou celles rattachées à la métropole de Grenoble.

¹ Termes originaux : "*Business Process and Notation*"(BPMN)

- **l'organisation d'un dispositif en SI** : organisation en charge de la mise à disposition, du maintien et de la mise à jour de l'Open Data Canvas. Il peut intervenir passivement auprès de l'organisation meta-designer en se contentant de lui indiquer où se trouve la documentation du Canvas détaillant les différentes étapes de déploiement du Canvas ; il peut également intervenir activement en accompagnant l'organisation meta-designer tout au long de la mise en place, du déploiement du Canvas et du développement des usages. Il peut s'agir par exemple d'une société coopérative comme Dataactivist, agissant traditionnellement comme prestataire des collectivités ou de réseaux de réseaux d'innovation à l'image de la French Tech in the Alps, qui accompagne les collectivités dans la transformation numérique.

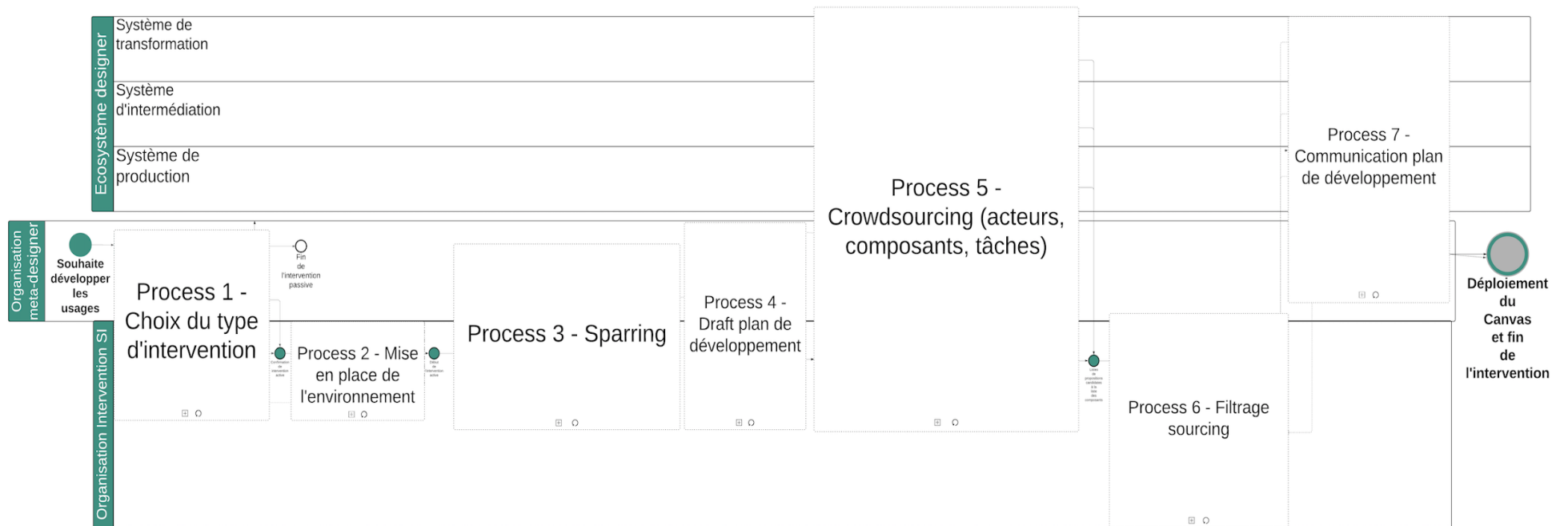


Figure 6.1 : cartographie de l'architecture métier de la version 0 du prototype de l'Open Data Canvas

Au sein des trois piscines, il existe sept séquences d'activités mises en oeuvre par chacun des trois groupes d'utilisateurs. Contrairement à ce que la représentation linéaire suggère, cette séquence d'activité est itérative et circulaire. En effet, il faut rappeler au lecteur que le Canvas s'insère dans un écosystème *open data*, où les modifications d'un seul élément entraîne des effets sur l'ensemble de l'écosystème (Nardi et O'Day, 1999). De fait, tout processus est nécessairement itératif et circulaire.

6.3.1.1.2 Cartographie de l'architecture fonctionnelle de l'Open Data Canvas

La cartographie de l'architecture fonctionnelle de l'Open Data Canvas décrit les interactions entre les fonctions sélectionnées dans la partie précédente. Elle a pour principale fonction de décrire les fonctions qui permettent la mise en place des sept processus métiers mentionnés ci-dessous.

Elle a été conçue à partir de l'architecture métier et plus précisément en regroupant les objets métiers (symbolisés par les feuilles pliées) par affinité. Autrement dit, j'ai regroupé les données qu'exploitent le Canvas dans des zones fonctionnelles. Son formalisme respecte quant à lui celui d'un Système d'Information urbanisé avec des zones fonctionnelles pré-définies comme la zone des "données et référentiels" ou encore celle des "échanges" entre zones fonctionnelles.

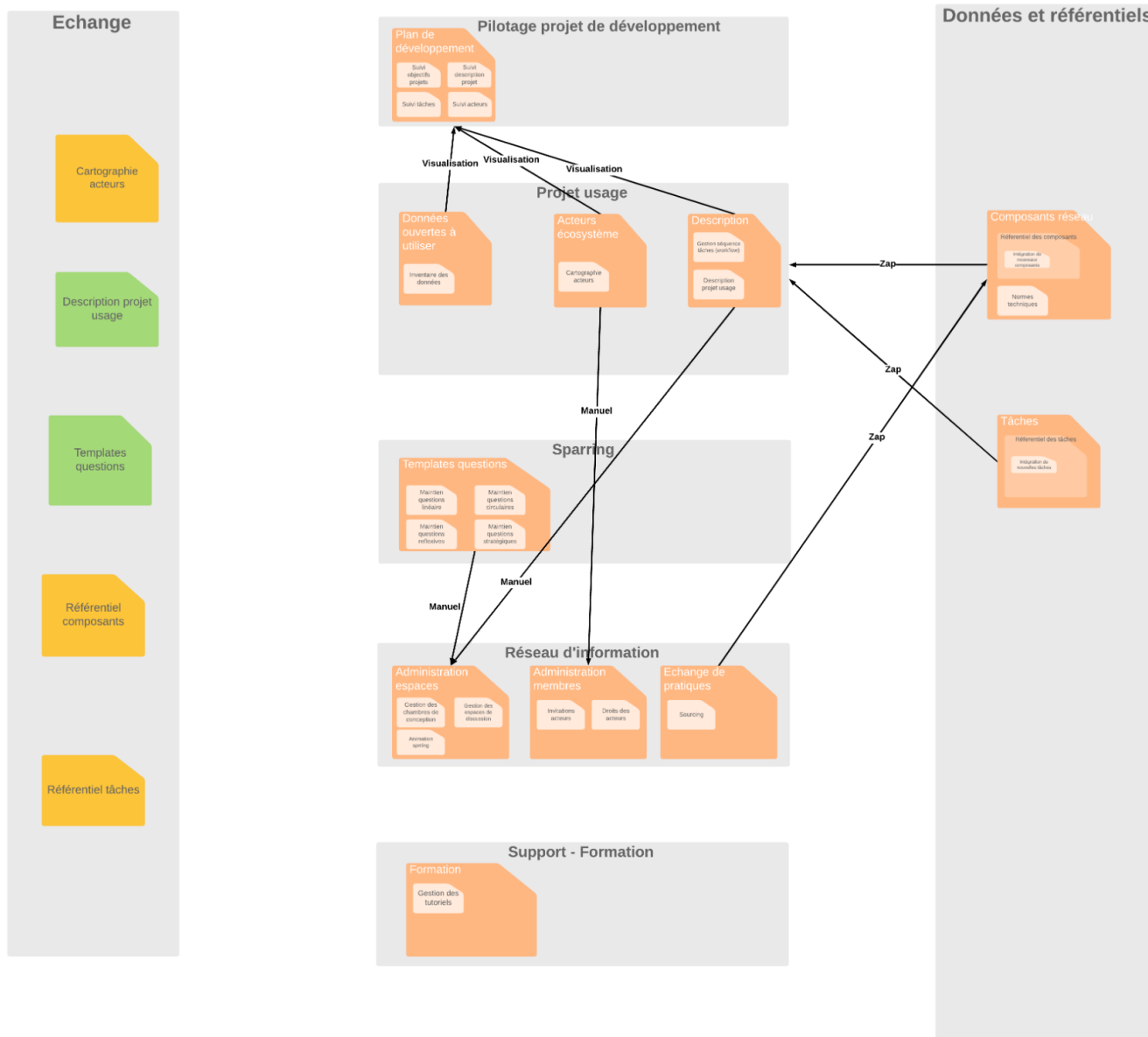


Figure 6.2: cartographie de l'architecture fonctionnelle de la version 0 du prototype de l'Open Data Canvas

Il convient de décrire cette cartographie en suivant l'ordre des sept processus métiers :

1. Choix du type d'intervention
2. Mise en place de l'environnement
3. Sparring
4. Draft du plan de développement
5. Crowdsourcing (acteurs, composants, tâches)
6. Filtrage sourcing
7. Communication du plan de développement

Pour supporter le processus du choix du type d'intervention (processus 1), le Canvas permet de stocker la description des projets de développement des usages des données ouvertes fourni par l'organisation meta-designer. En fonction de cette description, l'intervenant en SI (si choix d'une intervention active) ou l'organisation meta-designer (si choix d'une intervention passive) créé dans le réseau d'information du Canvas la chambre de conception et les espaces de discussion (un espace par projet) (processus 2).

Alors, en cas d'intervention active, l'intervenant en SI peut initier le processus de sparring (processus 3), et challenger les projets de développement des usages décrits par l'organisation. Pour cela, l'intervenant s'appuie sur les templates de questions qui s'affichent automatiquement dans chacun des espaces de discussion. En fonction des réponses apportées par l'organisation meta-designer, l'intervenant décide d'avancer dans le processus de sparring.

Une fois le processus de sparring terminé, l'organisation meta-designer est en mesure de préciser : l'inventaire des données à utiliser, les objectifs précis du projet, et la cartographie d'acteurs, formant un écosystème, à solliciter (processus 4). Le Canvas permet ensuite d'associer à la liste de données et des objectifs fixés, des séquence de tâches. La liste et description de tâches à la disposition de l'organisation meta-designer est issue d'un référentiel de tâches que la coopérative Dataactivist considère comme récurrente dans les projets de développement des usages des données. Aussi le Canvas permet-il de relier ces tâches aux "composants du réseau d'information", c'est-à-dire à des connaissances et bonnes pratiques normées. Ces connaissances et bonnes pratiques prennent la forme d'outils libre d'utilisation et qui permette de mettre en place les tâches identifiées. L'ensemble de ces éléments permet de rédiger un premier *draft* du plan de développement et ainsi de conclure le quatrième processus.

Le cinquième processus peut se mettre en œuvre grâce à une fonction du réseau d'information du Canvas qui permet d'inviter les acteurs de l'écosystème cartographié à rejoindre le réseau d'information du

Canvas, et octroie différents droits en fonction de leur rôle dans le projet. En parallèle, l'intervenant en Systèmes d'Information va aussi solliciter l'écosystème et faire un appel à contribution pour identifier d'éventuels composants qui permettraient de réaliser certaines tâches listées dans le plan de développement. À la fin du processus, le Canvas est peuplé des acteurs invités et doit traiter toutes les connaissances proposées pour répondre à l'appel à contribution.

C'est alors que le Canvas offre la possibilité de filtrer les propositions (processus 6). Il met à disposition un *workflow* qui permet de trier les composants "candidats", en fonction des normes techniques fixées par l'intervenant en SI (se référer au chapitre 5, section 5.3.2.1.1 (p.378) pour une description des normes). L'intervenant peut ensuite actualiser le référentiel de composants et le plan de développement avec les composants sélectionnées.

Enfin, la mise en place du septième processus est assurée en diffusant le référentiel de composants actualisés sur le réseau d'information du Canvas et le plan de développement actualisé dans l'espace de discussion de l'organisation meta-designer.

6.3.1.1.3 Cartographie de l'architecture applicative de l'Open Data Canvas

Enfin, l'architecture applicative décrit les blocs applicatifs, la structure des données utilisées et leurs interactions, qui correspondent directement à l'organisation fonctionnelle décrite dans l'architecture ci-dessous.

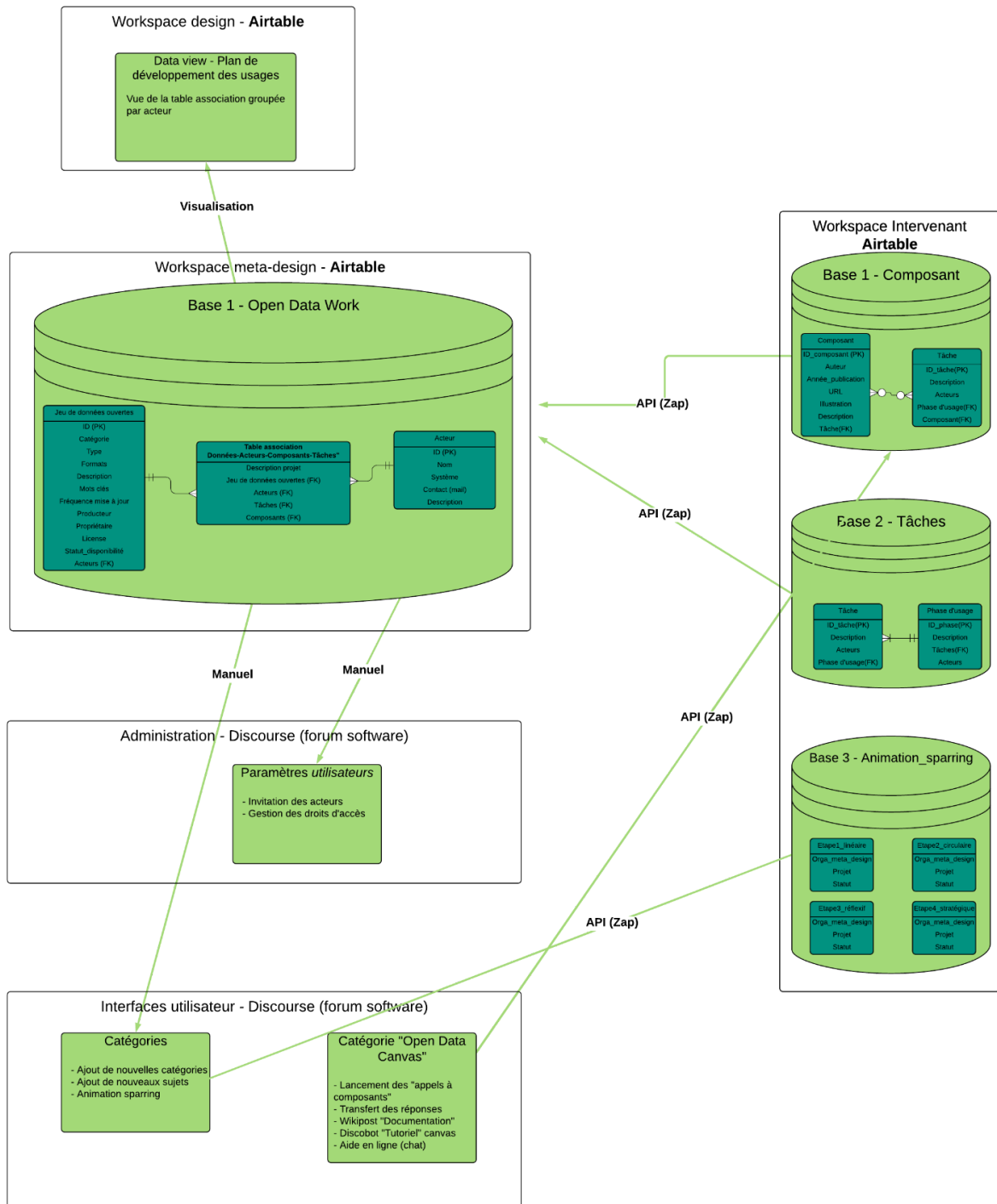


Figure 6.3 : cartographie de l'architecture applicative de la version 0 du prototype de l'Open Data Canvas

Il convient brièvement de rappeler que dans cette version 0, l'Open Data Canvas fait l'utilisation de deux outils logiciels : Airtable, une base de données en ligne collaborative en ligne, et Discourse, un logiciel de forum open-source. Les échanges entre les blocs applicatifs sont assurés soit manuellement par l'intervenant en SI, soit par un troisième outil logiciel, Zapier qui permet de configurer des ponts techniques (API) entre des applications.

6.3.1.2 Scénario d'utilisabilité de la version 0

Pour cette version 0 du prototype, le scénario d'utilisabilité est simple. J'ai utilisé comme support lors de *focus group* avec des collectivités supra-communales visant à évaluer la pertinence du design du Canvas vis à vis de leurs besoins fonctionnels (voir chapitre 7, Table 7.5, p.472 pour le détail des focus group réalisés ; voir chapitre 2, section 2.2.1.6 pour le détail de la méthodologie utilisée). Le lecteur trouvera au chapitre suivant une description détaillée de la manière dont le prototype a été utilisé pour atteindre cette fin.

6.3.2 Construction de la version alpha du prototype

La version *alpha* du prototype est un dispositif en Systèmes d'Information construit à partir de trois outils logiciels :

- Discourse, un logiciel libre de conception de forum de discussion. Il s'agira de prototyper la version alpha directement sur le domaine #teamopendata qui appartient à Dataactivist et qui rassemble déjà plus de 800 experts des données ouvertes.
- Airtable, application permettant de designer des bases de données relationnelles à partir de tableurs
- Zapier, application permettant de faire communiquer des logiciels et applications web, en l'occurrence Discourse et Airtable, par l'intermédiaire de la configuration d'API.

6.3.2.1 Architecture de la version alpha du prototype

L'architecture de la version alpha a été décrite dans la cartographie de l'architecture fonctionnelle et applicative ci-dessus (Figures 6.2 et 6.3).

6.3.2.2 Scénario d'utilisabilité de la version alpha

Je décrirai l'utilisation de la version alpha du prototype en me basant sur un exemple fictif d'organisation souhaitant développer l'usage des données ouvertes sur son territoire. Il suivra l'ordre des processus présentés dans la cartographie de l'architecture métier de l'Open Data Canvas (Figure 6.1).

6.3.2.2.1 Processus 1 : Choix du type d'intervention

Dans le cadre du déploiement de sa stratégie "Smartcity", une métropole, nommée "MetroMeta", souhaite développer l'usage des données ouvertes pour favoriser l'attractivité durable d'un de ses quartiers. À cet égard, elle investit des moyens pour développer des services et des applications numériques qui, sur la base de données ouvertes, permettent de favoriser l'attractivité économique de ce quartier tout en prenant en compte les contraintes liées à la transition énergétique et à la crise écologique.

à ce titre, pour disposer d'un accompagnement méthodologique, MetroMeta contacte l'intervenant en Systèmes d'Information, nommé "Intervenant". MetroMeta lui transmet les informations relatives aux deux projets de développement des usages qu'elle souhaite mettre en oeuvre :

1. Projet "Ecoquartier", qui vise à augmenter la production et la consommation d'énergie durables (solaire, éolien, etc.) des bâtiments du quartier.
2. Projet "SmartNeighbor", qui vise à fournir aux commerçants un maximum d'informations recelant un potentiel commercial (trafic routier et piéton dans le environs, travaux publics à venir sur les voiries, nouvelles infrastructures en cours de construction). Ce "comptoir de données" permet également aux commerçants de consulter des analyses des retours laissés par leurs clients sur les réseaux sociaux.

L'Intervenant lui envoie une proposition d'intervention, qui précise qu'il existe deux modalités d'interventions possibles (intervention passive/intervention active). Dans cet exemple, je pars du principe que MetroMeta choisit l'intervention active.

6.3.2.2 Processus 2 : Mise en place de l'environnement

Suite à la confirmation du choix de l'intervention active de la part de MetroMeta, l'Intervenant crée une chambre de conception "MetroMeta Smart-City" sur le logiciel Discourse, dotée de deux espaces de discussion. Le premier espace est dédié au projet "Ecoquartier", le deuxième au projet "SmartNeighbor".

Une fois la chambre et les espaces de discussion créés, MetroMeta est invité à fournir une description des deux projets "Ecoquartier" et "SmartNeighbor" ainsi que de la feuille de route SmartCity dans laquelle s'insère les deux projets (interface ci-dessous).

{#teamopendata} 🔍 ☰ 👤

Bienvenue dans la chambre de conception de la métropole MétroMeta
Description de la feuille de route Smart City

SMART CITY

🔒 MetroMeta_SmartCity ▾ all ▾ **Latest** Top 🔧 Edit + New Topic 🔔



🔒 Eco-quartier

Description du projet Eco-quartier



🔒 SmartNeighbor

Description du projet SmartNeighbor

☰ Topic	Replies	Views	Activity
<p>🚩 A propos de la feuille de route SmartCity</p> <p>🔒 MetroMeta_SmartCity</p> <p>Bienvenue dans la chambre de conception de la métropole MétroMeta Description de la feuille de route Smart City</p>	3	2	4d

Figure 6.4 : interface d'accueil de la chambre de conception de MetroMeta.

6.3.2.2.3 Processus 3 : Sparring

Dans chacun des deux espaces de discussion, MetroMeta va dialoguer avec l'Intervenant dans le but d'identifier tous les éléments nécessaires à la construction d'un plan de développement des usages, à savoir : - une liste d'objectifs et de sous-objectifs les plus précis possibles, traduisibles en sous-ensembles de tâches concrètes - la cartographie des acteurs essentiels à la mise en oeuvre de chacun des projets - un premier inventaire des données ouvertes dont il serait question de faire usage et/ou des données qu'il s'agirait d'acquérir.

Ce dialogue se déroule selon les quatre phases du processus de sparring désigné (voir chapitre 5, section 5.3.1.1), chacune des phases étant menée dans un fil de discussion indépendant (interface ci-dessous).

Topic	Replies	Views	Activity
★ A props du projet Eco-quartier	2	3	4d
🔗 Etape 1 : questionnaire linéaire du processus de sparring	1	2	4d
🔗 Etape 2 : questionnaire circulaire du processus de sparring	4	1	4d
🔗 Etape 3 : questionnaire réflexif du processus de sparring	2	1	4d
🔗 Etape 4 : questionnaire stratégique du processus de sparring	2	1	4d
Plan de développement du projet	0	1	4d

Figure 6.5: interface de l'espace de discussion du projet Eco-quartier contenant les quatre fils de discussion du processus de sparring

En fonction de la description des projets fournie par MetroMeta, l'Intervenant va adapter les templates des quatre types de questions qu'il posera à MetroMeta. Il effectue ce travail directement sur une base de données relationnelle faisant partie de son espace de travail Airtable (rectangle noir sur l'interface ci-dessous).

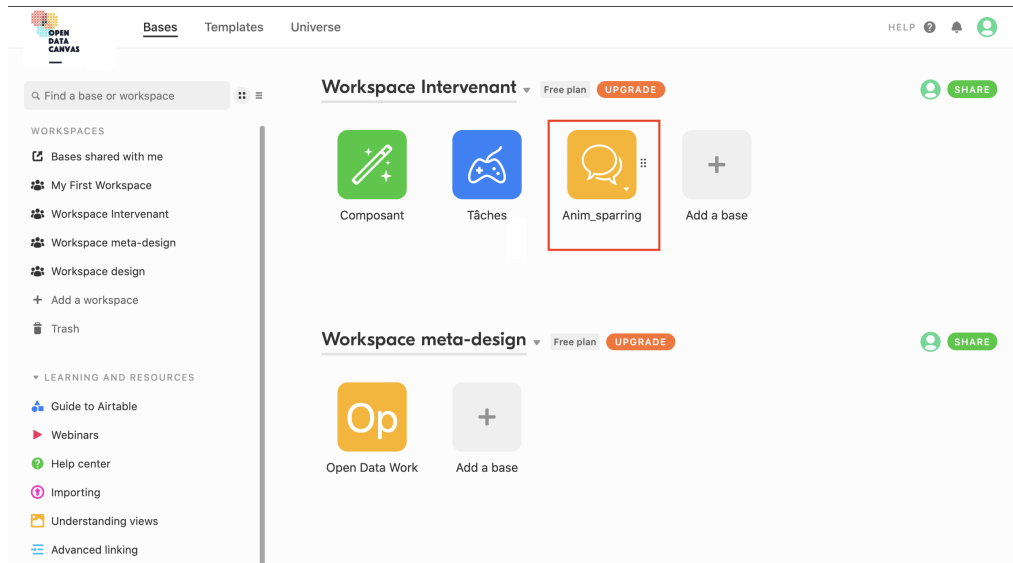


Figure 6.6 : interface de l'espace de travail (workspace) de l'Intervenant et de la base lui permettant d'animer le sparring avec MetroMeta

Dans cette base, l'Intervenant trouve les templates des quatres types de questions (linéaires, circulaires, réflexives et stratégiques), disposés dans quatre tables différentes. Dans la table correspondant à l'étape à laquelle se situe MetroMeta, l'Intervenant ajoute une ligne dans laquelle il adapte le template des questions (figure 6.7 ci-dessous).

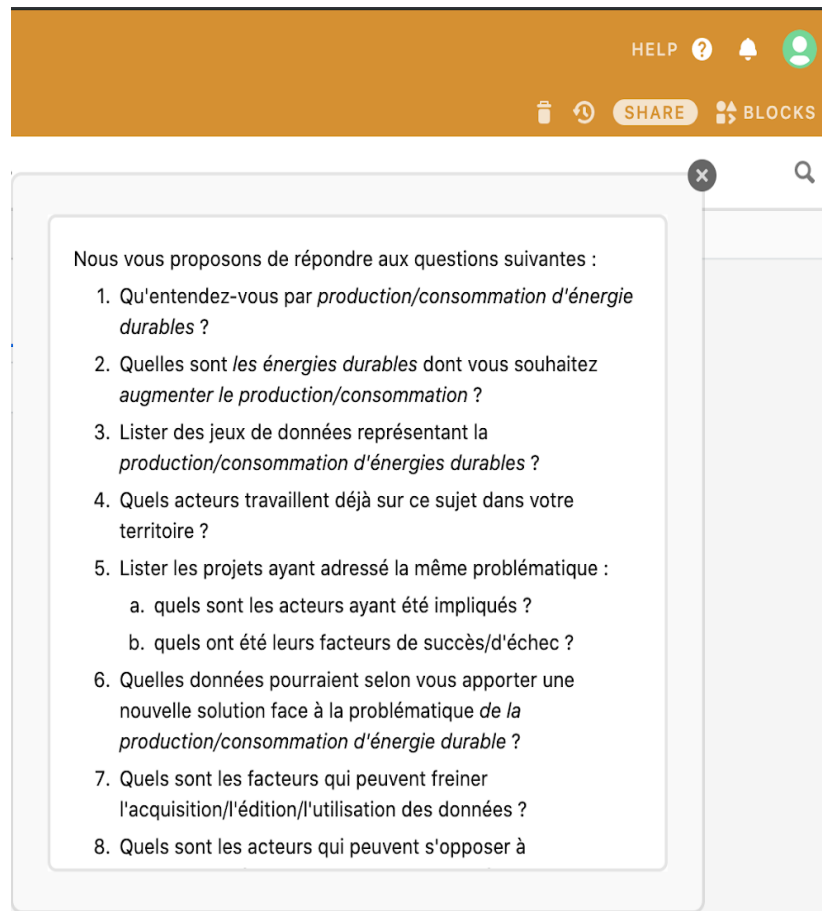


Figure 6.7 : interface de la table d'animation de l'étape 1 du processus de sparring avec zoom sur le template de question linéaire pour le projet Eco-quartier

Le fait d'ajouter une ligne dans une des tables déclenche automatiquement une API, configurée sur Zapier, qui transfère la question de l'espace de travail de l'Intervenant vers le fil de discussion lié. Ainsi une nouvelle ligne ajoutée dans la table "Etape1_lineaire" permet de transférer les questions vers le fil de discussion "Etape 1 : questionnement linéaire du processus de sparring" dans l'espace de discussion du projet Eco-quartier.

🔗 Etape 1 : questionnement linéaire du processus de sparring ✎

■ MetroMeta_SmartCity ■ Eco-quartier



ArthurSz participants_HdT

Aug '20

Vous trouverez dans ce fil de discussion le dialogue entre MetroData et l'Intervenant correspondant à la première étape du process de sparring

🔗 ✎ ... ↩ Reply

created	last reply	2	25	1		▼
Aug '20	Aug '20	replies	views	user		



ArthurSz participants_HdT

1 ✎ Aug '20

Nous vous proposons de répondre aux questions suivantes :

1. Qu'entendez-vous par *production/consommation d'énergie durable* ?
2. Quelles sont *les énergies renouvelables* dont vous souhaitez augmenter *la production et la consommation* ?
3. Lister des jeux de données représentant la *production/consommation d'énergie renouvelable* ?
4. Quels acteurs travaillent déjà sur ce sujet dans votre territoire ?
5. Lister les projets ayant adressé la même problématique :
 1. quels sont les acteurs ayant été impliqués ?
 2. quels ont été leurs facteurs de succès/d'échec ?
6. Quelles données pourraient selon vous apporter une nouvelle solution face à la problématique *de la production et consommation d'énergie renouvelable* ?
7. Quels sont les facteurs qui peuvent freiner l'acquisition/l'édition/l'utilisation des données ?
8. Quels sont les acteurs qui peuvent s'opposer à l'acquisition/l'édition/l'utilisation des données ?
9. Qui sont les utilisateurs finaux de votre solution ? Quel est leur besoin ?
10. Qu'est-ce qui vous permet de croire que les données ouvertes vont permettre de répondre aux besoins des utilisateurs finaux ?

Figure 6.8 : Interface du fil de discussion de l'étape 1 du processus de sparring pour le projet Eco-quartier après l'ajout d'une nouvelle ligne dans la table d'animation

À la fin de cette première étape, l'Intervenant et MetroMeta se mettent d'accord sur une série de problèmes clairs auxquels il s'agira de répondre lors du projet. La formulation de ces problèmes clôt la première étape du processus de sparring.

Dans notre exemple, le problème peut être de savoir comment utiliser l'open data pour encourager la production/consommation d'énergies durables, sachant que les données disponibles sont à un niveau de granularité qui ne permet pas de faire des analyses à l'échelle d'un individu ou d'une entreprise.

La deuxième étape du processus de sparring démarre alors. Le prototype fonctionne de la même manière que pour la première étape, à la différence près que les questions ne visent plus à formuler des problèmes clairs mais des dilemmes entre des éléments contradictoires au sein du projet. Pour cela, l'Intervenant pose des questions circulaires qui l'invite à voir les problèmes formulés lors de la première étape du point de vue des acteurs présents dans les autres systèmes de l'écosystème *open data* (voir interface ci-dessous). Dans le cas fictif, je pars du principe que MetroMeta est avant tout un producteur de données. Les questions de l'Intervenant l'invite donc à se mettre dans la peau d'intermédiaire ou de réutilisateurs de données.

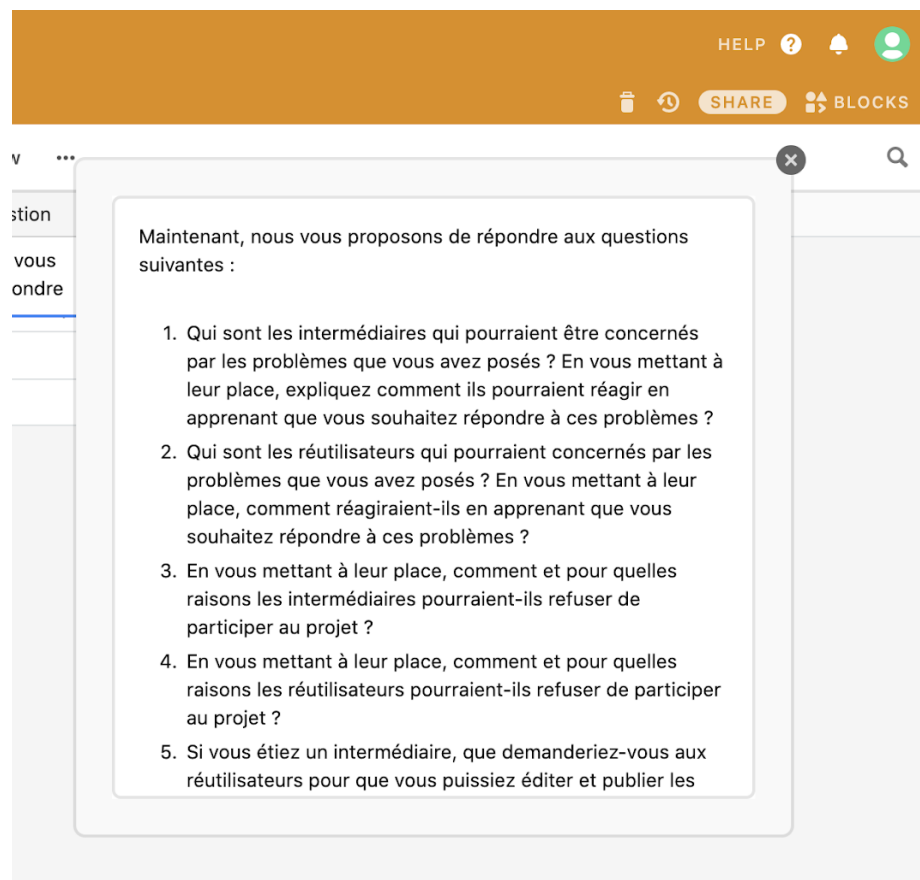


Figure 6.9 : Interface de la table d'animation de l'étape 2 du processus de sparring avec zoom sur le template de questions circulaires

À la fin de cette deuxième étape, l'Intervenant et MetroMeta formulent ensemble une série de dilemme et peuvent clôturer la deuxième étape du processus de sparring.

Dans notre exemple, un des dilemmes peut être de savoir comment répondre à la fois au besoin des réutilisateurs de données (développeurs, citoyens, etc.) et des producteurs de données (producteurs, fournisseurs d'énergie). Alors que les premiers cherchent à créer des applications qui puissent sensibiliser les individus/entreprises à la production/consommation d'énergies renouvelables ; les seconds ne peuvent mettre à disposition des données aussi précises, au risque de dévoiler des données personnelles et de trahir la confiance de leurs clients ?

La troisième étape du processus de sparring démarre alors. L'intervenant pose désormais des questions réflexives. Il invite MetroMeta à formuler des solutions aux dilemmes mentionnés dans l'étape précédente, puis à se positionner en observateur extérieur pour pouvoir observer les effets de ces solutions. Pour cela, l'Intervenant pose des questions réflexives (voir interface ci-dessous)

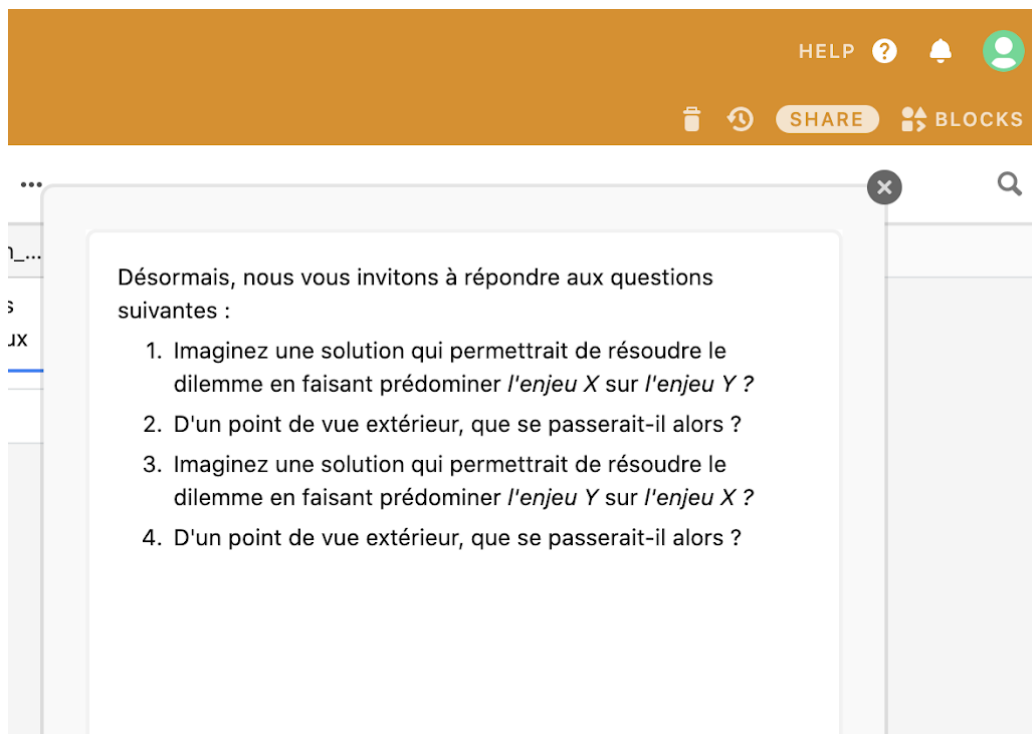


Figure 6.10 : interface de la table d'animation de l'étape 3 du processus de sparring avec zoom sur le template de questions réflexives

À la fin de cette troisième étape, l'Intervenant et MetroMeta formulent ensemble les paradoxes auxquels le projet répondra.

La quatrième et dernière étape du processus de Sparring commence alors. L'intervenant va alors poser des questions stratégiques. Il va demander à MetroMeta d'imaginer les objectifs à atteindre et les tâches et actions à réaliser pour répondre aux paradoxes formulés (interface ci-dessous). Aussi demandera-t-il à MetroMeta de s'assurer que celles-ci n'attisent pas la contradiction entre les éléments contradictoires des paradoxes. Les tâches et actions qui rentrent dans ce cas seront intégrées dans le plan de développement des usages.

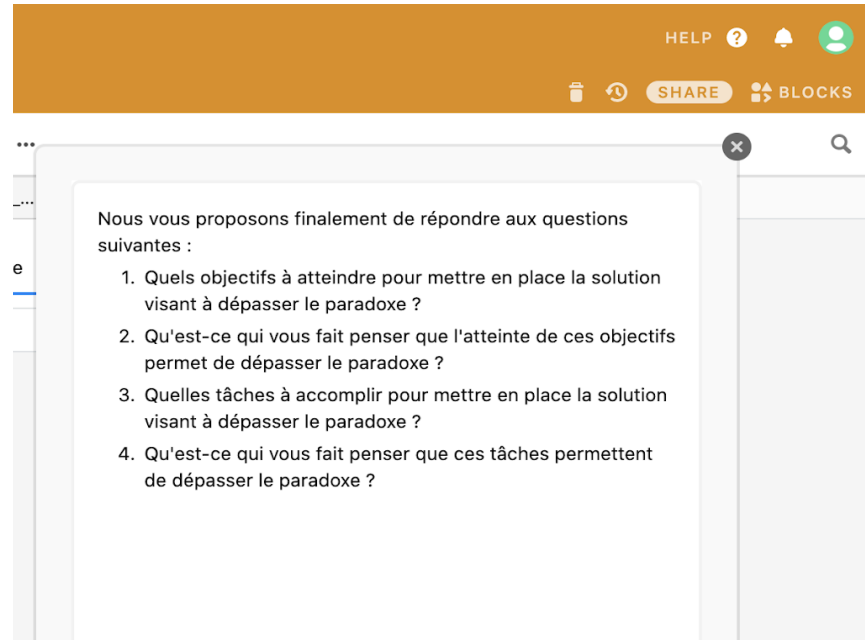


Figure 6.11 : interface de la table d'animation de l'étape 4 du processus de sparring avec zoom sur le template de questions stratégiques

À la fin de cette troisième étape, l'Intervenant et MetroMeta seront en mesure de formuler une liste d'objectifs à atteindre et de tâches à accomplir. Ces éléments seront utilisés pour créer une première version du plan de développement des usages.

6.3.2.2.4 Processus 4 : Draft du plan de développement

Grâce au processus de sparring, l'Intervenant et MetroMeta ont pu identifier tous les éléments nécessaires à la construction d'un plan de développement des usages, à savoir : - une liste d'objectifs et de sous-objectifs les plus précis possibles, traduisibles en sous-ensembles de tâches concrètes - la cartographie des acteurs essentiels à la mise en oeuvre de chacun des projets - un premier inventaire des données ouvertes dont il serait question de faire usage et/ou des données qu'il s'agirait d'acquérir.

L'Intervenante met alors à disposition de MetroMeta un espace de travail de travail (*workspace*) et d'une base de données ouverte dite "Open Data Work" (voir interface ci-dessous) dans laquelle elle va pouvoir construire un plan de développement des usages à partir des éléments mentionnés ci-dessus.

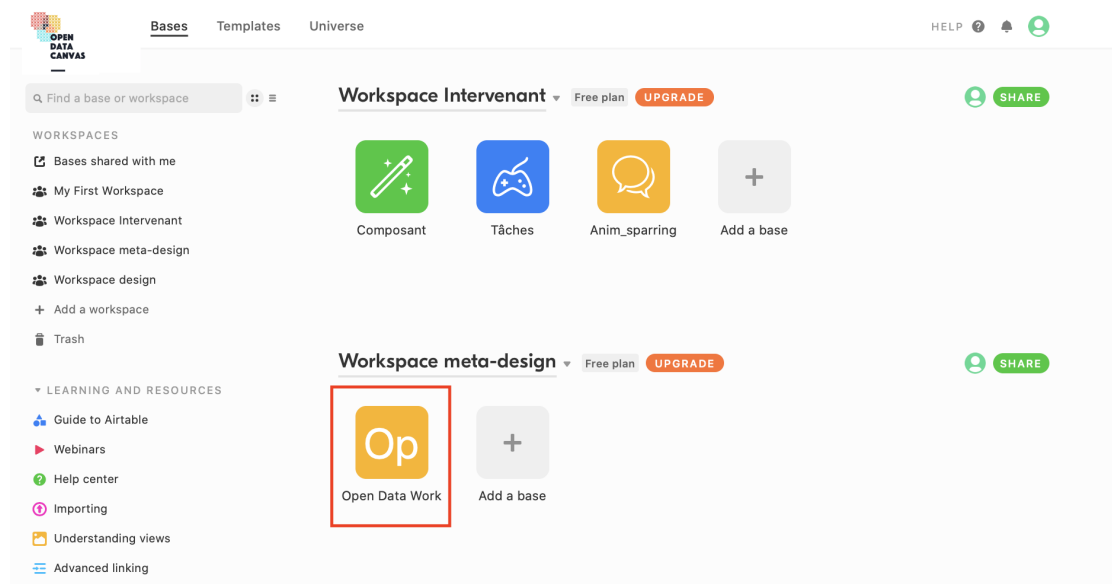


Figure 6.12 : Interface de l'espace de travail (*workspace*) de MetroMeta et de la base lui permettant de construire un plan de développement des usages

Dans cette base seront consignés, dans les trois premières tables (interface ci-dessous) : - les données ouvertes documentées dont il serait question de faire usage et/ou des données qu'il s'agirait d'acquérir (Table "Donnees_ouvertes"). Dans notre cas, le processus de sparring nous a permis de recenser deux jeux de données : le cadastre solaire, qui recense et cartographie le potentiel de production d'énergie solaire de tous les bâtiments sur le territoire de MetroMeta ; le diagnostic DPE, qui recense la consommation énergétique de tous les bâtiments sur le sol français.

- la cartographie des acteurs à impliquer dans le projet (Table "Acteurs"). En l'occurrence, j'ai identifié trois acteurs : la Direction de l'Innovation et du Numérique, qui produit les données du cadastre solaire ; le laboratoire d'innovation qui pilote une partie du projet Ecoquartier ; et Ekwater qui est une start-up fournisseur d'énergie durable sur le territoire de MetroMeta.
- les tâches concrètes à réaliser pour atteindre les objectifs du projet (Table "Tache"). En l'occurrence, j'ai identifié deux tâches qu'il s'agirait d'effectuer en priorité : demander les données du cadastre

solaire qui ne sont pas disponibles sous un format ouvert ; délimiter les coordonnées géographiques du quartier qui fait l'objet des deux projets.

ID_donnees	Categorie
1 cadastre_solaire	Energie Environne
2 dpe_batiments	Energie Occupati

Figure 6.13 : interface de la base Open Data Work de MetroMeta avec un focus (rectangle noir) sur les trois premières tables

Avec ces éléments, l'Intervenant va pouvoir construire la table d'association qui est le coeur même du plan de développement des usages (interface ci-dessous). Dans cette table, l'Intervenant va proposer une séquence de transformation (colonne "ID_transfo") des données documentées (colonne "Jeu_donnees"). Dans le cas de MetroMeta, l'Intervenant a identifié deux transformations : 1. Le première qui vise à demander les données brutes du cadastre solaire auprès du producteur (transformation de données visualisées en données brutes) 2. Le deuxième qui vise à délimiter les coordonnées géographiques de l'éco-quartier (ces coordonnées géographiques permettent de filtrer les diagnostic DPE pour ne garder que les bâtiments appartenant à l'éco-quartier)

L'Intervenant précise également :

- les acteurs qui devront effectués chacune de ces transformations (colonne "Acteurs") - le ou les tâches à effectuer pour accomplir cette transformation (colonne "Tâches") - les composants qui les guideront dans la réalisation de ces tâches (colonne "Composant") - une date de début (colonne Date_debut) et une date de fin (colonne "Date_fin") pour cette transformation - les objectifs auxquels se rattachent cette

transformation (colonne “Objectifs_projet”) - et enfin une description de la transformation (colonne “Description_transfo”) qui traduit textuellement les associations réalisées dans la table.

A ID_transfo	Jeu_donnees	Acteurs	Tâches
1	demande_donnees	cadastre_solaire	lab
2	perimetre_quartier	dpe_batiments	lab
3			
+			

Figure 6.14: Extrait de l’interface de la Table d’association

Dans notre cas fictif, le plan de développement des usages contient les deux transformations décrites ci-dessous :

Première transformation : “Dans la première transformation, il s’agira pour le laboratoire d’innovation, porteur du projet d’usage des données, d’effectuer une demande de données. Cette demande de données concerne le cadastre solaire et doit être adressée à la DINSI et à la mission énergie de la métropole. Pour réaliser cette demande, le laboratoire pourra s’aider du guide de demande de données auprès des collectivités”.

Deuxième transformation : “Dans la deuxième transformation, le laboratoire d’innovation devra définir les coordonnées géographiques du quartier en question. Pour cela, il pourra s’aider de l’application GeoJSON qui permet de délimiter graphiquement un périmètre et de voir ce périmètre en coordonnées géographiques. Précisons que ces coordonnées permettront de filtrer le jeu de données de diagnostic de production énergétique des bâtiments.”

Avant de continuer ce scénario d’utilisabilité, je tiens à détailler deux spécificités dans le fonctionnement de la table d’association, véritable élément central de l’Open Data Canvas. La première spécificité tient au fait que cette table est constituée de quatre champs ou colonnes liées. Elles agissent comme les *foreign keys* d’une base de donnée relationnelle dans la mesure où elles permettent de lier les tables de la base

Open Data Work entre elles. Ainsi, la colonne “Jeu_donnees” lie-t-elle la table d’association avec la table “Donnees ouvertes”. Autrement dit, la colonne “Jeu_donnees” appelle les valeurs présentes dans la table “Donnees ouvertes” comme le montre l’interface ci-dessous.

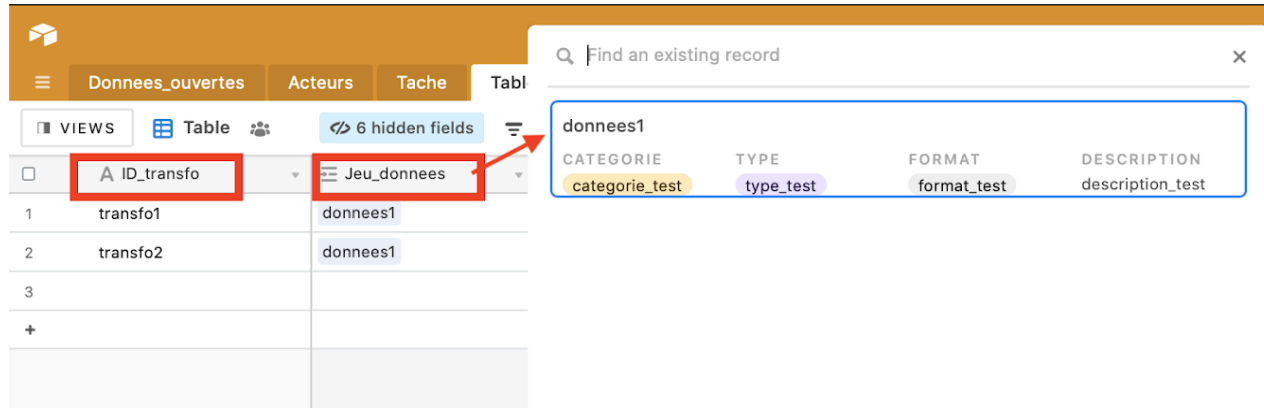


Figure 6.15: Interface présentant le fonctionnement de la colonne liant la Table d’association avec la Table “Donnees_ouvertes”

Il en est de même pour les colonnes suivantes :

- la colonne Acteurs qui fait appel à table “Acteurs”
- la colonne liée Tâches qui fait appel à la table “Tache”
- la colonne liée Composants fait appel à la table “Composant”

La deuxième spécificité provient du fait que les tables “Tâche” et “Composant” sont connectées, par API, à deux référentiels qui stockent l’ensemble des tâches et connaissances (les composants) reconnues par la communauté #teamopendata comme permettant le développement des usages. Ces tables sont présentes dans l’espace de travail de l’Intervenant (interface ci-dessous).

Ainsi, lorsqu’une nouvelle ligne est ajoutée dans la référentiel “Tâche”, celle-ci apparaît automatiquement dans la table “Tache” de la base OpenDataWork des projets. De la même manière, lorsqu’une nouvelle ligne est ajoutée dans le référentiel “Composants”, celle-ci apparaît automatiquement dans la table “Composant” de la base OpenDataWork des projets. C’est grâce à ces liens entre bases de données que l’Intervenant a immédiatement à disposition un grand nombre de tâches et de composants reconnus comme efficaces par la communauté d’expert #teamopendata.

L'intégration de nouvelles tâches et composants dans les référentiels se déroule en deux temps : le sourcing auprès de la communauté #teamopendata et l'évaluation des tâches et composants sourcés.

À propos du sourcing, celui-ci s'effectue sur une chambre spécifique du logiciel de forum en ligne Discourse (interface ci-dessous). Dans cette chambre, les membres du réseau #teamopendata peuvent, sur un certain nombre de thèmes, proposer des méthodes et applications qu'ils ont créées ou utilisées pour favoriser l'usage des données ouvertes (espace "Contribution au Canvas" rectangle noir, Figure 6.17).

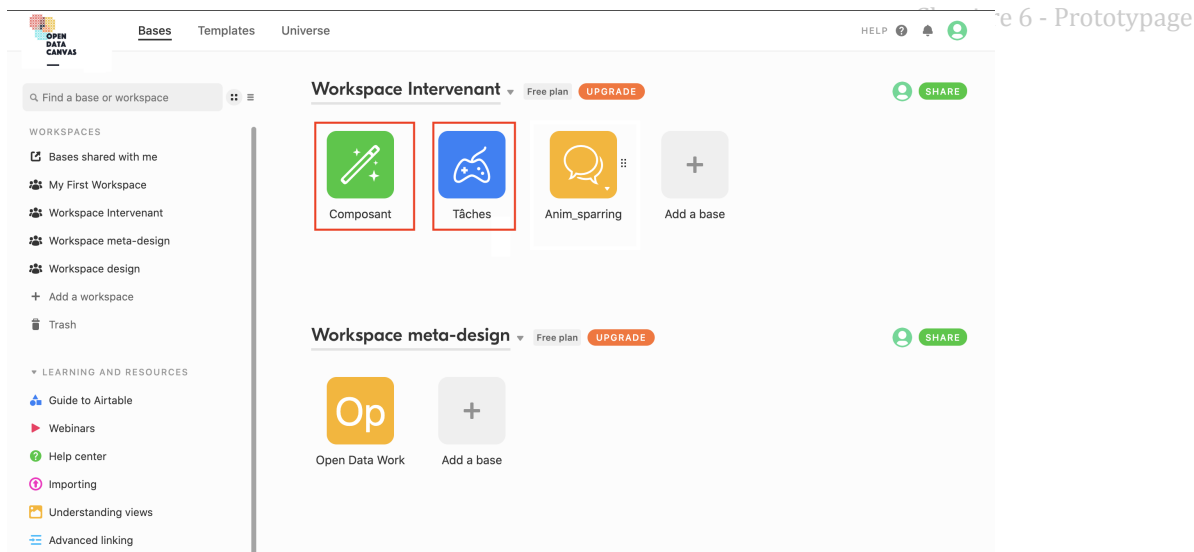


Figure 6.16 : interface présentant l'espace de travail de l'Intervenant avec les bases des référentiels (rectangle noir) de composants et de tâches entretenus par la communauté #teamopendata



Figure 6.17 : interface présentant la chambre dédiée à l'OpenDataCanvas et l'espace dédié au sourcing de composants et de tâches.

Par exemple, avec Dataactivist, nous avons fait appel à la communauté #teamopendata pour sourcer des composants listant les arguments courants utilisés pour s'opposer à l'ouverture des données. Des

arguments auxquels tout acteur impliqué dans un projet de développement des usages doit être préparé à répondre. Nous avons reçu à cet égard 23 propositions de composants.

The screenshot shows the Open Data Canvas forum interface. At the top, there is a search bar and a user profile icon. Below that, there are navigation tabs for 'Open Data Canvas' and 'Contribution au Canvas', with 'Latest' and 'Top' filters. A 'Topic' section is visible, containing a post about the 'Contribution au Canvas' category. The post text describes the category's purpose and includes a 'read more' link. Below the post, there is a table of activity with columns for 'Replies', 'Views', and 'Activity'. A specific topic titled 'Palmarès des arguments anti open data et comment vous y avez répondu' is highlighted with a red box, showing 23 replies, 1.3k views, and a date of 12 Mar. Below the table, there is a message: 'There are no more Contribution au Canvas topics. Why not create a topic?'

Figure 6.18 : interface présentant l'espace de dédié au sourcing de composants et de tâches ainsi que le fil de discussion dédié au sourcing de composants relatifs aux "arguments anti open data"

Les personnes ayant formulé une proposition de composants sont ensuite invitées à déposer leur contribution via un formulaire (interface ci-dessous) présent sur le site web de l'Open Data Canvas (<https://opendatacanvas.org>).



Contribuer à l'Open Data Canvas

Vous connaissez une ressource qui n'est pas proposée dans l'Open Data Canvas ? Indiquez là dans ce formulaire et nous l'ajouterons si elle correspond aux principes du Manifeste. Il peut s'agir d'un guide méthodologique, d'un article de presse (qui a une valeur méthodologique), d'un site internet spécialisé...

Titre de la ressource *

Origine de la ressource *

Quelle organisation ou quelle personne a produit cette ressource ?

Lien

Indiquez le lien de la ressource (ou ajoutez-ci dessous une pièce jointe)

Figure 6.19 : interface du questionnaire de dépôt des propositions de nouveaux composants

Les réponses au formulaire sont automatiquement transférées dans une base de données sur le logiciel Airtable. Cette base va servir à l'évaluation des composants et tâches sourcées. Cette évaluation correspond au sixième processus que je décrirai ci-après.

Pour s'assurer que les composants et tâches sourcées vont pouvoir servir dans les tables d'association, les propositions de composants et de tâches doivent respecter quatre critères : - ils/elles doivent être

génériques, c'est à dire être capables de s'adapter à de nombreux contextes - ils/elles doivent être génératifs, c'est à dire capables d'entraîner une action et des résultats - ils/elles doivent être collaboratives, c'est à dire qu'ils doivent pouvoir être modifiables par la communauté #teamopendata - ils/elles doivent être récursifs, c'est-à-dire être en mesure de s'imbriquer avec d'autres composants ou tâches déjà proposés.

Ainsi, seuls les composants et tâches respectant ces quatre critères peuvent intégrer le référentiel de composants et de tâches. Ce filtrage est assuré via une base de données présente sur l'espace de travail de l'intervenant.

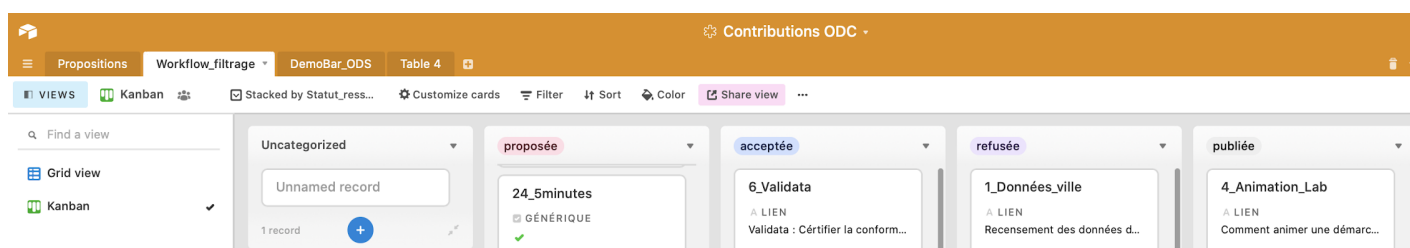


Figure 6.20 : interface de la base permettant le filtrage des composants et tâches sourcés dans la communauté #teamopendata

In fine, l'Intervenant a effectivement à disposition une large base de composants et de tâches lorsqu'il produit la première table d'association. Son travail consiste simplement à choisir les composants et tâches qui lui semblent être les plus pertinents par rapport aux projets de MetroMeta. Une fois la première table d'association produite, l'intervenant peut conclure le processus de *draft* d'un plan de développement des usages et passer au cinquième processus.

6.3.2.2.5 Processus 5 : crowdsourcing d'acteurs, de composants et de tâches

Au début de ce deuxième processus, l'Intervenant établit un diagnostic sur le niveau de complétude de la table d'association. Il évalue si pour chaque transformation envisagée correspond au moins une tâche et au moins un composant présent dans les référentiels sous-jacents.

Si la table d'association n'est pas complète, l'Intervenant lancera alors auprès de la communauté #teamopendata un appel à connaissances. Il précise à la communauté que cet appel a pour objectif de

sourcer des tâches et composants correspondant aux transformations pour lesquelles les référentiels “Tâche” et “Composants” n’ont aucun élément à proposer.

L’interface ci-dessous montre un exemple d’appel à composant que j’ai lancé conjointement avec Dataactivist, dans le but de récolter des tâches et composants pour contrer les arguments anti-open data que l’on peut rencontrer régulièrement sur le terrain.

{#teamopendata}

Palmarès des arguments anti open data et comment vous y avez répondu

■ Open Data Canvas ■ Contribution au Canvas

EtiennePichotDamon Nov '19

Bonjour à toutes et tous,

Hier, Dataactivist était à [Data On Board](#) pour tester l'[Open Data Canvas](#) et collecter des demandes de méthodes et outils à intégrer.

Une idée nous a été suggérée à plusieurs reprises : **lister les arguments que l'on pouvait entendre qui s'opposent à l'ouverture des données, ou qui justifient l'inaction.**

L'objectif serait de récolter ici tous ces arguments, qu'ils soient justifiés, ou au contraire tout à fait irrationnels, et de comprendre comment vous y avez répondu.

- Quel était l'ensemble de données en question ?
- Quel était l'argument pour ne pas le publier ?
- Comment avez-vous contré cet argument / ou avez-vous éprouvé des difficultés pour y répondre ?

Grâce à vos réponses, nous vous proposerons une synthèse : **palmarès des arguments, principaux types de données concernés, propositions de contres argumentations... et d'intégrer une composante dédiée dans l'Open Data Canvas.**

Merci d'avance pour vos contributions (*directement sous ce post*) : nous sommes certains que beaucoup d'entre-vous ont des histoires à raconter, et qu'elles seront d'une grande utilité !

5 ❤️ 23 1.3k 12 37 8
Nov '19 12 Mar replies views users likes links

Figure 6.21 : Interface montrant un exemple d’appel à connaissances

Si la table d’association est complète, alors MetroMeta peut lancer une invitation à tous les acteurs qu’elle a identifiés dans la table d’association. Cette invitation les encourage à venir rejoindre l’espace de discussion dédié au projet, dans lequel tous les échanges auront lieu.

À la fin de ce cinquième processus, tous les acteurs du projet sont présents sur l’espace de discussion du projet et (dans le cas où la table d’association était incomplète) la communauté a proposé un certain nombre de tâches et composants pour les transformations prévues par l’Intervenant.

6.3.2.2.6 Processus 6 : filtrage de sourcing

Le sixième processus a déjà été exposé ci-dessus, dans le quatrième processus, lorsque j'ai évoqué l'évaluation des composants sourcés. J'ai simplement rajouté qu'à la fin de ce processus, l'Intervenant et MetroMeta dispose d'un plan de développement qu'ils peuvent visualiser sous la forme d'un tableau Kanban (interface ci-dessous).

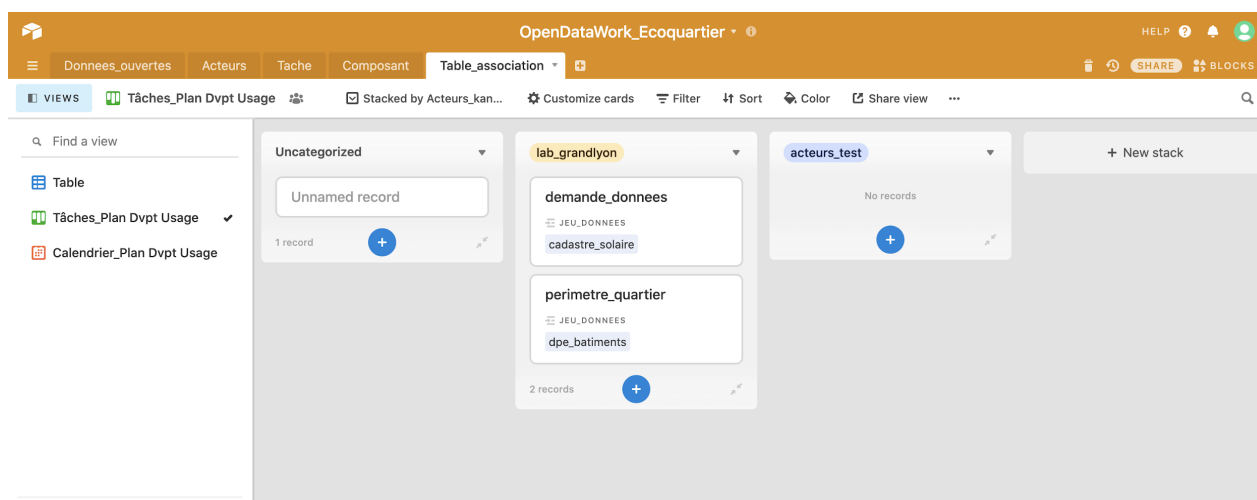


Figure 6.22 : interface montrant le tableau Kanban du plan de développement des usages.

Ce type de tableau permet de visualiser l'ensemble des transformations à réaliser par chacun des acteurs du projet et ainsi d'avoir une vision globale du projet. Dans l'exemple, le tableau Kanban indique que le laboratoire d'innovation doit effectuer deux transformations, à savoir demander les données aux collectivités et définir le périmètre de l'éco-quartier.

À noter également qu'à la fin de ce processus, MetroMeta dispose également du calendrier de son projet (interface ci-dessous).

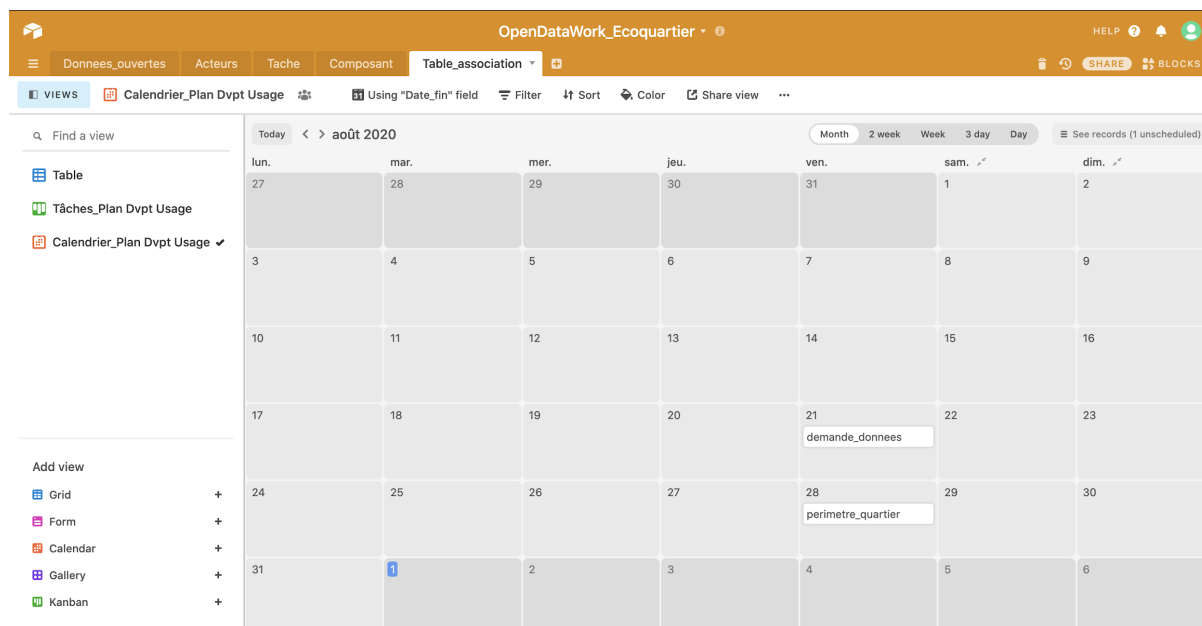


Figure 6.23 : interface montrant le calendrier du plan de développement des usages.

6.3.2.2.7 Processus 7 : communication du plan de développement

Le septième et dernier processus de l'OpenDataCanvas vise à communiquer aux acteurs parties prenantes du projet le plan de développement des usages.

Pour cela, il suffit que MetroMeta ouvre aux acteurs du projet un accès à leur espace de travail. Cet espace de travail est dédié à la mise en œuvre du plan de développement développé par l'intervenant et MetroMeta. Les acteurs trouveront dans cet espace de travail la base "Plan Dvpt Usage" (interface ci-dessous). Cette base est interconnectée, via API, avec la table d'association, si bien que lorsque MetroMeta ou l'Intervenant ajoute des éléments dans la table d'association, ceux-ci sont automatiquement transmis aux acteurs parties prenantes du projet.

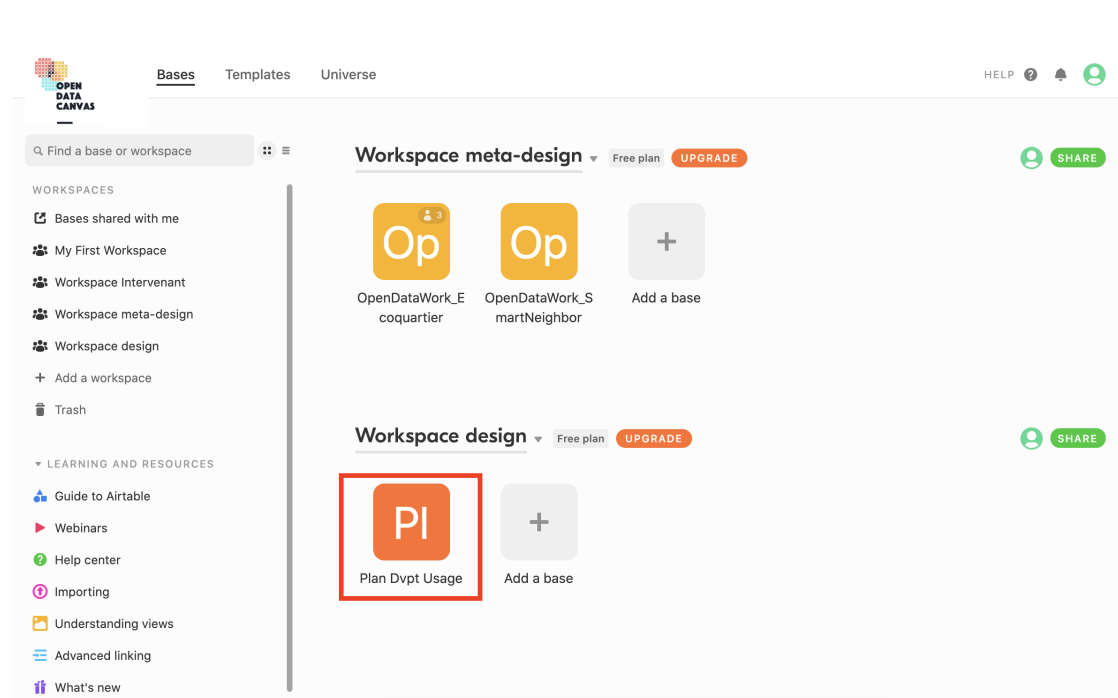


Figure 6.24: interface montrant l'espace de travail réservée aux acteurs du projet (workspace design) et la base permettant de mettre en oeuvre le plan de développement des usages

Dans cette base, chacun des acteurs partie prenante du projet a accès à un tableau Kanban dans lequel se trouve uniquement les transformations qui lui ont été attribuées. Dans l'interface ci-dessous, je prend pour exemple le tableau Kanban du laboratoire d'innovation dans lequel se trouvent les deux transformations déjà évoquées.

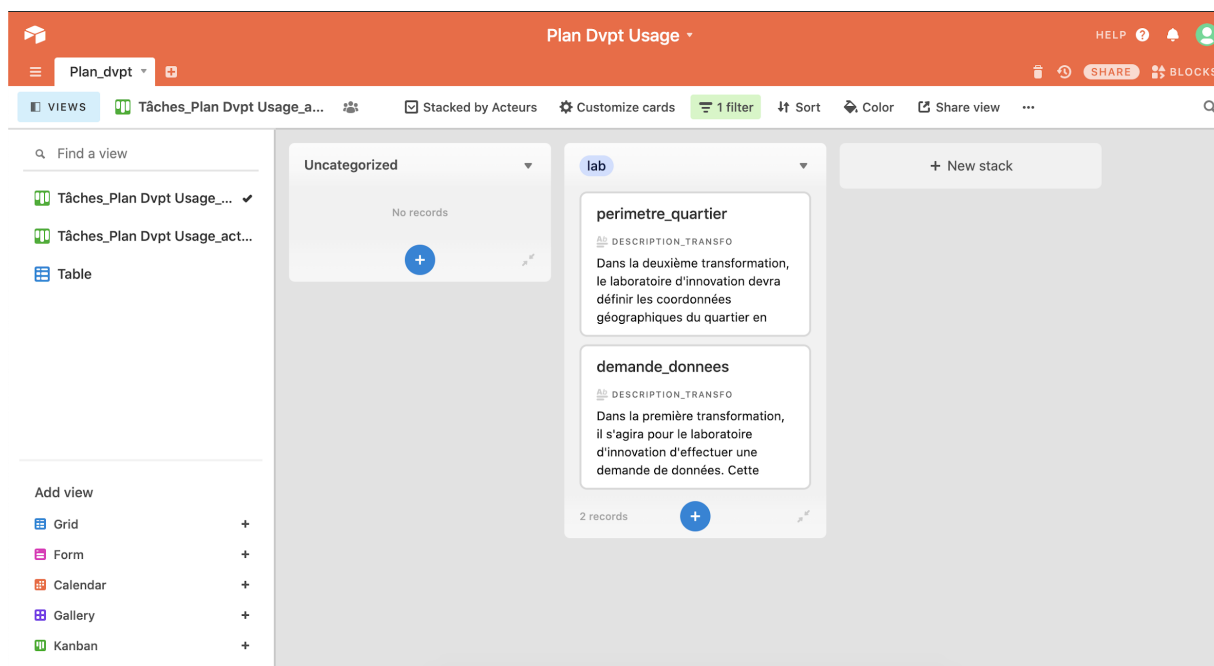














Figure 6.25 : interface montrant le tableau Kanban du laboratoire d'innovation avec les transformations qu'il doit réaliser.

À noter également que chacune des transformations menée par les acteurs fait l'objet d'un fil de discussion spécifique dans l'espace de discussion du projet (interface ci-dessous). Dans ce fil de discussion, ils pourront poser leurs questions, documenter leurs avancées et échanger avec les autres acteurs également en charge de la transformation.



Un projet pour augmenter la production et la consommation d'énergie durables (solaire, éolien, etc.) des bâtiments du quartier et ce, à travers l'usage des données ouvertes du territoire.

 MetroMeta_SmartCity  Eco-quartier Latest Top		 Edit + New Topic 		
Topic		Replies	Views	Activity
↳ Etape 1 : questionnaire linéaire du processus de sparring		2	2	7d
★ A props du projet Eco-quartier		2	5	6d
↳ Etape 2 : questionnaire circulaire du processus de sparring		4	2	6d
↳ Etape 3 : questionnaire réflexif du processus de sparring		3	1	2d
↳ Etape 4 : questionnaire stratégique du processus de sparring		3	1	2d
Plan de développement du projet		0	1	6d
Transformation1 : demande des données du cadastre solaire données-énergie		0	3	10m
Transformation2: délimiter le périmètre géographique de l'éco-quartier données-géographique		0	1	8m

There are no more Eco-quartier topics. [Why not create a topic?](#)

Figure 6.26 : Interface montrant les fils de discussions dédiées aux transformations à effectuer par les acteurs

Ce fil de discussion leur permettra également de trouver des connaissances complémentaires pour effectuer cette transformation. En effet, l'intervenant attachera à chacun de ces fils de discussion un tag correspondant au domaine de la donnée en cours de transformation (visible sous le titre du fil de discussion dans l'interface présentée ci-dessus). Ce tag fait partie d'un système qui liste un ensemble de domaine de données. Ce système est utilisé pour classifier l'ensemble des échanges de la communauté #teamopendata. De fait, en cliquant simplement sur le nom du tag présent sous le titre du fil de discussion, les acteurs auront accès à tous les échanges concernant le domaine de données avec lequel ils sont aux prises dans la transformation.

#teamopendata

Tags > données-énergie

Info + New Topic

Topic	Replies	Views	Activity
Transformation 1 : demande des données du cadastre solaire Eco-quartier données-énergie	0	3	28m
Transformation 1 : données test SmartNeighbor données-énergie	0	1	6d
Territoires et énergies : une dataviz open source collectivités, opendata, données-énergie	3	774	Aug '19
Opendata et énergie Veille collectivités, données-énergie	3	761	Feb '18
L'arrêté sur la publication des données des compteurs de consommation d'énergie est paru Veille données-énergie	1	251	Jan '18

Figure 6.27 : Interface montrant les échanges relatifs au tag “données-énergie”

Dans notre cas (interface ci-dessus), le fil de discussion relatif à la demande des données du cadastre solaire a été rattaché au tag “données-énergie”. En cliquant sur ce tag, l’interface ci-dessous apparaît. Elle liste toutes les discussions de la communauté #teamopendata sur les données d’énergie. Dans ces discussions, les acteurs pourront trouver des informations et connaissances précieuses pour mener à bien la demande des données.

La création des fils de discussion relatifs à chaque transformation marque le début du déploiement du Canvas et la fin de l’intervention active. Néanmoins, l’Intervenant assistera ponctuellement MetroMeta pour manager au mieux les itérations que suivra le plan de développement au fur et à mesure des retours que feront les acteurs de l’écosystème.

Il convient de rappeler que l’environnement de conception des usages qu’est l’Open Data Canvas a été désigné pour manager les boucles de rétroactions propres à la dynamique d’usage des données ouvertes. Dans ce cadre, une fois le plan de développement communiqué sur l’espace de discussion du projet et sur l’espace de travail, les acteurs seront invités à donner leurs impressions sur celui-ci et à suggérer des modifications. MetroMeta devra prendre en compte ces suggestions/modifications et lancer une nouvelle itération qui aboutira à la proposition d’une nouvelle version du plan de développement.

Pour proposer ces nouvelles versions, MetroMeta pourra alors suivre à nouveau l’ensemble des sept processus. Il lui suffira alors, dans l’espace de discussion d’archiver tous les fils de discussion relatifs à la

première itération et de créer de nouveaux fils de discussion pour accueillir les éléments de la deuxième itération ; aussi pourra-t-il, dans son espace de travail “Workspace meta-design” créer une nouvelle vue de la Table d’association (“Table itération2” dans l’interface ci-dessous) propre à la deuxième version du plan de développement.

The screenshot shows the OpenDataWork interface for a workspace named 'OpenDataWork_Ecoquartier'. The main view is 'Table_association', which displays a table with 3 records. The table has the following columns: ID_transfo, Jeu_donnees, Jeu_donnees_link, Acteurs, Acteurs_link, and Tâches. The data rows are:

ID_transfo	Jeu_donnees	Jeu_donnees_link	Acteurs	Acteurs_link	Tâches	
1	demande_donnees	cadastre_solaire	lab	lab	demande_donnees	
2	perimetre_quartier	dpe_batiments	dpe_batiments	lab	lab	delimitier_perimetre
3						

The left sidebar shows a list of views: 'Table', 'Tâches_Plan Dvpt Usage', 'Calendrier_Plan Dvpt Usage', and 'Table_itération2' (which is selected and highlighted with a red box). Below this list, there are options to 'Add view' with icons for Grid, Form, Calendar, Gallery, and Kanban.

Figure 6.28 : interface de la base OpenDataWork avec un focus sur la table d’association propre à la deuxième itération

Conclusion du chapitre 6

L'objectif de ce septième chapitre était de donner une forme matérielle et tangible à l'Open Data Canvas en suivant le design présenté dans le chapitre 6.

La première section de ce chapitre m'a permis de préciser dans quelle mesure le prototype permettrait de répondre à la question de recherche principale de cette thèse. Je conclus que ce prototype permettra dans un premier temps d'ajuster la loupe d'observation que j'utilise pour concevoir et problématiser les barrières socio-technique ; ensuite il maximisera mes chances de pouvoir observer si l'Open Data Canvas génère les effets souhaités dans un contexte réel ; enfin, le prototype me permettra d'évaluer quelle mesure l'artefact pourra s'adapter aux évolutions futures des besoins des usagers des données ouvertes. Aussi je précise que ce prototype prendra deux formes distinctes :

- Une version 0 prenant la forme d'une cartographie de l'architecture métier, fonctionnelle et applicative de l'Open Data Canvas
- Une version *alpha*, prenant la forme d'un prototype fonctionnel mais non-opérationnel construit à l'aide d'outils logiciels accompagné d'une méthode d'implémentation dans des contextes d'usages réel.

La deuxième section m'a permis quant à elle de sélectionner parmi les 28 fonctions possibles, celles qui seraient intégrées dans les différentes versions du prototype. Pour la version 0 et alpha, j'ai sélectionné 8 fonctions pour leur capacité à assurer l'essentiel du comportement imaginé de l'Open Data Canvas, c'est à dire :

- permettre aux futurs utilisateurs de décrire et questionner les usages de données ouvertes qu'ils souhaitent développer (cf mécanisme de sparring) ainsi que leur permettre d'obtenir des feedback de la part d'autres systèmes de l'écosystème en question (boucle de rétroaction).
- faire circuler des informations permettant aux acteurs de renforcer leurs connaissances sur l'accomplissement de tâches spécifiques et fixer un cadre commun adaptable et reproductible qui spécifie un ensemble de bonnes pratiques explicites et implicites à appliquer au sein et hors du réseau.
- organiser les métadonnées et permettre aux utilisateurs d'obtenir des séquences de tâches à appliquer pour transformer les données ouvertes

La liste de ces huit fonctions considérées comme centrales se trouve ci-dessous :

1. Création d'un chambre de conception dédiée à un territoire/une organisation
2. Création d'espaces de discussions dédiés à chacun des projets de construction des usages
3. Description de l'usage imaginé
4. Dialogue en ligne
5. Proposition de nouvelles connaissances pour la liste des composants
6. Évaluation du respect des normes techniques des connaissances
7. Adaptation du modèle de méta-données
8. Associer à la liste de données à utiliser une séquence de tâches

En ce qui concerne les 20 fonctions restantes (voir liste complète dans l'Annexe 12), bien qu'elles ne soient pas essentielles au comportement du Canvas, elles seront intégrées dans les versions ultérieures du prototype. Leur intégration permettra notamment de faciliter le travail de l'intervenant en automatisant certaines tâches (par exemple, la fonction 5 qui permettra la transmission automatique des réponses du processus de sparring dans une base de données) et d'instaurer une forme de gouvernance collaborative du Canvas (par exemple la fonction 12 qui permettra l'évaluation du respect des normes techniques des connaissances proposées pour devenir des composants par l'ensemble des membres du réseau).

Enfin, la troisième section me permet de présenter les trois cartographies d'architecture du prototype. Le lecteur pourra trouver une architecture métier qui décrit le processus que supporte l'Open Data Canvas, une architecture fonctionnelle qui montre l'imbrication des huit fonctions choisies et enfin une architecture applicative qui décrit les blocs applicatifs, la structure des données utilisées et leurs interactions. Je présente également, dans le cadre de scénarios d'utilisabilité, les interfaces utilisateurs de l'Open Data Canvas.

Si ce prototype apporte une réponse à la question de recherche centrale de cette thèse, elle n'est, en l'état, qu'hypothétique, transitoire. En effet, dans la mesure où la capacité du prototype à combler les douze besoins fonctionnels (chapitre 5) n'a pas été évaluée, je ne peux pas conclure que l'Open Data Canvas résout les tensions paradoxales qui font obstacle à la construction de nouveaux usages à partir des données ouvertes. Ce sera l'objet du chapitre suivant d'adresser cette question.

Chapitre 7

Évaluation du dispositif en Systèmes d'Information

Chapitre 7. Évaluations visant à estimer la capacité du dispositif à lever les barrières socio-techniques freinant l'usage des données ouvertes

Ce septième a pour objet d'évaluer la capacité de l'Open Data Canvas à lever les barrières socio-techniques qui freinent la construction des usages des données ouvertes. Il s'inscrit dans le cycle de pertinence de cette recherche dans la mesure où il vise à évaluer la capacité de l'Open Data Canvas à répondre aux situations problématiques rencontrées par les usagers des données ouvertes. Aussi sera-t-il guidé par la troisième sous-question de recherche :

Quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les situations problématiques au sein des écosystèmes open data ?

Pour rappel, j'ai détaillé dans le chapitre précédent le processus de prototypage qui m'a amené à concevoir l'artefact. Aussi ai-je décrit formellement la structure de cet artefact et expliqué en quoi ses différents éléments pourraient répondre aux différents aspects de la question. Ces explications sont autant de réponses temporaires dont il s'agit d'évaluer la pertinence empirique et la rigueur scientifique dans ce chapitre.

J'ai mené des évaluations dans le but d'estimer successivement la capacité des différentes versions de l'Open Data Canvas (i.e les artefacts) à lever les barrières socio-techniques qui freinent la construction des usages issus des données ouvertes.

7.1 Analyse de l'évaluation des prototypes

Le but de ma thèse est de trouver un mode de résolution des problèmes pernicieux qui font aujourd'hui obstacles à l'usage des données et ce dans le but de favoriser la création de nouveaux services qui en est issu. À cette fin, je propose au chapitre 6 le design d'un environnement de conception des usages comme réponse à ma problématique de recherche. Au chapitre suivant, je présente les différentes versions du prototype qui "matérialise" cet environnement de conception.

Ces étapes étant achevées, il est désormais question d'évaluer dans quelle mesure l'environnement de conception que je propose répond à la question de recherche. Pour cela, il s'agit d'estimer pragmatiquement (i.e sur le terrain et avec les futurs utilisateurs de l'environnement) la capacité des différentes versions du prototype à résoudre les tensions paradoxales qui font obstacle à l'usage des données ouvertes.

L'objet de cette partie relève de la première étape de notre stratégie d'évaluation des artefacts : l'analyse de l'évaluation. Elle consiste à analyser conjointement les fins que je recherche à travers l'évaluation, les différentes dimensions du prototype et les moyens dont je dispose pour évaluer la capacité du prototype à répondre à ma question de recherche.

Les résultats de cette analyse permettront de choisir, lors de la prochaine partie, une stratégie d'évaluation qui soit efficace (i.e compatible avec les finalités mon projet de recherche), cohérente vis à vis du prototype disponible et efficiente (i.e en optimisant les moyens à disposition pour mener à bien cette évaluation).

Venable et al. (2012) préconisent de subdiviser l'évaluation du prototype en sept critères :

1. La forme du prototype à évaluer
2. La nature du prototype à évaluer
3. Les propriétés du prototype à évaluer
4. Les objectifs de l'évaluation
5. Les contraintes de l'environnement de recherche
6. La rigueur à appliquer. Ce critère permet principalement de déterminer si l'évaluation est une étape préliminaire, auquel cas un niveau de rigueur modérée est suffisant, ou s'il s'agit

d'une évaluation définitive qui va donner lieu à une décision finale, auquel cas un niveau de rigueur élevé est nécessaire.

7. Pondération des critères précédents de manière à prioriser les critères essentiels des critères d'importance secondaire.

Dans cette partie, je présente et détaille un tableau qui résume l'analyse de l'évaluation des différentes versions du prototype, en suivant les six premiers critères nommés ci-dessus.

Critère	Forme du prototype	Nature du prototype	Propriété de l'évaluation	Objectif de l'évaluation	Contraintes de moyens	Rigueur à appliquer
Pondération (sur 5)	4	2	5	5	4	3
Version 0	Modèle	Processus	Utilité du prototype i.e sa capacité à résoudre les tensions paradoxales qui bloquent l'usage des données	Evaluer si l'Open Data Canvas permet de répondre aux besoins fonctionnels des acteurs	Accès au terrain privilégié Absence de budget Temps et compétences limitées	Modérée
Version alpha	Instantiation	Processus	Efficacité i.e fiabilité technique du prototype et facilité de son implémentation dans des projets concrets	Evaluer la fiabilité technique des outils logiciels choisis et identifier les principaux défauts de la méthode d'implémentatio n.	Accès au terrain privilégié ; absence de budget et absence de temps (fin de contrat de thèse) et compétences techniques limitées	Modérée

Figure 7.1 : Synthèse de l'analyse de l'évaluation

Le lecteur trouvera ci-dessous la description de chacune des lignes du tableau correspondant à chaque version du prototype.

7.1.1 Analyse de l'évaluation de la version 0 du prototype

7.1.1.1 La forme du prototype à évaluer

La version 0 du prototype consiste en trois cartographies de l'architecture de l'environnement de conception des usages. Ces cartographies décrivent la structure de l'environnement ainsi son comportement dans des contextes réels, en utilisant des notations spécifiques. En ce sens, il constitue dans l'interprétation de Venable et al. (2012) un modèle ou design, c'est-à-dire une version "non-instanciée" ou immatérielle d'un artefact.

7.1.1.2 La nature du prototype à évaluer

Je considère comme évident que la version 0 du prototype constitue un artefact process, et non un artefact produit (voir section 2.2.1.6.1, p.192-193 pour plus d'informations sur la distinction), en ce qu'elle combine les aspects techniques de l'environnement de conception via architecture fonctionnelle et applicative) ; et une dimension socio-technique via l'architecture métier (Annexe 19) qui décrit les interactions entre l'environnement et les utilisateurs humains.

7.1.1.3 Les propriétés du prototype à évaluer

Je spécifie au cours du chapitre 7 que la version 0 du prototype vise à tester auprès des futurs utilisateurs la pertinence du design vis à vis de leurs besoins fonctionnels, et par là de préciser la compréhension des besoins fonctionnels. Selon l'interprétation de Venable et al. (2012), il s'agit d'évaluer l'utilité du prototype, c'est-à-dire dans quelle mesure le prototype atteint son objectif central : résoudre les problèmes pernicieux qui font obstacle à l'usage des données ouvertes.

7.1.1.4 Les objectifs de l'évaluation

Au regard des propriétés du prototype à évaluer mentionnés ci-dessus, l'objectif de l'évaluation est par conséquent d'établir si l'environnement de conception tel que matérialisé par la version 0 du prototype permet d'atteindre les objectifs fixés au préalable. En l'occurrence il s'agit d'évaluer si le prototype permet de répondre aux 12 besoins fonctionnels établis à la fin du chapitre 5.

7.1.1.5 Les contraintes de l'environnement de recherche

Pour l'évaluation du prototype 0, j'ai disposé d'un accès privilégié au terrain. En effet, j'ai eu l'opportunité de pouvoir faire évaluer le prototype 0 par les membres du réseau Open Data France. Cet accès rapide et étendu au terrain a permis de contrebalancer l'absence de budget ainsi que le temps et les compétences limitées à ma disposition au cours du projet de thèse.

7.1.1.6 La rigueur à appliquer

La version 0 du prototype est une forme intermédiaire du prototype final qui viendra conclure ma thèse. Son évaluation vise à préciser la compréhension des besoins fonctionnels pour pouvoir construire une nouvelle version du prototype. Dès lors, je considère que l'évaluation à suivre est une évaluation préliminaire, qui justifie un niveau de rigueur modéré.

7.2 Choix d'une stratégie d'évaluation

7.2.1 Choix d'une stratégie d'évaluation pour la version 0 du prototype

Dans cette partie, il est question de choisir une stratégie d'évaluation en reportant dans le Framework de Sélection des Stratégies d'Évaluation en Méthodologie de recherche en design science (Venable et al., 2012, p.11) les sept critères de la version 0 du prototype de l'Open Data Canvas décrits dans la partie précédente. Suite à ce report d'information, le framework permet de déduire la stratégie d'évaluation adaptée.

Dans cette partie, je présente le framework de sélection de la stratégie d'évaluation de la version 0 du prototype (tableau ci-dessous) et en déduit le choix de la stratégie d'évaluation.

Framework de Selection des Stratégies d'Evaluation		Ex ante (avant le prototypage)	Ex post (après le prototypage)
		<input checked="" type="checkbox"/> Evaluation formative <input checked="" type="checkbox"/> Cout moins élevé <input checked="" type="checkbox"/> Plus rapide <input checked="" type="checkbox"/> Evaluation du design, d'un prototype partiel ou complet <input type="checkbox"/> Moins de risque pour les participants durant l'évaluation <input type="checkbox"/> Risque plus élevé de faux positifs	<input type="checkbox"/> Evaluation sommative <input type="checkbox"/> Cout plus élevé <input type="checkbox"/> Moins rapide <input type="checkbox"/> Evaluation des instances <input type="checkbox"/> Plus de risques pour les participants durant l'évaluation <input type="checkbox"/> Risque moins élevé de faux positifs
Naturaliste	<input checked="" type="checkbox"/> Plusieurs parties prenantes hétérogènes <input checked="" type="checkbox"/> Conflits substantiels <input checked="" type="checkbox"/> Artefacts process <input type="checkbox"/> Cout plus élevé <input type="checkbox"/> Evaluation longue <input checked="" type="checkbox"/> Accès au terrain requis <input checked="" type="checkbox"/> Evaluation de l'utilité de l'artefact <input checked="" type="checkbox"/> Niveau de rigueur : "On juge l'arbre à ses fruits" <input type="checkbox"/> Risque plus élevé pour les participants <input type="checkbox"/> Risque de faux positif plus faible	<input checked="" type="checkbox"/> Utilisateurs réels, problèmes réels et système irréal <input checked="" type="checkbox"/> Cout faible-moyen <ul style="list-style-type: none"> • Vitesse correcte • Risque faible pour les participants • Risque élevé de faux positifs 	<ul style="list-style-type: none"> • Utilisateurs réels, problèmes réels et systèmes réels • Cout le plus élevé • Risque le plus élevé pour les participants • Meilleure évaluation de l'utilité • Identification des effets collatéraux • Risque le moins élevé de faux positifs
Artificielle	<input type="checkbox"/> Peu de parties prenantes similaires <input type="checkbox"/> Peu ou pas de conflits <input type="checkbox"/> Artefacts purement techniques <input type="checkbox"/> Cout moins élevé <input type="checkbox"/> Evaluation courte <input type="checkbox"/> Niveau de rigueur : contrôle des variables <input type="checkbox"/> Evaluation de l'efficacité <input type="checkbox"/> Risque moins élevé pendant l'évaluation <input type="checkbox"/> Risque de faux positif plus élevé	<ul style="list-style-type: none"> • Utilisateurs irréels, problèmes irréels et système irréels • Cout le moins élevé • Vitesse la plus élevée • Risque le moins élevé pour les participants • Risque le plus élevé de faux positifs 	<ul style="list-style-type: none"> • Système réel, problème irréal et utilisateurs irréels • Cout moyen/élevé • Vitesse correcte • Risque faible-moyen pour les participants

Figure 7.2 : framework de sélection de la stratégie d'évaluation de la version 0 du prototype (d'après Venable et al., 2010, p.11)

Au regard des critères reportés dans le framework et de leur pondération, je suis amené à choisir dans le framework de Venable et al. (2010) une évaluation *ex ante* et naturaliste. Le choix de cette stratégie possède le double de pouvoir être menée à bien à une vitesse correcte et de limiter le risque pour les participants à l'évaluation. En revanche, cette stratégie comporte un risque plus élevé de faux positifs,

qu'il s'agira de prendre en compte dans le choix et le design de la méthode d'évaluation, ainsi que dans sa réalisation.

Ces points spécifiques font l'objet des parties suivantes.

7.2.2 Choix d'une stratégie d'évaluation pour la version alpha du prototype

Dans cette partie, il est question de choisir une stratégie d'évaluation en reportant dans le Framework de Sélection des Stratégies d'Évaluation en Méthodologie de recherche en design science (Venable et al., 2012, p.11) les sept critères de la version alpha du prototype de l'Open Data Canvas décrits dans la partie précédente. Suite à ce report d'information, le framework permet de déduire la stratégie d'évaluation adaptée.

Dans cette partie, je présente le framework de sélection de la stratégie d'évaluation de la version alpha du prototype (tableau ci-dessous) et en déduit le choix de la stratégie d'évaluation. Il existe deux stratégies possibles, chacune pouvant être appliquée avant (*ex ante*) ou après (*ex post*) la construction du prototype :

- Une stratégie d'évaluation naturaliste : type d'évaluation permettant d'explorer la performance d'un artefact dans un environnement réel (utilisateurs, systèmes et problématiques réelles). Ce type d'évaluation permet d'aborder toute la complexité des pratiques dans les organisations réelles et, à ce titre, bénéficie d'une forte validité interne (Venable et al., 2012).
- Une stratégie d'évaluation artificielle : type d'évaluation permettant de tester des hypothèses quant à la performance d'un artefact dans un environnement contrôlé (utilisateurs, systèmes et problématiques fictives). Du fait de la prédominance du paradigme scientifique positiviste, ce type d'évaluation jouit d'une plus grande fiabilité auprès de la communauté scientifique et ce, dans la mesure où elle respecte les critères de falsification et de reproductibilité (Venable et al., 2012).

Framework de Selection des Stratégies d'Evaluation		Ex ante (avant le prototypage)	Ex post (après le prototypage)
		<input type="checkbox"/> Evaluation formative <input type="checkbox"/> Cout moins élevé <input type="checkbox"/> Plus rapide <input type="checkbox"/> Evaluation du design, d'un prototype partiel ou complet <input type="checkbox"/> Moins de risque pour les participants durant l'évaluation <input type="checkbox"/> Risque plus élevé de faux positifs	<input checked="" type="checkbox"/> Evaluation sommative <input checked="" type="checkbox"/> Cout plus élevé <input checked="" type="checkbox"/> Moins rapide <input checked="" type="checkbox"/> Evaluation des instances <input checked="" type="checkbox"/> Plus de risques pour les participants durant l'évaluation <input type="checkbox"/> Risque moins élevé de faux positifs
Naturaliste	<input checked="" type="checkbox"/> Plusieurs parties prenantes hétérogènes <input checked="" type="checkbox"/> Conflits substantiels <input checked="" type="checkbox"/> Artefacts process <input checked="" type="checkbox"/> Cout plus élevé <input type="checkbox"/> Evaluation longue <input checked="" type="checkbox"/> Accès au terrain requis <input type="checkbox"/> Evaluation de l'utilité de l'artefact <input checked="" type="checkbox"/> Niveau de rigueur : "On juge l'arbre à ses fruits" <input type="checkbox"/> Risque plus élevé pour les participants <input type="checkbox"/> Risque de faux positif plus faible	<ul style="list-style-type: none"> • Utilisateurs réels, problèmes réels et système irréal • Cout faible-moyen • Vitesse correcte • Risque faible pour les participants • Risque élevé de faux positifs 	<input checked="" type="checkbox"/> Utilisateurs réels, problèmes réels et systèmes réels <input checked="" type="checkbox"/> Cout le plus élevé <input checked="" type="checkbox"/> Risque le plus élevé pour les participants <ul style="list-style-type: none"> • Meilleure évaluation de l'utilité • Identification des effets collatéraux • Risque le moins élevé de faux positifs
Artificielle	<input type="checkbox"/> Peu de parties prenantes similaires <input type="checkbox"/> Peu ou pas de conflits <input type="checkbox"/> Artefacts purement techniques <input type="checkbox"/> Cout moins élevé <input type="checkbox"/> Evaluation courte <input type="checkbox"/> Niveau de rigueur : contrôle des variables <input checked="" type="checkbox"/> Evaluation de l'efficacité <input type="checkbox"/> Risque moins élevé pendant l'évaluation <input type="checkbox"/> Risque de faux positif plus élevé	<ul style="list-style-type: none"> • Utilisateurs irréels, problèmes irréels et système irréels • Cout le moins élevé • Vitesse la plus élevée • Risque le moins élevé pour les participants • Risque le plus élevé de faux positifs 	<ul style="list-style-type: none"> • Système réel, problème irréel et utilisateurs irréels • Cout moyen/élevé • Vitesse correcte • Risque faible-moyen pour les participants

Figure 7.3 : framework de sélection de la stratégie d'évaluation de la version alpha du prototype (d'après Venable et al., 2010, p.11)

Au regard des critères reportés dans le framework et de leur pondération, je suis amené à choisir dans le framework de Venable et al. (2010) une évaluation *ex post* et naturaliste.

7.3 Choix de la méthode d'évaluation

7.3.1 Choix de la méthode d'évaluation de la version 0 du prototype : des focus group inspirés des méthodes utilisées en ergonomie informatique

Dans la partie précédente, j'établis qu'il sera question d'adopter une stratégie d'évaluation naturaliste et *ex ante*. Cette stratégie d'évaluation permet, selon le Framework de Sélection des Méthodes en Design Science (voir Table 2.15bis, p.195), deux méthodes : la recherche-action et les focus group (Venable et al., 2012, p.12).

J'ai décidé d'écarter l'option de la recherche-action dans la mesure où je n'avais ni les ressources matérielles ni le temps pour designer une recherche-action en lien avec les membres du réseau Open Data France auxquels j'ai eu accès. De fait, j'ai opté pour la mise en place de *focus group*.

Plus spécifiquement, la méthode l'évaluation de la version 0 du prototype a été inspirée des méthodes utilisées en ergonomie informatiquement. Comme spécifié précédemment, ils ont été réalisés avec les moyens disponibles, assez limités, ce qui explique le pragmatisme de la démarche. Ces *focus group* ont été effectués à distance. Cet écart par rapport au strict respect de la méthode des *focus group* a été contrebalancé par deux éléments : la rapidité et la facilité d'accès à des terrains éloignés, qui m'ont permis de recueillir des avis sur une bonne partie du territoire français en l'espace d'un mois ; la constitution d'un échantillon important de participant (14 chef(fes) de projet Open Data en France).

7.4 Réalisation de l'évaluation

7.4.1 Réalisation de l'évaluation de la version 0 du prototype

Il est désormais question dans cette partie de présenter les données récupérées, leur analyse et les résultats que je tire de leur interprétation.

L'objet de cette partie sera donc, dans un premier temps, de faire état des caractéristiques des cinq *focus group* menés (profil des participants, durée des focus group), puis, dans un second temps, de présenter des extraits de l'analyse des données récoltées avant, dans un troisième et dernier temps, de présenter une synthèse des résultats.

Le lecteur trouvera donc à la fin de cette partie la réponse à la question de recherche de ce chapitre, à savoir : dans quelle mesure l'Open Data Canvas est-il en capacité de lever les barrières socio-techniques qui freinent l'usage des données ouvertes, c'est à dire de répondre aux douze besoins fonctionnels ?

7.4.1.1 Profil des participants aux 5 focus group

Le tableau ci-dessous présente les caractéristiques démographiques des dix-neuf participants aux cinq *focus group*. Une grande majorité des participants travaillent pour une collectivité territoriale en charge d'animer l'usage des données ouvertes sur leur territoire : mise en place d'une plateforme, sensibilisation des acteurs aux enjeux de la donnée, transformation des données (réutilisation), mise en place d'une gouvernance, etc. Cette charge d'animation de l'usage des données ouvertes sur le territoire impliquant qu'ils travaillent directement ou indirectement avec les acteurs positionnés à l'échelon inférieur dans le découpage territorial. Par exemple, les participants agents d'une région ont tous témoigné de pratiques conjointes avec les agents des départements, les participants agents d'un département de pratiques conjointes avec les agents des métropoles, etc.

Il est pertinent de préciser que les collectivités desquelles sont issus les participants couvrent, dans une très grande majorité, un territoire de dimension supra-communale. On trouve ainsi trois métropoles, deux départements, trois régions, une collectivité spécifique (Corse), ainsi qu'un réseau de collectivité de rayonnement national. Il faut également préciser que les participants couvrent une partie significative du territoire français. De par ces éléments, je peux raisonnablement conclure que l'échantillon constitué est représentatif de la diversité des pratiques d'usage des données ouvertes à l'échelle supra-communale, ces dernières étant conçues en relation directe avec les pratiques situées à l'échelon territorial inférieur.

À propos du profil spécifique des participants, la totalité avait au moins une expérience passée de projets d'usage des données, conformément aux critères de recrutement fixés dans la partie précédente, cette expérience datant d'au moins six mois avant la tenue des *focus group*.

#	Focus group #	Profession	Ville/Collectivité	Durée (min)
14	Focus Group 1	<ul style="list-style-type: none"> • Délégué Général • Chef de projet 	Open Data France	104
8	Focus Group 1	Consultant Open data	Paris	"
3	Focus group 2	Chargé de mission Innovation numérique	Département de la Haute-Garonne	127
6	Focus group 2	Chef de projet Open Data	Région Occitanie	"
11	Focus group 2	Spécialiste open data	Département de la Haute-Garonne	"
13	Focus group 2	Chef de projet Open Data	Montpellier Méditerranée Métropole	"
12	Focus group 2	Directeur de projet Données et Transformation Digitale	Région Occitanie	"
2	Focus group 3	Chargée de mission Open data	Région Centre-Val de Loire	123
4	Focus group 3	<ul style="list-style-type: none"> • Chef de projet Open Data • Chef de projet 	Orléans Métropole	"
1	Focus group 4	<ul style="list-style-type: none"> • Chef de projet Open Data • Directeur 	Région Bourgogne-Franche-Comté	99
10	Focus group 4	Chargée de projets - Direction de l'Innovation	Agglomération de Nevers	"
5	Focus group 5 (PACA + Corse)	Cheffe de mission Données territoriales	Collectivité de Corse	130
7	Focus group 5 (PACA + Corse)	Chargé de Mission SIG & Open Data	Mairie d'Avignon	"
9	Focus group 5 (PACA + Corse)	Chargé de mission Open Data	Digne les-Bains	"
15	Focus Group 5(PACA + Corse)	<ul style="list-style-type: none"> • Cheffe du service Smart Services • Chef de projet open data 	Métropole Aix-Marseille	"
Total : 15				Total : 583

Table 7.1 : caractéristiques des focus group et des participants

7.4.1.2 Analyse des données

Pour établir dans quelle mesure l'environnement de conception s'avérait utile pour lever les barrières socio-techniques qui freinent l'usage des données ouvertes, je me suis basé sur les données relatives au code "Justification de l'utilité inutilité". Pour rappel et comme évoqué dans la partie précédente, les participants ont été questionnés sur les barrières socio-techniques qu'ils avaient rencontrées au cours de leur projet d'usage des données et dans quelle mesure le Canvas permettait d'y répondre.

Les données relatives au code "Justification de l'utilité inutilité" regroupent les propos des participants qui précisent en quoi le Canvas pourrait lever cette barrière ou, inversement, en quoi il ne le pourrait pas. Aussi ai-je inclus dans ce code tous les propos des participants qui évoquent explicitement un élément dont le design de l'Open Data Canvas est constitué, ce que je considère alors comme une preuve d'utilité. Inversement, j'ai considéré que tous les propos qui évoquent un élément déjà utilisé par les participants et qu'ils considèrent comme suffisant pour lever la barrière, comme une preuve de l'inutilité de l'Open Data Canvas.

Par exemple, lorsqu'un participant évoque au cours d'un échange sur la transformation des données ouvertes (3ème étape de l'usage) une barrière sous les termes suivants :

"La DSI m'a fait des extractions mais ça ne m'a pas apporté un niveau de satisfaction suffisant. Là, il faut un certain temps pour que je mette à jour la donnée, que je la retravaille, elle n'est pas très propre. La difficulté, c'est aussi que je suis toute seule et que je dois également m'occuper de certains aspects stratégiques" - Focus group 5

Et qu'il évoque dans la suite des échanges ce dont il aurait besoin pour lever cette barrière sous les termes suivants :

"Je suis inscrite au programme Open Data Locale parce que ce que je ne veux pas c'est m'épuiser par manque de moyens, et je veux qu'ils me donnent des outils : leur cadre, leur référentiel de données, leur boîte à outils pour remettre en forme les données correspondantes chez moi. Malheureusement, je n'ai pas encore reçu cet outillage et assistance." - Focus group 5

Alors, dans ce cas, le passage est affilié au code "Justification de l'utilité/inutilité" puisque l'environnement de conception des usages met à disposition un système de gestion des tâches dans lequel l'utilisateur a accès à un référentiel d'outil. Autrement dit, le participant a explicitement mentionné un élément constitutif du Canvas, ce qui vient prouver son utilité au regard de la barrière évoquée.

La table ci-dessous est un extrait de la synthèse du codage de la retranscription du premier *focus group*.

Étape d'usage	Participant	Tâches	Barrière	Utilité Canvas	Justification de l'utilité ou l'inutilité	Outils utilisés auparavant	Substitution ou combinaison possible
2-Evaluation potentiel des données	Directeur Open Data France	“Par exemple, tu abordes dans ton processus 4 à 7 la mise en place de l’open data et son utilisation et le Canvas semble adopter le raisonnement : "si j’ai une méthode sur le comment mettre en place alors ça marche". Déjà au fond cette approche méthodo par étape on en a mis une en place”	“Et il y a même quelque chose qui pourrait être dit différemment c’est qu’il y a des gens qui n’ont pas encore l’illumination le lundi matin qu’il faut faire de l’open data . Ils se demandent d’abord "pourquoi”	Oui	“Alors oui ça peut aider chaque collectivité avant d’aller voir les communes d’être tout de suite très pertinent en arrivant avec une histoire extrêmement efficace : "il y a un problème en centre ville on peut y répondre avec les données parce qu’on a un outil et un plan XY .”	Loi pour une République numérique	Oui .
3-Transformation des données	Directeur Open Data France	“Au début on a co-construit les solutions (aide à la transformation) avec les acteurs de terrain . Ca a pris un certain temps avant de créer ce référentiel de bonnes pratiques et il y a 2 ans ils n’avaient pas de bonnes pratiques et la documentation est arrivée en allant”	“Et nous on a eu nos difficultés, on a eu du mal à trouver des outils techniques pour faire la diffusion des outils . On ne savait faire la différence entre l’usine, la co-construction des outils et la galerie commerciale où on diffuse tout ça.”	Oui	Évidente. Le Canvas met en place un outil de recensement de filtrage et de diffusion d’outils permettant d’effectuer certaines tâches (voir processus métier)	GitBook (outil de documentation de projets informatiques)	Oui

Table 7.2 : extrait de la table faisant la synthèse du codage des transcriptions des focus group

7.4.1.3 Résultats

Je détaillerai ici l'utilité ou l'inutilité de l'Open Data Canvas pour lever les barrières socio-techniques relatives à chacune des cinq étapes de l'usage des données ouvertes. Pour une lecture plus rapide, le lecteur pourra se référer directement aux tables 7.3-7.14 (p.480-492, ci-dessous). L'utilité qui synthétise tous les éléments évoqués dans les paragraphes ci-dessous.

7.4.1.3.1 Étape 1 - Utilisation des fonctionnalités d'une plateforme open data

Aucune preuve n'a pu être établie de l'utilité du Canvas eu égard à cette étape. De fait, je considère que l'environnement de conception des usages est inutile pour lever les barrières socio-techniques qui freinent l'utilisation des fonctionnalités d'une plateforme *open data* (BF3).

7.4.1.3.2 Étape 2 - Évaluation du potentiel des données

Eu égard à la barrière du manque de sensibilisation et de compréhension technique des acteurs sur les enjeux des données ouvertes (BF4), trois *focus group* ont conclu à une utilité du Canvas, un *focus group* à l'inutilité, le dernier ne s'étant pas prononcé sur la question.

Les trois *focus group* ayant conclu à une utilité du Canvas fondent leur argument sur le fait qu'il permettrait de prendre contact avec des producteurs, intermédiaires ou réutilisateurs de données sur la base d'un plan de développement des usages qui fixent des objectifs précis. Ce plan permet de faire comprendre aux acteurs les finalités concrètes de l'usage des données, ce qui, selon les participants, permettrait d'effacer le manque de sensibilisation :

“Il faut tout replacer dans le cadre d'une stratégie globale qui explique le sens des actions sinon les acteurs se perdent. En ce sens, le fait que vous fixiez le questionnement des objectifs au début du Canvas et que ce soit le fil rouge, c'est une très bonne idée.” - Focus group 2

Si je me replace dans mon modèle théorique (Figure 3.10, p.284), je peux affirmer que le plan de développement intervient sur les conditions d'interaction avec les données propres aux acteurs (Légende e). Il les replace dans un cadre stratégique et pragmatique plus large, ce qui, selon les participants, facilite l'utilisation des données (Légende f) :

“Une idée intéressante. Il est certain que si on aborde l’open data en commençant par les questions de comprendre ce qu’on veut faire, si on commence par les enjeux data IA création de services, ça va être plus facile partout, et notamment pour ouvrir. On va pouvoir aller voir les personnes ou élus qui détiennent les données et leur disant qu’on a besoin de leurs données pour telle ou telle raison.” - Focus group 2

À noter que pour un participant (Focus group 5), cette utilité est à modérer dans la mesure où les plateformes *open data* peuvent jouer à elles seules ce rôle de sensibilisation.

Eu égard au besoin de disposer de flux de connaissances actualisées (BF5) sur ce qu’il est possible de faire avec les données et les méthodes pour libérer ce potentiel, trois *focus group* ont conclu à l’utilité du Canvas, les deux autres ne s’étant pas prononcés sur la question.

Les trois *focus group* ayant conclu à une utilité du Canvas fondent également leur argument sur le fait que le processus de *crowdsourcing* de connaissances permettrait de centraliser et de redistribuer les connaissances aux acteurs en demande de ces connaissances :

“Ca (processus de crowdsourcing de connaissances de l’Open Data Canvas) pourrait tout à fait nous aider à récupérer les outils et à les flécher vers les projets que l’on recense, que l’on connaît et qui pourraient en bénéficier” - Focus group 3

Si je me replace dans mon modèle théorique (Figure 1.17, p.284), je peux affirmer que le *crowdsourcing* de connaissances intervient sur la rencontre intra-système entre la représentation des données et les conditions d’interaction avec celles-ci (Légende *d* et *e*). Plus précisément, il informe les acteurs des rencontres ayant déjà abouti à l’émergence de séquences d’articulation, leur donnant ainsi des informations précises sur la manière de transformer leurs propres données. Pour certains participants, cette intervention aboutirait à un auto-renforcement du volume de connaissance contenu dans les flux :

“En tant que tel, du peu que j’ai vu, le Canvas va venir prendre ce qui a été fait à un endroit et l’amener à un autre. Vous savez, on se rend compte régulièrement que quand on veut innover dans un domaine, cette innovation existe très certainement dans un autre domaine. Et là, si vous faites des appels à sollicitation, que vous standardisez les ressources pour que tout le monde puisse les comprendre, qu’elles alimentent des objectifs concrets, alors ça va commencer à bouger et tout le monde va vouloir fédéraliser, mutualiser” - Focus group 2

Enfin, eu égard au besoin d’un outil d’interaction entre les producteurs et les intermédiaires d’un côté, et les réutilisateurs de l’autre (BF6), deux *focus group* ont conclu à l’utilité du Canvas, les trois autres ne s’étant pas prononcés sur la question.

Les deux *focus group* ayant conclu à une utilité du Canvas fondent également leur argument sur le fait que les chambres de discussion et espaces de conception du Canvas pourraient faciliter l'implication quotidienne des acteurs dans l'usage des données, et par là, les interactions au sein de l'écosystème :

“J'ai la sensation qu'il (le Canvas) pourrait adresser un point clé : (...) les impliquer (élus, agents et services) continuellement pour qu'à la fin, la démarche devienne naturelle.” - Focus group 4

Précisons que, d'après les participants, cette implication est conditionnée par le fait que les interactions se déroulent dans un cadre informel, sur la base du volontariat et vise à la co-production d'innovations entre les différents acteurs de l'écosystème :

“Ca pourrait bien s'intégrer avec ce qu'on essaye de faire . On a commencé avec un petit cercle de techniciens pour travailler ensemble. Groupe informel un peu structuré et on s'est organisé pour travailler sur 3 thématiques dont l'expérimentation d'usages de la donnée. On va essayer de co-produire des innovations. On lance un travail collectif là-dessus autour d'un sujet qui est : les données des bornes de wifi public et voit ce qu'on va en faire” - Focus group 3

Si je me replace dans mon modèle théorique, je peux affirmer que les chambres de conception interviennent sur la rencontre inter-système entre la représentation des données et les conditions d'interaction avec celles-ci. Plus précisément, elles permettent la rencontre entre la représentation d'une donnée “sortante” d'un système (*output*) ou d'une plateforme et les conditions d'interactions jouant le rôle de “barrières d'entrée” pour les données entrantes (*input*) d'un autre système.. Par exemple, ces chambres permettraient de dimensionner une offre de mise à disposition de données supportée par une plateforme *open data*, et ce, par la rencontre entre producteurs, intermédiaires et réutilisateurs.

7.4.1.3.3 Étape 3 - Transformation des données

Aucune preuve n'a pu être établie de l'utilité du Canvas eu égard au besoin d'un outil permettant de comparer la qualité des données entre elles (BF7). De fait, je considère que l'environnement de conception des usages est inutile pour lever les barrières socio-techniques relatives à la mauvaise qualité des données.

En revanche, eu égard au besoin de disposer d'outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité (BF8), trois des cinq *focus group* ont conclu à l'utilité du Canvas, les deux derniers ne s'étant pas prononcés sur le sujet.

Les trois *focus group* ayant conclu à une utilité du Canvas fondent leurs arguments sur le fait que le Canvas répondrait à la nécessité grandissante de fixer des thématiques prioritaires en matière de transformation des données :

“Il y a une demande à laquelle on a toujours pas répondu et qu'on est en train de construire : une de nos demandes a été de créer des groupes thématiques au sein d'ODF, dédiés aux département et dédiés aux régions, avec que des référents open data à ce niveau là pour qu'ils s'entraident. L'idée ça serait que untel dise "Moi, en tant que région, j'ai tel problème" et qui demandent si tout le monde à le même pour créer un groupe et qu'elles s'entraident et partagent entre elles. Elles ont aussi des problématiques et des marges d'amélioration avec de la normalisation sur des données spécifiques, des compétences pour rendre ces données plus exploitables.” - Focus group 1

En effet, le découpage des chambres de conception en espaces de discussions thématiques permettrait d'acter les données à transformer en priorité dans un écosystème, du fait du cadre stratégique sous-jacent.

Si je me replace dans mon modèle théorique, je peux affirmer que le découpage des chambres de conception en espaces de discussion thématiques intervient sur la structuration des échanges entre les trois systèmes, symbolisés dans le modèle par la convergence des trois systèmes vers un cadre cognitif partagé. Plus précisément, il permet de centrer les actions collectives de transformation des données sur certaines données considérées comme prioritaires.

En filigrane, d'après les participants, c'est le positionnement différentiel des écosystèmes et des territoires qui se joue, certains écosystèmes préférant se concentrer par exemple sur les données de la commande publique, environnementales ou les données de transport selon leurs enjeux propres. Un participant évoquait des évènements déjà organisés pour structurer l'écosystème en ce sens :

On va organiser la première matinée des Référents et Producteurs de données, l'idée étant de créer un réseau de tout un tas d'acteurs, pour que chaque territoire s'approprie les données et soit à l'initiative sur certains domaines - Focus group 2

7.4.1.3.4 Étape 4 - Création d'un cadre cognitif partagé

Eu égard au besoin d'un parcours de tâches avec objectifs annoncés (BF9) pour répondre au manque d'un cadre stratégique clair, l'utilité du Canvas a été reconnue par quatre des cinq *focus group*, le dernier ne s'étant pas prononcé sur la question.

Les quatre *focus group* ayant conclu à une utilité du Canvas fondent leurs arguments sur le fait que la construction de la table d'association "Données-Tâches-Composants-Acteurs" et plus généralement d'un plan de développement des usages par un intervenant extérieur permettrait de passer du discours politique sur l'ouverture des données à un véritable travail pratique et collectif sur l'ouverture des données :

"Oui, ce Canvas donne un moyen supplémentaire de travailler à plusieurs autour des données" - Focus group 2

Ainsi, le Canvas permettrait de donner à voir la stratégie en matière d'ouverture des données d'un territoire tout en la mettant en œuvre en faisant participer un maximum d'acteurs. Les participants précisent à cet égard l'importance pour la Canvas de lancer un travail sur l'ouverture des données qui soit le moins centralisé et le plus distribué possible et ce pour limiter les frictions politiques ou organisationnelles :

"L'idée c'est d'essayer de se répartir les tâches, et d'attribuer les tâches à ceux qui ont le plus de force pour les mener. Par exemple, on travaille avec la région sur ce point là, pour mettre en place un système fédéraliste de partage des tâches et des compétences, et ceux, en réussissant à ne jamais réintroduire un système central (...) et on essaye de leur (collectivités) faire quitter le côté leadership, centralisateur pour qu'ils prennent le pli de la coopération." - Focus group 2

Si je me replace dans mon modèle théorique, je peux affirmer que la construction d'un plan de développement des usages permet de créer une compréhension partagée des objectifs de l'usage des données et du rôle que chacun y joue et ce, par un travail concret de la donnée. En ce sens, ce plan de développement participe à la coordination de la transformation des données au sein de chacun des systèmes.. Un des participants précise le contexte dans lequel le plan de développement serait le plus utile :

"Pour des territoires ou EPCI beaucoup plus petits, avec beaucoup de communes et le gars qui doit s'occuper de l'open data a plein d'autres choses à faire : oui, il peut très clairement utiliser la ressource. Ca peut être soit des territoires complètement ruraux, et c'est le département qui doit bouffer les 200-300 communes, et donc là, une aide extérieure peut être complètement utile avec de la ressource pour traiter la donnée. Après, sur des EPCI plus gros, le mec de l'informatique, fait en même temps l'open data et le déploiement de la fibre optique, lui il va avoir besoin d'autant d'aide extérieure que possible." - Focus group 5

Eu égard à la barrière du coût élevé de l'évangélisation qui empêche sa mise en place (BF10), l'utilité du Canvas a été reconnue par deux *focus group* sur cinq, les trois derniers ne s'étant pas prononcés sur la question.

Les deux *focus group* ayant conclu à une utilité du Canvas fondent leurs arguments sur le fait que le plan de développement se situe à un niveau stratégique, avec des objectifs concrets et ce sans évoquer les aspects techniques d'ouverture des données qui freinent les acteurs et empêchent l'évangélisation, notamment des élus :

“Au niveau politique, si on parle technique, ça ne les intéresse plus” - Focus group 2

L'évangélisation se ferait alors, selon les participants, de manière autonome puisque le Canvas susciterait leur intérêt :

“J'ai la sensation qu'il(le Canvas) pourrait adresser un point clé : la construction de la confiance et de l'intérêt avec les services, il pourrait aider à effectuer un travail qui est énorme, à savoir amener les élus, les agents et services dans la démarche” " Focus group 4

Si je me replace dans mon modèle théorique, je peux affirmer que le plan de développement des usages permet, par l'accent porté sur la dimension stratégique, d'amener un maximum d'acteurs à s'intéresser à l'ouverture des données et donc de les attirer au sein du cadre cognitif partagé mentionné ci-dessus, au centre de l'écosystème *open data*.

7.4.1.3.5 Étape 5 - Institutionnalisation de l'usage des données

Eu égard au besoin de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes (BF11), deux *focus group* ont conclu à l'utilité du Canvas, un *focus group* ont conclu à l'inutilité du Canvas, les deux derniers ne s'étant pas exprimé sur la question. Par conséquent, l'utilité du Canvas est fortement contestable.

Les deux *focus group* ayant conclu à l'utilité du Canvas fondent leur argument sur le fait que le Canvas fonctionne en parallèle des structures organisationnelles ce qui correspond à une nécessité bien connue des participants :

“il va falloir aller vite, être autonome, bousculer ou effacer les chaînes de décision hiérarchiques, mobiliser l'argent, se donner le droit à l'erreur, avoir le droit à l'échec”. - Focus group 5

Aussi évoque-t-il le fait qu'il va permettre vise à développer des structures collectives, ce qui correspond également à une nécessité évoquée par les participants :

“Le frein pour moi, c’est l’absence de structures collectives. Tant que les régions n’embrassent l’ensemble de leurs utilisateurs publiques en leur disant : ”on prend en charge la dynamique territoriale”, ça ne marchera pas” - Focus group 2

A contrario, le *focus group* ayant conclu à l’inutilité du Canvas rappelle que des mécanismes fonctionnels existent déjà pour créer des structures collectives, à l’image des conventions signées entre collectivités pour la mise en oeuvre de prestations d’accompagnement relative à l’ouverture des données:

“On a la convention pour rassurer, qui fixe qui a le droit de faire quoi, les droits et les devoirs de chacun. On a essayé un petit peu de faire l’inventaire de ce qu’il était possible de faire. Ça cadre bien et ça rassure”” - Focus group 5

Enfin, aucune preuve n’a pu être établie de l’utilité du Canvas eu égard aux besoins de principes qui cadrent la manière de concevoir l’ouverture des données (BF12), excepté pour un *focus group*. De fait, je considère que l’environnement de conception des usages est inutile pour répondre à ce besoin.

Du fait que l’utilité du Canvas sur cette cinquième étape est contestable, je considère que l’environnement de conception des usages est inutile pour lever les barrières socio-techniques qui freinent l’institutionnalisation de l’usage des données.

1-BF1- Documentation production

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves					
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1	
1-BF1- Documentation production	NA	NA	NA	Non	NA	NA	NA	NA			NA	NA	NA	NA	NA	NA

Table 7.3 : Tableaux sur l'utilité de l'Open Canvas sur le premier besoin fonctionnel

1-BF3-Design plateforme double entrée

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
1-BF3-Design plateforme double entrée	NA	Non	NA	NA	NA	NA	Canvas répondrait à un autre besoin : “Le Canvas(...) faciliterait l’uniformisation des plateformes”	NA	NA	NA	NA	NA	NA	NA	NA

Table 7.4 : Tableaux sur l'utilité de l'Open Canvas sur le troisième besoin fonctionnel

2-BF4-Sensibilisation

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
2-BF4-Sensibilisation	Oui	NA	Non. Une plateforme open data suffit	Oui.	Oui.	Preuve 1 -“Là, je me dis que si on pouvait utiliser le Canvas pour chaque projet dans lequel il y a de la donnée ça serait sympa. A priori, je verrai l'utilité pour les projets qui viennent de démarrer et dans lesquels on peut coupler l'open data. Ca peut m'aider par exemple, en mettant tous les projets sur lesquels j'ai été sollicité, et pour me questionner sur comment on peut progressivement trouver des données, les utiliser avec pour but qu'elle soit utile à la fin au projet. Et ça permettrait en plus de faciliter l'acculturation des acteurs, par l'action et l'intégration dans tous les projets d'open data. C'est une façon efficace d'acculturer.” Preuve 2-	NA	“Mais si vous leur montrez la plateforme, vous pouvez leur dire : "l'open data, c'est ça", voilà ce que ça peut donner. Et puis en plus de ça, la plateforme montre à voir le boulot des agents, et donc vous allez dire aux agents que tout leur boulot va être transparent et que les citoyens vont savoir ce qu'il s'est passé. Et l'agent va aussi changer la perception de son boulot puisque le service public va être présent dès la donnée, puisqu'il va la produire pour un autre. Et finalement, ça aura un vrai impact sur l'organisation.”	Preuve 1 - D'après un participant, le Canvas permettrait de sensibiliser les acteurs sur la portée stratégique de l'open data : “Il faut tout replacer dans le cadre d'une stratégie globale qui explique le sens des actions sinon les acteurs se perdent. En ce sens, le fait que vous fixiez le questionnement des objectifs au début du Canvas et que ce soit le fil rouge, c'est une très bonne idée.”	“Alors oui, ça peut aider chaque collectivité, avant d'aller voir les communes, d'être tout de suite très pertinent, en arrivant avec une histoire extrêmement efficace : "il y a un problème en centre-ville, on peut y répondre avec les données parce qu'on a un outil" et un plan X,Y.”	““Et sur la difficulté d'intéresser les agents aux questions d'open data ? Le Canvas pourrait les aider ? Chef de projet open data/Montpellier - Très honnêtement, je ne pense pas parce que j'ai le temps de les voir un par un, de développer des stratégies, donc pour nos échelles, je ne suis pas sur.”	NA	NA	NA	NA

									<p>Preuve 2 - “Une idée intéressante. Il est certain que si on aborde l’open data en commençant par les questions de comprendre ce qu’on veut faire, si on commence par les enjeux data/IA/création de services, ça va être plus facile partout, et notamment pour ouvrir. On va pouvoir aller voir les personnes/élus qui détiennent les données et leur disant qu’on a besoin de leurs données pour telle ou telle raison.”</p>							
--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--

Table 7.5 : Tableaux sur l'utilité de l'Open Canvas sur le quatrième besoin fonctionnel

2-BF5-Flux de connaissance

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
2-BF5-Flux de connaissance	NA	Oui	Oui	Oui	NA	NA	“Le Canvas permettrait de mettre en commun des ressources”	“Ca (process 5 de l’Open Data Canvas) pourrait tout à fait nous aider à récupérer les outils et à les flécher vers les projets que l’on recense, que l’on connaît et qui pourraient en bénéficier”	“Mmm, en tant que tel, du peu que j’ai vu, le Canvas va venir prendre ce qui a été fait à un endroit et l’amener à un autre. Vous savez, on se rend compte régulièrement que quand on veut innover dans un domaine, cette innovation existe très certainement dans un autre domaine. Et là, si vous faites des appels à sollicitation, que vous standardisez les ressources pour que tout le monde puisse les comprendre, qu’elles alimentent des objectifs concrets, alors ça va commencer à bouger et tout le monde va vouloir fédéraliser, mutualiser.”	NA	NA	Non	L’utilité de l’outillage pour faire comprendre le potentiel des données est à relativiser selon les participants par rapport à celle d’avoir derrière ces outils des acteurs qui comprennent les enjeux de l’usage de l’open data.	NA	NA

Table 7.6 : Tableaux sur l’utilité de l’Open Canvas sur le cinquième besoin fonctionnel

2-BF6-Outil d'interaction

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
2-BF6-Outil d'interaction	NA	Oui	Oui	NA	NA	NA	“J’ai la sensation qu’il pourrait adresser un point clé : (...) les impliquer (élus, agents et services) continuellement pour qu’à la fin, la démarche devienne naturelle.	Preuve 1-D’après un participant : “là où on a du mal, c’est de dimensionner notre offre de service liée à notre portail(mise à disposition de données), puisqu’on ne connaît pas nos utilisateurs ou réutilisateurs, c’est difficile”. Or, le processus de sparring du Canvas implique d’identifier les contacts de réutilisateurs potentiels et le processus de crowdsourcing de les contacter pour comprendre les outils dont ils se servent pour faire usage des données. Ce qui répond directement à la barrière rencontrée par les acteurs. Preuve 2- D’après un participant : “On a commencé avec un petit cercle de techniciens pour travailler ensemble. Groupe informel un peu structuré et on s’est organisé pour travailler sur 3 thématiques (dont)expérimentation d’usages de la donnée. On va essayer de co-produire des innovations. On lance un travail collectif là-dessus autour d’un sujet qui est : les données des bornes de wifi public et voit ce qu’on va en faire”	NA	NA	NA	NA	Tout comme pour les plateformes open data, rien n’oblige les réutilisateurs (membres du système de transformation) à répondre aux sollicitations mises en place lors du processus de crowdsourcing ni à s’inscrire sur la chambre de conception dédiée pour interagir avec les producteurs ou intermédiaires de données.	NA	NA

Table 7.7 : Tableaux sur l'utilité de l'Open Canvas sur le sixième besoin fonctionnel

3-BF7-Comparaison qualité

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
#Focus group															
3-BF7-Comparaison qualité	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 7.8 : Tableaux sur l'utilité de l'Open Canvas sur le septième besoin fonctionnel

3-BF8-Tri

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
3-BF8-Tri	Oui	NA	NA	Oui	Oui.	<p>D’après un participant : “Après quelques galères au cas par cas, notamment avec l’anonymisation et les décrets qui sortent qui fait qu’on doit tout reprendre.”</p> <p>Or, le Canvas met à disposition un modèle de méta-données permettant de classer les ressources relatives à l’usage des données et une infrastructure open data qui met à disposition ces ressources.</p> <p>Ce qui répond directement au besoin exprimé par les acteurs.</p>	NA	NA	<p>D’après un participant : “on va organiser la première matinées des Référénts et Producteurs de données de Haute-Garonne, l’idée étant de créer un réseau de tout un tas d’acteurs, pour que chaque territoire s’approprie les données et soit à l’initiative sur certains domaines. (...) mais ça risque d’être difficile de trouver ne serait-ce qu’un groupe de 2 ou 3 individus dans la fonction publique qui soient près à tenter quelque chose. Il y a un problème d’ego-centrisme.”</p> <p>Or, le Canvas met justement en place un support pour des réseaux d’information, avec une entrée thématique, et qui peut rassembler les personnes expertes et intéressées par un domaine de données.</p> <p>Ce qui répond directement au besoin exprimé par les acteurs.</p>	<p>Preuve 1-D’après un participant : “on a eu nos difficultés, on a eu du mal à trouver des outils techniques pour faire la diffusion des outils. On ne savait faire la différence entre l’usine, la co-construction des outils et la galerie commerciale où on diffuse tout ça.”</p> <p>Or, le Canvas met en place un outil de recensement, de filtrage et de diffusion d’outils permettant d’effectuer certaines tâches (voir process métier)</p> <p>Ce qui répond directement au besoin exprimé par les acteurs.</p>	<p>“Mais on est ralenti, c’est tout. Pas de soucis majeurs sur le process.”</p>	NA	NA	NA	NA

											<p>Preuve 2-“Il y a une demande à laquelle on a toujours pas répondu et qu’on est en train de construire : une de nos demandes a été de créer des groupes thématiques au sein d’ODF, dédiés aux département et dédiés aux régions, avec que des référents open data à ce niveau là pour qu’ils s’entraident. L’idée ça serait que untel dise "Moi, en tant que région, j’ai tel problème" et qui demandent si tout le monde à le même pour créer un groupe et qu’elles s’entraident et partagent entre elles. Elles ont aussi des problématiques et des marges d’amélioration avec de la normalisation sur des données spécifiques, des compétences pour rendre ces données plus exploitables. Or, le Canvas met justement en place un support pour des réseaux d’information, avec une entrée thématique, et qui peut rassembler les personnes expertes et intéressées par un domaine de données.</p> <p>Ce qui répond directement au besoin exprimé par les acteurs.</p>						
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 7.9 : Tableaux sur l’utilité de l’Open Canvas sur le huitième besoin fonctionnel

4-BF9-Parcours de tâches avec objectifs annoncé

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
4-BF9-Parcours de tâches avec objectifs annoncé	Oui	Oui	Oui	Oui	NA	“Pour des territoires ou EPCI beaucoup plus petit, avec beaucoup de communes et le gars qui doit s’occuper de l’open data a plein d’autres choses à faire : oui, il peut très clairement utiliser la ressource. Ca peut être soit des territoires complètement ruraux, et c’est le département qui doit bouffer les 200-300 communes, et donc là, une aide extérieure peut être complètement utile avec de la ressource pour traiter la donnée. Après, sur des EPCI plus gros, le mec de l’informatique, fait en même temps l’open data et le déploiement de la fibre optique, lui il va avoir besoin d’autant d’aide extérieure que possible.”	“J’aime bien l’idée de mettre la question de l’utilisation final des données au tout début. Nous, sur l’Open data en général, on préfère se focus sur les usages. Ca nous évite d’aller voir les agents avec la loi pour une République numérique sous le bras. Le Canvas permet d’identifier précisément le besoin auquel l’open data vient répondre. On contourne "l’obligation vécue comme un frein", on arrive en parlant d’un besoin que les agents ont.”	Preuve 1- D’après un participant : “On imagine un accompagnement de mise en forme des données, pour que les collectivités s’auto-outillent. Alors on est pas tout seul, j’ai initié dans la région une coopération avec collectivités déjà avancées : les métropoles (Orléans et Tour sur les démarches de ville intelligente), départements et les agglos de Chartre, etc”. Or , le processus 4 du Canvas implique de créer un parcours de tâches avec les outils pour les réaliser. Ce qui répond directement au besoin exprimé par les acteurs. Preuve 2-“Quand on commence à donner à voir, avec un plan, alors on gagne un auditoire, et ça nous permet d’avancer, pourvu que ça soit utile, ça serait génial.”	Preuve 1 - “Oui, ce Canevas donne un moyen supplémentaire de travailler à plusieurs autour des données” Preuve 2 - D’après un participant : “L’idée c’est d’essayer de se répartir les tâches, et d’attribuer les tâches à ceux qui ont le plus de force pour les mener. Par exemple, on travaille avec la région sur ce point là, pour mettre en place un système fédéraliste de partage des tâches et des compétences, et ceux, en réussissant à ne jamais réintroduire un système central (...) et on essaye de leur (collectivites) faire quitter le côté leadership, centralisateur pour qu’ils prennent le pli de la coopération.”. Or , le Canvas prévoit la création d’une table d’association “Acteurs-Composant-Tâche” mise en oeuvre par intervenant tiers en SI, extérieur aux organisations centralisatrices dont l’objectif est de favoriser les interactions entre les acteurs. Ce qui répond directement au besoin exprimé par les acteurs.	NA	NA	Non	Non	NA	NA

Table 7.10 : Tableaux sur l’utilité de l’Open Canvas sur le neuvième besoin fonctionnel

4-BF10-Autonomiser travail évangélisation

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
4-BF10-Autonomiser travail évangélisation	NA	Oui	Oui	NA	NA	NA	“J’ai la sensation qu’il pourrait adresser un point clé : la construction de la confiance et de l’intérêt avec les services, il pourrait aider à effectuer un travail qui est énorme, à savoir amener les élus, les agents et services dans la démarche”	D’après un participant , difficile d’évangéliser les élus car “au niveau politique, si on parle technique, ça ne les intéresse plus”. Or , le draft du plan de développement des usages (process 4) se positionne à une échelle stratégique (du fait du lien entre tâches techniques et objectifs stratégiques) et remonte jusqu’aux élus. Ce qui répond directement à la barrière rencontrée par les acteurs.	NA	NA	NA	Le travail d’évangélisation est considéré par les participants comme principalement relationnel au sens d’un rapport quotidien entre un agent humain et un autre agent humain : “Un travail du quotidien, qui se base beaucoup dans le relationnel.”	NA	NA	NA

Table 7.11 : Tableaux sur l’utilité de l’Open Canvas sur le dixième besoin fonctionnel

5-BF11-Règles organisationnelles

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves				
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
#Focus group															
5-BF11-Règles organisationnelles	Non	NA	Oui	Oui	NA	D'après un participant, des mécanismes fonctionnels et utiles existent déjà pour cadrer l'organisation : "Ensuite on a la convention pour rassurer, qui fixe qui a le droit de faire quoi, les droits et les devoirs de chacun. On a essayé un petit peu de faire l'inventaire de ce qu'il était possible de faire. Ca cadre bien et ça rassure" Le Canvas est donc inutile	NA	D'après un participant, pour développer les usages au sein d'une institution : "il va falloir aller vite, être autonome, bousculer ou effacer les chaînes de décision hiérarchiques, mobiliser l'argent, se donner le droit à l'erreur, avoir le droit à l'échec". Or, le réseau d'information (mis en place au process 2, et support des process 5 et 7) se positionne justement hors des chaînes de décision hiérarchiques, permettant alors tout type d'entorses. Ce qui répond directement à la barrière rencontrée par les acteurs.	D'après un participant : "Le frein pour moi, c'est l'absence de structures collectives. Tant que les régions n'embrassent l'ensemble de leurs utilisateurs publics en leur disant : "on prend en charge la dynamique territoriale", ça ne marchera pas (...). Par contre, le risque c'est que tout le monde veuille jouer le rôle de fédérateur et que l'initiative perde en lisibilité, et qu'on retrouve une bataille pour la centralisation. Il faut une forme de neutralité." Or le réseau d'information est porté par une communauté neutre et hétérogène "teamopendata" qui alimente la dynamique de développement des usages. Ce qui répond directement au besoin exprimé par les acteurs.	NA	"Alors oui, ça (Canvas) pourrait servir à faire des feuilles de route avec un grand nombre de collectivités, de tisser des liens un à un avec ces partenaires. Mais pour l'instant, je ne communiquerai pas officiellement là-dessus, je ferai ça en one-to-one."	NA	L'absence de soutien hiérarchique officiel est considérée par les participants comme une entrave sérieuse à tout développement d'usage de l'open data.	NA	NA

Table 7.12 : Tableaux sur l'utilité de l'Open Canvas sur le onzième besoin fonctionnel

5-BF12-Principes cadres

Besoin Fonctionnel	Utilité du Canvas ?					Preuves					Contre-preuves					
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1	
#Focus group																
5-BF12-Principes cadres	NA	NA	NA	NA	Oui.	NA	NA	NA	NA		<p>D'après un participant : “Elles(région) cherchent aussi à savoir comment déployer et faire usage de l’open data à ces échelles là (échelles régionales). Il y a un vide à investiguer.”</p> <p>Or, le Canvas articule 3 niveaux d’action situés à différentes échelles territoriales : (intervenant/meta-designer /designer)</p> <p>Ce qui répond directement à la barrière rencontrée par les acteurs.</p>	NA	NA	NA	NA	NA

Table 7.13 : Tableaux sur l'utilité de l'Open Canvas sur le douzième besoin fonctionnel

7.5 Implications

7.5.1 Implication des résultats de l'évaluation de la version 0 du prototype

Dans la partie relative au design de l'évaluation de l'environnement de conception des usages (voir chapitre 2, section 2.2.1.6.4, p.195), j'ai spécifié ses trois objectifs :

1. Récupérer des données relatives à l'utilité de l'environnement de conception des usages sous la forme de feedbacks qualitatifs ;
2. Confirmer ou infirmer l'utilité du design de l'environnement de conception des usages, i.e sa capacité à répondre aux douze besoins fonctionnels ;
3. Récupérer des feedbacks permettant d'informer la construction d'une nouvelle version du prototype.

Le premier objectif de l'évaluation a été atteint et les *feedbacks* qualitatifs présentés dans la partie précédente. Dans cette partie, il sera question de confirmer ou d'infirmer l'utilité du design de l'environnement de conception des usages ainsi que de présenter les informations qui nous permettront de construire une nouvelle version du prototype.

7.5.1.1 Confirmation ou infirmation de l'utilité du design de l'environnement de conception des usages

D'après les résultats exposés dans la partie précédente, je peux tirer plusieurs conclusions relatives à l'utilité du design de l'environnement de conception des usages

En premier lieu, je peux raisonnablement affirmer que l'Open Data Canvas permet de répondre aux besoins fonctionnels présents dans le tableau ci-dessous.

Besoin Fonctionnel/Focus Group	FC 1	FC 2	FC 3	FC 4	FC 5	Utilité du Canvas
BF1	NA	NA	NA	Non	NA	Non
BF2	NA	NA	NA	NA	NA	Non
BF3	NA	Non	NA	NA	NA	Non
BF4	Oui	NA	Non	Oui	Oui	Oui
BF5	NA	Oui	Oui	Oui	NA	Oui
BF6	NA	Oui	Oui	NA	NA	Oui
BF7	NA	NA	NA	NA	NA	Non
BF8	Oui	NA	NA	Oui	Oui	Oui
BF9	Oui	Oui	Oui	Oui	NA	Oui
BF10	NA	Oui	Oui	NA	NA	Oui
BF11	Non	NA	Oui	Oui	NA	Non
BF12	NA	NA	NA	NA	Oui	Non

Table 7.14 : Utilité du Canvas relativement aux douze besoins fonctionnels

Autrement dit, l'environnement de conception des usages désigné dans ma thèse permet de lever toutes les barrières socio-techniques relatives à :

- l'évaluation du potentiel des données à partir de leur représentation ;
- la création d'un cadre cognitif partagé entre les systèmes.

Il permet également de lever certaines des barrières socio-techniques relatives à :

- la transformation des données grâce à des artefacts techniques.

En revanche, il ne permet pas de lever les barrières socio-techniques relatives à :

- l'utilisation des fonctionnalités des plateformes ;
- l'institutionnalisation de l'usage des données.

7.5.1.3 Présentation des *feedbacks* permettant d'informer la construction d'une nouvelle version du prototype.

Les *feedbacks* obtenus par l'intermédiaire des *focus group* m'ont permis de prendre trois décisions concernant le prototypage de la nouvelle version du prototype.

La première décision que je peux prendre sur la construction d'une nouvelle version du prototype concerne le troisième élément fonctionnel identifié lors du design de l'artefact : le manifeste et son codebook. Il n'est plus question de les prototyper. En effet les *focus group* ont mis en avant le fait qu'ils ne pouvaient adresser les trois besoins fonctionnels auxquels ils devait répondre (BF3, BF11, BF12).

Le deuxième décision concerne le prototypage du deuxième élément fonctionnel, à savoir le système de gestion des tâches basé sur un modèle de métadonnées. Aux deux fonctions déjà attribuées à cet élément, j'en ajoute une troisième : celle de générer automatiquement un plan de développement des usages des données. En effet, les participants des *focus group* ont évoqué à de nombreuses reprises l'utilité de pouvoir présenter aux acteurs d'un écosystème le plan de développement des usages et ce, pour initier tout dialogue avec eux (voir Table : utilité de l'Open Canvas sur chacun des 12 besoins fonctionnels). Dans ce cadre, il me paraît pertinent de permettre aux utilisateurs du Canvas de générer rapidement et par eux-mêmes un document écrit détaillant l'ensemble du plan de développement des usages. Dès lors, je rajouterai cette fonctionnalité dans la prochaine version du prototype.

Le troisième décision a trait au prototypage du premier élément fonctionnel, à savoir le réseau d'information. Plus précisément, elle concerne l'administration et le découpage des thématiques des chambres de conception dans un premier temps et des espaces de discussion, dans un second temps. Dans la prochaine version du prototype, je diviserai les chambres de conception en autant de projets de développement des usages que l'organisation meta-designer détient dans son portefeuille. Par exemple, il pourrait être question de diviser la chambre de conception d'une organisation en autant de projets *smart* qu'elle détient dans son portefeuille. Cette décision trouve sa justification dans le besoin des acteurs de rattacher l'usage des données ouvertes à des projets spécifiques existants ou à venir (voir Table 7.7).

A propos du découpage des espaces de discussion, je les diviserai dans la prochaine version du prototype en autant de domaine de données devant être utilisés dans le projet. Par exemple, un projet qui fait appel à des données démographiques et environnementales verra son espace de discussion divisé en deux fils de discussion dédiés chacun à un domaine de données. Cette décision trouve sa justification dans le fait que

les participants aux *focus group* ont fait part de leur besoin de pouvoir échanger avec d'autres acteurs sur la transformation de domaines de données spécifiques (voir Table 7.7). Il m'apparaît dès lors pertinent de permettre aux acteurs d'un écosystème d'avoir, dans chaque espace de discussion, des fils de discussion dédiés à chacun des domaines de données impliqués dans le projet.

Aussi sera-t-il question pour renforcer cet échange d'informations et de bonnes pratiques sur les domaines de données, de rendre obligatoire l'attachement d'une étiquette à tout fil de discussion créé sur le réseau d'information, choisies parmi une liste de domaines de données. En apposant une telle étiquette, il sera possible pour les utilisateurs du Canvas de trouver toutes les informations et bonnes pratiques disponibles relativement à un domaine de données avec lequel ils sont aux prises dans leur projet.

Conclusion du chapitre 7

L'objectif de ce chapitre était d'évaluer dans quelle mesure l'Open Data Canvas permettait de lever les barrières socio-techniques qui freinent la construction des usages issus des données ouvertes.

La première, deuxième, troisième et quatrième section de ce chapitre m'ont permis de choisir une stratégie et une méthode d'évaluation pertinente eu égard aux caractéristiques des prototypes (Chapitre 7). Eu égard à la version 0 du prototype, je conclus en choisissant une stratégie d'évaluation *ex post* et naturaliste, déployée grâce à cinq focus group rassemblant plus de 15 participants inspirés des méthodes utilisées en ergonomie informatique.

La cinquième section m'a permis de synthétiser les résultats des cinq *focus group*. Grâce à eux, je peux conclure que l'environnement de conception des usages designé dans ma thèse permet de lever toutes les barrières socio-techniques relatives à :

- l'évaluation du potentiel des données à partir de leur représentation ;
- la création d'un cadre cognitif partagé entre les systèmes.

Il permet également de lever certaines des barrières socio-techniques relatives à :

- la transformation des données grâce à des artefacts techniques.

En revanche, il ne permet pas de lever les barrières socio-techniques relatives à :

- l'utilisation des fonctionnalités des plateformes ;
- l'institutionnalisation de l'usage des données.

Enfin, la sixième question m'a permis d'établir les implications qu'ont les résultats de l'évaluation sur la construction de nouvelles versions du prototype qui auront lieu après cette thèse. Je conclus que les prochaines versions du prototype de l'Open Data Canvas peuvent faire l'économie de son troisième élément fonctionnel, à savoir son manifeste et son codebook ; aussi devraient-elles pouvoir générer automatiquement un plan de développement des usages des données ; encore, ils devraient permettre de diviser les chambres de conception en autant de projets de développement des usages que l'organisation meta-designer détient dans son portefeuille ; à propos du découpage des espace de discussion, elles devraient permettre de les diviser en autant de domaine de données devant être utilisés dans le projet ; enfin, elles devraient rendre obligatoire l'attachement d'une étiquette à tout fil de discussion créés sur le réseau d'information, choisie parmi une liste finie de domaines de données.

Chapitre 8

Discussion et conclusion générale

Chapitre 8. Discussion et conclusion générale

8.1 Résultats marquants de la recherche

8.1.1 Résultats empiriques de la recherche

Dans cette section, je présente les réponses apportées aux trois sous-questions de recherche, avant d'exposer ma réponse à la question centrale.

8.1.1.1 Première sous-question de recherche : état des lieux de la construction des usages issus des données ouvertes

En posant cette première question de recherche, mon ambition était d'identifier comment la littérature met en relation les quatre construits qui définissent la construction des usages des données ouvertes, à savoir les données ouvertes, l'usage, les plateformes et les écosystèmes de données ouvertes. La réponse à cette question devait me permettre d'identifier les zones d'ombre dans la littérature et les gaps que je choisirai d'exploiter.

J'ai construit un schéma qui montre la relation entre les différents concepts (voir Figure 1.2, chapitre 1, p.38). Il établit que :

- un écosystème *open data* constitue l'environnement dans lequel la construction des usages se déroule
- la plateforme *open data* apparaît comme une structure modulaire au sein de l'écosystème qui maximise la densité des données ouvertes et leur liquidité à un endroit donné
- les données ouvertes sont aux écosystèmes *open data* ce que les espèces ingénieurs sont aux écosystèmes biologiques : un matériau vivant dont la circulation structure et assure l'équilibre d'un écosystème.
- les usages issus des données ouvertes sont une construction socio-technique qui émerge des interactions entre les différentes sphères sociales présentes au sein des écosystèmes *open data*

De surcroît, j'ai proposé un modèle conceptuel permettant de comprendre les différentes couches ou niveau de compréhension qui composent la construction des usages :

- A propos de la structure statique de l'écosystème dans lequel la construction des usages a lieu (figures 3.1, 3.2 et 3.3), **je conclus que la construction des usages a lieu dans un écosystème alimenté par une plateforme *open data* qui diffuse des données agissant comme une espèce ingénieur en structurant les interactions entre trois systèmes (système de production, d'intermédiation et de réutilisation de données).**
- A propos de la structure statique de chacun des trois systèmes, c'est à dire des artefacts et acteurs impliqués, **je conclus que la construction des usages relève de transformations de jeux de données dans un état initial en jeux de données dans un état final et ce, par la mise en oeuvre de capacités d'action spécifiques d'artefacts sociaux et techniques dont le rôle est de faire passer les jeux de données d'un état à un autre.**
- A propos du comportement dynamique de chacun des trois systèmes, visant à comprendre comment les artefacts et acteurs impliqués se mettent en mouvement pour construire les usages, **je conclus que la construction résulte de capacités d'action qui s'articulent selon des séquences d'articulation définies comme une suite particulière dans lesquelles ces dernières interagissent, se renforcent mutuellement et mettent en place les conditions propices au changement d'état des jeux de données** (figure 3.4).
- Enfin, à propos de l'évolution de la structure globale de l'écosystème, **je conclus que les écosystèmes passeraient ainsi par quatre phases : une phase de genèse, une phase d'expansion, une phase de leadership et une phase d'auto-renouvellement** (figure 3.5).

Enfin, j'ai pu établir que si la littérature scientifique recense un grand nombre d'études qui permet de comprendre les dynamiques d'usage des données ouvertes sur une seule dimension (ou niveau d'analyse), rares sont les auteurs qui ont combiné les niveaux d'analyse et proposé une méthodologie pour attaquer l'usage depuis cet angle de vue. Je précise cette observation en établissant qu'encore plus rares sont les auteurs étant partis du point de vue que les niveaux d'analyse sont en tension paradoxale quand il est question d'aborder l'usage des données ouvertes. J'ai construit ma thèse en exploitant ce gap spécifique.

8.1.1.2 Deuxième sous-question de recherche : les situations problématiques face à la construction des usages

Derrière cette deuxième sous-question, il était question pour moi d'identifier les situations problématiques auxquelles les usagers font face et qui les bloquent dans leurs démarches de construction. Aussi, j'avais pour ambition d'identifier les besoins auxquels un dispositif devrait répondre pour constituer une solution aux situations problématiques.

Pour répondre à cette question j'ai pu préciser le modèle conceptuel et établir que la construction des usages se subdivise en cinq étapes :

1. Utilisation des fonctionnalités de la plateformes
2. Evaluation du potentiel des données à partir de leur représentation
3. Transformation des données grâce à des artefacts techniques
4. Création d'un cadre cognitif partagé entre les systèmes pour les coordonner
5. Institutionnalisation de l'usage des données

Je conclus qu'au cours de ces cinq étapes, les usagers des données ouvertes font face à quatorze situations problématiques (table 3.37) prenant la forme de barrières socio-techniques et qu'on peut les classer selon deux catégories : les barrières socio-techniques relatives aux interaction homme-machine¹ et les barrières socio-techniques relatives aux interactions sociales entre les opérateurs des données ouvertes.

Etape d'usage	Barrières socio-techniques/Situations problématiques
1.Utilisation des fonctionnalités de la plateformes <i>téléchargement/flux de données</i>	Manque de confiance dans la plateforme
	Plateforme disponible dans une seule langue
	Manque de lisibilité des interfaces
2.Evaluation du potentiel des données à partir de leur représentation	Manque de connaissance du domaine relatif aux données
	Manque de compétences et d'expertise technique
	Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation
3.Transformation des données grâce à des artefacts techniques	Format de données illisible par machine
	Mauvaise qualité des données

¹ Termes originaux : *Human-Computer Interaction*

	Hétérogénéité des condition d'utilisation des données
4.Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	Manque d'un cadre juridique et politique clair
	Manque d'une vision cohérente autour de la production et de l'utilisation des données
	Le classement de l'open data en bas de la liste des priorités
5.Institutionnalisation de l'usage des données	L'incompatibilité des structures organisationnelles avec l'ouverture des données
	Divergence culturelle de l'ouverture des données par rapport aux cultures organisationnelles

Table 8.1 : Framework d'identification des barrières face à l'usage de l'open data

J'ai également pu établir avec les usagers des données ouvertes une liste de douze besoins fonctionnels auxquels un dispositif en Systèmes d'Information devrait répondre pour constituer une solution aux quatorze situations problématiques :

1. Le dispositif en système d'information doit supporter la documentation du processus de production de donnée
2. Le dispositif en système d'information doit supporter la documentation de tous les mécanismes techniques ou juridiques qui fixent des restrictions sur les conditions d'utilisation des données
3. Le dispositif en système d'information doit faciliter le design de plateformes *open data* à double entrée, la première étant réservée aux développeurs et experts de la données et la deuxième réservée au grand public.
4. Le dispositif en système d'information doit faciliter la sensibilisation des acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre
5. Le dispositif en système d'information doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel.
6. Le dispositif en système d'information doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre
7. Le dispositif en système d'information, à savoir qu'il doit offrir des outils permettant de comparer la qualité des données entre elles et surtout dans le temps.

8. Le dispositif en système d'information doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité
9. Le dispositif en système d'information doit fournir un parcours de tâches concrètes et collectives à mettre en œuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles.
10. Le dispositif en système d'information doit offrir des outils permettant d'automatiser une partie du travail d'évangélisation.
11. Le dispositif en système d'information doit fournir un exemple de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes.
12. Le dispositif en système d'information doit fournir des exemples de principes qui cadrent la manière de concevoir l'ouverture des données.

8.1.1.3 Troisième sous-question de recherche : les éléments d'un dispositif en Systèmes d'Information pour résoudre les situations problématiques

Enfin, dans cette troisième sous-question il était question de savoir quels sont les éléments qui constituent un dispositif en Systèmes d'Information permettant de résoudre les situations problématiques.

Je propose d'intégrer trois éléments fonctionnels dans un dispositif en Systèmes d'Information prenant la forme d'un environnement de conception (ou artefact de *meta-design*) : un réseau d'information, un système de gestion des tâches basé sur un modèle de méta-données ainsi qu'un manifeste et son codebook.

Relativement à ces trois éléments fonctionnels, j'ai établi trois propositions de design :

1. Étant donné le processus de construction des usages des données ouvertes, et dans l'objectif de lever les barrières relatives à **l'utilisation des fonctionnalités d'une plateforme open data**, la **transformation des données grâce à des artefacts techniques** et la **création d'un cadre cognitif partagé entre les systèmes**, il est pertinent de mettre en oeuvre **un réseau d'information**.
2. Étant donné le processus de construction des usages des données ouvertes, et dans l'objectif de lever les barrières relatives relatives à **l'évaluation du potentiel des données à partir de leur représentation**, la **transformation des données grâce à des artefact techniques** et la **création d'un cadre cognitif partagé entre les systèmes**, il est pertinent de mettre en oeuvre **un système de gestion des tâches basé sur un modèle de métadonnées**.

3. Étant donné le processus de construction des usages des données ouvertes, et dans l'objectif de lever les barrières relatives à **l'utilisation des fonctionnalités de la plateforme *open data*** et à **l'institutionnalisation de l'usage des données**, il est pertinent de mettre en œuvre **un manifeste et son codebook**.

Sur la base de ces trois propositions de design, j'ai développé douze principes de design qui m'ont informés sur la manière d'implémenter les trois éléments de manière à ce qu'ils soient le plus efficace possible, c'est à dire en capacité de lever au maximum les barrières qui bloquent la construction des usages. Ces douze principes ont été construits grâce à quatre *kernel theory* et que j'ai défini avec Hanseth et Lyytinen (2004) comme "les théories issues des sciences naturelles et sociales qui gouvernent les exigences de conception ou le processus qui permet de les construire". Les théories du paradoxe, la théorie de la construction sociale de la technologie, la théorie de la cognition distribuée et la théorie des *middleground* ont permis de construire les douze besoins fonctionnels (voir Annexe 9).

En collaboration avec la société coopérative Dataactivist, j'ai pu construire le design d'un environnement de conception des usages nommé Open Data Canvas. Pris dans son ensemble, l'Open Data Canvas incorpore deux mécanismes :

- un processus de dialogue entre un intervenant en Systèmes d'Information et l'utilisateur du Canvas dit sparring
- une boucle de rétroaction permettant à l'utilisateur d'avoir des *feedbacks* sur ses actions de la part des membres de l'écosystème *open data* auquel il appartient.

Ensuite, j'ai développé avec Dataactivist un réseau d'information qui incorpore trois mécanismes :

- une surdétermination de l'usage du réseau qui permet aux utilisateurs du Canvas d'acquérir un certain savoir-faire en matière d'interaction autour des données ;
- un mécanisme de gestion collaborative des composants du réseau permettant aux utilisateurs d'acquérir une culture pratique de la donnée ;
- un mécanisme permettant de représenter la base de données relationnelle des composants du réseau sous la forme d'un tableur.

J'ai également développé avec Dataactivist un système de gestion des tâches à partir d'un modèle de métadonnées qui incorpore quatre mécanismes :

- Le modèle de métadonnées en lui-même qui permet de structurer le système de gestion des tâches
- Une flexibilité organisationnelle du système d'organisation des tâches qui permet de réorganiser le système de gestion des tâches de manière à ce qu'il corresponde aux priorités du processus de construction des usages
- Une plasticité des dispositions spatiales de la représentation des tâches, mécanisme qui permet de représenter la base de données relationnelle de plusieurs manières à partir d'une interface sous forme de tableur.
- Un mécanisme permettant de faire le lien entre les tâches présentes dans le système et des témoignages d'acteurs ayant effectué ces tâches.

Enfin, j'ai développé un manifeste et son codebook qui intègre deux mécanismes :

- un mécanisme de sourcing et d'intégration constante de nouvelles pratiques dans le manifeste et le codebook
- un mécanisme permettant à l'utilisateur de développer des produits (sites web par exemple) ou des services (assistance à réutilisateurs de données) à partir du manifeste et son codebook.

J'ai finalement développé avec Dataactivit deux prototypes de l'Open Data Canvas. La première version du prototype a été évaluée. Cette évaluation me permet de conclure que l'environnement de conception des usages construit dans ma thèse permet de lever toutes les barrières relatives à :

- l'évaluation du potentiel des données à partir de leur représentation ;
- la création d'un cadre cognitif partagé entre les systèmes.

Il permet également de lever certaines des barrières relatives à :

- la transformation des données grâce à des artefacts techniques.

En revanche, il ne permet pas de lever les barrières relatives à :

- l'utilisation des fonctionnalités des plateformes
- l'institutionnalisation de l'usage des données.

Je peux donc (étant donné l'état de développement du Canvas au moment de rédiger cette thèse) valider complètement la deuxième proposition de design, partiellement la première proposition de design et invalider la troisième proposition.

8.1.1.4 Question de recherche centrale : une réponse aux tensions paradoxales faisant obstacle à la construction des usages

J'ai commencé cette thèse avec pour objectif d'apporter une réponse aux tensions paradoxales faisant obstacle à la construction des usages issus des données ouvertes.

J'ai commencé avec une revue de littérature (chapitre 1) dans laquelle j'ai pu établir que les écosystèmes *open data* sont des environnements dans lesquels les liens et tensions entre acteurs sont en perpétuelle évolution. J'ai pu également noter que dans ces écosystèmes, les données ouvertes sont très peu utilisées et donc peu de nouveaux usages construits grâce à elles. Ce point apparaît aux praticiens comme une réelle problématique.

Grâce à une étude exploratoire de quatre cas d'usage des données ouvertes (chapitre 3), je peux affirmer et prouver que les données sont très peu utilisées du fait de blocages prenant la forme de barrières socio-techniques et qui bloquent d'une part les interactions entre acteurs, d'autre part les interactions homme-machine².

Deux études de cas exploratoires m'ont permis de trouver une solution locale à ces blocages, c'est-à-dire pouvant s'appliquer à deux écosystèmes particuliers : la métropole Grenoble-Alpes et la métropole du Grand Lyon (chapitre 4). **Cette solution consiste non pas à vouloir résoudre ou supprimer les tensions mais plutôt à les alimenter.** Pour cela, je propose de mettre en place un dispositif en Systèmes d'Information prenant la forme d'un environnement de conception ou artefact de meta-design.

J'ai ensuite généralisé cette solution de manière à ce qu'elle puisse s'appliquer à tous les écosystèmes *open data* (chapitre 5 et 6). Cette solution s'appelle Open Data Canvas et permet aux acteurs d'alimenter eux-mêmes les liens et tensions inhérents à leurs interactions d'une part, et aux interactions homme-machine d'autre part.

J'ai finalement évalué auprès de cinq écosystèmes (autres que celui de la métropole Grenoble-Alpes et de la métropole du Grand Lyon) la capacité de l'Open Data Canvas à permettre aux acteurs d'alimenter eux-même ces tensions (chapitre 7).

L'Open Data Canvas permettant aux praticiens d'alimenter eux-mêmes les liens et tensions inhérents aux interactions entre eux d'une part, et aux interactions homme-machine d'autre part, j'ai apporté une solution à question de recherche centrale de cette thèse.

² *Human-Computer interaction*

8.1.2 Résultats théoriques de la recherche

8.1.2.1 Une *design theory* pour les environnements de conception des usages

Pour aboutir à l'Open Data Canvas, j'ai sollicité une méthodologie de recherche en *design science*. Cette méthodologie m'a permis d'aboutir à une *design theory*, c'est-à-dire à une théorie permettant de concevoir et agir. Cette théorie permet en l'occurrence de concevoir un environnement de conception des usages issus des données ouvertes et d'agir pour favoriser la construction de ces mêmes usages.

D'après le littérature (Chapitre 1, figure 1.17) une *design theory* est composée de 8 éléments, décrits ci-dessous avec les contributions de ma thèse sur chacun d'entre eux.

8.1.2.1.1 Périmètre et objectif de la théorie

Le périmètre et objectif de la théorie vise à définir le type d'artefact à laquelle la théorie s'applique et les frontières du domaine auquel la théorie s'applique.

La théorie s'applique ici aux artefacts de meta-design (i.e, artefacts au service du design d'autres artefacts) qui visent à favoriser la construction des usages issus des données ouvertes. Elle s'applique à un processus particulier de construction des usages qui se subdivise en cinq étapes :

1. Utilisation des fonctionnalités de la plateformes
2. Evaluation du potentiel des données à partir de leur représentation
3. Transformation des données grâce à des artefacts techniques
4. Création d'un cadre cognitif partagé entre les systèmes pour les coordonner
5. Institutionnalisation de l'usage des données

Aussi cette théorie n'aborde-t-elle donc pas toutes les étapes que les producteurs de données doivent suivre avant la publication de jeux de données sur une plateforme. Elle exclut toutes les étapes que les producteurs de données doivent suivre pour identifier en interne les jeux de données ouverts, choisir une plateforme *open data* pour diffuser leurs données, mettre en place cette plateforme, choisir une licence pour les données.

8.1.2.1.2 Construits

Les construits sont les entités conceptuelles qui représentent l'objet de la *design theory*, en l'occurrence la construction des usages issus des données ouvertes.

Les construits de cette *design theory* sont les artefacts de meta-design, les données ouvertes, les plateformes *open data* et les écosystèmes *open data* et les usages des données ouvertes.

Les artefacts de meta-design sont définis comme des *infrastructures techniques et sociales dans laquelle de nouvelles formes de design collaboratifs peuvent prendre forme.* ; les données ouvertes comme des *objets informationnels vivant dont la circulation structure et assure l'équilibre d'un écosystème* ; les plateformes *open data* comme *une structure modulaire au sein de l'écosystème qui maximise la densité des données ouvertes et leur liquidité à un endroit donné* ; les écosystèmes *open data* comme les environnements dans lesquels la **construction des usages** se déroule ; enfin les usages issus des **données ouvertes** comme des constructions socio-technique qui émergent des interactions entre les différentes sphères sociales présentes au sein des **écosystèmes open data**.

8.1.2.1.3 Principes de structure, de forme et de fonction

Les principes de structure, de forme et de fonction sont les éléments constitutifs d'un schéma directeur pour décrire et produire la forme et la fonction finale de l'artefact produit.

J'ai identifié quatorze besoins fonctionnels auxquels l'environnement de conception doit répondre. J'ai proposé trois éléments fonctionnels pour y répondre : un réseau d'information, un système de gestion des tâches basé sur un modèle de métadonnées, un manifeste et son codebook. Douze principes de design ont été construits pour faciliter l'implémentation de ces éléments.

L'Open Data Canvas est l'artefact qui conclut cette thèse. Il est composé d'un design de son système (voir partie 5.3), de trois scénarios d'usage ainsi que de 28 fonctions qui décrivent son fonctionnement.

8.1.2.1.4 Plasticité de l'artefact

La plasticité de l'artefact définit la mesure dans laquelle l'artefact peut s'adapter à diverses situations.

L'Open Data Canvas peut s'adapter à trois scénarios d'usage : il peut être utilisé pour rallier les communautés autour de définitions socles de l'usage des données ouvertes, pour centraliser et mutualiser

les ressources utiles à la construction des usages et enfin pour mettre à disposition du plus grand nombre une forge des usages des données ouvertes.

8.1.2.1.5 Propositions testables

Les propositions testables constituent les vérités concernant la *design theory* pouvant être testées.

Ma thèse propose trois propositions testables :

1. Il est pertinent de mettre en oeuvre un réseau d'information pour lever les barrières relatives à l'utilisation des fonctionnalités d'une plateforme open data , la transformation des données grâce à des artefacts techniques et la création d'un cadre cognitif partagé entre les systèmes;
2. Il est pertinent de mettre en oeuvre un système de gestion des tâches basé sur un modèle de métadonnées pour lever les barrières relatives relatives à l'évaluation du potentiel des données à partir de leur représentation, la transformation des données grâce à des artefact techniques et la création d'un cadre cognitif partagé entre les systèmes,
3. Il est pertinent de mettre en œuvre un manifeste et son codebook pour lever les barrières relatives à l'utilisation des fonctionnalités de la plateforme *open data* et à l'institutionnalisation de l'usage des données.

Cette thèse ajoute qu'il faut combiner les mécanismes des réseaux d'information, du système de gestion des tâches basé sur un modèle de métadonnées et du manifeste et du codebook pour aboutir à un environnement de conception qui puisse répondre aux trois propositions de design. L'évaluation du dispositif (chapitre 7) supporte partiellement la première proposition, totalement la deuxième et rejette la troisième. En effet, l'évaluation prouve que l'Open Data Canvas, dans le version 0 de son prototype, permet de lever toutes les barrières relatives à l'évaluation du potentiel des données à partir de leur représentation et à la création d'un cadre cognitif partagé entre les systèmes. Il permet également de lever certaines des barrières relatives à la transformation des données grâce à des artefacts techniques. En revanche, il ne permet pas de lever les barrières relatives à l'utilisation des fonctionnalités des plateformes et à l'institutionnalisation de l'usage des données.

8.1.2.1.6 Connaissances justificatives

Connaissances justificatives sont les théories ou connaissances provenant des sciences naturelles, sociales ou de design et qui donnent une base et une explication à la capacité de l'artefact à agir dans le sens

souhaité par le concepteur. J'utilise quatre *kernel theory* : les théories du paradoxe, théorie de la construction sociale des technologies, théorie de la cognition distribuée, théorie des *middleground*.

J'ai extrait des théories du paradoxe deux mécanismes qui permettent aux usagers des données ouvertes d'identifier les tensions paradoxales auxquelles ils font face et d'alimenter leur évolution. De la théorie de la construction sociale des technologies, j'ai extrait trois mécanismes qui permettent de supporter la réalisation des tâches collaboratives relatives à la construction des usages. De la théorie de la cognition distribuée, j'ai extrait quatre mécanismes permettant de faciliter la distribution et l'absorption des informations et connaissances nécessaires à la réalisation des tâches relatives à la construction des usages. Enfin, j'ai extrait de la théorie des *middleground* deux mécanismes permettant d'institutionnaliser les nouveaux usages construits.

8.1.2.1.7 Principes d'implémentation

Les principes d'implémentation correspondent à une description des processus permettant d'implémenter la *design theory* dans des contextes de conception particuliers.

J'ai développé douze principes de design. Quatre principes de design adressent le fonctionnement global de l'Open Data Canvas. Trois principes adressent le fonctionnement du réseau d'information. Trois autres principes adressent le système de gestion des tâches. Les deux derniers principes adressent le fonctionnement du manifeste et de son codebook.

8.1.2.1.8 Prototype exposé

Le prototype exposé correspond à la forme physique de l'artefact. Il incarne la *design theory* et permet de l'évaluer auprès des praticiens. J'ai développé deux versions du prototype de l'Open Data Canvas. La première version (dite version 0) prend la forme de trois cartographies d'architecture de l'Open Data Canvas :

- **La cartographie de l'architecture métier** qui décrit, sous forme de diagramme, les processus que l'Open Data Canvas vient supporter. Chaque processus se subdivise en un enchaînement d'activités représenté par des diagrammes de workflow.
- **La cartographie de l'architecture fonctionnelle** qui décrit les interactions entre les vingt-huit fonctions incorporées dans l'Open Data Canvas. Elle a pour principale utilité de décrire les fonctions qui permettent la mise en place des sept processus métiers mentionnés dans la cartographie de l'architecture métier.

- **La cartographie de l'architecture applicative** qui décrit les blocs applicatifs, la structure des données utilisées et leurs interactions, qui correspondent directement à la cartographie de l'architecture fonctionnelle.

Cette version du prototype a été évaluée. Des implications de cette évaluation j'ai construit une deuxième version du prototype dite version alpha et ce à partir de trois outils logiciels :

- *Discourse* : un logiciel libre de conception de forum de discussion. Il s'agira de prototyper la version alpha directement sur le domaine #teamopendata qui appartient à Dataactivist et qui rassemble déjà plus de 800 experts des données ouvertes. **Ce logiciel permet d'assurer toutes les fonctions relatives au réseau d'information conçu.**
- *Airtable* : application permettant de designer des bases de données relationnelles à partir de tableurs. **Ce logiciel permet d'assurer toutes les fonctions relatives au système de gestion des tâches basé sur un modèle de métadonnées.**
- *Zapier* : application permettant de faire communiquer des logiciels et applications web, en l'occurrence Discourse et Airtable, par l'intermédiaire de la configuration d'API. **Ce logiciel permet d'assurer toutes les fonctions qui nécessitent une interopérabilité entre Discourse et Airtable.**

8.1.2.2 Les théories du paradoxe pour articuler et opposer les trois *kernel theory* sollicitées

La revue de littérature initiale nous a montré que la recherche sur la construction des usages a tendance à étudier séparément les différents niveaux d'analyse impliqués ou à les superposer. De même, les auteurs tendent à affilier à chaque niveau d'analyse une théorie qu'ils étudient séparément. Or, ma thèse tend à prouver qu'il est essentiel, pour développer un environnement de conception, d'articuler les différents niveaux d'analyse ainsi que les théories. En l'espèce, ma thèse articule trois *kernel theory* et de fait, elle articule les niveaux d'analyse que chacune d'entre elle sollicite : il s'agit de la théorie de la cognition distribuée, de la théorie de la construction sociale de la technologie et de la théorie du *middleground* :

- La théorie de la cognition distribuée sollicite une combinaison entre le niveau d'analyse intra-organisationnel (pour étudier les processus cognitifs) et le niveau inter-organisationnel (pour étudier l'environnement dans lequel la cognition est distribuée)
- La théorie de la construction sociale de la technologie sollicite majoritairement un niveau d'analyse extra-organisationnel (pour étudier les groupes significatifs qui influent sur la

construction d'une technologie) et de manière subsidiaire le niveau d'analyse sociétal (pour étudier l'impact sur la société des luttes entre groupes significatifs pour imposer une conception de la technologie construite)

- La théorie du *middleground* qui combine le niveau extra-organisationnel (pour étudier les communautés de l'*underground* et le *middleground*), le niveau organisationnel et le niveau industriel et sociétal (pour étudier le passage du *middleground* vers l'*upperground*)

Pour articuler ces trois *kernel theory*, je sollicite les théories du paradoxe. Cette articulation, sous l'égide de la théorie des paradoxes, s'incarne à travers le fonctionnement et l'utilisation de l'Open Data Canvas. En l'espèce, toute utilisation du Canvas suppose qu'un acteur réfléchisse aux tensions paradoxales qui animent et freinent l'écosystème *open data* dans lequel il se situe. C'est sur la base de la compréhension que l'acteur acquiert sur les tensions paradoxales (et donc sur l'acquisition de connaissances sur les combinaisons et oppositions entre niveaux d'analyse) qu'il conçoit un plan de développement des usages.

8.1.3 Implications de mes résultats

8.1.3.1 Apports pratiques

Le premier apport pratique se traduit par l'utilisation qu'un leader d'un écosystème *open data* pourrait faire du framework d'identification des barrières socio-techniques (Table 3.37). En interrogeant les membres de son écosystème sur les 5 dimensions et sur les quatorze critères que le framework comporte, il peut identifier les barrières socio-techniques qui freinent la construction des usages et les utiliser comme levier pour construire des solutions.

Le deuxième apport s'adresse à d'autres praticiens, en l'espèce toute organisation qui chercherait à développer dans un écosystème un environnement de conception des usages. Il se traduit à travers la liste de besoins fonctionnels auxquels un dispositif visant à faciliter la construction des usages doit répondre. Cette liste peut faciliter, renforcer et accélérer le processus d'identification des besoins à travers lequel ces organisations doivent passer avant d'envisager tout développement technique. De même ces praticiens peuvent-ils utiliser les principes de design à disposition pour guider leur développement technique et de manière à s'assurer de l'utilité et de l'efficacité des dispositifs techniques créés.

Enfin, l'apport pratique majeur de cette thèse s'incarne à travers les cartographies d'architectures métier, fonctionnelle et applicative. Elles peuvent être transmises en qualité de *blueprint* à des développeurs informatiques pour développer des environnements de conception des usages.

8.1.3.2 Apports méthodologiques

Toutes les études identifiées dans cette thèse et qui sollicitent la méthodologie de recherche en *design science* (Zuiderwijk., 2015 ; Alexopoulos et al., 2014 ; Yu, 2016 ; Pascal, 2006 ; Pascal et al, 2013) ont un point commun : elles s'inscrivent dans le cadre de grands projets de recherche financés par de grandes institutions, impliquant des collaborations entre plusieurs laboratoires de recherche et industriels. Ceci implique que les chercheurs ont à leur disposition des moyens techniques et financiers substantiels pour mener à bien leur recherche. Corollaire direct, ces méthodologies semblent peu accessibles aux équipes de recherche moins équipées.

Adressant la question de l'accessibilité des méthodologies de recherche *design science*, ma thèse est la première à proposer une méthodologie de recherche *design science* qui ne suppose aucune compétence technique, ni moyens financiers substantiels. Pour utiliser une image, ma méthodologie est à la recherche dans le domaine des Systèmes d'Information ce que le *bootstrapping* est au monde de l'entrepreneuriat, c'est à dire un "processus de création et de développement qui nécessite beaucoup d'efforts et aucun investissement de la part de parties prenantes extérieures". (Cambridge dictionary). Autrement dit, ma méthodologie suppose d'exploiter toutes les potentialités des outils et compétences du chercheur sans avoir besoin d'investissements matériels ou immatériels de la part d'autres personnes situées en dehors du projet de recherche.

Je rends cette optimisation des moyens alloués possible en proposant l'idée selon laquelle des cartographies d'architecture sont des résultats suffisants pour une recherche en *design science*. Dès lors, aucune connaissance ou assistance technique (programmation, développement informatique) n'est obligatoire pour mener à bien une recherche en *design science*.

Je contreviens alors à l'idée selon laquelle une méthodologie de recherche en *design science* doit nécessairement aboutir à un artefact fonctionnel et opérationnel ; par là, je rejoins l'école anglo-saxonne de la recherche en *design science* qui estime l'output de la recherche peut revêtir toutes les formes qui permettent d'accroître la base de connaissance du domaine : construit, modèle, méthode, artefact ou expérience.

Néanmoins, dans le cas où un développement technique est nécessaire, je propose de rendre cette économie de moyens possibles en utilisant des outils logiciens (Ince et Hekmatpour, 1987, p.11) qui permettent de concevoir des objets techniques sans disposer de compétences de programmation ou de compétences de programmation particulières.

In fine, mon apport méthodologique est de permettre aux chercheurs avec des moyens simples d'adresser des problèmes plus complexes, en l'occurrence des problèmes pernicieux³. Cet apport m'apparaît d'autant plus important que l'issue de cette méthodologie économe en moyen (cartographies d'architecture, prototypes créés avec des outils logiciens) peut ensuite attirer des financements et compétences plus importants pour développer des projets de recherche d'envergure.

8.1.3.3 Apports théoriques

En premier lieu, ma thèse vient contrebalancer la tendance des praticiens et chercheurs à se focaliser sur l'ouverture des données et peu sur leur usage (Goëta, 2016 ; Safarov et al., 2017). Il complète également les travaux des auteurs s'étant intéressés à l'usage, mais qui se sont restreint à l'étude des premières phases de l'usage des données (Davies, 2010 ; Zuiderwijk, 2015). En effet, les connaissances théoriques proposées dans ma thèse couvrent la totalité du processus de construction des usages, de l'utilisation des plateformes *open data* jusqu'à la création d'une organisation qui exploite la valeur ajoutée des nouveaux usages.

Autrement dit, ma thèse vient compléter les modèles théoriques établis jusqu'à présent sur la construction des usages.

Dans leur étude, Kuk et Davies (2011) proposent d'étudier au niveau intra-organisationnel le processus de construction des usages. Ils introduisent à cet égard les concepts de séquence d'articulation et de complémentarités. Je complète ces travaux sur deux aspects. Premièrement, je précise comment le passage d'une articulation à l'autre s'effectue et ce en introduisant les notions complémentaires de représentation des données, de conditions d'interaction avec les données et d'utilisation des données introduites par Mortier et al. (2014). J'affirme qu'une donnée passe d'une articulation à une autre à condition que la représentation des données et des conditions d'interaction des données s'accordent. Deuxièmement, je montre que ces séquences d'articulations sont différentes en fonction du système dans lequel elles se trouvent, et que c'est de la combinaison des séquences d'articulation de chaque système que dépend la construction des usages.

³ Termes originaux : *wicked problems*

Ici, je viens également compléter les travaux de Harrison et al. (2012) et Ruijer et al., (2017) qui se sont restreints à étudier les écosystèmes au niveau sociétal sans aborder les autres niveaux d'analyse. Je capitalise spécifiquement sur cette deuxième étude qui appelle les chercheurs à se concentrer sur la notion d'Open Data Work (Ruijer et al., 2017, p1) plutôt que sur les impacts. Ils invitent à laisser de côté les études des bienfaits des données ouvertes en termes de transparence ou de développement économique pour se concentrer sur ce que les individus "font avec les données ouvertes dans leur quotidien" (Ruijer et al., 2017, p.4). Les différentes strates de compréhension de la construction des usages introduits au chapitre 3 apporte à mon sens une contribution sur la compréhension de l'Open Data Work.

Dans une autre perspective, la *design theory* que je propose peut être considérée comme un complément de celle proposée par Zuiderwijk (2015). Celle-ci vise à informer la construction d'infrastructures *open data* qui permettent de faciliter la coordination de l'usage des données ouvertes tandis que la mienne informe la construction d'environnement pour faciliter la construction des usages. L'imbrication des deux est une piste intéressante de recherche.

Enfin, par l'utilisation d'une méthodologie de recherche en *design science*, ma thèse vient compléter toutes les études analytiques sur le processus d'ouverture des données (voir partie 1.3 de la revue de littérature). Plus précisément, elle exploite toutes les connaissances accumulées depuis le début du mouvement d'ouverture des données pour les intégrer dans un dispositif en Systèmes d'Information.

8.1.3.4 Validité interne et externe de la démarche

La fiabilité de la recherche est assurée par la validité interne et externe de la méthodologie mise en place, elles-mêmes garanties par l'architecture de recherche mise en place (Chapitre 2, Figure 2.3, p.162) et le respect des principes opérationnels de la méthodologie établis à la section 2.1.2 (p.162). **Du point de vue de la validité interne**, elle est garantie par le fait que le développement du produit a été réalisé en collaboration avec la coopérative Dataactivist, dont c'est le métier de développer de dispositifs pour le compte des opérateurs des données ouvertes. Leur participation aux différentes phases de construction du Canvas a permis de corriger au fur et à mesure mes erreurs ou incohérences. Aussi, la validité interne de la méthodologie est-elle garantie par l'évaluation que des chercheurs en *design science* ont pu en faire lors des consortium doctoraux comme celui de l'Association Information et Management ou de conférences scientifiques (Dominguez-Péry et al., 2018; Dominguez-Péry et al., 2019). **Du point de vue de la validité externe**, elle est assurée par le fait que la technologie informatique qui sous-tend l'Open Data Canvas a été testé dans le cadre des activités de prototype (voir chapitre 6) et s'est avérée opérationnelle. Aussi est-elle assurée par les activités d'évaluation (voir chapitre 7) qui ont permis d'établir, auprès d'un panel

d'acteurs hétérogènes, que l'usage de cette technologie est utile pour lever les barrières socio-techniques qui freinent la construction des usages.

8.2 Limites et ouverture de la recherche

8.2.1 Limites de la recherche

8.2.1.1 Du point de vue de la validité interne

La principale limite du point de vue de la validité interne provient du choix que j'ai fait de consolider en priorité la réalisation du premier cycle de la méthodologie *design science* pour aller jusqu'à la version alpha du prototype. De fait, je n'ai pu effectuer un deuxième cycle qui aurait permis de faire évoluer le design, de construire une nouvelle version du prototype, de l'évaluer, de donc d'obtenir un artefact qui valide les trois propositions de *design*.

Cette limite sera dépassée dans les suites de la thèse dans la mesure où je continuerai le travail de design mené avec la société coopérative. *In fine*, en me concentrant sur le premier cycle, j'ai échangé la complétude du Canvas contre de solides fondations sur lesquelles de nombreuses versions du Canvas pourront se développer.

8.2.1.2 Du point de vue de la validité externe

En premier lieu, je note que les caractéristiques des quatre cas exploratoires d'usage des données ouvertes ont bien évidemment influencé l'identification des barrières socio-techniques qui freinent la construction des usages. Il est important pour le lecteur de savoir que le type de données utilisées, le profil des usagers des données, les usages visés, les artefacts utilisés et leurs capacités d'action sont autant de facteurs qui m'ont influencé, dans l'identification des barrières. Cette limite est d'autant plus importante que l'identification des barrières a déterminé le reste de la thèse.

Ensuite, pour la phase d'évaluation du dispositif, les moyens dont je disposais m'ont amené à choisir un nombre limité le nombre de participants-évaluateurs (15 participants). Ce faible nombre, bien qu'il ait été contrebalancé en multipliant le nombre de *focus group* et donc le nombre de contextes à la lumière duquel

l'Open Data Canvas a été évalué, constitue une limite que les suites de cette thèse auront pour objectif de dépasser. Cette limite vient limiter l'applicabilité des résultats à des contextes extérieurs à la thèse.

Aussi, dans cette même phase d'évaluation, j'ai choisi de solliciter des panels d'acteurs majoritairement constitués des usagers provenant majoritairement du système de production de données, dans la mesure où ils sont les usagers de l'open data les plus expérimentés. Leur expérience fut un gage de crédibilité de leur évaluation. Cette uniformité des panels vient limiter la validité externe de l'Open Data Canvas. Les recherches futures viseront à évaluer le Canvas avec des *focus group* mêlant usagers expérimentés et usagers moins expérimentés, comme les nouveaux chefs de projet *open data* des collectivités n'ayant encore jamais fait usage ni ouvert leurs jeux de données, ainsi que des usagers provenant des trois systèmes (production, intermédiation, transformation)

8.2.2 Ouverture : recommandations pour les recherches futures sur les usages des données ouvertes

Du point de vue des perspectives théoriques, les recherches futures pourront se concentrer sur l'intégration de ma *design theory* avec la *design theory* développée par Zuiderwijk (2015). Cet intégration aurait le double avantage :

- de combiner les connaissances pratiques sur la construction d'infrastructures *open data* visant à coordonner l'usage des données et celles visant à construire des environnements de conception visant à construire les usages. Le résultat pourrait se traduire par la conception d'artefacts renforçant la consistance et le niveau d'organisation des démarches de construction des usages.
- de croiser l'utilisation qu'elle fait de la théorie de la coordination (Malone et Crowston, 1990) avec l'utilisation que j'ai fait des théories des paradoxes.

Aussi, il s'agirait d'investir la piste de la mise en continuité de mes recherches (et plus particulièrement de mon modèle théorique) par rapport au modèle de formation des usages liés aux technologies de l'information et de la communication numérique développée par Paquien-séguy (2007). La continuité est prometteuse dans la mesure où elle montrerait l'évolution entre la construction des usages d'objets matériels comme la télévision ou le téléphone et la construction des usages d'objets informationnels "pur" comme les données ouvertes.

Du point de vue des perspectives pratiques, les recherches futures pourront se concentrer sur le développement de la version bêta du prototype de l'Open Data Canvas, qui intégrerait l'ensemble des vingt-huit fonctions. Une feuille de route sera établie avec la société coopérative Dataactivist en ce sens. Aussi sera-t-il question d'étendre considérablement le périmètre des utilisateurs-évaluateurs du Canvas avec pour objectif d'atteindre l'ensemble des organisations ayant publié un premier jeu de données sans avoir encore développé aucun usage. Les PME, collectivités infra-métropolitaines (communes et autres) sont les cibles privilégiées pour la recherche future.

Bibliographie

- Abbate, T., De Luca, D., Gaeta, A., Lepore, M., Miranda, S., & Perano, M. (2015). Analysis of Open Innovation Intermediaries Platforms by Considering the Smart Service System Perspective. *PROMFG Procedia Manufacturing*, 3, 3575–3582.
- Abrams, B., Cwalina, K. (2008). *Framework design guidelines: Conventions, idioms, and patterns for reusable. Net libraries*. Pearson Education.
- Agarwal, R., Selen, W., Roos, G., & Green, R. (2015). *The handbook of service innovation*. Springer.
- Aguilera, U., Peña, O., Belmonte, O., & López-de-Ipiña, D. (2017). Citizen-centric data services for smarter cities. *Future Generation Computer Systems*, 76, 234–247.
<https://doi.org/10.1016/j.future.2016.10.031>
- Ahlers, D., Wienhofen, L. W., Petersen, S. A., & Anvaari, M. (2019). A Smart City Ecosystem Enabling Open Innovation. *International Conference on Innovations for Community Services*, 109–122.
- Ahonen, M., University of Tampere, & School of Information Sciences. (2011). *Designing an information system for open innovation: Bridging the gap between individual and organisational creativity : academic dissertation*. Acta Universitatis Tamperensis 1606 : Tampere University Press.
- Aken, J. E. van. (2004). Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules. *Journal of Management Studies*, 41(2), 219–246.
- Akrab, H. (2016). la réorganisation des activités info-communicationnelles et des modes de coopération dans la gestion des dispositifs numériques territoriaux. *Pyramides. Revue Du Centre d'études et de Recherches En Administration Publique*, 26/27, 191–214.
- Aler, R., Borrajo, D., Camacho, D., & Sierra-Alonso, A. (2002). A knowledge-based approach for business process reengineering, SHAMASH. *Knowledge-Based Systems*, 15(8), 473–483.
- Alexopoulos, C., Zuiderwijk, A., Charapabidis, Y., Loukis, E., & Janssen, M. (2014). Designing a second generation of open data platforms: Integrating open data and social media. *International Conference on Electronic Government*, 230–241.
- Ali, L., & Avdic, A. (2015). A knowledge management framework for sustainable rural development: The case of Gilgit-Baltistan, Pakistan. *Electronic Journal of Knowledge Management*, 13(2), 103.
- Allen, M., & Cervo, D. (2015). *Multi-domain master data management: Advanced MDM and Data Governance in Practice*. Morgan Kaufmann.
- Allen, P. M., & Varga, L. (2006). A co-evolutionary complex systems perspective on information systems. *Journal of Information Technology*, 21(4), 229–238.
- Allen, P., & Varga, L. (2006). A co-Evolutionary Complex Systems Perspective on Information Systems. *JIT*, 21, 229–238. <https://doi.org/10.1057/palgrave.jit.2000075>
- Allison-Cassin†, S., & Scott.‡, D. (2018). Wikidata: A platform for your library's linked open data. *The Code4Lib Journal*, 40. <https://journal.code4lib.org/articles/13424>
- Al-Ruithe, M., & Benkhelifa, E. (2017). Analysis and Classification of Barriers and Critical Success Factors for Implementing a Cloud Data Governance Strategy. *PROCS Procedia Computer Science*, 113, 223–232.
- Amabile, S., Peneranda, A., & Haller, C. (2018). Management des biens communs de la connaissance: Principes de conception et gouvernance de l'action collective. *Systèmes d'information & Management*, 23(1), 11–40.
- An, Y., Dalrymple, P. W., Rogers, M., Gerrity, P., Horkoff, J., & Yu, E. (2009). Collaborative social modeling for designing a patient wellness tracking system in a nurse-managed health care

- center. *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, 2.
- Anttiroiko, A.-V., Valkama, P., & Bailey, S. J. (2014). Smart cities in the new service economy: Building platforms for smart services. *AI & SOCIETY*, 29(3), 323–334. <https://doi.org/10.1007/s00146-013-0464-0>
- Ardestani, S. B., Håkansson, C. J., Laure, E., Livenson, I., Stranák, P., Dima, E., Blommesteijn, D., & Sanden, M. v d. (2015). B2SHARE: An Open eScience Data Sharing Platform. *2015 IEEE 11th International Conference on E-Science*, 448–453. <https://doi.org/10.1109/eScience.2015.44>
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399–418.
- Attour, A., & Rallet, A. (2014). Le rôle des territoires dans le développement des systèmes trans-sectoriels d'innovation locaux: Le cas des smart cities. *Innovations*, 1, 253–279.
- Autio, E., & Thomas, L. (2014). Innovation ecosystems. *The Oxford Handbook of Innovation Management*, 204–288.
- Auyang, S. Y. (1998). *Foundations of complex-system theories: In economics, evolutionary biology, and statistical physics*. Cambridge University Press.
- Avenier, M.-J. (2009). Par le paradigme des sciences de l'artificiel, déployer la pensée complexe dans l'interaction de pratiques et recherches. *Synergies Monde*, 6, 51–81.
- Badii, C., Bellini, P., Cenni, D., Difino, A., Nesi, P., & Paolucci, M. (2017). Analysis and assessment of a knowledge based smart city architecture providing service APIs. *Future Generation Computer Systems*, 75, 14–29. <https://doi.org/10.1016/j.future.2017.05.001>
- Badouard, R. (2014). La mise en technologie des projets politiques. Une approche «orientée design» de la participation en ligne. *Participations*, 1, 31–54.
- Baker, K. S., & Bowker, G. C. (2007). Information ecology: Open system environment for data, memories, and knowing. *Journal of Intelligent Information Systems*, 29(1), 127–144. <https://doi.org/10.1007/s10844-006-0035-7>
- Baldwin, C. Y., & Woodard, C. J. (2008). *The architecture of platforms: A unified view*. Harvard Business School. <http://www.hbs.edu/research/pdf/09-034.pdf>
- Balena, P., Bonifazi, A., & Mangialardi, G. (2013). Smart communities meet urban management: Harnessing the potential of open data and public/private partnerships through innovative e-governance applications. *International Conference on Computational Science and Its Applications*, 528–540.
- Barns, S. (2018). Smart cities and urban data platforms: Designing interfaces for smart governance. *City, Culture and Society*, 12, 5–12. <https://doi.org/10.1016/j.ccs.2017.09.006>
- Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity*, 19(1, 2), 129–152.
- Baskerville, R., Pries-Heje, J., & Venable, J. (2009). *Soft design science methodology*. 1–11.
- Basso, O., Legrain, T., & Institut de l'entreprise (France). (2004). *La dynamique entrepreneuriale dans les grands groupes*. Institut de l'entreprise.
- Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., & Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics*, 214(1), 481–518. <https://doi.org/10.1140/epjst/e2012-01703-3>
- Baudet, C., Termine, F., & Plumez, J. (2015). *Co-design multidisciplinaire: Un cas de ludification touristique selon une approche de Design Science*. 3ème journée de recherche AIM-Serious Games et Co-design: Faites vos jeux!
- Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1), 4–17.
- Beaud, M., Gravier, M., & de Toledo, A. (2013). *L'art de la thèse: Comment préparer et rédiger un mémoire de master, une thèse de doctorat ou tout autre travail universitaire à l'ère du Net*. La Découverte.

- Becker, H. S. (2010). *Writing for social scientists: How to start and finish your thesis, book, or article*. ReadHowYouWant. com.
- Behkamal, B., Kahani, M., Bagheri, E., & Jeremic, Z. (2014). A Metrics-Driven Approach for Quality Assessment of Linked Open Data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 64–79. <https://doi.org/10.4067/S0718-18762014000200006>
- Bellini, P., Nesi, P., Paolucci, M., & Zaza, I. (2018). *Smart City architecture for data ingestion and analytics: Processes and solutions*. 2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService).
- Benbasat, I., & Zmud, R. W. (1999). Empirical research in information systems: The practice or relevance. *MIS Quarterly*, 23(1), 3–16.
- Benbya, H., & McKelvey, B. (2006). Toward a complexity theory of information systems development. *Information Technology & People*, 19(1), 12–34.
- Bergadaa, M. M. (1990). The role of time in the action of the consumer. *Journal of Consumer Research*, 17(3), 289–302.
- Berners-Lee, T. (2018, March 12). *The web can be weaponised – and we can't count on big tech to stop it* | Tim Berners-Lee. The Guardian. <http://www.theguardian.com/commentisfree/2018/mar/12/tim-berners-lee-web-weapon-regulation-open-letter>
- Berners-Lee, T. J., & Cailliau, R. (1990). *WorldWideWeb: Proposal for a HyperText project*.
- Bertot, J. C., Butler, B. S., & Travis, D. M. (2014). Local big data: The role of libraries in building community data infrastructures. *Proceedings of the 15th annual international conference on digital government research*, 17–23.
- Bharosa, N. (2011). *Netcentric Information Orchestration: Assuring information and system quality in public safety networks*.
- Bhaskar, R. (1978). *A realist theory of science* Harvester Press. Hassocks, England.
- Boden, M. A. (1996). Creativity. In *Artificial intelligence* (pp. 267–291). Elsevier.
- Bogers, M., Zobel, A.-K., Afuah, A., Almirall, E., Brunswicker, S., Dahlander, L., Frederiksen, L., Gawer, A., Gruber, M., & Haefliger, S. (2017). The open innovation research landscape: Established perspectives and emerging themes across different levels of analysis. *Industry and Innovation*, 24(1), 8–40.
- Böhm, C., Naumann, F., Freitag, M., George, S., Höfler, N., Köppelmann, M., Lehmann, C., Mascher, A., & Schmidt, T. (2010). *Linking open government data: What journalists wish they had known*. 1–4.
- Boice, R. (1982). Increasing the writing productivity of 'blocked' academicians. *Behaviour Research and Therapy*, 20(3), 197–207.
- Bonvoisin, J., Buchert, T., Preidel, M., & Stark, R. G. (2018). How participative is open source hardware? Insights from online repository mining. *Design Science*, 4.
- Borgman, C. L., Scharnhorst, A., & Golshan, M. S. (1802). Digital data archives as knowledge infrastructures: Mediating data sharing and reuse. *Journal of the Association for Information Science and Technology*.
- Borzeix, A., & Cochoy, F. (2008). Travail et théories de l'activité: Vers des workspace studies?. Introduction. *Sociologie Du Travail*, 50(3), 273–286.
- Bothorel, E. (2020). *Pour une politique publique de la donnée*. Mission Bothorel, Paris, La Documentation Française.
- Boudières, V., & George-Marcelpoil, E. (2005). *Les remontées mécaniques: Quelle place dans la genèse et l'évolution des stations de montagne?* colloque SAPRR: Objectifs, méthodes, résultats et enjeux des observatoires socio-économique et environnementaux des infrastructures linéaires.
- Boudreau, K. J. (2007). Does Opening a Platform Stimulate Innovation? The Effect on Systemic and Modular Innovations. *SSRN Journal SSRN Electronic Journal*.

- Boullier, D. (2016). *Sociologie du numérique*. Armand Colin.
- Bourcier, D. (2013). Tele-communs versus tele-services publics: Vers des services publics collaboratifs en ligne. *Revue Française d'administration Publique*, 2, 271–284.
- Bowen, S., Durrant, A., Nissen, B., Bowers, J., & Wright, P. (2016). The value of designers' creative practice within complex collaborations. *Design Studies*, 46, 174–198.
- Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2009). Toward information infrastructure studies: Ways of knowing in a networked environment. In *International handbook of internet research* (pp. 97–117). Springer.
- Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012). OPEN—Enabling non-expert users to extract, integrate, and analyze open data. *Datenbank-Spektrum*, 12(2), 121–130.
- Briner, K., & Rome, J. (2018). The Need for and Value of Data Governance. In *Building Capacity in Institutional Research and Decision Support in Higher Education* (pp. 67–80). Springer.
- Briscoe, G., & De Wilde, P. (2006). *Digital ecosystems: Evolving service-orientated architectures*. 17-es.
- Brodie, M. L. (2010). Data integration at scale: From relational data integration to information ecosystems. *24th IEEE International Conference on Advanced Information Networking and Applications*, 2–3.
- Brown, V. A., Harris, J. A., & Russell, J. Y. (2010). *Tackling wicked problems through the transdisciplinary imagination*. Earthscan.
- Bruckner, R., List, B., & Scheifer, J. (2001). Developing requirements for data warehouse systems with use cases. *AMCIS 2001 Proceedings*, 66.
- Bullet, N. S. (1987). Essence and accidents of software engineering. *IEEE Computer*, 20(4), 10–19.
- Bunakov, V., & Jeffery, K. (2013). *Licence management for public sector information*. 277–287.
- Bunge, M. (2012). *Scientific research II: The search for truth*. Springer Science & Business Media.
- Burns, T., & Stalker, G. M. (1994). *The management of innovation*. Oxford University Press.
- Buscher, V., & Doody, L. (2013). Global innovators: International case studies on smart cities. *BIS Research Paper*, 135.
- Buzan, T. (n.d.). *La lecture rapide*.
- Byers, J. E., Cuddington, K., Jones, C. G., Talley, T. S., Hastings, A., Lambrinos, J. G., Crooks, J. A., & Wilson, W. G. (2006). Using ecosystem engineers to restore ecological systems. *Trends in Ecology & Evolution*, 21(9), 493–500.
- Cao, Q. H., Giyyarpuram, M., Farahbakhsh, R., & Crespi, N. (2017). Policy-based usage control for a trustworthy data sharing platform in smart cities. *Future Generation Computer Systems*. <https://doi.org/10.1016/j.future.2017.05.039>
- Carey, T., & Mason, R. (1983). Information system prototyping: Techniques, tools, and methodologies. *INFOR: Information Systems and Operational Research*, 21(3), 177–191.
- Carlsson, S. A. (2010). Design science research in information systems: A critical realist approach. In *Design Research in Information Systems* (pp. 209–233). Springer.
- Carlsson, S. A. (2006). Towards an information systems design research framework: A critical realist perspective. *Proceedings of the First International Conference on Design Science Research in Information Systems and Technology, Claremont, CA*, 192–212.
- Carrara, W., Chan, W., Fischer, S., & Steenbergen, E. van. (2015). Creating value through open data: Study on the impact of re-use of public data resources. *European Commission*.
- Carroll, J. M. (2000). *Making use: Scenario-based design of human-computer interactions*. MIT press.
- Carvalho, L., & Vale, M. (2018). From participation to start-up urbanisation? Re-situating open data in Lisbon. In *Inside Smart Cities* (pp. 211–226). Routledge.
- Castano, S., Ferrara, A., Genta, L., & Montanelli, S. (2016). Combining crowd consensus and user trustworthiness for managing collective tasks. *Future Generation Computer Systems*, 54, 378–388.

- Ceccagnoli, M., Forman, C., Huang, P., & Wu, D. J. (n.d.). *COCREATION OF VALUE IN A PLATFORM ECOSYSTEM: THE CASE OF ENTERPRISE SOFTWARE*. 29.
- Cecilio, J., Caldeira, F., & Wanzeller, C. (2018). CityMii—An integration and interoperable middleware to manage a Smart City. *Procedia Computer Science*, 130, 416–423. <https://doi.org/10.1016/j.procs.2018.04.062>
- Chan, C. M. (2013). From open data to open innovation strategies: Creating e-services using open government data. *46th Hawaii International Conference on System Sciences*, 1890–1899.
- Chau, P. Y. K., & Tam, K. Y. (2000). Organizational adoption of open systems: A ‘technology-push, need-pull’ perspective. *Information & Management*, 37(5), 229–239. [https://doi.org/10.1016/S0378-7206\(99\)00050-6](https://doi.org/10.1016/S0378-7206(99)00050-6)
- Checkland, P., & Poulter, J. (2010). *Systems Approaches to Managing Change: A Practical Guide*. Luettu, 12, 2016.
- Chen, P. P.-S. (1976). The entity-relationship model—Toward a unified view of data. *ACM Transactions on Database Systems (TODS)*, 1(1), 9–36.
- Chen, X., Dallmeier-Tiessen, S., Dasler, R., Feger, S., Fokianos, P., Gonzalez, J. B., Hirvonsalo, H., Kousidis, D., Lavasa, A., & Mele, S. (2018). Open is not enough. *Nature Physics*, 1.
- Chesbrough, H., & Bogers, M. (2014). Explicating open innovation: Clarifying an emerging paradigm for understanding innovation. *New Frontiers in Open Innovation*. Oxford: Oxford University Press, Forthcoming, 3–28.
- Churchman, C. W. (1967). Guest Editorial: wicked problems. *Management science*, 14(4), 141–142
- Cicirelli, F., Guerrieri, A., Spezzano, G., & Vinci, A. (2017). An edge-based platform for dynamic Smart City applications. *Future Generation Computer Systems*, 76, 106–118. <https://doi.org/10.1016/j.future.2017.05.034>
- CNIL. (n.d.). *La plateforme d'une ville—Les données personnelles au coeur de la fabrique des Smart city*.
- Cofer, R., & Harding, B. F. (2006). *Rapid System Prototyping with FPGAs: Accelerating the Design Process*. Elsevier.
- Cohendet, P., Grandadam, D., Simon, L., & Capdevila, I. (2014). Epistemic communities, localization and the dynamics of knowledge creation. *Journal of Economic Geography*, 14(5), 929–954.
- Cohendet, P., & Simon, L. (2015). Introduction to the special issue on creativity in innovation. *Technology Innovation Management Review*, 5(7).
- Colebatch, H., Hoppe, R., & Noordegraaf, M. (2010). Understanding Policy Work. In R. Hoppe, M. Noordegraaf, & H. K. Colebatch (Eds.), *Working for Policy* (pp. 11–26). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.4>
- Collin, S.-O., & Smith, E. (2003). Corporate governance and corporate entrepreneurship in public associations: The case of a public riding school. *Public and Private Sector Partnerships : Sustainable Success*, 201–216.
- Commission européenne. (n.d.). *Re-using open data a study on companies transforming open data into economic and social value*.
- Commission européenne—L'ouverture des données: Moteur de croissance. (n.d.). Retrieved November 20, 2017, from <http://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:52011DC0882&from=FR>
- Conein, B. (2004). Cognition distribuée, groupe social et technologie cognitive. *Réseaux*, 2, 53–79.
- Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*. <https://doi.org/10.1016/j.giq.2014.01.003>
- Conradie, P., & Choenni, S. (2012). Exploring process barriers to release public sector information in local government. *Proceedings of the 6th international conference on theory and practice of electronic governance*, 5–13.

- Cordasco, G., De Donato, R., Malandrino, D., Palmieri, G., Petta, A., Pirozzi, D., Santangelo, G., Scarano, V., Serra, L., & Spagnuolo, C. (2017). *Engaging citizens with a social platform for open data*. 242–249.
- Correa, A. S., Zander, P.-O., & da Silva, F. S. C. (2018). Investigating open data portals automatically: A methodology and some illustrations. *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, p.82.
- Cosgrave, E., Arbuthnot, K., & Tryfonas, T. (2013). Living Labs, Innovation Districts and Information Marketplaces: A Systems Approach for Smart Cities. *Procedia Computer Science Procedia Computer Science*, 16, 668–677.
- Courmont, A. (2016). *Politiques des données urbaines: Ce que l'open data fait au gouvernement urbain* [PhD Thesis]. <http://www.theses.fr/2016IEPP0042/document>
- Cowan, D., Alencar, P., McGarry, F., & Palmer, R. M. (2015). Adapting to Climate Change-An Open Data Platform for Cumulative Environmental Analysis and Management. *International Conference on Computational Science and Its Applications*, 3–15.
- Crosnier, H. L., & Vidal, P. (2017). Le rôle du numérique dans la redéfinition des communs urbains. *Netcom. Réseaux, Communication et Territoires*, 31-1/2, 09–32.
- Cuno, S., Bruns, L., Tcholtchev, N., Lämmel, P., & Schieferdecker, I. (2019). Data Governance and Sovereignty in Urban Data Spaces Based on Standardized ICT Reference Architectures. *Data*, 4(1), 16.
- Curts, R. J., & Campbell, D. E. (2001). Avoiding information overload through the understanding of OODA loops, a cognitive hierarchy and object-oriented analysis and design. *Annapolis, MD: C4ISR Cooperative Research Program (CCRP)*.
- Dameri, R. P. (2017). Using ICT in smart city. In *Smart City Implementation* (pp. 45–65). Springer.
- Danneels, L., Viaene, S., & Van den Bergh, J. (2017). Open data platforms: Discussing alternative knowledge epistemologies. *Government Information Quarterly*, 34(3), 365–378.
- Datta, A. (1998). Automating the discovery of as-is business process models: Probabilistic and algorithmic approaches. *Information Systems Research*, 9(3), 275–301.
- Davic, R. D. (2003). Linking keystone species and functional groups: A new operational definition of the keystone species concept. *Conservation Ecology*, 7(1).
- Davies, T. (2010). Open data, democracy and public sector reform. *A Look at Open Government Data Use from Data. Gov. Uk*.
- Davies, T. (2011). Open Data: Infrastructures and ecosystems. *Open Data Research*, 1–6.
- Davies, T., & Perini, F. (2016). Researching the emerging impacts of open data: Revisiting the ODDC conceptual framework. *The Journal of Community Informatics*, 12(2).
- Dawes, S. S., Pardo, T. A., & Cresswell, A. M. (2004). Designing electronic government information access programs: A holistic approach. *Government Information Quarterly*, 21(1), 3–23.
- Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*, 33(1), 15–27.
- de Klerk, P. (n.d.). *Linked Open Government Data*. 89.
- de Vries, A., Halffman, W., & Hoppe, R. (2010). Policy Workers Tinkering with Uncertainty: In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 91–110). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.10>
- Deakin, M. (2009). A Community-Based Approach to Sustainable Urban Regeneration. *Journal of Urban Technology*, 16(1), 91–112.
- Dedehayir, O., Mäkinen, S. J., & Ortt, J. R. (2018). Roles during innovation ecosystem genesis: A literature review. *Technological Forecasting and Social Change*, 136, 18–29.
- Définition Socle Commun des données locales*. (n.d.).
- Dekkers, M., Polman, F., te Velde, R., & de Vries, M. (2006). Measuring European Public Sector Information Resources. *Final Report of Study on Exploitation of Public Sector Information–Benchmarking of EU Framework Conditions*.

- Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: Disturbing the ‘one size fits all’ approach to data governance. *International Data Privacy Law*.
- Denis, J. (2011). Le travail de l’écrit en coulisses de la relation de service. *Activités*, 8(8–2).
- Denis, J. (2018). Le travail invisible des données: Éléments pour une sociologie des infrastructures scripturales. In *Le travail invisible des données: Éléments pour une sociologie des infrastructures scripturales*. Presses des Mines. <http://books.openedition.org/pressesmines/3934>
- Denis, J., & Goëta, S. (2013). La fabrique des données brutes. Le travail en coulisses de l’open data. *Penser l’écosystème Des Données. Les Enjeux Scientifiques et Politiques Des Données Numériques*.
- Denis, J., & Pontille, D. (2012). Travailleurs de l’écrit, matières de l’information. *Revue d’anthropologie Des Connaissances*, 6(1), 1–20.
- Denyer, D., Tranfield, D., & Van Aken, J. E. (2008). Developing design propositions through research synthesis. *Organization Studies*, 29(3), 393–413.
- Desouza, K. C., & Bhagwatwar, A. (2012). Citizen apps to solve complex urban problems. *Journal of Urban Technology*, 19(3), 107–136.
- Dessi, N., Milia, G., Pascariello, E., & Pes, B. (2016). COWB: A cloud-based framework supporting collaborative knowledge management within biomedical communities. *Future Generation Computer Systems*, 54, 399–408.
- D’Ignazio, C., & Bhargava, R. (2016). DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics*, 12(3).
- Dillenbourg, P. (1999). *Collaborative learning: Cognitive and computational approaches. Advances in learning and instruction series*. ERIC.
- Direction Interministérielle des systèmes d’information et de communication. (2012). *Cadre Commun d’Urbanisation des Système d’Information de l’Etat*.
- Sarazin A. (2018) *L’innovation de services issue de l’open data dans les villes intelligentes, Consortium doctoral*. AIM.
- Dominguez-Péry C., Bouabdallah K., Sarazin A. (2019) *Open Data Ecosystems : what models to co-create service innovation in smart cities ?*, 3rd RGCS Symposium. RGCS.
- Dong, H., Singh, G., Attri, A., & El Saddik, A. (2017). Open data-set of seven Canadian cities. *IEEE Access*, 5, 529–543.
- Doray, P. (2017). Construction sociale des technologies. In F. Bouchard & J. Prud’homme (Eds.), *Sciences, technologies et sociétés de A à Z* (pp. 57–61). Presses de l’Université de Montréal. <http://books.openedition.org/pum/4276>
- Dorr, D. A., Cohen, D. J., & Adler-Milstein, J. (2018). Data-driven diffusion of innovations: Successes and challenges in 3 large-scale innovative delivery models. *Health Affairs*, 37(2), 257–265.
- Dorst, K. (2011). The core of ‘design thinking’ and its application. *Design Studies*, 32(6), 521–532.
- Douay, N. (2016). *Planifier à l’heure du numérique*. Colin
- Douglas, J. D. (1976). *Investigative social research: Individual and team field research*. Sage.
- Drechsler, A., & Dörr, P. (2014). *What kinds of artifacts are we designing? An analysis of artifact types and artifact relevance in IS journal publications*. 329–336.
- Dumesny, R. (2016). Design et data: Retour sur une première année de doctorat. *Sciences Du Design*, 2, 128–135.
- Dumez, H. (2011). Faire une revue de littérature: Pourquoi et comment? *Le Libellio d’Aegis*, 7(2-Eté), 15–27.
- Dyer, J. H. (2000). *Collaborative advantage: Winning through extended enterprise supplier networks*. Oxford University Press.
- Dyer, J. H., & Hatch, N. W. (2006). Relation-specific capabilities and barriers to knowledge transfers: Creating advantage through network relationships. *Strategic Management Journal*, 27(8), 701.

- Dyer, J. H., & Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *The Academy of Management Review*, 23(4), 660–679. <https://doi.org/10.2307/259056>
- Dymytrova, V., Larroche, V., & Paquiénseguy, F. (2018). *Cadres d'usage des données par des développeurs, des data scientists et des data journalistes Livrable n° 3*.
- Eckelberg, R., Calixto, V. B., Pimentel, M. H., Del Fabro, M. D., Sunyé, M., Peres, L., Todt, E., Alves, T., Dragone, A., & Schneider, G. (2018). *Educational Open Government Data: From requirements to end users*. 463–470.
- Eco, U. (2018). *Comment écrire sa thèse*. Flammarion.
- Edwards, P. N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*. MIT Press.
- Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.
- Eisenmann, T., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard Business Review*, 84(10), 92.
- Eisenmann, T. R., Parker, G., & Van Alstyne, M. (2009). Opening platforms: How, when and why? *Platforms, Markets and Innovation*, 131–162.
- Eisenmann, T. R., Parker, G., Van Alstyne, M., & Harvard Business School. (2008). *Opening platforms: How, when and why?* Harvard Business School.
- Ellison, A. M. (2019). Foundation species, non-trophic interactions, and the value of being common. *Isience*, 13, 254.
- Finholt, T. A., & Olson, G. M. (1997). From laboratories to collaboratories: A new organizational form for scientific collaboration. *Psychological Science*, 8(1), 28–36.
- Fischer, G., & Giaccardi, E. (2006). Meta-design: A framework for the future of end-user development. In *End user development* (pp. 427–457). Springer.
- Fischer, G., & Herrmann, T. (2011). Socio-technical systems: A meta-design perspective. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 3(1), 1–33.
- Fischer, G., & Scharff, E. (2000). Meta-design: Design for designers. *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, 396–405.
- Flyverbom, M., & Murray, J. (2018). Datastructuring—Organizing and curating digital traces into action. *Big Data & Society*, 5(2), 2053951718799114.
- Fogli, D., & Provenza, L. P. (2012). A meta-design approach to the development of e-government services. *Journal of Visual Languages & Computing*, 23(2), 47–62.
- Foulonneau, M., Martin, S., & Turki, S. (2014). How open data are turned into services ?. *International Conference on Exploring services science*, 31–39.
- Fountain, J. E. (2004). *Building the virtual state: Information technology and institutional change*. Brookings Institution Press.
- Framework for Responsible Sharing of Genomic and Health-Related Data*. (n.d.). Retrieved May 27, 2020, from <https://www.ga4gh.org/genomic-data-toolkit/regulatory-ethics-toolkit/framework-for-responsible-sharing-of-genomic-and-health-related-data/>
- Franke T, Lukowicz P, & Blanke U. (2015). Smart crowds in smart cities: Real life, city scale deployments of a smartphone based participatory crowd management platform. *J. Internet Serv. Appl. Journal of Internet Services and Applications*, 6(1), 1–19.
- Fransman, M. (2010). *The new ICT ecosystem: Implications for policy and regulation*. Cambridge University Press.
- Fu, Z., & Lin, X. (2014). Building the Co-design and Making Platform to Support Participatory Research and Development for Smart City. *Cross-Cultural Design*, 609–620. https://doi.org/10.1007/978-3-319-07308-8_58

- Fung*, Y. Y. (2004). Collaborative online learning: Interaction patterns and limiting factors. *Open Learning: The Journal of Open, Distance and e-Learning*, 19(2), 135–149.
- Gagliardi, D., Schina, L., Sarcinella, M. L., Mangialardi, G., Niglia, F., & Corallo, A. (2017). Information and communication technologies and public participation: Interactive maps and value added for citizens. *Government Information Quarterly*, 34(1), 153–166.
- Ganapati, S., & Reddick, C. G. (2012). Open e-government in US state governments: Survey evidence from Chief Information Officers. *Government Information Quarterly*, 29(2), 115–122.
- Gardey, D. (2001). *La Dactylographe et l'expéditionnaire: Histoire des employés de bureau 1890-1930*. Belin.
- Garrett, L. E., Spreitzer, G. M., & Bacevice, P. A. (2017). Co-constructing a sense of community at work: The emergence of community in coworking spaces. *Organization Studies*, 38(6), 821–842.
- Garud, R., Jain, S., & Tuertscher, P. (2008). Incomplete by design and designing for incompleteness. *Organization Studies*, 29(3), 351–371.
- Gascó-Hernández, M., Martín, E. G., Reggi, L., Pyo, S., & Luna-Reyes, L. F. (2018). Promoting the use of open government data: Cases of training and engagement. *Government Information Quarterly*, 35(2), 233–242.
- Gavshon, D., & Gorur, E. (2019). Information overload: How technology can help convert raw data into rich information for transitional justice processes. *International Journal of Transitional Justice*, 13(1), 71–91.
- Gawer, A. (2011). *Platforms, markets and innovation*. Edward Elgar Publishing.
- Geuijen, K., & Hart, P. 't. (2010). Flying Blind in Brussels: In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 171–190). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.16>
- Giaccardi, E. (2003). *Principles of Metadesign: Processes and Levels of Co-Creation in the New Design Space*. Ph.D. Dissertation. CAiiA-STAR, School of Computing, Plymouth, UK
- Gibbons, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. Sage.
- Giffinger, R., & Pichler-Milanović, N. (2007). *Smart cities: Ranking of European medium-sized cities*. Centre of Regional Science, Vienna University of Technology.
- Gil-Garcia, J. R., Pardo, T. A., & Nam, T. (2015). What makes a city smart? Identifying core components and proposing an integrative and comprehensive conceptualization. *Information Polity*, 20(1), 61–87. <https://doi.org/10.3233/IP-150354>
- Giordano, Y., & Jolibert, A. (2012). Spécifier l'objet de la recherche. In *Méthodologie de la recherche. Réussir son mémoire ou sa thèse en sciences de gestion*. Pearson Education.
- Gitelman, L. (2014). *Paper knowledge: Toward a media history of documents*. Duke University Press.
- Goeta, S. (2016). *Instaurer des données, instaurer des publics: une enquête sociologique dans les coulisses de l'open data* (Doctoral dissertation, Paris, ENST)
- Gouarné, J.-M. (1998). Le projet décisionnel. *Enjeux, Modèles, Architectures Du Data Warehouse (in French)*. Eyrolles.
- Goy, A., Magro, D., Petrone, G., Picardi, C., & Segnan, M. (2016). Ontology-driven collaborative annotation in shared workspaces. *Future Generation Computer Systems*, 54, 435–449.
- Graves, A., & Hendler, J. (2013). *Visualization tools for open government data*. 136–145.
- Gray, E., O'Connell, P., Jensen, C., Weber, S., Seigneur, J.-M., & Yong, C. (2002). Towards a framework for assessing trust-based admission control in collaborative ad hoc applications. *Dept. of Computer Science, Trinity College Dublin, Technical Report*, 66.
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 337–355.
- Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems*, 8(5), 312.

- Grenier, C., & Pauget, B. (n.d.). Qu'est-ce que la recherche en management ? *Recherche en soins infirmiers*, 91, 12–23.
- Greuter, M., Gévart, P., & Leroy-Terquem, É. (2014). *Réussir son mémoire et son rapport de stage*. l'Étudiant.
- Grover, V., & Kohli, R. (2012). Cocreating IT Value: New Capabilities and Metrics for Multifirm Environments. *MANAGEMENT INFORMATION SYSTEMS QUARTERLY*, 36(1), 225–232.
- Grover, Varun, & Kohli, R. (2012). Cocreating IT value: New capabilities and metrics for multifirm environments. *Mis Quarterly*, 225–232.
- Gribnau, A. (2014). The structural adoption of open data in governmental organisations: Technology and organisation in practice. *Conference for E-Democracy and Open Government*, 283.
- Guenther, M. (2012). *Intersection: How enterprise design bridges the gap between business, technology, and people*. Newnes.
- Gupta, A., & Panagiotopoulos, P. (2019). *Operational Capabilities for Smart City Data Management*. 504–506.
- Gupta, A., Panagiotopoulos, P., & Bowen, F. (2020). An orchestration approach to smart city data ecosystems. *Technological Forecasting and Social Change*, 153, 119929.
- Gurstein, M. B. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*.
- Gurzick, D., & Lutters, W. G. (2009). Towards a design theory for online communities. *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, 1–20.
- Haefliger, S., Von Krogh, G., & Spaeth, S. (2008). Code reuse in open source software. *Management Science*, 54(1), 180–193.
- Hall, W., & Pesenti, J. (2017). Growing the artificial intelligence industry in the UK. *Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy. Part of the Industrial Strategy UK and the Commonwealth*.
- Hanseth, O., & Lyytinen, K. (2004). Theorizing about the design of Information Infrastructures: Design kernel theories and principles. *Sprouts: Working Papers on Information Environments, Systems and Organizations*, 4(4), 207–241.
- Hardin, G. (1998). Extensions of "the tragedy of the commons". *Science*, 280(5364), 682–683.
- Hargrave, T. J., & Van de Ven, A. H. (2017). Integrating dialectical and paradox perspectives on managing contradictions in organizations. *Organization Studies*, 38(3–4), 319–339.
- Harrison, C., Eckman, B., Hamilton, R., Hartswick, P., Kalaganam, J., Paraszczak, J., & Williams, P. (2010). Foundations for smarter cities. *IBM Journal of Research and Development*, 54(4), 1–16.
- Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating open government ecosystems: A research and development agenda. *Future Internet*, 4(4), 900–928.
- Hassan, M. I. A., & Twinomurizi, H. (2018). A systematic literature review of open government data research: Challenges, opportunities and gaps. *2018 Open Innovations Conference (OI)*, 299–304.
- Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a global data space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1(1), 1–136.
- Heimstädt, Maximilian. (2014). The British Open Data Ecosystem. (M.Sc. Management & Information Technology), University of St Andrews.
- Heimstädt, M., Saunderson, F., & Heath, T. (2014). Conceptualizing Open Data ecosystems: A timeline analysis of Open Data development in the UK. *Conference for E-Democracy and Open Government*, 245.
- Hein, A., Böhm, M., & Kremer, H. (2018). Platform configurations within information systems research: A literature review on the example of IoT platforms. *Multikonferenz Wirtschaftsinformatik, Lüneburg, Germany*, 465–476.

- Heise, A., & Naumann, F. (2012). Integrating open government data with stratosphere for more transparency. *Journal of Web Semantics*, 14, 45–56.
- Henriksen, H. Z. (2015). Scrutinizing open government data to understand patterns in egovernment uptake. *International Conference on Electronic Government*, 144–155.
- Herrmann, T. (2010). Support of collaborative creativity for Co-located Meetings. In *From CSCW to Web 2.0: European Developments in Collaborative Design* (pp. 65–95). Springer.
- Hess, C. (2012). The unfolding of the knowledge commons. *St Antony's International Review*, 8(1), 13–24.
- Hess, C., & Ostrom, E. (2007). *Understanding knowledge as a commons*. The mit press.
- Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: Theory and practice* (Vol. 22). Springer Science & Business Media.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, 28(1), 6.
- Höchtl, J., & Reichstädter, P. (2011). *Linked open data-a means for public sector information management*. 330–343.
- Hodgkinson, G. P., & Healey, M. P. (2008). Toward a (pragmatic) science of strategic intervention: Design propositions for scenario planning. *Organization Studies*, 29(3), 435–457.
- Hogan, M., Ojo, A., Harney, O., Ruijter, E., Meijer, A., Andriessen, J., Pardijs, M., Boscolo, P., Palmisano, E., & Satta, M. (2017). Governance, transparency and the collaborative design of open data collaboration platforms: Understanding barriers, options, and needs. In *Government 3.0-Next Generation Government Technology Infrastructure and Services* (pp. 299–332). Springer.
- Holling, C. S., & Meffe, G. K. (1996). Command and control and the pathology of natural resource management. *Conservation Biology*, 10(2), 328–337.
- Huijboom, N., & Van den Broek, T. (2011). Open data: An international comparison of strategies. *European Journal of EPractice*, 12(1), 4–16.
- Hutchel, A., & Molet, H. (1986). Rational modelling in understanding and aiding human decision-making: About two case studies. *European Journal of Operational Research*, 24(1), 178–186.
- Hutchins, E. (1995). *Cognition in the Wild*. MIT press.
- Huysman, M. (2004). Design requirements for knowledge-sharing tools: A need for social capital analysis. *Social Capital and Information Technology*, 187–207.
- Iansiti, M., & Levien, R. (2004). Strategy as ecology. *Harvard Business Review*, 82(3), 68–78.
- Iivari, J., & Venable, J. R. (2009). *Action research and design science research – Seemingly similar but decisively dissimilar*, 17th European Conference on Information Systems, Verona
- Immonen, A., Palviainen, M., & Ovaska, E. (2014). Requirements of an open data based business ecosystem. *IEEE Access*, 2, 88–103.
- Ince, D., & Hekmatpour, S. (1987). Software prototyping—Progress and prospects. *Information and Software Technology*, 29(1), 8–14.
- Intarakumnerd, P., & Chaoroenporn, P. (2013). The roles of intermediaries in sectoral innovation system in developing countries: Public organizations versus private organizations. *Asian Journal of Technology Innovation*, 21(1), 108–119.
- Ireland, R. D., Covin, J. G., & Kuratko, D. F. (2009). Conceptualizing Corporate Entrepreneurship Strategy. *Entrepreneurship Theory and Practice*, 33(1), 19–46.
- Ishida, T., Isbister, K., & Kyoto Meeting on Digital Cities (Eds.). (2000). *Digital cities: Technologies, experiences, and future perspectives*. Springer.
- Jaakola, A., Kekkonen, H., Lahti, T., & Manninen, A. (2015). Open data, open cities: Experiences from the Helsinki metropolitan area. Case Helsinki Region Infoshare www. Hri. Fi. *Statistical Journal of the IAOS*, 31(1), 117–122.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446–456.

- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268.
- Janssen, M., Chun, S. A., & Gil-Garcia, J. R. (2009). Building the next generation of digital government infrastructures. *Government Information Quarterly*, 26(2), 233–237.
- Janssen, M., Matheus, R., & Zuiderwijk, A. (2015). Big and open linked data (BOLD) to create smart cities and citizens: Insights from smart energy and mobility cases. *International Conference on Electronic Government*, 79–90.
- Janssen, M., & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review*, 32(5), 694–711.
- Jarke, M., Bui, X. T., & Carroll, J. M. (1998). Scenario management: An interdisciplinary approach. *Requirements Engineering*, 3(3–4), 155–173.
- Jarzabkowski, P., Lê, J. K., & Van de Ven, A. H. (2013). Responding to competing strategic demands: How organizing, belonging, and performing paradoxes coevolve. *Strategic Organization*, 11(3), 245–280.
- Jeffery, K. G. (2000). Metadata: The future of information systems. *Information Systems Engineering: State of the Art and Research Themes*. London: Springer Verlag.
- Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2014). Data-driven innovation through open government data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 100–120.
- Jin, D., Hannon, C., Li, Z., Cortes, P., Ramaraju, S., Burgess, P., Buch, N., & Shahidehpour, M. (2016). Smart street lighting system: A platform for innovative smart city applications and a new frontier for cyber-security. *The Electricity Journal*, 29(10), 28–35.
<https://doi.org/10.1016/j.tej.2016.11.011>
- Johnson, D. W., & Johnson, R. T. (1996). Cooperation and the use of technology. *Handbook of Research for Educational Communications and Technology: A Project of the Association for Educational Communications and Technology*, 1017–1044.
- Jones, M. R., & Karsten, H. (2008). Giddens's structuration theory and information systems research. *MIS Quarterly*, 32(1), 127–157.
- Jouison-Laffitte, E. (2013). La recherche action: Oubliée de la recherche dans le domaine de l'entrepreneuriat. *Revue de l'Entrepreneuriat*, 8(1), 1–35.
- Juell-Skielse, G., Hjalmarsson, A., Johannesson, P., & Rudmark, D. (2014). Is the public motivated to engage in open data innovation? *International Conference on Electronic Government*, 277–288.
- Smith WK., Erez M, Jarvenpaa S., et al. (2017). Adding complexity to theories of paradox, tensions, and dualities of innovation and change: Introduction to organization studies special issue on paradox, tensions, and dualities of innovation and change. *Organization Studies* 38(3–4): 303–317.
- Kalampokis, E., Tambouris, E., & Tarabanis, K. (2013). Linked open government data analytics. *International Conference on Electronic Government*, 99–110.
- Kale, P., Singh, H., & Perlmutter, H. (2000). Learning and protection of proprietary assets in strategic alliances: Building relational capital. *SMJ Strategic Management Journal*, 21(3), 217–237.
- Kampars, J., Zdravkovic, J., Stirna, J., & Grabis, J. (2019). Extending organizational capabilities with Open Data to support sustainable and dynamic business ecosystems. *Software and Systems Modeling*, 1–28.
- Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly*, 30(4), 508–513.
- Katz, P., Lunze, T., Feldmann, M., Röhrborn, D., & Schill, A. (2011). *System architecture for handling the information overload in enterprise information aggregation systems*. 148–159.
- Kets de Vries, M. F. R. *The Dark Side of entrepreneurship. America in the postindustrial age: readings from business journals / Edited by Giovanni Baldi.*

- Khan, S., & Foti, J. (2015). Aligning supply and demand for better governance: Open data in the open government partnership. *Open Government Partnership Independent Reporting Mechanism*.
- Khayyat, M., & Bannister, F. (2017). Towards a model for facilitating and enabling co-creation using open government data. *Information Polity*, 22(4), 211–231.
- King, R. D., Liakata, M., Lu, C., Oliver, S. G., & Soldatova, L. N. (2011). On the formalization and reuse of scientific research. *Journal of The Royal Society Interface*, 8(63), 1440–1448.
- Klievink, B., Bharosa, N., & Tan, Y. H. (2016). The collaborative realization of public values and business goals: Governance and infrastructure of public-private information platforms. *Government Information Quarterly*, 33(1), 67–79.
- Koch, M. (2002). Interoperable community platforms and identity management in the university domain. *International Journal on Media Management*, 4(1), 21–30.
- Koontz, T. M., & Bodine, J. (2008). Implementing ecosystem management in public agencies: Lessons from the US Bureau of Land Management and the Forest Service. *Conservation Biology*, 22(1), 60–69.
- Korzybski, A. (1998). *Une carte n'est pas le territoire: Prolégomènes aux systèmes non-aristotéliens et à la sémantique générale*. Éditions de l'Éclat.
- Krob, D. (2009). Éléments d'architecture des systèmes complexes. *Gestion de La Complexité et de l'information Dans Les Grands Systèmes Critiques*, 179–207.
- Kučera, J. (2015). Open government data publication methodology. *Journal of Systems Integration*, 6(2), 52–61.
- Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: Anatomy of a research project. *European Journal of Information Systems*, 17(5), 489–504.
- Kuhn, K. (2011). Open government data and public transportation. *Journal of Public Transportation*, 14(1), 5.
- Kuk, G., & Davies, T. (2011). The roles of agency and artifacts in assembling open data complementarities. *Thirty Second International Conference on Information Systems*.
- Lakomaa, E., & Kallberg, J. (2013). Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. *IEEE Access*, 1, 558–563.
- Lane, T. (1994). La circularité du questionnement et des questions dans l'entrevue circulaire. *Service Social*, 43(3), 59–77.
- Lanfranchi, V., & Ireson, N. (2009). User requirements for a collective intelligence emergency response system. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, 198–203.
- Laousse, D., & Hooge, S. (2018). Refaire société par la création de communautés d'innovation-Le cas des ateliers SpotLAB sur les nouvelles mobilités en régions. *Revue Française de Gestion*, 44(272), 85–102.
- Larroche, v., Peyrelong, m.-f., & beaune, p. (2018). L'ouverture des données publiques: Un bien commun en devenir? *Tic&société*, 12(1), 255–281.
- Larroche, V., & VILA, M. (2015). Urban Data et stratégies dans le secteur des services: Le cas de la métropole lyonnaise. *Évelyne Broudoux, Ghislaine Chartron, Big Data-Open Data, Quelles Valeurs*, 183–196.
- Lazarotti, V., Manzini, R., Nosella, A., & Pellegrini, L. (2016). Innovation ambidexterity of open firms. The role of internal relational social capital. *Technology Analysis & Strategic Management Technology Analysis & Strategic Management*, 29(1), 105–118.
- Le Clust'R. (2015). *Le projet d'ouverture des données publiques et la mise en œuvre du portail « Open data » grenoblois*. <https://www.youtube.com/watch?v=iqepxjVbxYM>
- Le Corf, J.-B. (2016). Les pratiques d'innovation de services des développeurs web dans les territoires: Le cas des projets Open Data. *Communication & Organisation*, 2, 123–136.
- Le Doze, F., & Krumb, C. (2015). *La force de la confiance: Une thérapie pour s'unifier*. O. Jacob.

- Leal, M. (2016). *L'innovation de services grâce aux pratiques d'open data: Opportunités et risques*. Université Grenoble-Alpes.(Mémoire de Master 2 MSIO, Grenoble IAE School of Management)
- Lee, D. (2014). Building an open data ecosystem: An Irish experience. *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance*, 351–360.
- Lee, G., & Kwak, Y. H. (2012). An open government maturity model for social media-based public engagement. *Government Information Quarterly*, 29(4), 492–503.
- Lee, S. U., Zhu, L., & Jeffery, R. (2018). Designing Data Governance in Platform Ecosystems. *Proceedings of the 51st Hawaii International Conference on System Sciences*.
- Lee, T., Lee, B.-K., & Lee-Geiller, S. (2020). The effects of information literacy on trust in government websites: Evidence from an online experiment. *International Journal of Information Management*, 52, 102098. <https://doi.org/10.1016/j.ijinfomgt.2020.102098>
- Leonardi, P. M. (2010). Digital materiality? How artifacts without matter, matter. *First Monday*, 15(6).
- Lessig, Lawrence (2000), *Code and Other Laws of Cyberspace*. New York: Basic Books.
- Léveillé, V., & Timms, K. (2015). Through a Records Management Lens: Creating a Framework for Trust in Open Government and Open Government Information/Les objectifs visés par les systèmes de gestion documentaires: La mise en place d'un cadre de confiance et de la transparence de l'information dans un gouvernement ouvert. *Canadian Journal of Information and Library Science*, 39(2), 154–190.
- Li, M. (2015). The Study of English Learning Community in Network Based on SNS. *2015 2nd International Conference on Electrical, Computer Engineering and Electronics*.
- Li, N., Guo, X., Chen, G., & Luo, N. (2015). Reading behavior on intra-organizational blogging systems: A group-level analysis through the lens of social capital theory. *INFMAN Information & Management*, 52(7), 870–881.
- Licoppe, C. (2008). Dans le «carré de l'activité»: Perspectives internationales sur le travail et l'activité. *Sociologie Du Travail*, 50(3), 287–302.
- Lin, C. J., Chen, F. F., & Chen, Y. M. (2013). Knowledge kanban system for virtual research and development. *Robotics and Computer-Integrated Manufacturing*, 29(3), 119–134.
- Linder, L., Vionnet, D., Bacher, J.-P., & Hennebert, J. (2017). Big Building Data—A Big Data Platform for Smart Buildings. *Energy Procedia*, 122, 589–594. <https://doi.org/10.1016/j.egypro.2017.07.354>
- Linders, D. (2013). Towards open development: Leveraging open data to improve the planning and coordination of international aid. *Government Information Quarterly*, 30(4), 426–434.
- Lindman, J., Kinnari, T., & Rossi, M. (2014). Industrial open data: Case studies of early open data entrepreneurs. *2014 47th Hawaii International Conference on System Sciences*, 739–748.
- Lipton, V. (2020). The Staged Model for Open Scientific Data. In *Open Scientific Data-Why Choosing and Reusing the RIGHT DATA Matters*. IntechOpen.
- Liyang, W. G. Z. (2016). A comparative study on the construction of open data platform between Chinese and foreign governments. *Information and Documentation Services*, 37(6), 75–79.
- Lnenicka, M., & Komarkova, J. (2019). Big and open linked data analytics ecosystem: Theoretical background and essential elements. *Government Information Quarterly*, 36(1), 129–144.
- Loeber, A. (2010). Evaluation as Policy Work: In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 131–152). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.13>
- Lofi, C., & Krestel, R. (2012). iParticipate: Automatic tweet generation from local government data. *International Conference on Database Systems for Advanced Applications*, 295–298.
- Loveluck, B. (2015). *Réseaux, libertés et contrôle: Une généalogie politique d'internet*. Armand Colin.

- Luft, T., & Wartzack, S. (2012). Requirement analysis for contextual management and supply of process-and design knowledge—a case study. *DS 70: Proceedings of DESIGN 2012, the 12th International Design Conference, Dubrovnik, Croatia*.
- Lusch, R. F., & Nambisan, S. (2015). Service Innovation: A Service-Dominant Logic Perspective. *MISQ MIS Quarterly*, 39(1), 155–175.
- Lüscher, L. S., & Lewis, M. W. (2008). Organizational change and managerial sensemaking: Working through paradox. *Academy of Management Journal*, 51(2), 221–240.
- Lv, Z., Li, X., Wang, W., Zhang, B., Hu, J., & Feng, S. (2018). Government affairs service platform for smart city. *Future Generation Computer Systems*, 81, 443–451. <https://doi.org/10.1016/j.future.2017.08.047>
- Mabi, C. (2015). La plate-forme «data. Gouv. Fr» ou l’open data à la française. *Informations Sociales*, 5, 52–59.
- Maciulienė, M. (2018). Mapping Digital Co-Creation for Urban Communities and Public Places. *SYSTEMS*, 6(2).
- Magalhaes, G., Roseira, C., & Manley, L. (2014). *Business models for open government data*. 365–370.
- Mainka, A., Hartmann, S., Meschede, C., & Stock, W. G. (2015). Mobile application services based upon open urban government data. *ICConference 2015 Proceedings*.
- Malone, T. W., & Crowston, K. (1990). *What is coordination theory and how can it help design cooperative work systems?* 357–370.
- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30–37.
- Maramieri, J. (2014). Open Government Data: A citizen’s right or a concession of public authorities? *Journal of E-Learning and Knowledge Society*, 10(2).
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251–266.
- Markus, M. L., Majchrzak, A., & Gasser, L. (2002). A design theory for systems that support emergent knowledge processes. *MIS Quarterly*, 179–212.
- Martin, C. (2014). Barriers to the open government data agenda: Taking a multi-level perspective. *Policy & Internet*, 6(3), 217–240.
- Martin, L., & Ute, S. (2017, January). Measuring employee innovation: a review of existing scales and the development of the innovative behavior and innovation support inventories across cultures. *International Journal of Entrepreneurial Behaviour & Research*.
- Martin, S., Foulonneau, M., & Turki, S. (2013). 1-5 stars: Metadata on the openness level of open data sets in Europe. *Research Conference on Metadata and Semantic Research*, 234–245.
- Martin, S., Turki, S., & Renault, S. (2017). Open data ecosystems. *International Conference on Electronic Government and the Information Systems Perspective*, 49–63.
- McDermott, P. (2010). Building open government. *Government Information Quarterly*, 27(4), 401–413.
- McKay, J., & Marshall, P. (2005). *A review of design science in information systems*. EJ.
- McLeod, M., & McNaughton, M. (2016). Mapping an emergent Open Data eco-system. *The Journal of Community Informatics*, 12(2).
- McNaughton, M., McLeod, M. T., & Boxill, I. (2016). An Actor Network Perspective of Tourism Open Data. In *Tourism and Hospitality Management* (pp. 47–60). Emerald Group Publishing Limited.
- McQueary, B., DeFrancesco, A., Lehman, L., & Dziegiel Jr, R. (2016). Adaptive Workspace to Reduce Analyst Information Overload. *Proceedings on the International Conference on Artificial Intelligence (ICAI)*, 437.
- Meier, O. (2012). *Management du changement changement culturel et organisationnel*. Dunod.

- Meijer, A., & Thaens, M. (2009). Public information strategies: Making government information available to citizens. *Information Polity*, 14(1, 2), 31–45.
- Merkel, J. (2015). Coworking in the city. *Ephemera*, 15(2), 121–139.
- Metze, T. (2010). New Life for Old Buildings: In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 75–90). Amsterdam University Press; JSTOR.
<http://www.jstor.org/stable/j.ctt46n2ww.9>
- Mezghani, E., Exposito, E., & Drira, K. (2016). A collaborative methodology for tacit knowledge management: Application to scientific research. *Future Generation Computer Systems*, 54, 450–455.
- Michelle Bergadaà, & Simon Nyeck. (1992). Recherche en marketing: Un état des controverses. *Recherche et Applications en Marketing*, 7(3), 23–44.
- Miles, M. B., Huberman, A. M., Huberman, M. A., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
- Millette, C., & Hosein, P. (2016). A consumer focused open data platform. *2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)*, 1–6.
- Mina, A., Bascavusoglu-Moreau, E., & Hughes, A. (2014). Open service innovation and the firm's search for external knowledge. *Research Policy*, 43(5), 853–866.
- Misuraca, G., & Viscusi, G. (2014). Is open data enough?: E-governance challenges for open government. *International Journal of Electronic Government Research (IJEGR)*, 10(1), 18–34.
- Mom, T. J. M., van Neerijnen, P., Reinmoeller, P., & Verwaal, E. (2015). Relational Capital and Individual Exploration: Unravelling the Influence of Goal Alignment and Knowledge Acquisition. *Organization Studies (01708406)*, 36(6), 809–829. buh.
- Mom T.J.M, van Neerijnen P, Reinmoeller P, & Verwaal E. (2015). Relational Capital and Individual Exploration: Unravelling the Influence of Goal Alignment and Knowledge Acquisition. *Organ. Stud. Organization Studies*, 36(6), 809–829.
- Monino, J.-L., & Sedkaoui, S. (2016). Data Development Mechanisms. In *Big Data, Open Data and Data Development* (pp. 43–72). John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781119285199.ch3>
- Mortier, R., Haddadi, H., Henderson, T., McAuley, D., & Crowcroft, J. (2014). Human-data interaction: The human face of the data-driven society. *Available at SSRN 2508051*.
- Nahon, K., Peled, A., & Shkabatur, J. (2015). OGD heartbeat: Cities' commitment to open data. *JeDEM*, 7(2), 116–136.
- Nam, T. (2015). Challenges and concerns of open government: A case of government 3.0 in Korea. *Social Science Computer Review*, 33(5), 556–570.
- Nambisan S, & Nambisan P. (2008). How to Profit from a Better Virtual Customer Environment'. *MIT Sloan Manage. Rev. MIT Sloan Management Review*, 49(3), 53–61.
- Nardi, B. A., & O'Day, V. (1999). *Information ecologies: Using technology with heart*. Mit Press.
- National Science Foundation, C. C. (2007). Cyberinfrastructure vision for 21st century discovery. *Arlington, VA: National Science Foundation*). <Http://Www.Nsf.Gov/Pubs/2007/Nsf0728/Nsf0728.Pdf>.
- Neumaier, S., Thurnay, L., Lampoltshammer, T. J., & Knap, T. (2018). Search, filter, fork, and link open data: The adequate platform: Data-and community-driven quality improvements. *Companion Proceedings of the The Web Conference 2018*, 1523–1526.
- Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104). Prentice-Hall Englewood Cliffs, NJ.
- Ng S.T, Xu F.J, Yang Y, Lu M, & Creative Construction Conference, C. 2017. (2017). A Master Data Management Solution to Unlock the Value of Big Infrastructure Data for Smart, Sustainable and Resilient City Planning. *Procedia Eng. Procedia Engineering*, 196, 939–947.
- Ngomo, A.-C. N., Auer, S., Lehmann, J., & Zaveri, A. (2014). *Introduction to linked data and its lifecycle on the web*. 1–99.

- Noordegraaf, M. (2010). Academic Accounts of Policy Experience. In M. Noordegraaf, H. K. Colebatch, & R. Hoppe (Eds.), *Working for Policy* (pp. 45–68). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.7>
- ODI report: Mapping data ecosystems | Map | Ecosystem. (n.d.). Scribd. Retrieved April 30, 2018, from <https://www.scribd.com/document/374682869/ODI-report-Mapping-data-ecosystems>
- OECD Recommendation on Health Data Governance—OECD. (n.d.). Retrieved May 27, 2020, from <https://www.oecd.org/els/health-systems/health-data-governance.htm>
- Offermann, P., Blom, S., Schönherr, M., & Bub, U. (2010, June). Artifact types in information systems design science—a literature review. In *International Conference on Design Science Research in Information Systems* (pp. 77-92). Springer, Berlin, Heidelberg.
- O'Hara, K. (2012, June). Transparency, open data and trust in government: shaping the infosphere. In *Proceedings of the 4th annual ACM web science conference* (pp. 223-232).
- Ohemeng, F. L., & Ofosu-Adarkwa, K. (2015). One way traffic: The open data initiative project and the need for an effective demand side initiative in Ghana. *Government Information Quarterly*, 32(4), 419–428.
- Ojo, A., Curry, E., & Zeleti, F. A. (2015, January). A tale of open data innovations in five smart cities. In *2015 48th Hawaii International Conference on System Sciences* (pp. 2326-2335). IEEE.
- Ojo, A., Stasiewicz, A., Porwol, L., Petta, A., Pirozzi, D., Serra, L., ... & Vicidomini, L. (2018, May). A comprehensive architecture to support Open Data access, co-creation, and Dissemination. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (pp. 1-2).
- O'Reilly, T. (2011). Government as a Platform. *Innovations: Technology, Governance, Globalization*, 6(1), 13–40.
- Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science*, 3(3), 398–427.
- Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science*, 11(4), 404–428.
- Osagie, E., Waqar, M., Adebayo, S., Stasiewicz, A., Porwol, L., & Ojo, A. (2017). *Usability evaluation of an open data platform*. 495–504.
- Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, 14(3), 137–158.
- Ostrom, E. (2015). *Governing the commons*. Cambridge university press.
- Oxman, R. (2017). Thinking difference: Theories and models of parametric design thinking. *Design Studies*, 52, 4–39.
- Paprica, P. A., Sutherland, E., Smith, A., Brudno, M., Cartagena, R. G., Crichlow, M., Courtney, B. K., Loken, C., McGrail, K. M., Ryan, A., Schull, M. J., Thorogood, A., Virtanen, C., & Yang, K. (2020). Essential requirements for establishing and operating data trusts: Practical guidance based on a working meeting of fifteen Canadian organizations and initiatives. *ArXiv:2005.06604 [Cs]*. <http://arxiv.org/abs/2005.06604>
- Paquiénéguy, F. (2007). Comment réfléchir à la formation des usages liés aux technologies de l'information et de la communication numériques? *Les Enjeux de l'information et de La Communication*, 2007(1), 63–75.
- Paquiénéguy, F. (2018). L'usage au fil des Tic: Une genèse à raviver pour mieux le repenser? *Interfaces Numériques*, 6(3), 481–481.
- Paquiénéguy, F. (2013, September). La Notion de dispositif appliquée aux TIC et aux médias numériques. In " *Les Actes du colloque d'Agadir 2012 sont publiés dans le cadre de l'action intégrée Phc Volubilis 2012 Projet N26866ZD* (p. 378).
- Paquiénéguy, F., Larroche, V., Peyrelong, M.-F., Vila-Raimondi, M., & Dymytrava, V. (2016). *Document 1L - Synthèse des résultats de l'enquête auprès des réutilisateurs de données ouvertes*

- Paraponaris, C. (2017). *Plateformes numériques, conception ouverte et emploi* (p. 148). Editions de l'IREES.
- Parker, G. G., & Van Alstyne, M. W. (2005). Two-Sided Network Effects: A Theory of Information Product Design. *Management Science*, 51(10), 1494–1504. <https://doi.org/10.1287/mnsc.1050.0400>
- Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You*. WW Norton & Company.
- Parker, G., Van Alstyne, M. W., & Jiang, X. (2016). Platform ecosystems: How developers invert the firm. *Boston University Questrom School of Business Research Paper*, (2861574).
- Parycek, P., Hocht, J., & Ginner, M. (2014). Open government data implementation evaluation. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 80–99.
- Pascal, A. (2006). *Conception d'une solution TIC pour favoriser l'émergence de projets innovants: une approche usage-L'expérience KMP* (Doctoral dissertation, Université Nice Sophia Antipolis).
- Pascal, A. (2012). Le design science dans le domaine des systèmes d'information: Mise en débat et perspectives. *Systèmes d'information & Management*, 17(3), 7–31.
- Pascal, A., & Rouby, É. (2017). Les scénarios d'usage comme support aux méthodologies de Recherche en Design Science dans le cas d'invention. *Systèmes d'information & Management*, 22(4), 47–78.
- Pascal, A., Thomas, C., & Romme, A. G. L. (2013). Developing a human-centred and science-based approach to design: The knowledge management platform project. *British Journal of Management*, 24(2), 264–280.
- Paskaleva, K., Evans, J., Martin, C., Linjordet, T., Yang, D., & Karvonen, A. (2017). Data Governance in the Sustainable Smart City. *Informatics Informatics*, 4(4), 41.
- Patrick J. Lynch, S. H. (n.d.). *Organizing Your Information Web Style Guide, Third edition*. Retrieved January 27, 2019, from <https://webstyleguide.com/wsg3/3-information-architecture/2-organizing-information.html>
- Peixoto, T. (2012). The Uncertain Relationship between Open Data and Accountability: A Response to Yu and Robinson's the New Ambiguity of Open Government. *UCLA L. Rev. Discourse*, 60, 200.
- Pelloux, S. (2016). Cartographie des données: Donner à voir l'open data SNCF. *Sciences Du Design*, 2, 11–15.
- Pereira, G. V., Macadar, M. A., Luciano, E. M., & Testa, M. G. (2017). Delivering public value through open government data initiatives in a Smart City context. *Inf Syst Front Information Systems Frontiers: A Journal of Research and Innovation*, 19(2), 213–229.
- Perrin, A. (2008). *Les pratiques des gestionnaires des connaissances en entreprise* (Doctoral dissertation, Nice).
- Pesqueux, Y. (2008). Un modèle organisationnel du changement ? *Communication et organisation*, 86–92.
- Petychakis, M., Vasileiou, O., Georgis, C., Mouzakis, S., & Psarras, J. (2014). A state-of-the-art analysis of the current public data landscape from a functional, semantic and technical perspective. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 34–47.
- Pfenninger, S., DeCarolis, J., Hirth, L., Quoilin, S., & Staffell, I. (2017). The importance of open data and software: Is energy research lagging behind? *Energy Policy*, 101, 211–215.
- Pflügler, C., Schreieck, M., Hernandez, G., Wiesche, M., & Krcmar, H. (2016). A concept for the architecture of an open platform for modular mobility services in the smart city. *Transportation Research Procedia*, 19, 199–206.
- Pflügler, C., Schreieck, M., Hernandez, G., Wiesche, M., & Krcmar, H. (2016). A Concept for the Architecture of an Open Platform for Modular Mobility Services in the Smart City. *Transportation Research Procedia*, 19, 199–206. <https://doi.org/10.1016/j.trpro.2016.12.080>

- Picon, A. (2013). *Smart Cities: Théorie et critique d'un idéal auto-réalisateur*. B2.
- Picon, A. (2015). *Smart cities: A spatialised intelligence*. John Wiley & Sons.
- Picon, A. (2018). Villes et systèmes d'information: De la naissance de l'urbanisme moderne à l'émergence de la smart city. *Flux, 1*, 80–93.
- Pitt, J., Diaconescu, A., & Bourazeri, A. (2017). *Democratisation of the SmartGrid and the active participation of prosumers*. 1707–1714. <https://doi.org/10.1109/ISIE.2017.8001505>
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2016). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 1461444816661553.
- Poikola, A., Kola, P., & Hintikka, K. A. (2011). Public data. *An Introduction to Opening Information Sources. Helsinki, Finland.: Ministry of Transport and Communications*.
- Pollock, R. (2011). Building the (open) data ecosystem. *Open Knowledge Foundation Blog*, 31.
- Ponte, D. (2015). Enabling an Open Data Ecosystem. *Twenty-Third European Conference on Information Systems, Münster*.
- Power, R., Robinson, B., Rudd, L., & Reeson, A. (2015, March). Scenario planning case studies using open government data. In *International Symposium on Environmental Software Systems* (pp. 207-216). Springer, Cham.
- Press, N. (2004). National Information Standards Organization. *Understanding Metadata*.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). *Strategies for Design Science Research Evaluation*. 255–266.
- Proulx, M.-U. (1995). *Réseaux d'information et dynamique locale*. Université du Québec à Chicoutimi.
- Proulx, S. (2000). *La construction sociale des objets informationnels: Matériaux pour une ethnographie des usages*. Colloque Comprendre les usages d'Internet. <http://barthes.enssib.fr/atelier/articles/proulx2000.html>
- Proulx, S. (2001). Usages des technologies d'information et de communication: Reconsidérer le champ d'étude. *Émergences et Continuité Dans Les Recherches En Information et Communication*, 10–13.
- Proulx, S. (2005). Penser les usages des TIC aujourd'hui: Enjeux, modèles, tendances. *Enjeux et Usages Des TIC: Aspects Sociaux et Culturels, 1*, 7–20.
- Proulx, S. (2006). Pour comprendre l'usage des objets communicationnels,(re) penser le constructivisme. *Degrés, Signe, Culture et Lien Social à l'ère Des Réseaux, 126–127*, B1–B18.
- Prud'homme, J., Doray, P., & Bouchard, F. (2015). *Sciences, technologies et sociétés de A à Z*. Presses de l'Université de Montréal.
- Puech, L., Durand, T., Montmorillon, B. de, Dumez, H., Lecocq, X., Basso, O., Kaczmarek, D., Conservatoire national des arts et métiers (France), École doctorale Abbé Grégoire (Paris), & Laboratoire interdisciplinaire de recherche en sciences de l'action (Paris). (2014). *Processus intrapreneurial: Entre temps alloué et temps autosaisi*. éditeur inconnu.
- Purao, S., Storey, V. C., & Han, T. (2003). Improving analysis pattern reuse in conceptual design: Augmenting automated processes with supervised learning. *Information Systems Research, 14*(3), 269–290.
- Qin, J., Ball, A., & Greenberg, J. (2012). Functional and architectural requirements for metadata: Supporting discovery and management of scientific data. *International Conference on Dublin Core and Metadata Applications*, 62–71.
- Rabari, C., & Storper, M. (2014). The digital skin of cities: Urban theory and research in the age of the sensed and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions, Economy and Society, 8*(1), 27–42.
- Radl, W., Skopek, J., Komendera, A., Jäger, S., & Mödritscher, F. (2013). And data for all: On the validity and usefulness of open government data. *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies*, p.1–4.
- Ramirez et Hachiya (2006). how do firm-specific organizational capital and other intangibles affect sales, value and productivity? evidence from japanese firm-level data. *Int. J. Innovation*

- Technol. Management International Journal of Innovation and Technology Management*, 03(03), 265–282.
- Rathore, M. M., Paul, A., Hong, W.-H., Seo, H., Awan, I., & Saeed, S. (2018). Exploiting IoT and big data analytics: Defining Smart Digital City using real-time urban data. *Sustainable Cities and Society*, 40, 600–610. <https://doi.org/10.1016/j.scs.2017.12.022>
- Ravier, V. (2017). *L'élaboration d'un SIG métropolitain: Analyse comparée entre les métropoles Aix-Marseille-Provence et Grenoble* (p. 72). Institut de Géographie Alpine.
- Recupero, D. R., Castronovo, M., Consoli, S., Costanzo, T., Gangemi, A., Grasso, L., Lodi, G., Merendino, G., Mongiovi, M., & Presutti, V. (2016). An innovative, open, interoperable citizen engagement cloud platform for smart government and users' interaction. *Journal of the Knowledge Economy*, 7(2), 388–412.
- Reggi, L., & Dawes, S. (2016). *Open government data ecosystems: Linking transparency for innovation with transparency for participation and accountability*. 74–86.
- Reix, R. (2011). *Systèmes d'information et management des organisations*. Vuibert.
- Relly, J. E., & Sabharwal, M. (2009). Perceptions of transparency of government policymaking: A cross-national study. *Government Information Quarterly*, 26(1), 148–157.
- Rheingold, H. (1993). *The virtual community: Finding connection in a computerized world*. Addison-Wesley Longman Publishing Co., Inc.
- Ritchie, F. (2017). The "Five Safes": A framework for planning, designing and evaluating data access solutions. *Data for Policy 2017*, London, UK.
- Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), 155–169.
- Roberts, T. L., & Moran, T. P. (1983). The evaluation of text editors: Methodology and empirical results. *Communications of the ACM*, 26(4), 265–283.
- Roger, É. (2014). *Le contrôle de la gestion des connaissances d'un fournisseur stratégique: une application aux projets de la Direction des applications militaires du Commissariat à l'énergie atomique* (Doctoral dissertation, Aix-Marseille).
- Roland, L. K., Sanner, T. A., Sæbø, J. I., & Monteiro, E. (2017). P for Platform: Architectures of large-scale participatory design. *Scandinavian Journal of information systems*, 29(2).
- Romme, A. G. L. (2003). Making a difference: Organization as design. *Organization Science*, 14(5), 558–573.
- Rowe, F. 2014. "What literature review is not: Diversity, boundaries and recommendations," *European Journal of Information Systems* (23:3), pp. 241- 255
- Royer, I., & Zarlowski, P. (2014). Le design de la recherche. In R.-A. T. et al (Ed.), *Méthodes de recherche en Management* (pp. 168–196). Dunod.
<https://halshs.archives-ouvertes.fr/halshs-01074017>
- Ruijter, E., Grimmelikhuijsen, S., Hogan, M., Enzerink, S., Ojo, A., & Meijer, A. (2017). Connecting societal issues, users and data. Scenario-based design of open data platforms. *Government Information Quarterly*, 34(3), 470–480.
- Ruijter, E., Grimmelikhuijsen, S., & Meijer, A. (2017). Open data for democracy: Developing a theoretical framework for open data use. *Government Information Quarterly*, 34(1), 45–52.
- Ruijter, E., Grimmelikhuijsen, S., van den Berg, J., & Meijer, A. (2018). Open data work: Understanding open data usage from a practice lens. *International Review of Administrative Sciences*, 0020852317753068.
- Ruppert E (2015) Doing the transparent state: Open government data as performance indicators. In: Rottenburg R, Merry SE, Park SJ, et al. (eds) *A World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge: Cambridge University Press, pp. 1–18.
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, 22(1), 1–24.

- Sandler, T. (1992). *Collective action: Theory and applications*. University of Michigan Press.
- Satoshi, I. (2017). *Impact of new technologies on the organization of global value chains (1): Dynamic transformations within global value chain typology*. Institute of Developing Economies, Japan External Trade Organization (JETRO).
- Scacchi, W. (2002). Understanding the requirements for developing open source software systems. *IEE Proceedings-Software*, 149(1), 24–39.
- Schafer, V. (2012). Internet et avant? *Les Réseaux*, 73–84.
- Scharmer, O., & Kaufer, K. (2014). *Leading from the emerging future: From ego-system to eco-system economies*. <http://www.credoreference.com/book/bkpleading>
- Schmidt, K. (1994). Cooperative work and its articulation: Requirements for computer support. *Le Travail Humain*, 345–366.
- Scholes, J., & Checkland, P. (1990). Soft systems methodology in action. *Chichester, Wiley*, 876, 910.
- Schön, D. A. (1967). *Technology and change the new Heraclitus*. Dell Pub. Co. <http://books.google.com/books?id=PVZIAQAAIAAJ>
- Schweik, C. M. (2007). Free/open-source software as a framework for establishing commons in science. *Understanding Knowledge as a Commons*, 277.
- Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly*, 37–56.
- Servigne, S. (2008). Conception, architecture et urbanisation des systèmes d'information. *Encyclopædia Universalis, Encyclopædia Britannica*.
- Shen, Y. (2019). Create synergies and inspire collaborations around the development of intelligent infrastructure for human-centered communities. *Journal of the Association for Information Science and Technology*.
- Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly*, 28(2–3), 6–6.
- Shkabatur, J. (2012). Transparency with (out) accountability: Open government in the United States. *Yale L. & Pol'y Rev.*, 31, 79.
- Shore, C. (2010). Locating the Work of Policy. In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 211–224). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.19>
- Shum, S. B., Aberer, K., Schmidt, A., Bishop, S., Lukowicz, P., Anderson, S., Charalabidis, Y., Domingue, J., de Freitas, S., & Dunwell, I. (2012). Towards a global participatory platform. *The European Physical Journal Special Topics*, 214(1), 109–152.
- Shumpeter, J. A. (1961). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. Transl. by redvers opie. Oxford University Press.
- Silvia, P. J. (2007). *How to write a lot: A practical guide to productive academic writing*. American Psychological Association.
- Smith, G., Ofé, H. A., & Sandberg, J. (2016). *Digital service innovation from open data: Exploring the value proposition of an open data marketplace*. 1277–1286.
- Smith, K. K., & Berg, D. N. (1987). *Paradoxes of group life: Understanding conflict, paralysis, and movement in group dynamics*. Jossey-Bass.
- Sobczak, A. (2017). Data Governance as an Element of Implementation Open Government Concept. *Collegium of Economic Analysis Annals*, 46, 215–228.
- Söderman, A. (2016). *How to improve value towards third-party developers: An analysis of the open data platform Trafiklab*. <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-304053>
- Solar, M., Concha, G., & Meijueiro, L. (2012, September). A model to assess open government data in public agencies. In *International Conference on Electronic Government* (pp. 210-221). Springer, Berlin, Heidelberg.

- Song, P., Xue, L., Rai, A., & Zhang, C. (2018). The ecosystem of software platform: A study of asymmetric cross-side network effects and platform governance. *Mis Quarterly*, 42(1), 121-142.
- Spohrer, J., Maglio, P. P., Bailey, J., & Gruhl, D. (2007). Steps toward a science of service systems. *Computer Computer*, 40(1).
- Srinivasan, S. (2012). Mapping the Tso Kar basin in Ladakh. *The Journal of Community Informatics*, 8(2).
- Star, S. L., & Bowker, G. C. (2002). How to infrastructure? In L. A. Lievrouw & S. L. Livingstone (Eds.), *The handbook of new media. Social shaping and consequences of ICTs* (pp. 151–162). London: Sage Publications
- Sterrenberg, L. (2010). Managing the Problematic in Policy Work. In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 115–130). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.12>
- Stevenson, H. H., Jarillo, J. C., & International Institute for Management Development (Lausanne). (1992). *A paradigm of entrepreneurship: Entrepreneurial management*. International Institute for management Development. Research and Development.
- Stolfi D.H, Alba E, Yao X, & 2nd International Conference on Smart Cities, S.-C. 2017. (2017). Predicting car park occupancy rates in smart cities. *Lect. Notes Comput. Sci. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10268 LNCS, 107–117.
- Strömberg, F., & Näslund, E. (2017). *Open Data within a Smart City Initiative: A case study exploring how collaboration can foster innovation within a smart city initiative*.
- Susha, I., Grönlund, Å., & Janssen, M. (2015). Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries. *Information Polity*, 20(1), 19–34.
- Susha, I., Zuiderwijk, A., Janssen, M., & Grönlund, Å. (2015). Benchmarks for evaluating the progress of open data adoption: Usage, limitations, and lessons learned. *Social Science Computer Review*, 33(5), 613–630.
- Tanriverdi, H., Rai, A., & Venkatraman, N. (2010). Research commentary—Reframing the dominant quests of information systems strategy research for complex adaptive business systems. *Information Systems Research*, 21(4), 822–834.
- Teece, D. J., & Consortium on Competitiveness & Cooperation. (1996). *Firm organization, industrial structure, and technological innovation*. University of California at Berkeley, Center for Research in Management, Consortium on Competitiveness & Cooperation.
- Thévenot, J. (2011). *Master systèmes d'information*. Éditions Eska.
- Thévenot, L. (1993). Essai sur les objets usuels. Propriétés, fonctions, usages. *Raisons Pratiques*, 4, 85–111.
- Thiétart, R. A., & Forgues, B. (1995). Chaos Theory and Organization. *Orsc Organization Science*, 6(1), 19–31.
- Tien, J. M. (2007). Services Innovation: Decision Attributes, Innovation Enablers, and Innovation Drivers. In *Service Enterprise Integration* (pp. 39–76). Springer, Boston, MA. https://doi.org/10.1007/978-0-387-46364-3_2
- Tinati, R., Carr, L., Halford, S., & Pope, C. (2012). *Exploring the impact of adopting open data in the UK government*.
- Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics. *Information Systems Research*, 21(4), 675–687. <https://doi.org/10.1287/isre.1100.0323>
- Toivonen, T. (2016). What is the social innovation community? Conceptualizing an emergent collaborative organization. *Journal of Social Entrepreneurship*, 7(1), 49–73.
- Townsend, A. M. (2013). *Smart cities: Big data, civic hackers, and the quest for a new utopia*. WW Norton & Company.

- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Trilles, S., Calia, A., Belmonte, Ó., Torres-Sospedra, J., Montoliu, R., & Huerta, J. (2017). Deployment of an open sensorized platform in a smart city context. *Future Generation Computer Systems*, 76, 221–233. <https://doi.org/10.1016/j.future.2016.11.005>
- Turki, S., & Foulonneau, M. (2015). Valorisation des données ouvertes: acteurs, enjeux et modèles d'affaires. *Big data–Open data: Quelles valeurs*, 113.
- Turki, S., Martin, S., & Renault, S. (2017, September). How open data ecosystems are stimulated?. In *Proceedings of the International Conference on Electronic Governance and Open Society: Challenges in Eurasia* (pp. 179-187).
- Tuti, T., Bitok, M., Malla, L., Paton, C., Muvinga, N., Gathara, D., Gachau, S., Mbevi, G., Nyachiro, W., & Ogero, M. (2016). Improving documentation of clinical care within a clinical information network: An essential initial step in efforts to understand and improve care in Kenyan hospitals. *BMJ Global Health*, 1(1), e000028.
- Ubaldi, B. (2013). Open government data: towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance No. 22*. OECD. Retrieved from dx.doi.org/10.1787/5k46bj4f03s7-en
- Vaishnavi, V. K., Buchanan, G. C., & Kuechler, W. (1997). A data/knowledge paradigm for the modeling and design of operations support systems. *IEEE Transactions on Knowledge and Data Engineering*, 9(2), 275–291.
- Vaishnavi, V. K., & Kuechler, W. (2015). *Design science research methods and patterns: Innovating information and communication technology*. Crc Press.
- Vaishnavi, V., & Kuechler, W. (2004). Design research in information systems. *AISWorld*.
- Van de Ven, A. H., & Johnson, P. E. (2006). Knowledge for theory and practice. *Academy of Management Review*, 31(4), 802–821.
- Van Merriënboer, J. J., Jelsma, O., & Paas, F. G. (1992). Training for reflective expertise: A four-component instructional design model for complex cognitive skills. *Educational Technology Research and Development*, 40(2), 23–43.
- Van Schalkwyk, F., Willmers, M., & McNaughton, M. (2016). Viscous open data: The roles of intermediaries in an open data ecosystem. *Information Technology for Development*, 22(sup1), 68–83.
- Vander Sande, M., Dimou, A., Colpaert, P., Mannens, E., & Van de Walle, R. (2013). Linked Data as enabler for Open Data Ecosystems. *Open Data Week, 2013*.
- Veeckman, C., & Van Der Graaf, S. (2015). The city as living laboratory: Empowering citizens with the citadel toolkit. *Technology Innovation Management Review*, 5(3).
- Venable, J. (2006). A framework for design science research activities. In *Emerging Trends and Challenges in Information Technology Management: Proceedings of the 2006 Information Resource Management Association Conference* (pp. 184-187). Idea Group Publishing.
- Venable, J., Pries-Heje, J., & Baskerville, R. (2012, May). A comprehensive framework for evaluation in design science research. In *International Conference on Design Science Research in Information Systems* (pp. 423-438). Springer, Berlin, Heidelberg.
- Verdier, H. (2015). Les données au service de la transformation de l'action publique. *Rapport Au Premier Ministre Sur La Gouvernance de La Donnée. Administrateur Général Des Données, Secrétariat Général Pour La Modernisation de l'action Publique, Services Du Premier Ministre, Paris, La Documentation Française*.
- Vert, S., & Vasiu, R. (2015). Integrating linked open data in mobile augmented reality applications—a case study. *Tem Journal*, 4(1), 35.
- Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. CRC Press.

- Victorelli, E. Z., dos Reis, J. C., Hornung, H., & Prado, A. B. (2019). Understanding Human-Data Interaction: Literature Review and Recommendations for Design. *International Journal of Human-Computer Studies*.
- Viljoen, M., Dutka, L., Kryza, B., & Chen, Y. (2016). *Towards European Open Science Commons: The EGI Open Data Platform and the EGI DataHub*. 148–152.
- Villum, C. (2014). ‘Open-Washing’—The Difference between Opening Your Data and Simply Making Them Available. *Open Knowledge Blog*. October, 3.
- Viscusi, G., Castelli, M., & Batini, C. (2014). Assessing social value in open data initiatives: A framework. *Future Internet*, 6(3), 498–517.
- Vlaisavljevic, V., Cabello-Medina, C., & Pérez-Luño, A. (2016). Coping with Diversity in Alliances for Innovation: The Role of Relational Social Capital and Knowledge Codifiability. *BJOM British Journal of Management*, 27(2), 304–322.
- Volpi, V., Ingresso, A., Pazzola, M., Opromolla, A., & Medaglia, C. M. (2014). *Roma crash map: An open data visualization tool for the municipalities of Rome*. 284–295.
- Von Alan, R. H., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105.
- Walker, J., & Simperl, E. (2018). Open Data and Entrepreneurship. Analytical Report n10. Retrieved from https://www.europeandataportal.eu/sites/default/files/analytical_report_10_open_data_and_entrepreneurship.pdf
- Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an information system design theory for vigilant EIS. *Information Systems Research*, 3(1), 36–59.
- Walravens, N., Breuer, J., & Ballon, P. (2014). Open data as a catalyst for the smart city as a local innovation platform. *Communications & Strategies*, 96, 15.
- Wang, H.-J., & Lo, J. (2016). Adoption of open government data among government agencies. *Government Information Quarterly*, 33(1), 80–88.
- Wang, J.-C., & Chen, C.-L. (2004). An automated tool for managing interactions in virtual communities-using social network analysis approach. *Journal of Organizational Computing and Electronic Commerce*, 14(1), 1–26.
- Webb, G. S., Fawcett, A. M., Fawcett, S. E., Brockhaus, S., & Knemeyer, A. M. (2015). Relational Evolution: The Role of Managerial Commitment on Collaborative Transformation. *Academy of Management Proceedings Academy of Management Proceedings*, 2015(1), 18636.
- Webb, M., Finighan, R., Buscher, V., Doody, L., Cosgrave, E., Giles, S., Hawes-Hewitt, J., Walt, N., & Mulligan, C. (2011). Information marketplaces: The new economics of cities. *The Climate Group, Arup, Accenture, Horizon*.
- Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: An empirical view from the Citizen’s perspective. *Information Systems Frontiers*, 19(2), 285–300.
- Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*.
- Welle Donker, F., & van Loenen, B. (2017). How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, 10(3), 284–306.
- Whitmore, A. (2014). Using open government data to predict war: A case study of data and systems challenges. *Government Information Quarterly*, 31(4), 622–630.
- Williams, A. (2010). Is Evidence-Based Policy Making Really Possible? In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 195–210). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.18>
- Winter, R., Zhao, J. L., & Aier, S. (2010). *Global Perspectives on Design Science Research: 5th International Conference, DESRIST 2010, St. Gallen, Switzerland, June 4-5, 2010. Proceedings* (Vol. 6105). Springer.
- Wirtz, B. W., & Birkmeyer, S. (2015). Open government: Origin, development, and conceptual perspectives. *International Journal of Public Administration*, 38(5), 381–396.

- Wirtz, B. W., Piehler, R., Thomas, M.-J., & Daiser, P. (2016). Resistance of public personnel to open government: A cognitive theory view of implementation barriers towards open government data. *Public Management Review*, 18(9), 1335–1364.
- Wiseli, D., Tanusetiawan, R., & Purnomo, F. (2017). Simulation Game as a Reference to Smart City Management. *PROCS Procedia Computer Science*, 116, 468–475.
- Woeltjes, T. (2010). Policy Work Between National and International Contexts: In H. K. Colebatch, R. Hoppe, & M. Noordegraaf (Eds.), *Working for Policy* (pp. 159–170). Amsterdam University Press; JSTOR. <http://www.jstor.org/stable/j.ctt46n2ww.15>
- Woodard C. J. (n.d.). *Modeling Architectural Strategy Using Design Structure Networks*. 21.
- Worthy, B. (2015). The impact of open data in the UK: Complex, unpredictable, and political. *Public Administration*, 93(3), 788–805.
- Wu, C., Kao, S.-C., Shih, C.-H., & Kan, M.-H. (2018). Open data mining for Taiwan's dengue epidemic. *Acta Tropica*, 183, 1–7.
- Wurman, R. S., Leifer, L., Sume, D., & Whitehouse, K. (2001). *Information anxiety 2*. Que.
- Yaffee, S. L. (1999). Three faces of ecosystem management. *Conservation Biology*, 13(4), 713–725.
- Yan, J., Liu, J., & Tseng, F. M. (2020). An evaluation system based on the self-organizing system framework of smart cities: A case study of smart transportation systems in China. *Technological Forecasting and Social Change*, 153, 119–371.
- Yang, D., Fu, C.-S., Smith, A. C., & Yu, Q. (2017). Open land-use map: A regional land-use mapping strategy for incorporating OpenStreetMap with earth observations. *Geo-Spatial Information Science*, 20(3), 269–281. <https://doi.org/10.1080/10095020.2017.1371385>
- Yang, T.-M., Lo, J., & Shiang, J. (2015). To open or not to open? Determinants of open government data. *Journal of Information Science*, 41(5), 596–612.
- Yang, T. M., Lo, J., Wang, H. J., & Shiang, J. (2013, October). Open data development and value-added government information: Case studies of taiwan e-government. In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance* (pp. 238–241).
- Yi Y, Wei Z, Li Y, Liu Y, & Hitt M.A. (2016). The influence of resource bundling on the speed of strategic change: Moderating effects of relational capital. *Asia Pac. J. Manage. Asia Pacific Journal of Management*, 33(2), 435–467.
- Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research commentary—The new organizing logic of digital innovation: An agenda for information systems research. *Information Systems Research*, 21(4), 724–735.
- Yu, C.-C. (2016). A value-centric business model framework for managing open data applications. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 80–115.
- Yu, H., & Robinson, D. G. (2011). The new ambiguity of open government. *UCLA L. Rev. Discourse*, 59, 178.
- Zeleti, F. A., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, 33(3), 535–551.
- Žemaitis, E. (2014). Knowledge Management in Open Innovation Paradigm Context: High Tech Sector Perspective. *Procedia - Social and Behavioral Sciences Procedia - Social and Behavioral Sciences*, 110, 164–173.
- Zhang, J., Dawes, S. S., & Sarkis, J. (2005). Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing. *Journal of Enterprise Information Management*.
- Zhang, X., & Yuhao, L. (2012). Knowledge Building in E-Learning. In *E-Learning-Theories, Design, Software and Applications*. IntechOpen.
- Zhu, X., & Freeman, M. A. (2018). An evaluation of US municipal open data portals: A user interaction framework. *Journal of the Association for Information Science and Technology*.
- Zotano, M. A. G., & Bersini, H. (2017). A Data-driven Approach to Assess the Potential of Smart Cities: The Case of Open Data for Brussels Capital Region. *Energy Procedia Energy Procedia*, 111(2), 750–758.

- Zuiderwijk-van Eijk, A. M. G. (2016). The Acceptance and Use of Open Data infrastructures-Drawing upon UTAUT and ECT. *Electronic Government and Electronic Participation*, 23.
- Zuiderwijk, A. M. G. (2015). "Open Data Infrastructures: The Design of an Infrastructure to Enhance the Coordination of Open Data Use." Ph.D. Dissertation, Delft University of Technology.
- Zuiderwijk, Anneke, & Janssen, M. (2014). Barriers and development directions for the publication and usage of open data: A socio-technical view. In *Open government* (pp. 115–135). Springer.
- Zuiderwijk, A., & Janssen, M. (2013). A coordination theory perspective to improve the use of open data in policy-making. In *International Conference on Electronic Government* (pp. 38-49). Springer, Berlin, Heidelberg.
- Zuiderwijk, A., & Janssen, M. (2015, June). Participation and data quality in open data use: Open data infrastructures evaluated. In *Proceedings of the 15th European Conference on E-Government* (pp. 351-359).
- Zuiderwijk, Anneke, Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of E-Government*, 10(2).
- Zuiderwijk, Anneke, Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. *Information Polity*, 19(1, 2), 17–33.
- Zuiderwijk, A., Janssen, M., & Jeffery, K. (2013, May). Towards an e-infrastructure to support the provision and use of open data. In *Conference for E-Democracy and Open Government* (p. 259).
- Zuiderwijk, Anneke, Janssen, M., & Susha, I. (2016). Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 116–146.
- Zuiderwijk, A., Jeffery, K., & Janssen, M. (2012). The necessity of metadata for linked open data and its contribution to policy analyses. In *Proceedings CeDEM2012 Conference, Donau-Universität, Krems* (pp. 281-94).
- Zuiderwijk, A., Loukis, E., Alexopoulos, C., Janssen, M., & Jeffery, K. (2014, May). Elements for the development of an open data marketplace. In *Conference for E-Democracy and Open Government* (pp. 309-322).
- Zwicky, F. (1967). The morphological approach to discovery, invention, research and construction. In *New methods of thought and procedure* (pp. 273–297). Springer.

Glossaire

Analyse de documents - p.176

Travail qui consiste à appliquer une grille de catégorisation de données pour analyser du texte produit par les organisations observées par le chercheur.

Analyse morphologique (Zwicky, 1967) - p.36

Analyse qui consiste à créer des structures ou formes en interconnectant les informations disponibles sur un sujet et ce de manière à pouvoir exposer les zones qui n'ont pas reçues une attention suffisante de la part des chercheurs.

Approche design science - p.19

Approche scientifique visant à créer un artefact utile pour la résolution d'un problème pratique. Cette approche inscrit le chercheur et ses outils dans un rôle de médiation et de facilitation dont bénéficie les praticiens impliqués dans le recherche. Les connaissances produites par cette approche, prescriptives, viennent compléter les connaissances descriptives existantes, permettent d'aborder des questions inaccessibles aux outils méthodologiques traditionnels et permettent de réduire l'écart entre le théorie et la pratique. Par là, cette approche permet d'améliorer la pertinence et la notoriété de la recherche.

Architecture de recherche - p.156

Aussi appelée design de la recherche, elle constitue la trame qui permet d'articuler les différents éléments d'une recherche : problématique, littérature, données, analyse et résultats (Thiétart et al., 2014). Nous considérons dans cette thèse que cette trame doit découler logiquement d'un positionnement ontologique, épistémologique et méthodologique déterminé.

Architecture d'un prototype - p.190

Design structurel qui facilite la construction du prototype et s'assure de sa robustesse

Artefact - p.40

Résultat d'une recherche en design science créé dans le but de répondre à un problème organisationnel important. De nombreux débats existent dans la communauté des chercheurs en système d'information autour de la forme et de la nature que peut prendre cet artefact. D'un côté, pour l'école américaine (Hevner et al., 2004 ; March et Smith., 1995) un artefact se définit par sa dimension matérielle et technologique (*IT artefact*). Il peut prendre la forme d'un modèle, d'un construit/concept, d'une méthode ou d'une instanciation. De l'autre, l'école européenne et son porte-parole Carlsson (2006, 2010, Venable et al., 2012) adoptent une définition plus large. Ils considèrent qu'un artefact doit prendre en compte les individus et les organisations qui gravitent autour des objets matériels et technologiques (*IS intervention*) et peut aussi prendre la forme d'une *design theory*. Offermann et al. (2010), en accord avec cette typologie, distingue les artefacts structurels des artefacts comportementaux.

Artefact de meta-design ou environnement de conception - p.151

Infrastructure technique et sociale dans laquelle de nouvelles formes de design collaboratifs peuvent prendre forme. Il dépasse la notion traditionnelle d'artefact en ce qu'il intègre un processus co-adaptatif entre l'artefact et l'utilisateur, élevant ainsi l'utilisateur au rang de co-développeur ou de co-designer de l'artefact (Fischer et Giaccardi, 2006, p.1).

Barrières socio-techniques - p.175

Éléments ayant une influence négative sur les interactions entre un système technique et un système social et de fait, sur les résultats de ces interactions (Bostrom et Heinen, 1977; Vonk et Ligtenberg, 2010)

Besoins fonctionnels - p.22

Besoins auxquels une intervention en système d'information doit répondre pour passer d'une situation problématique à une situation désirée. Ces besoins fonctionnels comportent trois éléments : 1. une description du contexte (C) social et technique 2. une description des différents éléments de la situation problématique (ESP) 3. une description des solutions (S) entrevues par les opérateurs de l'open data 4. une description de la situation désirée (SD).

Cadre technologique (Orlikowski, 1994) - p.54

Sous-ensemble des cadres organisationnels qui contient les hypothèses, attentes et connaissances relatives à la place de la technologie dans les organisations. Ce sous-ensemble est utilisé pour utiliser la technologie dans les organisations

Codebook - p.184

Document listant un certain nombre de lois explicatives et un code de conduite pour réaliser le potentiel de la vision d'un manifeste et créer un surplus de valeurs (Ladley, 2020)

Connaissance théorique/connaissance pratique - p.93

Tandis que les connaissances théoriques visent à comprendre la nature des Systèmes d'Informations et précisent ce qu'ils font, les connaissances pratiques cherchent à concevoir de nouveaux Systèmes d'Information qui répondent à un but précis. On retrouve la distinction d'Alexander (1969, p.124) entre "les scientifiques qui essayent d'identifier les composants des structures existantes et le designer qui essayer de façonner les composants des structures".

Contexte (Situation Problématique) - p.19

Circonstances qui conditionnent la mise en place d'un artefact

Contexte parents - p.174

Contextes différents de celui qui a permis de faire émerger un objet théorique ou un artefact et qui vise à le mettre à l'épreuve. (Passeron, 2006).

Stabilisation temporaire du mode d'usage d'un artefact dans une direction donnée

Culture de la donnée - p.40

Capacité à comprendre et utiliser les données de manière à informer des décisions. Elle se subdivise en un ensemble spécifique de compétences et de connaissances qui permettent de transformer les données en information et finalement en connaissance actionnable. Ces compétences incluent l'identification, la collecte, l'organisation, l'analyse, la synthèse et la priorisation des données. Elles incluent également le développement d'hypothèses, l'identification des problèmes, l'interprétation des données, la construction, la planification et le suivi de projets relatifs aux données.

Cycle de design - p.20

Coeur du projet de recherche en *design science*. Selon Simon (1996) ce cycle permet de générer et d'évaluer des versions alternatives du design jusqu'à atteindre une version satisfaisante.

Cycle de pertinence - p.19

Cycle de recherche qui initie une recherche en *design science*. Il établit le domaine d'application de la recherche, les besoins auxquels elle va répondre ainsi que les critères acceptables pour évaluer la forme de l'artefact (Hevner et Chatterjee, 2010)

Cycle de rigueur - p.19

Cycle de recherche qui fournit au chercheur des recherches existantes visant à s'assurer du caractère innovant du projet de recherche (Hevner et Chatterjee, 2010).

Data trust - p.86

Ensemble de relations, supporté par un cadre commun adaptable et reproductible, entraînant ses parties à partager les données de manière équitable, juste et sécurisée (Hall-Prentiss report, 2017).

Design de recherche - p.162

Trame qui permet d'articuler les différents éléments d'une recherche : problématique, littérature, données, analyse et résultat (Thiéart et al., p. 168)

Design d'un système - p.184

Documents et graphiques qui décrivent la structure et le comportement de ce système (Offermann et al., 2010)

Design theory - p.96

Design augmenté par un cadre théorique et des propositions qui précisent son comportement dans un contexte réel (Offermann et al., 2010, p. 86)

Dispositif en Systèmes d'Information - p.19

Concept à une double entrée alliant un dispositif social, lieu d'échange, et un dispositif technique le permettant. Le dispositif technique, toujours plus facile à lire, formalise, cristallise le dispositif social d'autant plus que celui-ci a intégré dans ses pratiques de fonctionnement la souplesse et l'interconnexion des technologies (Paquien-séguy, 2012)

Données (*data*) - p.13

Selon l'Administrateur général des données (2015), une donnée numérique est la description élémentaire de nature numérique, représentée sous forme codée, d'une réalité (chose, événement, mesure, transaction, etc.) en vue d'être :
- collectée, enregistrée - traitée, manipulée, transformée - conservée, archivée - échangée, diffusée, communiquée.

Selon leurs destinations, les données peuvent être fermées (réservées à quelques personnes ou à des organisations), partagées (leur usage est conditionné par une licence contractuelle ou des conditions générales d'utilisation) ou ouvertes (ouvertes à tous utilisateurs et à tous usages légaux).

Données métiers - p.214

Données dont l'usage est réservé aux membres d'une seule organisation

Données ouvertes (*open data*) - p.13

Données générées par tout acteur public, privé ou civil et qui respectent les huit principes fondateurs de l'open data, aussi connu sous le nom des huit principes de Sébastopol :

- Instantané : les données doivent être immédiatement disponibles afin de préserver la valeur de la donnée.
- Accessible : les données doivent être disponibles pour le plus grand nombre d'utilisateurs possible et pour différents buts. Les données doivent permettre une large utilisation pour un large public.
- Autorisant une automatisation : les données sont structurées de manière à permettre une automatisation. Il doit donc s'agir de chaînes de caractère
- Non-discriminatoire : les données sont disponibles pour tous sans besoin d'inscription.
- Non-propriétaire : les données doivent être disponibles dans un format ne nécessitant pas l'utilisation d'un logiciel spécifique
- Sans-licence : les données ne doivent pas être sujettes à l'utilisation d'une licence spécifique.

Ecosystème open data - p.16

Écologie informationnelle dont l'équilibre repose sur la circulation de jeux de données ouverts qui agissent à l'image d'une "espèce ingénieur". Cette écologie est constituée de 3 systèmes : le système de production de données, le système de transformation de données en services informationnels et le système d'intermédiation qui coordonne les interactions entre les deux via une plateforme.

Éléments fonctionnels - p.20

Parties d'un ensemble qui forme le dispositif en Systèmes d'Information conçu, à savoir l'environnement de conception des usages dit *Open Data Canvas*.

Entretiens semi-directifs - p.30

Technique d'entretien qui consiste à utiliser un guide structuré pour aborder une série de thèmes préalablement définis (Thiétart et al., 2014).

Evaluation naturaliste - p.194

Type d'évaluation permettant d'explorer la performance d'un artefact dans un environnement réel (utilisateurs, systèmes et problématiques réelles). Ce type d'évaluation permet d'aborder toute la complexité des pratiques dans les organisations réelles et, à ce titre, bénéficie d'une forte validité interne (Venable et al., 2012).

Evaluation artificielle - p.194

Type d'évaluation permettant de tester des hypothèses quant à la performance d'un artefact dans un environnement contrôlé (utilisateurs, systèmes et problématiques fictives). Du fait de la prédominance du paradigme scientifique positiviste, ce type d'évaluation jouit d'une plus grande fiabilité auprès de la communauté scientifique et ce, dans la mesure où elle respecte les critères de falsification et de reproductibilité (Venable et al., 2012).

Evaluation formative - p.195

Type d'évaluation utilisé quand l'artefact est en cours de développement et qu'il s'agit de déterminer où il est question d'apporter des précisions et des améliorations (Venable et al., 2012)

Evaluation sommative - p.195

Type d'évaluation visant à établir la capacité d'un artefact à atteindre les objectifs initialement fixés (utilité, efficacité, etc.)

Exploration hybride utilisant la méthode des cas - p.21

Méthode de recherche qui allie, de manière itérative des observations participantes et des entretiens.(Hlady-Rispal,2002)

Fonctions/fonctionnalités - p.33

Tâches pouvant être exécutées par l'artefact et configurées par le chercheur

Feedback et transformation - p.358

Reproduction, révision ou transformation de tensions dynamiques suite à une action de management délibérée (*feedback loop*) donnant lieu à de l'innovation" (Hargrave et van de Ven, 2017, p.323)

Formes d'un artefact - p.189

Un artefact peut prendre les formes suivantes :

- Des concepts/construits Conceptualisations permettant de décrire le problème au sein d'un domaine et de spécifier les solutions possibles. Ils forment ainsi le vocabulaire du domaine (March et Smith, 1995). Ces conceptualisations peuvent atteindre un degré de formalisation élevé à l'instar des construits utilisés dans la modélisation de données sémantiques (entités, attributs, etc.) ou plus informel à l'instar des concepts utilisés pour analyser le travail coopératif (consensus, participation, satisfaction)
- Un modèle Ensemble de propositions exprimant les relations entre construits. Il permet de représenter la situation problématique et de tirer des conjectures sur les formes possibles de la solution (March et Smith, 1995).

- Une méthode Série d'étapes pour réaliser certaines tâches. Elle se base sur des construits et des modèles (Offerman et al., 2010)
- Une instanciation Réalisation d'un artefact dans un environnement. Elle opérationnalise les construits, les méthodes et les concepts et démontrer leur faisabilité (March et Smith, 1995)
- Une design theory

Gap (research gap) - p.19

Formule qui résume le regard critique porté sur la littérature scientifique dans un domaine et qui propose une réorientation de la recherche qui permettrait une meilleure compréhension d'un problème ou d'un construit. Le gap peut pointer du doigt ce qui n'a pas été réalisé de façon rigoureuse ou pertinente. Il peut également venir identifier des biais méthodologiques ou théoriques dans la littérature. (Rowe, 2014, p.243)

Heuristique - p.21

Connaissance ou artefact caractérisé par le fait qu'il "marche" dans un environnement donnée sans qu'on puisse expliquer dans le détail pourquoi il fonctionne. Cette connaissance ou artefact est validé par le fait qu'il s'insère comme convenu dans un environnement donnée;

Instanciation - p.125

Processus qui permet "d'implémenter des fonctionnalités dans un objet à partir d'une structure prédéfinie et pré-optimisée" (Cofer et Harding, 2006, p.120).

Intervention en Systèmes d'Information - p.135

Design et implémentation d'un artefact dans un système socio-technique où les systèmes d'information sont cruciaux pour atteindre les objectifs visés par l'artefact.(Carlsson, 2010)

Instruments de l'étude de cas - p.176

Appareils de mesure qui découle d'un concept et qui permet de récolter des données sur celui-ci (Thé)

Impératif de décentration/impératif de participation - p.173

Il s'agit des deux tendances à équilibrer dans la posture dite du balancier (Hlady-Rispal, 2002, p.109), utilisé dans la collecte de données qualitatives. L'impératif de participation stipule que pour pouvoir être adopté par le milieu observé et récupérer des informations riches et pertinentes, le chercheur doit adopter les codes de communication et comportements qui lui sont propres. Cette tendance à l'immersion doit être contrebalancée par l'impératif de décentration qui stipule de prendre du recul et ce afin d'éviter l'absorption, synonyme d'arrêt du projet de recherche. Échantillon d'observation. Ensemble d'éléments observés et sur lesquels des données seront rassemblées. (Thiéart et al., 2014)

Intervenant en système d'information - p.376

Chercheur en charge de l'[intervention en Systèmes d'Information](#)

Kernel theories - p.22

Théorie choisie parmi les quatre premiers types de théorie de la classification de Jones (les théories pour analyser, les théories pour expliquer, les théories pour prédire, et les théories pour expliquer et prédire) et qui informe le cinquième type, à savoir les théories pour agir et designer.

Ces kernel theories jouent le rôle de “noyau” (*kernel* en anglais) théorique autour duquel on peut construire le design de l’artefact, résultat final de cette thèse. Hanseth et Lyytinen (2004, p.208) les définissent comme : “les théories issues des sciences naturelles et sociales qui gouvernent les exigences de conception ou le processus qui permet de les construire”.

La carte et le territoire (Situation Problématique) - p.288-289

Une carte n’est pas le territoire signifie que de la même manière que les cartes nous servent à nous repérer dans le réel sans avoir vocation à (ni être en mesure de) remplacer la réalité, nous devons nous souvenir dans l’emploi des modèles heuristiques n’ont pour but que de nous orienter dans la complexité de la réalité, et qu’ils ne représentent pas l’état des choses.

Loi malthusienne des données (Pool, 1984) - p.343

Loi selon laquelle l’écart entre le volume de données produites et le volume de données consommés est voué à s’agrandir du fait que la production suit une croissance exponentielle tandis que la consommation suit une consommation linéaire.

Manifeste - p.184

Artefact, implicite ou explicite, qui “fait valoir la positionnement stratégique déterminé d’une nouvelle idée, exprimé en termes de différenciation et de valeurs (...) et qui assure l’expression d’une orientation partagée dans une dimension symbolique (Cohendet et Simon, 2015, p.2)

Mécanisme - p.22

Dans la perspective réaliste de la science établie par Bhaskar (1978), un mécanisme est “élément constitutif du “domaine du réel”. Les acteurs peuvent s’appuyer sur ces mécanismes pour produire des pattern connus d’évènements dont ils feront éventuellement l’expérience” (Carlsson, 2010, p.215)

Modèle de design (voir Design d’un système)

Ensemble de propositions ou d’affirmations, illustrées graphiquement, et qui expriment les relations entre les différents éléments fonctionnels.

Nature d’un artefact - p.461

Ensemble de composants essentiels d’un artefact. On distingue deux natures possibles parmi les artefacts :

- Les artefacts produits ou structurels Ils sont constitués exclusivement d’objets technologiques (outils, diagrammes, logiciels) dont les individus peuvent s’emparer pour accomplir une tâche
- Les artefacts process ou comportementaux Ils sont constitués également d’objets technologiques mais englobent également des processus qui supposent des interactions directes avec des individus et un changement dans leur comportement.

Objectif de prototypage - p.414

Il correspond à ce que le prototype apporte au projet de recherche. Il peut préciser les besoins fonctionnels, améliorer les interactions entre l'artefact et les utilisateurs ou encore venir assurer la capacité de l'artefact à s'adapter aux évolutions des besoins des utilisateurs.

“On juge un arbre à ses fruits”(Venable et al., 2006, p.5) (mode d'évaluation) - p.460

Mode d'évaluation naturaliste d'un artefact permettant d'explorer la mesure dans laquelle une solution technologique s'adapte à son environnement réel. En ce qu'elle met l'artefact à l'épreuve de la complexité des pratiques humaines au sein d'organisations réelles, cette évaluation permet de juger effectivement l'arbre, c'est à dire la recherche, qui a donné les fruits, à savoir l'artefact (Venable et al., 2006).

Opérateurs de l'open data - p.66

Individus dont le travail consiste à mettre à disposition des autres diverses formes de connaissances relatives à la production, l'édition ou la réutilisation de l'open data.

Outils logiciels - p.189

Outils qui ne nécessitent pas d'avoir des compétences de programmation mais à partir desquels il est possible de construire un logiciel (Ince et Hekmatpour, 1987)

Pattern de design ou de recherche - p.25

Une technique particulière ou générale pour approcher des classes ou types de problèmes similaires à un certain niveau d'abstraction bien que ces problèmes ne soient jamais apparus dans la réalité de manière identique. (Vaishnavi et Kuelcher, 2015)

Pertinence (Research relevance) - p.19

La pertinence d'une recherche interroge sa capacité à pouvoir être appliquée dans le travail quotidien des praticiens qu'elle vise à éclairer. Selon Benbasat et Zmud (1999), elle est déterminée par quatre critères : * le choix d'un sujet en lien avec les problèmes, dilemmes et challenges que rencontre les organisations au moment où la recherche est produite * la forme des prescriptions établies : elles doivent être formulées de telle manière à pouvoir être applicables en l'état par les praticiens * la concordance entre les technologies et problèmes adressés dans la recherche et ceux utilisés par les praticiens au moment de la publication. * le style et le ton des contenus qui facilitent leur accès aux praticiens. **Voir Cycle de pertinence**

Plateforme open data - p.14

Structure modulaire composée d'éléments tangibles et intangibles qui facilitent l'interaction des acteurs et des données ouvertes. Selon Lusch et Nambisan (2015), une plateforme permet de maximiser la densité des données ouvertes à un endroit donné et d'augmenter leur liquidité. (Eisenmann et al., 2008 ; Lusch et Nambisan, 2015)

Posture épistémologique - p.129

Positionnement du chercheur vis à vis de la nature de la connaissance. Cette thèse considère avec Romme que les connaissances doivent être définies par rapport aux objectifs qu'elles visent. Elle considère que la connaissance vise “à produire ou servir la production de systèmes qui n'existent pas encore et par là, induire des changements dans les systèmes existants pour obtenir ceux que l'on souhaite” Romme (2003, p.560)

Posture méthodologique - p.91

Positionnement du chercheur vis à vis de la nature des résultats auxquels ses travaux aboutiront et du mode d'interaction à adopter avec le terrain. A propos de la nature des résultats, cette thèse considère "qu'est contribution tout ce qui accroît la base de connaissance du domaine : approfondissement théorique, méthodologique ou nouvelle expérience acquise par la recherche" (Pascal et al., 2012, p.18) . A propos du mode d'interaction avec le terrain, cette thèse s'immergera dans le terrain pour co-développer des artefacts matériels et des connaissances pratiques pour les mettre à disposition des praticiens. En revanche, il s'occupera seul de la démonstration de la capacité des résultats à provoquer le changement souhaité.

Posture scientifique - p.155

Positionnement du chercheur vis à vis de la nature de la réalité (ontologie) et de la manière d'interagir avec elle. Cette thèse considère avec Simon (1996), que la réalité est constituée d'objets qui ont été façonnés par l'homme dans un but précis (i.e. on dit que ces objets ont des propriétés téléologiques). Pour interagir avec elle, le chercheur doit "être capable de comprendre l'enchevêtrement de (ces) objets dans l'optique de concevoir des artefacts qui pourront s'inscrire dans cet enchevêtrement et le faire évoluer"(Avenier, 2010, p.56)

Principe de design - p.22

Ligne directrice et normative permettant au chercheur-concepteur d'agir. Il se formule selon la logique "PIMCO" : étant donné le contexte théorique C(t), pour résoudre la classe de problème P, il est pertinent de mettre en œuvre l'intervention I dans la mesure où elle déclenchera les mécanismes naturels, sociaux ou techniques M, moteurs pour atteindre l'objectif O.

Principe d'opportunisme méthodologique - p.173

Stratégie adoptée par le chercheur en réponse à la limite de ses ressources temporelles et qui l'amène à faire des compromis entre l'exhaustivité nécessaire (en termes de validité interne et externe) et la volonté de produire des résultats. En se concentrant sur les unités d'analyse les plus accessibles, il va réviser ses ambitions et adapter sa question de recherche. Il peut, à ce titre, réduire les échantillons, préférer des populations exemplaires pour construire une théorie ou encore tester seulement une partie des théories initialement envisagées (Thiétart et al., p.128)

Principes opérationnels (*mise en place de la méthodologie*) - p.155

Aussi appelés principes de construction, ils se définissent comme « un ensemble cohérent de propositions, bâties sur l'état de l'art en sciences des organisations, pour produire de nouvelles conceptions et formes organisationnelles et redévelopper celles existantes» (Romme et Endenburg, 2006, p.288). Ils régissent ainsi le processus de production des artefacts et des connaissances qui constitueront les résultats de nos travaux de thèse.

Cette thèse adopte deux types de principes : ceux relatifs à l'évaluation du produit de la thèse et ceux relatifs à l'évaluation du processus de la thèse. Concernant l'évaluation du produit, cette thèse suivra un double mouvement : - d'un côté, c'est l'artefact, c'est-à-dire la technologie informatique en elle-même qui sera évaluée. - de l'autre, c'est l'usage de cette technologie par des membres d'écosystèmes open data qui sera évaluée.

Concernant le processus, celui-ci sera évalué à deux égards : - par l'intermédiaire de présentations faites aux organisations qui développent des artefacts informatiques pour le compte des travailleurs de l'open data. - en

participant à des conférences scientifiques comme la conférence DESRIST (Design Science Research in Information Systems and Technology) dédiée au design science dans le domaine des systèmes d'information.

Problème pernicieux (*Wicked problem*) - p.15

Problème scientifique caractérisé par : - Des supposés flous et des contraintes basées sur des contextes environnementaux aux contours mal définis - Des interactions complexes entre les sous-composantes du problème et sa solution - Une forte sensibilité au changement aussi bien concernant le processus de design que l'artefact - Une forte dépendance aux capacités cognitives humaines pour produire des solutions efficaces - Une dépendance aux habiletés sociales des êtres humains pour produire des solutions efficaces.

Processus de prototypage - p.413

Un processus de prototypage est une méthodologie de développement d'un système basé sur la construction et l'évaluation de prototypes d'un système et ce dans le but de clarifier, en collaboration avec les futurs utilisateurs, leurs besoins (Baskerville et al., 2009, p.2)

Proposition de design - p.22

Élément permettant d'informer et de guider les praticiens dans la réponse à apporter à leurs besoins opérationnels. Il se formule selon la logique de la règle technologique de Bunge (1967) : étant donné le contexte C et dans l'objectif d'atteindre l'objectif O, alors utiliser l'intervention I.

Propriétés d'un artefact - p.193

Aspects d'un artefact devant être évalués au cours de la recherche. Checkland et Scholes (1990) en distingue cinq :

- **Efficiences** : capacité d'un artefact à résoudre avec un minimum de moyen un problème organisationnel
- **Utilité** : capacité d'un artefact à répondre à la question de recherche qui lui a donné naissance et à résoudre un problème organisationnel en situation d'usage réelle
- **Efficacité** : capacité d'un artefact à produire précisément les effets escomptés, pris indépendamment de situations d'usage réelles.
- **Ethique** : degré de prise en compte des normes morales et éthiques dans la conception et les usages supposés de l'artefact
- **Elegance** : caractère esthétique d'un artefact aidant à sa prise en main par les futurs utilisateurs.

Protocole de l'étude de cas - p.296

méthode systématique de collecte et d'analyse de données de l'étude de cas (Miles et Huberman, 2014)

Question/Objet de recherche - p.17

Une question articulant des objets théoriques, empiriques ou méthodologiques et qui permettra de créer ou découvrir d'autres objets théoriques, empiriques ou méthodologiques pour expliquer, prédire, comprendre, changer la réalité sociale (Thiétart et al., 2014, p.52)

Réseau d'information - p.338

Ensemble de voies relationnelles aptes à faire circuler des flux d'informations entre des acteurs (Proulx, 1995)

Rigueur scientifique (Research rigor) -p.109

Usage précis de méthodes et analyses adaptées à la production de connaissance visée par le chercheur (Benbasat et Zmud, 1999) **Voir** [Cycle de rigueur](#)

Situation problématique/Situation désirée - p.20

Présentation des problèmes pratiques abordés par la recherche (situation problématique) permettant d'entrevoir les solutions que les praticiens souhaitent leur apporter (Carlsson, 2010)

Sparring - p.374

Dans le sens commun, le *sparring* renvoie au *sparring partner*, défini comme l'individu qui participe à l'entraînement d'un sportif avant une confrontation. Tout en gardant l'idée originelle du terme, nous empruntons la conception de *sparring* élaborée par Lüscher et Lewis (2008) pour lesquelles ce terme renvoie à un "dialogue visant à mettre à l'épreuve, par des questions exigeantes et stimulantes, les problématiques que rencontrent les acteurs qui se trouvent face à des tensions paradoxales" (p.225)

Structure médiatrice (Berger, 1976) - p.365

Institution ou artefact qui se tient à la frontière entre la sphère privée de l'individu et la sphère publique.

Substance informationnelle - p.343

Contenu matériel et immatériel transitant à travers les réseaux. Proulx (1995) distingue sept types de substances : l'information naïve, l'information descriptive, l'information impulsive, l'information normative, l'information d'orientation, l'information effective et l'information exécutive.

Surdétermination de l'usage (Proulx, 2000) - p.361

Stabilisation temporairement des modes d'usage des données dans une direction donnée

Synergie - p.64

Manière de coordonner des éléments contradictoires visant à satisfaire chacun d'entre eux (Hargrave et van de Ven, 2017, p.324)

Technique de la boule de neige - p.173

La technique de la boule de neige est une procédure utilisée pour les populations difficiles à identifier. Elle consiste à trouver un premier répondant qui vérifie les critères de sélection définis par le chercheur. On demande à ce premier interlocuteur d'en désigner d'autres, qui seront, eux aussi, susceptibles de présenter les caractéristiques requises, et ainsi de suite (Thiétart et al., p.234).

Théorie des complémentarités - p.171

Théorie expliquant la manière dont des activités coordonnées amène à des coûts moindres et des retours sur investissement plus importants par rapport à des activités non coordonnées.

Théorie de l'apprentissage contextualisé (situated learning theory, Lave et Wenger, 1991) - p.361

Théorie sur l'apprentissage dont la focale se porte sur la relation entre l'apprentissage (considérée comme une co-construction de connaissance) et le contexte social dans lequel il a lieu. Cette théorie affirme la place centrale de l'environnement de travail dans les situations

Théorie de la cognition distribuée (Hutchins, 1995) - p.338

Théorie selon laquelle le processus mental qui permet d'effectuer des tâches concrètes ne se situe pas exclusivement dans le corps et l'esprit de l'utilisateur : il est distribué entre plusieurs agents humains et non-humains qui coopèrent.

Type d'un artefact - p.185

Ensemble auquel se rattache un artefact. Chaque ensemble est défini par l'utilisation finale de l'artefact et sa structure (Offermann et al., 2010).

Type de prototypage - p.414

Ensemble auquel se rattache la procédure de prototypage choisie. On distingue trois ensembles de procédures : *throw it away*, incrémental et évolutif.

Unité d'analyse - p.37

Niveau de compréhension de l'objet de recherche

Usage des données ouvertes (inspiré de Ruijter et al., 2018) - p.13

Construction socio-techniques qui émerge dans un cadre de compréhension partagée et en conséquence des interactions entre trois sphères sociales : la sphère de production des données ouvertes, la sphère de réutilisation des données ouvertes et la sphère d'édition ou d'intermédiation des données ouvertes **Voir** [Design theory](#)

Utilisabilité/Utilité - p.40

L'utilisabilité a trait au fonctionnement d'un artefact et la facilité de son utilisation tandis que l'utilité regarde la capacité d'un artefact à atteindre l'objectif qu'il s'est fixé.

Annexes

<u>Annexe 1</u> : Exemple de pattern de recherche en design science (Vaishnavi et Kuelcher, 2015)	558
<u>Annexe 2</u> : Grille d'observation de projets de développement des usages des données ouvertes	559
<u>Annexe 3</u> : Codebook des unités de sens utilisés sous NVivo (chapitre 3)	560
<u>Annexe 4</u> : Guide d'entretien semi-directif	561
<u>Annexe 5</u> : Présentation du modèle théorique de construction des usages issus des données ouvertes	565
<u>Annexe 6</u> : Liste des thèmes abordés avec les interviewés pendant la première phase de l'étude de cas	567
<u>Annexe 7</u> : Guide d'entretien semi-directif	569
<u>Annexe 8</u> : Exemple de codage assisté par NVivo	570
<u>Annexe 9</u> : Synthèse de l'ensemble des besoins fonctionnels identifiés	571
<u>Annexe 10</u> : Principes de design des éléments fonctionnels	574
<u>Annexe 11</u> : Tableau récapitulatif des items des scénarios d'usage	577
<u>Annexe 12</u> : Liste des fonctions de l'Open Data Canvas	580
<u>Annexe 13</u> : Interfaces du site de l'Open Data Canvas	585
<u>Annexe 14</u> : Conducteur des focus group	586
<u>Annexe 15</u> : Codebook des unités de sens utilisés sous NVivo (chapitre 8)	587
<u>Annexe 16</u> : Extrait du tableau de synthèse des résultats des focus group	588
<u>Annexe 17</u> : Exemple d'interaction au sein de la communauté transport.data.gouv.fr développant les usages des données ouvertes de transport	589
<u>Annexe 18</u> : Architecture métier de l'Open Data Canvas	590
<u>Annexe 19</u> : Extrait de la liste des documents utilisés en données secondaires lors de l'étude de cas (chapitre 4)	591

Annexe 1 : Exemple de pattern de recherche en design science (Vaishnavi et Kuelcher, 2015)

Pattern d'analyse des systèmes complexes (Chapitre 3)

Intention

Analyser un système complexe pour trouver les domaines dans lesquels la recherche est nécessaire pour améliorer la performance ou l'efficacité du système.

Contexte et applicabilité

On connaît ou on a accès à un système complexe. Vous êtes intéressé par la conduite de recherches susceptibles d'améliorer les performances ou l'efficacité du système complexe.

Description

Il convient de rester attentif aux lacunes et aux problèmes tout en procédant à l'analyse suivante du système complexe :

1. Analyser la structure statique du système complexe. Découvrez quels sont les sous-systèmes du système, ainsi que leurs relations entre eux, et appliquez la même analyse aux sous-systèmes de manière récursive. Le plus souvent, on constate que le système est un système hiérarchique ou un "système presque décomposable". (La différence entre un système presque décomposable et un système hiérarchique est que si les interactions entre les sous-systèmes du premier sont faibles par rapport à celles qui existent au sein des sous-systèmes, ces interactions ne sont pas négligeables).
2. Analyser le comportement dynamique du système et étudier comment ce comportement est produit.
3. Étudier l'évolution du système. Les systèmes complexes sont généralement le résultat d'un long processus d'évolution à partir d'un système relativement simple.
4. Tenter une représentation du système, de préférence simple. La représentation des systèmes complexes n'a pas besoin d'être complexe.

Conséquences

On comprendra mieux le système complexe et la façon dont il gère sa complexité. On pourra également voir les domaines problématiques qui peuvent être un point de départ pour formuler un problème de recherche pertinent pour le système complexe.

Exemples

Le problème de recherche a été identifié au cours du processus d'analyse d'un environnement d'opérations complexes et de sa modélisation (Vaishnavi et al., 1997)

Sources et références

Simon, H. (1996). *The Sciences of the Artificial*, troisième édition. Cambridge, MA : MIT Press.

Vaishnavi, V., Buchanan, G., et Kuechler, W. (1997). A Data/Knowledge Paradigm for the Modeling and Design of Operations Support Systems. *IEEE Transactions on Knowledge and Data Engineering*, 9(2), 275-291.

Annexe 2 : Grille d'observation de projets de développement des usages des données ouvertes

Strate	Element	Question empirique	Influence sur l'observation
Statique/globale	Plateforme open data	<ul style="list-style-type: none"> • Quels sont les éléments tangibles et intangibles qui définissent les plateformes open data ? • Quel est le rôle des plateformes open data dans la construction de nouveaux usages à partir des données ouvertes ? 	Etablir et observer la composition des plateformes open data et leur rôle dans la construction de nouveaux usages à partir des données ouvertes
Statique/globale	Parties prenantes	<ul style="list-style-type: none"> • Qui sont les parties prenantes des plateformes ? • Quel(s) rôle(s) jouent ces parties prenantes ? 	Etablir qui participe à l'évolution des plateformes et comment
Statique/globale	Données ouvertes	<ul style="list-style-type: none"> • Quels sont les outils informatiques structurés par les données ouvertes ? 	Etablir et comprendre les outils qui constituent l'écosystème open data
Statique/globale	Ecosystème open data	<ul style="list-style-type: none"> • Qui sont les acteurs qui constituent les systèmes de production, d'intermédiation et de transformation des données ouvertes ? • Quel est l'état initial des données en entrée du système et l'état final des données en sortie du système ? 	Etudier l'influence des acteurs d'un système sur l'état des données
Statique/systémique	Artefacts	<ul style="list-style-type: none"> • Quels sont les types et les objectifs des artefacts technologiques et sociaux à l'oeuvre dans l'usage de l'open data ? 	Comprendre la nature et les objectifs des artefacts à l'oeuvre dans l'usage des données ouvertes
Statique/systémique	Capacité d'action	<ul style="list-style-type: none"> • Quelles sont les capacités d'action des artefacts ? • Quel est leur influence sur l'état des données ouvertes ? 	Etudier l'influence des capacités d'action sur l'état des données ouvertes
Statique/systémique	Etats	<ul style="list-style-type: none"> • Quelles sont les différentes configurations dans lesquelles se trouvent les jeux de données ouvertes ? 	Comprendre les différentes configurations prises par les données ouvertes au cours de leur usage
Dynamique/systémique	Séquence d'articulation	<ul style="list-style-type: none"> • Quelles sont les différentes séquences d'articulation identifiables dans l'usage des données ouvertes ? 	Comprendre les différentes séquences qui définissent l'usage des données ouvertes
Dynamique/systémique	Complémentarités	<ul style="list-style-type: none"> • Quel est le résultat de chacune des séquences d'articulation ? • Quelle est la valeur ajoutée émergente de chacun de ces résultats ? 	Etudier le résultat des séquences d'articulation et la valeur ajoutée générée par chacune d'entre elle

Annexe 3 : Codebook des unités de sens utilisés sous NVivo (chapitre3)

Thème	Code	Description	Catégorie	Description	Code
Artefacts	artefacts	Comprendre la nature et les objectifs des artefacts à l'oeuvre dans l'usage des données ouvertes	Type	Type des artefacts (technologiques ou sociaux)	artefacts_type
			Objectif	Objectifs des artefacts	artefacts_objectif
Capacité d'action	capacite_action	Etudier l'influence des capacités d'action sur l'état des données ouvertes	idem	idem	idem
			Influence	Influence de la capacité d'action sur l'état des données	capacite_action_influence
Complémentarités	complementarites	Etudier le résultat des séquences d'articulation et la valeur ajoutée générée par chacune d'entre elle	Résultat	Résultat des séquences d'articulation	complementarites_resultat
			Valeur ajoutée	Valeur ajoutée émergeant des résultats	complementarites_valeur_ajoutee
Données ouvertes	donnees_ouvertes	Etablir et comprendre les outils qui constituent l'écosystème open data	idem	idem	idem
Ecosystèmes open data	ecosysteme_OD	Etudier l'influence des acteurs d'un système sur l'état des données	Acteurs	Acteurs qui constituent chacun des système	ecosysteme_OD_acteurs_systeme
			Etat initial/final des données	Etat initial des données en entrée du système et état final en sortie du système	ecosysteme_OD_etat_donnees
Etat des données	etats_donnees	Comprendre les différentes configurations prises par les données ouvertes au cours de leur usage	idem	idem	idem
Parties prenantes	parties_prenantes	Etablir qui participe à l'évolution des plateformes et comment	Identifité	Identifité des parties prenantes	parties_prenantes_identite
			Rôle des parties prenantes		roles_parties_prenantes
Plateforme open data	plateforme_OD	Etablir et observer la composition des plateformes open data et leur rôle dans l'innovation de services	Elements tangibles et intangibles de la plateforme		plateforme_elements
			Rôle de la plateforme dans la construction des usages		plateforme_role
Séquence d'articulation	sequence_articulation	Comprendre les différentes séquences qui définissent l'usage des données ouvertes	idem	idem	idem

Annexe 4 : Guide d'entretien semi-directif

Construire les usages des données ouvertes

Présentation

- Étudiant en thèse en Sciences de Gestion à l'UGA
- Ma thèse est financée par la Région Rhône-Alpes Auvergne Ma thèse s'inscrit dans une démarche d'investigation visant à estimer quels peuvent être les retours sur investissement des démarches d'ouverture des données menées sur le territoire de la région.

Objectif

- Évaluer la pertinence d'hypothèses sur la construction des usages issus des données ouvertes
- Ces entretiens nous permettront d'affiner un modèle empirique visant à comprendre comment les usages se construisent
- Aussi est-il question de trouver les zones du processus de construction dans lesquelles la recherche est nécessaire pour améliorer la performance ou l'efficacité de cette construction ou simplement pour la faire advenir
- Nous *designerons* un dispositif qui viendra adresser ces zones précises

Date de l'entretien :

Anonymat souhaité :

Lieu :

Autorisation d'enregistrement :

Interviewer(s) :

Partie 1 : Généralités

Nom complet:

Nom de la structure:

Domaine de données adressé au quotidien:

Historique professionnel:

Compétences technique en manipulation de données:

Nom des projets impliquant des données ouvertes:

Dates de lancement des projets:

Partie 2 : Déroulement des projets d'usage (Séquence d'articulation)

- *Conjecture/Hypothèse* : des blocages d'ordre technique bloquent les séquences d'articulation permettant l'usage des données ouvertes.
- *Question* : quels sont les blocages d'ordre technique dont vous avez fait l'expérience ?
- *Conjecture/Hypothèse* : des blocages d'ordre sociaux (relations entre parties prenantes du projet) bloquent les séquences d'articulation permettant usage des données ouvertes.
- *Question* : quels sont les blocages d'ordre sociaux dont vous avez fait l'expérience ?

Partie 3 : Organisation/Gouvernance des projets d'usage (Capacité d'action des artefacts sociaux et techniques)

- *Conjecture/Hypothèse* : il existe un(des) blocage(s) social qui empêche(ent) aux instruments de gouvernance, de coordination, de direction des projets (artefacts sociaux) de remplir leurs missions ;
- *Question* : quels sont selon vous les blocages qui freinent l'organisation, la gouvernance des projets d'usage ?
- *Conjecture/Hypothèse* : il existe un(des) blocage(s) socia(ux)l qui empêche(ent) aux artefacts techniques d'exprimer leurs capacités d'action.
- *Question* : quels sont selon vous les blocages qui freinent l'utilisation des plateforme et autres outils techniques ?

Partie 4 : Utilisation des plateformes open data (Plateforme open data)

- *Conjecture/Hypothèse* : il existe un(des) blocage(s) social qui empêche(ent) la mise en place d'une organisation sociale autour des plateformes open data.
- *Question* : quels sont selon vous les blocages qui freinent la mise en place d'une organisation sociale autour des plateformes open data ?
- *Conjecture/Hypothèse* : il existe un(des) blocage(s) qui empêche(nt) d'activer les fonctionnalités techniques des plateformes open data.
- *Question* : quels sont selon vous les blocages qui freinent l'activation des fonctionnalités techniques des plateformes open data ?

Annexe 5 : Présentation du modèle théorique de construction des usages issus des données ouvertes

2a. Concepts constitutifs de la formation des usages de l'open data (1/3)

Données brutes
Matière informationnelle "pure" n'ayant pas subi de modification ou d'agrégation (Goeta, 2016)

Écosystème open data
Écologie informationnelle dont l'équilibre repose sur la circulation des données entre trois sphères sociales (production, réutilisation, intermédiation)

Usages de l'open data
Constructions socio-techniques qui émergent dans un cadre de compréhension partagée et en conséquence des interactions entre les trois sphères sociales

Le diagramme illustre un cycle de données. Trois cercles gris sont disposés en triangle : 'Production' en haut à gauche, 'Édition' en haut à droite, et 'Réutilisation' en bas. Au centre se trouve un cercle orange 'Données brutes' entouré d'une zone pointillée 'Plateforme open data'. Des flèches rouges indiquent le flux de données brutes de Production vers la plateforme, de la plateforme vers Édition, et de Réutilisation vers la plateforme. Des flèches grises indiquent des interactions entre Production et Édition, et entre Édition et Réutilisation.

Rhône Alpes Thèse – Arthur Sarazin – 8 octobre 2020 CERAG 8

2a. Concepts constitutifs de la formation des usages de l'open data (2/3)

Représentation des données
Combinaison des caractéristiques substantielles (i.e contenu et signification) et formelles (i.e formes de sa présentation) des données

Conditions d'interaction avec les données
Éléments qui façonnent la manière dont les agents humains interagissent avec des jeux de données et les artefacts permettant de les manipuler

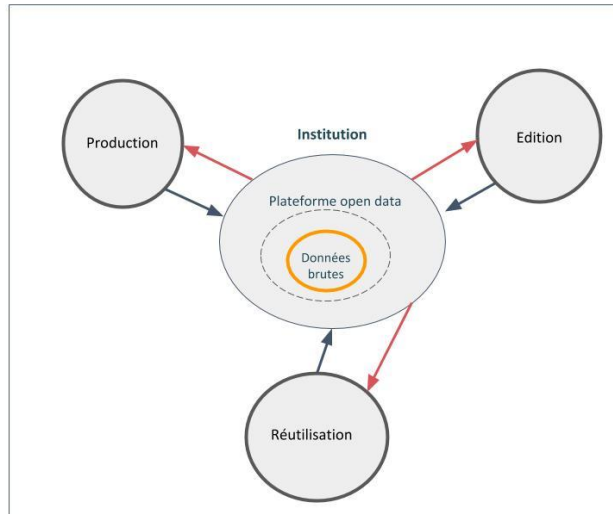
Utilisation des données
Transformation de l'état d'un jeu de données grâce à la capacité d'action d'artefacts sociaux et technologiques

Le diagramme illustre des interactions plus complexes. Les mêmes trois cercles gris ('Production', 'Édition', 'Réutilisation') et le cercle orange 'Données brutes' dans la 'Plateforme open data' sont présents. Ici, les flèches sont bidirectionnelles et de couleur mixte (bleu et rouge). Des flèches bleues relient Production et la plateforme, la plateforme et Réutilisation, ainsi que Réutilisation et Édition. Des flèches rouges relient Édition et la plateforme, la plateforme et Production, ainsi que Édition et Réutilisation. Des flèches grises indiquent également des interactions entre Production et Édition, et entre Édition et Réutilisation.

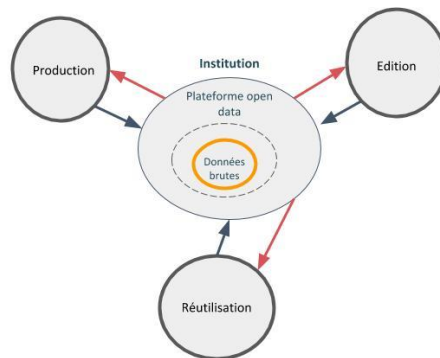
Rhône Alpes Thèse – Arthur Sarazin – 8 octobre 2020 CERAG 9

2a. Concepts constitutifs de la formation des usages de l'open data (3/3)

Institutionnalisation de l'usage des données
Création d'une structure tierce exploitant les usages et produisant de nouvelles données



2b. Présentation du modèle de compréhension de la formation des usages des données



Annexe 6 : Liste des thèmes abordés avec les interviewés pendant la première phase de l'étude de cas

	Producteurs de données	Intermédiaire de données	Réutilisateur de données
Première liste étendue	<p>Contexte :</p> <ul style="list-style-type: none"> - lois et règlements entourant et influençant l'accès et l'utilisation des données dans la ville - type d'organisation impliqués dans la collecte et la gestion des données ouvertes - écosystème autour des données ouvertes - volonté politique/Leadership dans la démarche de production des données - ressources et compétences techniques mises à disposition de la production de données ouvertes - contexte économique de production des données <p>Données produites :</p> <ul style="list-style-type: none"> - définition d'une donnée ouverte - dimensions techniques, légales et sociales des données ouvertes - jeux disponibles et jeux à ouvrir - procédure d'analyse de qualité des données <p>Utilisation des données produits</p> <ul style="list-style-type: none"> - nom et caractéristiques des utilisateurs - but des utilisations <p>Impact des données produites</p> <ul style="list-style-type: none"> - bénéfices sociaux - bénéfices environnementaux - optimisation de la gouvernance 	<p>Objectifs de la démarche open data</p> <ul style="list-style-type: none"> - ambition - dispositifs mis en place - indicateurs, méthodes et suivi de performances <p>Situation actuelle</p> <ul style="list-style-type: none"> - nombre de jeux de données - nombre de réutilisations de jeux de données <p>Actions pour réduire l'écart entre les objectifs et la situation actuelle</p> <ul style="list-style-type: none"> - animation de la communauté - appel à projet <p>Caractéristiques de la plateforme open data</p> <ul style="list-style-type: none"> - type de plateforme - documentation mise à disposition - fonctionnalités à disposition de l'utilisateur <p>Guide d'utilisation de la plateforme</p> <ul style="list-style-type: none"> - Prise en main de la plateforme par des acteurs externes - Phases de prise en main de la plateforme - Standard de la méthode de prise en main <p>Evaluation de la méthode de déploiement de la démarche open data</p> <ul style="list-style-type: none"> - Précision de la méthode - Evolutions de la méthode 	<p>Présentation de la réutilisation</p> <ul style="list-style-type: none"> - Origine de l'idée - Objectif de la réutilisation - Créateur de la réutilisation - Fonctionnalités <p>Vision et stratégie de la réutilisation</p> <ul style="list-style-type: none"> - Stratégie commerciale ? sociale ? - Articulation entre différents parties prenantes <p>Gouvernance de la donnée</p> <ul style="list-style-type: none"> - cadre d'usage de la donnée - actions autour de la donnée - collaborations autour de la donnée <p>Impact de la réutilisation</p> <ul style="list-style-type: none"> - Indicateurs de suivi de la réutilisation

Annexe 7 : Guide d'entretien semi-directif

Construire les usages des données ouvertes

Présentation

- Etudiant en thèse en Sciences de Gestion à l'UGA
- Ma thèse est financé par la Région Rhône-Alpes Auvergne
- Ma thèse s'inscrit dans une démarche d'investigation visant à estimer quels peuvent être les retours sur investissement des démarches d'ouverture des données menées sur le territoire de la région.
- Ma thèse vise à construire un dispositif pour vous assister dans la construction des usages des données ouvertes

Objectif

- Comprendre les barrières socio-techniques qui freinent la construction des usages issus des données ouvertes à deux niveaux :
 - interactions homme-machine
 - interactions sociales avec les autres opérateurs des données ouvertes
- Identifier les besoins qu'un dispositif d'assistance à la construction des usages devrait combler

Date de l'entretien :

Anonymat souhaité :

Lieu :

Autorisation d'enregistrement :

Interviewer(s) :

Partie 1 : Généralités

Nom complet:

Nom de la structure:

Domaine de données adressé au quotidien:

Position dans l'écosystème : producteur/intermédiaire/réutilisateur

Historique professionnel:

Compétences technique en manipulation de données:

Nom des projets impliquant des données ouvertes:

Dates de lancement des projets:

Partie 2 : Présentation du modèle et des différentes barrières identifiées jusqu'à présent

Etape d'usage	Barrières socio-techniques
Utilisation des fonctionnalités de la plateformes <i>téléchargement/flux de données</i>	Manque de confiance dans la plateforme
	Plateforme disponible dans une seule langue
	Manque de lisibilité des interfaces
Evaluation du potentiel des données à partir de leur représentation	Manque de connaissance du domaine relatif aux données
	Manque de compétences et d'expertise technique
	Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation
Transformation des données grâce à des artefacts techniques	Format de données illisible par machine
	Mauvaise qualité des données
	Hétérogénéité des conditions d'utilisation des données
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	Manque d'un cadre juridique et politique clair
	Manque d'une vision cohérente autour de la production et de l'utilisation des données
	Le classement de l'open data en bas de la liste des priorités
Institutionnalisation de l'usage des données	L'incompatibilité des structures organisationnelles avec l'ouverture des données
	Divergence culturelle de l'ouverture des données par rapport aux cultures organisationnelles

Partie 3 : Discussion autour des barrières et solutions potentielles

- *Question* : dans l'ordre des barrières évoquées, quelles sont les expériences qui se rattachent à elles ?
- *Question* : quel dispositif/mécanisme/solution permettrait de lever ces barrières ?

Annexe 8 : Exemple de codage assisté par NVivo

The screenshot shows the NVivo interface with a document titled "Retranscription - Entretien informel - Vincent Köber". The text in the document is as follows:

Pour prendre l'exemple de la ville de Vars, on (équipe animation open data de la métropole de Grenoble) a été faire une animation pour voir quels jeux de données ils pouvaient avoir et on s'est rendu compte qu'ils avaient la budget. Or, le budget, sur les 22 communes au dessus de 3500 habitants dans la métropole, il y a 10 logiciels différents de gestion budgétaires. Donc, on va demander au logiciel de nous extraire cette donnée, ce qui voudrait dire qu'elle n'est pas au bon format, donc il faut la transformer. Donc quel outil pour prendre un document dans un format source, dire qu'on va basculer le nom des colonnes, je mets la date au bon format et je génère un csv qui correspond idéalement au Socle Commun des Données Locales. Là se présente 2 alternatives :

1. La première solution qui est de faire appel à une entreprise de logiciel comme Oxyd qui propose sur les délibérations une license Interface open data : 1200 euros TTC + une prestation de paramétrage : 1140 euros TTC + transfert de compétences : 2700 euros + compétences hors site : 900 euros = 5900 euros TTC pour la ville de Vincesmes. Et uniquement pour la brique logicielle qui permet d'avoir un bouton Open data pour sortir les délibération au format. Et ce qu'ils peuvent faire, c'est reprendre le schéma du Socle Commun des Données locales et sortir les délibérations selon ce format.
2. Au niveau d'Open Data France, on dit qu'on peut faire des extractions à partir d'Oxyd, au format brut donc pas au format SCDL mais quels sont les outils que l'on peut utiliser, par exemple sur R avec un script. On a listé plusieurs cas possibles :
 - a. **Source de la donnée** : soit j'extrais la donnée du logiciel ; soit récupération en API ; soit fichier CSV en téléchargement
 - b. **Transformation de la donnée** : script Bash en R ou Python ou un ETL et alors tu peux penser que tous les sorties Oxyd devraient avoir le même format, ce qui veut dire que si une commune le fait, toutes les communes le font.

The coding panel on the right shows a code strip for "Transformation des données" with a density of 1.0.

The screenshot shows the NVivo interface with a list of files on the left. The selected file is "Transformation des donn...". The document content is as follows:

Files\Cas Grand Lyon\Documents\Paquiénéguy et al. - 2016 - SYNTHESE DES RESULTATS DE L'ENQUETE AUPRES DE RE-U
3 references coded, 0.89% coverage

Reference 1: 0.21% coverage
Toutefois, les données en temps réel « sont des OPENSENSING CITY – Livrable n°1 – ELICO – Septembre 2016 – dir. F. Paquiénéguy 6 données difficiles à gérer, ça demande de gros volumes de données, des connections entre machine en temps réel, des contraintes de la chaîne technique » [Entretien 6].

Reference 2: 0.57% coverage
«Nous ce qu'on essaie de faire, c'est de travailler beaucoup avec OpenStreetMap parce qu'on croit beaucoup à ce projet-là, cela pourrait être finalement une base commune, vraiment commune des données ouvertes. Aujourd'hui la grande difficulté avec tout open data c'est finalement la mise à jour, de quand datent ces données, est-ce que ce sont les dernières, est-ce qu'on va pouvoir les croiser avec d'autres données. Or si tout le monde faisait l'effort aujourd'hui de pousser tout cela vers OpenStreetMap, c'est vrai que cela permettrait d'avoir une base des données commune et cohérente. Nous, en interne, c'est vraiment ce qu'on essaye de faire, garder finalement moins des données chez nous et d'avoir le plus des données possibles sur OpenStreetMap parce que c'est plus facilement maintenable » [Entretien 14].

Reference 3: 0.11% coverage
« La complexité de certains formats et la structuration métier de certaines données les rendent bien souvent inintelligibles au citoyen » [Entretien 5].

Annexe 9 : Synthèse de l'ensemble des besoins fonctionnels identifiés

Barrières socio-techniques	Exemples de barrières dans les cas de Grenoble et Lyon	Besoins fonctionnels
Manque de confiance dans la plateforme	Absence de garantie sur la caractère brut des données	1. Le dispositif en système d'information doit supporter la documentation du processus de production de donnée
	Manque d'informations sur la production des données <i>métadonnées</i>	2. Le dispositif en système d'information doit supporter la documentation de tous les mécanismes techniques ou juridiques qui fixent des restrictions sur les conditions d'utilisation des données
	Doute sur les restrictions d'usage apposées aux données par les licences (restriction économique, conditions de réutilisation)	
Manque de lisibilité des interfaces	Plateforme perçue comme des vitrines et non comme un point d'accès aux données	3. Le dispositif en système d'information doit faciliter le design de plateformes <i>open data</i> à double entrée, la première étant réservée aux développeurs et expert de la données et la deuxième réservé au grand public.
Manque de compétences et d'expertise technique	Format illisible par machine	4. Le dispositif en système d'information doit faciliter la sensibilisation des acteurs sur ce qu'est une donnée et comment la présenter de manière à la rendre exploitable par le plus grand nombre
	Complexification constante des données et accroissement du volume	5. Le dispositif en système d'information doit fournir des flux de connaissances actualisés sur ce qu'il est possible de faire avec les données et les méthodes pour libérer ce potentiel.
	Manque d'expertise technique spécifique aux données	
Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation	Méconnaissance des réutilisateurs de données du fait de la libre utilisation des données\Effet bouteille à la mer	6. Le dispositif en système d'information doit fournir un outil d'interaction entre les producteurs et les intermédiaires d'un côté, et les réutilisateurs de l'autre

Mauvaise qualité des données	Données non-maintenues dans le temps	7. Le dispositif en système d'information, à savoir qu'il doit offrir des outils permettant de comparer la qualité des données entre elles et surtout dans le temps.
Hétérogénéité des conditions d'utilisation des données	Complexité des dispositifs réglementaires	8. Le dispositif en système d'information doit offrir des outils permettant à tout acteur de filtrer, trier et structurer les données selon des critères de transformabilité
	Spectre des données à traiter trop large par rapport aux ressources disponibles	
Manque d'un cadre juridique et politique clair	Manque d'un cadre stratégique clair <i>objectifs moyens alloués</i>	9. Le dispositif en système d'information doit fournir un parcours de tâches concrètes et collectives à mettre en oeuvre, nécessitant le moins de moyens possibles et permettant d'atteindre des résultats visibles.
Manque d'une vision cohérente autour de la production et de l'utilisation des données	Coût trop élevé du travail d'évangélisation par rapport aux ressources humaines disponibles	10. Le dispositif en système d'information doit offrir des outils permettant d'automatiser et d'autonomiser une partie du travail d'évangélisation.
L'incompatibilité des structures organisationnelles avec l'ouverture des données	Incompatibilité de l'architecture des systèmes d'information avec l'institutionnalisation de la production	11. Le dispositif en système d'information doit fournir un exemple de règles clés à suivre dans le développement d'une structure adaptée à l'usage des données ouvertes.
	Incompatibilité de l'ouverture des données avec la démarche de développement des système d'information dites "applicative" (un besoin, une application, une base de données)	
	Opposition entre la silotage organisationnel et la transversalité portée par l'ouverture des données	

Divergence culturelle de l'ouverture des données par rapport aux organisations	Opposition entre exigences de résultats à court\-terme des organisations et les changements structurels portés par l'ouverture des données.	12. Le dispositif en système d'information doit fournir des exemples de principes qui cadrent la manière de concevoir l'ouverture des données.
	Opposition entre culture organisationnelle averse aux risques et la culture opportuniste portée par l'ouverture des données.	
	Opposition entre culture organisationnelle portée sur le secret des affaires et la culture de la transparence portée par l'ouverture des données.	

Annexe 10 : Principes de design des éléments fonctionnels

Élément fonctionnel	Principes de design
Réseau d'information	<p>Premier principe : lorsque plusieurs groupes constitutifs sont en phase de débat pour stabiliser l'usage d'un objet technologique dans une direction donnée (C(t)), pour éviter de cristalliser le débat (P), il est pertinent d'incorporer dans le design du réseau d'information un "cadre technologique" de l'objet sous la forme de prescription d'interdiction, des dispositifs contraignants et des impositions de normes de "bon usage" du réseau(I). Ces trois moyens vont permettre de surdéterminer l'usage du réseau et, par extension, de l'objet dans une direction donnée (M) pour finalement stabiliser le mode d'usage de l'objet dans une direction donnée.</p>
	<p>Deuxième principe : lorsqu'il est question pour des acteurs de mener à bien des tâches relatives à l'usage des données C(t), bien qu'ils ne possèdent pas le savoir-faire nécessaire (P), il est pertinent de rendre le réseau d'information interactif (I) dans la mesure où cette interactivité va obliger les acteurs à acquérir un minimum de savoir-faire technique (M), leur permettant d'acquérir alors un minimum de culture de la donnée (O) moteur pour la réalisation des tâches.</p>
	<p>Troisième principe : lorsque des acteurs isolés (P) doivent coopérer pour effectuer certaines tâches (C(t)), il est pertinent de présenter le réseau de manière complémentaire à la représentation que les acteurs ont de l'usage du réseau (I), dans la mesure où cela facilitera la distribution de la cognition (M), motrice pour permettre l'accomplissement de tâches collectives.</p>
Système de gestion des tâches	<p>Premier principe : lorsqu'un groupe d'acteurs peu nombreux souhaite transmettre des savoir-faire pratiques à tout un environnement constitué d'un grand nombre d'acteurs (C(t), P), il est pertinent de permettre au système basé sur un modèle de métadonnées d'être organisé de manière différente (I) dans la mesure où cela lui permettra de s'adapter aux</p>

	<p>différentes architectures cognitives des futurs utilisateur (M), et par là, de faciliter la propagation d'une forme d'organisation et de savoir-faire d'une partie de l'environnement cognitif vers toutes les autres parties (O).</p>
	<p>Deuxième principe : lorsque des acteurs se trouvent en situation de devoir réaliser certaines tâches pour lesquelles ils ne sont pas qualifiés (C(t)), et donc qu'ils ne savent que faire et dans quel ordre (P), il est pertinent de permettre au système basé sur un modèle de méta-données d'emprunter différentes représentations spatiales traduisant l'ordre des tâches à effectuer et la manière de les mener (I), dans la mesure où cela facilitera la coordination entre la procédure suggérée par un artefact médiateur et la procédure mentale invoquée par les acteurs pour l'utiliser (M). Cela permettra aux acteurs de comprendre l'ordre dans lequel il s'agit d'effectuer une séquence d'action et le nombre de fois que chaque action doit être menée (O).</p>
	<p>Troisième principe : lorsque des acteurs ne sont pas familiers avec le domaine d'application d'un l'artefact (C(t)), et qu'ils ne peuvent donc avoir accès au sens des tâches suggérées par l'artefact, ce qui comporte un risque de délaissement de l'artefact par ces acteurs (P), il est pertinent de mentionner le monde réel auquel est confronté l'utilisateur à chacune des tâches suggérées par le système basée sur le modèle de métadonnées (I) dans la mesure où cela assurera la coordination entre l'artefact, son sens et le monde (M), coordination importante pour maximiser les chances que l'artefact suscite une action réelle chez ses utilisateurs (O).</p>
<p>Manifeste & codebook</p>	<p>Premier principe : lorsque les membres d'une communauté épistémique sont confrontés au risque de voir leur manifeste mal interprété (C(t)), de sorte que les procédures, méthodes et expériences promues dans le codebook ne sont pas adoptées en dehors de la communauté (P), il est pertinent de concevoir le manifeste et codebook de manière à ce qu'ils puissent intégrer de nouvelles conceptions et pratiques (I) dans la mesure où cela leur permettra de faire appel au mécanisme de traduction/inscription</p>

	(M), moteurs pour attirer de plus en plus d'acteurs et de communautés et d'organisation dans leur giron (O).
	Deuxième principe : lorsqu'une communauté épistémique cherche à convaincre la hiérarchie d'une structure que la connaissance novatrice peut donner lieu à de l'innovation et de la création de valeur, et qu'en sus, cette innovation peut s'insérer dans la structure (C(t), P), il est pertinent d'intégrer les connaissances du manifeste et du codebook, sous formes de prototypes, dans des actions concrètes (I) dans la mesure où cela active le mécanisme de "creatifizing" qui permet que la connaissance atterrisse dans une structure pré-existante (O).

Annexe 11 : Tableau récapitulatif des items des scénarios d'usage

	Scenario 1 - Rallier	Scenario 2 - Mutualiser	Scenario 3 - Mettre à disposition une forge
<i>Items - environnement d'utilisation de l'artefact</i>			
Utilisateurs	Acteurs institutionnels ou associatifs oeuvrant au niveau territorial et/ou national pour favoriser l'usage des données ouvertes.	Tout acteur individuel impliqué dans un projet de développement des usages des données ouvertes	Groupes de réutilisateurs de données répartis par territoire, par organisation ou par thématique
Partenaires	Acteurs locaux coordonnés à l'échelle nationale	Chef(fes) de projet utilisant des données ouvertes	
Tâches	<ul style="list-style-type: none"> - Co-construction du manifeste et du codebook - Développer des produits ou services à partir du manifeste et du codebook 	<ul style="list-style-type: none"> - Déposer des connaissances dans le système de gestion des tâches - Placer les ressources dans le modèle de métadonnées - Utiliser les connaissances relatives aux tâches permettant de développer les usages - Organiser/Visualiser les tâches à réaliser 	<ul style="list-style-type: none"> - Décrire les usages qu'ils souhaitent développer - Être questionné pour formuler le paradoxe qu'ils contiennent - Etablir un plan de développement des usages - Solliciter des feedback sur ce plan de la part des acteurs de l'écosystème

		<ul style="list-style-type: none"> - Obtenir des retours sur les tâches planifiées de la part de l'écosystème - Acquisition d'un savoir-faire homogène 	<ul style="list-style-type: none"> - Inventorier les données à utiliser, les documenter et interagir avec l'écosystème autour d'elles - Utiliser le réseau d'information pour s'orienter dans la réalisation des tâches nécessaires à l'usage
Ressources	Communautés en ligne <i>et/ou</i> réseau de collectivités	Accès privilégié pour récupérer les tâches et ressources à la source.	Connaissances couvrant l'ensemble des étapes propres à l'usage des données et pouvant être utilisées avec un minimum d'accompagnement
Information ou flux d'information			
Outputs	<ul style="list-style-type: none"> - Diffusion du manifeste et du codebook - Renforcer l'adhésion des communautés vav du manifeste et du codebook 	<ul style="list-style-type: none"> - Mutualiser et diffuser un savoir-faire homogène 	<ul style="list-style-type: none"> - Permettre à un maximum d'acteurs de se lancer dans l'usage des données ouvertes
<i>Items - environnement structurel des utilisateurs de l'artefact</i>			
Schémes	Dispersion des	Pont/Liaison entre les	Lancer la dynamique

	communautés	systèmes d'un écosystème	d'usage des données ouvertes
Normes	Agir <i>a minima</i> à l'échelle nationale	Désigner la mutualisation des ressources comme partie intégrante de projets concrets	Permettre à des néophytes de l'open data de mener une démarche de bout en bout en bénéficiant d'un accompagnement frugal sur l'ensemble de la démarche
Facilités	Moyens et autorités d'un réseau d'acteur national	Notoriété d'un "tiers de confiance" pour mutualiser les connaissances	Sceau ou tampon officiel d'une entité reconnue affiché sur la "forge"

Annexe 12 : Liste des fonctions de l'Open Data Canvas

Étapes de l'usage des données	Fonction	Description de la fonction
Toutes les étapes	1) Création d'un chambre de conception dédiée à un territoire/une organisation	Les utilisateurs peuvent créer une chambre de conception dédiée à leur territoire dans laquelle les projets de développement des usages peuvent être débattus
	2) Création d'espaces de discussions dédiés à chacun des projets	Les utilisateurs peuvent créer des espaces dédiés à un projet qui accueillera quatre fils de discussion correspondant aux 4 étapes du processus de sparring
	3) Description de l'usage imaginé	Les utilisateurs peuvent librement décrire des usages qu'il souhaite mettre en place et des difficultés qu'ils rencontrent
	4) Dialogue en ligne	L'utilisateur-intervenant questionne l'acteur sur la base de template de questionnement linéaire qu'il adapte. L'utilisateur répond aux questions
	5) Transfert automatique des réponses	Les réponses de l'utilisateur sont transférées dans une base de données relationnelles
	6) Création de questionnaires	Les utilisateurs peuvent créer des questionnaires qui vont venir peupler la base de données relationnelles dans le but d'obtenir des feedback de la part

		des acteurs des autres systèmes
	7) Création de vues de la base de données relationnelles	Les utilisateurs vont pouvoir prendre acte des feedback en créant de nouvelles vues de la base de données pour montrer les évolutions du plan d'action suite aux retours
Utilisation des fonctionnalités de la plateformes (téléchargement/flux de données)	8) Création de chambre de discussion pour chaque plateforme open data française	Les utilisateurs-concepteurs peuvent créer des chambres de discussion pour accueillir toutes les demandes d'informations relatives aux données présentes sur les plateformes open data françaises
	9) Demande de précisions sur les données des plateformes	Les utilisateurs peuvent poser leurs questions sur des données spécifiques présentes sur les plateformes open data
	10) Mettre à jour la base de données de gestion des composants du réseau en collaboration avec d'autres membres	Les utilisateurs peuvent modifier à plusieurs la base de données de gestion des composants.
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	11) Proposition de nouvelles connaissances pour la liste des composants	Les utilisateurs peuvent proposer de nouvelles connaissances comme candidates à la liste des composants du réseau d'information
	12) Évaluation du respect des normes techniques des connaissances	Les utilisateurs-concepteurs peuvent évaluer les connaissances proposées. Ces évaluations sont

		transparentes et disponibles à l'ensemble du réseau
Transformation des données grâce à des artefacts techniques	13) Post de questions spécifiques à l'accomplissement des tâches spécifiques à l'usage	Les utilisateurs peuvent poser des questions pratiques quant à la réalisation de certaines tâches relatives à l'usage de la donnée et faire appel aux membres du réseau
	14) Evaluation du respect des normes sociales	L'utilisateur-administrateur surveillent le réseau et s'assure du bon respect des règles sociales.
Utiliser les fonctionnalités de la plateforme open data	15) Télécharger le catalogue des données ouvertes d'une plateforme	L'utilisateur peut télécharger le catalogue de données ouvertes dans le système de gestion des tâches
	16) Documentation des données du catalogue	L'utilisateur peut renseigner les informations qu'il considère utile sur les données du catalogue
Evaluation du potentiel des données à partir de leur représentation	17) Télécharger les jeux de données à utiliser	L'utilisateur peut télécharger les données ouvertes dont il souhaite faire l'usage sur le système de gestion des tâches
	18) Visualiser les données ouvertes	L'utilisateur peut visualiser les données pour évaluer son potentiel
	19) Réorganisation des tables listant les tâches relatives à l'usage des données ouvertes	L'utilisateur peut réorganiser les tâches à réaliser en fonction des critères de son choix et ainsi créer un plan de développement de l'usage.

	20) Visualiser la séquence de tâches à réaliser sur les données	L'utilisateur peut visualiser de différentes manière l'ordre des tâches à effectuer sur le jeu de données
	21) Mise à disposition de témoignages de réalisation de tâches en situation réelle	L'utilisateur peut accéder à des témoignages concernant la mise en oeuvre réelle d'une tâche
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	22) Adaptation du modèle de méta-données	L'utilisateur-concepteur peut modifier le modèle de métadonnées qui structure le système de tâches en fonction d'ajouts de nouvelles pratiques dans le codebook ou de composants dans le réseau d'information
Transformation des données grâce à des artefacts techniques	21) Mise à disposition de témoignages de réalisation de tâches en situation réelle	L'utilisateur peut accéder à des témoignages concernant la mise en oeuvre réelle d'une tâche
	17) Télécharger les jeux de données à utiliser ou la liste de ces jeux	L'utilisateur peut télécharger les données ouvertes dont il souhaite faire l'usage sur le système de gestion des tâches
	24) Associer à la liste de données à utiliser une séquence de tâches	L'utilisateur peut relier les données qu'il compte utiliser à un ensemble de tâche qu'il effectue sur ces données
Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	25) Proposition de nouvelles pratiques dans le codebook	Les utilisateurs peuvent proposer l'ajout de nouvelles pratiques au sein du codebook
	26) Evaluation du respect de l'esprit du manifeste par les	Les utilisateurs-concepteurs évaluent si les nouvelles

	nouvelles pratiques proposées	pratiques proposées correspondent à l'esprit du manifeste et communique leur choix d'intégration ou de rejet des nouvelles pratiques
	27) Déclinaison du manifeste dans un domaine de données spécifiques	Les utilisateurs peuvent proposer des déclinaisons du manifeste pour l'adapter à des domaines de données spécifiques
Institutionnalisation de l'usage des données ouvertes	28) Propositions de service et produits compatibles avec le codebook	Les utilisateurs se voient proposer et redirigés vers les services ou produits dans lequel ils peuvent injecter le codebook

Annexe 13 : Interfaces du site de l'Open Data Canvas



The screenshot shows the homepage of the Open Data Canvas website. The header includes navigation links for "Le Manifeste", "Les composants", and "Contribuer". The main content area features a large graphic of a grid of colored triangles (red, yellow, blue, white) forming the letters "OPEN DATA CANVAS". To the right of the graphic, the text reads "UN CANVAS POUR DES DONNEES COUSUES MAIN". Below this, there are three paragraphs of text explaining the project's goals and the challenges of open data. At the bottom, there is a button labeled "Le manifeste".

UN CANVAS POUR DES DONNEES COUSUES MAIN

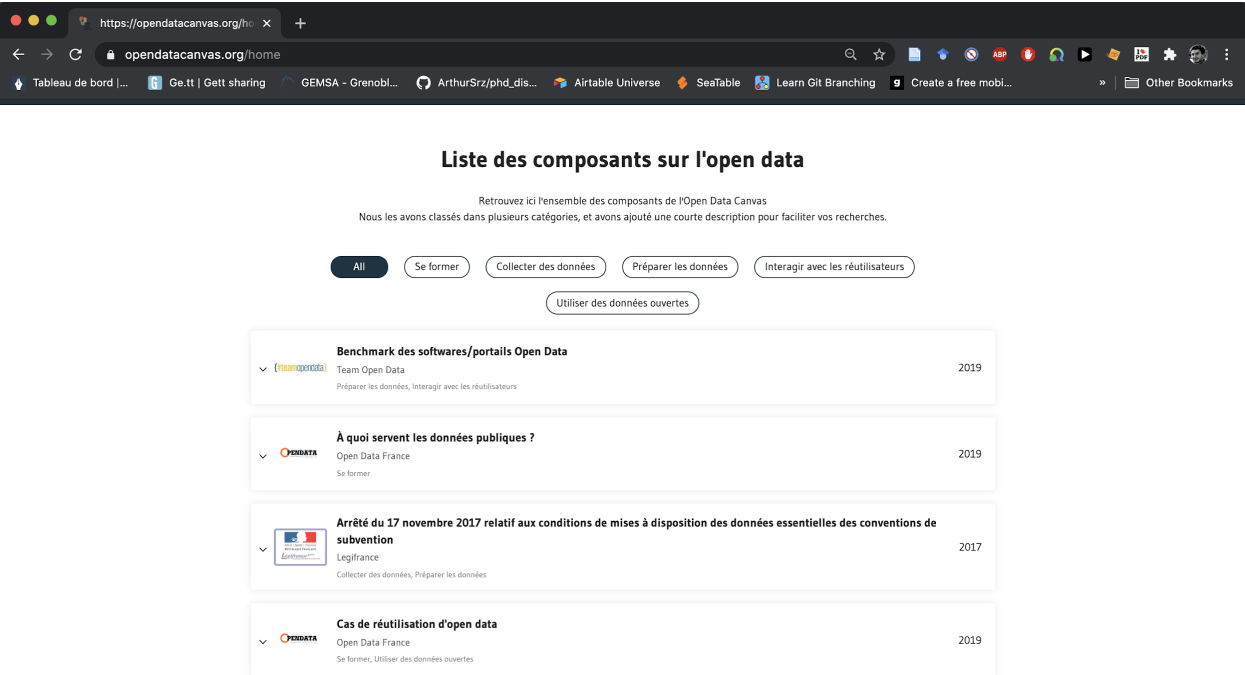
Par où commencer un projet open data ?
Et comment lui donner le plus de chances d'avoir de l'impact ?

La Loi Pour une République Numérique oblige les administrations de plus de 50 agents et collectivités de plus de 3500 habitants à ouvrir leurs données dans des formats réutilisables. Cela pose un défi de taille à toutes les institutions qui n'ont pas encore initié une stratégie. Et pour celles et ceux qui ont la responsabilité de conduire un projet open data, il peut être difficile de s'y retrouver parmi la jungle de documents, méthodes, articles de presse, textes législatifs sur le sujet.

C'est tout l'objet de l'Open Data Canvas, ici en version alpha, qui propose une liste de ressources appelées composants (françaises et internationales) sur l'open data : de la collecte au traitement des données, et de leur publication à leur réutilisation.

L'Open Data Canvas, c'est aussi un manifeste de l'open data, qui explique les principes essentiels pour qu'une stratégie open data crée de l'impact.

[Le manifeste](#)



The screenshot shows the "Liste des composants sur l'open data" page. The page title is "Liste des composants sur l'open data". Below the title, there is a sub-header "Retrouvez ici l'ensemble des composants de l'Open Data Canvas" and a description: "Nous les avons classés dans plusieurs catégories, et avons ajouté une courte description pour faciliter vos recherches." There are five filter buttons: "All", "Se former", "Collecter des données", "Préparer les données", and "Interagir avec les réutilisateurs". Below the filters, there is a list of components, each with a dropdown arrow, a title, a description, and a year.

Liste des composants sur l'open data

Retrouvez ici l'ensemble des composants de l'Open Data Canvas
Nous les avons classés dans plusieurs catégories, et avons ajouté une courte description pour faciliter vos recherches.

[All](#) [Se former](#) [Collecter des données](#) [Préparer les données](#) [Interagir avec les réutilisateurs](#)

[Utiliser des données ouvertes](#)

- Benchmark des logiciels/portails Open Data**
Team Open Data 2019
Préparer les données, Interagir avec les réutilisateurs
- À quoi servent les données publiques ?**
Open Data France 2019
Se former
- Arrêté du 17 novembre 2017 relatif aux conditions de mises à disposition des données essentielles des conventions de subvention**
Legifrance 2017
Collecter des données, Préparer les données
- Cas de réutilisation d'open data**
Open Data France 2019
Se former, Utiliser des données ouvertes

Annexe 14 : Conducteur des focus group

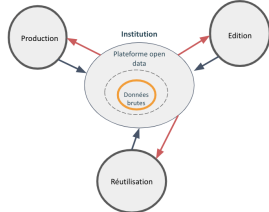


Focus group #1

Points abordés au cours de ce focus group

- Présentation du projet de thèse
- Découpage de votre projet de construction des usages sur votre territoire en différentes **tâches + barrières** rencontrées
- **Classement** des barrières par ordre d'importance
- Démonstration du Canvas et de ses **fonctionnalités**
- Questions/Réponses sur l'**utilité perçue** du Canvas vis à vis des barrières

Présentation du projet de thèse : un modèle pour comprendre la **construction des usages**



Tâches et barrières rencontrées au cours de votre projet

Exemple de découpage

#	Type de tâche	Tâche	Exemple
Contexte de lancement de la démarche :			
1	Open - behavioral and observable & Covert - mental (cognitive) and not observable (Fact)	Designier un responsable de la démarche open data	
2	Covert - mental (cognitive) and not observable (Fact)		C'est quoi l'open data et pourquoi ? - donner de sens à la démarche - Une définition avec des exemples de démarches - Des exemples de use-case de l'open data
3	Covert - mental (cognitive) and not observable (Fact)		Se poser la question du quel libérer et des premières briques à mettre en place ? - pour se lancer et capter la démarche ! - Savoir quelles données il faut avoir - Quelles licences il faut choisir - Stratégie de diffusion - créer une plateforme ?

+ barrières rencontrées

Tâches et barrières rencontrées au cours de votre projet

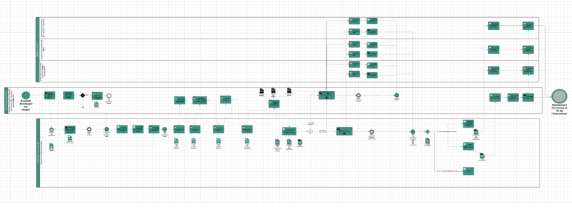
Exemple de barrières rencontrées

Cas de Grenoble et Lyon
Absence de garantie sur la caractère brut des données
Manque d'informations sur la production des données métadonnées
Doute sur les restrictions d'usage apposées aux données par les licences (restriction économique, conditions de réutilisation)
Plateforme perçue comme des vitrines et non comme un point d'accès aux données

Classement des barrières par ordre d'importance



Top Barrières	
1	
2	
3	
4	
5	

Démonstration du Canvas et de ses **fonctionnalités**



Utilité perçue : lesquelles de vos barrières le Canvas peut-il lever ?

Top Barrières	
1	
2	
3	
4	
5	

+  =  ?

Annexe 15 : Codebook des unités de sens utilisés sous NVivo (chapitre 8)

Name	Description
Création d'un cadre cognitif	cf Chapitre 3
Manque d'un cadre juridique et politique clair	cf Chapitre 3
Justification de l'utilité/inutilité	cf Chapitre 2
Manque d'une vision cohérente	cf Chapitre 3
Justification de l'utilité/inutilité	
Evaluation potentiel des données	cf Chapitre 3
Difficultés intégration	cf Chapitre 3
Justification de l'utilité/inutilité	
Manque expertise technique	cf Chapitre 3
Justification de l'utilité/inutilité	
Institutionnalisation	cf Chapitre 3
Divergence culturelle	cf Chapitre 3
Justification de l'utilité/inutilité	
Incompatibilité des structures organisationnelles	cf Chapitre 3
Justification de l'utilité/inutilité	
Transformation des données	cf Chapitre 3
Hétérogénéité des conditions d'utilisation	cf Chapitre 3
Justification de l'utilité/inutilité	
Mauvaise qualité	cf Chapitre 3
Justification de l'utilité/inutilité	
Utilisation des fonctionnalités d'une plateforme	cf Chapitre 3
Manque de confiance	cf Chapitre 3
Justification de l'utilité/inutilité	
Manque de lisibilité	cf Chapitre 3
Justification de l'utilité/inutilité	

Annexe 16 : Extrait du tableau de synthèse des résultats des focus group

Besoin Fonctionnel #Focus group	Utilité du Canvas ?			Preuves			
	3	2	1	5	4	3	2
1-BF1-Documentation production	NA	Non	NA	NA	NA	NA	<p>Outil déjà disponible et fonctionnel pour la production et la documentation des données :</p> <p>"L'idée c'est de dire : vous avez du Excel, maintenant vous avez ça en complément ou en substitution, avec de l'import de la saisie, les contrôles de cohérence de champs sont pris en compte dans le formulaire, y a un workflow de validation, de notification, un truc assez évolué. C'est aujourd'hui le seul accompagnement qu'on a à faire dans une collectivité, c'est de dire "vous devez ouvrir les données de subvention", donc on vous met à disposition un outil, à vous de voir en fonction de votre organisation, c'est de l'import, interfaçage avec SI, et tout de suite, ça leur paraît abordable. On leur demande pas d'ingénierie, on leur apporte des solutions clés en main.</p> <p>Ensuite, les données agrégées sont automatiquement agrégées à l'échelle du département, et ensuite elles sont redescendues sur les outils de communication de nos communes."</p>
1-BF2-Documentation mécanisme	NA	<i>idem</i>	NA	NA	NA	NA	<i>idem</i>
1-BF3-Design plateforme double entrée	NA	NA	NA	NA	Canvas répondrait à un autre besoin : "Le Canevas(...) faciliterait l'uniformisation des plateformes"	NA	NA
2-BF4-Sensibilisation	Non. Une plateforme open data suffit	Oui.	Oui.	<p>Preuve 1 - "Là, je me dis que si on pouvait utiliser le Canvas pour chaque projet dans lequel il y a de la donnée ça serait sympa. A priori, je verrai l'utilité pour les projets qui viennent de démarrer et dans lesquels on peut coupler l'open data. Ça peut m'aider par exemple, en mettant tous les projets sur lesquels j'ai été sollicité, et pour me questionner sur comment on peut progressivement trouver des données, les utiliser avec pour but qu'elle soit utile à la fin au projet. Et ça permettrait en plus de faciliter l'acculturation des acteurs, par l'action et l'intégration dans tous les projets d'open data. C'est une façon efficace d'acculturer."</p> <p>Preuve 2-</p>	NA	<p>"Mais si vous leur montrez la plateforme, vous pouvez leur dire : "l'open data, c'est ça", voilà ce que ça peut donner. Et puis en plus de ça, la plateforme montre à voir le boulot des agents, et donc vous allez dire aux agents que tout leur boulot va être transparent et que les citoyens vont savoir ce qu'il s'est passé. Et l'agent va aussi changer la perception de son boulot puisque le service public va être présent dès la donnée, puisqu'il va la produire pour un autre. Et finalement, ça aura un vrai impact sur l'organisation."</p>	<p>Preuve 1 - D'après un participant, le Canvas permettrait de sensibiliser les acteurs sur la portée stratégique de l'open data : "Il faut tout replacer dans le cadre d'une stratégie globale qui explique le sens des actions sinon les acteurs se perdent. En ce sens, le fait que vous fixiez le questionnement des objectifs au début du Canvas et que ce soit le fil rouge, c'est une très bonne idée. "</p> <p>Preuve 2 - "Une idée intéressante. Il est certain que si on aborde l'open data en commençant par les questions de comprendre ce qu'on veut faire, si on commence par les enjeux data/IA/création de services, ça va être plus facile partout, et notamment pour ouvrir. On va pouvoir aller voir les personnes/élus qui détiennent les données et leur dire qu'on a besoin de leurs données pour telle ou telle raison."</p>

Annexe 17 : Exemple d'interaction au sein de la communauté transport.data.gouv.fr développant les usages des données ouvertes de transport

#general ☆
http://transport.beta.gouv.fr/ 288

December 2nd, 2019

Victor Alexandre (Le Parisien) 11:19 AM
Bonjour à tous,
Je suis à la recherche d'un jeu de données géographique avec l'ensemble des entrées et sorties des stations de métro et de RER (sortie 1, sortie 2, etc..)

tristram 11:25 AM
Bonjour,
Elles sont disponibles dans le fichier d'horaire de idfm (<https://transport.data.gouv.fr/datasets/horaires-prevus-sur-les-lignes-de-transport-en-commun-dile-de-france-gtfs/>)
Dans le zip, prendre le fichier stops.txt (à ouvrir dans un tableur) et c'est ceux qui ont location_type=2 (Entrance/exit) :
<https://gtfs.org/reference/static#stopstxt>

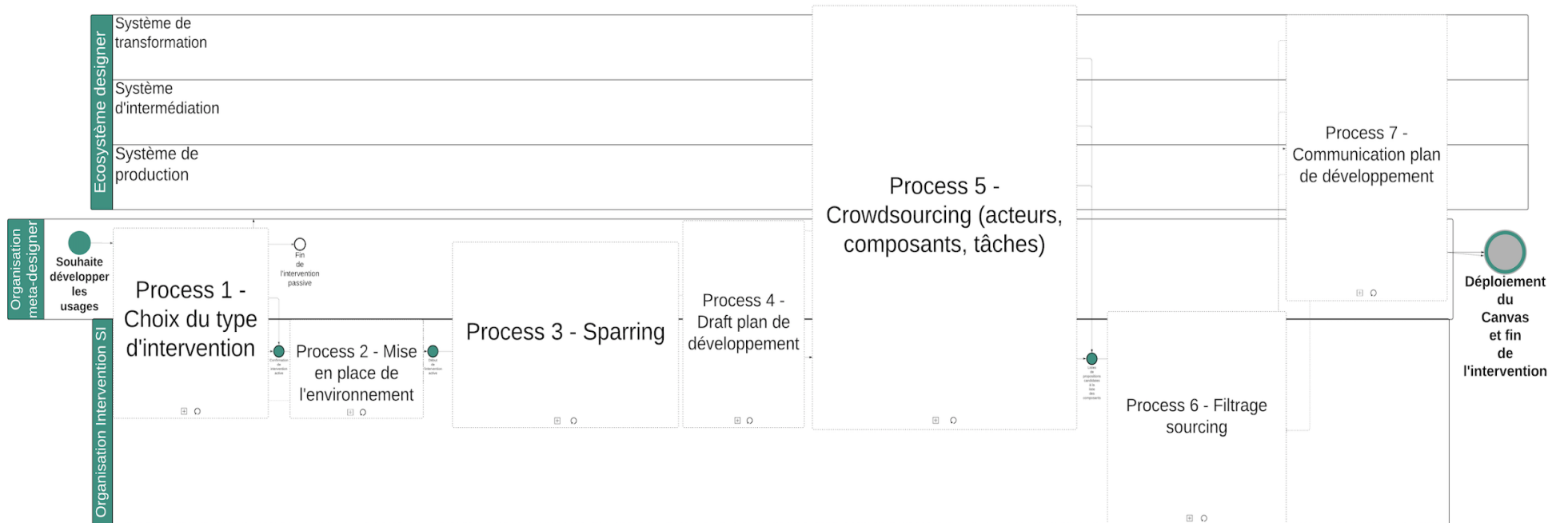
transport.data.gouv.fr
transport.data.gouv.fr
Rendre disponible, valoriser et améliorer les données transports

gtfs.org
General Transit Feed Specification
Sample

1

Sinon, il est aussi possible de les récupérer avec une requête sparql:

Annexe 18 : Cartographie de l'architecture métier



Annexe 19 : Extrait de la liste des documents utilisés en données secondaires lors de l'étude de cas (chapitre 4)

Les documents sont accessibles à l'adresse suivante : <https://bit.ly/3elwcLY>

Métropole Grenoble-Alpes

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes d'octobre 2018

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 24 novembre 2018

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 6 décembre 2018

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 20 décembre 2018

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 10 janvier 2019

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 24 janvier 2019

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 21 mars 2019

Compte Rendu de la réunion du Groupe de Travail Open Data de la métropole Grenoble-Alpes du 21 novembre 2021

Délibération du conseil municipal de la Ville de Grenoble sur la politique Open Data

1er scénario d'usage des données ouverte du portail data.metrogrenoble.fr par le projet "Ma Petite Poubelle"

2ème scénario d'usage des données ouverte du portail data.metrogrenoble.fr par le projet "Ma Petite Poubelle"

3ème scénario d'usage des données ouverte du portail data.metrogrenoble.fr par le projet "Ma Petite Poubelle"

Descriptif des projets lauréats du "Grenoble Civic Lab", concours de services réutilisant les données ouvertes sur le territoire grenoblois

Introduction au design de services publics à partir des données ouvertes dans le cadre de l'Atelier Design de services (SciencesPo Grenoble et IAE Grenoble)

Présentation de la démarche open data grenobloise auprès des étudiants participant à l'Atelier Design de services (SciencesPo Grenoble et IAE Grenoble)

Métropole du Grand Lyon

Document de référence pour la structuration des systèmes d'information du Grand Lyon : Cadre Commun d'Urbanisation des Systèmes d'Information de l'Etat

Enquête sur le quotidien de l'ouverture des données au sein de la métropole du Grand Lyon réalisée par Antoine Courmont (2016)

Délibération du conseil métropolitain du Grand Lyon sur la mise à disposition d'un service d'accompagnement à l'ouverture à destination des communes du territoire.

Compte Rendu de la Journée Thématix, rendez-vous des membres de l'écosystème open data lyonnais

Enquête sur les différents modes de réutilisation des données lyonnaises (Larroche and Vila, 2018)

Extraits des feedbacks des réutilisateurs sur les nouvelles fonctionnalités du portail open data du Grand Lyon (forum #TeamOpenData)

Synthèse des résultats d'une enquête menée auprès des réutilisateurs des données lyonnaises (Paquienréguy et al. 2016)

Support de formation des agents du territoire du Grand Lyon à l'open data utilisé par le CNFPT

Liste des freins à l'open data identifiés par la laboratoire d'innovation lyonnais TUBA

Liste des arguments utilisés pour justifier les freins à l'ouverture (TUBA Lyon)

Liste des figures

Figure 0.2 : Structure de la thèse	28
Figure 0.3 : Structure du chapitre 1	29
Figure 0.4 : Structure du chapitre 3	30
Figure 0.5 : Structure du chapitre 4	31
Figure 0.6 : Structure du chapitre 5	32
Figure 0.7 : Structure du chapitre 6	33
Figure 0.8 : Structure du chapitre 8	34
Figure 1.1 : Exemple d'analyse morphologique	37
Figure 1.2 : Les relations entre les termes clés de la revue de littérature	38
Figure 1.3 : Logigramme établissant le lien entre l'offre de données (conditions cadres), la demande de données et les résultats économiques de l'ouverture des données (impact)	41
Figure 1.4 : Visualisation des différents niveaux d'analyse de l'usage des données du référentiel géographique (Courmont, 2016, p.81)	50
Figure 1.5 : Les éléments de l'écosystème d'une plateforme selon Eisenmann et al. (2008, p.164)	58
Figure 1.6 : La typologie des parties prenantes impliquées dans les plateformes open data	63
Figure 1.7 : Chaîne de valeur d'un écosystème open data d'une ville intelligente	85
Figure 1.8 : Le système d'activité open data côté utilisateur (d'après Engeström, 2001, p.140)	87
Figure 1.9 : Modèle des écosystèmes selon la théorie de l'activité (d'après Ruijter et al., 2017, p.48)	88
Figure 1.10 : Analyse morphologique de la littérature (pour parcourir le tableau : lien)	90
Figure 1.11 : Phases d'analyse des données et des traitements de la méthode Merise (Servigne, 2008,p.3)	102
Figure 1.12 : Exemple d'un modèle un V de conception d'un système d'information (Firesmith, 2014, p.14)	104
Figure 1.13 : Exemple d'un modèle en W de conception d'un système d'information (Petit, 2006, p.155)	105
Figure 1.14 : Phases de conception selon la méthode RAD (adapté de Reix et al., 2009, p.85)	106
Figure 1.15 : Tableau récapitulatif des facteurs expliquant le déclin de pertinence des recherches en systèmes d'information (Benbasat et Zmud, 1999, p.10)	110
Figure 1.16 : La boucle de rigueur dans les Méthodologies de recherche en <i>design science</i> (Hevner et al., 2004, p.80)	115
Figure 1.17 : Représentation des similitudes entre les composants de la nouvelle démarche scientifique évoquée par Pascal (2006, p.241) et ceux de la design theory (Gregor et Jones, 2007, p.322)	125
Figure 1.18 : Les relations entre la design theory et les artefact	126
Figure 1.19 : Processus nominal comportant les différentes activités	136
Figure 1.20 : Lignes directrices en design science (Hevner et al., 2004, p.83)	138
Figure 1.21 : Le réseau nomologique des artefact technologiques (Benbasat et Zmud, 2003, p.130)	142
Figure 1.22 : Architecture de recherche des recherches de Zuiderwijk (2015, p.29)	149
Figure 2.1 : Présentation des choix qui influencent l'architecture de recherche	156
Figure 2.2 : Intervention en Systèmes d'Information (Carlsson, 2010, p.220)	159
Figure 2.3 : L'architecture de la recherche	162
Figure 2.4 : Exemple de structure statique d'un système d'objets intelligents (Vaishnavi et al., 1997, p.277)	166
Figure 2.5 : Exemple de structure statique d'un objet intelligent (Vaishnavi et al., 1997, p.281)	166
Figure 2.6 : Représentation d'un processus selon Datta (1998, p.279)	167
Figure 2.7 : Exemple de comportement dynamique d'un système d'objets intelligents (Vaishnavi et al., 1997, p.277)	167
Figure 2.8 : La représentation d'un écosystème open data (d'après Harrison et al., 2012, p.908)	170
Figure 2.9 : Catégorisation récursive des articles par groupes d'éléments fonctionnels	180
Figure 2.10: Mots clés choisis dans les différentes revues de littérature	182
Figure 2.11 : Design du système de l'environnement de conception des usages et des trois éléments fonctionnels	185
Figure 2.12 : Pattern de coordination de l'environnement de conception des usages.	187

Figure 2.13. Tableau de correspondance entre les objectifs et les types de prototypage.	190
Figure 3.1 : Place de la plateforme open data “data.metro.fr”	209
Figure 3.2 : Fonctionnement de l'écosystème open data autour de l'Annuaire	213
Figure 3.3 : Représentation graphique de la structure statique des systèmes à l'oeuvre	216
Figure 3.4 : Représentation graphique du comportement dynamique au sein des systèmes	218
Figure 3.5 : Évolution de l'écosystème open data	220
Figure 3.6 : Illustration de la première séquence d'articulation du cas d'usage “Quartier”	228
Figure 3.7 : La représentation de la première séquence d'articulation du cas d'usage “Transport”	238
Figure 3.8 : La représentation de la première séquence d'articulation du cas d'usage “Logement”	247
Figure 3.9 : La représentation de la première séquence d'articulation du cas d'usage “Collaboration”	258
Figure 3.10 : Un modèle conceptuel affiné de l'usage de l'open data	284
Figure 4.1 : Illustration des frontières et unités d'analyse choisie pour le cas de grenoble	291
Figure 4.2 : Illustration des frontières et unités d'analyse choisie pour le cas du Grand Lyon	294
Figure 4.3 : La liste des sources de données	300
Figure 4.4 : La liste des entretiens semi-directifs conduits	302
Figure 4.5 : Les besoins fonctionnels relatifs à l'utilisation des fonctionnalités d'une plateforme open data.	322
Figure 4.6 : Les besoins fonctionnels relatifs à l'évaluation du potentiel des données à partir de leur représentation	324
Figure 4.7 : Les besoins fonctionnels relatifs à la transformation des données grâce à des artefacts techniques.	326
Figure 4.8 : Les besoins fonctionnels relatifs à la création d'un cadre cognitif partagé entre les systèmes	329
Figure 4.9 : Les besoins fonctionnels relatifs à l'institutionnalisation de l'usage des données	332
Figure 5.1 : Le rôle des propositions dans le design du dispositif	337
Figure 5.2 : Le rôle des mécanismes, principes et fonctions dans le design du dispositif	338
Figure 5.3 : La place du dispositif et de ses différentes aires fonctionnelles dans le modèle théorique de l'usage des données ouvertes	372
Figure 5.4 : Design des deux mécanismes principaux définissant le comportement du système de l'artefact de meta-design.	375
Figure 5.5 : Design des 3 mécanismes principaux définissant le comportement du système du réseau d'information	380
Figure 5.6 : Design des 3 mécanismes principaux définissant le comportement du système de modèle de méta-données	386
Figure 5.7 : Design des 2 mécanismes principaux définissant le comportement du manifeste et du codebook	390
Figure 5.8 : Description des fonctions du l'environnement de conception des usages pris dans son ensemble	403
Figure 5.9 : Description des fonctions du réseau d'information	405
Figure 5.10 : Description des fonctions du système de gestion des tâches	406
Figure 5.11 : Description des fonctions du manifeste et du codebook	407
Figure 5.12 : Les fonctions relatives à l'ensemble de l'environnement	409
Figure 5.13 : Les fonctions relatives au réseau d'information	409
Figure 5.14 : Les fonctions relatives à un système de gestion des tâches basée sur un modèle open data	410
Figure 5.15 : Les fonctions relatives au manifeste et son codebook	411
Figure 6.1 : Cartographie de l'architecture métier de la version 0 du prototype de l'Open Data Canvas	424
Figure 6.2 : Cartographie de l'architecture fonctionnelle de la version 0 du prototype de l'Open Data Canvas	426
Figure 6.3 : Cartographie de l'architecture applicative de la version 0 du prototype de l'Open Data Canvas	429
Figure 6.4 : Interface d'accueil de la chambre de conception de MetroMeta.	432

Figure 6.5 : Interface de l'espace de discussion du projet Eco-quartier contenant les quatre fils de discussion du processus de sparring	433
Figure 6.6 : Interface de l'espace de travail (workspace) de l'Intervenant et de la base lui permettant d'animer le sparring avec MetroMeta	434
Figure 6.7 : Interface de la table d'animation de l'étape 1 du processus de sparring avec zoom sur le template de question linéaire	435
Figure 6.8 : Interface du fil de discussion de l'étape 1 du processus de sparring après l'ajout d'une nouvelle ligne dans la table d'animation	436
Figure 6.9 : Interface de la table d'animation de l'étape 2 du processus de sparring avec zoom sur le template de questions circulaires	437
Figure 6.10 : Interface de la table d'animation de l'étape 3 du processus de sparring avec zoom sur le template de questions réflexives	438
Figure 6.11 : Interface de la table d'animation de l'étape 4 du processus de sparring avec zoom sur le template de questions stratégiques	439
Figure 6.12 : Interface de l'espace de travail (workspace) de MetroMeta et de la base lui permettant de construire un plan de développement des usages	440
Figure 6.13 : Interface de la base Open Data Work de MetroMeta avec un focus (en rouge) sur les trois premières tables	441
Figure 6.14 : Interface de la Table d'association	442
Figure 6.15 : Interface présentant le fonctionnement de la colonne liant la Table d'association avec la Table "Donnees_ouvertes"	443
Figure 6.16 : Interface présentant l'espace de travail de l'Intervenant avec les bases des référentiels (en noir) de composants et de tâches entretenus par la communauté #teamopendata	445
Figure 6.17 : Interface présentant la chambre dédiée à l'OpenDataCanvas et l'espace dédié au sourcing de composants et de tâches.	445
Figure 6.18 : Interface présentant l'espace de dédié au sourcing de composants et de tâches ainsi que le fil de discussion dédié au sourcing de composants relatifs aux "arguments anti open data"	446
Figure 6.19 : Interface du questionnaire de dépôt des propositions de nouveaux composants	447
Figure 6.20 : Interface de la base permettant le filtrage des composants et tâches sourcés dans la communauté #teamopendata	448
Figure 6.21 : Interface montrant un exemple d'appel à connaissances	449
Figure 6.22 : Interface montrant le tableau Kanban du plan de développement des usages.	450
Figure 6.23 : Interface montrant le calendrier du plan de développement des usages.	451
Figure 6.24 : Interface montrant l'espace de travail réservée aux acteurs du projet (workspace design) et la base permettant de mettre en oeuvre le plan de développement des usages	452
Figure 6.25 : Interface montrant le tableau Kanban du laboratoire d'innovation avec les transformations qu'il doit réaliser.	453
Figure 6.26 : Interface montrant les fils de discussions dédiées aux transformations à effectuer par les acteurs	454
Figure 6.27 : Interface montrant les échanges relatifs au tag "données-énergie"	455
Figure 6.28 : Interface de la base OpenDataWork avec un focus sur la table d'association propre à la deuxième itération	456
Figure 7.1 : Synthèse de l'analyse de l'évaluation	462
Figure 7.2 : Framework de sélection de la stratégie d'évaluation de la version 0 du prototype	465
Figure 7.3 : Framework de sélection de la stratégie d'évaluation de la version alpha du prototype	467

Liste des tables

Table 0.1 : Aperçu de l'ensemble des chapitres de la thèse	26
Table 1.1 : Les conditions et impact économique des données ouvertes	43
Table 1.2 : Les barrières face à l'utilisation des jeux de données ouvertes	78
Table 2.1 : La grille d'observation des cas exploratoires	172
Table 2.2 : Framework de Sélection des Stratégies d'Évaluation	195
Table 2.3 : Framework de Sélection des Méthodes d'Évaluation	196
Table 3.1 : Présentation de l'écosystème du cas "Logement"	224
Table 3.2 : Présentation des artefacts du cas "Logement"	226
Table 3.3 : La capacité d'action des artefacts du cas "Logement"	227
Table 3.4 : Présentation de l'écosystème du cas "Transport"	233
Table 3.5 : Présentation des artefacts du cas d'usage "Transport".	235
Table 3.6 : Les capacités d'action des artefacts du cas d'usage "Transport"	237
Table 3.7 : L'écosystème du cas d'usage "Logement"	243
Table 3.8 : Les artefacts du cas d'usage "Logement"	245
Table 3.9 : Les capacités d'action des artefacts du cas d'usage "Logement"	246
Table 3.10 : L'écosystème du cas d'usage "Collaboration"	252
Table 3.11 : Les artefacts du cas d'usage "Collaboration"	256
Table 3.12 : Les capacités d'action du cas d'usage "Collaboration"	257
Table 3.13 : Les composants des plateformes open data	262
Table 3.14 : Les parties prenantes des plateformes open data	263
Table 3.15 : Les outils structurés par les données ouvertes	264
Table 3.16 : Les éléments des écosystèmes open data	265
Table 3.17 : Les artefacts sociaux et techniques impliqués dans l'usage de l'open data	266
Table 3.18 : Les capacités d'action des artefacts et leur influence sur l'état des jeux de données ouverts	267
Table 3.19 : Les différentes configurations des jeux de données ouverts	268
Table 3.20 : Les séquences d'articulation identifiables dans l'usage de l'open data	268
Table 3.21 : Les complémentarités formées lors des séquences d'articulation	269
Table 3.22 : L'échantillon de parties prenantes	274
Table 3.23 : Confrontation des éléments empiriques et des verbatims des entretiens	276
Table 3.24 : Extrait du tableau récapitulatif des barrières socio-techniques freinant l'usage des données	282
Table 3.25 : Framework d'identification des barrières face à l'usage de l'open data	283
Table 4.1 : Thèmes abordés avec les personnes interrogées	298
Table 4.2 : Barrière freinant l'utilisation des fonctionnalités de la plateforme open data	306
Table 4.3 : Barrière freinant l'évaluation du potentiel des données	310
Table 4.4 : Barrière freinant la transformation des données grâce à des artefacts techniques	312
Table 4.5 : Barrière freinant la création d'un cadre cognitif partagé entre les systèmes	315
Table 4.6 : Barrière freinant l'institutionnalisation de l'usage des données	318
Table 6.1 : Table de correspondance entre les apports souhaités du prototype et le type de prototypage à effectuer (adapté d'Ince et Hekmatpour, 1987, p.12)	415
Table 6.2 : Intégration et amélioration des fonctions dans les différentes versions du prototype	417
Table 6.3 : Fonctions relatives à l'ensemble de l'environnement de conception sélectionnées pour intégrer la Version 0	418
Table 6.4 : Fonctions relatives au réseau d'information sélectionnées pour intégrer la Version 0	419

Table 6.5 : Fonctions relatives au système de gestion des tâches sélectionnées pour intégrer la Version 0	420
Table 7.1 : Caractéristiques des focus group et des participants	470
Table 7.2 : Extrait de la table faisant la synthèse du codage des transcriptions des focus group	472
Table 7.3 : Tableaux sur l'utilité de l'Open Canvas sur le premier besoin fonctionnel	480
Table 7.4 : Tableaux sur l'utilité de l'Open Canvas sur le troisième besoin fonctionnel	481
Table 7.5 : Tableaux sur l'utilité de l'Open Canvas sur le quatrième besoin fonctionnel	483
Table 7.6 : Tableaux sur l'utilité de l'Open Canvas sur le cinquième besoin fonctionnel	484
Table 7.7 : Tableaux sur l'utilité de l'Open Canvas sur le sixième besoin fonctionnel	485
Table 7.8 : Tableaux sur l'utilité de l'Open Canvas sur le septième besoin fonctionnel	486
Table 7.9 : Tableaux sur l'utilité de l'Open Canvas sur le huitième besoin fonctionnel	488
Table 7.10 : Tableaux sur l'utilité de l'Open Canvas sur le neuvième besoin fonctionnel	489
Table 7.11 : Tableaux sur l'utilité de l'Open Canvas sur le dixième besoin fonctionnel	490
Table 7.12 : Tableaux sur l'utilité de l'Open Canvas sur le onzième besoin fonctionnel	491
Table 7.13 : Tableaux sur l'utilité de l'Open Canvas sur le douzième besoin fonctionnel	492
Table 7.14 : Utilité du Canvas relativement aux douze besoins fonctionnels	494
Table 8.1 : Framework d'identification des barrières face à l'usage de l'open data	502

Table des matières

Construire les usages des données ouvertes : une approche par le design science	1
Sommaire	10
Introduction générale	12
1. Avant-propos	13
1.1 L'ouverture des données : le nouvel or noir ?	13
1.2 L'insuffisante ouverture des données pour créer de nouveaux usages	13
2. Problématique de recherche	15
2.1 Cas d'usage des données ouvertes	15
2.2 Les écosystèmes open data comme espace-problème de la recherche	16
2.3 La question de recherche, un besoin à la fois théorique et pratique	17
3. Propositions de la recherche	21
3.1 Méthodologie adaptée	21
3.2 Plan de lecture de la thèse	23
Table 0.1 - Aperçu de l'ensemble des chapitres de la thèse	26
Figure 0.2 - Structure de la thèse	28
Figure 0.3 - Structure du chapitre 1	29
Figure 0.4 - Structure du chapitre 3	30
Figure 0.5 - Structure du chapitre 4	31
Figure 0.6 - Structure du chapitre 5	32
Figure 0.7 - Structure du chapitre 6	33
Figure 0.8 - Structure du chapitre 8	34
Chapitre 1	35
Revue de la littérature	35
<i>Quel est l'état des lieux concernant la construction des usages ? Comment les différents construits qui définissent la construction de nouveaux usages à partir des données ouvertes sont-ils mis en relation dans la littérature scientifique ?</i>	36
Figure 1.1 : Exemple d'analyse morphologique	37
Figure 1.2 : Les relations entre les termes clés de la revue de littérature	38
1.1 Usages des données ouvertes	39
1.1.1 Les usages des données ouvertes au niveau intra-organisationnel	39
1.1.1.1 Les difficultés d'usage relatives au contenu et au format des données	39
1.1.1.2 Les difficultés d'usage relatives aux relations et collaborations entre usagers	40
1.1.2 Les usages des données ouvertes au niveau organisationnel	41
Figure 1.3 : Logigramme établissant le lien entre l'offre de données (conditions cadres), la demande de données et les résultats économiques de l'ouverture des données (impact)	41
Table 1.1 : Les conditions et impact économique des données ouvertes	43

1.1.2.1 Les obstacles économiques à la construction de nouveaux usages issus des données ouvertes : la question du business model	43
1.1.2.2 Les obstacles structurels à la construction de nouveaux usages issus des données ouvertes	44
1.1.3 Les usages des données ouvertes au niveau extra-organisationnel	44
1.1.3.1 Sur les freins à la prise en main de l'usage par les communautés	45
1.1.3.2 Sur les moyens d'améliorer la prise en main de l'usage par les communautés	46
1.1.4 Les usages des données ouvertes au niveau inter-organisationnel	46
1.1.4.1 Les solutions sociales proposées pour favoriser la formation d'alliance, de réseau ou d'écosystème	46
1.1.4.2 Les solutions techniques proposées pour favoriser la formation d'alliance, de réseau ou d'écosystème	47
1.1.5 Les usages des données ouvertes au niveau industriel, territorial et sociétal	48
1.1.5.1 Les usages des données ouvertes pour améliorer la performance des Etats et des services public	49
1.1.5.2 Les usages des données ouvertes et le risque d'aggravation des déséquilibres sociétaux	49
1.1.6 Les usages des données ouvertes vus depuis la superposition de plusieurs niveaux d'analyse	50
Figure 1.4 : visualisation des différents niveaux d'analyse de l'usage des données du référentiel géographique	50
1.1.7 Les usages des données ouvertes vus depuis les tensions paradoxales entre plusieurs niveaux d'analyse	51
1.2 Plateformes open data	53
1.2.1 Les plateformes open data au niveau intra-organisationnel	53
1.2.1.1 Le travail de production des données ouvertes et d'alimentation des plateformes open data	54
1.2.1.2 Le travail de maintenance des données ouvertes sur les plateformes open data	55
1.2.1.3 Le travail de coordination autour des plateformes open data	56
1.2.1.4 Le travail de coopération autour des plateformes open data	56
1.2.2 Les plateformes open data au niveau organisationnel	57
1.2.2.1 L'architecture technique des plateformes open data	57
1.2.2.2 La gouvernance des plateformes open data	58
Figure 1.5 : Les éléments de l'écosystème d'une plateforme selon Eisenmann et al. (2008, p.164)	58
1.2.3 Les plateformes open data au niveau extra-organisationnel	60
1.2.3.1 La licence d'utilisation de la plateforme comme facteur différenciant des communautés rassemblées autour des plateformes open data	61
1.2.3.2 Les protocoles techniques comme facteur différenciant des communautés rassemblées autour des plateformes open data	62
1.2.4 Les plateformes open data au niveau inter-organisationnel	62
Figure 1.6 : La typologie des parties prenantes impliquées dans les plateformes open data (Lusch et Nambisan, 2015)	63
1.2.5 Les plateformes open data au niveau industriel, territorial et sociétal	64
1.2.6 Les plateformes open data vus depuis la superposition de plusieurs niveaux d'analyse	66
1.2.7 Les plateformes vus depuis les tensions paradoxales entre plusieurs niveaux d'analyse	67
1.3 L'ouverture des données	68
1.3.1 L'ouverture des données vue depuis le niveau intra-organisationnel	68
1.3.2 L'ouverture des données vue depuis le niveau organisationnel	69
1.3.3 L'ouverture des données vue depuis le niveau extra-organisationnel	70
1.3.3.1 Les services d'agrégation de données par les communautés	71
1.3.3.2 Les collaborations des communautés sur des projets publics	72

1.3.3.3 La production de données par les communautés	72
1.3.4 L'ouverture des données vue depuis le niveau inter-organisationnel	73
1.3.4.1 Les données ouverte comme possibilités de nouvelles alliances entre organisations	73
1.3.4.2 Des nouvelles collaborations comme fonction de l'interopérabilité des données ouvertes	74
1.3.5 L'ouverture des données vue depuis le niveau industriel, territorial et sociétal	75
1.3.5.1 Histoire du mouvement de l'ouverture des données	75
1.3.5.2 Le mouvement d'ouverture des données : un progrès sociétal aux multiples potentialités	76
1.3.6 L'ouverture des données vue depuis la superposition de plusieurs niveaux d'analyse	77
Table 1.2 : Les barrières face à l'utilisation des jeux de données ouvertes (Ruijter et al., 2017)	78
1.3.7 L'ouverture des données vue depuis les tensions paradoxales entre plusieurs niveaux d'analyse	78
1.4 Les écosystèmes open data	79
1.4.1 Les écosystèmes open data vus depuis le niveau intra-organisationnel	79
1.4.1.1 Le rôle des acteurs dans les écosystèmes open data	79
1.4.1.2 Des modèles de management pour les équipes d'animation des écosystèmes open data.	80
1.4.2 Les écosystèmes open data vus depuis le niveau organisationnel	80
1.4.2.1 L'influence d'une organisation sur un écosystème open data	80
1.4.2.2 L'influence d'un écosystème open data sur une organisation	81
1.4.3 Les écosystèmes open data vus depuis le niveau extra-organisationnel	82
1.4.4 Les écosystèmes open data vus depuis le niveau inter-organisationnel	83
1.4.4.1 Les différents types d'interactions entre organisations membres d'un écosystème open data	83
1.4.4.2 La modélisation des interactions inter-organisationnelles sous forme d'une chaîne de valeur	84
Figure 1.7 Chaîne de valeur d'un écosystème open data d'une ville intelligente	85
1.4.5 Les écosystèmes open data vus depuis le niveau industriel, territorial et sociétal	85
1.4.6 Les écosystèmes open data vus depuis la superposition de plusieurs niveaux d'analyse	86
Figure 1.8. Le système d'activité open data côté utilisateur (Engeström, 2001, p.140)	87
Figure 1.9 : Modèle des écosystèmes selon la théorie de l'activité (d'après Ruijter et al., 2017, p.48)	88
1.4.7 Les écosystèmes open data vus depuis les tensions paradoxales entre plusieurs niveaux d'analyse	88
Figure 1.10 - Analyse morphologique de la littérature	89
1.5 Identification Gap	91
Un faible volume de recherches étudiant la construction des usages depuis les tensions paradoxales entre les niveaux d'analyses.	91
1.6 Une absence de recherches utilisant la méthodologie de recherche design science pour aborder l'intégralité du processus de construction des usages issus des données ouvertes	93
1.6.1 L'opportunité de solliciter la méthodologie de recherche design science pour combler un gap dans la littérature	93
Quelles sont les situations problématiques qui freinent la construction des usages issus des données ouvertes ?	95
Quels sont les éléments d'un dispositif en Systèmes d'Information permettant de lever les barrières de ces situations problématiques au sein des écosystèmes open data ?	95
1.6.2 L'approche design science dans le domaine des systèmes d'information : une activité scientifique innovante pour créer des connaissances pratiques et utiles aux praticiens de l'open data	95
1.6.2.1 Les fondements et le développement de l'approche design science au sein de la communauté scientifique	96

1.6.2.1.1 Le développement historique des méthodologies de design science dans la communauté scientifique : de l'étude de la conception des systèmes d'information à la conception comme démarche d'étude scientifique	96
a. L'étude de la conception de systèmes d'information	98
b. Le cadre d'analyse classique de la conception : le cycle de vie d'un système d'information	99
c. L'analyse préalable : définir le projet de conception	99
d. La conception générale	100
e. Le développement	100
f. L'implantation	101
g. La maintenance	101
1.6.2.1.2 La méthode MERISE : une méthode pour modéliser les systèmes d'information et faciliter leur conception	101
Figure 1.11: Phases d'analyse des données et des traitements de la méthode Merise (Servigne, 2008, p.3)	102
1.6.2.1.3 Les dernières innovations en matière d'analyse de la conception : RAD, approches sociotechniques	103
a. Le modèle en V face au besoin de systèmes d'information intégrés	103
Figure 1.12: Exemple d'un modèle en V de conception d'un système d'information (Firesmith, 2014, p.14)	103
b. Le modèle en W pour pallier à l'insuffisance de l'expression des besoins métiers	104
Figure 1.13 : Exemple d'un modèle en W de conception d'un système d'information (Petit, 2006, p.155)	105
c. La méthodologie RAD, cheffe de file des méthodes collaboratives : une réponse aux contraintes de temps et à l'évolution dynamique des besoins métiers	105
Figure 1.14 : phases de conception selon la méthode RAD (adapté de Reix et al., 2009, p.85)	106
d. Les approches sociotechniques : la prise en compte de l'environnement extérieur dans le design et l'analyse de la conception	107
1.6.2.1.4 Le basculement des recherches en système d'information : passer du mode 1 au mode 2 de recherche pour gagner en pertinence	108
Figure 1.15 : tableau récapitulatif des facteurs expliquant le déclin de pertinence des recherches en systèmes d'information (Benbasat et Zmud, 1999, p.10)	110
a. Le mode 1 de recherche : accroître la pertinence des recherches en SI	111
b. Le mode 2 de recherche : faire participer activement les praticiens à la conduite de la recherche	111
1.6.2.1.5 La conception comme démarche scientifique	113
a. La reconsidération de la conception en sciences de gestion comme véritable champ scientifique	113
b. Les sciences de l'artificiel : un nouveau cadre scientifique pour développer des connaissances permettant la conception et l'intégration d'artefact dans la réalité (Avenier, 2009)	114
1.6.2.1.6 Apports pratiques et théoriques de la conception comme démarche scientifique	115
Figure 1.16 : la boucle de rigueur dans les Design Science Research Methodology (Hevner et al., 2004, p.80)	115
1.6.2.2 La construction du cadre théorique propre à l'approche design science	117
1.6.2.2.1 Attaches théoriques de l'approche design science	118
1.6.2.2.2 L'apport de la théorie de la traduction : l'émergence de supports matériels à la médiation entre parties prenantes d'une recherche scientifique	119
a. Les Objets Intermédiaires de Conception (OIC) : une représentation partagée de la situation de conception	119
b. Les Objets Frontières (OF)	120

1.6.2.2.3 L'apport de la théorie de la structuration : l'instauration des interactions entre concepteurs et utilisateurs	122
1.6.2.2.4 L'élaboration d'une design theory indépendante pour consolider les recherches de design science	123
a. Les similitudes entre les nouvelles démarches scientifiques de recherche-intervention et la design theory soutenant les démarches de design science : preuve des liens de filiation entre les deux approches.	124
Figure 1.17 : représentation des similitude entre les composants de la nouvelle démarche scientifique évoquée par Pascal (2006, p.241) et ceux de la design theory (Gregor et Jones, 2007, p.322)	125
b. L'anatomie d'une design theory indépendante pour appuyer un nouveau mode de recherche	126
Figure 1.18 : Les relations entre la design theory et les artefact	126
1.6.2.3 Les débats et controverses autour des méthodologies design science aujourd'hui	128
1.6.2.3.1 L'articulation entre méthodologie et épistémologie : la place controversée de l'approche design science	129
a. Les diverses acceptions de l'approche design science en fonction de la posture épistémologique choisie	129
b. L'idée d'un nouveau courant épistémologique indépendant dit design pour s'articuler avec les méthodologies de design science	131
1.6.2.3.2 La complémentarité entre recherche-action et design science en débat	133
a. La recherche-action comme guide des recherches sollicitant les méthodologies design science	134
b. Concevoir les méthodologies design science indépendamment de la recherche-action	133
Figure 1.19 : Processus nominal comportant les différentes activités	136
Figure 1.20 : Lignes directrices en design science (Hevner et al., 2004, p.83)	137
1.6.2.3.3 La finalité des recherches en design science	138
a. Une conception étroite de la finalité des recherches en design science : la construction de théories nouvelles comme contribution essentielle	139
b. Une conception large de la finalité des recherches en design science : les construits, modèles et méthodes comme contributions potentielles	140
1.6.2.3.4 Les limites du concept de système d'information	141
a. Une conception technologique du concept de système d'information	141
Figure 1.21 : Le réseau nomologique des artefact technologiques (Benbasat et Zmud, 2003, p.130)	141
b. Une conception systémique du concept de système d'information	142
1.6.3 L'approche design science à l'appui de la recherche et de la pratique sur la construction des usages issus des données ouvertes.	143
1.6.3.1 La pertinence de l'approche design science pour aborder la transformation de l'open data en innovation de services	144
1.6.3.1.1 L'Action Design Research (ADR)	144
1.6.3.1.2 La Soft Systems Methodology (SSM)	145
1.6.3.1.3 La Recherche-Action (RA)	145
1.6.3.1.4 Contextes d'application de l'approche design science	146
1.6.3.2 Les travaux sur la transformation de l'open data sollicitant l'approche design science	146
1.6.3.2.1 Zuiderwijk (2015) : la Design Science Research Methodology pour construire une infrastructure facilitant la coordination de l'usage de l'open data.	147
Figure 1.22 : Architecture de recherche des recherches de Zuiderwijk (2015, p.29)	149
1.6.4 Les framework pour la construction d'usages issus des données ouvertes	148
1.6.4.1 Pertinence du choix d'un framework comme artefact central de la thèse	149

1.6.4.2 Les caractéristiques d'un framework : un artefact de meta-design d'écosystèmes open data	151
1.6.4.3 Le framework comme infrastructure d'écosystèmes	151
<u>Conclusion du chapitre 1</u>	<u>153</u>
Chapitre 2	154
<u>Méthodologie de recherche design science</u>	154
<i>Comment résoudre les tensions paradoxales qui font obstacle à la construction de nouveaux usages à partir des données ouvertes ?</i>	155
2.1 Présentation de l'architecture de recherche	156
Figure 2.1 : Présentation des choix qui influencent l'architecture de recherche	156
2.1.1 De la posture scientifique et épistémologique à la posture méthodologique	156
2.1.1.1 Posture ontologique et scientifique : le choix du paradigme des sciences de l'artificiel	158
2.1.1.2 Posture épistémologique	158
2.1.1.2.1 La choix du paradigme épistémologique de design research établi par Romme (2003)	158
2.1.1.2.2 Une posture épistémologique inspirée des travaux de Carlsson (2010)	158
Figure 2.2 . Intervention en Systèmes d'Information (Carlsson, 2010, p.220)	159
2.1.1.3 Posture méthodologique	159
2.1.1.3.1 Une méthodologie pour contribuer aux efforts de construction d'une théorie sur l'usage et la transformation des données ouvertes	159
2.1.1.3.2 L'utilisation d'une méthodologie design science sans emprunt à la recherche-action	160
2.1.1.3.3 Une méthodologie de thèse au croisement des deux écoles de recherche en design science	160
2.1.2 Principes opérationnels	161
2.1.2.1 Validité externe : les principes relatifs à l'évaluation du produit de la thèse	161
2.1.2.2 Validité interne : les principes relatifs à l'évaluation du processus de la thèse	161
2.2 Présentation des cinq étapes du design de recherche	162
Figure 2.3. L'architecture de la recherche	162
2.2.1 Les méthodes utilisées	164
2.2.1.1 Méthode de revue de littérature	164
2.2.1.2 Méthode d'exploration hybride	165
2.2.1.2.1 Phase d'invention et de conceptualisation du phénomène	165
Figure 2.4 : Exemple de structure statique d'un système d'objets intelligents (Vaishnavi et al., 1997, p. 277)	166
Figure 2.5 : Exemple de structure statique d'un objet intelligent (Vaishnavi et al., 1997, p.281)	166
Figure 2.6 : Représentation d'un processus selon Datta (1998, p.279)	167
Figure 2.7 : Exemple de comportement dynamique d'un système d'objets intelligents (Vaishnavi et al., p.277)	167
2.2.1.2.2 Choix des observations participantes à effectuer et réalisation	168
2.2.1.2.2.1 Une plateforme open data	168
2.2.1.2.2.2 Les parties prenantes des plateformes open data	168
2.2.1.2.2.3 Les données ouvertes en tant qu'espèce ingénieur	169
2.2.1.2.2.4 L'écosystème open data	169

Figure 2.8 : La représentation d'un écosystème open data (d'après Harrison et al., 2012, p.908)	170
2.2.1.2.2.5 Les artefacts technologiques et sociaux	170
2.2.1.2.2.6 Leurs capacités d'action (material et human agency)	170
2.2.1.2.2.7 Les états des jeux de données ouvertes	171
2.2.1.2.2.8 Les séquences d'articulation	171
2.2.1.2.2.9 Les complémentarités formées	171
Table 2.1 : La grille d'observation des cas exploratoires	172
2.2.1.2.3 Confrontation des observations avec les praticiens et formulation d'un nouveau cadre conceptuel	173
2.2.1.3 Méthode d'étude de cas	174
2.2.1.3.1 Définition de la sous-question de recherche adressée	175
2.2.1.3.2 Sélection des cas	176
2.2.1.3.3 Construction du protocole et des instruments de l'étude de cas	176
2.2.1.3.4 Immersion dans le terrain	176
2.2.1.3.5 Analyse intra-cas	177
2.2.1.3.6 Recherche de pattern inter-cas	177
2.2.1.3.7 Ébauche d'hypothèses	177
2.2.1.3.8 Confrontation avec la littérature	178
2.2.1.3.9 Atteinte du point de clôture	178
2.2.1.4 Méthode de design	178
2.2.1.4.1 Identification d'éléments fonctionnels et construction des propositions de design	179
Figure 2.9 : Catégorisation récursive des articles par groupes d'éléments fonctionnels	180
Figure 2.10: Mots clés choisis dans les différentes revues de littérature	182
2.2.1.4.2 Construction des principes de design	183
2.2.1.4.3 Construction du design de l'artefact	183
Figure 2.11 : Design du système de l'environnement de conception des usages et des trois éléments fonctionnels	185
Figure 2.12 : Pattern de coordination de l'environnement de conception des usages.	187
2.2.1.5 Méthode de prototypage	188
2.2.1.5.1 Définir les objectifs du prototype vis à vis de la recherche et le type de prototypage	189
2.2.1.5.2 Le choix d'un type de prototypage (throw-it-away, incremental, evolutionary)	189
Figure 2.13. Tableau de correspondance entre les objectifs et les types de prototypage.	191
2.2.1.5.3 Sélectionner les fonctions du prototype (classées par élément fonctionnel)	191
2.2.1.5.4 Construction du prototype	191
2.2.1.5.4.1 Architecture (modèle simple) - rappel phase design	192
2.2.1.5.4.1 Opérations/Prototypage des interfaces	192
2.2.1.6 Méthode d'évaluation	192
2.2.1.6.1 Construction de la stratégie d'évaluation : analyse de l'évaluation	193
2.2.1.6.2 Construction de la stratégie d'évaluation : choix d'une stratégie	194
Table 2.2 : Framework de Sélection des Stratégies d'Évaluation	195
2.2.1.6.3 Construction de la stratégie d'évaluation : choix des méthodes	196
Table 2.3 : Framework de sélection des méthodes d'évaluation	196
2.2.1.6.4 Construction de la stratégie d'évaluation : design du détail de l'évaluation	196
2.2.1.6.5 Construction de la stratégie d'évaluation : réalisation des évaluations	202
2.2.1.6.6 Construction de la stratégie d'évaluation : implication sur la prochaine itération	203

Chapitre 3	204
Les barrières à la construction des usages issus des données ouvertes	204

<i>Quelles sont les situations problématiques qui freinent la construction de nouveaux usages issus des données ouvertes ?</i>	205
--	-----

3.1 Invention et conceptualisation des difficultés rencontrées dans l'innovation de services issue des données ouvertes.	206
---	------------

3.1.1 Invention de la structure globale statique à l'œuvre dans l'usage des données ouvertes.	208
Figure 3.1 : Place de la plateforme open data "data.metro.fr" dans la structure globale de l'usage des données ouvertes.	209
Figure 3.2 : Fonctionnement de l'écosystème open data autour de l'Annuaire	213
3.1.2 Invention de la structure statique des systèmes à l'œuvre dans l'usage des données ouvertes	214
Figure 3.3: Représentation graphique de la structure statique des systèmes à l'œuvre (adapté de Kuk et Davies, 2011)	216
3.1.3 Invention du comportement dynamique des systèmes et de la manière dont ce comportement est produit	217
Figure 3.4 : Représentation graphique du comportement dynamique au sein des systèmes	218
3.1.4 Invention de l'évolution historique de la structure globale	219
Figure 3.5 : Évolution de l'écosystème open data	220

3.2 Choix des observations à effectuer et réalisation	221
--	------------

3.2.1 Présentation et analyse du cas d'usage 1 "Quartier" : Utiliser les données ouvertes smart living et smart mobility d'une ville pour faciliter les interactions entre des salariés et leur quartier	221
3.2.1.1 Genèse et évolution du cas d'usage "Quartier"	221
3.2.1.2 Plateforme open data	222
3.2.1.3 Parties prenantes	223
3.2.1.4 Écosystème	223
Table 3.1 : Présentation de l'écosystème du cas "Logement"	224
3.2.1.5 États des données	224
3.2.1.6 Données ouvertes et outils	225
3.2.1.7 Artefacts	226
Table 3.2 : Présentation des artefacts du cas "Logement"	226
3.2.1.8 Capacité d'action	226
Table 3.3 : La capacité d'action des artefacts du cas "Logement"	227
3.2.1.9 Séquence d'articulation	227
Figure 3.6 : Illustration de la première séquence d'articulation du cas d'usage "Quartier"	228
3.2.1.10 Complémentarité	230
3.2.3 Présentation et analyse du cas d'usage 2 "Transport" - Utiliser les données smart mobility pour créer une plateforme de gestion et d'innovation autour des mobilités	231
3.2.3.1 Genèse et évolution du cas d'usage	231
3.2.3.2 Plateforme open data	232
3.2.3.3 Parties prenantes	232
3.2.3.4 Écosystème	232
Table 3.4 : Présentation de l'écosystème du cas "Transport"	233
3.2.3.5 États des données	233
3.2.3.6 Données ouvertes et outils	234
3.2.3.7 Artefacts	235
Table 3.5 : Présentation des artefacts du cas d'usage "Transport".	235
3.2.3.8 Capacité d'action	236
Table 3.6 : Les capacités d'action des artefacts du cas d'usage "Transport"	237

3.2.3.9 Séquence d'articulation	237
Figure 3.7 : La représentation de la première séquence d'articulation du cas d'usage "Transport"	238
3.2.3.10 Complémentarité	240
3.2.4 Présentation et analyse du cas d'usage 3 "Logement" - Utiliser les données smart economy pour maximiser l'activité commerciale d'une zone touristique	241
3.2.4.1 Genèse et évolution du cas d'usage	241
3.2.4.2 Plateforme open data	242
3.2.4.3 Parties prenantes	243
3.2.4.4 Écosystème	243
Table 3.7 : L'écosystème du cas d'usage "Logement"	243
3.2.4.5 États des données	243
3.2.4.6 Données ouvertes et outils	244
3.2.4.7 Artefacts	244
Table 3.8 : Les artefacts du cas d'usage "Logement"	245
3.2.4.8 Capacité d'action	245
Table 3.9 : Les capacités d'action des artefacts du cas d'usage "Logement"	246
3.2.4.9 Séquence d'articulation	246
Figure 3.10 : représentation de la première séquence d'articulation	247
3.2.4.10 Complémentarité	249
3.2.5 Présentation et analyse du cas d'usage 4 "Collaboration" - Utiliser les données de smart governance pour créer un service public de la donnée	250
3.2.5.1 Genèse et évolution du cas d'usage	250
3.2.5.2 Plateforme open data	250
3.2.5.3 Parties prenantes	251
3.2.5.4 Ecosystème	252
Table 3.10 : L'écosystème du cas d'usage "Collaboration"	252
3.2.5.5 États des données	253
3.2.5.6 Données ouvertes et outils	254
3.2.5.7 Artefacts	255
Table 3.11 : Les artefacts du cas d'usage "Collaboration"	256
3.2.5.7 Capacité d'action	256
Table 3.12 : Les capacités d'action du cas d'usage "Collaboration"	257
3.2.5.8 Séquence d'articulation	257
Figure 3.9 : La représentation de la première séquence d'articulation	258
3.2.5.9 Complémentarité	260
3.2.3 Analyse transversale des quatre cas exploratoires d'usage des données ouvertes	261
3.2.3.1 Une plateforme open data	262
Table 3.13 : Les composants des plateformes open data	262
3.2.3.2 Les parties prenantes des plateformes open data	263
Table 3.14 : Les parties prenantes des plateformes open data	263
3.2.3.3 Les données ouvertes en tant qu'espèce ingénieur	264
Table 3.15 : Les outils structurés par les données ouvertes	264
3.2.3.4 Les écosystèmes open data	264
Table 3.16 : Les éléments des écosystèmes open data	265
3.2.3.5 Les artefacts techniques et sociaux et leur capacité d'action	266
Table 3.17 : Les artefacts sociaux et techniques impliqués dans l'usage de l'open data	266
Table 3.18 : Les capacités d'action des artefacts et leur influence sur l'état des jeux de données ouverts	267
Table 3.19 : Les différentes configurations des jeux de données ouverts	268
3.2.3.7 Les séquences d'articulation	268
Table 3.20 : Les séquences d'articulation identifiables dans l'usage de l'open data	268

3.2.3.8 Les complémentarités formées	269
Table 3.21: Les complémentarités formées lors des séquences d'articulation	269
3.2.3.9 Synthèse des résultat de l'analyse transversale	269
3.2.3.9.1 Les blocages d'ordre technique	270
3.2.3.9.1 Les blocages d'ordre social	270
3.3 Confrontation des observations avec les praticiens et formulation d'un nouveau cadre conceptuel	272
3.3.1 Constitution de l'échantillon de parties prenantes	272
Table 3.22 : L'échantillon de parties prenantes	274
3.3.2 Méthode de collecte et d'analyse des entretiens	274
3.3.2.1 Collecte	274
3.3.2.2 L'analyse	275
Table 3.23 : Extrait de la confrontation des éléments empiriques et des verbatims des entretiens	276
3.3.3 Résultats et présentation du modèle théorique affiné	277
3.3.3.1 Les barrières techniques (human-computer interactions)	277
3.3.3.1.1 Les barrières techniques inhibant les fonctions des plateformes open data	277
3.3.3.1.2 Les barrières techniques inhibant les capacités d'action des artefacts techniques	277
3.3.3.1.3 Les barrières techniques forçant l'arrêt des séquences d'articulation entre artefact technique et artefact social	278
3.3.3.2 Les barrières sociales	279
3.3.3.2.1 Les barrières sociales empêchant la mise en place d'une organisation sociale dotée d'une gouvernance propre autour des plateformes	279
3.3.3.2.2 Les barrières sociales empêchant la mise en place d'une coordination entre les systèmes	279
3.3.3.2.3 Les barrières sociales inhibant les capacités d'action des artefacts sociaux	280
3.3.3.2.3 Les barrières sociales forçant l'arrêt des séquences d'articulation entre artefact social et un artefact technique	281
Table 3.24 : Extrait du tableau récapitulatif des barrières socio-techniques freinant l'usage des données	282
3.3.3.3 Présentation du modèle théorique affiné	282
3.3.3.3.1 Un framework pour identifier les barrières à lever pour mettre en oeuvre l'usage des données ouvertes	283
Table 3.35 : framework d'identification des barrières face à l'usage de l'open data	283
3.3.3.3.2 Un modèle théorique affiné pour comprendre l'usage des données ouvertes	284
Figure 3.10 : Un modèle conceptuel affiné de l'usage de l'open data	284
Conclusion du chapitre 3	285

Chapitre 4 **287**

Les attentes des opérateurs de la donnée vis-à-vis d'un dispositif en Systèmes d'Information **287**

Quels sont les besoins fonctionnels auxquels un dispositif en Systèmes d'Information doit répondre pour lever les barrières socio-techniques qui freinent la construction des usages issus des données ouvertes ? 288

4.1 Sélection et présentation des cas	290
4.1.1 Sélection des écosystèmes de la métropole Grenoble Alpes et de la métropole du Grand Lyon	290
4.1.2 Présentation des cas	290

4.1.2.1 Cas 1 - L'usage des données ouvertes dans la métropole grenobloise	290
Figure 4.1 : Illustration des frontières et unités d'analyse choisie pour le cas de grenoble	291
4.1.2.2 Cas 2 - L'usage des données ouvertes dans la métropole du Grand Lyon	293
Figure 4.2 : Illustration des frontières et unités d'analyse choisie pour le cas du Grand Lyon	294
4.2 Construction du protocole et instruments de l'étude de cas	296
4.2.1 Approche du terrain et guides d'entretien	296
Table 4.1 : Thèmes abordés avec les personnes interrogées	298
4.2.2 Guide pour la réalisation de rapports d'études de cas	299
4.2.3 Sources et méthodes de collecte de données	299
Figure 4.3 : La liste des sources de données	300
Figure 4.4 : La liste des entretiens semi-directifs conduits	302
4.3 Analyse des données : codage des entretiens en utilisant un système permettant d'identifier les barrières socio-techniques et les moyens imaginés pour les lever	302
4.3.1 Utilisation des fonctionnalités de la plateformes (téléchargement/flux de données)	303
Table 4.2 : Barrière freinant l'utilisation des fonctionnalités de la plateforme open data	306
4.3.2 Évaluation du potentiel des données à partir de leur représentation	306
Table 4.3 : Barrière freinant l'évaluation du potentiel des données	310
4.3.3 Transformation des données grâce à des artefacts techniques	310
Table 4.4 : Barrière freinant la transformation des données grâce à des artefacts techniques	312
4.3.4 Création d'un cadre cognitif partagé entre les systèmes pour les coordonner	312
Table 4.5 : Barrière freinant la création d'un cadre cognitif partagé entre les systèmes	315
4.3.5 Institutionnalisation de l'usage des données	315
Table 4.6 : Barrière freinant l'institutionnalisation de l'usage des données	318
4.4 Résultats et discussion : un besoin d'avoir des connaissances partagées entre opérateurs pour lever plus facilement les barrières	318
4.4.1 Les besoins fonctionnels pour lever les obstacles dans l'utilisation des fonctionnalités d'une plateforme open data	319
4.4.1.1 Manque de confiance dans la plateforme	320
4.4.1.2 Manque de lisibilité des interfaces	320
Figure 4.5 : Les besoins fonctionnels relatifs à l'utilisation des fonctionnalités d'une plateforme open data.	322
4.4.2 Les besoins fonctionnels pour lever les obstacles dans l'évaluation du potentiel des données à partir de leur représentation	322
4.4.2.1 Manque de compétences et d'expertise technique	322
4.4.2.2 Difficulté d'intégration de profils experts et techniques dans les processus d'utilisation	323
Figure 4.6 : Les besoins fonctionnels relatifs à l'évaluation du potentiel des données à partir de leur représentation	324
4.4.3 Les besoins fonctionnels pour lever les obstacles dans la transformation des données grâce à des artefacts techniques	324
4.4.3.1 Manque de qualité des données	325
4.4.3.2 Hétérogénéité des conditions d'utilisation	325
Figure 4.7 : Les besoins fonctionnels relatifs à la transformation des données grâce à des artefacts techniques.	326
4.4.4 Les besoins fonctionnels pour lever les obstacles dans la création d'un cadre cognitif partagé entre les systèmes	327
4.4.4.1 Manque d'un cadre juridique et politique clair	327
4.4.4.2 Manque d'une vision cohérente autour de la production et de l'utilisation des données	328
Figure 4.8 : Les besoins fonctionnels relatifs à la création d'un cadre cognitif partagé entre les systèmes	329
4.4.5 Les besoins fonctionnels pour lever les obstacles dans l'institutionnalisation de l'usage des données	329
4.4.5.1 Incompatibilité des structures organisationnelles avec l'open data	329
4.4.5.2 Divergence culturelle de l'ouverture des données par rapport aux organisations	330

Figure 4.9 : Les besoins fonctionnels relatifs à l'institutionnalisation de l'usage des données	332
<u>Conclusion du chapitre 4</u>	<u>333</u>
Chapitre 5	335
Design d'un dispositif en Systèmes d'Information	335
<i>De quels éléments un dispositif visant à répondre aux douze besoins fonctionnels identifiés est-il constitué ?</i>	336
Figure 5.1 : Le rôle des propositions dans le design du dispositif	337
Figure 5.2 : Le rôle des mécanismes, principes et fonctions dans le design du dispositif	338
5.1. Construction des propositions de design	340
5.1.1 Première proposition : Un réseau d'information autour des données ouvertes pour lever les barrières face à la construction des usages	341
5.1.2 Proposition 2 : Une infrastructure open data basée sur un modèle de métadonnées pour lever les barrières face à la construction des usages	344
5.1.3 Troisième proposition : Un manifeste et sa déclinaison opérationnelle, le codebook pour lever les barrières face à la construction des usages	348
5.2 Construction des principes de design	353
5.2.1 Principes relatifs à l'ensemble du design construits à partir de la littérature sur le concept d'usage	354
5.2.1.1 Premier principe	356
5.2.1.2 Deuxième principe	356
5.2.1.3 Troisième principe	357
5.2.1.4 Quatrième principe	358
5.2.2 Principes relatifs au design d'un réseau d'information construits à partir de la littérature sur la théorie de la construction sociale de la technologie (Pinch et Bijker, 1984)	359
5.2.2.1 Premier principe	361
5.2.2.2 Deuxième principe	362
5.2.2.3 Troisième principe	363
5.2.3 Principes relatifs au design d'un modèle de métadonnées construit à partir de la littérature sur la théorie de la cognition distribuée (Hutchins, 1995)	364
5.2.3.1 Premier principe	365
5.2.3.2 Deuxième principe	366
5.2.3.3 Troisième principe	366
5.2.4 Principes relatifs au design d'un manifeste et d'un codebook construits à partir de la littérature sur la théorie du middleground (Cohendet et al., 2014)	368
5.2.4.1 Premier principe	369
5.2.4.2 Deuxième principe	369
5.3 Construction du design de l'artefact	371
Figure 5.3 : La place du dispositif et de ses différentes aires fonctionnelles dans le modèle théorique de l'usage des données ouvertes	372
5.3.1 Construction du design du système de l'environnement de conception des usages dans son ensemble	373
Figure 5.4 : Design des deux mécanismes principaux définissant le comportement du système de l'artefact de meta-design.	375
5.3.1.1 Le sparring : un processus collaboratif permettant d'aller au bout/de traverser (work through) les paradoxes et tendre vers l'innovation (Luscher et Lewis, 2008)	375
5.3.1.2 La boucle de rétroaction pour alimenter la transformation	377
5.3.2 Construction du design du système du réseau d'information	379
Figure 5.5 : Design des 3 mécanismes principaux définissant le comportement du système du réseau d'information	380

5.3.2.1 La surdétermination de l'usage du réseau par la mise en oeuvre de normes techniques et sociales implémentées dans l'objet technologique (Lessig, 1999)	380
5.3.2.1.1 Normes techniques guidant le management des composants du réseau	381
5.3.2.1.2 Normes sociales guidant les interactions entre membre du réseau et l'accomplissement de tâches collectives	381
5.3.2.2 Un mécanisme de gestion collaboratif des composants du réseau à travers une base de données collaborative	382
5.3.2.2 Le tableur comme interface de gestion des composants du réseau	383
5.3.3 Construction du design du système de gestion des tâches à partir d'un modèle de méta-données	385
Figure 5.6 : Design des 3 mécanismes principaux définissant le comportement du système de modèle de méta-données	386
5.3.3.1 Le modèle de métadonnées	386
5.3.3.2 La flexibilité organisationnelle du système d'organisation des tâches basé sur le modèle	387
5.3.3.3 La plasticité des dispositions spatiales de la représentation des tâches basé sur le modèle	388
5.3.3.4 Le rattachement des entités du modèle à des exemples réels et concrets	388
5.3.4 Construction du design du système du manifeste et de son codebook	389
Figure 5.7 : Design des 2 mécanismes principaux définissant le comportement du manifeste et du codebook	390
5.3.4.1 Le sourcing et l'intégration constante de nouvelles pratiques dans le manifeste et le codebook	390
5.3.4.2 Le développement et l'expérimentation de produits et services à partir du manifeste et du codebook	391
5.4 Construction de scénarios d'usage compilant les pattern de coordination de l'artefact avec un contexte réel de développement des usages des données ouvertes	392
5.4.1 Rallier les diverses communautés autour de piliers définissant l'usage des données ouvertes : construction du scénario 1	393
5.4.2 Mutualiser et classer les ressources permettant de développer les usages : construction du scénario 2	396
5.4.3 Mettre à disposition du plus grand nombre une forge des usages des données ouvertes : construction du scénario 3	399
5.5 Description des fonctions du design de l'environnement de conception des usages des données ouvertes.	402
5.5.1 Description des fonctions de l'environnement de conception pris dans son ensemble	402
Figure 5.8 : Table décrivant les fonctions de l'environnement de conception des usages pris dans son ensemble	403
5.5.2 Description des fonctions du réseau d'information	403
Figure 5.9 : Table décrivant les fonctions du réseau d'information	405
5.5.3 Description des fonctions du système de gestion des tâches basé sur le modèle de méta-données	405
Figure 5.10 : Table décrivant les fonctions du système de gestion des tâches	406
5.5.4 Description des fonctions du manifeste et du codebook	407
Figure 5.11 : Table décrivant les fonctions du manifeste et du codebook	407
Conclusion du chapitre 5	408
Figure 5.12 : Fonctions relatives à l'ensemble de l'environnement	409
Figure 5.13 : Fonctions relatives au réseau d'information	409
Figure 5.14 : Fonctions relatives à un système de gestion des tâches basée sur un modèle open data	410
Figure 5.15 : Fonctions relatives au manifeste et son codebook	411
Chapitre 6	412
Prototypage d'un dispositif en Systèmes d'Information	412
<i>Sous quelle forme le dispositif en systèmes d'information se matérialise-t-il ?</i>	413
<i>Arthur Sarazin - « Construire les usages des données ouvertes » - Thèse CERAG de Grenoble - 2021</i>	610

6.1 Objectifs du prototype	414
6.1.1 Les trois apports du prototype vis à vis de la résolution de ma problématique de recherche	414
6.1.2 Le choix d'un type de prototypage évolutif	415
Table 6.1 : Table de correspondance entre les apports souhaités du prototype et le type de prototypage à effectuer (adapté d'Ince et Hekmatpour, 1987, p.12)	415
6.2 Sélection des fonctions	416
Table 6.2 : intégration et amélioration des fonctions dans les différentes versions du prototype	417
6.2.1 Sélection des fonctions à intégrer dans la Version 0 du prototype	417
6.2.1.1 Fonctions de l'environnement de conception des usages pris dans son ensemble	418
Table 6.3 : Fonctions relatives à l'ensemble de l'environnement de conception sélectionnées pour intégrer la Version 0	418
6.2.1.2 Fonctions du réseau d'information	419
Table 6.4 : Fonctions relatives au réseau d'information sélectionnées pour intégrer la Version 0	419
6.2.1.3 Fonctions du système de gestion des tâches basé sur un modèle de méta-données	420
Table 6.5 : Fonctions relatives au système de gestion des tâches sélectionnées pour intégrer la Version 0	420
6.2.1.4 Fonctions du manifeste et de son codebook	420
6.2.2 Sélection des fonctions à intégrer dans la Version alpha du prototype	421
6.3 Construction du prototype	421
6.3.1 Construction de la version 0 du prototype	421
6.3.1.1 Architecture de la version 0	422
6.3.1.1.1 Cartographie de l'architecture métier de l'Open Data Canvas	422
Figure 6.1 : Cartographie de l'architecture métier de la version 0 du prototype de l'Open Data Canvas	424
6.3.1.1.2 Cartographie de l'architecture fonctionnelle de l'Open Data Canvas	425
Figure 6.2: Cartographie de l'architecture fonctionnelle de la version 0 du prototype de l'Open Data Canvas	426
6.3.1.1.3 Cartographie de l'architecture applicative de l'Open Data Canvas	428
Figure 6.3 : Cartographie de l'architecture applicative de la version 0 du prototype de l'Open Data Canvas	429
6.3.1.2 Scénario d'utilisabilité de la version 0	430
6.3.2 Construction de la version alpha du prototype	430
6.3.2.1 Architecture de la version alpha du prototype	431
6.3.2.2 Scénario d'utilisabilité de la version alpha	431
6.3.2.2.1 Processus 1 : Choix du type d'intervention	431
6.3.2.2.2 Processus 2 : Mise en place de l'environnement	432
Figure 6.4 : Interface d'accueil de la chambre de conception de MetroMeta.	432
6.3.2.2.3 Processus 3 : Sparring	433
Figure 6.5: Interface de l'espace de discussion du projet Eco-quartier contenant les quatre fils de discussion du processus de sparring	433
Figure 6.6 : Interface de l'espace de travail (workspace) de l'Intervenant et de la base lui permettant d'animer le sparring avec MetroMeta	434
Figure 6.7 : Interface de la table d'animation de l'étape 1 du processus de sparring avec zoom sur le template de question linéaire pour le projet Eco-quartier	435
Figure 6.8 : Interface du fil de discussion de l'étape 1 du processus de sparring pour le projet Eco-quartier après l'ajout d'une nouvelle ligne dans la table d'animation	436
Figure 6.9 : Interface de la table d'animation de l'étape 2 du processus de sparring avec zoom sur le template de questions circulaires	437
Figure 6.10 : Interface de la table d'animation de l'étape 3 du processus de sparring avec zoom sur le template de questions réflexives	438
Figure 6.11 : Interface de la table d'animation de l'étape 4 du processus de sparring avec zoom sur le template de questions stratégiques	439
6.3.2.2.4 Processus 4 : Draft du plan de développement	439
Figure 6.12 : Interface de l'espace de travail (workspace) de MetroMeta et de la base lui permettant de construire un plan de développement des usages	440

Figure 6.13 : Interface de la base Open Data Work de MetroMeta avec un focus (rectangle noir) sur les trois premières tables	441
Figure 6.14: Extrait de l'interface de la Table d'association	442
Figure 6.15: Interface présentant le fonctionnement de la colonne liant la Table d'association avec la Table "Donnees_ouvertes"	443
Figure 6.16 : Interface présentant l'espace de travail de l'Intervenant avec les bases des référentiels (rectangle noir) de composants et de tâches entretenus par la communauté #teamopendata	445
Figure 6.17 : Interface présentant la chambre dédiée à l'OpenDataCanvas et l'espace dédié au sourcing de composants et de tâches.	445
Figure 6.18 : Interface présentant l'espace de dédié au sourcing de composants et de tâches ainsi que le fil de discussion dédié au sourcing de composants relatifs aux "arguments anti open data"	446
Figure 6.19 : Interface du questionnaire de dépôt des propositions de nouveaux composants	447
Figure 6.20 : Interface de la base permettant le filtrage des composants et tâches sourcés dans la communauté #teamopendata	448
6.3.2.2.5 Processus 5 : Crowdsourcing d'acteurs, de composants et de tâches	448
Figure 6.21 : Interface montrant un exemple d'appel à connaissances	449
6.3.2.2.6 Processus 6 : Filtrage de sourcing	450
Figure 6.22 : Interface montrant le tableau Kanban du plan de développement des usages.	450
Figure 6.23 : Interface montrant le calendrier du plan de développement des usages.	451
6.3.2.2.7 Processus 7 : Communication du plan de développement	451
Figure 6.24: Interface montrant l'espace de travail réservée aux acteurs du projet (workspace design) et la base permettant de mettre en oeuvre le plan de développement des usages	452
Figure 6.25 : Interface montrant le tableau Kanban du laboratoire d'innovation avec les transformations qu'il doit réaliser.	453
Figure 6.26 : Interface montrant les fils de discussions dédiées aux transformations à effectuer par les acteurs	454
Figure 6.27 : Interface montrant les échanges relatifs au tag "données-énergie"	455
Figure 6.28 : Interface de la base OpenDataWork avec un focus sur la table d'association propre à la deuxième itération	456
Conclusion du chapitre 6	457

Chapitre 7 **459**

Évaluation du dispositif en Systèmes d'Information **459**

Dans quelle mesure le dispositif en système d'information permet-il de lever les barrières socio-techniques qui freinent la construction des nouveaux usages issus des données ouvertes ? 460

7.1 Analyse de l'évaluation des prototypes	461
Figure 7.1 : Synthèse de l'analyse de l'évaluation	462
7.1.1 Analyse de l'évaluation de la version 0 du prototype	463
7.1.1.1 La forme du prototype à évaluer	463
7.1.1.2 La nature du prototype à évaluer	463
7.1.1.3 Les propriétés du prototype à évaluer	463
7.1.1.4 Les objectifs de l'évaluation	463
7.1.1.5 Les contraintes de l'environnement de recherche	464
7.1.1.6 La rigueur à appliquer	464
7.2 Choix d'une stratégie d'évaluation	464
7.2.1 Choix d'une stratégie d'évaluation pour la version 0 du prototype	464

Figure 7.2 : Framework de sélection de la stratégie d'évaluation de la version 0 du prototype (d'après Venable et al., 2010, p.11)	465
7.2.2 Choix d'une stratégie d'évaluation pour la version alpha du prototype	466
Figure 7.3 : Framework de sélection de la stratégie d'évaluation de la version alpha du prototype (d'après Venable et al., 2010, p.11)	467
7.3 Choix de la méthode d'évaluation	468
7.3.1 Choix de la méthode d'évaluation de la version 0 du prototype : des focus group inspirés des méthodes utilisées en ergonomie informatique	468
7.4 Réalisation de l'évaluation	468
7.4.1 Réalisation de l'évaluation de la version 0 du prototype	468
7.4.1.1 Profil des participants aux 5 focus group	469
Table 7.1 : Les caractéristiques des focus group et des participants	470
7.4.1.2 Analyse des données	471
Table 7.2 : Extrait de la table faisant la synthèse du codage des transcriptions des focus group	472
7.4.1.3 Résultats	473
7.4.1.3.1 Étape 1 - Utilisation des fonctionnalités d'une plateforme open data	473
7.4.1.3.2 Étape 2 - Évaluation du potentiel des données	473
7.4.1.3.3 Étape 3 - Transformation des données	475
7.4.1.3.4 Étape 4 - Création d'un cadre cognitif partagé	476
7.4.1.3.5 Étape 5 - Institutionnalisation de l'usage des données	478
Table 7.3 : Tableau sur l'utilité de l'Open Canvas sur le premier besoin fonctionnel	480
Table 7.4 : Tableau sur l'utilité de l'Open Canvas sur le troisième besoin fonctionnel	481
Table 7.5 : Tableau sur l'utilité de l'Open Canvas sur le quatrième besoin fonctionnel	483
Table 7.6 : Tableau sur l'utilité de l'Open Canvas sur le cinquième besoin fonctionnel	484
Table 7.7 : Tableau sur l'utilité de l'Open Canvas sur le sixième besoin fonctionnel	485
Table 7.8 : Tableau sur l'utilité de l'Open Canvas sur le septième besoin fonctionnel	486
Table 7.9 : Tableau sur l'utilité de l'Open Canvas sur le huitième besoin fonctionnel	488
Table 7.10 : Tableau sur l'utilité de l'Open Canvas sur le neuvième besoin fonctionnel	489
Table 7.11 : Tableau sur l'utilité de l'Open Canvas sur le dixième besoin fonctionnel	490
Table 7.12 : Tableau sur l'utilité de l'Open Canvas sur le onzième besoin fonctionnel	491
Table 7.13 : Tableau sur l'utilité de l'Open Canvas sur le douzième besoin fonctionnel	492
7.5 Implications	493
7.5.1 Implication des résultats de l'évaluation de la version 0 du prototype	493
7.5.1.1 Confirmation ou infirmation de l'utilité du design de l'environnement de conception des usages	493
Table 7.14 : Utilité du Canvas relativement aux douze besoins fonctionnels	494
7.5.1.3 Présentation des feedbacks permettant d'informer la construction d'une nouvelle version du prototype.	494
<u>Conclusion du chapitre 7</u>	<u>497</u>
Chapitre 8	498
Discussion et conclusion générale	498
8.1 Résultats marquants de la recherche	499
8.1.1 Résultats empiriques de la recherche	499

8.1.1.1 Première sous-question de recherche : état des lieux de la construction des usages issus des données ouvertes	499
8.1.1.2 Deuxième sous-question de recherche : les situations problématiques face à la construction des usages	501
Table 8.1 : Framework d'identification des barrières face à l'usage de l'open data	502
8.1.1.3 Troisième sous-question de recherche : les éléments d'un dispositif en Systèmes d'Information pour résoudre les situations problématiques	503
8.1.1.4 Question de recherche centrale : une réponse aux tensions paradoxales faisant obstacle à la construction des usages	506
8.1.2 Résultats théoriques de la recherche	507
8.1.2.1 Une design theory pour les environnements de conception des usages	507
8.1.2.1.1 Périmètre et objectif de la théorie	507
8.1.2.1.2 Construits	508
8.1.2.1.3 Principes de structure, de forme et de fonction	508
8.1.2.1.4 Plasticité de l'artefact	509
8.1.2.1.5 Propositions testables	509
8.1.2.1.6 Connaissances justificatives	510
8.1.2.1.7 Principes d'implémentation	510
8.1.2.1.8 Prototype exposé	510
8.1.2.2 Les théories du paradoxe pour articuler et opposer les trois kernel theory sollicitées	511
8.1.3 Implications de mes résultats	512
8.1.3.1 Apports pratiques et fiabilité	512
8.1.3.2 Apport méthodologique et fiabilité	513
8.1.3.3 Apports théoriques et fiabilité	514
8.1.3.4 Fiabilité des apports : validité interne et externe de la démarche	515
8.2 Limites et ouverture de la recherche	516
8.2.1 Limites de la recherche	516
8.2.1.1 Du point de vue de la validité interne	516
8.2.1.2 Du point de vue de la validité externe	516
8.2.2 Ouverture : recommandations pour les recherches futures sur les usages des données ouvertes	517
Bibliographie	519
Glossaire	545
Annexes	557
Liste des figures	593
Liste des tables	596
Table des matières	598

