Aide à la gestion de l'impact des stratégies IT sur la maîtrise du risque réglementaire

Guillaume Delorme

N° d'ordre NNT : 2023LYO30008

THÈSE DE DOCTORAT DE L'UNIVERSITÉ JEAN MOULIN LYON 3

Membre de l'université de Lyon

École doctorale n° 485 - Infomaths, Informatique, mathématiques

Discipline : **Informatique**

Soutenue publiquement le 28/03/2023, par

Guillaume DELORME

Aide à la gestion de l'impact des stratégies IT sur la maîtrise du risque réglementaire

---

Laboratoire de recherche : **Magellan**

Directrice de thèse : **Mme Guilaine TALENS**

Devant le jury composé de :

Mme Guilaine TALENS
Maîtresse de conférences HDR, université Jean Moulin Lyon 3. Directrice de thèse

Mme Frédérique BIENNIER
Professeure des universités,
institut national des sciences appliquées de Lyon, Villeurbanne. Présidente du jury

Mme Kathia MARCAL DE OLIVEIRA
Professeure des universités,
université polytechnique Hauts-de-France, Valenciennes. Rapporteure

M. Benjamin NGUYEN
Professeur des universités,
institut national des sciences appliquées du Centre Val de Loire, Bourges. Rapporteur

M. Éric DISSON
Maître de conférences, université Jean Moulin Lyon 3. Examinateur

M. Ludovic MÉ
Advance Research Position HDR, centre Inria de l'université de Rennes. Examinateur

Remerciements

Au terme de ce travail je tiens à remercier :

Madame Kathia Marcal de Oliveira, Professeur à l'université Polytechnique Hauts-de-France, pour avoir accepté d'être rapporteur scientifique de ce mémoire.

Monsieur Benjamin NGUYEN, Professeur à l'INSA Centre Val de Loire, pour avoir contribué à l'évaluation de mes travaux, apporté des remarques et conseils pertinents tout au long de ces travaux et ainsi que pour avoir accepté d'être rapporteur scientifique de ce mémoire.

Je tiens également à remercier particulièrement Madame Frédérique BIENNIER, Professeur à l'INSA Lyon, pour avoir contribué à l'évaluation de mes travaux en sa qualité de rapporteur.

Monsieur Ludovic Mé, Adjoint au directeur scientifique de l'Institut national de recherche en sciences et technologies du numérique (Inria).

Je tiens également à remercier ceux qui ont su me tenir motivé tout au long de ces années :

Madame Guilaine TALENS et Monsieur Eric DISSON respectivement directeur et co-directeur de thèse pour m'avoir accueilli, accompagné et supporté tout au long de ce parcours. Ce mémoire concrétise un travail de plusieurs années pendant lesquelles la patience, bienveillance et rigueur dont ils ont su et du faire preuve m'ont permis de mener cette aventure à son terme.

Je tiens aussi à remercier les membres de Solvay et plus particulièrement ceux de l'équipe « Information Risk & Security » qui ont su m'accueillir, me soutenir et m'aiguiller lors de la conduite de ces travaux : Patrice, Guillaume, Virginia et Thomas. Bien que dans d'autres équipes, j'ai également eu l'occasion de travailler avec d'autres belles personnes chez Solvay : Elise, Alexis, Thierry, Stéphanie.

J'ai également une pensée toute particulière pour ma famille et mes amis qui ont subi sans relâche les contraintes d'un thésard.

Sommaire

Table des figures

Liste des Tableaux

Introduction générale

Depuis l'apparition de l'ENIAC (Electronic Numerical Integrator and Computer) en 1946, les technologies de l'information (IT) se sont graduellement immiscées dans les moindres méandres des organisations publiques et privées, parvenant même jusqu'au cœur des foyers. Leur place centrale dans les administrations publiques, les industries et les sociétés de services ont créé une dépendance universelle. L'assurance de leur bon fonctionnement est devenue une priorité, il y va de la protection et la stabilité des sociétés, Etats et individus.

Les conséquences de ce rapport aux technologies de l'information et de la communication et des nouvelles dimensions socio-économiques qu'elles introduisent sont à l'origine des prémices de la gestion du risque de l'information et de sa sécurité. Conscients de ces enjeux sociétaux, les législateurs et les Etats se sont intéressés à la gestion et sécurisation de l'information en promulguant des règlementations et des lignes directrices de plus en plus strictes, précises et contraignantes. La gestion de ces nouvelles contraintes se traduit pour les responsables de traitements et les propriétaires de données par une gestion du risque IT plus complexe et compliquée.

« Nul n'est censé ignorer la loi »

De cette fiction juridique découle la problématique de l'étendue et l'absolutisme de la conformité qui motive nos travaux. Traduction d'une maxime latine autrefois présente dans le code civil français, celle-ci représente fidèlement les enjeux auxquels les organisations font face : assurer leur conformité et respecter les différentes réglementations aux contraintes parfois divergentes.

Nous sommes aujourd'hui témoins de l'émergence d'un nouveau cycle de gestion du risque de l'information en cybersécurité, celui de la conformité. Celui-ci vient supplémenter les risques traditionnels par la responsabilisation des organisations quant aux manquement de moyens et mesures nécessaires à la prévention d'infraction aux règles juridiques et à la protection des données. Par l'utilisation des réglementations et autres éléments du corpus juridique, les législateurs et politiques peuvent également contraindre les organisations par leur contrôle et encadrement. Ceux-ci sont alors à l'origine d'un risque pour toute organisation qui manquerait de se conformer à ces obligations.

Notre propos s'attachera aux risques et impacts liés à la conformité des règlementations régissant le traitement des données et/ou les gouvernances et processus des technologies de l'information et communication et/ou les technologies et services de l'information.

L'objectif de nos travaux de recherche est la conception d'un système capable de représenter de manière intelligible et précise les contraintes et obligations réglementaires qu'une organisation doit respecter pour assurer sa conformité. Ce système doit également permettre la considération de l'évolution des exigences réglementaires, leurs disparités et similarités ainsi que la reconnaissance du champ législatif applicable à une organisation. Celui-ci doit en outre être en mesure de fournir les informations nécessaires à la prise de décision stratégique dans le cadre d'une mise en conformité avec une loi régissant le traitement de données.

Les travaux de cette thèse ont pour ambition de répondre à deux objectifs principaux:

- Comment représenter un risque règlementaire multidisciplinaire ?
- Comment permettre une meilleure gestion d'un tel risque ?

Notre sujet est multidisciplinaire. Il est à la jonction de deux domaines distincts : le domaine juridique et le domaine des technologies de l'information et de la communication.

Afin de répondre aux enjeux de notre proposition de support à la gestion du risque règlementaire multidisciplinaire, nous avons organisé cette thèse comme suit :

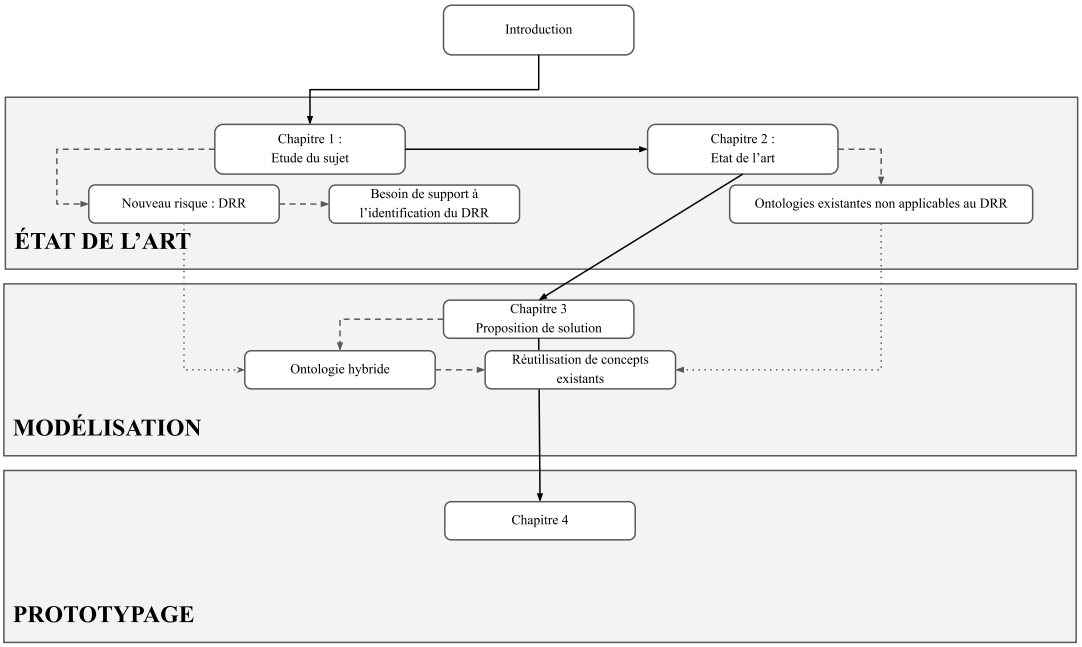
Le **Chapitre 1** est consacré à la présentation du contexte de notre domaine d'étude. La première section présente l'origine du « Data Regulation Risk » (DRR), une classe de risque liée à la conformité des règlementations spécifiques. Il sera également étudié l'importance et l'impact que ces réglementations peuvent avoir sur les organisations. La seconde section présente les principes fondamentaux du risque et introduit des méthodologies de gestion du risque. Enfin, nous concluons ce premier chapitre par une analyse d'adaptabilité de différentes méthodologies au DRR.

Le **Chapitre 2** est composé de deux sections principales. La première présente les ontologies comme outils de gestion de connaissances. Il sera également étudié leurs objectifs, leurs typologies et leur conception. La seconde section étudie des travaux existants et présentant des ontologies. Il sera notamment question d'ontologies de gestion de la sécurité de l'information et d'ontologies de conformité ou juridiques.

Le **Chapitre 3** est consacré à la présentation et l'élaboration de notre modèle qui repose sur une ontologie hybride. Ce chapitre est composé de trois sections principales. La première présente les objectifs et les contextes d'utilisation de notre modèle. La deuxième introduit les concepts et relations nécessaire à la conception de notre ontologie. Elle présente également les fonctionnalités prévues par notre proposition de solution. Nous concluons ce troisième chapitre par des exemples d'utilisation de notre modèle en réalisant les instanciations de deux articles de lois.

Le **Chapitre 4,** composé de deux sections principales, est consacré à la présentation de notre prototype, le protocole de test et de retours d'expérience ainsi que son éventuelle mise en production et les contraintes associées.

L'organisation de ce mémoire est présentée dans la figure suivante :



Chapitre 1 – Pour une meilleure prise en compte du risque lié à la conformité des données réglementées

Ce chapitre sera consacré à l'étude du contexte de notre domaine d'étude. D'une part nous présentons l'apparition d'une nouvelle classe de risque, d'autre part, nous étudions les méthodologies de gestion du risque.

Il convient d'étudier l'origine de ce risque qui trouve ses racines dans le corpus juridique moderne. L'importance grandissante des données et des systèmes d'informations au sein de nos sociétés a graduellement amené les politiques et législateurs à établir des cadres juridiques. Ceux-ci présentent un risque pour les organisations échouant à se conformer aux exigences réglementaires.

La première section présente cette nouvelle classe de risque, son contexte et son impact sur les organisations. La seconde section introduit le concept de risque et sa gestion. Son objectif est d'analyser différentes méthodes afin de déterminer leur applicabilité.

Section I. Les enjeux règlementaires sur les données et l'information

Cette section est organisée comme suit : les deux premières parties présentent l'importance de la donnée et de sa sécurisation alors que la troisième et quatrième étudient l'existence des réglementations, leur ambition et leur impact. Cette section termine par la présentation d'un nouveau risque que nous appelons Data Regulation Risk[1].

1.1. De la gestion de l'information…

Cette partie a pour intention d'alimenter notre réflexion en étudiant la place des systèmes d'information, leur importance et leur complexité. Il nous parait important d'apporter du contexte et de nous attarder sur l'importance grandissante des systèmes d'information et de la gestion de donnée. Cette partie nous permet de comprendre les enjeux auxquels sont confrontés non seulement les organisations privées et publiques mais aussi les individus et les gouvernements.

1.1.1. L'information et les systèmes au cœur des organisations

Longtemps considérées par les économistes comme simples outils, fonction de productions ou courbe de coût, l'étude de la croissance des entreprises était traditionnellement délaissée pour l'étude du marché. Les travaux d'Edith Penrose sont parmi les premiers à défendre qu'une firme doit être étudiée comme un ensemble dont les ressources uniques et leur partage sont un facteur de croissance (Penrose & Penrose, 2009).

Parmi ces ressources uniques, on retrouve les systèmes d'information qui ont un rôle capital dans la croissance d'une organisation. La littérature en management stratégique considère une nouvelle perspective de l'entreprise centrée sur la connaissance (Snyder et

[1] Afin de simplifier la lecture de ce mémoire, nous garderons l'expression « Data Regulation Risk » ou celle de son acronyme « DRR » dans sa version originale anglaise. Une traduction française possible serait « risque lié à la réglementation de donnée ».

al.,1998), (Nonaka & Takeuchi, 1995), (Spender, 1996). C'est dans la connaissance que réside les critères différenciants d'une entreprise et donc la clé de sa croissance à long terme.

La gestion de la connaissance par l'utilisation de systèmes d'information performants et adaptés vient expliquer les divergences de croissances entre des entreprises se développant dans des industries et conditions similaires. Cependant, la seule existence de connaissances et de savoirs au sein d'une entreprise ne permet cependant pas de garantir et de maintenir un quelconque avantage compétitif. En effet, d'après Alavi et al. (Alavi & Leidner, 2007) la capacité d'une firme à manipuler, stocker et distribuer sa connaissance existante pour capitaliser sur celle-ci prévaut sur la « quantité et la qualité de connaissance » disponible à un moment donné.

Les technologies de l'information et de la communication doivent alors être employées à cet effet en synthétisant, manageant des quantités importantes de connaissances. Si l'on se réfère à l'approche classique des systèmes d'information dans le contexte d'un traitement de donnée (Alavi & Leidner, 2007), (Ghernaouti, 2016), celui-ci peut être divisé en quatre axes principaux qui sont :

- La recherche et l'extraction de données
- Le traitement, la mise à jour et l'exploitation de données pour créer de l'information
- La distribution et communication de données et/ou information
- Le stockage

Le National Institute of Standards and Technology (NIST) (National Institute of Standards and Technology, 2020) définit les systèmes d'information comme « un ensemble discret de ressources d'information organisées pour la collection, le traitement, la maintenance, l'utilisation, le partage, la diffusion ou la disposition des informations ». Ceux-ci sont des systèmes sociotechniques composés de sous-systèmes sociaux et techniques. La structure organisationnelle et les individus liés au système d'information forment le sous-système social alors que les processus fonctionnels (aussi nommé processus d'affaire) forment le sous-système technique en impliquant les composants technologiques (Piccoli, 2007).

La « Law of Limits » de Warfield (Mar, 1997) présente les délimitations des systèmes et la nécessité de données entrantes qualitatives (inputs) pour fournir des résultats qualitatifs (outputs). Il est alors impératif de toujours optimiser l'utilisation des données et particulièrement leur mise à jour afin de pouvoir atteindre les objectifs espérés. Les

organisations ne peuvent espérer des résultats qualitatifs et pertinents que si elles mettent en place des solutions de gestion de données efficaces.

Conscientes de ce fait, les organisations considèrent aujourd'hui les données créées, collectées, traitées et diffusées au cours de leurs activités comme des ressources clés et un levier potentiel de croissance. Il en résulte l'émergence de l'étude du management de la connaissance dans les branches de management stratégique qui repose grandement sur les travaux de Penrose (Penrose & Penrose, 2009) tout en venant compléter ceux-ci.

De plus en plus, l'importance est donnée au capital reposant sur la connaissance au détriment du capital reposant sur la propriété (Miller & Shamsie, 1996). Celui-ci permet en effet d'obtenir un avantage compétitif long terme sous réserve de la capacité d'une entreprise à pleinement utiliser les connaissances disponibles et à en créer de nouvelles. La prochaine partie présente les principes de création et gestion de la connaissance.

1.1.2. Création et gestion de la connaissance

La théorie de la connaissance définie la connaissance et sa gestion comme fonction de productivité, de compétitivité, de différenciation d'une firme et par conséquent une fonction de sa croissance (Alavi & Leidner, 2007). Les technologies de l'information et de la communication sont donc cruciales pour garantir l'efficience de la gestion de la connaissance d'une entreprise, son amélioration continue et au final pour assurer la pérennité de celle-ci.

L'approche contemporaine de la littérature en informatique et technologie de l'information est de distinguer la connaissance de l'information et de la donnée. Cette distinction permet de justifier un intérêt nouveau pour la gestion de la connaissance qui diffère de la gestion de l'information ou encore de la donnée (Fahey & Prusak, 1998). Cette approche introduit notamment une hiérarchie entre la donnée, l'information et la connaissance.

Traditionnellement, la connaissance repose sur l'information, qui elle-même repose sur la donnée. Les travaux de Zimmermann & al. (Zimmermann et al., n.d.) représente cette hiérarchie sous forme pyramidale à trois niveaux. La conceptualisation et la catégorisation des données à la base de cette pyramide permettent d'atteindre l'information. La contextualisation et personnalisation de l'information permet alors d'atteindre la connaissance qui représente le sommet de la pyramide.

Il y a néanmoins de nombreuses définitions et divergences quant aux définitions de donnée, information et connaissance. Certains définissent l'information comme des données interprétées en cadriciel et la connaissance comme de l'information authentique et perçue comme vraie (D. Vance, 1997). D'autres considèrent la donnée comme des chiffres bruts et des faits, l'information comme de la donnée traitée et la connaissance est de l'information actionnable (Maglitta, 1996). D'autres défendent que la différence entre information et connaissance réside dans la confiance d'un individu receveur. Celui-ci doit avoir confiance en la source d'une information pour la qualifier de connaissance.

Bien que largement utilisé, le principe de hiérarchie entre la donnée, l'information et la connaissance est également source de nombreuses critiques (Alavi & Leidner, 2007), (Mcqueen, 1998), (Tuomi, 2000), (Newell et al., 1999). L'une des raisons de ces critiques résulte du constat que la présomption de hiérarchie varie en fonction du contexte, de l'interprétabilité et de son utilité finale (Alavi & Leidner, 2007). Le critère différenciant entre la connaissance et l'information ne se trouverait alors pas dans le contenu, la structure, utilité ou encore la confiance en une information ou connaissance. Il se trouverait plutôt dans l'esprit des individus. Ce serait le fait de personnaliser une information qui en ferait une connaissance. Cette approche ne prend plus en compte les critères précédents mais repose sur l'observation, le jugement et l'interprétation des individus. Celle-ci est en cohérence avec la conceptualisation de la connaissance de Churchman (Churchman, 1971) qui avance qu'elle réside dans les utilisateurs et non dans la collection de l'information. Il découle de cette définition que les systèmes de gestion de connaissances devraient se rapprocher des systèmes d'informations traditionnels tout en permettant aux utilisateurs d'individuellement transformer de l'information en connaissance.

S'opposant au courant occidental traditionnel qui prône les « connaissances dures et quantifiables» telles que les données brutes, les procédures codifiées et qui reposent sur l'efficience, les coûts et les retours sur investissement, Nonaka (Nonaka & Takeuchi, 1995) a étudié le management de la création de connaissances dans les firmes japonaises. Cette étude avance la combinaison de deux types différents de connaissances : tacites et explicites. Les premières sont personnelles, issues de l'expérience d'un individu et difficiles à partager. La pertinence des connaissances tacites est alors dépendante du contexte de son utilisation. Elles comprennent à la fois des éléments cognitifs et techniques propre à chaque individu. Les connaissances explicites sont quant à elles articulées, systématiques et formelles. A ces deux types de connaissances, le travail de Nonaka met également en avant l'existence de

connaissances individuelles et sociales. On trouve une forme similaire de classification de la connaissance à quatre dimensions dans la matrice des types de connaissance de Spender (Spender, 1996). Cette dernière différencie les connaissances organisationnelles implicites et explicites ainsi que les connaissances individuelles et collectives (« social »).

Considérant le potentiel de croissance pour les organisations qui repose sur les données et par extension les systèmes d'information, il est alors naturel de se poser les questions de leur sécurisation. Dans la suite de cette section, nous examinons les principes de sécurité de l'information et de cybersécurité.

1.2. … A la sécurité de l'information et la cybersécurité

L'importance et la criticité des systèmes d'informations nous poussent à nous questionner quant aux moyens de les sécuriser, d'assurer leur bon fonctionnement et de les pérenniser afin de capitaliser sur leur utilisation et le coût de leur mise en place. Les données et les systèmes qui les gèrent n'ont pas tous la même importance, les mêmes besoins, les mêmes objectifs ni les mêmes contraintes ou valeurs perçues pour leur propriétaire à un moment donné. Dans cette partie, nous détaillons en premier les objectifs de sécurité de l'information avant de s'intéresser au concept de cybersécurité.

1.2.1. La sécurité de l'information

Il existe de nombreux contrôles et moyens techniques qui peuvent être implémentés au sein d'une organisation pour assurer la sécurité de l'information et de ses systèmes d'information. Traditionnellement, la sécurité informatique repose sur le maintien de la triptyque :

- Confidentialité : le fait de préserver les restrictions d'accès et de divulgation au travers de moyens de protection de donnée personnelle ou confidentielle (National Institute of Standards and Technology, 2020) ;
- Intégrité : le fait d'assurer que les données (ou systèmes) n'ont pas été modifiées de manière non autorisée. L'intégrité des données couvre entre autre les données stockées, en cours de traitement et en transit (National Institute of Standards and Technology, 2020) ;

- Disponibilité : le fait d'assurer qu'une donnée (ou système) est accessible et utilisable à la demande d'un utilisateur autorisé (National Institute of Standards and Technology, 2020).

Venant compléter le célèbre modèle CIA (Confidentiality, Integrity, Availability), d'autres points d'attention sont de plus en plus cités tels que l'authentification, la non-répudiation (impossibilité de nier les faits ou actions), l'imputabilité (responsabilité des faits ou actions) ou encore la traçabilité. Ces fonctions additionnelles permettent, par le biais de procédures de contrôle d'accès, d'assurer que seuls les acteurs identifiés et authentifiés sont en mesure de réaliser certaines actions ; que ces actions peuvent être analysées et qu'elles ne peuvent pas être réfutées (Ghernaouti, 2016).

L'intérêt donné à la sécurité d'une donnée est directement liée à son cycle de vie et sa valeur qui ne nécessitera donc pas les mêmes efforts à différentes périodes du cycle. Si l'on prend des données relatives à un brevet technologique, celles-ci voient leur importance grandir exponentiellement jusqu'au dépôt du dit brevet. En effet, garantir leur confidentialité est impératif pendant la période pré-brevet avant de devenir optionnel une fois le brevet déposé. La classification des données est alors un outil nécessaire pour garantir et optimiser la sécurité de celles-ci tout au long de leur cycle de vie. Pareillement, Ghernaouti (Ghernaouti, 2016) regroupent les technologies de l'information et de la communication en plusieurs sphères d'activité dont chacune est concernée par la sécurité de l'information. Celle-ci se décline alors en :

- Sécurité matérielle, physique et environnementale ;
- Sécurité logique ;
- Sécurité applicative ;
- Sécurité de l'exploitation ;
- Sécurité des réseaux.

La technologie n'est cependant pas l'unique point d'attention pour maintenir et garantir la sécurité de l'information. L'aspect humain de la sécurité de l'information doit également être adressé (Sohrabi Safa et al., 2016). Il est en effet nécessaire de considérer l'erreur humaine, les comportements opportunistes néfastes ou encore le manque de confiance (Flowerday & von Solms, 2006). Les organisations doivent alors inculquer une culture de la sécurité à leurs employés au travers de formations, de politiques internes et de culture d'entreprise.

Le positionnement, la direction et l'ambition de l'organisation quant à sa sécurité informatique sont traditionnellement regroupés dans un jeu de documents appelés politiques. Ce sont ces politiques qui retranscrivent la gouvernance de chaque organisation. L'International Information Systems Security Certification Consortium ((ISC)²)[2] définit la gouvernance par « le processus de gestion d'une organisation. Cela inclut tous les aspects de la façon dont les décisions sont prises pour cette organisation et peut (et inclut généralement) la politique, les rôles et les procédures que l'organisation utilise pour prendre ces décisions. ».

En sécurité de l'information, la gouvernance est traditionnellement décrite dans un document intitulé politique de sécurité de l'information. Celui-ci est généralement complété par d'autres politiques, standards, procédures ou code de conduite (Ghernaouti, 2016) .

La criticité et l'entendue de ce jeu de documents peut grandement complexifier sa rédaction, son déploiement et son maintien dans le temps. Les organisations peuvent se tourner vers de nombreux standard internationaux dont chacun adresse des éléments clés des politiques de sécurité de l'information (Höne & Eloff, 2002). Stanton et al. (Stanton et al., 2005) décrivent cette démarche qui permet d'influencer positivement « l'hygiène basique » et les « erreurs naïves » des employées. L'hygiène basique englobe les comportements qui ne requièrent pas d'expertises techniques mais qui incluent des intentions de préserver et protéger les ressources informatiques de l'entreprise. Les erreurs naïves sont des comportements qui ne requièrent pas ou peu d'expertise technique et qui n'ont pas d'intention de causer du mal à l'entreprise. Selon Bagchi et al. (Bagchi & Udo, 2003), en moyenne les violations de politiques de sécurité sont la cause d'une faille de sécurité par an pour les organisations et plus de la moitié des failles sont directement ou indirectement causées par un manquement des employées quant au respect des procédures de sécurité (Stanton et al., 2005). Selon Vance et al. (A. Vance et al., 2012), l'adhérence des employées aux politiques de sécurité est donc primordiale et les entreprises doivent s'assurer que ceux-ci prennent conscience des risques et menaces auxquels ils doivent collectivement faire face.

L'utilisation des données faites par les organisations évolue et pose alors le problème d'évolutivité et d'adéquation des mesures de sécurité. L'intensification de l'utilisation d'internet ou encore la multiplication des sources et formats des données utilisées en des

[2] L' International Information Systems Security Certification Consortium ((ISC)²) est une organisation à but non lucratif chargée de certifier des professionnels dont le site est accessible via ce lien : https://www.isc2.org/about.

volumes conséquents viennent défier les mesures traditionnelles comme le montre les travaux de Tang et Pan (Tang & Pan, 2015). Le maintien de la confidentialité, l'intégrité ou la disponibilité des données peut difficilement reposer sur une stratégie de sécurité traditionnelle. Une approche plus large parait nécessaire. Cette approche s'incarne dans le concept de cybersécurité que nous détaillons dans la partie suivante.

1.2.2. La cybersécurité

Du fait de l'importance et de la valeur de l'information, il est légitime de se poser la question de sa sécurisation. En effet, le nombre de failles de sécurité informatique augmente chaque année (Lukasik, 2000) et celles-ci restent peu communiquées (Ullman & Ferrera, 1998), couteuses et difficiles à estimer (Martin et al., 2014). L'augmentation relative des incidences d'attaque oblige les organisations et les gouvernements à trouver des moyens de les prévenir ou de les réduire (Bagchi & Udo, 2003).

Le paysage des menaces liées à l'utilisation de technologies ou aux activités et à l'environnement d'une organisation évolue rapidement. Par manque de mesures efficaces, d'information ou d'outils, les organisations peinent encore à correctement comprendre et estimer les cybermenaces et les préjudices qui leur sont associées.

Agrafiotis et al. (Agrafiotis et al., 2018) proposent une classification des préjudices liés à l'utilisation de technologies en cinq catégories :

- Préjudice physique ou digital ;
- Préjudice économique ;
- Préjudice psychologique ;
- Préjudice réputationnel ;
- Préjudice social et sociétal.

On peut voir depuis peu une utilisation massive du terme cybersécurité malgré le fait que celui-ci souffre d'un manque de consensus quant à sa définition. Il est encore trop souvent utilisé pour définir la sécurité de l'information (von Solms & van Niekerk, 2013). La cybersécurité qui est désormais considérée comme un pilier de la sécurité informatique (Bier & Lin, 2013), (Schatz et al., 2017) est fondée sur la définition de la cybernétique de Norbert Wiener (Wiener, 2014). Elle est aujourd'hui un sujet d'importance pour de nombreuses entreprises, gouvernements et individus. Certains définissent la cybersécurité par l'intérêt de

l'humain quant à la sécurité (Kemmerer, 2003), prennent en considération l'environnement cyber et le contrôle de la propriété ou du fonctionnement de cet espace (Canongia & Mandarino, 2012) tandis que d'autres se concentrent sur le contrôle de l'information et de ses systèmes (Amoroso, 2010).

Craigen et al. (Craigen et al., 2014) définissent la cybersécurité comme « the organization and collection of resources, processes and structures used to protect cyberspace and cyberspace-enabled systems from occurrences that misalign *de jure* from *de facto* property rights ».

Celle-ci permet de mettre en avant l'interdisciplinarité de la cybersécurité et l'étendue de son champ d'action plus vaste que la sécurité de l'information (von Solms & van Niekerk, 2013), (Collard et al., 2017). En omettant volontairement de spécifier les ressources, processus ou structures, cette définition permet de mettre en avant les interactions entre humains, les interactions entre machines et enfin, les interactions entre humains et machines. Les auteurs précisent également que, dans le cadre de la cybersécurité, la protection doit prévoir de traiter l'ensemble des événements intentionnels, accidentels et dangers naturels. La définition sous-entend également que certains de ces événements sont imprévisibles. Enfin, cette définition permet d'intégrer les notions de contrôle et de propriété en précisant que tout événement ou activité qui écarte les droits de propriété réels (*de facto*) des droits de propriété perçus (*de jure*), que ce soit intentionnellement ou par accident, connu ou inconnu, constitue un incident de cybersécurité.

La cybersécurité ne s'arrête alors pas à la protection des données et de l'information. Elle incorpore également la protection d'actifs non basés sur l'information dès lors qu'ils sont vulnérables à des menaces liées aux technologies de l'information et de la communication.

La structure organisationnelle et les individus qui sont liés au système d'information forment le sous-système social du système d'information. Ils peuvent alors être considérés comme des actifs sous le prisme d'une organisation. Il en demeure qu'ils ne sont pas « basés sur l'information » à la différence du sous-système technique qui implique les composants technologiques. La cybersécurité s'étend alors au-delà de la sécurité de l'information et de la sécurité des technologies de l'information et communication.

La Figure 1 Les relations entre la sécurité de l'information, la sécurité des technologies de l'information et communication, et la cybersécurité issue du travail de Solms et Van Niekerk (von Solms & van Niekerk, 2013) met en avant l'étendue du champ de la cybersécurité.

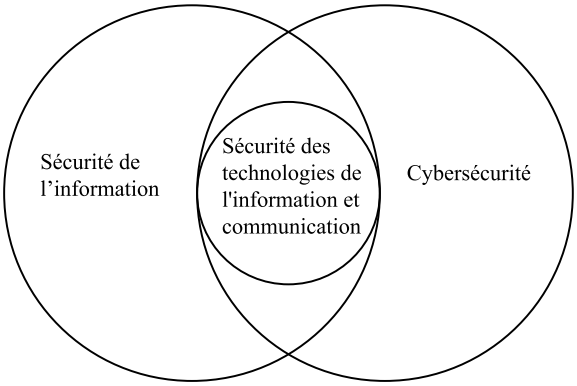


Figure 1 Les relations entre la sécurité de l'information, la sécurité des technologies de l'information et communication, et la cybersécurité

La cybersécurité est non seulement devenue un sujet d'intérêt pour les gouvernements, les politiques et les régulateurs, mais également un enjeu stratégique pour les entreprises (Ciolan, 2014). En 2012, il y avait déjà plus de 50 pays qui avaient publiés des documents démontrant leur positionnement, leur intérêt et l'importance que les politiques accordaient à la cybersécurité, au cyberespace et aux cybercrimes (Klimburg, 2012).

Un exemple de cet intérêt est la récente parution du Code de la cybersécurité, qui rassemble des contributions d'universitaires et de praticiens de la cybersécurité, en France en 2022 (Collectif Dalloz, 2022). Ce code présente les textes en vigueur (nationaux, européens ou internationaux), des commentaires, ainsi que des annotations bibliographiques et jurisprudentielles de nature à traiter juridiquement la menace et les conséquences d'une cyberattaque. Il s'adresse notamment à tous les acteurs de la cybersécurité tels que : « délégué à la protection des données, responsable de la sécurité informatique, membre des forces de l'ordre, mais aussi, avocat, magistrat, juriste d'entreprise ou professionnel d'une administration centrale ou d'une collectivité territoriale ».

La première partie de cette section nous a permis de mettre en avant l'importance de la gestion de l'information pour les organisations. Le positionnement stratégique des technologies de l'information et de la communication au sein de celles-ci implique pour ces dernières la nécessité de pouvoir leur faire confiance. Autrement dit, les risques liés aux technologies de l'information et de la communication ont accru conjointement avec leur utilisation. Pour la suite de ce mémoire, nous utiliserons la définition de Craigen et al. (Craigen et al., 2014) pour le terme cybersécurité. Nous allons maintenant voir comment la confiance joue un rôle dans l'adoption et l'utilisation de système d'information.

1.3. Un concept central, la notion de confiance : définition et spécificités dans le contexte IT

Dans cette partie, nous détaillons en premier les relations entre la gestion de l'incertitude et la confiance. Nous présentons dans un second temps comment la sécurité informatique, la confidentialité de l'information et la fiabilité des systèmes (informatiques) peuvent répondre au besoin de confiance des organisations.

1.3.1. Le management de l'incertitude et la confiance

La définition du risque de Knight est l'une des plus populaires dans la sphère scientifique (Holton, 2004). Écrite durant une période où la recherche se concentrait sur la compréhension et l'étude de la probabilité, Knight (Knight & Kelley, 1964) distingue deux types de probabilité : la probabilité « a priori » dérivée de symétrie ou de récurrence d'évènement et la probabilité « statistique » déduite d'analyse de données.

La recherche s'est penchée sur la question de la subjectivité contre l'objectivité de la probabilité. Les interprétations objectives défendent que la probabilité d'un risque soit réelle et qu'elle puisse être calculée grâce à la logique et les statistiques. Les interprétations subjectives sous-entendent au contraire que la probabilité est déduite des croyances des individus. Son utilité s'en trouve par conséquent limitée car elle ne permet au final que de quantifier l'incertitude perçue (Holton, 2004). Le hasard, et donc la chance d'occurrence d'un risque serait basé sur l'ignorance et par conséquent la détermination de la cause réelle de tout évènement et de notre compréhension de celui-ci sont affectés (Baggini, 2002).

D'après Knight (Knight & Kelley, 1964), lorsqu'un individu manque de données ou de symétrie pour calculer la probabilité d'occurrence d'un risque, celui-ci tend à quantifier la probabilité en se basant sur sa propre incertitude. La multitude d'éléments à prendre en compte lors du calcul d'une probabilité concernant un risque relatif à une organisation et son cœur de métier est telle, qu'il est alors impossible de le réaliser. Le terme risque revient alors à désigner une incertitude qui peut être calculée et incertitude, une incertitude qui ne peut l'être.

La théorie du portfolio pose les fondations du risque lié à l'investissement. Développée par Markowitz en 1952 (Markowitz, 1952), elle suggère qu'un investisseur considère les retours sur investissements espérés comme désirables et la variance comme indésirable. Le risque est alors un terme désignant la variance du retour sur investissement. Il déclare également que les analyses d'optimisation doivent prendre en considération et « combiner les techniques statistiques et le jugement pratique d'un individu ». On peut alors se demander si l'analyse des risques et le management du risque qui en découle peuvent être considéré comme une analyse d'optimisation d'investissement. En effet, investir pour minimiser l'impact ou l'occurrence d'un risque ne promet pas de retours sur investissements à proprement parlé mais limite des dommages éventuels.

Ce besoin croissant de confiance aux systèmes d'information justifie que les organisations investissent pour garantir leur sécurité et par conséquent leur maintien et leur résilience (Haimes, 2009). La gestion des vulnérabilités et du bon fonctionnement des systèmes d'information permet non seulement de supporter les différentes fonctions d'une organisation et d'atteindre les objectifs fixés mais également d'assurer la continuité des échanges et la confiance avec les acteurs tiers. Cette continuité est clé pour minimiser l'incertitude dans le temps entre les différents acteurs d'un marché (Mayer et al., 1995).

1.3.2. La confiance, un enjeu pour les organisations

Comme énoncé par Gefen et al. (Gefen et al., 2003), la confiance évolue et peut grandir au fil du temps et des échanges. Les organisations sont alors incitées à assurer une gestion des vulnérabilités de leurs systèmes. Haimes (Haimes, 2006) définit une vulnérabilité comme une « référence aux états inhérents d'un système donné (par exemple, physique, technique, organisationnel et culturel) qui peuvent être exploités par un adversaire pour affecter négativement (causer ou endommager) ce système ». Une autre manifestation de l'état d'un système est la résilience (Haimes, 2009). Celle-ci peut être définie comme la capacité à absorber un stress d'origine externe (Holling, 1973), la capacité à prévoir, à reconnaître, à anticiper et à se défendre contre la forme changeante du risque avant que des conséquences négatives ne surviennent (Woods, 2017), ou encore la capacité inhérente et les réponses adaptatives des systèmes permettant d'éviter les pertes potentielles (Rose & Liao, 2005).

La confiance peut alors être considérée comme une exigence pour garantir la stabilité d'une relation en impactant l'appétence pour le risque des parties (Mayer et al., 1995). Les causes, natures et effets de la confiance sont la source de nombreux travaux de recherche à travers diverses disciplines. Divers concepts de la confiance ont été développés résultant en des définitions distinctes en fonction du domaine d'application et de la discipline de recherche (Gambetta, 2000). Parmi elles, les définitions opérationnelles et internes de la confiance dominent. La première se rapporte à la théorie des jeux (Myerson, 1991) impliquant un processus décisionnel rationnel et une aversion au risque basée sur les gains et les pertes prévus. Ce dernier décrit la confiance comme un état de croyance se référant à son acceptation de la vulnérabilité basée sur sa croyance ou ses attentes positives en ce qui concerne les motivations et les comportements des autres (Flowerday & von Solms, 2006). La baisse de l'incertitude est alors corrélée avec l'augmentation de la communication et l'échange d'information (Pearce, 1974).

Camp (Camp, 2002) présente le concept tri-dimensionnel de la confiance « *three dimensional concept of trust* » qui définit la confiance comme l'intersection de la « *privacy, security and reliability* ». En se concentrant sur l'existence même du risque et non sa quantification, cette définition que l'on peut qualifier d'opérationnelle se base sur le risque et non sur la perception du risque.

D'après le concept tri-dimensionnel de la confiance, la sécurité n'est pas la confidentialité « privacy » mais un moyen de garantir la confidentialité en fournissant les outils et moyens pour contrôler les données et l'information.

Prise à part, la sécurité seule ne permet donc pas de garantir la confiance ou la confidentialité (Flowerday & von Solms, 2006). Avec un raisonnement similaire, on peut conclure que la sécurité n'équivaut pas nécessairement à la fiabilité « reliability » mais constitue un moyen d'y parvenir et donc de contribuer à la perception d'intégrité et d'autorité des parties de confiance (Camp, 2002).

Enfin, on peut affirmer que la sécurité n'est pas dissociable de la confiance et que cette dernière nécessite une grande compréhension des interactions et des motivations des différentes parties. Cela est d'autant plus impactant dans le contexte d'informatique en nuage « Cloud computing », de services Cloud ou de traitement des données externalisés.

Bacharach et Gambetta (Collegio & Alberto, 2001) définissent l'action de faire confiance comme "une personne fait confiance à une autre pour faire X si elle agit dans l'attente

qu'elle fera X quand tous deux savent que deux conditions sont remplies : si elle ne fait pas X, elle aurait mieux fait d'agir autrement, et agir comme elle le fait lui donne une raison égoïste de ne pas faire X ». Dans une telle situation, la confiance peut alors se définir comme la croissance positive de la prédictibilité. La première condition implique que la personne qui accorde sa confiance accepte de s'exposer alors que la seconde condition implique une tentation de tirer profit de la violation de la confiance. Cela établit que la confiance dépend de trois conditions : l'incertitude, l'exposition et la tentation (Guerra et al., 2003). Celle-ci est alors atteinte lorsque les niveaux de sécurité, résilience, transparence et responsabilité sont acceptables pour les parties.

Nous pouvons donc traduire la confiance par l'ambition d'atteindre un niveau de risque perçu suffisamment faible pour qu'il soit acceptable pour une organisation d'utiliser les services d'un tiers ou pour un utilisateur final d'accepter le partage et la gestion de ses données personnelles.

Atteindre un état absolu de confiance n'est dans les faits pas réalisable. L'existence d'une incertitude est causée par l'incapacité générale d'atteindre une compétition pure et parfaite. Les organisations sont alors forcées de réduire cette incertitude en augmentant la prédictibilité. Elles peuvent y parvenir par leurs engagements, transparences et mesures de sécurité (Humphrey & Hubert Schmitz, 1998). Il est alors nécessaire pour elles de développer et de mettre en place des contrôles pour se prévenir d'évènements non désirés afin de renforcer la confiance de leur clients (Flowerday & von Solms, 2006).

Dans la continuité de cette volonté et face à la nécessité de mettre en place des contrôles et divers engagements pour assurer la confiance des différentes parties impliquées dans la gestion des données, les législateurs et responsables politiques ont promulgué de nombreuses lois afin de garantir la protection des droits et de la vie privée des individus, les intérêts économiques et la sécurité nationale dans chaque région du monde (Kosseff, 2017).

Dans la prochaine partie, nous détaillerons comment les réglementations peuvent être utilisées comme outils de gestion de l'incertitude. Nous étudions également comment celles-ci sont paradoxalement à l'origine d'un nouveau risque que nous appelons risque réglementaire.

1.4. Les réglementations comme outil pour la gestion de l'incertitude

D'après Kosseff (Kosseff, 2017), l'apparition de nouvelles lois, qu'il nomme lois de cybersécurité (Cybersecurity law), répond au but premier de garantir la protection des droits et de la vie privée des individus, les intérêts économiques et la sécurité nationale. Celles-ci ont pour ambition d'assurer la confiance des acteurs publics et privés du marché en garantissant la confidentialité, l'intégrité et la disponibilité des informations, des systèmes d'information et des réseaux.

1.4.1. Le rôle des réglementations

Par l'utilisation des réglementations et autres éléments du corpus juridique, les législateurs et politiques assurent le niveau souhaitable de confiance au sein de leur marché. Ils tentent d'atteindre ce niveau par entre autre le contrôle et l'encadrement des traitements des données et l'utilisation de certaines technologies et services.

Souvent perçues comme couteuses, contraignantes ou encore comme un frein au développement des entreprises, les règlementations sont cependant également bénéfiques. En outre, elles imposent un cadre structurant pour les organisations au travers de contraintes de gestion, classification, protection et minimisation des données.

La croissance du nombre de réglementation peut s'expliquer avec la théorie de l'acteur stratégique (Crozier & Friedberg, 1977). Celles-ci auraient alors pour but premier de limiter ou d'encadrer les stratégies individuelles des acteurs (les organisations) pour prévenir des comportements et des intérêts individuels parfois contradictoires. Elles peuvent alors limiter les risques liés à une utilisation inefficiente des systèmes d'information en imposant un niveau minimum de qualité par la réglementation des marchés.

Une utilisation inefficiente des systèmes d'information correspond, par exemple, à la collecte, traitement et stockage des données sans réel besoin et valeur ajoutée. Cette multiplication des données entraine in fine une complexité de recherche et d'extraction de valeur ainsi que des coûts augmentés de gestion. En somme, des réglementations telles que le Règlement Général sur la Protection des Données (RGPD) (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council*

of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1., n.d.) peuvent alors s'apparenter à des cadres permettant une optimisation de la gestion des données et donc de création de valeur à long terme. Par exemple, l'article 32 du RGPD permet d'assurer la sécurité des données à caractère personnel, l'article 35 introduit la nécessité d'effectuer des analyses de risques et l'article 15 assure la mise à jour des données.

Dans son texte, le RGPD stipule que les «différences peuvent dès lors constituer un obstacle à l'exercice des activités économiques au niveau de l'Union, fausser la concurrence et empêcher les autorités de s'acquitter des obligations qui leur incombent en vertu du droit de l'Union» [3], il est également mentionné que «pour que le marché intérieur fonctionne correctement, il est nécessaire que la libre circulation des données à caractère personnel au sein de l'Union ne soit ni limitée ni interdite pour des motifs liés à la protection des personnes physiques à l'égard du traitement des données à caractère personnel». Il y a donc une volonté forte de renforcer un marché digital unique au sein de l'Union. Cela permet en outre un accès aux marchés des différents pays simplifié par une uniformisation et une diminution des disparités réglementaires. Il y aurait donc également un intérêt économique pour l'Union Européenne au travers de ce règlement présenté comme outil de protection des citoyens européens et de leurs données personnelles.

Cependant, la volonté de l'Europe au travers du RGPD n'est pas d'assurer l'utilisation de technologies européennes ni de restreindre l'accès sur des bases de protectionnisme ou de souveraineté nationale. Cela se traduit par un certain degré de liberté pour les organisations dans leur mise en conformité. Le choix des mesures de sécurité est à la discrétion de chaque entité assujettie au RGPD dans la mesure où elles répondent à un niveau estimé « suffisant ». Cela n'est pas forcément le cas pour toutes les réglementations. En effet, certaines ont pour ambitions de garantir des objectifs économiques et de protection de marchés intérieurs. Celles-ci se traduisent par une volonté de protéger la souveraineté des données et par des limitations technologiques. Elles ne présentent pas des limitations qualitatives au sens où la réglementation limiterait l'accès aux technologies avancées et de pointes des entreprises sur son territoire. Elles

[3] Les différences font référence aux écarts dans les niveaux de protection entre les États membres.

prennent cependant la forme d'encadrement des fournisseurs et typologies de solutions répondant à certains critères.

Les réglementations peuvent également dans certains cas limiter l'export et l'utilisation de technologies issues de leur marché. En effet, les méthodes récentes de protections des données ont pour but de faciliter le contrôle des informations et technologies qui impactent les intérêts nationaux (Crook, 2009). Par exemple, le congrès américain décrit dans le « Arm Export Control Act » que le Président américain est autorisé à contrôler l'import et l'export des biens et services de défenses et est aussi en charge de fournir des directives pour les citoyens américains impliqués dans l'export ou l'import de ces biens et services (Barker, 2008).

Par l'utilisation des réglementations et autres éléments du corpus juridique, les législateurs et politiques contraignent les organisations par leur contrôle et encadrement. Ceux-ci sont à l'origine d'un risque de pénalité pour toute organisation qui manquerait de se conformer à ces obligations. La prochaine partie détaille comment la conformité crée un risque pour les organisations.

1.4.2. La conformité et son risque

En reprenant le cycle de vie de la donnée, on peut identifier les étapes clés de son traitement qui incluent sa collecte, sa consultation, son utilisation, son partage, son stockage et enfin sa suppression (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.). Les législateurs et responsables politiques qui peuvent impacter et contrôler ces différentes étapes de diverses manières, sont aussi à l'origine d'un risque règlementaire nouveau pour les organisations par l'apparition de nouvelles pénalités en cas de non-conformité.

Il est alors nécessaire pour les organisations de comprendre les différentes conditions et contraintes des lois et réglementations afin d'assurer leur conformité et d'atteindre les objectifs de celles-ci que ce soit des lois régissant les données, la chaine de production ou encore la gestion des comptes d'une entreprise. Les entreprises ont à leur disposition de nombreux « outils » pour répondre à ces contraintes règlementaires. Ces jeux de documents comprennent entre autre les cadriciels (Frameworks), standards, lignes directrices (Guidelines)

ou codes de conduite (Code of practice), contrôles (étapes mesurées à suivre pour atteindre un objectif spécifique), certifications et accréditations (Eloff & von Solms, 2000). Grâce à ces outils, elles sont en capacité de réduire leur marge d'erreur et de mauvaise compréhension des textes juridiques.

Les différents textes juridiques ne définissent pas forcément la manière dont les entreprises doivent répondre à leurs exigences ni comment implémenter les contrôles nécessaires. Ils laissent parfois à la discrétion des entreprises le choix technologique et organisationnel. Afin d'éviter des mauvaises interprétations du texte, des non-conformités et afin de donner la liberté aux organisations d'adapter continuellement leur système d'information selon les évolutions technologiques, les législateurs peuvent faire référence à des documents annexes tels que des cadriciels, certifications ou lignes directrices. Ces documents sont amenés à évoluer dans le temps et peuvent provenir de différentes entités publiques ou privées. Ils couvrent différents points nécessaires pour assurer une mise en conformité tels que les solutions techniques, le contrôle d'activité, la gouvernance, les rôles et responsabilités ou encore les méthodologies à suivre pour mener un audit.

La possibilité des pénalités prévues en cas de non-conformité ou conformité partielle force les organisations à considérer et à mettre à disposition les moyens nécessaires pour évaluer le risque réglementaire.

La conformité des organisations

Avant de nous attarder sur les moyens d'évaluation de la conformité, il est important d'apporter de la clarté quant à sa définition dans notre contexte. La conformité qui est parfois alternativement nommée compliance nous provient du mot français « complaire ». Sa définition reste complexe du fait de ses multiples interprétations comme le souligne Gaudemet (Gaudemet, 2019). En effet, la culture juridique française est attachée et s'est développée autour du concept de légalité qui renvoie au principe de règles qui dictent, interdisent ou obligent des actions et évènements. La compliance, ou conformité se différencie de la légalité par son périmètre étendue. Celle-ci englobe également la mise en œuvre de dispositifs nécessaires et efficaces pour se prévenir d'infraction des règles. En somme, la compliance prend également en compte les moyens d'assurer la légalité et non le seul fait de respecter des règles.

La conformité prend alors en considération l'impact au sein des organisations et les actions qui permettent de réduire le risque d'infraction aux règles (Gros, 2013). Allant plus loin que le concept de légalité, elle rend les organisations responsables de ne pas avoir mis en place les moyens et mesures nécessaires de prévention d'infraction aux règles juridiques.

Considérant les règles juridiques, qui sont les fondations sur lesquelles le principe de légalité repose, celles-ci peuvent prendre différentes formes. Nous nous intéressons aux modalités juridiques qu'incarnent les règles. Querler (le Querler, 1996) définit le concept de modalité comme l' « expression de l'attitude du locuteur par rapport au contenu propositionnel de son énoncé ». La modalité permet de présenter un fait énoncé comme nécessaire, possible ou vrai et permet alors l'expression de logique (KALINOWSKI, 1971). En linguistique, la notion de modalité couvre un champ étendu complexifiant sa compréhension et délimitation (Stage, 2002).

Dans le contexte de règlementations qui apportent des contraintes et exigences aux organisations quant à leur gestion de l'information et de leur système d'information, les règles juridiques peuvent être qualifiées de modalité déontique. Ce concept décrit le fait d'imposer la volonté d'un locuteur à son interlocuteur (Stage, 2002). Dans notre contexte, le locuteur correspond aux législateurs alors que l'interlocuteur fait référence aux organisations assujetties aux lois et réglementations.

En d'autres termes, les modalités juridiques sont des outils à disposition des législateurs dont la vocation est d'indiquer aux organisations le degré de possibilité de la réalisation d'un évènement ou d'une action en fonction de certaines circonstances. La compliance revient alors à la mise en œuvre de dispositifs nécessaires et efficaces répondant au degré de liberté prévu par les règles juridiques. Cette liberté peut s'illustrer par une obligation, une interdiction ou un libre choix (Stage, 2002). Un exemple simple de non-conformité serait pour une organisation de ne pas mettre en œuvre les moyens estimés nécessaires à la prévention de survenance d'un évènement interdit.

Les risques de la non-conformité

Afin d'assurer le respect et la conformité des organisations aux lois et autres textes juridiques, les législateurs et responsables politiques peuvent programmer des audits des organisations par des organismes publics ou privés. De tels audits peuvent être organisés à la discrétion d'une autorité publique ou suite à l'évidence d'une atteinte à la sécurité de données

réglementées. Ces audits permettent de déterminer si l'entreprise touchée avait mis en place les moyens et mesures nécessaires de prévention tels qu'exigés par les réglementations.

A ce jour, seuls des organismes approuvés et prévus dans les textes peuvent entreprendre de tels audits et ceux-ci n'ont pas toujours l'obligation de notifier à l'avance une organisation. Dans certains cas, les législateurs rendent responsables les entreprises de l'évaluation de la conformité de leurs fournisseurs de services en obligeant la présence de clause d'audit dans les contrats commerciaux. C'est par exemple le cas dans le RGPD, l'Export Administration Regulations (EAR) ou l' International Traffic in Arms Regulations (ITAR). Pareillement, certaines lois prévoient également des clauses d'auto-divulgation ou de dénonciation dans des délais impartis. Enfin, certains législateurs vont jusqu'à contraindre les individus à dénoncer leur entreprise en cas de non-conformité avérée.

Une fois la non-conformité mise en évidence, démontrée et connue des organismes en charge de l'application d'une loi, l'entreprise peut être sanctionnée. Selon l'interprétation de lois, la jurisprudence et les pénalités prévues par les textes, une entreprise peut être sanctionnées économiquement par une amende pécuniaire ou une interdiction d'activité dans un marché pour une période donnée. Une autre forme de sanction est l'obligation pour une entreprise de communiquer publiquement sur l'atteinte à la sécurité des données réglementées. Une fois publique, cette information peut causer des conséquences indirectes telles que la perte de confiance, impacter sa réputation et sa valeur boursière.

Enfin, certaines règlementations telles que l'ITAR ou l'EAR impactent directement les individus en engageant leur propre responsabilité s'ils sont la cause de la non-conformité de leur entreprise. Ceux-ci se retrouvent personnellement responsables pour les décisions prisent dans le cadre de leurs fonctions. Refuser de manière délibérée de se conformer ou contrevenir à une loi en connaissance de causes n'expose pas uniquement l'entreprise à des sanctions mais également l'individu à l'origine de la non-conformité.

Comme mentionné précédemment, la complexité de la gestion de ces règlementations réside dans leur périmètre étendu. Il est de ce fait nécessaire pour les organisations de traduire les contraintes et exigences réglementaires en termes techniques, organisationnels et opérationnels. Sans oublier que l'applicabilité des différents textes juridiques est spécifique au contexte d'une organisation, dépend de ses marchés, de la présence géographique et des juridictions de celle-ci, elle nécessite donc une analyse approfondie impliquant un large éventail de compétences. Ces compétences sont généralement fragmentées entre les

départements de l'organisation tels que le juridique, la sécurité des systèmes d'information, l'informatique opérationnel, finance, gestion des risques, ressources humaines, etc. En d'autres termes, une exigence réglementaire spécifique peut non seulement nécessiter une interprétation juridique, mais également l'implication de praticiens de la sécurité, de responsables informatiques, d'experts en gouvernance et métiers.

En nous basant sur le champ d'action des règlementations et notre terrain d'étude, nous avons identifié les principaux types d'acteurs qui sont :

- La direction des systèmes d'information,
- La direction des risques et de la sécurité,
- La direction juridique.

Tous trois ont besoin d'informations différentes extraites des lois pour exercer leurs fonctions tout en assurant la continuité des activités et la conformité de leur entreprise. D'autres utilisateurs peuvent être considérés tels que les membres de la direction de l'audit interne et externe ou de la direction financière.

1.4.3. L'unicité du risque réglementaire

Le risque et ses concepts sont la source de nombreux travaux de recherche pluridisciplinaires. En effet, la multiplicité des critères à prendre en considération lors de l'exercice de définition et de calcul d'un risque justifie la nécessité de multiples réflexions. Tous les concepts de risque se rejoignent dans leur précondition : la contingence des actions humaines (Renn, 2008). A chaque instant une entité (individu, organisation, etc.) fait face à plusieurs options d'action ou de non-action qui sont toutes associées à des conséquences potentiellement négatives ou positives.

Risque juridique

Parmi les travaux de recherche qui traitent du risque et de ces concepts, il existe la branche du risque juridique. Les normes Bâle II (Basel II) qui sont des recommandations sur les lois et règlementations bancaires issues par le comité de Bâle définissent le risque juridique comme : « Le risque juridique comprend, mais sans s'y limiter, l'exposition à des amendes, des pénalités ou des intérêts punitifs résultant d'actions de surveillance, ainsi que des règlements privés » (Mahler, 2007). Une des définitions du risque juridique réside également dans

l'incertitude du contexte juridique. Cette incertitude repose sur les modifications de textes juridiques existant ou encore la promulgation de nouveaux textes.

Alors que certains avancent le fait qu'une définition du risque juridique n'est pas nécessaire de par le fait que ce sont des risques que les juristes peuvent identifier et adresser (Kenny & Watson, 2004) et (Luhmann et al., 2017), Mahler (Mahler, 2007) décrit plusieurs risques juridiques selon le prisme de l'incertitude (juridique ou factuelle) et le prisme du type de texte juridique (déontique ou qualitatif).

Dans notre contexte, nous nous intéressons à des règlementations dites déontiques qui impactent les organisations et leur gestion des systèmes d'information et traitement de données. Celles-ci étant apparues il y a peu, les entreprises n'ont pas encore suffisamment de base de références adaptée au contexte pour les appréhender correctement. Le domaine bancaire reste le principal terrain pour ces recherches car il est historiquement sujet à de nombreuses crises réparties sur l'ensemble des pays (Barth et al., 2004). Cependant, celles-ci sont limitées à un domaine spécifique, une réglementation spécifique pour un cas d'étude spécifique.

Le risque de non-conformité est déterminé par deux types de facteurs qui peuvent être interne ou externe à une organisation. Dans ses travaux, Asenov (Asenov, 2015) présente sept causes de non-conformité qui peuvent être mises en avant:

- L'incapacité de se conformer aux exigences d'activité de la législation en vigueur.
- La disparité entre les textes législatifs et les procédures internes à une organisation qui résulte d'une incapacité à adapter ses activités.
- L'inefficience de l'organisation à former ses employés aux risques législatifs.
- La non-conformité de contrats existants dans lesquels l'organisation est impliquée.
- Les imperfections du système législatif d'un pays ou d'une zone (absence, contradiction, changements excessifs, supervision inadaptée ou application erronée).
- La non-conformité d'un tiers (fournisseurs de services ou de clients).
- La répartition géographique d'une entreprise qui l'assujettit à des juridictions différentes.

Ces nouvelles réglementations et leur non-conformité diffèrent des risques traditionnels par la création d'un nouveau risque que nous présentons ci-après.

Un nouveau risque: Data Regulation Risk

Les sous sections ci-dessus nous ont permis de mettre en avant que les différentes réglementations favorisent la protection des droits et de la vie privée des individus, de l'intérêt économique ou de la sécurité nationale. Ces règlementations diffèrent des autres de par leur champ d'action. Elles encadrent le traitement des données, la gouvernance et les processus internes des organisations, ainsi que l'utilisation de la technologie et des services. Enfin, ces lois et réglementations assurent leur application par divers moyens et obligent les organisations à se conformer ou à risquer la sanction d'un organisme de réglementation.

Ces règlementations prévoient deux types de sanctions en cas de non-conformité : les sanctions pour les organisations et les poursuites pénales pour les individus. Les sanctions ne peuvent être prononcées que par un organisme réglementaire, dépendent du contexte et d'une appréciation externe des textes.

Le risque, que nous appelons le risque lié à la réglementation de donnée ou Data Regulation Risk (DRR), dont l'origine provient de ces règlementations particulières est donc différent des autres risques.

En effet, le DRR ne provient pas des risques classiques liés à l'informatique tels que les violations de données mais de la possibilité d'être sanctionné suite à une non-conformité avérée. Seule la connaissance du non-respect des différentes restrictions relatives au traitement des données, à la gouvernance et aux processus internes, aux technologies et services de l'information peut entraîner une sanction. En d'autres termes, une faille de sécurité n'entraîne pas nécessairement des sanctions si le respect de la loi (légalité) est prouvé. De plus, selon la jurisprudence et les interprétations, la non-conformité n'entraîne pas nécessairement des sanctions. Par exemple, l'absence de préjudice causé par une violation de données peut ne pas déclencher de sanction RGPD. Une non-conformité qui n'entraîne pas de sanction ne peut donc pas constituer un DRR. Celui-ci découle alors de la possibilité d'une sanction et non du non-respect, ni d'une faille de sécurité ou d'une incertitude juridique.

Seules les sanctions émanant d'un organisme de réglementation peuvent être la source du DRR, bien que le risque soit basé sur des facteurs externes et internes. Les facteurs internes peuvent être la décision de ne pas se conformer à la norme, la divulgation volontaire ou le non-respect involontaire de la norme. Les facteurs externes peuvent être la divulgation d'une violation de données, le résultat d'un audit externe, etc.

Nous définissons alors le Data Regulation Risk comme « un risque provenant de la possibilité d'une sanction par un organisme réglementaire suite à la preuve de non-conformité d'une norme régissant le traitement des données et/ou les gouvernances et processus des technologies de l'information et communication et/ou les technologies et services de l'information » (Delorme et al., 2021).

Par l'analyse de cette définition, nous pouvons identifier ses implications et les prérequis nécessaires à sa gestion. La première implication est la nécessité d'une méthodologie de gestion de risque multidisciplinaire et transversale. En effet, le périmètre du DRR s'étend au-delà d'un risque juridique ou de conformité et repose également sur les risques traditionnels de l'information. La deuxième implication est le besoin d'étude du contexte de l'organisation. Ce contexte comprend non seulement l'ensemble du corpus juridique applicable mais aussi l'exhaustivité des processus d'affaire d'une organisation, la gestion et la sécurité des systèmes d'information. Enfin, pour répondre aux enjeux du DRR, une méthodologie de gestion de risque adaptée doit permettre la gestion de risque multi-scénario.

Conclusion

Cette section a présenté notre domaine d'étude : le Data Regulation Risk. L'importance grandissante des données dans les organisations a naturellement conduit celles-ci à s'intéresser à l'optimisation de leur système d'information et leur protection.

De façon similaire, il y a eu ces dernières décennies une prise de conscience chez les politiciens et législateurs dont la résultante est un nombre croissant de textes juridiques visant à encadrer le traitement des données, la gouvernance et les processus internes des organisations, ainsi que l'utilisation de la technologie et des services.

Considérant que ces nouvelles règlementations ne se limitent pas à contrôler leur infraction mais également la mise en œuvre de dispositifs nécessaires et efficaces pour s'y prévenir, il en résulte une nouvelle classe de risque juridique : le Data Regulation Risk.

Bien que celui-ci repose grandement sur un risque technologique, son périmètre est plus important que celui-ci.

La prochaine section de ce chapitre présente différentes méthodologies de gestion de risque. Son objectif est d'analyser les méthodes de management du risque afin de déterminer leur applicabilité au Data Regulation Risk.

Section II. En quoi le management du risque répond partiellement aux besoins de la gestion de la conformité des données réglementées

La première section de ce chapitre nous a permis d'étudier l'importance de la donnée et de l'information qui lui est associée. Afin d'être conformes aux lois qui sont à l'origine du DRR, il est alors nécessaire pour les organisations de se doter de moyens techniques, technologiques et humains adéquats. Ces moyens doivent alors être conjointement employés pour identifier, analyser et répondre au DRR.

On peut alors se demander si les méthodologies traditionnelles d'analyse de risque de l'information peuvent être utilisées en l'état ou nécessitent d'être adaptées afin de prendre en considération cette nouvelle classe de risque qu'est le DRR.

Cette section est organisée comme suit : la première partie étudie les principes fondamentaux de la gestion de risque et la seconde présente différentes méthodologies existantes.

2.1. Notion de management du risque

La place centrale du risque et de sa gestion au sein des organisations publiques et privées, de la recherche et des politiques a conduit à l'élaboration de multiples méthodologies. Avant de nous consacrer à leur étude et leur utilisabilité dans notre contexte, cette partie présente le concept du risque, ses définitions et les fondements de la gestion du risque.

2.1.1. Risque : concept et définition

Du fait de sa nature multidisciplinaire, l'exercice de définition du risque est complexe. Il en résulte de nombreuses définitions qui divergent selon le prisme sous lequel le risque est étudié. En effet, selon Aven et al., la pertinence d'une définition peut être influencée par son contexte d'utilisation (Aven & Renn, 2009). De plus, conjointement avec la croissance de l'importance et de l'intérêt accordé au domaine d'étude, il y a une évolution naturelle de la compréhension de ce dernier. Il en résulte une évolution de son sens qui explique l'existence de ses définitions multiples (Aven, 2012),(Gerber & von Solms, 2005) & (Douglas, 1990).

Face à cette problématique de multiplicité de définitions, Aven (Aven, 2012) propose un système de classification composé de 9 catégories :

- (1) Risque = valeur attendue (R=E) ;
- (2) Risque = probabilité d'un évènement (indésirable) (R=P) ;
- (3) Risque = Incertitude objective/mesurable (R=OU) ;
- (4) Risque = Incertitude (R=U) ;
- (5) Risque = Perte potentielle/possible (R=PO) ;
- (6) Risque = Probabilité & scénario/conséquence/sévérité (R=P&C) ;
- (7) Risque = Evènement/ conséquence (R=C) ;
- (8) Risque = Conséquence/dommage/sévérité & incertitude (R=C&U) ;
- (9) Risque = Effet de l'incertitude sur les objectifs (R=ISO).

Où R=Risque ; E= valeur attendue ; P= probabilité d'un évènement (indésirable) ; OU = incertitude mesurable ; U= incertitude ; PO = perte potentielle; C= conséquence ; ISO = effet de l'incertitude sur les objectifs.

La probabilité comme présentée dans ces définitions peut se rapporter à une mesure représentative de l'incertitude. D'après Kalpan & Garrick (Kaplan' & Garrick2, 1981), la probabilité est alors une mesure de l'état de connaissance, un degré de croyance et de confiance. Celle-ci ne doit alors pas se limiter à la fréquence d'occurrence qui résulte d'une expérience répétée.

La « probabilité fréquentiste » (Aven, 2012) sert à exprimer l'occurrence d'un évènement considérant une population infinie de situation ou scénarios similaires alors que la « probabilité subjective » sert à exprimer ce qui est connu et perçu par un individu. On parle alors de risque perçu (Kaplan' & Garrick2, 1981).

La valeur attendue correspond à la perte attendue « expected loss ». Celle-ci peut se déterminer de différentes manières. L'(ISC)² propose une méthodologie qui repose sur l'annualisation des pertes estimées (ALE) dont l'estimation provient du produit du taux annualisé d'exposition d'un bien à une perte (EF), de la valeur du bien (AV) et du taux d'occurrence annualisé de la perte (ARO).

$$ALE = AV*EF*ARO$$

Enfin, l'incertitude sert à exprimer le fait qu'un individu ne connait pas si la proposition est vraie ou fausse (Holton, 2004),(Aven, 2012). Cela revient à ne pas être conscient de la survenance de l'évènement ou non.

Ce système de classification de définition du concept du risque nous permet alors de poser les premiers critères nécessaire à sa gestion. En effet, une définition du risque se basant sur les probabilités fréquentistes (ou fréquence) nécessitera une méthodologie de gestion adaptée se basant sur un historique de résultats « d'expériences » ou d'évènements. A l'inverse, une définition du risque se basant sur les probabilités subjectives nécessitera une méthodologie laissant les gestionnaires du risque exprimer leur point de vue subjectif et reposant sur leurs connaissances et expertises.

Nous avons précédemment défini le Data Regulation Risk comme « un risque provenant de la possibilité d'une sanction par un organisme réglementaire suite à la preuve de non-conformité d'une norme régissant le traitement des données et/ou les gouvernances et processus des technologies de l'information et communication et/ou les technologies et services de l'information ».

Par déduction, nous pouvons écarter les catégories (3) R= OU ; (4) R=U ; (8) R= C&U) et (9) R=ISO qui reposent sur l'incertitude. En effet, si l'incertitude permet d'exprimer l'état de conscience de survenance d'un risque, alors celle-ci est incompatible avec le concept de conformité.

Les catégories (2) R=P ; (5) R=PO et (7) R=C peuvent partiellement s'appliquer à notre définition du DRR. Celles-ci correspondent à l'occurrence d'un évènement qui est la sanction d'une organisation par un organisme règlementaire. La limite de ces catégories est le manque de précision et de prise en compte des différents scénarios nécessaires pour représenter les conditions du DRR. En effet, le DRR ne se limite pas à l'occurrence d'un évènement (sanction d'une organisation par un organisme règlementaire) mais prend en considération l'état de conformité, la nécessité de preuve et l'applicabilité des règlementations.

Si l'on considère qu'un risque correspond à la probabilité de différents scénarios, de leur conséquence respective et de la sévérité de chaque conséquence, il est alors possible de représenter pleinement le DRR. Partant de ce postulat, la catégorie (6) R=P&C est celle qui s'applique le plus à notre définition.

Dans une optique de représentation du risque multi-scénario, Kalpan & Garrick (Kaplan' & Garrick2, 1981) proposent une méthodologie de calcul qui prend en compte ce qui peut arriver (les différents scénarios), l'occurrence de ces scénarios et leur conséquences.

$$R=\{<S_i,P_i,X_i>\},\ i=1,2,\ldots,N$$

S_i = scénario ; P_i,= probabilité du scénario ; X_i = conséquence du scénario

0 Classe de risque	(1) R=E	(2) R=P	(3) R=OU	(4) R=U	(5) R=PO	(6) R=P&C	(7) R=C	(8) R=C&U	(9) R=ISO
DRR	-	+	-	-	+	++	+	-	-

R=Risque ; E= valeur attendue ; P= probabilité d'un évènement (indésirable) ; OU = incertitude mesurable ; U= incertitude ; PO = perte potentielle; C= conséquence ; ISO = effet de l'incertitude sur les objectifs

- Non applicable
+ Partiellement applicable
++ Applicable

Tableau 1 Applicabilité des catégories de définitions du risque au DRR

Le Tableau 1 Applicabilité des catégories de définitions du risque au DRR reprend les différentes catégories de définition du risque et leur applicabilité au DRR. Celui-ci nécessite une méthodologie de gestion de risque qui permet d'illustrer différents scénarios, de représenter leurs impacts et leur probabilité d'occurrence.

Nous avons vu précédemment que la non-conformité implique que les organisations soient responsables de ne pas avoir mis en place les moyens et mesures nécessaires de prévention d'infraction aux règles juridiques. Il en convient alors qu'« un risque provenant de la possibilité d'une sanction par un organisme réglementaire suite à la preuve de non-conformité » répond aux critères de la probabilité subjective.

La mise en place de moyens et mesures nécessaires de prévention d'infraction aux règles juridiques revient à impacter la possibilité (ou probabilité) d'occurrence du DRR. La gestion d'un risque multidisciplinaire et multi-scénario tel que le DRR nécessite une méthodologie de management du risque adaptée. La prochaine partie étudie les fondamentaux du management du risque.

2.1.2. Fondamentaux du management du risque

La gestion du risque autrement appelée le management du risque (RM) correspond à un ensemble de différentes étapes qui permettent pour une organisation d'identifier, d'évaluer et de répondre à un risque (National Institute of Standards and Technology, 2018a). Comme souligné par Myagmar et al. (Myagmar et al., 2005), l'objectif du RM (et de la sécurité) n'est pas de garantir un environnement libéré de tout risque mais d'arriver à un état où l'ensemble des risques sont acceptables et acceptés. Le RM est alors utilisé pour répondre à des risques jusqu'à ce qu'ils soient suffisamment réduits et acceptables.

Outre l'acceptation, les réponses au risque sont la mitigation, le transfert et l'évitement d'un risque. La mitigation de risque correspond à la mise en place de mesures et contrôles afin de réduire la probabilité de survenance d'un risque ou sa conséquence (impact). Transférer un risque peut se faire par la collectivisation de celui-ci (assurance) ou son transfert lors d'une contractualisation. L'évitement d'un risque correspond à refuser l'existence du scénario du risque. Cela peut être par exemple le fait de ne pas effectuer une action résultant en un scénario dont la probabilité d'occurrence est nulle.

Enfin l'acceptation du risque correspond à une non-action. En d'autres termes, l'acceptation du risque revient à ne pas le mitiger, transférer ou l'éviter mais de prévoir et d'assumer pleinement le risque et les conséquences associées. L'acceptation survient généralement à la suite des autres réponses au risque lorsque le risque résiduel devient acceptable pour une organisation. Pour reprendre des termes génériques et populaires, suite à la réduction du risque par une organisation, cette dernière « prend le risque » et accepte le risque résiduel.

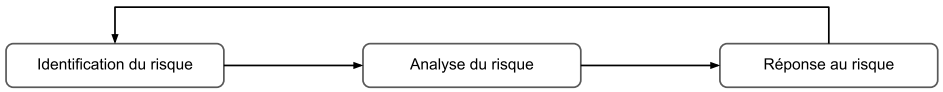


Figure 2 Etapes principales du management du risque

Comme illustré par la Figure 2 Etapes principales du management du risque, le RM est un cycle continue d'identification, analyse et réponse à un risque. Des étapes additionnelles peuvent être ajoutées selon la méthodologie employée. Par exemple, l'identification du risque peut être complétée par une étape d'étude du contexte, l'analyse du risque par l'estimation et

l'évaluation des risques. L'étape de réponse peut être complétée par des étapes de catégorisation, classification ou surveillance de l'efficacité des réponses (National Institute of Standards and Technology, 2018a), (Myagmar et al., 2005), (Anthony Cox, 2008).

La phase d'identification des risques est essentielle car elle pose les fondements de leur gestion. Elle correspond au processus d'établissement d'une liste des risques pouvant survenir de l'intérieur ou l'extérieur d'une organisation. Les données issues de la phase d'identification seront utilisées aux étapes suivantes et nécessitent donc d'être exhaustives et correctes (Jallow et al., 2007).

La phase d'analyse permet d'estimer et d'évaluer les risques en se basant sur les méthodologies de calcul du risque qui découlent des définitions vues précédemment. Cette étape est clé pour les organisations afin d'adapter les réponses, de les prioriser et d'assurer leur pertinence (Kaplan' & Garrick2, 1981).

Le rôle des méthodes de RM est de permettre la facilitation de ces étapes qui peuvent s'avérer complexes. De nombreuses méthodologies existent. Certaines sont issues d'organisations nationales ou internationales, d'autres proviennent d'organisations privées traditionnelles ou de travaux de recherches. Chacune de ces méthodes a été développée pour répondre à des besoins précis et différents (Saleh & Alfantookh, 2011).

Dans la prochaine partie de cette section, nous allons étudier l'applicabilité des principales méthodologies de management du risque au DRR.

2.2. Etude de différentes méthodologies de gestion du risque : prise en compte partielle du risque de conformité des données réglementées

La première partie est consacrée à la présentation de nos critères d'applicabilité et d'analyse. La deuxième partie présente différentes méthodologies de gestion du risque. Notre analyse est présentée dans la troisième partie.

2.2.1. Critères d'applicabilité

Lors de la première section, nous avons défini le Data Regulation Risk (DRR) comme « un risque provenant de la possibilité d'une sanction par un organisme réglementaire suite à la

preuve de non-conformité d'une norme régissant le traitement des données et/ou les gouvernances et processus des technologies de l'information et communication et/ou les technologies et services de l'information » (Delorme et al., 2021).

Le DRR est alors intrinsèquement lié à la conformité de l'organisation. Nous avons également vu que la conformité prend en considération les actions qui permettent de réduire le risque d'infraction aux règles et rend les organisations responsables de ne pas avoir mis en place les moyens et mesures nécessaires.

Pour qu'une méthodologie de RM soit applicable à la gestion du DRR, celle-ci doit prendre en considération la gouvernance, les mesures de sécurité et actions mises en place. Notre premier critère sera alors la multidisciplinarité de la méthodologie :

- Multidisciplinaire ; ce critère permet de différencier les méthodologies qui peuvent être appliquées à des risques non IT.

La possibilité de prendre en compte de multiples scénarios qui répond à notre définition du DRR doit aussi être présente dans la méthodologie de RM. Notre deuxième critère est la capacité de prendre en compte différents scénarios :

- Multi-scénario ; ce critère permet d'évaluer si plusieurs scénarios de survenance sont pris en compte pour un même risque.

L'applicabilité des différents textes juridiques étant spécifique au contexte d'une organisation, dépendant de ses marchés, de la présence géographique et des juridictions de celle-ci, la méthodologie de RM doit être en mesure de prendre en considération le contexte d'une organisation. Pour évaluer la prise en compte du contexte lors d'une analyse de risque, nous créons le critère suivant :

- Contexte ; ce critère permet d'évaluer l'existence d'une étude du contexte de l'organisation lors de l'analyse de risque.

Une exigence réglementaire spécifique peut non seulement nécessiter une interprétation juridique, mais également l'implication de praticiens de la sécurité, de responsables informatiques, d'experts en gouvernance et métiers. Ces experts sont généralement répartis entre les départements de l'organisation tels que le juridique, la sécurité des systèmes d'information, l'informatique opérationnelle, finance, gestion des risques, ressources humaines, etc. Nous pouvons en déduire que la méthodologie de RM doit prévoir l'implication

de différents experts pour permettre l'estimation d'une probabilité subjective précise. Ce prérequis se sera évalué par le critère suivant :

- Transversalité ; ce critère permet de préciser l'implication d'experts de différentes disciplines.

Reprenant les étapes présentées dans la Figure 2 Etapes principales du management du risque, chaque méthodologie de RM doit prévoir des étapes d'identification, d'analyse et de réponse au risque. Notre analyse se porte alors sur les critères suivants :

- Identification ; ce critère permet d'évaluer l'existence d'une étape d'identification du risque.
- Analyse ; ce critère permet d'évaluer l'existence d'une étape d'analyse du risque.
- Réponse ; ce critère permet d'évaluer l'existence d'une étape de réponse au risque.

Afin de différencier les méthodologies qui prennent en compte la conformité réglementaire dans les différentes étapes, nous créons le critère suivant :

- Conformité réglementaire ; ce critère permet d'évaluer la prise en compte de la conformité réglementaire.

Pour chacun des critères présentés, nous attribuons la valeur « oui » s'ils sont présents dans la méthodologie de RM étudiée. Leur absence est signifiée par la valeur « non ».

Enfin, pour chaque méthodologie étudiée, nous attribuons les valeurs « académique », « publique » ou « privée » selon si elles sont issues de travaux de recherche, d'organismes réglementaires ou d'organisations privées.

2.2.2. Présentation des méthodologies sélectionnées

Comme énoncé ci-avant, il existe de nombreuses définitions qui divergent selon le prisme sous lequel le risque est étudié. De manière similaire, les besoins et champs d'action divergents ont conduit à l'élaboration de méthodologies différentes.

Chacune dispose de ses spécificités, avantages et limites. Du fait de l'importance et de la variété des méthodologies disponibles, des travaux de comparaison des différents cadriciels de gestion des risques de sécurité de l'information ont été réalisés (Barlette & Fomin, 2009), (Labuschagne & Vorster, 2005),(Mellado et al., 2011),(Wangen, 2007),(Broderick, 2006). D'autres travaux se sont intéressés aux critères de sélection (Schlarman, 2007), (Al-Ahmad & Bassil Mohammad, 2012), ou encore à l'évaluation de leur complétude (Wangen et al., 2018).

L'objectif de cette partie n'est pas de présenter un état de l'art exhaustif des méthodologies à la disposition des organisations. Pour cela nous invitons les lecteurs à s'intéresser aux travaux existants et dédiés à ce sujet tels que Aven (Aven, 2016), Shameli-Sendi et al. (Shameli-Sendi et al., 2016), Cherdantseva et al. (Cherdantseva et al., 2016) et Tixier et al. (Tixier J et al., 2002).

Nous proposons dans cette partie d'étudier un panel représentatif des différentes méthodologies existantes. Pour cela nous avons sélectionné les méthodologies les plus citées parmi les auteurs ci-dessus[4]. Notre sélection consiste en des méthodologies issues :

- Du monde académique ;
- D'organisations privées (entreprises spécialisées, associations de professionnels, etc.) ;
- D'organisations publiques (organismes règlementaires ou rattachés à un gouvernement).

Les méthodologies divergent aussi selon leur perspective (orientée service, métier, ou technologique) et leur méthode d'évaluation du risque (quantitative, qualitative, hybride) (Shameli-Sendi et al., 2016).

Les méthodologies sont présentées par ordre alphabétique.

AIPD (Analyse d'Impact relative à la Protection des Données)

L'entrée en vigueur du RGPD a grandement influencé la sphère de la gestion du risque des données à caractère personnel. Suite à sa promulgation, différentes méthodologies de gestion du risque lié aux données à caractère personnel sont apparues. Celles-ci répondent à

[4] En supplément des méthodologies citées et issues d'organisations privées et publiques, nous avons sélectionné deux méthodologies issues de travaux académiques.

l’article 35 du RGPD qui stipule que « Lorsqu'un type de traitement, en particulier par le recours à de nouvelles technologies, et compte tenu de la nature, de la portée, du contexte et des finalités du traitement, est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes physiques, le responsable du traitement effectue, avant le traitement, une analyse de l'impact des opérations de traitement envisagées sur la protection des données à caractère personnel. Une seule et même analyse peut porter sur un ensemble d'opérations de traitements similaires qui présentent des risques élevés similaires. » (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.).

Communément appelé Data Privacy Impact Assessment (DPIA) ou Privacy Impact Assessment (PIA), ces méthodologies aident non seulement à construire des traitements de données respectueux de la vie privée, mais aussi à démontrer leur conformité au RGPD.

La Commission Nationale de l'Informatique et des Libertés (CNIL), un organisme public français qui agit au nom de l'Etat, est en charge de veiller à la protection des données personnelles. La CNIL propose depuis 2018 une analyse d’impact relative à la protection des données composée de quatre étapes :

- La définition du contexte ;
- L’analyse de la légitimité des traitements et des mesures de sécurité en place ;
- L’étude des risques liés à la sécurité des données ;
- La validation de l’analyse d’impact.

Selon les lignes directrices de la CNIL (Commission Nationale de l’Informatique et des Libertés, 2018), la définition du contexte nécessite d’établir la liste exhaustive des données impactées, de la description des traitements associés, des éléments du système d’information impliqués et des contraintes règlementaires (« codes de conduite approuvés »). Le DPIA permet alors la prise en compte de multiple scénarios. L’étude des risques nécessite d’établir l'origine, la nature, la particularité et la gravité de ces derniers.

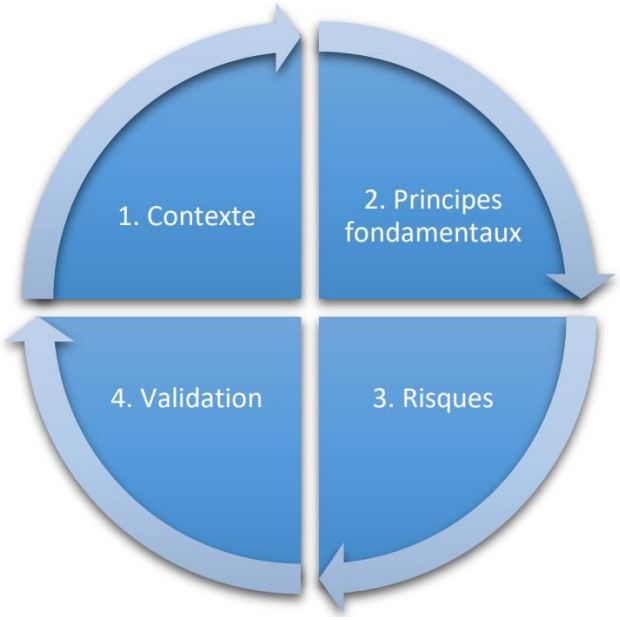


Figure 3 Analyse d'impact relative à la protection des données de la CNIL

Pour chaque risque, l'analyse doit prendre en considération : les impacts potentiels sur les droits et libertés des personnes concernées ; les menaces associées ; leur vraisemblance et gravité et les mesures envisagées pour y répondre. Malgré son périmètre d'application limité aux traitements de données à caractère personnel, le DPIA répond à notre critère de multidisciplinarité car il prend en considération la gouvernance, les mesures de sécurité et action mises en place.

La CNIL définit le DPIA comme un processus d'amélioration continue, itératif, impliquant différents experts métiers et composé de quatre étapes tel qu'illustré dans la

Figure 3 Analyse d'impact relative à la protection des données de la CNIL.

COBIT (Control OBjectives for Information and Related Technology)

Le Control Objectives for Information and Related Technology (COBIT) est un cadriciel de gouvernance informatique qui consiste en un ensemble d'outils de support permettant aux responsables de « combler le fossé entre les exigences de contrôle, les problèmes techniques et les risques métiers » (Sheikhpour & Modiri, 2012). Le COBIT est un

référentiel de bonnes pratiques pour faciliter l'élaboration de politiques et de bonnes pratiques en matière de sécurité et de contrôle. La version 5 du référentiel développé par l'Information Systems Audit and Control Association[5] est disponible depuis 2012. Cette version décline 37 processus regroupés en cinq domaines qui sont :

- Évaluer, diriger, et surveiller ;
- Aligner, planifier et organiser ;
- Bâtir, acquérir, et implanter ;
- Livrer, servir et soutenir ;
- Surveiller, évaluer et mesurer.

Les différents domaines permettent au COBIT d'être un cadriciel multi-scénario, multidisciplinaire et transversal. La conformité est un élément clé du COBIT. Celle-ci est présente dans le dernier domaine (surveiller, évaluer et mesurer). L'objectif 27 stipule qu'une organisation doit « assurer la conformité de l'informatique aux lois et aux règlements » (Moisand & Garnier de Labareyre, 2009). Il n'a cependant pas pour vocation de servir de référentiel de conformité à des exigences réglementaires ou contractuelles. Pour cela, le cadriciel COBIT doit être complété par d'autres plus adaptés.

EBIOS (Expression des Besoins et Identification des Objectifs de Sécurité)

Créée par l'Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI) en 1995, la méthode EBIOS se concentre sur l'appréciation et la gestion des risques relatifs à la sécurité des systèmes d'information (Agence Nationale de la Sécurité des Systèmes d'Information, 2010). La version EBIOS 2010 repose sur cinq étapes qui sont :

- L'étude du contexte ;
- L'étude des évènements redoutés ;
- L'étude des scénarios de menaces ;
- L'étude des risques ;

[5] L'Information Systems Audit and Control Association (ISACA) est une association professionnelle internationale dont l'objectif est d'améliorer la gouvernance des systèmes d'information accessible via le lien suivant : https://www.isaca.org/

- L’étude des mesures de sécurité.

L’ANSSI a fait évoluer cette méthodologie et propose désormais EBIOS Risk Manager. Cette nouvelle version est basée sur une approche par conformité supportée par la réalisation de scénarios (Agence Nationale de la Sécurité des Systèmes d’Information, 2018). EBIOS Risk Manager consiste en une démarche itérative de cinq étapes (« ateliers ») comme illustré dans la Figure 4 Méthodologie EBIOS Risk Management.

Cette approche multi-scénario repose sur des scénarios stratégiques conçus à l’échelle de l’écosystème et des valeurs métier de l’objet étudié mais aussi sur des scénarios opérationnels. Ces derniers reprennent les modes opératoires susceptibles d’être utilisés par les sources de risque pour réaliser les scénarios stratégiques.

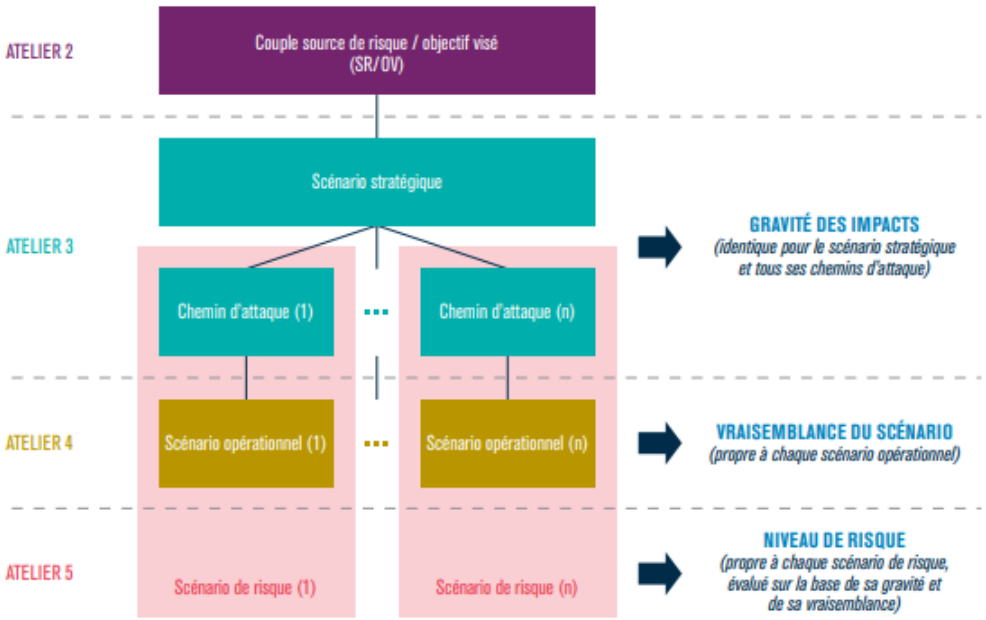


Figure 4 Méthodologie EBIOS Risk Management

Le dernier atelier a pour but la réalisation d’une synthèse des scénarios de risque identifiés et de définir une stratégie de traitement du risque multidisciplinaire et transversale. Cette stratégie aboutit à la définition de mesures de sécurité, recensées dans un plan d’amélioration continue de la sécurité.

EISRM (Enterprise Information Security Risk Management) de Saleh & Alfantookh

Les travaux de Saleh et Alfantookh (Saleh & Alfantookh, 2011) ont conduit à l'élaboration d'un cadriciel de gestion des risques de sécurité de l'information pour les organisations utilisant les technologies de l'information. Leur méthodologie nommée Enterprise Information Security Risk Management (EISRM) comporte deux parties principales : la partie structurelle et la partie procédurale. La première regroupe les critères et le périmètre alors que la seconde représente les processus et les outils. Leur méthodologie à quatre dimensions est présentée dans la Figure 5 Méthodologie de Saleh et Alfantookh. Les différentes dimensions permettent à l'EISRM d'être une méthodologie de gestion du risque multi-scénario, multidisciplinaire et transversal.

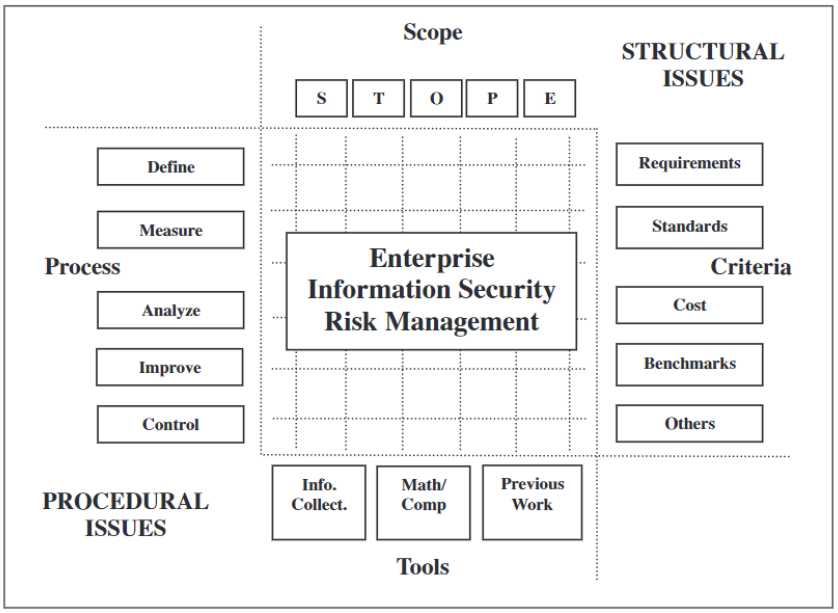


Figure 5 Méthodologie de Saleh et Alfantookh

Le champ d'action ou périmètre de EISRM est précisé par l'application des domaines de stratégie, technologie, organisation, individu et environnement (STOPE). Les critères correspondent aux contrôles et différentes documentations applicables (cadriciels, procédures, politiques, etc.). La dimension processus reprend les phases cycliques du modèle Six-Sigma (Montgomery & Woodall, 2008) qui sont la définition, la mesure, l'analyse, l'amélioration et le contrôle. Enfin, la dimension outil se réfère à l'ensemble des solutions et outils qui supportent l'application de la méthodologie. L'EISRM permet alors l'étude du contexte et de la conformité réglementaire.

ISO (International Organization for Standardization)

L'organisation Internationale de normalisation est une organisation non gouvernementale, indépendante regroupant des organismes nationaux de normalisation. Celle-ci est à l'origine de différents standards ou normes internationales. La force des documents ISO sont leur complétude et la diversité des sujets qu'ils adressent.

Le standard ISO/IEC 27005:2018 (Technologies de l'information, Techniques de sécurité, Gestion des risques liés à la sécurité de l'information) a pour but de supporter tout type d'organisation qui a l'intention de gérer des risques susceptibles de compromettre la sécurité de ses informations (International Organization for Standardization, 2018b).

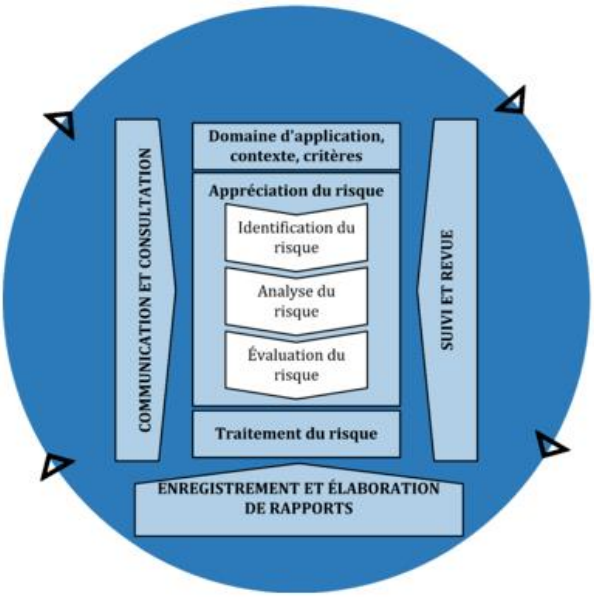


Figure 6 Processus de gestion des risques ISO

La Figure 6 Processus de gestion des risques ISO présente un processus cyclique qui comprend les étapes :

- D'appréciation d'un traitement des risques ;
- De décision de reconnaître si les niveaux de risque résiduel sont acceptables ;
- De génération d'un nouveau traitement des risques si les niveaux de risque ne sont pas acceptables ;
- D'appréciation de l'efficacité du traitement considéré.

Le processus de gestion des risques ISO repose sur l'étude du contexte afin d'apprécier et traiter les risques multidisciplinaires et transverses. Suffisamment générique, il est applicable à la conformité réglementaire cependant le standard ISO/IEC 27005:2018 ne spécifie pas de

méthodologie de gestion de risque mais présente la nécessité d'avoir un processus itératif basé sur l'ISO 31000:2018 (International Organization for Standardization, 2018a).

MEHARI (MEthod for Harmonized Analysis of RIsk)

MEHARI est une méthode d'analyse et de gestion des risques développée par le CLUSIF (Club de la sécurité de l'information français). Celle-ci se concentre sur les risques associés à la sécurité de l'information excluant les risques transversaux, multidisciplinaires et la conformité réglementaire. MEHARI est basée sur un modèle de risque qui comprend un aspect quantitatif et un aspect qualitatif (Club de la sécurité de l'information français, 2007) permettant l'étude du contexte de l'organisation. Cette méthodologie propose trois types de menaces qui sont d'origines volontaires, d'une erreur et accidentelle et ne traite que de risques liés aux composants technologiques.

NIST RMF (Risk Management Framework)

Le Risk Management Framework for Information Systems and Organizations (NIST RMF) fournit des directives pour la gestion de la sécurité, de la confidentialité et des risques inhérents aux systèmes d'information et aux organisations (National Institute of Standards and Technology, 2018b). Publié par le National Institute of Standards and Technology (NIST), le RMF est une méthodologie à plusieurs niveaux pour la gestion des risques à un niveau organisationnel, processus métier et information (donnée). Cette méthodologie repose sur sept étapes qui sont :

- La préparation qui permet d'établir le contexte et les priorités ;
- La catégorisation des éléments du système d'information ;
- La sélection des contrôles à mettre en place ;
- L'implémentation et la documentation des contrôles sélectionnés ;
- L'évaluation de la bonne mise en place et de l'efficacité des contrôles ;
- L'autorisation de déploiement des contrôles et mesures ;
- La surveillance de l'organisation et de son bon fonctionnement.

Cette méthodologie requiert l'identification et l'implication de différents experts afin de déterminer le contexte et les besoins de l'organisation pour arriver à l'état souhaité. Les différentes phases sont déclinées selon le prisme étudié (organisationnel, processus métier et information) afin de permettre une gestion du risque granulaire (National Institute of Standards

and Technology, 2018b). NIST RMF fournit une méthodologie multi-scénario, multidisciplinaire qui prend en considération la conformité réglementaire. En outre, elle prévoit l'étude du contexte organisationnel et sa transversalité supporte les différentes étapes de gestion du risque.

OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation)

L'Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE) est une méthodologie pour rationaliser et optimiser le processus d'évaluation des risques de sécurité de l'information. OCTAVE se concentre principalement sur l'information et les technologies de l'information (« information assets ») dans leur contexte d'utilisation (stockage, transport, traitement, etc.) (Caralli et al., 2007). Ce périmètre d'application qui n'est pas transversal, écarte de son utilisation les risques multidisciplinaires et la conformité réglementaire.

Cette méthodologie utilise des questionnaires de scénario de menace pour aider les utilisateurs à identifier les menaces et les risques multi-scénario associés. Ces questionnaires sont basés sur des arbres de menaces conçus autour du concept de « conteneur ». Cela revient à limiter l'analyse de risque au traitement de la donnée par un composant informatique (Wangen et al., 2018).

ROPE (Risk-Oriented Process Evaluation)

Développée par Jakoubi & al. (Jakoubi et al., 2007), la méthodologie ROPE (Risk-Oriented Process Evaluation) vise à sécuriser les processus métiers en intégrant la gestion des risques et de la continuité des activités comme partie intégrante. ROPE combine l'avantage de la modélisation des processus métier, de la gestion des risques et des concepts de continuité des activités. Selon ses auteurs (Jakoubi et al., 2007), « ROPE fournit une détermination de l'impact des menaces et des contre-mesures sur l'exécution des processus métier basée sur des simulations ».

Elle consiste en cinq processus itératifs, qui sont dérivés du Business Process Management Systems Concepts de Karagiannis et al. (Karagiannis et al., 1996), du diagramme

CARE (Condition, Action, Ressource et Environnement) et du diagramme TIP (Threat Impact Process) (Tjoa et al., 2008). Les cinq processus sont :

- Le processus de décision stratégique qui permet l'identification et la priorisation des processus métier à analyser ;
- Le processus de réingénierie qui en composé de sous processus afin d'obtenir le modèle cible défini lors de la première étape ;
- L'allocation des ressources ;
- Le management des processus métier qui consiste en l'exécution du processus afin de l'évaluer;
- L'évaluation des performances.

Ce modèle à trois couches (business process modelling layer, CARE et TIP) permet en outre une représentation des processus métiers d'une organisation, des éléments qui composent ces processus et les menaces qui leur sont propres. Ces trois couches nécessitent l'implication de différents experts et permet la simulation de différents scénarios simultanément. La méthodologie ROPE fournit une approche structurée, multi-scénario, transversale qui guide la représentation d'une entreprise sous le prisme opérationnel des menaces pouvant impacter la continuité d'activité ou la sécurité des processus métiers. Bien qu'elle permette la prise en compte du contexte d'une organisation, celle-ci ne couvre pas la conformité réglementaire. Un exemple d'application de cette méthodologie est détaillé dans les travaux et Tjoa et al. (Tjoa et al., 2008).

La prochaine partie présente notre analyse des méthodologies de gestion du risque présentées.

2.2.3. Analyse d'applicabilité des méthodologies au DRR

Nous avons vu que plusieurs méthodologies déjà existantes peuvent être utilisées pour la gestion du DRR.

Parmi les méthodologies étudiées, cinq répondent à nos critères d'applicabilité :

- AIPD (Analyse d'Impact relative à la Protection des Données) qui provient d'une organisation publique ;

- EISRM (Enterprise Information Security Risk Management), qui provient du monde académique ;
- EBIOS RM (Expression des Besoins et Identification des Objectifs de Sécurité Risk Manager), qui provient d'une organisation publique ;
- ISO (International Organization for Standardization) qui provient d'une organisation privée ;
- NIST RMF (Risk Management Framework), qui provient d'une organisation publique ;

Il demeure la problématique d'identification des risques. Nous rappelons que la phase d'identification est essentielle car elle pose les fondements de la gestion des risques indépendamment de la méthodologie choisie. Elle correspond au processus d'établissement d'une liste des risques utilisée aux étapes suivantes et nécessitent donc d'être exhaustive et correcte (Jallow et al., 2007). Elle permet dans un second temps d'estimer et d'évaluer les risques afin d'adapter les réponses, de les prioriser et d'assurer leur pertinence (Kaplan' & Garrick2, 1981).

Dans l'ensemble des méthodologies étudiées, la phase d'analyse repose sur les risques perçus (Kaplan' & Garrick2, 1981). Cela revient à limiter les analyses de risque à ce qui est connu et perçu par les individus impliqués dans celles-ci. Il en convient que l'exhaustivité des risques traités et la pertinence de la gestion des risques sont intrinsèquement liées et limitées par les connaissances des personnes impliquées.

Nom	Origine	Multi-scénario	Multidisciplinaire	Transversalité	Contexte	Identification du risque	Analyse du risque	Réponse au risque	Conformité règlementaire
AIPD	Publique	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
COBIT	Privée	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Non
EBIOS 2010	Publique	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Non
EBIOS RM	Publique	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
EISRM	Académique	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
ISO	Privée	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
MEHARI	Privée	Non	Non	Non	Oui	Oui	Oui	Oui	Non
NIST RMF	Publique	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
OCTAVE	Privée	Oui	Non	Non	Oui	Oui	Oui	Oui	Non
ROPE	Académique	Oui	Non	Oui	Oui	Oui	Oui	Oui	Non

Tableau 2 Analyse d'applicabilité de méthodologies de gestion du risque au DRR

Si l’on considère qu’une exigence réglementaire spécifique peut non seulement nécessiter une interprétation juridique, mais également l'implication de praticiens de la sécurité, de responsables informatiques, d'experts en gouvernance et métier, l’identification du DRR nécessite que chaque individu ou expert perçoive et connaisse la part du risque qui lui est propre.

La complexité de cet exercice de gestion du risque réside dans la nécessité de traduire les contraintes et exigences réglementaires en des termes techniques, organisationnels et opérationnels pour être en mesure de fournir des instructions pragmatiques et concrètes en se basant sur des textes juridiques génériques et parfois abstraits. À cette abstraction et généricité, s’adjoint la strate des modalités déontiques qui complexifie la compréhension des lois et l’appréhension des impacts. Cette compréhension doit également être partagée et acceptée par l’ensemble des individus impliqués dans la gestion de risque.

Il nous parait alors peu heuristique d’espérer que les analyses de risque puissent être exhaustives sans outil de gestion de connaissance. En effet, la compréhension des impacts métiers et opérationnels des exigences réglementaires est impérative pour permettre d’identifier, d’évaluer et de répondre à un DRR.

Conclusion

L'identification d'un risque multidisciplinaire et multi-scénario tel que le DRR nécessite la représentation des différentes modalités juridiques auxquelles une organisation est confrontée lors de sa mise en conformité avec une règlementation. Dans le cadre du DRR, la conformité décrit la mise en place de moyens et mesures nécessaires de prévention d'infraction aux règles juridiques régissant le traitement des données et/ou les gouvernances et processus des technologies de l'information et communication et/ou les technologies et services de l'information d'une organisation.

L'identification du DRR implique donc la nécessité de pouvoir représenter les différentes connaissances présentées dans la première section. Ceux-ci sont :

- Les connaissances relatives au corpus juridique (règlementations et modalités juridiques) ;
- Les connaissances relatives au système d'information (sous-systèmes sociaux et techniques) ;
- Les connaissances relatives au contexte (applicabilité des règlementations).

Partant de ce postulat, un système de gestion de connaissances apparait comme une solution pouvant subvenir aux besoins d'identification des risques et plus spécifiquement du DRR. L'accès à ces connaissances peut permettre à une organisation d'alimenter son processus de gestion des risques.

Nous avons néanmoins vu que les différents concepts impliqués dans le DRR interagissent entre eux (principe de conformité). Il en convient que pour pouvoir pleinement identifier le DRR, il est également nécessaire de pouvoir identifier les relations entre les différents concepts. Par conséquent, nous devons écarter les systèmes de gestion de connaissances de type taxonomie qui reposent uniquement sur la classification (Ponzetto & Strube, 2007).

Chapitre 2 - Représentation des connaissances en conformité : usage des ontologies, limites et perspectives

Ce chapitre sera consacré à l'étude des ontologies. Celles-ci sont des « spécifications formelles et explicites de conceptualisations partagées » (Gruber, 1993).

Nous présentons, dans la première section, les objectifs et les fondamentaux des ontologies. Nous discutons également de différentes méthodologies de construction et des principaux langages utilisés pour la conception d'ontologie.

La seconde section de ce chapitre nous permet d'étudier différentes ontologies existantes issues des domaines juridique et de la sécurité de l'information. Au total, 13 ontologies sont présentées et analysées selon une grille d'évaluation définie.

Section I. La représentation des connaissances dans les ontologies

Cette section est organisée comme suit : la première partie présente les principes des ontologies et leurs caractéristiques. La seconde partie étudie différentes méthodologies de conception et les principaux langages utilisés pour la conception d'ontologie.

1.1. Ontologie, concept et définition

La discipline de « Knowledge Engineering » (KE) trouve son origine dans l'insatisfaction générale liée aux problématiques de construction de « Knowledge Based System » (KBS). En effet, les premiers KBS se trouvaient limités par manque de scalabilité (Studer et al., 1998). La discipline de KE s'est naturellement tournée vers les ontologies pour répondre à ses objectifs de permettre le partage et la réutilisation de connaissances sous forme computationnelle.

L'origine des ontologies se situe néanmoins dans le domaine de la philosophie. Cette branche du domaine se ramifie autour de la définition de « l'être ». La question de l'être et de sa définition a intéressé les premiers penseurs tel que Platon. La signification moderne du terme n'est apparue que plus tard. Suite aux travaux d'Aristote, la question de l'être est mise de côté pour se concentrer sur les multiples significations de l'être.

Enfin, en 1991, Neches et al. (Neches et al., 1991) se posent la question de la réutilisation d'entités existantes lors de conception de bases de connaissance. Ils qualifient l'ontologie comme « définition d'un sous langage pour modéliser un domaine constitué d'objets, relations et contraintes ». Cette proposition est inscrite dans une volonté d'adresser quatre obstacles que rencontrent les systèmes de partage de connaissance :

- Les représentations hétérogènes de connaissance,
- Les disparités sémantiques,
- Le manque de protocole de communication (entre systèmes),
- Le manque d'interopérabilité entre les bases de connaissance.

Par la suite, Grüber (Gruber, 1993) définit une ontologie comme « une spécification formelle et explicite d'une conceptualisation partagée ». La « conceptualisation » fait référence à un modèle abstrait d'un phénomène dans le monde par l'identification des concepts pertinents

de ce phénomène. Le terme « explicite » signifie que le type de concepts utilisés et les contraintes sur leur utilisation sont explicitement définis. La caractéristique « formelle » fait référence au fait que l'ontologie doit être lisible par une machine, ce qui exclut le langage naturel. Enfin, « partagée » reflète la notion selon laquelle une ontologie capture la connaissance consensuelle, c'est-à-dire qu'elle n'est pas privée pour un individu, mais acceptée par un groupe.

Les ontologies permettent alors la modélisation d'un modèle avec un niveau d'abstraction souhaité par des conceptualisations explicites d'un domaine (Bench-Capon et al., 1998). Selon cette approche, elles sont reconnues pour faciliter le partage, la réutilisation, et l'utilisation des connaissances.

Fondamentalement, le rôle des ontologies dans le processus de KE est de faciliter la construction d'un modèle de domaine. Une ontologie fournit un vocabulaire de termes et de relations avec lesquels modéliser le domaine. Dans le cas où l'ontologie convient parfaitement au domaine, alors un modèle de domaine peut être obtenu en ne remplissant l'ontologie qu'avec des instances. Il existe plusieurs types d'ontologies, et chaque type remplit un rôle différent dans le processus de construction d'un modèle de domaine (Studer et al., 1998).

Dans cette section, nous aborderons les principes qui régissent les ontologies, leur création et leur utilisation. Cette section introduit également les différents types d'ontologie puis couvre les différentes méthodologies de construction d'ontologie. Enfin, cette section présente différents langages à disposition pour la création d'ontologie.

1.1.1. Principe des ontologies

Partant du postulat que les humains s'appuient de plus en plus sur le support informatique pour traiter les données en raison de l'augmentation du volume, de la complexité, et vitesse de création des données, Wilkinson et al. (Wilkinson et al., 2016) ont présenté des lignes directrices pour améliorer la possibilité de trouver « Findability », l'accessibilité « Accessibility », l'interopérabilité « Interoperability » et la réutilisation « Reusability » des données et actifs numériques.

Ces lignes directrices sont regroupées dans les « principes FAIR » dont le but est d'assurer ou de supporter la capacité des systèmes informatiques à trouver, accéder, interagir et réutiliser les données sans intervention humaine ou avec une intervention humaine minimale.

Findability

La première catégorie de principes correspond à la possibilité de trouver l'information ou la donnée. Quatre principes FAIR permettent de répondre à ces enjeux, ceux-ci sont :

- F1 : Un identifiant global unique et persistant doit être attribué à chaque donnée ;
- F2 : Chaque donnée doit être « richement » décrite avec des métadonnées ;
- F3 : Les métadonnées doivent clairement et explicitement inclure l'identifiant unique de la donnée décrite ;
- F4 : Les données (et métadonnées) doivent être indexées.

Accessibility

La deuxième catégorie de principes correspond à la capacité d'accéder à l'information trouvée. Deux principes FAIR permettent de répondre à ces enjeux, ceux-ci sont :

- A1 : Les données peuvent être accédées par des protocoles de communication standards en utilisant leur identifiant unique ;
    - A1.1 : Les protocoles utilisés doivent être libre d'accès, gratuits et universellement implémentable ;
    - A.1.2 : Les protocoles utilisés doivent permettre (le cas échéant) la mise en place de procédure d'authentification et d'autorisation ;
- A2 : Les métadonnées doivent rester accessibles indépendamment des données.

Interoperability

La troisième catégorie de principes correspond à la capacité d'interaction avec des applications pour divers traitements (e.g. intégration, stockage). Trois principes FAIR permettent de répondre à ces enjeux, ceux-ci sont :

- I1 : Les données doivent utiliser un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances ;
- I2 : Les données doivent utiliser un vocabulaire qui respecte les principes FAIR ;

- I3 : Les données doivent utiliser des références qualifiées à d'autres données. Une référence qualifiée est une cross-référence explicite qui spécifie son intention. En d'autres termes, les références ne doivent pas être génériques et vagues.

Reusability

La quatrième catégorie de principes correspond à la capacité de réutilisation des données. Un principe FAIR permet de répondre à ces enjeux :

- R1 : les données sont (richement) décrites avec une pluralité d'attributs précis et pertinents ;
    - R1.1 : Les données sont publiées avec une licence d'utilisation des données claire et accessible ;
    - R1.2 : La provenance des données est détaillée ;
    - R1.3 : Les données répondent aux normes communautaires applicables au domaine.

Les ontologies en tant qu'outils de conceptualisation de l'information et méthodes formelles de classification et représentation des conceptualisations doivent répondre aux principes FAIR (Guizzardi, 2020), (Collins et al., 2018; Poveda-Villalón et al., 2020). En outre, les ontologies jouent un rôle pertinent dans certains des principes FAIR, tels que l'interopérabilité et la réutilisation des données (Guizzardi, 2020).

Les principes FAIR datent de 2016 et sont donc relativement récents. Cependant des travaux antérieurs discutaient des problématiques de publication et partage des ontologies, du besoin d'identifiants uniques et permanent ou encore l'utilisation de protocoles standardisés (Janowicz et al., 2014), (Heath & Bizer, 2011), (Bizer & Heath, 2011).

En effet, les principes des données liées « Linked Data Principles » ont été proposés en 2006 comme un ensemble de lignes directrices pour la publication et la liaison de données sur le Web (Bizer & Heath, 2011). Les principes des données liées sont les suivants :

- L'utilisation d'identifiant de ressource unique (URI) ;
- L'utilisation de URI HTTP pour la recherche de ressource ;
- L'utilisation de langages normés pour fournir des informations utiles sur les URI ;

- La mise en relation des URI par des liens.

Ces principes ont par la suite été étendus en 2010, avec le système de « notation 5 étoiles » pour la publication de « Linked Open Data » (Janowicz et al., 2014). Ce système de notation permet d'assurer que les données sont disponibles sur le Web avec les critères suivants :

- L'utilisation d'une licence libre ;
- L'utilisation d'une forme lisible par machine ;
- L'utilisation d'un format non-propriétaire,
- L'utilisation du langage Ressource Descriptive Framework (RDF) pour identifier et décrire des ressources ;
- La création de liens vers d'autres ressources.

L'application et le respect de ces critères permettent d'assurer la construction d'un modèle de domaine et d'une ontologie utilisable et réutilisable. Toutes les ontologies ne se ressemblent pas et peuvent différer selon leur objectif, leur structure, leur domaine ou leur granularité. Nous présentons ci-après les différents types d'ontologie.

1.1.2. Typologie d'ontologie

Selon les travaux de Van Heijst et al. (van Heijst et al., 1997), les ontologies peuvent être classées selon deux dimensions : « 1) la quantité et le type de structure de la conceptualisation et 2) le sujet de conceptualisation » .

La première dimension est composée de :

- Les ontologies terminologiques, elles permettent de préciser les termes qui sont utilisés pour représenter les connaissances dans le domaine du discours ;
- Les ontologie d'information, elles spécifient la structure d'enregistrement des bases de données ;
- Les ontologies de modélisation de connaissance. Elles permettent de préciser les conceptualisations des connaissances. Par rapport aux ontologies d'information, les ontologies de modélisation de connaissance ont généralement une structure interne plus riche.

La seconde dimension regroupe :

- Les ontologies d'application, elles contiennent toutes les définitions nécessaires pour modéliser les connaissances requises pour une application particulière ;
- Les ontologies de domaine, elles permettent des conceptualisations spécifiques à un domaine ;
- Les ontologies génériques, elles sont similaires aux ontologies de domaine, mais les concepts qu'elles définissent sont considérés comme génériques dans de nombreux domaines. Les concepts dans les ontologies de domaine sont souvent définis comme des spécialisations de concepts dans des ontologies génériques ;
- Les ontologie de représentation, elles permettent d'expliciter les formalismes de représentation des connaissances qui décrivent les ontologies de domaine et les ontologies génériques.

Alors que la classification des ontologies par le sujet de conceptualisation est globalement acceptée, la dimension structurelle ne fait pas l'unanimité auprès de la communauté scientifique (Guarino & Gangemi, 1997). D'autres classifications existent. Par exemple, Uschold et Grüninger (Uschold & Gruninger, 1996) reposent sur le niveau de formalisme utilisé pour leur classification.

Quel que soit le type de structure et le sujet de la conceptualisation, la conception d'une ontologie représente un réel challenge. Différentes méthodologies de construction d'ontologie ont été proposées pour faciliter et guider leur création.

1.2. Conception d'ontologie : méthodes et outils

La conception d'une ontologie est un exercice passionnant mais complexe. Peu de méthodologies ou de lignes directrices expliquant comment construire une ontologie ont été développées, ce qui a entraîné des différences et des disparités importantes entre les ontologies existantes. Ces différences sont présentes même dans les ontologies construites à des fins similaires (Visser & Bench-Capon, 1998).

1.2.1. Méthodologie de construction

Avec l'ambition de faciliter la construction d'ontologies, des méthodologies de conception ont été présentées (Ferndndez et al., 1997), (Uschold & King, 1995), (Grüninger &

Fox, 1995) mais également des travaux de revue et d'analyse de ces différentes méthodologies (Bench-Capon et al., 1998), (Pinto & Martins, 2004).

Il ressort de ces analyses que les trois principales méthodologies de construction d'ontologie sont :

- Enterprise (Uschold & Gruninger, 1996; Uschold & King, 1995);
- Tove (Grüninger & Fox, 1995);
- Methontology (Ferndndez et al., 1997).

La popularité des méthodologies Tove et Enterprise est principalement historique. Ces méthodologies sont apparues en même temps que les premiers développements d'ontologies. Leurs concepteurs se concentraient alors principalement sur la création d'ontologies à partir de rien (« from scratch ») et n'accordaient que peu ou pas d'importance à la maintenance des ontologies (Pinto & Martins, 2004). L'importance était de ce fait donnée à l'acquisition de connaissances, leur formalisation et documentation.

Du fait du développement et par conséquent de l'accessibilité à de nombreuses ontologies, de nouvelles méthodologies ont commencé à apparaitre. L'accent est alors mis sur la réutilisation de l'existant et leurs maintenances (Pinto & Martins, 2004) les versions actuelles de Methontology sont un exemple de ces nouveaux intérêts (Ferndndez et al., 1997). Parallèlement à la croissance du nombre d'ontologies et de leurs partages, des méthodologies d'évaluation d'ontologies ont été proposées telle que OntoClean (Guarino & Welty, 2002). Elles permettent de supporter la structure des ontologies en vue de leurs potentielles intégrations ou réutilisations.

Enfin, des méthodologies spécifiques à des domaines commencent à être développées. Nous pouvons citer l'exemple de MeLOn (Methodology for building Legal Ontology) qui a été utilisée pour la conception d'une ontologie du domaine de conformité du RGPD (Palmirani et al., 2018).

Il convient alors de reconnaitre que ces différents efforts telles que des méthodologies de conception ou d'évaluation supportent les objectifs initiaux des ontologies qui sont la modélisation d'un modèle par des conceptualisations explicites d'un domaine. Bien que chaque méthodologie présente des nomenclatures différentes, des rapprochements entre les différentes étapes de conception peuvent être faits (Bench-Capon et al., 1998), (Pinto & Martins, 2004). Pinto et Martins (Pinto & Martins, 2004) proposent de regrouper ces étapes en cinq catégories principales :

- Les étapes de spécification, elles permettent de préciser les objectifs de l'ontologie et de définir leurs périmètres et ambitions ;
- Les étapes de conceptualisation, elles permettent de répondre aux besoins présenter dans l'étape de spécification. Ces étapes définissent les concepts et les relations du domaine d'étude ;
- Les étapes de formalisation, elles permettent de formaliser les concepts, les relations et contraintes afin d'ajouter des précisions pour garantir la compréhension de l'ontologie ;
- Les étapes d'implémentation, elles permettent la sélection du ou des langages de représentation et l'implémentation de l'ontologie ;
- Les étapes de maintenance pour la mise à jour et les éventuelles corrections de l'ontologie.

Ces différentes étapes peuvent être complétées par des activités additionnelles telles que l'acquisition de connaissance, l'évaluation ou la documentation.

1.2.2. Choix de langage

Au fil du temps, plusieurs langages ont été développés pour représenter les ontologies. Ces différents langages de spécifications d'ontologies sont issus des formalismes présentés ci-après. Chacun répond à des besoins spécifiques et peuvent être classés en trois catégories (Maniraj & Sivakumar, 2010):

- Langages basés sur des « frames » ;
- Langages basés sur des logiques de description ;
- Langages basés sur des graphes.

Les frames sont utilisés pour représenter des objets structurés. Présentés par Minsky (Minsky, 1974), ils représentent des « structures de donnée pour représenter des situations auxquelles sont attachées des informations». Les frames sont des réseaux de nœuds ou classes organisés hiérarchiquement selon des relations et liens de spécification. Composés de « slots », ceux-ci permettent de préciser les frames tels que des attributs. Les « slots » de haut niveau sont fixes et utilisés pour représenter ce qui est toujours vrai et inchangé dans les situations étudiées. Les « slots » de bas niveau, nommés « terminals », peuvent être vides et

sont spécifiques à une instance particulière d'une situation (Minsky, 1974). Les frames trouvent alors leur intérêt dans la représentation de la façon de penser d'experts par la fourniture d'une représentation structurée et concise des relations utiles (Fischer, 1998).

Les logiques de description sont issues de l'analyse des réseaux sémantiques structurés (Nardi & Brachman, 2003) et des frames. Les langages basés sur des logiques de description reposent sur trois éléments : les concepts, rôles et individus. Les concepts représentent des ensembles d'objets ou classes, les rôles illustrent les relations entre objets et les individus sont des instanciations de classe.

Ces éléments sont décrits par deux structures ou types de connaissances : la T-Box (boîte terminologique) et la A-Box (boîte assertionnelle). La boîte terminologique permet la description des concepts et des rôles. Elle exprime des faits relatifs s'apparentant à des définitions et lien hiérarchique. La boîte assertionnelle permet la description des individus et des descriptions et relations qui leurs sont propres. L'utilisation de ces deux boîtes permet de vérifier si une description est satisfaisable (i.e. absence de contradiction avec d'autres descriptions) et consistante (i.e. les individus instancient bel et bien les concepts qu'ils sont censés instancier). Décrites comme étant plus flexibles que les frames, les logiques de description reposent sur une syntaxe et une sémantique rigoureuses (Baader & Hollunder, 1991).

Présentés par Sowa (Sowa, 1984), les graphes conceptuels permettent la formalisation de relations entre prédicats et arguments (concepts). Ce formalisme de représentation de connaissances et de raisonnements prend la forme de graphes finis, connexes et bipartites ou chaque relation conceptuelle possède un arc ou plus devant être relié à un concept. Utilisant une notation à base de graphes ils sont notamment composés de types de nœuds permettant de représenter les concepts et les relations. Basés sur les graphes existentiels de Peirce (Peirce, 1978), ils disposent de différents niveaux de formalisme et peuvent être vus comme des schémas permettant de représenter graphiquement « des formules logiques, des schémas sans contraintes, des formules ou des opérations de graphes » (de Chalendar et al., 2000).

De ces formalismes de représentation de connaissances et de raisonnements sont issus différents langages de représentation d'ontologie. Pour servir notre propos, nous ne décrivons que certains d'entre eux qui sont orientés Web sémantique et recommandés par le World Wide

Web Consortium (W3C)[6], l'objectif de cette partie n'étant pas de présenter un état de l'art exhaustif des langages existants. Pour cela nous invitons les lecteurs à s'intéresser aux travaux existants et dédiés à ce sujet tels que (Su & Ilebrekke, 2002).

Extensible Markup Language (XML)

Le langage « Extensible Markup Language » (XML) permet de générer des balises servant la structuration de données et de documents (Bradley, 2002). XML est un composant ou un « sous-ensemble » du Standard Generalized Markup Language (SGML)[7] qui permet au SGML générique d'être transmis, reçu et traité sur le Web. Sa syntaxe est dite « extensible » car elle permet de définir différents langages.

```
<?xml version=”1.0”?>
<Security_Measure>
        <Measure>
                <name>Detect_Security_Measure</name >
                 <source>Security_Policy</source>
        </Measure>
        <Measure>
                <name>Identify_Security_Measure</name >
                 <source>Security_Policy</source>
         </Measure >
</ Security_Measure>
```

Figure 7 Exemple d'utilisation XML

La

Figure 7 Exemple d'utilisation XML présente comment le langage XML peut être utilisé pour illustrer deux mesures de sécurité respectivement nommées « Detect_Security_Measure » et « Identify_Security_Measure » par l'utilisation des balises <name></name >. Les balises

[6] Le World Wide Web Consortium (W3C), est un organisme de standardisation à but non lucratif chargé de promouvoir la compatibilité des technologies du World Wide Web accessible via le lien suivant : https://www.w3.org/

[7] Standard Generalized Markup Language (SGML) est un langage de balisage généralisé normalisé selon la norme ISO (ISO 8879:1986)

<source></source> permettent pour chaque mesure de nommer leur source « Security_Policy ».

Ressource Descriptive Framework (RDF)

Le « Ressource Descriptive Framework » (RDF) permet la gestion des métadonnées de documents XML. Reposant sur un ensemble de triplets qui sont :

- Le sujet (ressource à décrire) ;
- Le prédicat (propriété applicable à la ressource) ;
- L'objet (une ressource qui est la valeur de la propriété).

RDF est un modèle de données dont chaque ensemble de triplet forme un graphe RDF (Lassila & Swick, 1998). W3C préconise l'utilisation de XML pour décrire les ressources modélisées en RDF. Cette utilisation conjointe permet la représentation des déclarations de propriétés sur des ressources.

RDF Schema (RDFS) est une extension sémantique de RDF qui permet la représentation des connaissances par la structuration des ressources RDF.

RDFS permet entre autre de définir des triplets par la spécification de classe «rdfs:class », sous classe « rdfs:subClassOf », propriété «rdfs:property » et de sous propriété «rdfs:subPropertyOf ». RDFS permet également de préciser leur domaine « rdfs :domain » et leur intervalle de valeurs « rdfs : range ».

RDF et RDFS offrent la possibilité de décrire des entités ontologiques, mais trouvent leurs limites dans la description des relations entre des classes ou entre des propriétés car les axiomes ne peuvent pas être directement décrits et le type des relations ne peut être spécifié.

La

Figure 8 Exemple d'utilisation RDF/RDFS présente l'utilisation du langage RDF/RDFS pour illustrer notre exemple précédant. Par l'utilisation de RDFS, notre exemple peut être complété. Il est par exemple possible de spécifier le principe de sous classe via <rdfs:subClassOf> ou de spécifier la source d'une mesure de sécurité en associant différentes ressources.

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.semanticweb.org/DataRegulationRisk#"
                     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                     xmlns:xml="http://www.w3.org/XML/1998/namespace"
                    xmlns:SM="http://www.semanticweb.org/DataRegulationRisk#Security_Measure">
        <rdf:Description rdf:about="http://www.semanticweb.org/DataRegulationRisk#Security_Measure">
                <SM:source
rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Security_Policy">
                </SM:source>
        </rdf:Description>
        <rdf:Description  rdf:about="http://www.semanticweb.org/DataRegulationRisk#Security_Policy">
        </rdf:Description>
        <rdf:Description
rdf:about="http://www.semanticweb.org/DataRegulationRisk#Detect_Security_Measure">
    <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Security_Measure">
        </rdf:Description>
        <rdf:Description
rdf:about="http://www.semanticweb.org/DataRegulationRisk#Identify_Security_Measure">
    <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Security_Measure">
        </rdf:Description>
</rdf:RDF>
```

Figure 8 Exemple d'utilisation RDF/RDFS

Darpa Agent Markup Language + Ontology Inference Layer (DAML + OIL)

Darpa Agent Markup Language (DAML) est un langage permettant l'expression d'ontologies par l'utilisation de syntaxe RDF et une représentation basée sur des frames.

Similairement, Ontology Inference Layer (OIL) (Fensel et al., 2000) est un langage permettant l'expression d'ontologie. Il combine les primitives des langages basées sur les frames avec une sémantique formelle. Reposant sur RDFS et XML, il offre également des possibilités de raisonnement issues de la logique de description.

DAML+OIL correspond à une version améliorée de DAML intégrant certaines propriétés de OIL (Baader et al., 2005; Fensel et al., 2000). Reposant sur RDF et RDFS, DMAL+OIL est un langage plus adapté aux ontologies.

Ontologie Web Language (OWL)

Ontologie Web Language (OWL) est un langage présenté comme le standard de représentation des ontologies par le W3C (McGuinness & van Harmelen, 2004). Basé sur des logiques de description, OWL est inspiré de DAML + OIL dont la syntaxe repose sur la combinaison RDF/XML.

OWL vient compléter la syntaxe RDF par la présence de constructeurs qui permettent d'ajouter plus de vocabulaire pour décrire les propriétés et les classes tels que : la disjonction, la cardinalité, l'égalité, des propriétés plus riches (transitivité, symétrie, …) ou encore les classes énumérées.

A ce jour, il existe trois versions du langage OWL :

- OWL Lite. Cette version répond aux besoins d'une hiérarchie de classification et de contraintes simples pour faciliter la définition des thésaurus et autres taxonomies ;
- OWL DL. Cette version répond aux besoins d'expressivité maximale tout en conservant l'exhaustivité des calculs et de la décidabilité (tous les calculs se termineront en un temps fini). OWL DL inclut toutes les constructions du langage OWL sous certaines restrictions ;
- OWL Full. Cette version répond aux besoins d'expressivité maximale couplée à la liberté syntaxique de RDF sans aucune garantie de calcul. Il n'impose pas de séparation entre classe, propriété, individu et valeur des données et permet à une ontologie d'augmenter la signification du vocabulaire prédéfini.

La

Figure 9 Exemple d'utilisation OWL présente l'utilisation du langage OWL pour illustrer notre exemple précédant. L'utilisation d'OWL permet de préciser que les mesures « Detect_Security_Measure » et « Identify_Security_Measure » sont disjointes mais aussi que les mesures de sécurité et la politique de sécurité (Security_Policy) sont reliées par la propriété « isSourceOf ». D'autres précisions telles que la transitivité ou encore les restrictions des propriétés peuvent également être ajoutées par le même principe.

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.semanticweb.org/DataRegulationRisk#"
    xml:base="http://www.semanticweb.org/DataRegulationRisk"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:xml="http://www.w3.org/XML/1998/namespace"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
   <owl:Ontology rdf:about="http://www.semanticweb.org/DataRegulationRisk"/>
   <owl:ObjectProperty rdf:about="http://www.semanticweb.org/DataRegulationRisk#isSourceOf"/>
   <owl:Class rdf:about="http://www.semanticweb.org/DataRegulationRisk#Detect_Security_Measure">
      <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Security_Measure"/>
      <owl:disjointWith
rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Identify_Security_Measure"/>
   </owl:Class>
   <owl:Class rdf:about="http://www.semanticweb.org/DataRegulationRisk#Identify_Security_Measure">
      <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Security_Measure"/>
   </owl:Class>
   <owl:Class rdf:about="http://www.semanticweb.org/DataRegulationRisk#Security_Measure"/>
   </owl:Class>
   <owl:Class rdf:about="http://www.semanticweb.org/DataRegulationRisk#Security_Policy">
      <rdfs:subClassOf>
         <owl:Restriction>
            <owl:onProperty rdf:resource="http://www.semanticweb.org/DataRegulationRisk#isSourceOf"/>
            <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Detect_Security_Measure"/>
         </owl:Restriction>
      </rdfs:subClassOf>
      <rdfs:subClassOf>
         <owl:Restriction>
            <owl:onProperty rdf:resource="http://www.semanticweb.org/DataRegulationRisk#isSourceOf"/>
            <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/DataRegulationRisk#Identify_Security_Measure"/>
         </owl:Restriction>
      </rdfs:subClassOf>
   </owl:Class>
</rdf:RDF>
```

Figure 9 Exemple d'utilisation OWL

Cette section a présenté les ontologies, leurs objectifs et leurs conceptions. Dans la prochaine section, nous nous intéressons à différentes ontologies existantes issues des domaines juridique et de la sécurité de l'information. Nous analyserons ces ontologies afin de déterminer si elles peuvent répondre aux enjeux du DRR présentés dans le premier chapitre.

Section II. Etat de l’art des ontologies pour la gestion de la conformité des données réglementées

La première section de ce chapitre nous a permis d’étudier les grands principes qui définissent les ontologies. Différents langages ont été présenté afin de répondre aux enjeux de conceptualisation formelle et partagée tel que le défini Gruber (Gruber, 1993). Enfin, la première section nous a également permis de mettre en avant différentes méthodologies de conception d’ontologies.

Alors que le nombre d'articles qui traitent des ontologies et de la gestion de la conformité n'a cessé de croître au cours des dernières années, on peut alors se demander si les ontologies existantes peuvent être utilisées en l’état ou nécessitent d’être adaptées afin de conceptualiser et représenter cette nouvelle classe de risque qu’est le DRR. Avant de nous lancer dans l’élaboration d’un nouveau modèle et la conception des différentes entités inhérentes à notre domaine, il est en effet primordial d’étudier l’existant (Ferndndez et al., 1997), (Pinto & Martins, 2004).

Notre revue se concentre sur deux principaux domaines de recherche qui sont étroitement liés au nôtre : les ontologies de gestion de la sécurité de l'information et les ontologies de conformité ou juridiques.

Cette section est organisée comme suit : la première partie introduit notre sélection et nos critères d’évaluation, la deuxième présente différentes ontologies du domaine juridique alors que la troisième partie présente différentes ontologies du domaine des technologies de l’information et de la sécurité. Enfin, nous présenterons notre analyse dans la quatrième partie.

2.1. Sélection et méthodologie d’analyse des besoins

Après la présentation de notre méthodologie de sélection d’ontologie, cette partie présente notre grille d’évaluation.

2.1.1. Sélection des ontologies

D'après Baker (Baker, 2000), la revue de littérature consiste en « une première étape et un fondement essentiel pour entreprendre un projet de recherche ». Cette étape permet de découvrir les sources pertinentes pour un sujet à l'étude. Elle apporte une contribution essentielle à la pertinence en évitant de réexaminer ce qui est déjà connu et à la rigueur de la recherche par une utilisation efficace de la base de connaissances existante (Hevner et al., 2004).

Afin de sélectionner les travaux existants, nous avons suivi les recommandations de Simons et. al. (Simons et al., 2009) décrites dans Reconstructing the Giant: On the Importance of Rigor in Documenting the Literature Search Process. Nous avons utilisé un large éventail de sources pour identifier des publications pertinentes provenant du milieu académique. Ces publications proviennent des bases de données Mendeley, Google Scholar et IEEE.

Cette revue de littérature a pour ambition d'identifier des ontologies existantes qui couvrent les domaines juridique et de la sécurité de l'information. Pour cela, nous avons effectué des requêtes en utilisant les mots-clefs suivants :

- Information Security Ontology ;
- Data Security Ontology ;
- Cybersecurity Ontology ;
- Legal Ontology ;
- Regulation Ontology ;
- Compliance Ontology ;
- IT / Information Technology Ontology.

Ces requêtes ont été effectuées en langue Française et Anglaise.

Selon Webster et Watson (Webster & Watson, 2002), une recherche documentaire comprend essentiellement l'interrogation de bases de données savantes à l'aide de mots-clés et des recherches en arrière (« backward search ») ou en avant (« foreward search ») sur la base d'articles pertinents.

Alors que la recherche vers l'arrière signifie l'étude des références des articles issus de la recherche, la recherche vers l'avant, fait référence à l'étude de sources supplémentaires qui

ont cité l'article. Suivant ces recommandations, nous avons donc effectué des recherches en avant et en arrière.

Enfin, nous nous sommes concentrés sur les articles publiés dans des journaux académiques et des conférences scientifiques afin d'assurer la sélection d'articles revus par des pairs avant leur publication. Cette contrainte permet de garantir la pertinence de la sélection comme le préconisent Webster et Watson (Webster & Watson, 2002).

2.1.2. Grille d'évaluation

Dans le premier chapitre, nous avons précisé le contexte de notre domaine d'étude. Par la définition du Data Regulation Risk, nous en avons déterminé sa multidisciplinarité. Celui-ci présente des caractéristiques qui trouvent leur origines dans les domaines juridiques, des technologies de l'information et de la conformité (Delorme et al., 2021).

Notre premier critère sera alors la spécification du domaine des ontologies étudiées :

- Juridique ;
- IT et sécurité.

Nous attribuons la classification « juridique » aux ontologies qui permettent la modélisation du corpus juridique et des modalités juridiques ; la classification « IT et sécurité » aux ontologies qui permettent d'illustrer les systèmes d'information et la sécurité de l'information.

Reprenant la définition de Piccoli (Piccoli, 2007), nos critères suivant illustrent les sous-systèmes sociaux et techniques qui composent les systèmes d'information. La structure organisationnelle et les individus liés au système d'information forment le sous-système social alors que les processus fonctionnels (aussi nommé processus d'affaire) forment le sous-système technique en impliquant les composants technologiques.

Pour cela, nous créons les critères suivants :

- Structure organisationnelle : ce critère regroupe les concepts d'organisation (entreprise), des rôles attribués aux individus et de la documentation interne ;
- Individu ;

- Composant informatique : ce critère regroupe les concepts permettant d'illustrer et de différencier les différents composants informatiques qui composent un système d'information ;
- Processus métier (traitement) : ce critère correspond à la présence de concept permettant d'illustrer des actions effectuées. Afin de correspondre à notre domaine d'étude, nous nous concentrons uniquement sur les actions effectuées sur des données (traitement) ;
- Sécurité : ce critère regroupe les concepts permettant d'illustrer et de différencier les différentes mesures de sécurité. Afin de correspondre à notre domaine d'étude, nous nous concentrons uniquement sur les mesures de sécurité de l'information.

De façon analogique, nous prévoyons différents critères pour illustrer l'application au domaine juridique :

- Corpus juridique : ce critère regroupe les concepts permettant d'illustrer et de différencier les différentes éléments du corpus juridique ;
- Modalité juridique : ce critère correspond à la présence de concept permettant d'illustrer les modalités juridiques ;
- Pluri lois : ce critère permet d'évaluer si une ontologie est spécifique à une réglementation, une famille de réglementation (e.g. réglementations traitant des données à caractère personnel) ou générique (i.e. indifférent de la famille de réglementation).

Pour chacun des critères présentés, nous attribuons la valeur « présent » si l'ontologie étudiée permet leur modélisation. L'absence de concept pouvant illustrer les composants des sous-systèmes sociaux et techniques ou le domaine juridique est signifiée par la valeur « absent ». La présence de concepts permettant une représentation partielle est représentée par la valeur « partiel ».

Enfin, notre dernier critère correspond à l'applicabilité de l'ontologie pour modéliser un risque. Ce critère peut avoir la valeur « applicable » si l'ontologie présente un ou plusieurs concepts qui illustrent le contexte et les actions qui permettent de réduire le risque d'infraction aux règles ainsi que son impact au sein des organisations.

Nous introduisons notre sélection d'ontologies par la présentation de travaux du domaine juridique.

2.2. Etat de l’art des ontologies du domaine juridique

Privacy Ontology for Legal Compliance (PrOnto)

Privacy Ontology for Legal Compliance (PrOnto) (Palmirani et al., 2018) est une ontologie qui modélise les principaux concepts du RGPD tels que les types de données et les documents, les agents et les rôles, les finalités de traitement et les bases juridiques.

Développé grâce à une approche interdisciplinaire, Methodology for building Legal Ontology (MeLOn), PrOnto modélise les principaux concepts du RGPD (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.) : types de données, documents, agents, rôles, finalités de traitement, bases juridiques, traitements et modalités déontiques pour la modélisation des droits et devoirs. Ces concepts, nommés « modules » par les auteurs, permettent la représentation de documentations internes d’une organisation, des individus, des traitements, du corpus juridique et des modalités juridiques.

Le concept « agent » permet la représentation « d’individu, d’organisation privée, de systèmes d’information, d’application, intelligence artificielle ou robot » (Palmirani et al., 2018) sans distinction. Le concept de rôle permet la modélisation des rôles qu’ont les agents dans le cadre du RGPD et ne correspond pas au rôle d’un individu dans le cadre de ses fonctions. Le concept de rôle correspond alors au responsable de traitement (« Controller »), au représentant juridique (« Representative ») et au sous-traitant (« Processor »).

Afin de représenter les traitements, PrOnto prévoit le concept « action ». Ce concept a besoin d’un « InformationObject » et provient d’un « agent ». Les auteurs ne détaillent pas le concept « InformationObject », les systèmes d’information ou les mesures de sécurité.

Enfin, les modalités juridiques sont représentées par le concept « DeonticSpecification » qui est un composant du concept « LegalRule ». PrOnto permet la modélisation de différents opérateurs déontiques qui sont le droit, l'obligation, la permission et l'interdiction. Cette ontologie permet également l’illustration de « la violation et la

conformité » en dotant les opérateurs déontiques de paramètres (propriétés) temporels et de juridiction pour considérer les droits qui ne sont effectifs que dans une réglementation nationale spécifique.

GDPR Text Extension (GDPRtEXT)

GDPR Text Extension (GDPRtEXT) (Pandit et al., 2018) permet de représenter le texte du RGPD (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.) comme un catalogue DCAT[8] contenant le texte officiel par l'utilisation de ressources RDF et d'une ontologie SKOS[9] définissant les concepts du règlement.

Le but principal des auteurs est de fournir un moyen de se référer aux concepts et termes exprimés dans le RGPD. Ceux-ci ont été liés aux points pertinents du texte GDPR en utilisant la propriété d'objet « isDefinedByproperty ». Des liens supplémentaires entre les termes ont été créés par la propriété « implies ».

Cette ontologie permet la représentation exhaustive du texte mais pas l'applicabilité du texte à un contexte. En effet, les concepts permettent une « lecture facilitée » du texte par la mise en avant des relations et référence entre les termes. Néanmoins, l'ontologie GDPRtEXT n'est pas prévue pour prendre en compte les impacts du règlement, le contexte d'une organisation ou la conformité de celle-ci.

Legal Knowledge Interchange Format Core (LKIFC)

Legal Knowledge Interchange Format Core (LKIFC) (Hoekstra et al., 2009) est une ontologie juridique présentée comme un formalisme de représentation des connaissances qui permet la traduction entre différentes bases juridiques. Basée sur deux langages OWL 2 DL (Motik et al., 2009), et LKIFRules, LKIFC permet la « modélisation de nombreux formalismes

[8] Data Catalog Vocabulary (DCAT) est un vocabulaire RDF conçu pour faciliter l'interopérabilité entre les catalogues de données publiés sur le Web.

[9] Simple Knowledge Organization System (SKOS) est une recommandation du W3C pour représenter des thésaurus documentaires, classifications ou d'autres types de langages documentaires.

utilisés dans la pratique de la construction de systèmes de connaissances juridiques » (Hoekstra et al., 2009).

LKIFC est une ontologie à trois niveaux (« top level », niveau intentionnel et niveau juridique) qui regroupe dix modules ontologiques : expression, norme, processus, action, rôle, lieu (« place »), temporalité, action juridique et rôle juridique. Chaque module représente « un cluster de concepts relativement indépendants » (Hoekstra et al., 2009).

Chaque niveau additionnel introduit une extension des définitions existantes. A chaque niveau, les concepts sont définis à un niveau d'abstraction supérieur, en ajoutant de nouvelles structures d'organisation telles que des propriétés. Selon les auteurs, cette méthodologie assure une configuration modulaire qui améliore la réutilisabilité et permet de compléter l'ontologie à n'importe quel niveau.

Le premier niveau permet d'instancier le contexte d'une action par la présence de concepts représentant les espaces spatio-temporels, les processus et les objets physiques.

Le deuxième niveau précise l'action par les concepts « agent », « organisation » et « person » pour représenter les « auteurs de l'action ». Les actions sont liées à des processus alors que les agents sont liés à un rôle pour représenter respectivement les changements et la justification de l'action. Enfin, ce niveau permet également la représentation des croyances et des intentions des agents.

Le troisième niveau est utilisé pour représenter le contexte juridique principalement par le concept « norm ». Celui-ci s'applique à des situations précises et « qualifie » celle-ci. La qualification représente les modalités juridiques par les concepts illustrant l'autorisation, l'interdiction et l'obligation.

LKIFC permet alors une représentation exhaustive du corpus juridique, du contexte d'application de ce corpus et des actions ou évènements réglementés. Néanmoins, LKIFC peine à représenter des actions pouvant survenir dans un contexte informatique par le manque de concept représentant les organisations. Par ailleurs, nous n'avons pas trouvé de concepts relatifs aux impacts des lois en cas de non-conformité.

Frame-Based Ontology (FBO)

La « Frame-Based Ontology » (FBO) de Van Kralingen (van Kralingen, 1997) cherche à adresser le problème de réutilisation des ontologies en divisant le domaine juridique en trois

concepts distincts : « norm », « act » et « concept description». Construite pour être suffisamment générique, les trois concepts de cette ontologie sont supposés permettre la modélisation de sous domaines juridiques.

Le concept « norm » correspond aux règles générales, standards et principes de comportement que les sujets se doivent de respecter. L'auteur propose huit propriétés pour établir les « norm » : un identificateur, un type, une source, un champ d'action, les conditions d'applications, le sujet correspondant à l'entité assujettie, les modalités d'applications et enfin un identificateur pour aider à la différentiation.

Le concept « act » illustre les changements et effets impliqués ou causés par les réglementations. Ce concept dispose de treize propriétés : un identificateur, une source, l'agent responsable, un type, la modalité de moyen, la modalité de manière, les aspects temporels, les aspects spatiaux, les aspects circonstanciels, la raison ou cause, le but, l'intentionnalité et enfin l'état final.

La signification et description des deux concepts précédents se trouve dans le « concept description ». Elles peuvent être des conditions nécessaires et suffisantes (dispositions déterminatives) ou encore l'applicabilité du concept. Le concept « concept description » comprend sept propriétés : le concept à décrire, le type de concept, la priorité, une source, le champ d'action, les conditions d'applicabilité et une énumération de cas d'applicabilité du concept.

Cette ontologie a été conçue pour répondre à des besoins de représentation de lois et règlementations couvrant tous les domaines existants. Cette contrainte résulte en des entités disposant de nombreuses propriétés pour chaque concept.

Privacy compliance and enforcement ontology (PC&EO)

Privacy compliance and enforcement on European healthgrids: An approach through ontology de Rahmouni et al. (Rahmouni et al., 2010) présente une ontologie pour la protection de la confidentialité des données de santé. Celle-ci repose sur les réglementations européennes.

Néanmoins, l'ontologie ne permet pas la représentation du corpus juridique. En effet, les auteurs présentent des concepts pour illustrer des éléments spécifiques des réglementations tels que le besoin de consentement ou la légitimité du traitement d'une donnée de santé (« purpose »). Les concepts de l'ontologie sont justifiés par des extraits de textes juridiques

mais le modèle n'est que partiellement réutilisable. En effet, afin d'être plus précis, les auteurs ont réécrit les réglementations en utilisant « un langage naturel », mais, en « ne gardant que le vocabulaire minimal représentant les concepts et les classifications nécessaires ». Cela conduit à un vocabulaire spécifique et non générique.

Compliance Management Ontology (CoMOn)

La Compliance Management Ontology de Abdullah et al. (CoMOn) (Abdullah et al., 2016), (Syed Abdullah et al., 2012) une ontologie destinée à fournir une conceptualisation du domaine de la gestion de la conformité pour les différentes parties prenantes. Fondée sur une analyse de documentations, sondages et interviews d'experts des domaines juridiques et de la conformité, l'ontologie a ensuite été validée par une approche d'étude de cas avec un ensemble ciblé de questions.

CoMOn est constituée de sept concepts de haut niveau qui illustrent : la culture du management, la gestion du risque, la gestion des processus d'affaire, les solutions (outils de conformité), les ressources, les obligations et le programme (de conformité).

Bien que les éléments qui constituent un système d'information ne soient que partiellement représentés, CoMOn prévoit des concepts pour adresser la gestion de risque, la représentation des documents (internes et externes), la représentation des individus et de leur rôles, du corpus juridique et des modalités qui lui sont associées.

Du fait du manque de représentation des composant informatiques, la représentation des mesures de sécurité n'est pas suffisamment granulaire et se limite à un concept nommé « Control Identification & Definition ». En effet, la représentation des composants informatiques se limitent aux outils et solutions de conformité. Il en résulte une incapacité de représenter fidèlement les processus fonctionnels.

2.3. Etat de l'art des ontologies IT et sécurité

La deuxième partie de notre sélection d'ontologies se concentre sur des travaux du domaine des technologies de l'information et leurs sécurités.

Simulating Threats to Corporate Assets (STCA)

La Security Ontology: Simulating Threats to Corporate Assets (STCA) (Ekelhart et al., 2006) permet la représentation des connaissances du domaine de la sécurité. Cette ontologie est composée de trois sous-ontologies : sécurité, infrastructure (IT) et entreprise (rôles et individus).

La sous-ontologie de la sécurité repose sur la taxonomie de la sécurité et de la fiabilité de Landwehr (Avizienis et al., 2004). Celle-ci dispose de trois concepts principaux : les attributs, les menaces et les moyens (mesure de prévention). Les attributs correspondent aux enjeux de la sécurité présentés dans le premier chapitre tels que la confidentialité, l'intégrité, etc.

La sous-ontologie de l'infrastructure permet la représentation des bâtiments, des composants informatiques qui s'y trouvent et les mesures associées à ces derniers.

La troisième sous-ontologie permet la représentation des individus de l'organisation et de leur rôle. Un individu représenter par le concept « person » est lié à un rôle par la relation « hasRole ».

L'ontologie SCTA ne prévoit pas de concepts pour illustrer les jeux de documents qui regroupent les différentes mesures de sécurité ou la gouvernance de l'organisation. Celle-ci permet néanmoins d'identifier les risques associés aux menaces présentes dans la première sous-ontologie. Enfin, les auteurs ne présentent pas de propriété de données pour les différents concepts.

Formalizing Information Security Knowledge (FISK)

Formalizing Information Security Knowledge (FISK) (Fenz & Ekelhart, 2009) est une ontologie de sécurité qui fournit une structure ontologique pour la connaissance du domaine de

la sécurité de l'information. FISK repose sur trois sous-ontologies représentant les domaines de la sécurité, de l'entreprise et la localisation.

La sous-ontologie de la sécurité présente des concepts illustrant les menaces, les vulnérabilités, les contrôles (mesures de sécurité), l'évaluation de l'efficacité des contrôles (« rating »). Les contrôles disposent d'une propriété permettant de spécifier leur type tel que correctif, détectif ou préventif ainsi que de la propriété « implementedBy » pour représenter la relation avec le concept « asset ». Afin de représenter leur origine telle qu'une réglementation ou une politique de sécurité, ils disposent de la propriété « standard control ».

La sous-ontologie de l'entreprise présente des concepts illustrant l'organisation, les individus et les « asset » qui regroupent les ressources tangibles et intangibles tels que les données, rôles des individus, logiciels (« software »), etc. La relation « OwnedBy » permet de représenter l'appartenance d'une ressource à l'organisation.

La dernière sous-ontologie n'est pas précisée afin de laisser « l'utilisateur du modèle décider du niveau de granularité nécessaire » (Fenz & Ekelhart, 2009). Le concept « location » n'est utilisé que pour représenter une liste d'emplacement selon le contexte d'utilisation.

Du fait du concept de représentation des composants informatiques, la représentation des mesures de sécurité est suffisamment granulaire cependant, l'ontologie FISK ne prévoit pas de concepts pour illustrer les jeux de documents qui regroupent les différentes mesures de sécurité ou la gouvernance de l'organisation. L'ontologie ne permet également pas la représentation du corpus juridique.

An Ontology of Information Security (OIS)

Herzog et al. présente An Ontology of Information Security (OIS) (Herzog et al., 2011), une ontologie de la sécurité de l'information qui modélise les ressources (« assets »), les menaces, les vulnérabilités, les objectifs de sécurité, la stratégie de défense et les contre-mesures. L'ontologie se veut être très fournie pour répondre à son objectif « d'être utilisée comme vocabulaire général, feuille de route et dictionnaire extensible du domaine de la sécurité de l'information » (Herzog et al., 2011).

Le concept « asset » est relié aux mesures de sécurité par la relation « protected by ». Celui-ci permet également de représenter les mesures qui protège les objectifs de sécurité. En outre, il est également utilisé pour la représentation des composants informatiques, des individus et des données.

L’ontologie se veut générique et par conséquent ne prévoit pas la représentation du contexte d’une organisation tels que la gouvernance, les processus fonctionnels, le rôle des individus ou le contexte juridique.

Ontology of Information Security in Enterprises (OISE)

L’Ontology of Information Security in Enterprises (OISE) de Shiavone et al. (Schiavone et al., 2014) diffère des autres ontologies présentées dans cette partie. Les auteurs présentent la création d'une ontologie spécifique à l'entreprise qui décrit celle-ci comme un système complexe afin d’aider à l’identification des mesures de sécurité nécessaires.

OISE repose sur les concepts de haut niveau « Enterprise », « Business Capability », « Enterprise Domaine Ressource », « Enterprise Extended », « Enterprise Reference Architecture » et « Enterprise Value » pour représenter les organisations. Les deux premiers permettent la représentation des missions, de la vision et des objectifs de l’organisation.

Le concept « Enterprise Domaine Ressource » est composé de trente-quatre sous classes permettant la représentation des processus d’affaire, des éléments d’un système d’information et les autres ressources nécessaires à l’exécution des objectifs et missions.

Le concept « Enterprise Extended » permet la représentation du contexte de l’organisation par la représentation des prestataires et autres entités en interaction avec l’organisation concernée.

Alors que le concept « Enterprise Value » permet d’illustrer la valeur monétaire des autres concepts, le concept « Enterprise Reference Architecture » est utilisé pour illustrer des standards, politiques et cadriciels.

L’ontologie ne couvre donc pas ni le domaine de la sécurité informatique ni le domaine juridique mais présente une ontologie complète pour la représentation des organisations.

Ontology of Cybersecurity Operational Information (OCOI)

Afin de structurer les informations du domaine de la cybersécurité et les spécifications de l'industrie, Takahashi & Kadobayashi propose l’Ontology of Cybersecurity Operational Information OCOI (Takahashi & Kadobayashi, 2014). L’objectif de OCOI est de décrire « qui (rôle) utilise quel type d'informations (informations du domaine de la cybersécurité) dans quel but (opération) ».

OCOI repose sur trois concepts de haut niveau qui représente les opérations, les rôles et les informations du domaine de la cybersécurité. Les opérations correspondent à la gestion des composants informatiques, des incidents et de la collecte et génération d'information du domaine de la cybersécurité. Les rôles permettent la représentation d'individus impliqués dans la réponse à incident tels que les « ResponseTeam » ; « Administrator » ou « Coordinator ». De manière similaire, les informations du domaine de la cybersécurité se limitent aux informations relatives à la gestion d'incident de sécurité informatique.

Cette ontologie ne présente pas de concept permettant la représentation de corpus juridique, de la structure organisationnelle, des processus fonctionnels d'une organisation et se limite aux concepts relatifs aux individus, à certains composants informatiques et à leur implication dans la gestion d'incidents.

Security Asset-Vulnerability Ontology (SAVO)

Security Asset-Vulnerability Ontology (SAVO) est une ontologie présentée par Vorobiev & Bekmamedova (Vorobiev & Bekmamedova, 2010). SAVO repose sur des sous-ontologies d'attaque, de défense, d'algorithme de sécurité et de fonction de sécurité. Du fait de ses nombreux concepts, SAVO permet la représentation des différents éléments techniques d'un système d'information. Par exemple, SAVO propose les concepts de donnée, de compte, des composants relatifs au applications (tels que la mémoire, le processeur, le compte administrateur).

Cette ontologie ne présente pas de concept permettant la représentation de corpus juridique ou des processus fonctionnels d'une organisation. Du fait de son domaine (les logiciels), SAVO ne présente pas de concepts liés à la structure organisationnelle et se limite aux concepts relatifs aux individus et à leur implication dans l'utilisation d'un logiciel.

Cloud Security And Compliance Ontology (CSCO)

La Compliance Cloud Security And Compliance Ontology (CSCO) (Hendre & Joshi, 2015) de Hendre & Joshi est une ontologie pour modéliser les menaces, les politiques et mesures de sécurité de l'informatique en nuage.

L'ontologie propose des concepts pour illustrer les modèles de sécurité, les contrôles de sécurité, les menaces et les fournisseurs de services en nuage. Les modèles de sécurités

correspondent aux cadriciels, guides de bonnes pratiques ou autres jeux de documents disponibles. Ils sont reliés aux mesures de sécurité par la propriété « support ».

CSCO ne permet ni la représentation de corpus juridique, ni celle d'une organisation et de son système d'information. Elle se concentre, dans le domaine de l'informatique en nuage, sur les mesures de sécurité, les menaces qu'elles adressent et les documents qui les décrivent.

2.4. Analyse et résultats

Alors que le nombre d'articles qui traitent des ontologies et de la gestion de la conformité n'a cessé de croître au cours des dernières années, il n'existe toujours pas, à notre connaissance, de travaux existants qui correspondent parfaitement à nos besoins.

Parmi les ontologies du domaine juridique, seules les ontologies PrOnto et CoMon permettent de représenter fidèlement les structures organisationnelles, la gouvernance et les documentations internes des entreprises. La conceptualisation de la sécurité et des mesures de sécurité n'est que partiellement possible ou totalement absente des ontologies juridiques. Contrairement aux ontologies du domaine des technologies de l'information et de la sécurité, toutes les ontologies du domaine juridique permettent la représentation (parfois partiellement) des modalités juridiques.

Parmi les ontologies du domaine des technologies de l'information et de la sécurité, nous n'en avons pas trouvé qui permettent de représenter les corpus juridiques ou les modalités juridiques.

Le Tableau 3 Analyse d'applicabilité d'ontologie au DRR présente pour chaque ontologie présentée, notre analyse en détail de chacun des critères d'évaluation.

Nom	Domaine	Système d'information					Juridique			Risque
		Structure organisationnelle (documentation interne)	Individu	Composant informatique	Processus métier : traitement	Sécurité	Corpus juridique	Modalité juridique	Multi lois	
PrOnto	Juridique	Présent	Présent	Partiel	Présent	Absent	Présent	Présent	Absent	Présent
GDPRtEXT	Juridique	Absent	Absent	Absent	Absent	Absent	Présent	Absent	Absent	Absent
LKIFC	Juridique	Absent	Présent	Absent	Partiel	Absent	Présent	Présent	Présent	Partiel
FBO	Juridique	Absent	Absent	Absent	Absent	Absent	Présent	Présent	Présent	Absent
PC&EO	Juridique	Absent	Présent	Absent	Absent	Partiel	Non	Partiel	Absent	Absent
CoMOn	Juridique	Présent	Présent	Partiel	Partiel	Partiel	Présent	Présent	Présent	Présent
STCA	IT et sécurité	Absent	Présent	Présent	Absent	Présent	Absent	Absent	Absent	Présent
FISK	IT et sécurité	Absent	Présent	Partiel	Absent	Présent	Absent	Absent	Absent	Partiel
OIS	IT et sécurité	Absent	Partiel	Présent	Absent	Présent	Absent	Absent	Absent	Présent
OISE	IT et sécurité	Présent	Présent	Présent	Présent	Absent	Absent	Absent	Absent	Absent
OCOI	IT et sécurité	Absent	Partiel	Présent	Absent	Présent	Absent	Absent	Absent	Présent
SAVO	IT et sécurité	Absent	Partiel	Partiel	Absent	Présent	Absent	Absent	Absent	Présent
CSCO	IT et sécurité	Partiel	Absent	Absent	Absent	Présent	Absent	Absent	Absent	Partiel

Tableau 3 Analyse d'applicabilité d'ontologie au DRR

Nous comprenons que les domaines respectifs des ontologies étudiées ne sont pas nécessairement adaptés à un domaine transversal qu'est celui du DRR. A la croisée entre les domaines juridiques, des technologies de l'information et de la sécurité, la représentation du domaine du DRR nécessite alors la conceptualisation d'un nouveau modèle. Des ontologies présentées ci-avant, nous avons sélectionné les concepts qui correspondent à des éléments de notre définition du DRR et qui permettent sa représentation. Ces concepts permettent la représentation (totale ou partielle) :

- D'une entreprise et de son système d'information ;
- De la sécurité (de l'information) ;
- Des réglementations ;
- De contexte de l'entreprise (applicabilité des réglementations) ;

Notre modèle repose néanmoins sur des concepts déjà existants. Certains nécessitent une adaptation pour répondre à nos besoins.

2.4.1. Les concepts issues des ontologies juridiques

Nous utiliserons les concepts « Norm » et « Act » présents dans la « Frame-Based Ontology » de Van Kralingen (van Kralingen, 1997) pour représenter respectivement le corpus juridique et les modalités juridiques.

Pour le concept « Norm », nous conservons les propriétés de données suivantes :

- « Norm identifier » par la création de la propriété de données « NormeName » ;
- « Promulgation » par la création de la propriété de données « NormeSource ».

Les propriétés de périmètre d'application « scope », conditions d'applications, modalité juridique et « act identifier » seront transformées en propriétés d'objet (relation) avec les autres concepts de notre modèle. Ceux-ci seront décrits dans la troisième section de ce chapitre.

Pour le concept « Act », nous conservons les propriétés de données suivantes :

- « Act identifier » par la création de la propriété de données « ActName » ;
- « Act type » pour représenter le type de modalité juridique.

Les propriétés de source de promulgation, de périmètre d'application « scope », de moyens, de temporalité, de causalité et d'agent seront transformées en propriétés d'objet

(relation) avec les autres concepts de notre modèle. Ceux-ci seront décrit dans la troisième section de ce chapitre.

De la Privacy Ontology for Legal Compliance (PrOnto) (Palmirani et al., 2018), nous utiliserons les concepts suivants :

- « Action » pour représenter les traitements effectués sur les données ;
- « Role » pour représenter les actions et activités qu'un individu est missionné et autorisé d'effectuer ;
- « Document » que nous complèterons pour une représentation exhaustive des jeux de documents d'une organisation.

Alors que PrOnto propose le concept « Agent » pour la représentation « d'individu, d'organisation privée, de systèmes d'information, d'application, intelligence artificielles ou robot », nous proposons la création de plusieurs concepts pour améliorer la granularité de notre modèle et limiter le risque de mauvaise utilisation.

Nous détaillerons dans le prochain chapitre la création des concepts permettant la représentation des organisations, individus et éléments d'un système d'information.

De l'ontologie Legal Knowledge Interchange Format (LKIF) (Hoekstra et al., 2009), nous utiliserons les concepts :

- « Organisation » ;
- « Person » ;
- « Process ».

Nous modifierons et compléterons le concept « Process » pour représenter un ensemble d'activités corrélées ou en interaction qui implique des composants informatiques et des individus.

De manière similaire, nous modifierons et compléterons le concept « Organisation » pour représenter une personne physique ou morale, un organisme d'autorité publique exerçant une activité quelle que soit sa forme juridique, y compris les sociétés de personnes ou les associations.

2.4.2. Les concepts issues des ontologies IT et sécurité

L'ontologie Security Ontology: Simulating Threats to Corporate Assets (STCA) (Ekelhart et al., 2006) a été conçue pour répondre à des besoins de représentation de menaces aux ressources d'une entreprise. Ces besoins ont conduit à une utilisation de concepts qui ne sont pas directement liés à notre domaine.

De ce fait, nous n'utiliserons que les concepts :

- « CounterMeasure » ;
- « ElectronicDevice ».

Nous modifierons et adapterons le concept « CounterMeasure » afin de représenter les mesures de protection et de réponse. Nous modifierons et adapterons également le concept « ElectronicDevice » afin de représenter les différents composants informatiques impliqués dans les processus fonctionnels.

Nous utiliserons les trois sous-ontologies représentant les domaines de la sécurité, de l'entreprise et la localisation présentées dans Formalizing Information Security Knowledge (FISK) (Fenz & Ekelhart, 2009). Cependant, cette ontologie a été conçue pour répondre à des besoins de représentation de la sécurité de l'information d'une entreprise. Ces besoins ont conduit à une utilisation de concepts qui ne sont pas directement liés à notre domaine.

De ce fait, au sein de ces sous-ontologies, nous ne reprenons que les concepts :

- « Control » par la création du concept « Security_Measure » ;
- « Asset » par la création de plusieurs concepts permettant une plus grande granularité ;
- « Location » par la création du concept « Country ».

Le concept « Security_Measure » nous permet, en sus des mesures de protection et de réponse, de représenter les mesures de sécurité d'identification, de détection, et de reprise par la création des sous concepts respectifs.

Conclusion

Les ontologies sont reconnues pour faciliter le partage, la réutilisation, et l'utilisation des connaissances et permettent la modélisation d'un modèle par des conceptualisations explicites de celui-ci. Le rôle des ontologies est de faciliter la construction d'un modèle de domaine en fournissant un vocabulaire de termes et de relations avec lesquels il est possible de le modéliser. Pour se faire, différents langages ont été développés pour répondre à des besoins spécifiques. Ceux-ci permettent notamment d'illustrer et d'expliciter les relations entre différents concepts.

Afin de modéliser fidèlement le domaine du DRR, il est nécessaire d'avoir un modèle présentant les concepts présents dans sa définition. Il n'existe toujours pas, à notre connaissance, de travaux existants qui correspondent parfaitement à ces besoins. Ceux-ci s'illustrent par la nécessité de permettre la représentation :

- D'une entreprise et de son système d'information ;
- De la sécurité (de l'information) ;
- Des réglementations ;
- De contexte de l'entreprise (applicabilité des réglementations).

Seule la présence de l'ensemble de ces concepts dans un modèle unique permet d'adresser pleinement la problématique du DRR et de sa gestion. Nous proposons dans le prochain chapitre un modèle qui tente à satisfaire ces exigences.

Chapitre 3 - Proposition d’une ontologie multi-domaines pour l’aide à la gestion du risque de conformité des données réglementées

Ce chapitre sera consacré à la proposition de notre solution. D’une part nous présentons les objectifs et les fondamentaux de notre modèle, d’autre part, nous expliquerons notre solution et sa conception.

Nous avons précédemment mis en avant la nécessité de proposer un moyen efficient et fiable de fournir les informations nécessaires au management du risque. Ces informations sont notamment utilisées aux étapes d’identification et d’évaluation du risque. Il convient de proposer une solution qui réponde aux attentes des utilisateurs présentés dans le premier chapitre, à savoir :

- La direction des systèmes d’information ;
- La direction des risques et de la sécurité ;
- La direction juridique.

Le Chapitre 2 - Représentation des connaissances en conformité : usage des ontologies, limites et perspectives a pour sa part présenté les ontologies comme solutions adaptées pour faciliter la gestion de l’impact des stratégies IT sur la maîtrise du risque réglementaire. Ce second chapitre a toutefois exposé le manque d’adéquation entre les ontologies existantes et la problématique que nous souhaitons adresser. En effet, malgré d'importantes contributions soutenant les organisations dans l'évaluation et la gestion de leurs risques, il existe encore un besoin de méthodologies et de modèles pour identifier les risques multidisciplinaires comme le Data Regulation Risk.

Le Chapitre 3 - Proposition quant à lui, présente notre proposition de solution reposant sur une ontologie. Ce chapitre est divisé en trois sections.

La première section définit les objectifs de notre modèle. Ceux-ci sont doubles. Notre modèle doit être en mesure de retranscrire précisément les règlementations. Cela implique d’être capable d’illustrer les modalités juridiques, leurs évolutions et leurs impacts sur les organisations. Nous devons également nous assurer dès les premières étapes de conception que

notre modèle est pertinent, utilisable et réutilisable. Enfin, cette section pose le contexte d'utilisation de notre modèle par la description des utilisateurs et les attentes envisagées lors de sa conception.

La deuxième section explique la conception de notre modèle. Elle présente l'ensemble des concepts utilisés pour représenter fidèlement notre domaine. Ceux-ci sont au nombre de 15 et sont répartis en quatre sous domaines :

- Le sous domaine de l'entreprise ;
- Le sous domaine de la sécurité ;
- Le sous domaine réglementaire ;
- Le sous domaine de la localisation.

Cette section introduit également les relations ou propriétés d'objet qui permettent de relier ces concepts. Celles-ci sont d'autant plus importantes qu'elles permettent de mettre en place notre logique sémantique et par conséquent la logique de notre modèle. Cette deuxième section a également pour but de présenter les fonctionnalités de notre modèle qui sont les processus d'instanciation, de consultation et de recherche d'information.

La troisième et dernière section de ce chapitre nous permet de présenter des exemples d'instanciation et d'utilisation de notre modèle par la représentation de deux lois :

- L'instanciation de l'EAR Supplement No. 18 to part 734 ;
- L'instanciation de l'article 32 du RGPD.

Avec l'aide de l'instanciation de l'EAR Supplement No. 18 to part 734, cette section également des exemples d'utilisation des processus de recherche, de consultation et d'instanciation.

Section I. Contexte et objectif de notre proposition

Comme mentionnée par Uschold et al. (Uschold & King, 1995), l'étape fondamentale du développement d'une ontologie consiste à déterminer, spécifier et valider les raisons pour laquelle celle-ci est construite et quelles sont ses utilisations prévues. Définir l'objectif d'une ontologie est également présenté comme l'étape clé des activités de spécification présentées dans la méthodologie Methonlogy (Ferndndez et al., 1997).

Dans cette section, nous allons dans un premier temps définir les objectifs de notre modèle. Ceux-ci sont doubles. Notre modèle doit être en mesure de retranscrire précisément les règlementations. Cela implique d'être capable d'illustrer les modalités juridiques, leurs évolutions et leurs impacts sur les organisations. Enfin, nous devons nous assurer dès les premières étapes de conception que notre modèle est pertinent, utilisable et générique.

Dans un second temps, nous poserons le contexte d'utilisation de notre modèle par la description des utilisateurs et leurs attentes envisagées lors de sa conception.

1.1. Objectif de notre proposition

Afin de répondre à la recrudescence de promulgation de réglementations visant à réglementer le traitement des données qui induit une complexification de la gestion de la conformité pour les entreprises réglementées, notre modèle doit être en mesure de retranscrire précisément les règlementations. Cela implique d'être capable d'illustrer :

- Les modalités juridiques qui relèvent de l'interdiction, l'autorisation et le choix libre (Stage, 2002) ;
- Leurs évolutions dans le temps ;
- Leurs impacts pour appréhender les impacts métiers et opérationnels des exigences réglementaires.

Notre ontologie doit être structurée pour gérer le contexte juridique des différentes entreprises tout en se voulant indépendante des méthodologies de gestion des risques. Pour cela, nous devons nous assurer dès les premières étapes de conception que notre modèle est :

- Pertinent, en reposant sur une méthodologie documentée, éprouvée et validée scientifiquement ;

- Utilisable et réutilisable en reposant sur une base de données lexicales tierce et libre d'accès.

Ces deux points nous permettent de garantir l'adoption de notre ontologie et son utilisation en tant que support à la gestion des risques des organisations. En effet, les entreprises qui souhaitent mettre en œuvre une gestion des risques efficace et rentable tout en assurant leur conformité ont besoin d'une compréhension des exigences réglementaires facilitée.

1.1.1. Les objectifs de retranscription

Retranscrire les modalités juridiques

Notre travail consiste en la conception d'un système capable de précisément et intelligiblement représenter les différentes modalités juridiques auxquelles une organisation est confrontée lors de la gestion d'un Data Regulation Risk (DRR). Notre ontologie se doit d'être structurée pour gérer le contexte juridique de diverses entreprises tout en restant indépendante des méthodologies de gestion des risques.

La modalité correspond à un évènement qui modifie la naissance ou l'exigibilité d'une obligation. Dans le cadre de la gestion du DRR, les modalités juridiques sont des outils à disposition des législateurs dont la vocation est d'indiquer aux organisations le degré de possibilité de la réalisation d'un évènement ou d'une action en fonction de certaines circonstances. Cette liberté peut s'illustrer par une obligation, une interdiction ou un libre choix (Stage, 2002).

Les législateurs qui souhaitent enjoindre à des organisations de faire ou ne pas faire une action doit alors exprimer de manière claire et précise les différentes obligations. En d'autres termes, les modalités juridiques apparaissent alors comme des indicateurs de marge de manœuvre.

D'un point de vue pragmatique, les modalités déontiques s'instrumentalisent en trois typologies de volonté, celles-ci sont présentées dans le

Tableau 4 Modalités déontiques, marge de manœuvre et de survenance :

- Une marge de manœuvre nulle et une marge de survenance nulle. La possibilité pour évènement ou une action d'avoir lieu est de 0%. Dans tel contexte, un évènement ou

une action ne peut pas avoir lieu. Autrement dit, dans ce contexte, le législateur interdit l'évènement ou l'action.

- Une marge de manœuvre nulle et une marge de survenance maximale. La possibilité pour évènement ou une action d'avoir lieu est de 100%. Dans tel contexte, un évènement ou une action doit avoir lieu. Autrement dit, dans ce contexte, le législateur force l'évènement ou l'action.
- Une marge de survenance et de manœuvre comprises entre 0 et 1. Dans tel contexte, un évènement ou une action peut avoir lieu. Autrement dit, dans ce contexte, le législateur n'interdit pas ou ne force pas l'évènement ou l'action.

	Interdiction	Obligation	Choix Libre
Marge de Manœuvre (M^1)	$M^1=0$	$M^1=0$	$0<M^1<1$
Marge de survenance (M^2)	$M^2=0$	$M^2=1$	$0<M^2<1$

Tableau 4 Modalités déontiques, marge de manœuvre et de survenance

En se plaçant en tant qu'organisation ou entreprise assujettie à une règlementation, les modalités juridiques peuvent alors être représentées par une interdiction (marge de manœuvre et marge de survenance nulles), une obligation (marge de manœuvre nulle et une marge de manœuvre égale à 1) ou de libre choix (marge de manœuvre et marge de survenance comprises entre 0 et 1).

Le

Tableau 4 Modalités déontiques, marge de manœuvre et de survenance illustre ces différentes modalités juridiques. Il en convient que, dans le cadre d'une marge de manœuvre nulle, la gestion du système d'information et sa sécurité sont fortement impactées par la gestion du DRR et de la conformité.

Retranscrire les évolutions juridiques

Nous avons comme objectif de conceptualiser un modèle qui permet de retranscrire fidèlement les contrôles divergents ou spécifiques à une loi ainsi qu'être en mesure de se focaliser sur un environnement informatique propre à une organisation.

A cela, le modèle doit également être conçu pour intégrer l'évolution rapide du paysage réglementaire. Ces changements peuvent par exemple prendre la forme d'une modification d'une modalité juridique tel qu'un changement de marge de survenance, l'ajout ou la suppression de modalités. De manière similaire, la promulgation d'une nouvelle règlementation ou le changement de champ d'application d'une règlementation existante doivent pouvoir être pris en compte par notre modèle.

Par exemple, les restrictions concernant les modules de chiffrement dans le cadre de l'EAR sont spécifiées dans Part 734 - Scope of the Export Administration Regulations, section 18. Celle-ci se réfère au Federal Information Processing Standard 140 qui spécifie les exigences de sécurité qu'un module cryptographique doit satisfaire. La prise en compte de l'évolution des versions de ce standard par une organisation est nécessaire pour assurer sa conformité. Un des impacts technologiques qu'implique le passage à la version actuelle est l'interdiction d'utilisation « Triple Data Encryption Algorithm » d'ici 2023 [10]. De ce fait, toute organisation utilisant cet algorithme après 2022 s'expose à un risque de non-conformité reposant sur un choix technique de sécurité de l'information.

Représenter les impacts métiers et opérationnels

Les utilisateurs ont besoin d'un accès simple aux informations nécessaires pour comprendre les impacts métiers et opérationnels des exigences réglementaires. La complexité de cet exercice de mise en conformité réside dans la nécessité de traduire les contraintes et exigences réglementaires en des termes techniques, organisationnels et opérationnels. En effet, le modèle doit être en mesure de fournir aux utilisateurs des informations et instructions pragmatiques et concrètes en se basant sur des textes juridiques génériques et parfois abstraits. À cette abstraction et généricité, s'adjoint la strate des modalités déontiques qui complexifie la compréhension des lois et l'appréhension des impacts.

Nous pouvons illustrer la complexité d'interprétation et traduction par l'exemple de la gestion des droits d'accès aux données réglementées par deux lois d'export de biens à double usage : le règlement européen précise que le contrôle d'accès aux données réglementées doit être basé uniquement sur l'emplacement individuel et les frontières géographiques, tandis que

[10] Les problématiques de migration du Federal Information Processing Standard (140-1) au Federal Information Processing Standard (140-3) sont adressées sur le site du COMPUTER SECURITY RESOURCE CENTER accessible via le lien suivant : https://csrc.nist.gov/projects/.

l'Export Arm Regulations (EAR) des États-Unis inclut également des contrôles basés sur la citoyenneté. Cela implique, pour une entreprise souhaitant être conforme, de mettre en place une gestion des identités et accès correspondant à un ensemble des processus mis en œuvre pour la gestion des habilitations de ses utilisateurs à son système d'information. Dans ce cas précis, celle-ci doit se baser sur la localisation des individus au moment T de l'accès et d'une vérification de l'identité pour garantir qui a accès à quelle information à travers le temps. D'un point de vue opérationnel, cela implique ainsi une gestion de la création, la modification, et les droits d'accès de chaque identité numérique interagissant avec les différentes ressources informatiques.

Nous allons maintenant décrire comment garantir dès les premières étapes de conception que notre modèle est pertinent, utilisable et générique.

1.1.1. Pertinence du modèle

Afin de garantir l'adoption de notre ontologie et son utilisation en tant que support à la gestion des risques des organisations, notre modèle doit être pertinent, utilisable et générique. Nous avons avancé dans le premier chapitre que la gestion et représentation des connaissances se heurte à plusieurs difficultés telles que la gestion de sources de connaissance différentes ou la complexité de permettre un partage et un accès simplifié aux informations pertinentes pour les différents utilisateurs.

Afin de pouvoir appréhender les réglementations de la manière la plus optimale possible, il est nécessaire pour les organisations d'avoir une vue d'ensemble :

- des besoins et actions requises pour assurer leur mise en conformité ;
- de l'état des lieux et avancement de la mise en conformité ;
- des spécificités techniques des systèmes d'information ;
- des guides, supports et autres formes de support d'aide à la conformité.

Afin de pouvoir développer une ontologie qui réponde à nos objectifs, il nous faut nous poser certaines questions tout au long de son élaboration afin de les atteindre. En reprenant la liste de questions établie par Visser et Bench-Capon (Visser & Bench-Capon, 1998) utilisée lors de la comparaison d'ontologies juridiques, nous pouvons nous poser les questions suivantes :

- Est-ce que l’ontologie définit clairement les relations et les concepts ?
- Est-ce que le vocabulaire utilisé répond aux besoins et intuitions des experts du domaine concerné ?
- Est-ce que l’ensemble des concepts et relations sont pertinents pour les tâches, méthodes et domaines ?
- Est-ce que l’exhaustivité du domaine est couverte ?
- Est-ce que l’ensemble des concepts et définitions sont cohérents ?
- Est-ce que l’ontologie fournit une base utilisable pour une représentation informatique ?
- Est-ce que l’ontologie est réutilisable, et jusqu’à quelles limites, pour différentes méthodes, tâches, domaines et sous domaines ?

Cette liste a pour ambition de poser les bases de notre réflexion en regroupant les travaux de recherches antécédents. L’ensemble de ces points de références seront utilisés tout au long de l’élaboration de notre modèle pour assurer sa pertinence et sa pérennité. Nous reposons sur une méthodologie éprouvée pour assurer que notre modèle puisse être compris, utilisé et accepté par les différents utilisateurs.

Enterprise Model Approach

De surcroît, nous avons décidé d'adopter l'approche « Enterprise Model Approach » (Pinto & Martins, 2004) présentée dans le Chapitre 2. Le fait de reposer sur une méthodologie documentée, éprouvée et validée scientifiquement par les pairs nous apporte un dégré de confiance quant à la poursuite des objectifs fixés. Cette approche par étape, largement répandue, nous offre une liberté de représentation suffisante et appropriée à un domaine pluridisciplinaire tel que le risque réglementaire. Articulée autour de quatre étapes principales, elle consiste en un squelette méthodologique qui comprend :

- l'identification de la finalité ;
- la construction de l’ontologie ;
- l'évaluation ;
- la documentation.

Enfin, par opposition à l'approche classique « bottom-up » et « top-down » pour identifier les principaux termes de notre ontologie, nous optons pour l'approche middle out

présentée dans (Uschold & King, 1995). Cette approche nous permet entre autre, d'identifier les concepts primaires de notre ontologie avant de passer à la spécialisation ou à la généralisation des termes, uniquement si cela est nécessaire. L'approche « middle-out » conduit implicitement à des concepts plus stables, à moins de retravail et d'efforts tout en augmentant la clarté du document, en particulier pour les utilisateurs possédant moins de compétences techniques.

Wordnet

Nous souhaitons que notre modèle réponde aux besoins d'amélioration des communications entre des acteurs de milieux différents, utilisant des outils différents, ayant des connaissances et des vocabulaires différents (Uschold & Gruninger, 1996). Le modèle a alors pour but de définir une base ou langage de communication accepté par les différentes parties, d'organiser et de faciliter l'accès à l'information pour assurer un partage efficient et transverse des connaissances au sein d'une organisation.

En ce qui concerne la clarté, qui est le fondement de l'utilisabilité et de la réutilisabilité d'une ontologie, nous avons besoin d'une base de données terminologique mondialement connue, facilement accessible, éprouvée et acceptée. Cela permet d'atteindre notre ambition de définir une base ou langage de communication accepté par les différentes parties, d'organiser et de faciliter l'accès à l'information pour assurer un partage efficient et transverse des connaissances au sein d'une organisation.

Nous avons décidé d'utiliser, lorsque cela est possible, les terminologies de la base de données WordNet développée par l'Université de Princeton [11]. Word-Net est « une grande base de données lexicales de noms, verbes, adjectifs et adverbes regroupés en ensembles de synonymes cognitifs (Synsets), chacun exprimant un concept distinct. Les Synsets sont interconnectés au moyen de relations conceptuelles-sémantiques et lexicales. » Chaque concept, relation ou attribut de notre ontologie est identifié et différencié par un Synset unique à l'aide de l'ID Synset.

Par exemple, la relation Gouvern peut faire référence à : « exercer une autorité sur ; tel qu'une nation » (Synset ID : 202586619), « oriente ou influence fortement le comportement de

[11] La base de donnée Wordnet est disponible sur le site de Princeton accessible via le lien suivant : https://wordnet.princeton.edu.

» (Synset ID : 202442205) ou encore « met en conformité avec des règles, des principes ou des usages ; imposer des réglementations » (Synset ID : 202511551). En spécifiant le Synset ID, le risque de mauvaise interprétation est donc fortement réduit tout en préservant l'interopérabilité sémantique. Par exemple, nous utiliserons le Synset ID : 202511551 pour la relation Govern dans notre ontologie.

Ce choix d'utiliser une terminologie tierce et libre d'accès présente deux avantages principaux.

Premièrement, il permet de garantir une réutilisabilité forte de notre modèle en garantissant une compréhension des concepts et relations par un utilisateur externe. Cela est d'autant plus important pour notre modèle car il répond à une volonté future de fusion avec différentes ontologies de domaines spécifiques existantes telles que des ontologies de sécurité informatique ou de gestion de risque transverses.

Dans un second temps, ce choix nous permet d'anticiper les problématiques linguistiques liés aux textes juridiques. Nous en avons identifié deux :

Premièrement, les textes juridiques et autres documentations ne sont pas tous disponibles dans chaque langue et notamment la langue anglaise. Ce manque de traduction officielle peut introduire des risques d'incapacité de compréhension des textes ou de mauvaise interprétation des différents textes.

Secondement, les textes juridiques et autres documentations peuvent introduire des syntaxes et terminologies hétérogènes. Dans le premier cas, plusieurs termes sont utilisés pour représenter le même concept. En effet, l'hétérogénéité terminologique peut se traduire par différents noms utilisés pour définir la même entité (synonymie). Dans le second cas, le même nom est utilisé pour désigner des entités distinctes (polysémie).

Concept, propriété et instance

Notre modèle repose sur une ontologie composée d'un ensemble de concepts représentant notre domaine d'étude, de propriétés (ou relations) et d'instances de classe. Chaque entité utilisée pour représenter notre domaine d'étude possède un IRI[12]

[12] Les Internationalized Resource Identifier sont présentés par le World Wide Web Consortium sur leur site accessible via le lien suivant : https://www.w3.org/TR/owl2-syntax/#IRIs.

(Internationalized Resource Identifier), un nom, une description, une version, des propriétés d'objet et de donnée qui lui sont propres ainsi que zéro à plusieurs instances.

Instance

Les instances sont utilisées pour représenter les objets du domaine d'étude. Elles sont spécifiques au contexte d'utilisation (ici, celui d'une organisation) et reposent sur les bases de connaissances présentées précédemment. En effet, les instances des composants informatiques ou des processus fonctionnels sont liés au système d'information et aux modèles opératoires propre à chaque organisation et seront naturellement uniques et spécifiques. Par conséquent, la création, suppression et modification des instances reposent entièrement sur les utilisateurs d'une organisation.

Propriétés et relations

Il existe trois types de propriétés : les propriétés d'objet (object properties), les propriétés de données (datatype properties) et les propriétés d'annotation (annotation properties).

Les propriétés d'objet sont des relations binaires entre deux instances qui correspondent aux relations. Ces propriétés peuvent avoir des caractéristiques qui leur sont propres et qui permettent d'ajouter la partie logique à notre modèle.

Les propriétés de données sont utilisées pour lier des instances à des valeurs de donnée. Par exemple, une propriété de données peut lier une instance d'un individu « Héloïse » à la donnée « 2 » de type integer pour représenter son âge. Ces propriétés sont aussi parfois appelées attribut de concept.

Les propriétés d'annotation peuvent être utilisées pour ajouter des informations (métadonnées) aux classes, aux individus et aux propriétés d'objet et de donnée.

En premier lieu, il s'agit de déterminer les propriétés ou relations utilisées pour notre modèle. Celui-ci sera constitué de deux types de propriété d'objet : les propriétés de caractéristique et les propriétés d'action. Dans les premières, le concept patient n'effectue pas d'action directement sur le concept agent mais permet de préciser les attributs de ce dernier. Les relations de caractéristique permettent de représenter les liens entre les différents concepts du modèle. Les relations d'action sont utilisées lorsqu'un concept agent effectue une action directe sur un concept patient. La Figure 10 Typologie des relations du modèle illustre la hiérarchie des relations de caractéristique et des relations d'action.

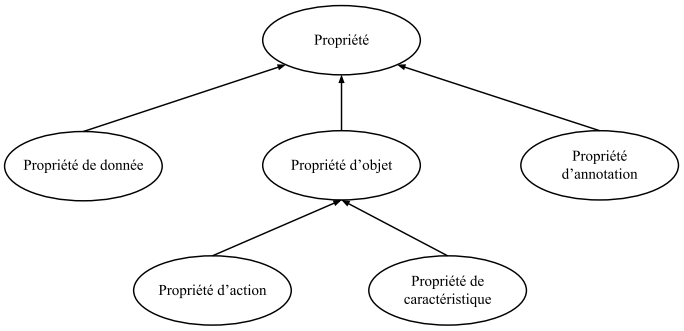


Figure 10 Typologie des relations du modèle

Afin de répondre à nos objectifs d'interopérabilité et d'intégration avec d'autres ontologies de domaine ou de haut niveau, nous avons utilisé les termes définis par WordNet. Afin d'assurer l'élaboration d'un modèle suffisamment granulaire pour adresser pleinement les problématiques du DRR et d'élaboration d'un modèle indépendant, non spécifique et adaptable à différents contextes juridiques, nous avons ponctuellement ajouté des précisions quant aux définitions des Synsets.

Concepts et classes

Les concepts (aussi nommés classes) sont utilisés pour représenter des ensembles contenant des instances. Ils sont décrits à l'aide de descriptions reposant sur des formules formelles qui énoncent précisément les conditions d'appartenance à la classe.

Dans la prochaine partie, nous détallions le contexte d'utilisation de notre modèle.

1.2. Contexte et terrain

Par la proposition de notre solution, nous souhaitons apporter le support nécessaire aux activités des individus impliqués dans la gestion du DRR. Dans le premier chapitre nous avons mentionné ces trois types d'utilisateur qui sont :

- La direction des systèmes d'information ;
- La direction des risques et de la sécurité ;
- La direction juridique.

La première partie sera consacrée à leur étude.

Dans un second temps, cette partie détaille quelles sont les attentes de notre modèle par les différents utilisateurs et quelles sont les bases de connaissances nécessaires pour répondre aux dites attentes.

Notre modèle repose sur trois bases de connaissance distinctes dont le regroupement d'informations dispersées au sein de ces trois bases de connaissance permet de répondre aux besoins des utilisateurs de manière optimale. Ces trois bases sont :

- Base de connaissance interne qui correspond aux informations relatives à une organisation spécifique ;
- Base de connaissance externe qui repose sur le corpus de textes et documentations disponibles pour les organisations afin de les aider et guider dans leur mise en conformité ;
- Base de connaissance réglementaire qui correspond au corpus juridique composé de différents textes et actes juridiques qui composent la législation auquel une organisation doit se conformer.

Les différentes bases de connaissance seront détaillées en détail dans la partie suivante.

1.2.1. Les utilisateurs

Comme mentionné précédemment, la complexité de la gestion du Data Regulation Risk réside dans la nécessité de traduire les contraintes et exigences réglementaires en termes techniques, organisationnels et opérationnels. Notre modèle doit faciliter la compréhension aux différents experts du domaine des exigences auxquelles une organisation doit se conformer.

Sans oublier que le DRR est spécifique au contexte d'une entreprise, dépend de ses marchés, de la présence géographique et des juridictions d'une organisation, il nécessite donc une analyse approfondie impliquant un large éventail de compétences qui sont généralement fragmentées entre les départements de l'organisation tels que le juridique, la sécurité des systèmes d'information, l'informatique opérationnel, finance, gestion des risques, ressources humaines, etc. En d'autres termes, une exigence réglementaire spécifique peut non seulement nécessiter une interprétation juridique, mais également l'implication de praticiens de la sécurité, de responsables informatiques, d'experts en gouvernance et métiers.

Enfin, dans le cadre d'organisations de taille importante, le partage de ces informations clés est primordial pour assurer une conformité globale et une continuité des activités. En effet,

au sein d'un même département, les connaissances peuvent être disparates, segmentées ou fragmentées.

Nous avons identifié les principaux types d'utilisateurs qui sont :

- La direction des systèmes d'information,
- La direction des risques et de la sécurité,
- La direction juridique.

Tous trois ont besoin d'informations différentes extraites des lois pour exercer leurs fonctions tout en assurant la continuité des activités et la conformité de leur entreprise. D'autres utilisateurs peuvent être considérés tels la direction de l'audit interne et externe ou la direction financière. Bien que ceux-ci peuvent trouver dans notre modèle une solution d'accès à des informations utiles pour mener à bien leurs missions, leur utilisation reste marginale et ponctuelle. Nous concentrerons donc nos efforts sur les trois principaux types d'utilisateurs.

Par exemple, la direction des systèmes d'information aura besoin des modalités déontiques et des exigences réglementaires pour construire et gérer l'ensemble du parc informatique tandis que la direction des risques et de la sécurité se concentrera sur les contrôles de sécurité obligatoires qui doivent être mis en œuvre. Enfin, les directions juridiques et des audits auront besoin d'une vue d'ensemble du paysage réglementaire, du périmètre du système d'information global et son état de mise en conformité pour mener à bien leur missions.

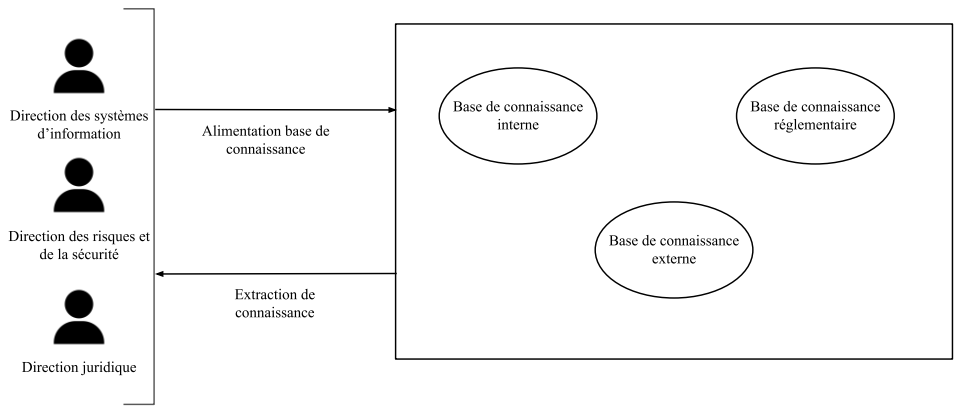


Figure 11 Utilisateurs et utilisation du modèle

Tous trois possèdent également des informations limitées quant aux exigences réglementaires, à la situation actuelle du système ou de la mise en conformité globale de l'organisation. L'objectif est donc double : permettre l'accès aux informations manquantes et pertinentes pour chaque type de fonction et permettre le partage des informations possédées

avec les autres groupes ou dans un même groupe. Ce partage s'effectue par l'alimentation de la base de connaissance ou l'extraction d'information comme illustré dans la Figure 11 Utilisateurs et utilisation du modèle.

Cette alimentation repose préalablement sur une étape d'analyse. Si l'on prend l'exemple de la base de connaissance règlementaire, la direction juridique doit analyser le contexte de l'entreprise pour déterminer le corpus juridique qui est applicable. Ensuite, celle-ci doit analyser les réglementations et en extraire les modalités juridiques afin d'alimenter la base de connaissance. Cette base vient alors répondre aux besoins des autres utilisateurs afin de les orienter sur les actions nécessaires pour assurer la conformité de l'organisation.

Le premier groupe d'utilisateurs correspond à des membres de la direction des systèmes d'information que nous détaillons ci-après.

La direction des systèmes d'information

La direction des systèmes d'information a la responsabilité de définir et de mettre en œuvre les politiques de gestion du parc informatique ainsi que de définir les orientations stratégiques et évolutions des systèmes d'information. En outre, ils ont à charge de superviser la conception et la mise en œuvre du système d'information et le maintenir en conditions opérationnelles et par conséquent d'opérer et de garantir la continuité des systèmes tout en assurant la prise en compte des besoins des utilisateurs, de l'organisation et des diverses contraintes.

On retrouve dans ce groupe d'utilisateurs des directeurs de systèmes d'information, des chefs de projet informatique, des administrateurs réseau, des architectes réseau, des gestionnaires d'identité et accès, etc. Les principales missions de la direction des systèmes d'information sont :

- Comprendre les besoins et les attentes métier ;
- Définir la stratégie du système d'information et un plan d'action associé en accord avec les besoins et attentes ;
- Implémenter un modèle d'organisation du système d'information pour être capable de mettre en place la stratégie et le plan d'action (organisation des équipes internes, externes, mise en place des partenariats nécessaires) ;
- Définir et implémenter l'architecture de haut niveau du système d'information (infrastructure et applications, externalisation, etc.) ;

- Définir et implémenter l'architecture de bas niveau du système d'information par couche (réseau, système, applicatif, donnés, etc.) ;
- Assurer l'évolution et la gestion du cycle de vie du système d'information ;
- Assurer le maintien en condition opérationnelle du système d'information (construction et suivi d'indicateurs de qualité et de performance ; gestion des sauvegardes, inventaires et incidents).

Nous allons maintenant présenté le deuxième groupe d'utilisateurs qui est composé d'individus appartenant à la direction risque et sécurité. Ceux-ci sont en charge de la sécurisation des systèmes et des données.

La direction des risques et de la sécurité

La direction risque et sécurité est entre autre en charge de définir les axes et les objectifs stratégiques en matière de cybersécurité, de définir et maintenir la politique de sécurité des systèmes d'information. En charge de la sécurisation des systèmes et des données, elle est responsable de la définition des mesures organisationnelles et techniques à mettre en œuvre pour atteindre les objectifs de sécurité. Cela comprend également les programmes de formation et de sensibilisation à la cybersécurité.

On retrouve dans ce groupe d'utilisateurs des Responsable de la Sécurité des Systèmes d'Information (RSSI), des architectes sécurité, des coordinateurs sécurité, ou encore des spécialistes sécurité d'un domaine technique. Les principales missions de la direction risque et sécurité sont :

- Comprendre le fonctionnement et la criticité des processus métiers de l'organisation ;
- Définir et implémenter la politique de sécurité de l'information et documents associés ;
- Etablir la stratégie sécurité du système d'information et l'intégrer avec la stratégie du système d'information ;
- Implémenter un modèle d'organisation de la Sécurité (organisation des équipes internes, externes, mise en place des partenariats nécessaires) ;
- Définir les mesures de sécurité ;
- Assurer l'évolution et la gestion du cycle de vie des mesures de sécurité.

La partie suivante présente le troisième groupe d'utilisateurs correspondant aux membres de la direction juridique qui assurent la connaissance, compréhension et respect des réglementations.

La direction juridique

La direction juridique a pour responsabilité de déterminer le contexte réglementaire de l'organisation et d'en mesurer l'impact métier. C'est elle qui effectue la veille et qui a la pleine connaissance des textes. Elle doit faire en sorte d'essaimer les obligations réglementaires au sein des différentes fonctions en leur indiquant leur rôle et responsabilité et en leur apportant un support dans les actions à mettre en place. La direction juridique est aussi celle en contact avec les autorités locales pour assister l'organisation dans sa mise en conformité ou en cas d'audit. Les principales missions de la direction juridique sont :

- Comprendre les besoins et les attentes métier ;
- Analyser et maîtriser les réglementations en vigueur ;
- Définir la politique de conformité selon le contexte de l'organisation;
- Assurer le suivi de la conformité de l'organisation et son évolution.

Limites d'utilisation

Notre modèle se doit d'être structuré pour gérer le contexte juridique de diverses entreprises tout en restant indépendant des méthodologies de gestion des risques. Afin de réaliser cela, il est nécessaire que le modèle soit applicable en l'état et ne nécessite pas de développements ou adaptations nécessaires à chaque organisation souhaitant l'implémenter.

Les utilisateurs présentés ci-avant sont des experts dans leur domaine respectif mais n'ont pas pour ambition d'effectuer des modifications au niveau de la structure de l'ontologie. Ceux-ci sont donc légitimes pour l'instanciation des concepts qui sont propres à leur contexte, leur organisation et leurs besoins.

La modification, création ou suppression des concepts et relations nécessitent des compétences spécifiques mais aussi un niveau de connaissance et de compréhension du modèle qui ne sont pas nécessairement adaptés aux utilisateurs présentés. Par conséquent, laisser la possibilité aux utilisateurs non-formés et non-experts d'effectuer ces changements expose notre

modèle au risque de modifications non souhaitées, non pertinentes qui peuvent altérer et compromettre la pertinence de ce dernier.

Les implications de cette décision sont doubles :

- La nécessité d'avoir un modèle unique ;
- La non-nécessité d'avoir en place une gestion de modification des concepts et propriétés (d'objet et donnée).

En d'autres termes, nous souhaitons limiter les utilisateurs aux actions effectuées sur les instances et ne pas autoriser les actions effectuées sur les autres entités de notre ontologie (concepts et propriétés). Il convient alors non seulement de nous assurer que notre modèle soit suffisamment générique pour couvrir l'ensemble des contextes existants mais aussi de convenir d'une granularité suffisante pour refléter fidèlement les problématiques de notre domaine.

Le

Tableau 5 Type d'action autorisé par notre modèle reprend les actions qui sont effectuées au niveau du modèle par les concepteurs du modèle et les actions autorisées et effectuées par les utilisateurs. L'autorisation et la répartition des actions sont représentées par un « X » dans le tableau.

Afin de définir la liste d'action possible dans notre ontologie, nous nous sommes basés sur la taxonomie des changements atomiques proposée par Stojanovic (Stojanovic, 2004).

Entité	Action	Utilisateur	Concepteur
Concept	Ajout/Suppression d'un concept à l'ontologie		x
Hiérarchie de concept	Ajout/Suppression d'un sous-concept		x
Propriété	Ajout/Suppression/Modification d'une propriété à un concept		x
Hiérarchie de propriété	Ajout/Suppression/Modification d'une sous-propriété		x
Domaine de propriété	Ajout/Suppression/Modification d'un domaine à une propriété		x
Symétrie de propriété	Ajout/Suppression/Modification d'une caractéristique de symétrie à une propriété		x
Transitivité de propriété	Ajout/Suppression/Modification d'une caractéristique de transitivité à une propriété		x
Inverse de propriété	Ajout/Suppression/Modification d'une caractéristique d'inverse (ou non) à une ou plusieurs propriétés		x
Max / Min des cardinalités	Ajout/Suppression/Modification d'une contrainte sur la cardinalité maximale ou minimale d'une propriété		x
Instance	Ajout/Suppression d'une instance	x	
Propriété d'instance	Ajout/Suppression/Modification d'une propriété à une instance	x	
Ontologie	Ajout/Suppression d'une ontologie		x

Tableau 5 Type d'action autorisé par notre modèle

La conception de notre modèle doit permettre de répondre aux besoins des différents utilisateurs. Pour cela, nous présentons ci-après, les attentes auxquelles notre modèle doit répondre.

1.2.2. Attentes du modèle

La partie précédente a présenté les différents utilisateurs de notre modèle. Il s'agit maintenant de détailler quelles sont leurs attentes. Cette partie présente également quelles sont les connaissances nécessaires pour répondre aux dites attentes et comment celles-ci sont organisées.

Questions de compétence

Par souci de clarté, nous désignerons le traitement des données comme défini dans le Règlement Général sur la Protection des Données (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.) à savoir : des opérations effectuées sur les données, à la fois par des moyens manuels ou automatisés. Le traitement comprend la collecte, l'enregistrement, l'organisation, la structuration, le stockage, l'adaptation ou la modification, la récupération, la consultation, l'utilisation, la divulgation par transmission, diffusion ou autre mise à disposition, l'alignement ou la combinaison, la restriction, l'effacement ou la destruction de données. Afin de répondre aux objectifs mentionnés précédemment et d'assurer une réponse aux besoins variés des différents utilisateurs, notre modèle repose sur des questions de compétences présentées dans la Figure 12 Question de compétence.

Ces questions sont regroupées en trois questions principales appelées « questions mères » :

- Qui/Quoi ?
- Pourquoi ?
- Comment ?

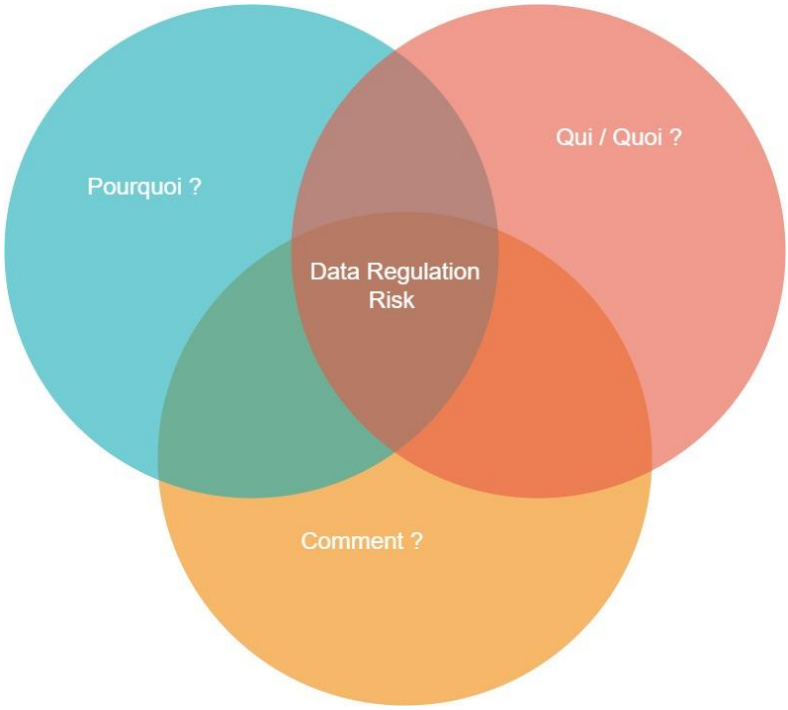


Figure 12 Question de compétence

Qui / Quoi ?

En répondant à la question « qu'est ce qui est réglementé ? », notre modèle permet de délimiter le sujet des réglementions et leur champ d'action respectif. Cela peut correspondre à une donnée, un individu, un traitement effectué ou un système qui traite une donnée réglementée. Dans le cadre de donnée personnelle, cela peut correspondre à un individu qui possède une donnée.

L'objectif de cette question est de concentrer l'attention sur ce qui est réglementé et d'écarter ce qui ne l'est pas afin d'optimiser et rationaliser les efforts.

Comment ?

En répondant à la question « comment est-ce réglementé ? », notre modèle permet d'expliciter les modalités juridiques. Cela correspond aux interdictions, obligations et libre choix qui s'appliquent à gestion de données réglementées ou de leur traitement.

L'objectif de cette question est de concentrer l'attention sur le degré de possibilité de la réalisation d'un évènement ou d'une action en fonction de certaines circonstances et ainsi d'exprimer de manière claire et précise les différentes obligations qui incombe aux organisations de respecter.

Pourquoi ?

Enfin, en répondant à la question « pourquoi est-ce réglementé ? », notre modèle permet d'expliciter et de référencer les textes juridiques qui régissent les données et leur traitement.

L'objectif de cette question est de justifier les réponses aux deux questions précédentes afin d'éviter la sur-conformité tout en assurant les utilisateurs de la véracité des informations présentes dans l'outil.

Le regroupement des réponses à ces trois questions permet à une organisation d'appréhender le Data Regulation Risk de manière optimale. La première définit le périmètre du risque, son contexte et ce qui doit être pris en compte dans la phase d'analyse de la gestion du risque. La deuxième définit ce qui est attendu, ce qui doit être fait pour traiter le risque ou ce qui justifie l'existence du risque en cas d'absence. La troisième permet de définir l'origine du risque.

La combinaison de « ces trois questions mères » nous permet également de formuler des « questions filles » davantage granulaires qui répondent à des besoins métiers et opérationnels spécifiques.

Les différents types d'utilisateurs présentés ci-avant présentent des besoins d'accès à des informations variés qui se matérialisent par des questions filles parfois différentes. Il n'est pas réaliste de prévoir, dès l'étape de conception du modèle, l'ensemble des questions auxquelles les différents utilisateurs vont chercher à obtenir les réponses. En effet, celles-ci vont dépendre du contexte de l'organisation tels que le champ législatif applicable, son système d'information mais également des connaissances de l'utilisateur ou encore de son activité professionnelle. Par exemple, il est peu probable qu'un utilisateur appartenant à la direction des systèmes d'information spécialisé dans la gestion des réseaux dans une entreprise multinationale et qu'un utilisateur en charge du maintien du système d'information d'une PME[13] aient besoin des mêmes informations pour effectuer leur missions respectives.

[13] Petites et Moyennes Entreprises (PME) sont celles qui, d'une part, occupent moins de 250 personnes, d'autre part, ont un chiffre d'affaires annuel n'excédant pas 50 millions d'euros ou un total de bilan n'excédant pas 43 millions d'euros comme définie par le décret d'application (n°2008-1354) de l'article 51 de la loi de modernisation de l'économie.

Nous avons cependant déterminer les questions que nous estimons les plus fréquentes et les plus communes. Ces questions sont supposées évoluer selon les retours et besoins des utilisateurs.

Les « questions filles » ou questions de compétences sont les suivantes :

- Quels sont les éléments du système d'information qui sont impliqués dans le traitement d'une donnée réglementée ?
- Dans le cadre de plusieurs réglementations, il y a-t-il des modalités juridiques similaires, opposées ou incompatibles ?
- Quelles sont les mesures de sécurité applicables aux différents éléments du système d'information ?
- Dans le cadre d'une donnée réglementée, est-ce qu'un traitement / composant peut être utilisé ?
- Pour une organisation, est-ce que sa stratégie du système d'information et le plan d'action associé permettent de répondre à ses exigences réglementaires ?
- Pour une organisation, est-ce que sa stratégie de sécurité du système d'information et son modèle d'organisation de la sécurité permettent de répondre à ses exigences réglementaires ?

Afin de pouvoir répondre à ces questions, notre modèle doit pouvoir disposer des informations nécessaires que nous présentons ci-après.

Les bases de connaissance

Afin de répondre à nos objectifs d'intelligibilité et de précision de représentation des différentes modalités juridiques, leurs impacts métiers et opérationnels, notre modèle repose sur trois bases de connaissances distinctes :

- Base de connaissance interne ;
- Base de connaissance externe ;
- Base de connaissance réglementaire.

Le regroupement d'informations dispersées au sein de ces trois bases de connaissance permet à une organisation d'appréhender le Data Regulation Risk de manière efficace.

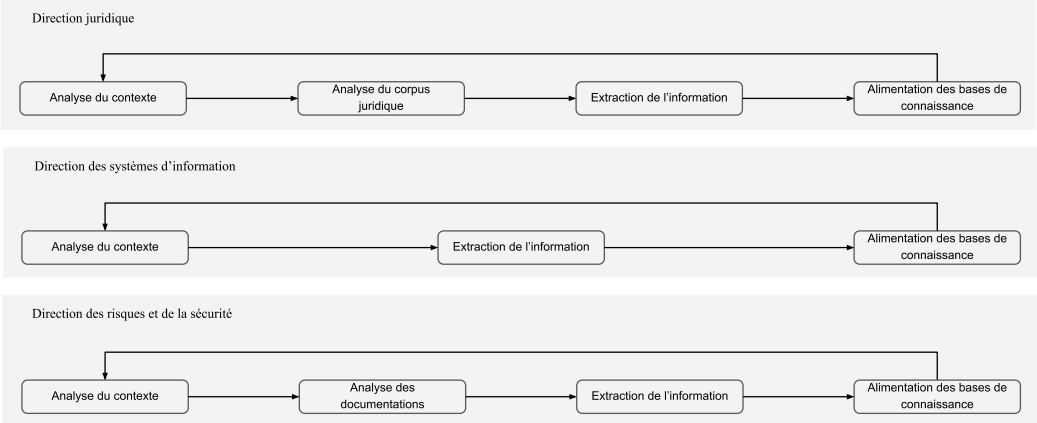


Figure 13 Alimentation des bases de connaissance

La Figure 13 Alimentation des bases de connaissance représente les étapes d'alimentation des bases de connaissance pour les différents types d'utilisateurs. Après l'étude du contexte propre à l'expertise de chacun, les utilisateurs extraient les connaissances dont ils ont la maîtrise et alimentent les bases de connaissance respectives. Par exemple, fournir les informations relatives au système d'information de l'organisation est un tache effectuée par la direction des systèmes d'information alors que la direction des risques et de la sécurité est responsable d'extraire les mesures de sécurité disponibles dans les cadriciels externes.

Il n'y a pas de limitations d'accès ou d'écriture des différentes bases de connaissance selon le type d'utilisateur. En effet, un utilisateur peut utiliser des connaissances présentes dans les différentes bases du fait de ses fonctions. Par exemple, une mesure de sécurité peut être d'origine interne, provenir d'un cadriciel externe ou d'un texte juridique.

Base de connaissance interne

La base de connaissance interne correspond aux informations relatives à une organisation spécifique. Cette base de connaissance permet à notre modèle et aux utilisateurs de notre ontologie de s'adapter au contexte propre à une organisation. Pour cela, il est nécessaire pour les utilisateurs d'être en capacité de premièrement décrire ce contexte puis, dans un second temps, d'être en mesure d'extraire des informations qui sont spécifiques à leur organisation.

Ces informations se rapportent au système d'information et aux éléments qui le composent tels que les traitements des données, la gouvernance, les processus et services relatifs aux technologies de l'information et communication de l'organisation.

Bien que le périmètre d'un système d'information puisse varier d'une organisation à une autre, il comprend généralement l'ensemble des éléments qui participent à la gestion, au traitement, au transport et à la diffusion de l'information. Il est important de noter que les frontières du système d'information ne sont pas définies par les limites de possession ou de gestion réelle des composants. Celles-ci peuvent s'étendre au-delà sous le prisme d'organisation ou de système étendu. Ce prisme comprend quant à lui l'ensemble des partenaires et clients d'une organisation pour une représentation et prise en compte holistique du système d'information.

In fine, la base de connaissance interne correspond aux :

- Processus fonctionnels ;
- Individus et leur rôle ;
- Systèmes et composants technologiques ;
- Données ;
- Traitements de données ;
- Gouvernance (telles que les mesures de sécurité et documentation interne).

Base de connaissance externe

La base de connaissance externe repose sur le corpus de textes et documentations disponibles pour les organisations afin de les aider et guider dans leur mise en conformité. Ces jeux de documents sont généralement établis par consensus, approuvés par des entités tierces et indépendantes qui prévoient des règles ou des lignes directrices pour des activités précises et spécifiques et visant à atteindre le résultat optimal dans un contexte donné.

Ces jeux de documents comprennent entre autre les cadriciels (Framework), standards, lignes directrices (Guidelines) ou code de conduite (Code of practice), contrôle (étapes mesurées à suivre pour atteindre un objectif spécifique), certification et accréditation (Eloff & von Solms, 2000).

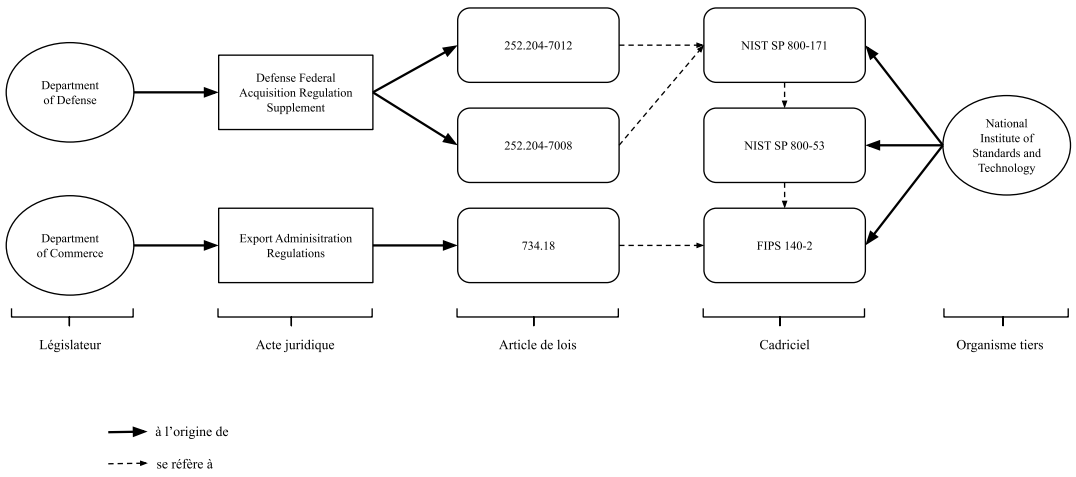


Figure 14 Interactions entre cadriciels et règlementations

Bien que leur élaboration et adoption sont souvent indépendantes de la volonté des législateurs, ces documents sont primordiaux pour toutes organisations qui souhaitent assurer sa bonne interprétation des différents actes juridiques. Ces derniers peuvent également s'y référer dans leur texte. Enfin, bien que rédigés dans une optique de facilitation et de lignes directrices à l'intention des organisations, l'utilisation de cadriciels et standards peut s'avérer complexe du fait des multiples interactions que ceux-ci peuvent avoir entre eux.

La Figure 14 Interactions entre cadriciels et règlementations illustre cette complexité au travers d'un exemple composé de deux lois américaines distinctes qui sont « Defense Federal Acquisition Regulation Supplement » et « Export Adminisitration Regulations ». Dans les articles mentionnés, ces réglementations font directement référence à différents documents publiés par l'agence américaine NIST. Certains de ces documents font également référence à d'autres ce qui augmente la complexité de leur utilisation par les organisations.

Base de connaissance réglementaire

La base de connaissance réglementaire correspond au corpus juridique composé de différents textes et actes juridiques qui composent la législation auquel une organisation doit se conformer. Bien que les corpus juridiques peuvent légèrement varier selon les pays ou zones juridiques, ceux-ci reprennent une structure et articulation d'actes composée de traités, règlements, directives, recommandations, actes délégués et actes d'exécution.

Les traités définissent les objectifs poursuivis ainsi que les règles de fonctionnement des institutions, les processus décisionnels et les relations entre différents États.

Les règlements sont des actes législatifs qui s'appliquent, dès leur entrée en vigueur, de manière automatique et uniforme dans tous les pays signataires, sans devoir être transposés dans la législation nationale. Les organisations présentes dans ces pays sont par conséquent obligées de se conformer dans tous les éléments des règlements dès leur entrée en vigueur.

Les directives diffèrent des règlements par la liberté accordée aux Etats signataires de fixer les modalités pour parvenir aux obligations de résultat présentes dans celles-ci. Les organisations présentes dans ces pays sont par conséquent obligées de se conformer à chacune des mesures présentes dans les différentes législations nationales. Cela peut conduire à une complexité accrue du fait du degré de liberté accordé aux Etats.

Les recommandations, actes délégués et actes d'exécution sont des instruments qui permettent respectivement de suggérer une ligne de conduite sans imposer d'obligation légale à leurs destinataires, de compléter ou de modifier des éléments non essentiels des actes législatifs et de fixer des conditions garantissant l'application uniforme de la législation.

	Base de connaissance interne	Base de connaissance externe	Base de connaissance réglementaire
Domaine de connaissance	Système d'information	Cadriciel	Traité
	Processus fonctionnel	Standard	Règlement
	Composant informatique	Ligne directrice	Directive
	Individus	Code de conduite	Recommandation
	Donnée	Contrôle	Acte délégué
	Documentation	Certification	Acte d'exécution
	Mesure de sécurité	Accréditation	

Tableau 6 Les 3 bases de connaissance de notre modèle

Le

Tableau 6 Les 3 bases de connaissance de notre modèle regroupe les différents domaines de connaissance des trois bases présentées ci-avant. Le regroupement d'informations dispersées au sein de ces trois bases permet à une organisation de couvrir le domaine du DRR présenté dans le premier chapitre. En effet, nous avons précisé précédemment que pour modéliser fidèlement le domaine du DRR, il était nécessaire d'avoir un modèle présentant les concepts présents dans sa définition. Ceux-ci doivent permettre la représentation :

- D'une entreprise et de son système d'information ;
- De la sécurité (de l'information) ;
- Des réglementations ;
- De contexte de l'entreprise (applicabilité des réglementations).

La base de connaissance interne a donc pour ambition de regrouper les instances de concepts qui représentent une entreprise et son système d'information ainsi que son contexte alors que la base de connaissance externe permet d'organiser et regrouper le corpus de textes et documentations disponibles pour les organisations afin de mettre en place les mesures de sécurité nécessaires pour assurer leur conformité. Enfin, la base de connaissance réglementaire a pour ambition de regrouper les instances du corpus juridique.

Cette section a présenté les objectifs de notre modèle : retranscrire précisément les règlementations par l'illustration des modalités juridiques, de leurs évolutions et de leurs impacts sur les organisations et le contexte d'utilisation de notre modèle par la description des utilisateurs et leurs attentes envisagées lors de sa conception.

Dans la prochaine section, nous nous intéressons au principes fondamentaux de notre modèle, la réutilisation de travaux existants et l'interaction des utilisateurs avec notre proposition de solution.

Section II. Conception du modèle

La création d'une ontologie nécessite de déterminer quelles seront les entités étudiées, comment elles interagissent entre elles et quelles sont leurs spécificités. Cette détermination se doit d'être précise afin de permettre une bonne compréhension de l'ontologie et de justifier son utilisation par la suite. Elles peuvent alors répondre aux besoins d'amélioration des communications entre des acteurs de milieux différents, utilisant des outils différents, ayant des connaissances et des vocabulaires différents.

Nous avons déjà identifié des concepts et relations présents dans des travaux existants. Nous avons également posé les fondements de notre modèle, nos objectifs ainsi que les utilisateurs envisagés.

Cette section présente l'ensemble des concepts utilisés pour représenter notre domaine. Ceux-ci sont au nombre de 15 et sont répartis en quatre sous domaines. Afin de représenter notre domaine de manière explicite et juste, chaque concept possède des propriétés qui lui sont propres. Celles-ci sont aussi décrites dans cette section.

Dans un second temps, cette section présente également les relations ou propriété d'objet qui permettent de relier ces concepts. Celles-ci sont d'autant plus importantes qu'elles permettent de mettre en place notre logique sémantique et par conséquent la logique de notre modèle.

Enfin, cette section conclut par une synthèse de notre modèle conceptuel (Delorme et al., 2022) afin de pouvoir tester et valider celui-ci dans la section suivante.

2.1. Création de l'ontologie

La première étape de la conception de notre modèle consiste à identifier l'ensemble des concepts utilisés pour représenter assidûment notre domaine. Chaque concept possède une à plusieurs propriétés de données qui nous permettent de les différencier et de les spécifier. Nous allons ensuite présenter les relations ou propriété d'objet qui seront utilisées pour représenter les interactions entre nos concepts.

2.1.1. Les concepts de notre modèle

Cette section présente l'ensemble des concepts utilisés pour représenter notre domaine. Ceux-ci sont au nombre de 15 et sont répartis en quatre sous domaines :

- Le sous domaine de l'entreprise ;
- Le sous domaine de la sécurité ;
- Le sous domaine réglementaire ;
- Le sous domaine de la localisation ;

La Figure 15 Concepts clés de l'ontologie représente l'ensemble des concepts de notre modèle répartis par sous domaine. Nous avons repris des travaux de Ekelhart et al. le principe des sous domaines de l'entreprise, de la sécurité et de la localisation (Ekelhart et al., 2006). Nous avons complété ces trois sous domaines par celui du réglementaire afin de pouvoir représenter pleinement notre sujet.

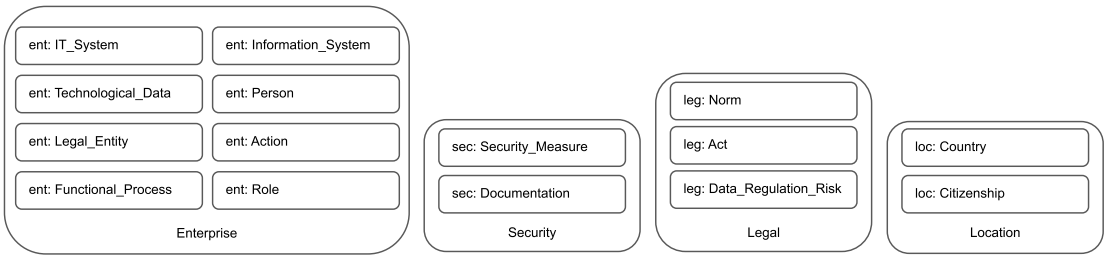


Figure 15 Concepts clés de l'ontologie

Sous domaine de l'entreprise

Notre premier sous domaine est celui de l'entreprise qui a pour ambition de représenter :

- L'entreprise de référence à laquelle les utilisateurs appartiennent ;
- Les entreprises et organisations impliquées dans la mise en conformité de l'entreprise de référence ;
- Les entreprises et organisations impliquées dans les activités et la gestion du système d'information de l'entreprise de référence.

Les organisations légales : (ent: Legal Entity ⊑ T)

Le concept *Legal_Entity* représente une personne physique ou morale, un organisme d'autorité publique exerçant une activité quelle que soit sa forme juridique, y compris les sociétés de personnes ou les associations.

SynSet ID	100001740
SynSet définition	That which is perceived or known or inferred to have its own distinct existence (living or nonliving).

Ce concept est complété par des sous concepts afin de représenter au mieux la variété d'entités impliquées dans notre domaine comme illustré dans la Figure 16 La classe Legal_Entity et ses sous classes. Ces sous classes sont :

- Les entreprises (ent : Business_Organization ⊑ Legal_Entity) ;
- Les organismes publics indépendants (ent : Independant_Organization ⊑ Legal_Entity) ;
- Les agences réglementaires (ent : Regulatory_Agency ⊑ Legal_Entity).

Ses trois sous classes sont disjointes.

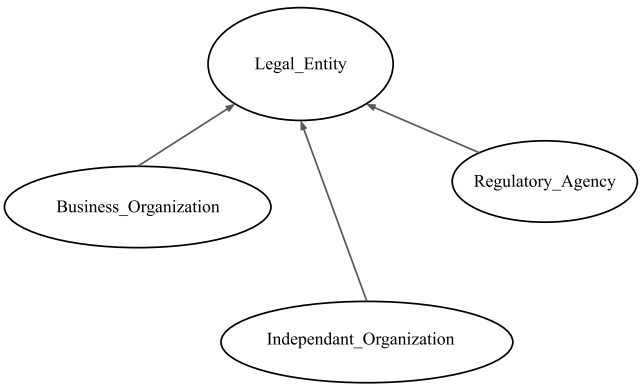


Figure 16 La classe Legal_Entity et ses sous classes

Le concept *Legal_Entity* et ses sous concepts dispose d'un nom, d'un IRI, d'une version et d'une description.

La sous classe *Business_Organization* possède la propriété de données « RefOrganization ». Cette propriété est de type booléen et permet d'identifier si l'entreprise

correspond à l'entreprise de référence. Bien qu'elle n'ait pas d'impact sur la pertinence de notre modèle, cette propriété à des fins d'administration et gestion des connaissances simplifiées.

Les systèmes d'information : (ent: Information_System ⊑ T)

Le concept *Information_System* représente un ensemble organisé de ressources (matériels, logiciels, personnel, données et procédures) qui permet de regrouper, de classifier, de traiter et de diffuser de l'information sur un environnement donné.

SynSet ID	103164344
SynSet définition	System consisting of the network of all communication channels used within an organization.

Dans notre modèle, les systèmes d'information sont composés d'au moins un processus fonctionnel qui implique des composants informatiques et des individus. La création de ce concept nous permet de regrouper différents processus fonctionnels en un système d'information. Cette possibilité nous donne la granularité suffisante pour représenter des processus fonctionnels impliquants des composants informatiques ou des individus externes à l'entreprise de référence. Le concept *Information_System* dispose d'un nom, d'un IRI, d'une version et d'une description.

Les processus fonctionnels : (ent : Functional_Process ⊑ T)

Le concept *Functional_Process* représente un ensemble d'activités corrélées ou en interaction qui utilise des éléments d'entrée pour produire un résultat escompté.

SynSet ID	101023820
SynSet définition	A particular course of action intended to achieve a result

Dans notre modèle, les processus fonctionnels impliquent des composants informatiques et des individus. La création de ce concept nous permet d'illustrer des activités qui regroupent des individus internes ou externes à l'entreprise de référence ou des composants techniques qui n'appartiennent pas ou ne sont pas gérés par celle-ci.

Un exemple serait l'implication de différentes applications dans un même processus fonctionnel. Ces applications impliquent des utilisateurs internes et des administrateurs

externes sur des applications internes hébergées sur un Data Center externe de type PaaS (Plateform as a Service) et des applications hébergées sur le système d'information de l'enterprise. Par la création du concept *Functional_Process,* il nous est alors possible de retranscrire les scénarios complexes d'architecture informatique impliquant des composants ou utilisateurs appartenant à une ou plusieurs *Legal_Entity.*

Le concept *Functional_Process* dispose d'un nom, d'un IRI, d'une version et d'une description.

Les composants informatiques : (ent: IT_System ⊑ T)

Le concept *IT_System* représente une combinaison d'éléments interactifs (ressources) organisés pour atteindre un ou plusieurs objectifs escomptés.

SynSet ID	104377057
SynSet définition	Instrumentality that combines interrelated interacting artifacts designed to work as a coherent entity

Les composants informatiques peuvent être classés selon leur utilisation. Ces classes correspondent aux différents éléments d'un système d'information tels que représentés dans la Figure 17 La classe IT_System et ses sous classes. Ces sous classes sont disjointes.

En se basant sur le modèle OSI et les modèles utilisés dans les environnements d'informatique en nuage (cloud computing) , nous avons établi les sous concepts suivants :

- Data_Center (ent: Data_Center ⊑ IT_System);
- Network (ent: Network ⊑ IT_System);
- Physical_Server (ent: Physical_Server ⊑ IT_System);
- Virtualization (ent: Virtualization ⊑ IT_System);
- Operating_System (ent: Operating_System ⊑ IT_System);
- DataBase (ent: DataBase ⊑ IT_System);
- Application (ent: Application ⊑ IT_System);
- Device (ent: Device ⊑ IT_System).

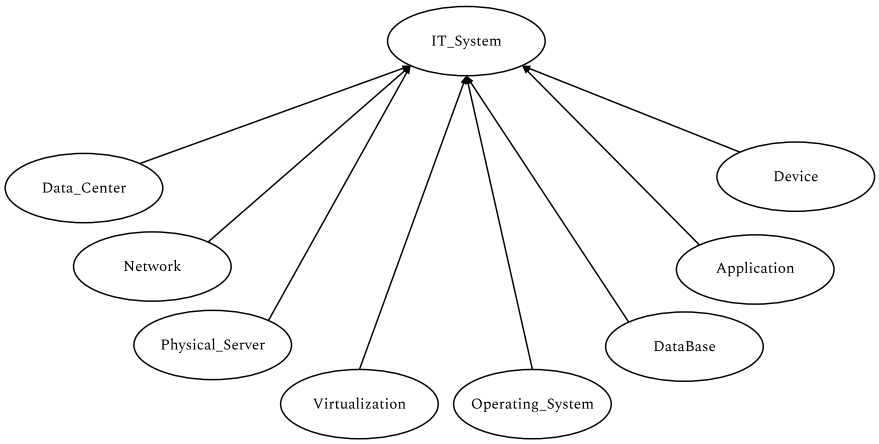


Figure 17 La classe IT_System et ses sous classes

La création de ce concept nous permet d'illustrer des interactions de différents composants impliqués dans le traitement de données réglementées. Cela est d'autant plus important que les réglementations peuvent atteindre une granularité telle que certaines modalités juridiques se concentrent sur des typologies de composants précises. Par exemple, les exigences de chiffrement diffèrent selon la couche du modèle OSI considérée.

Le concept *IT_System* dispose d'un nom, d'un IRI, d'une version et d'une description. Celui-ci nécessite également la propriété de données « IsOwnedInternally». Cette propriété est de type booléen et permet d'identifier les composants qui appartiennent à l'entreprise de référence. De façon similaire, ce concept nécessite la propriété de données « IsManagedInternally». Cette propriété est de type booléen et permet d'identifier les composants qui sont gérés et administrés par l'entreprise de référence.

Les actions (ent: Action ⊑ T)

Le concept *Action* représente un traitement effectué sur une donnée.

SynSet ID	100030358
SynSet définition	something that people do or cause to happen.

Si l'on se réfère à l'approche classique des systèmes d'information, un traitement de donnée peut être divisé en quatre principaux types qui sont :

- La recherche et l'extraction de données ;
- Le traitement, la mise à jour et l'exploitation de données pour créer de l'information ;

- La distribution et communication de données et/ou information ;
- Le stockage.

Le concept *Action* dispose d'un nom, d'un IRI, d'une version et d'une description. A des fins de facilitation d'utilisation du modèle et avec une ambition d'obtenir un niveau de granularité suffisant pour répondre aux attentes des utilisateurs, nous avons créé la propriété de données « ActionType » de type « Enumerated datatype ». Elle a pour « range » : « RetrieveAction », « ExploitAction », « TransferAction » et « StorageAction ».

Les individus : (ent: Person ⊑ T)

Le concept *Person* représente une personne physique identifiable. Celle-ci peut faire partie du capital humain d'une entité ou être indépendante de cette dernière.

SynSet ID	100007846
SynSet définition	A human being

Le concept *Person* dispose d'un nom, d'un IRI, d'une version et d'une description. Celui-ci nécessite également la propriété de données « EmployeeInternal». Cette propriété est de type booléen et permet d'identifier les individus qui sont des employés de l'entreprise de référence.

Les rôles : (ent: Role ⊑ T)

Le concept *rôle* représente les actions et activités qu'un individu est missionné et autorisé d'effectuer.

SynSet ID	100720565
SynSet définition	The actions and activities assigned to or required or expected of a person or group.

Le concept *Role* dispose d'un nom, d'un IRI, d'une version et d'une description.

Les données : (ent : Technological_Data ⊑ ⊤)

Le concept *Technological_Data* représente une donnée sous format numérique. Une donnée est un item d'information factuelle.

SynSet ID	105816622
SynSet définition	An item of factual information derived from measurement or research.

Ce concept est complété par deux sous concepts afin de différencier les données personnelles des données dites « business » qui sont inhérentes au fonctionnement d'une organisation. Les sous concepts sont représentés dans la Figure 18 La classe Technological_Data et ses sous classes, celles-ci sont :

- Business_Data (ent: Business_Data ⊑ Technological_Data);
- Personal_Data (ent: Personal_Data ⊑ Technological_Data).

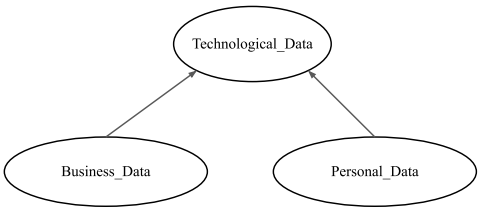


Figure 18 La classe Technological_Data et ses sous classes

Le concept *Technological_Data* et ses sous concepts dispose d'un nom, d'un IRI, d'une version et d'une description.

La Figure 19 Instanciation des concepts du sous domaine de l'entreprise illustre une instanciation de ces concepts par un utilisateur autorisé d'une organisation qui a le rôle de sous-traitant (« processor ») accédant à des données à caractère personnel réglementées par le RGPD. Les relations utilisées dans cette figure sont expliquées dans la partie 2.1.2.

Cela se traduit par un accès à des données réglementées par le RGPD : *GDPRData is_a (ent: Personal_Data ⊑ Technological_Data ⊑ ⊤)* par un utilisateur *UserGDPRApplication is_a (ent : Person ⊑ ⊤).*

Cet utilisateur ayant un rôle autorisé *GDPRUserRole is_a (ent : Role ⊑ ⊤).*

Cette action d'accès à une donnée *AccessGDPRData is_a (ent: Action ⊑ ⊤)* est effectuée par une application *GDPRApplication is_a (ent: Application ⊑ IT_System).*

Cette application est impliquée dans un processus fonctionnel *GDPRProcess is_a (ent: Functional_Process ⊑ ⊤)* qui compose le système d'information *EnterpriseIS is_a (ent: Information_System ⊑ ⊤)* de l'entreprise *EnterpriseProcessor is_a (ent:Business_Organization ⊑ Legal_Entity)* auquel appartient l'utilisateur.

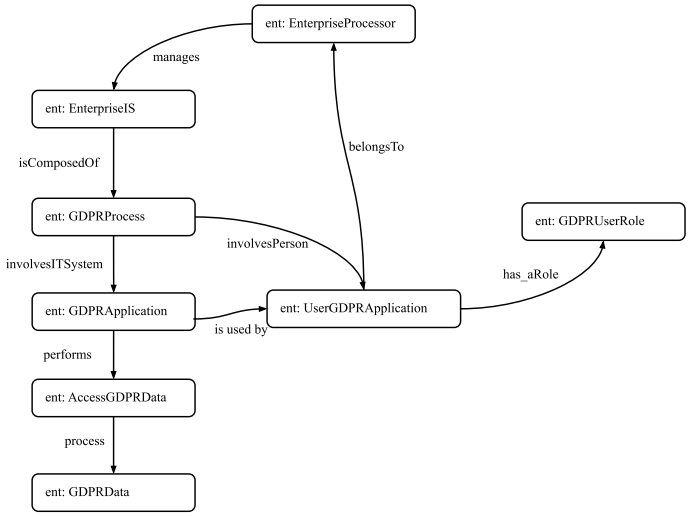


Figure 19 Instanciation des concepts du sous domaine de l'entreprise

Sous domaine de la sécurité

Notre second domaine est celui de la sécurité qui a pour ambition de représenter :

- L'ensemble des mesures de sécurité appliquées au domaine d'étude,
- L'ensemble des documents qui décrivent ces mesures de sécurité.

Le sous domaine de la sécurité est composé des classes Security_Measure (sec : Security_Measure ⊑ ⊤) et Documentation (sec : Documentation ⊑ ⊤).

La Figure 20 Instanciation des concepts du sous domaine de la sécurité illustre une instanciation de ces deux concepts par une politique de sécurité interne *InternalSecurityPolicy is_a (sec : Policy ⊑ Documentation ⊑ ⊤)* qui définit une mesure de sécurité de gestion des accès *AccessManagementSecurityMeasure is_a (sec : Identify_Security_Measure ⊑ Security_Measure ⊑ ⊤).*

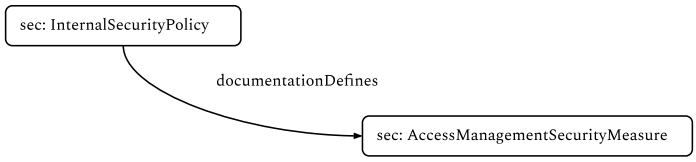


Figure 20 Instanciation des concepts du sous domaine de la sécurité

Les mesures de sécurité : (sec: Security_Measure ⊑ T)

Le concept *Security_Measure* représente une contre-mesure prescrite pour un système d'information ou une organisation. Celles-ci ont pour but de protéger la confidentialité, l'intégrité et la disponibilité de ses informations et de répondre à un ensemble d'exigences de sécurité définies.

SynSet ID	100823316
SynSet définition	Measures taken as a precaution against theft or espionage or sabotage etc.

Les mesures de sécurité se différencient selon leur nature (mesures procédurales et managériales, outils technologiques matériels ou logiciels, compétences humaines) et leur but (proactive ou réactive) (Ghernaouti, 2016). La classe *Security_Measure* nous permet de représenter deux principales sources de mesures de sécurité : les mesures mises en place au sein d'une organisation et les mesures décrites dans les documentations ou mentionnées dans le corpus juridique.

Nous avons basé la création de la classe *Security_Measure* et des sous classes qui la composent sur les informations fournies par le NIST Framework for Improving Critical Infrastructure Cybersecurity (NIST CSF)[14]. Publié en 2014 et révisé en 2017 et 2018, ce cadriciel est le résultat d'un effort de collaboration continu impliquant des organisations industrielles, le milieu universitaire et gouvernemental. Celui-ci met l'accent sur l'utilisation de critères commerciaux pour guider les activités de cybersécurité et prendre en compte les risques associés dans le cadre des processus de gestion des risques de l'organisation. Bien que ce document ait été élaboré pour améliorer la gestion des risques de cybersécurité dans les

[14] Le NIST Cybersecurity Framework Version 1.1 est accessible via ce lien : https://www.nist.gov/cyberframework/framework.

infrastructures, il peut être utilisé par des organisations de n'importe quel secteur ou communauté. Il permet aux organisations indépendamment de leur taille, des risques de cybersécurité ou de leur maturité d'appliquer les principes fondamentaux et les meilleures pratiques de gestion des risques visant à améliorer leur sécurité et leur résilience.

En se basant sur ce document, nous avons établi les sous concepts suivants :

- Mesures de sécurité d'identification (sec : Identify_Security_Measure ⊑ Security_Measure) ;
- Mesures de sécurité de protection (sec : Protect_Security_Measure ⊑ Security_Measure) ;
- Mesures de sécurité de détection (sec : Detect_Security_Measure ⊑ Security_Measure) ;
- Mesures de sécurité de réponse (sec : Respond_Security_Measure ⊑ Security_Measure) ;
- Mesures de sécurité de reprise (sec : Recover_Security_Measure ⊑ Security_Measure).

L'ensemble des sous classes sont présentées dans la Figure 21 La classe Security_Measure et ses sous classes qui sont disjointes.

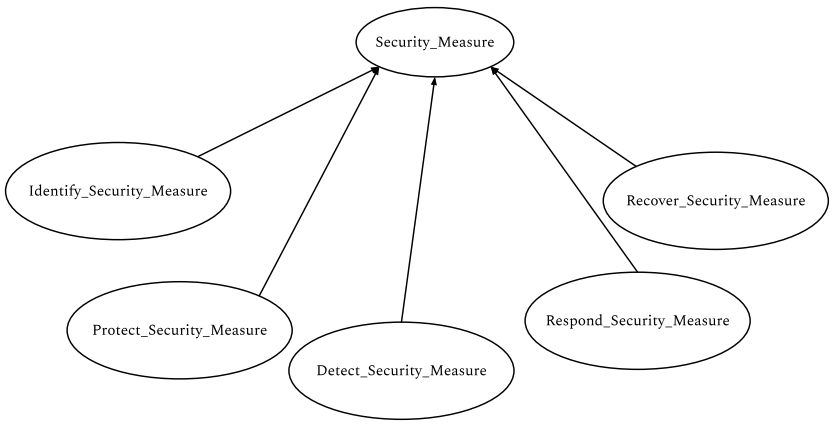


Figure 21 La classe Security_Measure et ses sous classes

Le concept *Security_Measure* et ses sous concepts dispose d'un nom, d'un IRI, d'une version et d'une description. Nous avons également doté ce concept de deux propriétés de données : «HasSecurityMeasureNature » et « HasSecurityMeasurePurpose ». Les deux propriétés sont de type « Enumerated datatype »[15]. La première propriété permet de

[15] Les Enumerated Datatype sont présentés par le World Wide Web Consortium sur leur site accessible via le lien suivant : https://www.w3.org/TR/owl-ref/#rdf-datatype

différencier les mesures de sécurité par leur nature : mesures procédurales et managériales, outils technologiques matériels ou logiciels, compétences humaines. Elle a pour « range » : « GovernanceType », « TechnologicalType » et « HumanType ». La seconde propriété permet de spécifier si les mesures ont pour but premier la protection proactive ou réactive. Elle a pour « range » : « Proactive » et « Reactive ».

La documentation : (sec: Documentation ⊑ T)

Le concept *Documentation* représente l'ensemble de textes et documents qui composent la base de connaissance externe présentée dans la première section de ce chapitre. Celui-ci représente également les textes et documents internes d'une organisation.

SynSet ID	106588326
SynSet définition	Program listings or technical manuals describing the operation and use of programs.

La classe *Documentation* nous permet de représenter les deux principales sources de documents : le corpus de textes et documentations d'origine externe et disponible pour les organisations afin de les aider et guider dans leur mise en conformité ainsi que les informations formalisées qui se rapportent au système d'information et les éléments qui le composent telles que les politiques, les processus et descriptions de services propres à une organisation. Afin de proposer suffisamment de granularité et une organisation efficace des connaissances, nous avons établi les sous concepts suivants :

- Politique (sec : Policy ⊑ Documentation) ;
- Procédure (sec : Process ⊑ Documentation) ;
- Standard (sec : Standard ⊑ Documentation) ;
- Ligne directrice (sec : Guidelines ⊑ Documentation) ;
- Cadriciel (sec : Framework ⊑ Documentation) ;
- Code de conduite (sec : CodeofPractice ⊑ Documentation) ;
- Contrôle (sec : Control ⊑ Documentation) ;
- Certification (sec : Certification ⊑ Documentation) ;
- Accréditation (sec : Accreditation ⊑ Documentation).

Ce concept et ses sous concepts, basés sur les travaux de Eloff et Von Solms (Eloff & von Solms, 2000), sont présentés dans la Figure 22 La classe Documentation et ses sous classes doivent répondre aux besoins des utilisateurs en permettant à l'organisation l'accès aux informations décrites précédemment qui correspondent en tout ou partie aux bases de connaissance internes et externes. Ces sous classes sont disjointes.

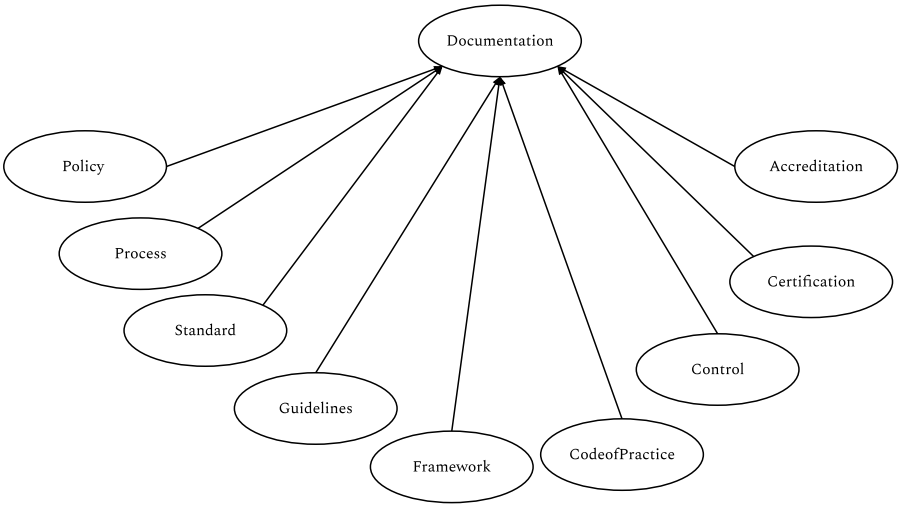


Figure 22 La classe Documentation et ses sous classes

Le concept *Documentation* et ses sous concepts dispose d'un nom, d'un IRI, d'une version et d'une description. Celui-ci nécessite également la propriété de données «IsDocumentationInternal». Cette propriété est de type booléen et permet d'identifier les documents qui sont originaires de l'entreprise de référence.

Sous domaine réglementaire

Notre troisième sous domaine est celui du réglementaire qui a pour ambition de représenter :

- Le corpus juridique ;
- Les modalités juridiques ;
- Le risque réglementaire inhérent aux règlementations.

Le sous domaine règlementaire est composé des classes Norm (leg: Norm $\sqsubseteq \top$), Act (leg: Act $\sqsubseteq \top$) et Data_Regulation_Risk (leg: Data_Regulation_Risk $\sqsubseteq \top$). La Figure 23 illustre une instanciation de ces trois classes par le règlement RGPD *GDPRNorm is_a (leg :Regulation* $\sqsubseteq$ *Norm* $\sqsubseteq \top$*)*, l'article 32 qu'il définit *GDPRArt32 is_a (leg: Act* $\sqsubseteq \top$*)* et le risque qu'il crée auprès des organisations *GDPRRisk is_a (leg: Data_Regulation_Risk* $\sqsubseteq \top$*)*.

Ce sous domaine doit répondre aux besoins des utilisateurs en permettant à l'organisation l'accès aux informations décrites précédemment qui correspondent à la base de connaissance réglementaire. Nous avons basé la création de notre domaine, de ces classes et sous classes qui le composent, sur les informations fournies par la commission européenne quant à la structure, typologie et hiérarchie des actes législatifs en vigueur dans l'Union Européenne [16].

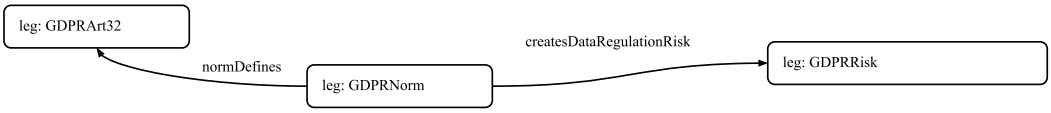


Figure 23 Instanciation des concepts du sous domaine juridique

Le corpus juridique: (leg: Norm ⊑ T)

Le concept *Norm* représente l'ensemble des lois, règlementations et textes juridiques qui régissent le traitement de données et/ou la gouvernance et les processus des technologies de l'information et communication et/ou les services et technologies de l'information.

SynSet ID	106532330
SynSet définition	Legal document setting forth rules governing a particular kind of activity.

Ce concept est complété par des sous concepts afin de représenter au mieux le corpus juridique auquel une organisation doit se conformer. Ces sous classes, qui reposent sur la structure, typologie et hiérarchie des actes législatifs en vigueur dans l'Union Européenne, représentent les actes juridiques et sont composés de :

- Traités (leg: Treaty ⊑ Norm) ;
- Règlements (leg: Regulation ⊑ Norm) ;
- Directives (leg: Directives ⊑ Norm) ;
- Recommandations (leg: Recommendation ⊑ Norm) ;
- Actes délégués (leg: Delegated_Act ⊑ Norm) ;

[16] Les types d'actes législatifs de l'Union Européenne sont accessibles via ce lien : https://ec.europa.eu/info/law/law-making-process/types-eu-law_fr

- Actes d'exécution (leg: Implementing_Act ⊑ Norm).

Ces sous classes sont disjointes et représentées dans la Figure 24 La classe Norm et ses sous classes.

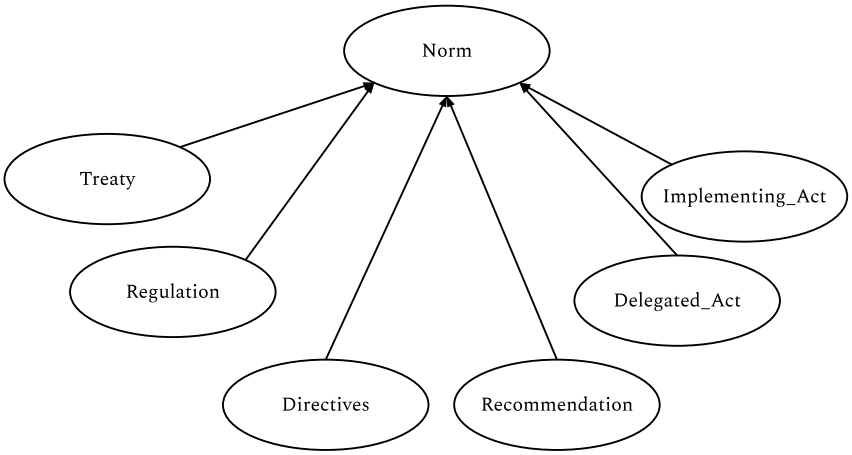


Figure 24 La classe Norm et ses sous classes

Le concept *Norm* et ses sous concepts dispose d'un nom, d'un IRI, d'une version et d'une description. Celui-ci nécessite également la propriété de données «HasNormSource». Cette propriété est de type string. Nous nous sommes basés sur la « Frame-Based Ontology » de Van Kralingen pour l'élaboration de ce concept (van Kralingen, 1997).

Les modalités juridiques : (leg: Act ⊑ T)

Le concept *Act* représente les modalités juridiques qui sont représentées par une interdiction, une obligation ou de libre choix concernant une action régie par une règlementation.

SynSet ID	100805034
SynSet définition	The act of controlling or directing according to rule

Ce concept permet de différencier les actions et évènements réglementés en définissant leur cadre. Les modalités juridiques sont des outils à disposition des législateurs dont la vocation est d'indiquer aux sujets le degré de possibilité de la réalisation d'un évènement ou d'une action en fonction de certaines circonstances qui apparaissent alors comme des

indicateurs de marge de manœuvre. En se plaçant en tant qu'organisation ou entreprise assujettie à une Data Regulation, les modalités déontiques peuvent alors être représentées par une interdiction, une obligation ou de libre choix.

Le concept *Act* dispose d'un nom, d'un IRI, d'une version et d'une description. Nous nous sommes également basés sur la «Frame-Based Ontology» pour l'élaboration de ce concept. Nous avons volontairement écartés les attributs liés : à la source, l'agent responsable, la modalité de moyen, la modalité de manière, les aspects temporels, les aspects spatiaux, les aspects circonstanciels, la raison ou cause, le but, l'intention et enfin l'état final. Ces attributs sont remplacés par les concepts de notre modèle et leurs interactions (relations). Par exemple les modalités de moyen et de manière sont remplacés par les concepts du sous domaine de la sécurité.

Nous avons néanmoins doté ce concept d'une autre propriété de donné: «hasActType » de type « Enumerated datatype ». Celle-ci permet de différencier les modalités juridiques par leur nature : interdiction, obligation ou libre choix. Elle a pour « range » : « Prohibition », « Obligation» et « Permission ».

Le Data Regulation Risk : (leg: Data_Regulation_Risk ⊑ T)

Le concept Data Regulation Risk représente un risque provenant de la possibilité d'une sanction donnée par une agence réglementaire à la suite d'évidence de non-conformité avec une loi et/ou texte juridique régissant le traitement de données et/ou la gouvernance et les processus des technologies de l'information et communication et/ou les services et technologies de l'information.

SynSet ID	114541852
SynSet définition	A source of danger; a possibility of incurring loss or misfortune

Ce concept au cœur de notre modèle dispose d'un nom, d'un IRI, d'une version et d'une description.

Sous domaine de la localisation

Notre quatrième et dernier sous domaine est celui de la localisation qui a pour ambition de représenter :

- Les pays ;
- La citoyenneté des individus.

Le sous domaine de la localisation est composé des classes Country (loc: Country ⊑ ⊤) et Citizenship (loc: Citizenship ⊑ ⊤). La Figure 25 Instanciation des concepts du sous domaine de la localisation illustre une instanciation de ces deux classes par la citoyenneté américain *USCitizenship is_a (loc: Citizenship ⊑ ⊤)* et les Etats-Unis *US is_a (loc: Country ⊑ ⊤).*

Figure 25 Instanciation des concepts du sous domaine de la localisation

Le pays : (loc: Country ⊑ ⊤)

Le concept *Country* représente le territoire d'une nation délimité par des frontières et constituant une entité géographique.

SynSet ID	108544813
SynSet définition	The territory occupied by a nation

Le concept *Country* dispose d'un nom, d'un IRI, d'une version et d'une description.

La nationalité : (loc: Citizenship ⊑ ⊤)

Le concept *Citizenship* permet de représenter la ou les nationalités d'un individu.

SynSet ID	113953467
SynSet définition	The status of a citizen with rights and duties.

Le concept *Citizenship* dispose d'un nom, d'un IRI, d'une version et d'une description.

2.1.2. Les relations de notre modèle

Cette partie présente également les relations ou propriétés d'objet qui permettent de relier ces concepts. Celles-ci sont d'autant plus importantes qu'elles permettent de mettre en place notre logique sémantique et par conséquent la logique de notre modèle.

Au total, notre modèle dispose de 15 propriétés d'objet constituées de deux types de propriété d'objet : les relations de caractéristique et les relations d'action. Dans les premières, le concept patient n'effectue pas d'action directement sur le concept agent mais permet de préciser les attributs de ce dernier. Les relations de caractéristique permettent de représenter les liens entre les différents concepts du modèle. Les relations d'action sont utilisées lorsqu'un concept agent effectue une action directe sur un concept patient.

Relations de caractéristique

Relation de gouvernance : Govern

Le lien de gouvernance entre deux concepts signifie qu'un concept agent met en conformité un concept patient avec des règles, principes ou usages. Le patient voit ses traitements de données et/ou sa gouvernance et ses processus des technologies de l'information et communication et/ou ses services et technologies de l'information régis et soumis à des règles, principes ou usages définis par un concept agent.3

SynSet ID	202511551
SynSet définition	Bring into conformity with rules or principles or usage; impose regulations

La propriété d'objet *Govern* modélise le fait qu'une instance de la classe *Act* indique aux organisations le degré de possibilité de la réalisation d'un évènement ou d'une action en fonction de certaines circonstances. Cette propriété d'objet dispose de plusieurs sous propriétés correspondant aux classes gouvernées par les règlementations comme illustré dans la Figure 26 Utilisation de la relation "gouvernance" et son inverse.

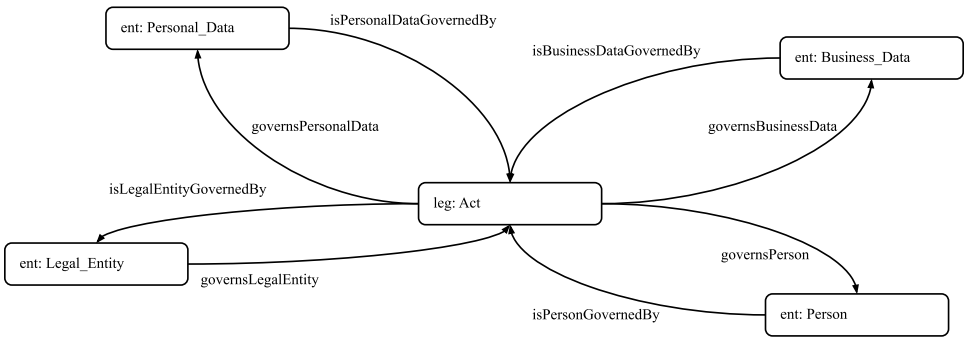


Figure 26 Utilisation de la relation "gouvernance" et son inverse

Les instances concernées sont les instances des classes *Person*, *Legal_Entity*, *Business_Data* et *Personal_Data* et sont présentées dans le Tableau 7 Tableau de sous propriétés de la relation Govern.

Propriété d'objet	Domain	Range	Inverse
governsPerson	Act	Person	isPersonGovernedBy
governsLegalEntity	Act	Legal_Entity	isLegalEntityGovernedBy
governsBusinessData	Act	Business_Data	isBusinessDataGovernedBy
governsPersonalData	Act	Personal_Data	isPersonalDataGovernedBy

Tableau 7 Tableau de sous propriétés de la relation Govern

Relation de détention : Have

Le lien de possession entre deux concepts signifie qu'un concept agent possède un concept patient.

SynSet ID	202203362
SynSet définition	Have or possess, either in a concrete or an abstract sense

La propriété d'objet *Have* modélise le fait qu'une instance de la classe Person détient une ou plusieurs citoyennetés et un ou plusieurs rôles tel que présenté dans la Figure 27 Utilisation de la relation "détention" et son inverse.

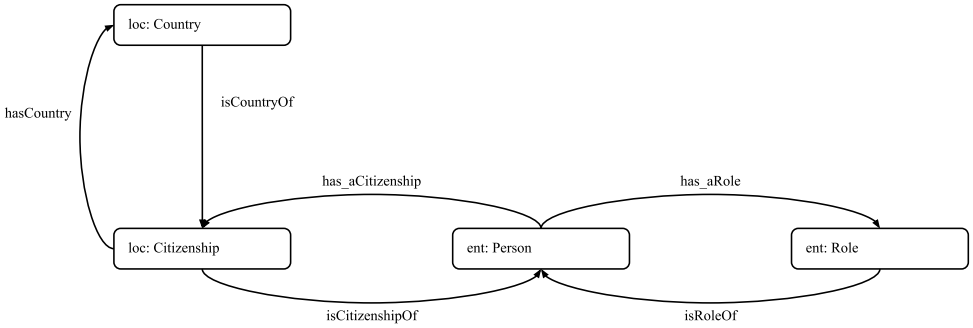


Figure 27 Utilisation de la relation "détention" et son inverse

Cette propriété d'objet dispose de plusieurs sous propriétés correspondant aux classes détenues par les individus comme illustré dans le Tableau 8 Tableau de sous propriétés de la relation Have.

Propriété d'objet	Domain	Range	Inverse
has_aCitizenship	Person	Citizenship	isCitizenshipOf
has_aRole	Person	Role	isRoleOf
hasCountry	Citizenship	Country	isCountryOf

Tableau 8 Tableau de sous propriétés de la relation Have

Relation de localisation : Located

La propriété d'objet *Located* signifie la présence d'un concept agent dans un endroit particulier. L'endroit est déterminé par le concept patient *Country*.

SynSet ID	302126430
SynSet définition	Situated in a particular spot or position

La présence peut être physique (physiquement présent dans un territoire) ou abstraite (activités réalisées dans un territoire) comme illustré dans la Figure 28 Utilisation de la relation "localisation".

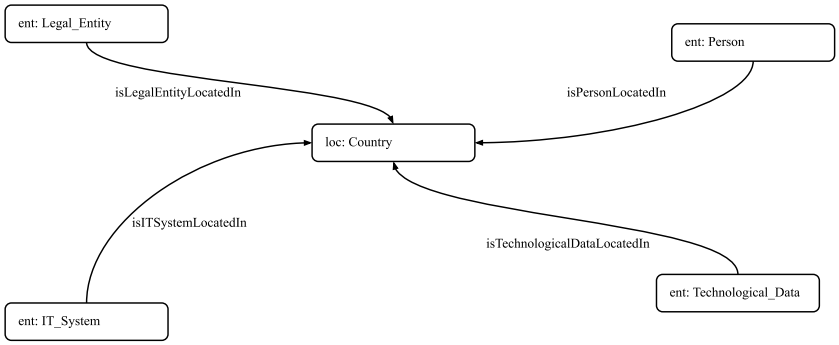


Figure 28 Utilisation de la relation "localisation"

Cette propriété d'objet dispose de plusieurs sous propriétés correspondant aux classes localisées dans les différents pays comme illustré dans le Tableau 9 Tableau de sous propriétés de la relation Located.

Propriété d'objet	Domain	Range	Inverse
isPersonLocatedIn	Person	Country	N/A
isITSystemLocatedIn	IT_System	Country	N/A
isLegalEntityLocatedIn	Legal_Entity	Country	N/A
isTechnologicalDataLocatedIn	Technological_Data	Country	N/A

Tableau 9 Tableau de sous propriétés de la relation Located

Relation d'appartenance : Belong

La propriété d'objet *Belong* entre deux concepts signifie qu'un concept patient est un membre, adhérent, employé, etc. d'un concept agent.

SynSet ID	202756359
SynSet définition	Be a member, adherent, inhabitant, etc.

Dans le cadre de notre modèle, le lien *Belong* correspond à un individu contractuellement engagé auprès d'un organisation pour une mission donnée, définie et délimitée. Les actions du patient sont subordonnées par l'agent qui s'en voit responsable comme illustré dans la Figure 29 Utilisation de la relation "appartenance" et le Tableau 10 Tableau de sous propriétés de la relation Belong.

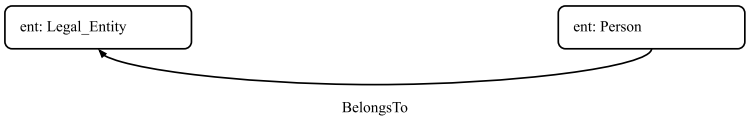


Figure 29 Utilisation de la relation "appartenance"

Propriété d'objet	Domain	Range	Inverse
belongsTo	Person	Legal_Entity	N/A

Tableau 10 Tableau de sous propriétés de la relation Belong

Relation d'implication : Involve

La propriété d'objet *Involve* signifie qu'un concept patient est engagé dans les actions d'un concept agent.

SynSet ID	202677567
SynSet définition	Engage as a participant

Dans notre modèle, les processus fonctionnels impliquent des composants informatiques et des individus comme l'illustre la Figure 30 Utilisation de la relation "implication" et son inverse.

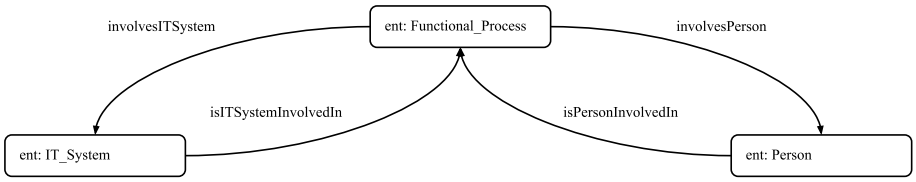


Figure 30 Utilisation de la relation "implication" et son inverse

Cette propriété d'objet dispose de plusieurs sous propriétés correspondant aux classes impliquées dans les différents processus fonctionnels comme illustré dans le

Tableau 11 Tableau de sous propriétés de la relation Involve.

Propriété d’objet	Domain	Range	Inverse
involvesPerson	Functional_Process	Person	isPersonInvolvedIn
involvesITSystem	Functional_Process	IT_System	isITSystemInvolvedIn

Tableau 11 Tableau de sous propriétés de la relation Involve

Relation de protection : Protect

La propriété d’objet *Protect* signifie qu’un concept agent protège un concept patient contre les dangers, blessures, destructions ou dommages.

SynSet ID	201128193
SynSet définition	Shield from danger, injury, destruction, or damage

Dans le cadre de notre modèle, il s’agit de protéger la confidentialité, l'intégrité et la disponibilité des concepts technologiques (*IT_System*, *Technological_Data*, etc.) en répondant à un ensemble d'exigences de sécurité définies. La Figure 31 Utilisation de la relation "protection" et son inverse montre le champ d’action des mesures de sécurité.

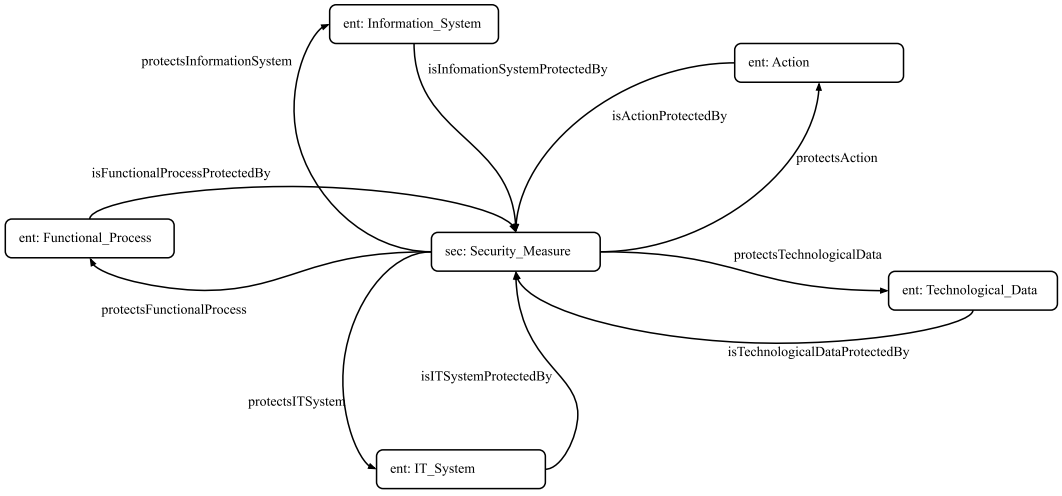


Figure 31 Utilisation de la relation "protection" et son inverse

Cette propriété d’objet dispose de plusieurs sous propriétés correspondant aux classes protégées par les différentes mesures de sécurité comme illustré dans le

Tableau 12 Tableau de sous propriétés de la relation Protect.

Propriété d'objet	Domain	Range	Inverse
protectsInformationSystem	Security_Measure	Information_System	isInfomationSystemProtectedBy
protectsFunctionalProcess	Security_Measure	Functional_Process	isFunctionalProcessProtectedBy
protectsTechnologicalData	Security_Measure	Technological_Data	isTechnologicalDataProtectedBy
protectsAction	Security_Measure	Action	isActionProtectedBy
protectsITSystem	Security_Measure	IT_System	isITSystemProtectedBy

Tableau 12 Tableau de sous propriétés de la relation Protect

Relation de définition : Define

La propriété d'objet *Define* signifie qu'un concept Agent fixe les paramètres d'un concept patient.

SynSet ID	200947077
SynSet définition	Decide upon or fix definitely

Dans le cadre de notre modèle, la classe *Documentation* nous permet de représenter les deux principales sources de documents : le corpus de textes et documentations d'origine externe ainsi que les informations formalisées qui se rapportent au système d'information et les éléments qui le composent. Ces documents ont pour mission de décrire les mesures de sécurité. De façon similaire, la classe *Act* définit les exigences de sécurité par les modalités juridiques. Enfin, la classe *Norm*, qui reprend le corpus juridique, définit les modalités juridiques comprises dans la classe *Act*. La Figure 32 Utilisation de la relation "définition" et son inverse représente les différentes utilisation de la propriété *Define*.

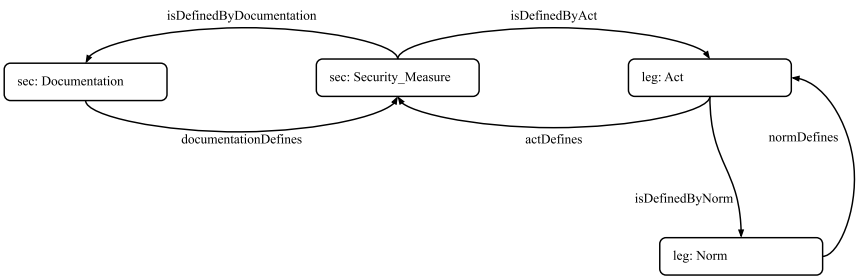


Figure 32 Utilisation de la relation "définition" et son inverse

Cette propriété d'objet dispose de plusieurs sous propriétés correspondant aux classes définies par d'autres comme illustré dans le Tableau 13 Tableau de sous propriétés de la relation Define.

Propriété d'objet	Domain	Range	Inverse
documentationDefines	Documentation	Security_Measure	isDefinedByDocumentation
actDefines	Act	Security_Measure	isDefinedByAct
normDefines	Norm	Act	isDefinedByNorm

Tableau 13 Tableau de sous propriétés de la relation Define

Relation de gestion : Manage

La propriété d'objet *Manage* signifie qu'un concept agent est en charge, agit sur ou dispose d'un concept patient.

SynSet ID	202436349
SynSet définition	Be in charge of, act on, or dispose of

Dans le cadre de notre modèle, le concept agent est légitimement et légalement responsable de contrôler et/ou organiser un concept patient. Dans notre modèle, les systèmes d'information sont gérés par des organisations tel qu'illustré dans la Figure 33 Utilisation de la relation "gestion" et son inverse.

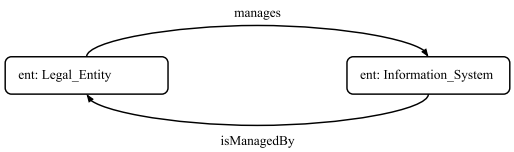


Figure 33 Utilisation de la relation "gestion" et son inverse

La relation *Manage* permet représenter des systèmes d'information gérés par des organisations comme indiqué dans le

Tableau 14 Tableau de sous propriétés de la relation Manage.

Propriété d'objet	Domain	Range	Inverse
manages	Legal_Entity	Information_System	isManagedBy

Tableau 14 Tableau de sous propriétés de la relation Manage

Relation de possession : Own

La propriété d'objet *Own* signifie qu'un concept agent possède un concept patient.

SynSet ID	63041172
SynSet définition	Have ownership or possession of

Dans notre modèle, la relation de possession d'une donnée correspond à la définition du RGPD et de la relation entre la « personne concernée » (data subject) et les données à caractères personnel qui lui sont propres [17]. Cette relation n'a lieu d'être que pour les Personal_Data (ent: Personal_Data ⊑ Technological_Data).

La Figure 34 Utilisation de la relation "possession" et son inverse illustre cette relation.

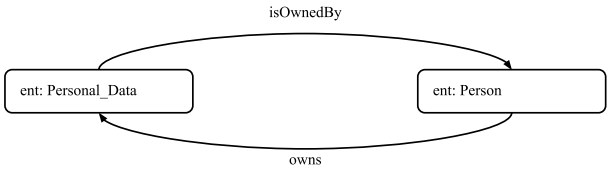


Figure 34 Utilisation de la relation "possession" et son inverse

La relation *Own* permet de représenter la relation entre « une personne concernée » et ses données à caractère personnel comme indiqué dans le

Tableau 15 Tableau de la relation Own.

[17] L'article 4 du Règlement général sur la protection des données définit le concept de « personne concernée » et est accessible via ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32016R0679

Propriété d'objet	Domain	Range	Inverse
owns	Person	Personal_Data	isOwnedBy

Tableau 15 Tableau de la relation Own

Relation de composition : Compose

La propriété d'objet *Compose* signifie qu'un ou plusieurs concepts agents forment un concept patient.

SynSet ID	201626138
SynSet définition	Put together out of existing material

Dans notre modèle, les systèmes d'information sont composés d'au moins un processus fonctionnel. La création de ce concept nous permet de regrouper différents processus fonctionnels en un système d'information comme présenté dans la Figure 35 Utilisation de la relation "composition" et son inverse.

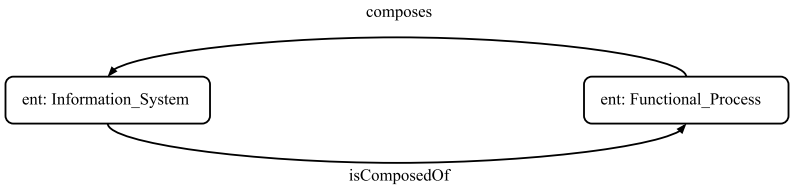


Figure 35 Utilisation de la relation "composition" et son inverse

La relation *Compose* permet représenter des systèmes d'information composés de processus fonctionnels comme indiqué dans le

Tableau 16 Tableau de la relation Compose.

Propriété d'objet	Domain	Range	Inverse
isComposedOf	Information_System	Functional_Process	composes

Tableau 16 Tableau de la relation Compose

Relation d’action

Relation d’effet : Impact

La propriété d’objet *Impact* signifie qu’un concept agent a un effet sur un concept patient.

SynSet ID	200137313
SynSet définition	Have an effect upon

Dans notre modèle, le risque règlementaire est impacté par la création, modification et implémentation de documentations et de mesures de sécurité comme l’illustre la Figure 36 Utilisation de la relation "impact" et son inverse.

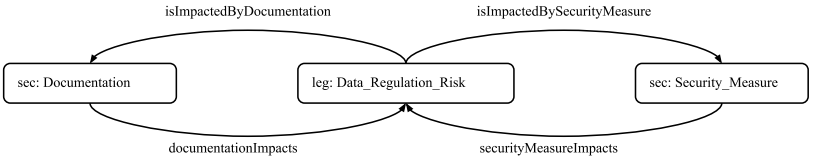


Figure 36 Utilisation de la relation "impact" et son inverse

Cette propriété d’objet dispose de plusieurs sous propriétés correspondant aux classes qui ont un impact sur le risque réglementaire comme illustré dans le

Tableau 17 Tableau de sous propriétés de la relation Impact.

Propriété d’objet	Domain	Range	Inverse
securityMeasureImpacts	Security_Measure	Data_Regulation_Risk	isImpactedBySecurityMeasure
documentationImpacts	Documentation	Data_Regulation_Risk	isImpactedByDocumentation

Tableau 17 Tableau de sous propriétés de la relation Impact

Relation de création : Create

La propriété d'objet *Create* permet d'illustrer la création d'un concept patient par un concept agent.

SynSet ID	201617192
SynSet définition	Make or cause to be or to become

Notre modèle dispose de deux types de création représentés dans la Figure 37 Utilisation de la relation "création" et son inverse. Le premier type correspond à une organisation qui crée un document et le second représente la création du risque réglementaire inhérent aux lois et réglementations.

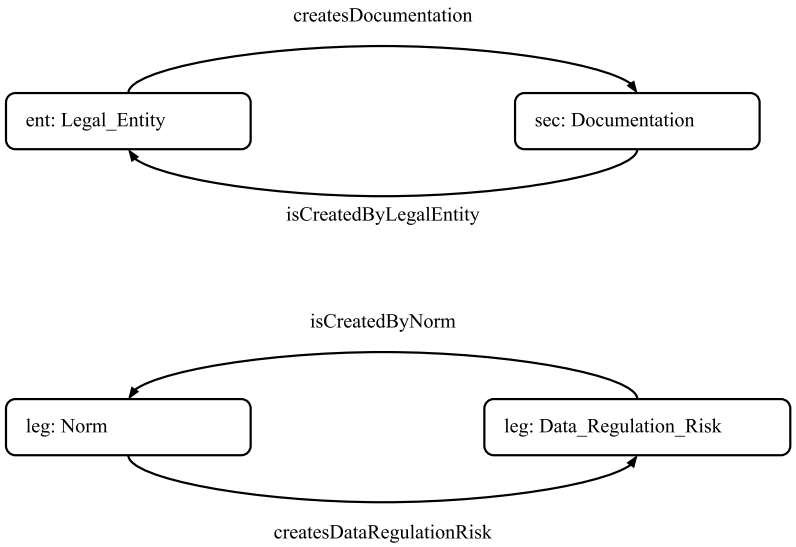


Figure 37 Utilisation de la relation "création" et son inverse

Cette propriété d'objet dispose de plusieurs sous propriétés correspondant aux classes qui sont créées et les classe créatices comme illustré dans le

Tableau 18 Tableau de sous propriétés de la relation Create.

Propriété d'objet	Domain	Range	Inverse
createsDataRegulationRisk	Norm	Data_Regulation_Risk	isCreatedByNorm
createsDocumentation	Legal_Entity	Documentation	isCreatedByLegalEntity

Tableau 18 Tableau de sous propriétés de la relation Create

Relation de traitement : Process

La propriété d'objet *Process* signifie qu'un concept agent effectue des actions particulières destinées à obtenir un résultat.

SynSet ID	101023820
SynSet définition	A particular course of action intended to achieve a result

Dans le cadre de notre modèle, les traitements concernent les opérations sur des données, y compris par des moyens manuels ou automatisés comme l'illustre la Figure 38 Utilisation de la relation "traitement" et son inverse. Si l'on se réfère à l'approche classique des systèmes d'information, un traitement de donnée peut être divisé en quatre principaux types qui sont :

- La recherche et l'extraction de données ;
- Le traitement, la mise à jour et l'exploitation de données pour créer de l'information ;
- La distribution et communication de données et/ou information ;
- Le stockage.

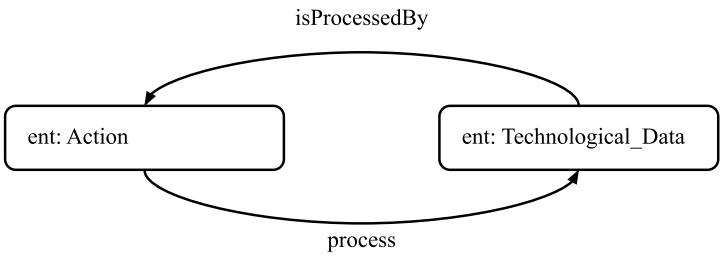


Figure 38 Utilisation de la relation "traitement" et son inverse

La relation *Process* permet de représenter des traitements effectués sur des données comme indiqué dans le

Tableau 19 Tableau de la relation Process.

Propriété d'objet	Domain	Range	Inverse
process	Action	Technological_Data	isProcessedBy

Tableau 19 Tableau de la relation Process

Relation d'utilisation : Use

La propriété d'objet *Use* signifie qu'un concept agent utilise un concept patient pour réaliser une action.

SynSet ID	100947128
SynSet définition	The act of using

Cette propriété permet d'illustrer l'utilisation d'un composant informatique du système d'information par un individu comme montré dans la Figure 39 Utilisation de la relation "utilisation" et son inverse.

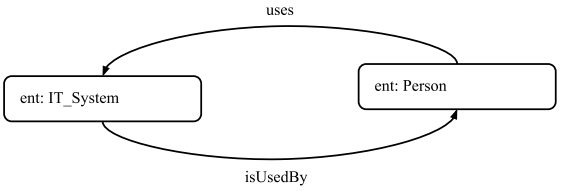


Figure 39 Utilisation de la relation "utilisation" et son inverse

La relation *Use* permet représenter l'utilisation d'un composant informatique par un individu comme indiqué dans le Tableau 20 Tableau de la relation Use.

Propriété d'objet	Domain	Range	Inverse
uses	Person	IT_System	isUsedBy

Tableau 20 Tableau de la relation Use

Relation d'effectuation : Perform

La propriété d'objet *Perform* signifie qu'un concept agent réalise une action (l'action correspond au concept patient).

SynSet ID	201712704
SynSet définition	Carry out or perform an action

Dans notre modèle, les actions sont réalisées par les composants informatiques comme illustré dans la Figure 40 Utilisation de la relation "effectuation" et son inverse.

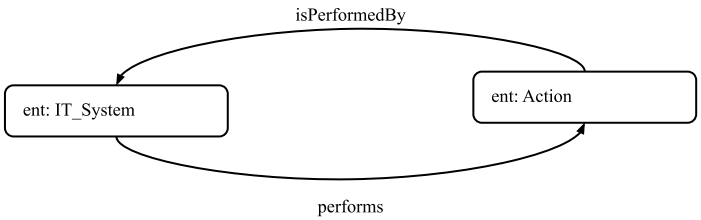


Figure 40 Utilisation de la relation "effectuation" et son inverse

La relation *Perform* permet de représenter l'effet d'une action sur un composant informatique comme indiqué dans le

Tableau 21 Tableau de la relation Perform.

Propriété d'objet	Domain	Range	Inverse
performs	IT_System	Action	isPerformedBy

Tableau 21 Tableau de la relation Perform

La prochaine partie présente comment les utilisateurs peuvent créer et rechercher des instances.

2.2. Fonctionnalités du système proposé

Nous avons défini dans la première section de ce chapitre quelles étaient les utilisations prévues de notre modèle par les utilisateurs. Pour rappel, nous souhaitons limiter les utilisateurs finaux aux actions effectuées sur les instances et ne pas autoriser les actions effectuées sur les autres entités de notre ontologie (concepts et propriétés).

Cette partie présente les processus d'instanciation et de consultation à disposition des utilisateurs. Nous utilisons des diagrammes de séquence pour illustrer les étapes de ces différents processus.

Nous présentons également le processus de recherche par un diagramme d'activité en fin de section.

2.2.1. Processus de consultation

Le processus de consultation correspond à la recherche non assistée des informations. Cela correspond à permettre à un utilisateur de consulter les informations, telles que les instances créées, par la sélection de critère de recherche et requête.

Le processus de consultation repose sur plusieurs étapes :

1. L'expert métier envoie au système une requête de consultation.
2. Le système va rechercher les sous-ontologies puis les importer.
3. L'expert métier sélectionne l'ontologie qu'il souhaite consulter.
4. Le système va rechercher les concepts de la sous-ontologie sélectionnée puis les importer.
5. L'expert métier sélectionne le concept qu'il souhaite consulter.
6. Le système va rechercher les instances du concept sélectionné puis les importer.

La Figure 41 Processus de consultation représente ces différentes étapes par un diagramme de séquence.

L'objectif de notre modèle est de répondre aux questions :

- Qu'est ce qui est réglementé ?
- Comment est-ce réglementé ?
- Pourquoi est-ce réglementé ?

Afin de fournir ces informations aux utilisateurs, le processus de consultation doit disposer de plusieurs étapes qui viennent compléter celles mentionnées ci-avant:

7. L’expert métier sélectionne l’instance qu’il souhaite consulter.
8. Le système va rechercher les propriétés d’objet de l’instance sélectionnée puis les importer.
9. L’expert métier sélectionne les propriétés d’objet qu’il souhaite consulter.
10. Le système va rechercher les instances qui sont liées à cette propriété puis les importer.

Ces quatre étapes supplémentaires peuvent être répétées jusqu’à ce que l’utilisateur ait obtenu les informations souhaitées. En réitérant les étapes 7 à 10, un utilisateur peut « remonter » les instances de l’ontologie pour obtenir les informations souhaitées.

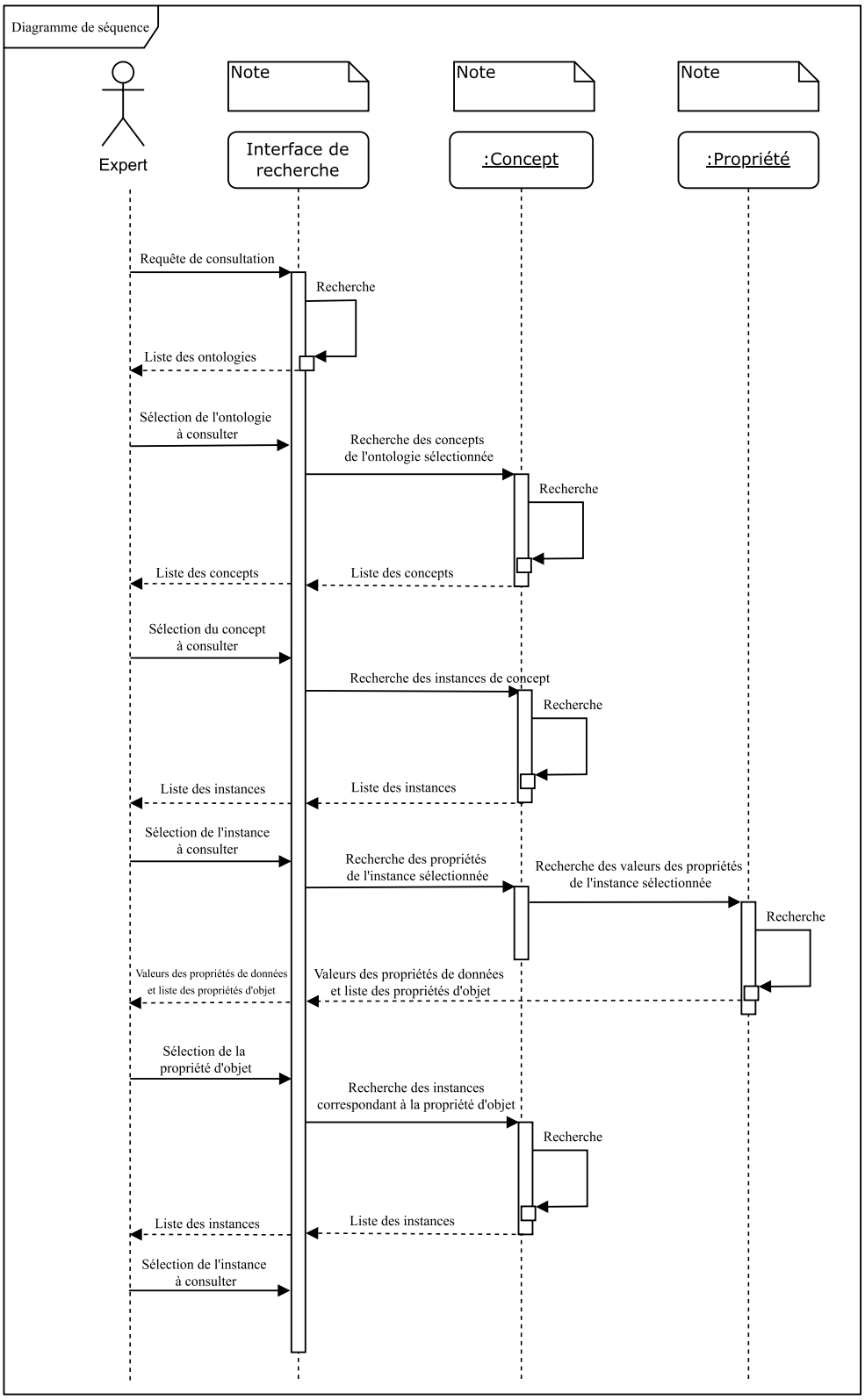


Figure 41 Processus de consultation

La troisième section de ce chapitre présente un exemple détaillé de consultation et comment un utilisateur peut trouver les réponses aux différentes questions de compétences.

2.2.2. Processus d'instanciation

Le processus d'instanciation correspond à la création d'instance de concept. Cela permet un utilisateur de consulter les instances existantes et d'en créer de nouvelles si besoin.

Le processus d'instanciation repose sur plusieurs étapes :

1. L'expert métier envoie au système une requête d'instanciation.
2. Le système recherche les sous-ontologies puis les importe.
3. L'expert métier sélectionne l'ontologie dans laquelle il souhaite créer une instance.
4. Le système recherche les concepts de la sous-ontologie sélectionnée puis les importe.
5. L'expert métier sélectionne le concept qu'il souhaite instancier.
6. Le système recherche les instances du concept sélectionné puis les importe.
7. L'expert métier valide qu'il souhaite créer une nouvelle instance et saisit les propriétés de données du concept.
8. Le système valide la création d'instance et recherche les propriétés d'objet du concept sélectionné puis les importe.
9. L'expert métier sélectionne la propriété d'objet qu'il souhaite saisir.
10. Le système recherche les instances des concepts liés à la propriété d'objet sélectionnée puis les importe.
11. L'expert métier sélectionne une instance existante ou crée une nouvelle instance comme valeur de propriété d'objet.

La onzième étape diffère selon si l'instance qui doit être reliée n'existe pas. Dans le cas échéant, l'expert métier doit créer une nouvelle instance et retourne à l'étape sept. La Figure 42 Processus d'instanciation représente ces différentes étapes par un diagramme de séquence.

La troisième section de ce chapitre présente un exemple détaillé du processus par l'instanciation de lois.

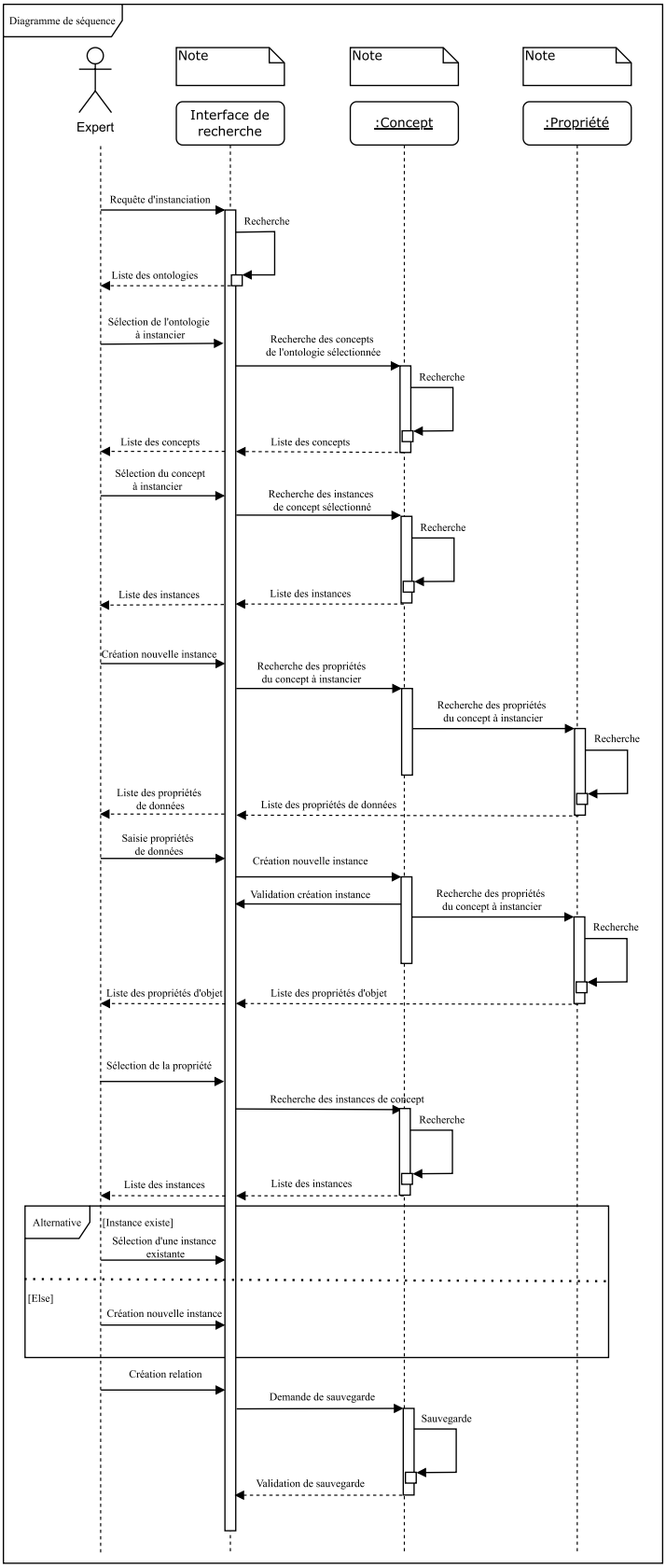


Figure 42 Processus d'instanciation

2.2.3. Processus de recherche

Le processus de recherche correspond à la recherche assistée d'information. Cela correspond à guider un utilisateur souhaitant consulter des informations par le biais de requêtes prédéfinies. La même problématique que les questions de recherche existe pour le processus de recherche, à savoir l'impossibilité d'être exhaustif. Il nous parait en effet compliqué de prévoir l'ensemble des recherches des futurs utilisateurs.

Nous avons cependant déterminé le fonctionnement du processus de recherche pour les questions que nous estimons les plus fréquentes et les plus communes.

Par exemple, si l'on reprend la « question fille » suivante :

• Quels sont les éléments du système d'information qui sont impliqués dans le traitement d'une donnée réglementée ?

L'objectif de la recherche assistée est alors de fournir la liste des instances d'une classe sélectionnée par un utilisateur et pour laquelle les instances sont directement ou indirectement reliée à une instance de donnée réglementée.

Les éléments du système d'information pouvant être impliqués dans le traitement d'une donnée réglementée sont :

- Les processus fonctionnels ;
- Les composants informatiques ;
- Les individus.

Selon l'élément sélectionné par l'utilisateur, le processus de recherche doit alors être en mesure de rechercher les instances correspondant au concept sélectionné puis les importer.

Le processus de recherche repose sur plusieurs étapes. Pour cet exemple, celles-ci sont :

1. L'expert métier sélectionne la question fille : « Quels sont les éléments du système d'information qui sont impliqués dans le traitement d'une donnée réglementée ? ».
2. Le système demande à l'utilisateur de sélectionner l'élément souhaité parmi les processus fonctionnels, les composants informatiques (et sous classes) et les individus.
3. L'expert métier sélectionne l'élément qu'il souhaite rechercher.

4. Le système demande à l'utilisateur de préciser si la recherche doit être effectuée sur une instance de donnée ou sur une instance de modalité juridique « act ».
5. L'expert métier sélectionne le critère de recherche souhaité.
6. Le système recherche les instances correspondantes puis les importe.

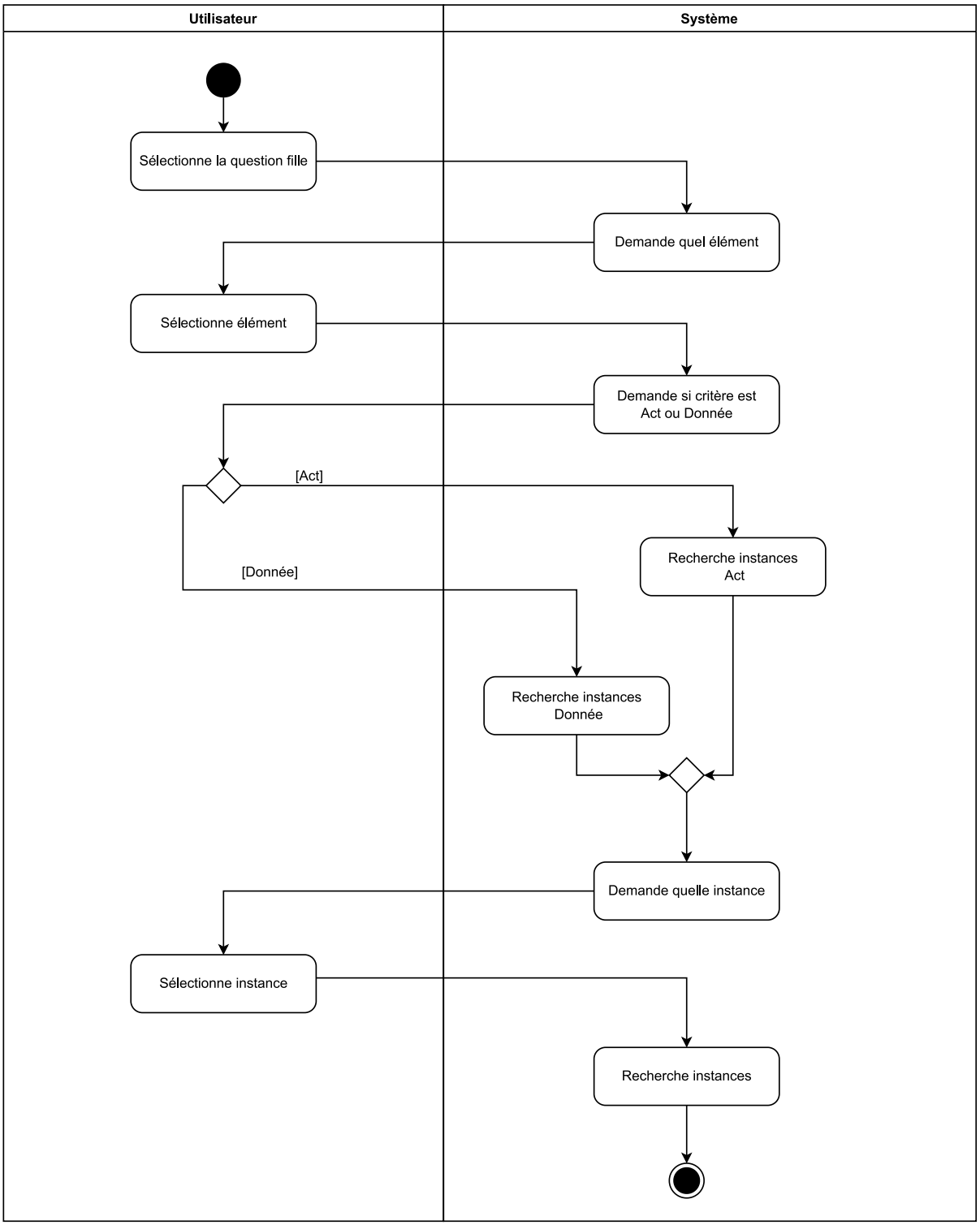


Figure 43 Processus de recherche

La Figure 43 Processus de recherche illustre les différentes activités du processus. La recherche d'instance de la quatrième étape est transparente pour l'utilisateur. Celle-ci correspond à :

- Dans le cas d'une recherche de composant :

IT_System ⊑ performs.**Action** ⊓ process.**Technological_Data**

- Dans le cas d'une recherche d'individu :

Person ⊑ uses.**IT_System** ⊓ performs.**Action** ⊓ process.**Technological_Data**

- Dans le cas d'une recherche de processus fonctionnel :

Functional_Process ⊑ involvesITSystem.**IT_System** ⊓ performs.**Action** ⊓ process.**Technological_Data**

Ensuite, selon la réponse de l'utilisateur, le système complète la requête en spécifiant l'instance de la donnée ou en ajoutant les critères de recherches spécifiques à la modalité juridique :

Business_Data ⊑ isBusinessDataGovernedBy.**Act**

Personal_Data ⊑ isPersonalDataGovernedBy.**Act**

Enfin, le système regroupe les informations dans une requête finale afin de fournir la liste des instances qui correspondent aux critères.

Conclusion Section 2

Cette deuxième section nous a permis d'expliquer la conception de notre modèle. La Figure 44 Ontologie principale présente l'ensemble des concepts utilisés pour représenter assidûment notre domaine. Ceux-ci sont au nombre de 15 et sont répartis en quatre sous domaines :

- Le sous domaine de l'entreprise ;
- Le sous domaine de la sécurité ;
- Le sous domaine réglementaire ;
- Le sous domaine de la localisation.

Dans un second temps, cette section a introduit les relations ou propriétés d'objet qui permettent de relier ces concepts. Celles-ci sont d'autant plus importantes qu'elles permettent de mettre en place notre logique sémantique et par conséquent la logique de notre modèle.

L'Appendix 1 Liste des relations regroupe l'ensemble des propriétés d'objet utilisée par notre modèle.

L'Appendix 2 Liste des classes de l'ontologie apporte également davantage de précision quant aux modifications des SynSet nécessaires pour adapter les définitions à notre modèle. Enfin, l'Appendix 3 Liste des propriétés de données par classe regroupe l'ensemble des propriétés de données présente dans notre modèle.

La prochaine section de ce chapitre présente différents exemples d'instanciation et d'utilisation de notre modèle.

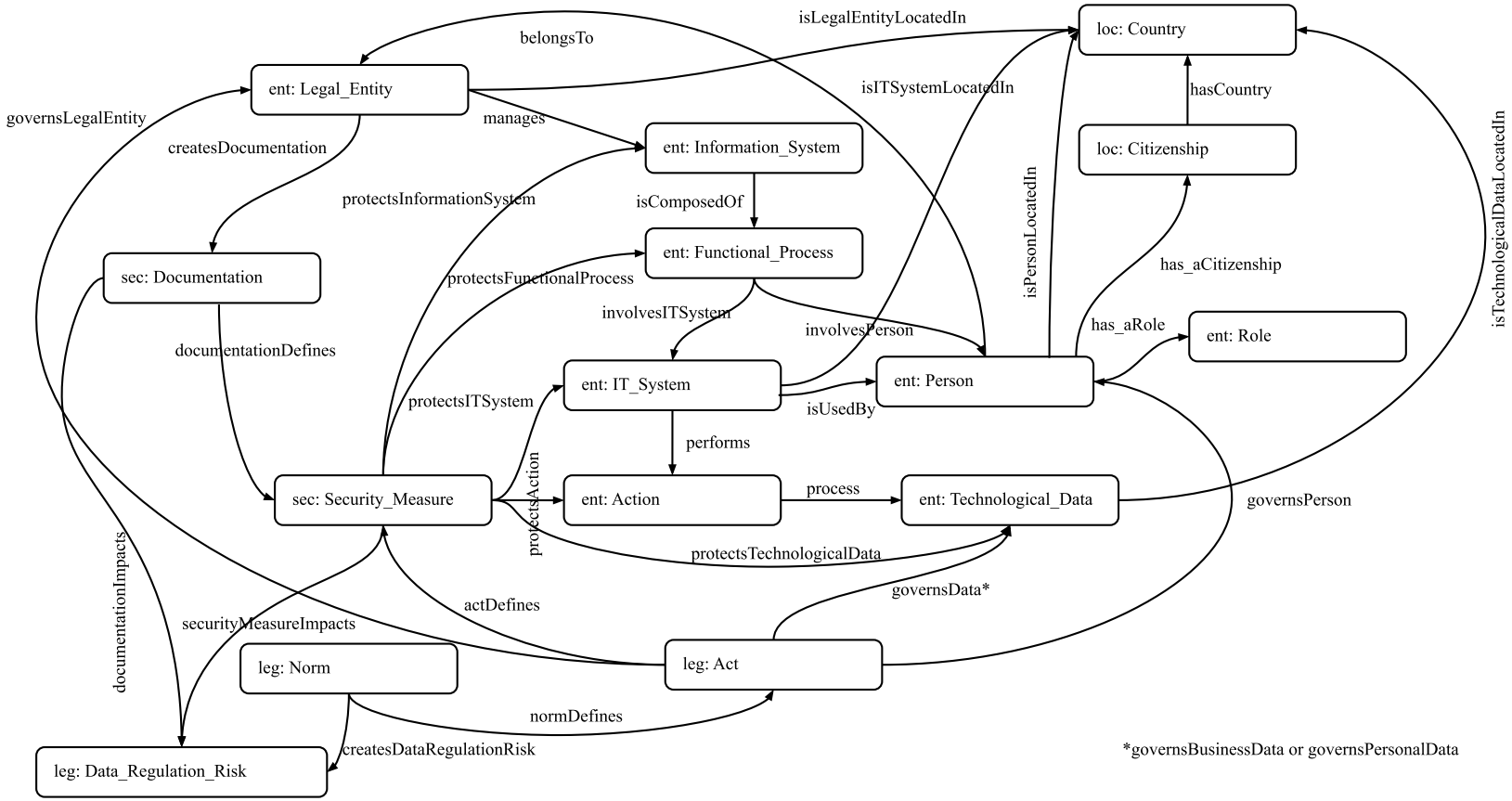


Figure 44 Ontologie principale

Section III. Quelques exemples d'instanciation et d'utilisation

Cette troisième et dernière section nous permet de présenter des exemples d'instanciation et d'utilisation de notre modèle par la représentation de deux lois (Delorme et al., 2022a) :

- L'instanciation de l'EAR Supplement No. 18 to part 734 ;
- L'instanciation de l'article 32 du RGPD.

3.1. Modélisation de la loi américaine EAR Supplement No. 18 to part 734

Notre première instanciation sera la règlementation américaine Export Arm Regulations (EAR). Celle-ci est émise par le Bureau de l'Industrie et de la Sécurité (Bureau of Industry and Security) rattaché au département de commerce des Etats-Unis[18].

Il existe un vaste corpus juridique sous-jacent à l'EAR qui sont répertoriés dans les documents du Federal Register. Ce sont ces derniers qui promulguent l'EAR. Afin d'instancier notre modèle, nous allons nous concentrer sur l'un d'entre eux et formaliser le Supplément No. 18 to part 734 (ACTIVITIES THAT ARE NOT EXPORTS, REEXPORTS, OR TRANSFERS) de l'EAR.

[18] Le Bureau of Industry and Security (BIS) est présenté sur le site web : https://www.bis.doc.gov/.

Celui-ci stipule :

§ 734.18 ACTIVITIES THAT ARE NOT EXPORTS, REEXPORTS, OR TRANSFERS

(a) The following activities are not exports, reexports, or transfers:

(1) Launching a spacecraft, launch vehicle, payload, or other item into space.

(2) Transmitting or otherwise transferring “technology” or “software” to a person in the United States who is not a foreign person from another person in the United States.

(3) Transmitting or otherwise making a transfer (in-country) within the same foreign country of “technology” or “software” between or among only persons who are not “foreign persons,” so long as the transmission or transfer does not result in a release to a foreign person or to a person prohibited from receiving the “technology” or “software.”

(4) Shipping, moving, or transferring items between or among the United States, the District of Columbia, the Commonwealth of Puerto Rico, or the Commonwealth of the Northern Mariana Islands or any territory, dependency, or possession of the United States as listed in Schedule C, Classification Codes and Descriptions for U.S. Export Statistics, issued by the Bureau of the Census.

(5) Sending, taking, or storing “technology” or “software” that is:

(i) Unclassified;

(ii) Secured using 'end-to-end encryption;'

(iii) Secured using cryptographic modules (hardware or “software”) compliant with Federal Information Processing Standards Publication 140-2 (FIPS 140-2) or its successors, supplemented by “software” implementation, cryptographic key management and other procedures and controls that are in accordance with guidance provided in current U.S. National Institute for Standards and Technology publications, or other equally or more effective cryptographic means; and

(iv) Not intentionally stored in a country listed in Country Group D:5 (see Supplement No. 1 to part 740 of the EAR) or in the Russian Federation.

Note 1 to paragraph (a)(5)(iv): Data in-transit via the Internet is not deemed to be stored.

(b) Definitions

For purposes of this section, End-to-end encryption means (i) the provision of cryptographic protection of data such that the data is not in unencrypted form between an originator (or the originator's in-country security boundary) and an intended recipient (or the recipient's in-country security boundary), and (ii) the means of decryption are not provided to any third party. The originator and the recipient may be the same person.

(c) The ability to access “technology” or “software” in encrypted form that satisfies the criteria set forth in paragraph (a)(5) of this section does not constitute the release or export of such “technology” or “software.”

Nous pouvons commencer par utiliser les concepts Norm et Act pour instancier l’EAR et le Supplément No. 18 to part 734 :

EARNorm *is_a (leg: Regulation ⊑ Norm ⊑ T)*

EAR734.18Act *is_a (leg: Act ⊑ T)*

Ces deux instances interagissent avec la fonction de définition:

EARNorm *normDefines* ***EAR734.18Act***

Le Supplément No. 18 to part 734 définit des obligations auxquelles les organisations doivent répondre. Pour illustrer cela, nous utilisons la propriété de données: «hasActType » qui permet de différencier les modalités juridiques par leur nature : interdiction, obligation ou libre choix. L’instance ***EAR734.18Act*** est alors dotée d’un type qui a pour valeur « Obligation ».

Cette section de l’EAR mentionne également des données réglementées. Celles-ci comprennent les données liées aux technologies et logiciels mais aussi les données « classified » (contrôlées) :

EARBusiness_Data *is_a (ent: Business_Data ⊑ Technological_Data)*

Seuls les points numéro 2, 3 et 5 du Supplément No. 18 to part 734 concerne les données réglementées. Les points 2 et 3 étant similaires à l’exception de la contrainte du pays, nous nous concentrerons uniquement sur l’instanciation des points 2 et 5 pour cet exemple d’instanciation.

3.1.1. Point 2 du Supplement No. 18 to part 734

Cette section mentionne deux individus, le premier recevant l'accès à une donnée (réglementée) et le second étant à l'origine de l'envoi de celle-ci.

PersonReceivingEARData *is_a (ent: Person ⊑ T)*

PersonSendingEARData *is_a (ent: Person ⊑ T)*

Il s'agit ensuite d'utiliser les concepts Country et Citizenship pour illustrer les contraintes de nationalités et de localisations des deux individus instanciés précédemment :

US *is_a (loc: Country ⊑ T)*

USCitizenship *is_a (loc: Citizenship ⊑ T)*

Enfin, nous pouvons lier ces différentes instances via les propriétés d'objet présentées dans la section précédentes :

PersonReceivingEARData *isPersonLocatedIn.****US*** *⊓ has_aCitizenship.****USCitizenship***

PersonSendingEARData *isPersonLocatedIn.****US***

Pour illustrer cela dans notre modèle, le concept Action sera utilisé pour représenter le transfert de données contrôlées :

TransferEARData *is_a (ent: Action ⊑ T)*

Pour ajouter une couche supplémentaire de granularité, nous pouvons créer des instances supplémentaires telles que l'octroi de l'accès et son inverse, la réception de l'accès :

GrantAccessEARData *is_a (ent: Action ⊑ T)*

ReceiveAccessEARData *is_a (ent: Action ⊑ T)*

En utilisant la propriété de données « ActionType », nous pouvons spécifier que ces deux instances du concept Action sont de type « TransferAction »

Nous pouvons en déduire que transmettre ou transférer serait instancié comme suit :

An person that uses an ***EARSystem*** *is_a (ent:IT_System ⊑ T) to perform the action* ***TransferEARData*** *that process* ***EARBusiness_Data****.*

Pour conclure, le deuxième point du Supplement No. 18 to part 734 s'illustre par les règles suivantes:

EAR734.18Act *⊑ (governsPerson.****PersonReceivingEARData)*** *⊓ (isPersonLocatedIn.****US)*** *⊓ (has_aCitizenship.****USCitizenship)*** *⊓ (uses.****EARSystem)*** *⊓ (performs.****ReceiveAccessEARData)*** *⊓ (process.****EARBusiness_Data)***

EAR734.18Act $\sqsubseteq$ *(governsPerson.**PersonSendingEARData**)* $\sqcap$ *(isPersonLocatedIn.**US**)* $\sqcap$ *(uses.**EARSystem**)* $\sqcap$ *(perform.**GrantAccessEARData**)* $\sqcap$ *(process.**EARBusiness_Data**)*

3.1.2. Point 5 du Supplement No. 18 to part 734

Cette section mentionne une mesure de sécurité de protection, le chiffrement de bout en bout (« end-to-end encryption » (E2EE)) :

E2EE *is_a (sec: Protect_Security_Measure* $\sqsubseteq$ *Security_Measure* $\sqsubseteq$ *T)*

Cette mesure protège les données réglementées :

E2EE *protectsTechnologicalData.**EARBusiness_Data***

Ou inversement :

EARBusiness_Data *isTechnologicalDataProtectedBy.**E2EE***

Il s'agit ensuite d'utiliser les concepts Legal_Entity et Documentation pour représenter respectivement le U.S. National Institute for Standards and Technology (NIST) et le Federal Information Processing Standards Publication 140-2 (FIPS 140-2) :

NIST *is_a (ent:Independant_Organization* $\sqsubseteq$ *Legal_Entity)*

FIPS140-2 *is_a (sec : Standard* $\sqsubseteq$ *Documentation)*

NIST *createsDocumentation.**FIPS140-2***

Ou inversement :

FIPS140-2 *isCreatedByLegalEntity.**NIST***

Enfin, nous pouvons illustrer que la mesure de sécurité E2EE est décrite par le FIPS 140-2 :

FIPS140-2 *documentationDefines.**E2EE***

Concernant les restrictions de stockage de données réglementées, nous utilisons le concept Country pour représenter les pays du groupe D5[19] et $\neg$ pour représenter la négation dans une logique formelle:

D5 *is_a (loc: Country* $\sqsubseteq$ *T)*

EARBusiness_Data *isTechnologicalDataLocatedIn.*$\neg$***D5***

[19] Dans la pratique, chaque pays doit être mentionné. Pour faciliter la lecture, nous créons uniquement une instance de pays « D5 » qui suffit à illustrer nos propos.

Pour conclure, le cinquième point du Supplement No. 18 to part 734 s'illustre par les règles suivantes:

$$\textbf{\textit{EAR734.18Act}} \sqsubseteq (governsBusinessData.\textbf{\textit{EARBusiness\_Data}}) \sqcap (isTechnologicalDataLocatedIn.\ \neg \textbf{\textit{D5}})$$

$$\textbf{\textit{EAR734.18Act}} \sqsubseteq (governsBusinessData.\textbf{\textit{EARBusiness\_Data}}) \sqcap (isTechnologicalDataProtectedBy.\textbf{\textit{E2EE}}) \sqcap (isDefinedByDocumentation.\textbf{\textit{FIPS140-2}}) \sqcap (isCreatedByLegalEntity.\textbf{\textit{NIST}}) \sqcap (documentationDefines.\textbf{\textit{E2EE}})$$

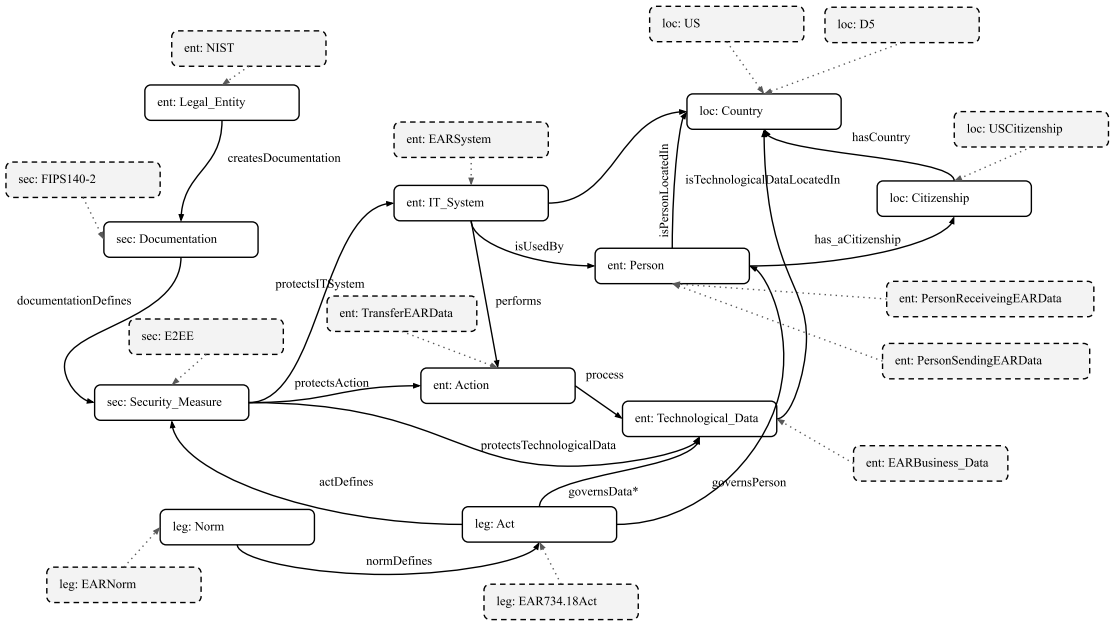


Figure 45 Supplement No. 18 to part 734

La prochaine partie présente l'instanciation de l'article 32 du Règlement Général sur la Protection des Données (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.).

3.2. Instanciation de l'article 32 du Règlement Général pour la Protection des Données

Notre seconde instanciation sera l'article 32 du Règlement Général sur la Protection des Données (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L*

119/1., n.d.). Alors que la section 2 du RGPD fixe les règles de sécurité des données personnelles, l'article 32 se concentre sur la sécurité du traitement en stipulant ce qui suit :

Article 32 : Sécurité du traitement

1. Compte tenu de l'état des connaissances, des coûts de mise en œuvre et de la nature, de la portée, du contexte et des finalités du traitement ainsi que des risques, dont le degré de probabilité et de gravité varie, pour les droits et libertés des personnes physiques, le responsable du traitement et le sous-traitant mettent en œuvre les mesures techniques et organisationnelles appropriées afin de garantir un niveau de sécurité adapté au risque, y compris entre autres, selon les besoins:

a) la pseudonymisation et le chiffrement des données à caractère personnel;
b) des moyens permettant de garantir la confidentialité, l'intégrité, la disponibilité et la résilience constantes des systèmes et des services de traitement;
c) des moyens permettant de rétablir la disponibilité des données à caractère personnel et l'accès à celles-ci dans des délais appropriés en cas d'incident physique ou technique;
d) une procédure visant à tester, à analyser et à évaluer régulièrement l'efficacité des mesures techniques et organisationnelles pour assurer la sécurité du traitement.

2. Lors de l'évaluation du niveau de sécurité approprié, il est tenu compte en particulier des risques que présente le traitement, résultant notamment de la destruction, de la perte, de l'altération, de la divulgation non autorisée de données à caractère personnel transmises, conservées ou traitées d'une autre manière, ou de l'accès non autorisé à de telles données, de manière accidentelle ou illicite.

3. L'application d'un code de conduite approuvé comme le prévoit l'article 40 ou d'un mécanisme de certification approuvé comme le prévoit l'article 42 peut servir d'élément pour démontrer le respect des exigences prévues au paragraphe 1 du présent article.

4. Le responsable du traitement et le sous-traitant prennent des mesures afin de garantir que toute personne physique agissant sous l'autorité du responsable du traitement ou sous celle du sous-traitant, qui a accès à des données à caractère personnel, ne les traite pas, excepté sur instruction du responsable du traitement, à moins d'y être obligée par le droit de l'Union ou le droit d'un État membre.

Un responsable de traitement (controller) est défini par le RGPD (*EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.*, n.d.) comme « la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement; lorsque les finalités et les moyens de ce traitement sont déterminés par le droit de l'Union ou le droit d'un État membre, le responsable du traitement peut être désigné ou les critères spécifiques applicables à sa désignation peuvent être prévus par le droit de l'Union ou par le droit d'un État membre ». Le « processor » correspond au sous-traitant agissant pour le compte d'un responsable de traitement.

Cette section mentionne deux Legal_Entity qui sont le processor et le controller :

***EnterpriseProcesso**r is_a (ent: Business_Organization ⊑ Legal_Entity)*

***EnterpriseController** is_a (ent: Business_Organization ⊑ Legal_Entity)*

L'article 32 présente ensuite la nécessité de mettre en place des mesures de sécurité qui seront définies dans une politique de sécurité de l'information. Pour cela, nous utilisons les concepts Documentation et Security_Measure :

***SecurityPolicy** is_a (sec: Policy ⊑ Documentation ⊑ T)*

***SecurityControl** is_a (sec: Security_Measure ⊑ T)*

SecurityPolicy** documentationDefines.**SecurityControl

Pour illustrer les systèmes d'information, processus fonctionnels et les composants informatiques protégés par les mesures de sécurité, nous utilisons les concepts Information_System, Functional_Process et IT_System:

***EnterpriseIS** is_a (ent: Information_System ⊑ T)*

GDPRProcess *is_a (ent: Functional_Process ⊑ ⊤)*

GDPRApplication *is_a (ent: Application ⊑ IT_System)*

Il en découle que :

EnterpriseIS *isComposedOf.****GDPRProcess***

GDPRProcess *involvesITSystem.****GDPRApplication***

Considérant une donnée personnelle réglementée par le RGPD, et l'action d'accéder à une donnée, nous pouvons en déduire :

GDPRData *is_a (ent: Personnal_Data ⊑ Technological_Data ⊑ ⊤)*

AccessGDPRData *is_a (ent: Action ⊑ ⊤)*

SecurityControl *⊑ (protectsInformationSystem.****EnterpriseIS)*** *⊓ (isComposedOf.****GDPRProcess)*** *⊓* (involvesITSystem.***GDPRApplication)*** *⊓ (performs.****AccessGDPRData)*** *⊓(process.****GDPRData***)[20]

Le quatrième point de cet article mentionne un individu (« personne physique ») qui a un rôle attribué (« agissant sous l'autorité »). Pour cela nous utilisons les concepts Rôle et Person :

GDPRUserRole *is_a (ent: Role ⊑ ⊤)*

Employee *is_a (ent: Person ⊑ ⊤)*

Employee *has_aRole.****GDPRUserRole*** *⊓ uses.****GDPRApplication****.*

Si besoin, l'individu peut être relié à l'une des instances d'entreprise (EnterpriseProcessor ou EnterpriseController) par la propriété d'objet belongsTo.

Enfin, il est nécessaire d'avoir des mesures de sécurité pour assurer le contrôle d'accès selon le rôle de l'individu :

SecurityControl *⊑ (protectsITSystem****GDPRApplication)*** *⊓ ((performs.****AccessGDPRData****) ⊓ (process.****GDPRData****)) ⊓ ((isUsedBy.****Employee)*** *⊓ (has_aRole.****GDPRUserRole****)))*

Nous pouvons conclure l'instanciation de l'article 32 du GPDR par la règle suivante :

GDPRNorm *is_a (leg: Norm ⊑ ⊤)*

GDPRArt32 *is_a (leg: Act ⊑ ⊤)*

GDPRNorm *normDefines.****GDPRArt32***

GDPRArt32 *⊑ governsLegalEntity (****EnterpriseController*** *⊔* ***EnterpriseProcessor****)*

[20] Dans la pratique, chaque mesure de sécurité doit être mentionnée. Pour faciliter la lecture, nous illustrons uniquement la relation « protectsInformationSystem » qui suffit à illustrer nos propos.

GDPRArt32 ⊑ *governsLegalEntity (**EnterpriseControlle**r* ⊔ ***EnterpriseProcessor**) ⊓ ((createsDocumentation.**SecurityPolicy**) ⊓ (documentationDefines.**SecurityControl**))*

GDPRArt32 ⊑ *(actDefines.**SecurityControl**)* ⊓ *(protectsITSystem**GDPRApplication**)* ⊓ *((performs.**AccessGDPRData**)* ⊓ *(process.**GDPRData**))* ⊓ *((isUsedBy.**Employee**)* ⊓ *(has_aRole.**GDPRUserRole**)))*

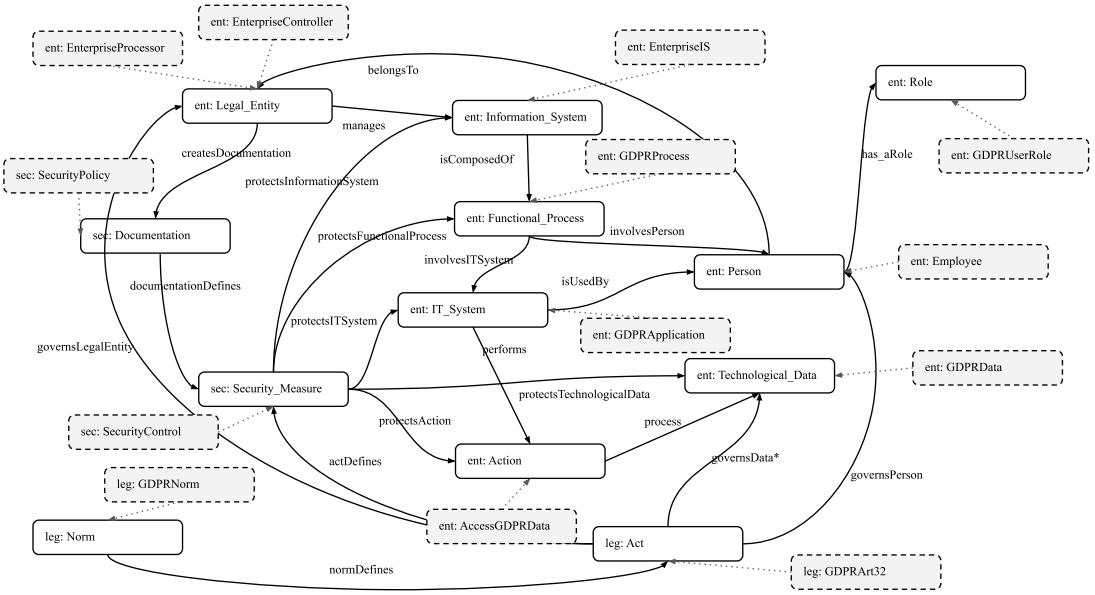


Figure 46 article 32 du RGPD

La prochaine partie présente les processus d'instanciation et de recherche d'information. Nous utiliserons l'instanciation de l'EAR Supplement No. 18 to part 734 comme support d'exemple.

3.3. Un exemple d'utilisation de notre modèle

Dans la section précédente, nous avons présenté les processus génériques d'instanciation, de recherche et consultation d'information. Avec l'aide de l'instanciation de l'EAR Supplément No. 18 to part 734, nous allons maintenant montrer des exemples d'utilisation de ces processus.

3.3.1. Exemple du processus d'instanciation

Pour cet exemple, nous simulons la création de l'instance :

EAR734.18Act *is_a (leg: Act* ⊑ *T)*

Une fois sur le concept « Act » sélectionné par l'utilisateur, celui-ci se voit retourner les instances existantes.

L’utilisateur peut alors décider de créer une nouvelle instance du concept et l’interface lui affiche les propriétés de données.

L’utilisateur est alors en charge d’attribuer les valeurs des propriétés de données suivantes :

- hasName : « *EAR734.18Act » ;*
- hasVersion : « Version1 » ;
- hasDescription : « Instance de l’EAR Supplément No. 18 to part 734 » ;
- hasActType : « Obligation ».

L’interface valide la création de l’instance lorsque les valeurs des propriétés de données sont attribuées.

L’outil retourne ensuite les propriétés d’objet. Dans notre cas, il retourne la propriété « isDefinedByNorm ». Celle-ci est l’inverse de la propriété « normDefines ». Comme défini dans la deuxième section de ce chapitre, cette propriété a pour « domain » la classe « Act » et pour « range » le concept « Norm ».

L’interface présente alors à l’utilisateur les instances existantes du concept « Norm » pour sélection.

Dans le cas où l’instance souhaitée n’existe pas, l’utilisateur peut la créer et lui attribuer les valeurs des propriétés de données.

Une fois l’instance EARNorm créée et sélectionnée, nous avons la relation :

EARNorm *normDefines* ***EAR734.18Act ;***

ou son inverse : ***EAR734.18Act*** *isDefinedByNorm* ***EARNorm***

L’instanciation des autres classes est similaire. La Figure 47 Exemple d'instanciation de la classe "Act" présente un diagramme de séquence illustrant ces étapes.

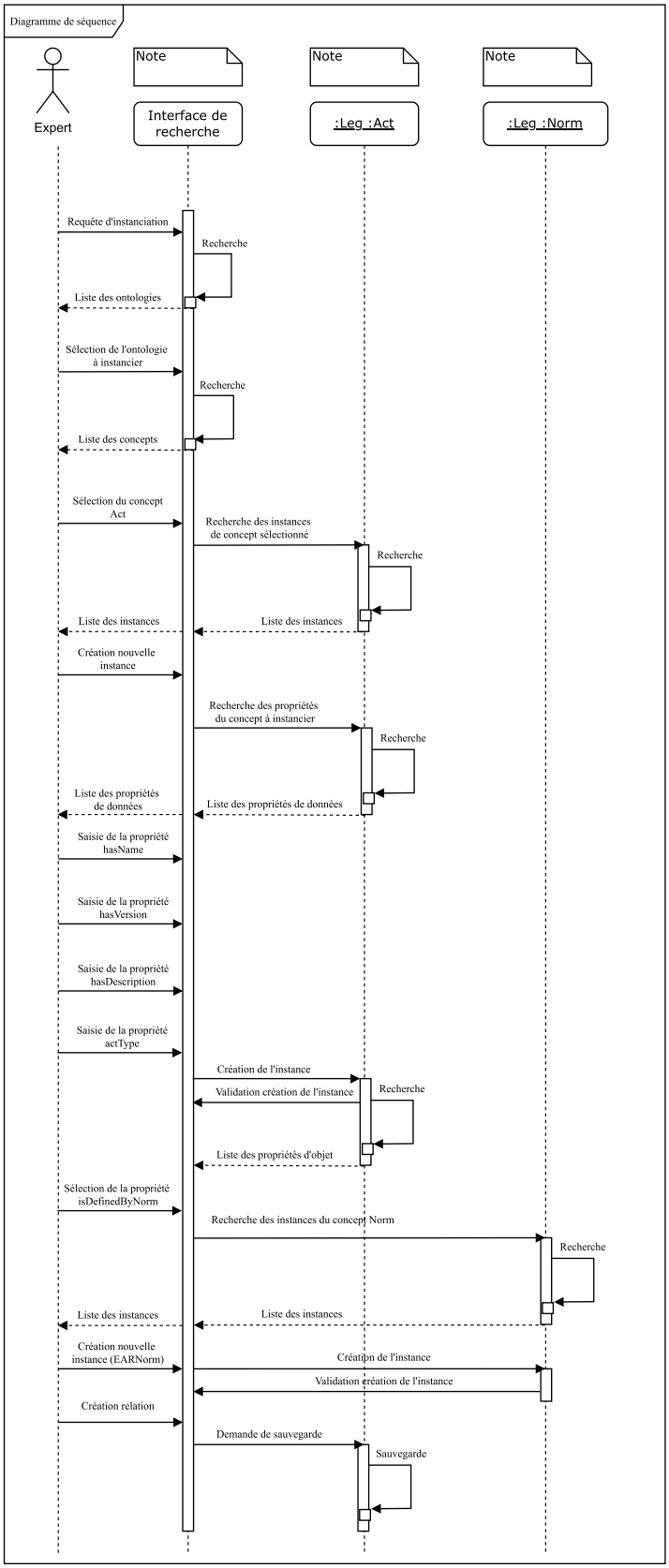


Figure 47 Exemple d'instanciation de la classe "Act"

3.3.2. Exemple du processus de recherche

Nous avons ultérieurement présenté plusieurs questions de compétences auquel notre modèle doit être capable de fournir les réponses pour répondre aux attentes de nos utilisateurs. Présentées dans la section 1.2.2 de ce chapitre, celles-ci sont :

- Qu'est ce qui est réglementé ?
- Comment est-ce réglementé ?
- Pourquoi est-ce réglementé ?

Pour cet exemple, nous simulons la recherche des individus qui sont impactés par la règlementation EAR.

Une fois la classe « Person » sélectionnée par l'utilisateur comme premier élément de recherche, le système demande si le critère de recherche doit être une instance de la classe « Donnée » ou de la classe « Act».

Une fois le critère précisé, l'outil recherche et présente l'ensemble des instances de « Norm ».

L'utilisateur sélectionne l'instance « EAR734.18Act ».

L'outil recherche alors l'ensemble des instances de la classe « Person » qui sont reliées à l'instance « EAR734.18Act ».

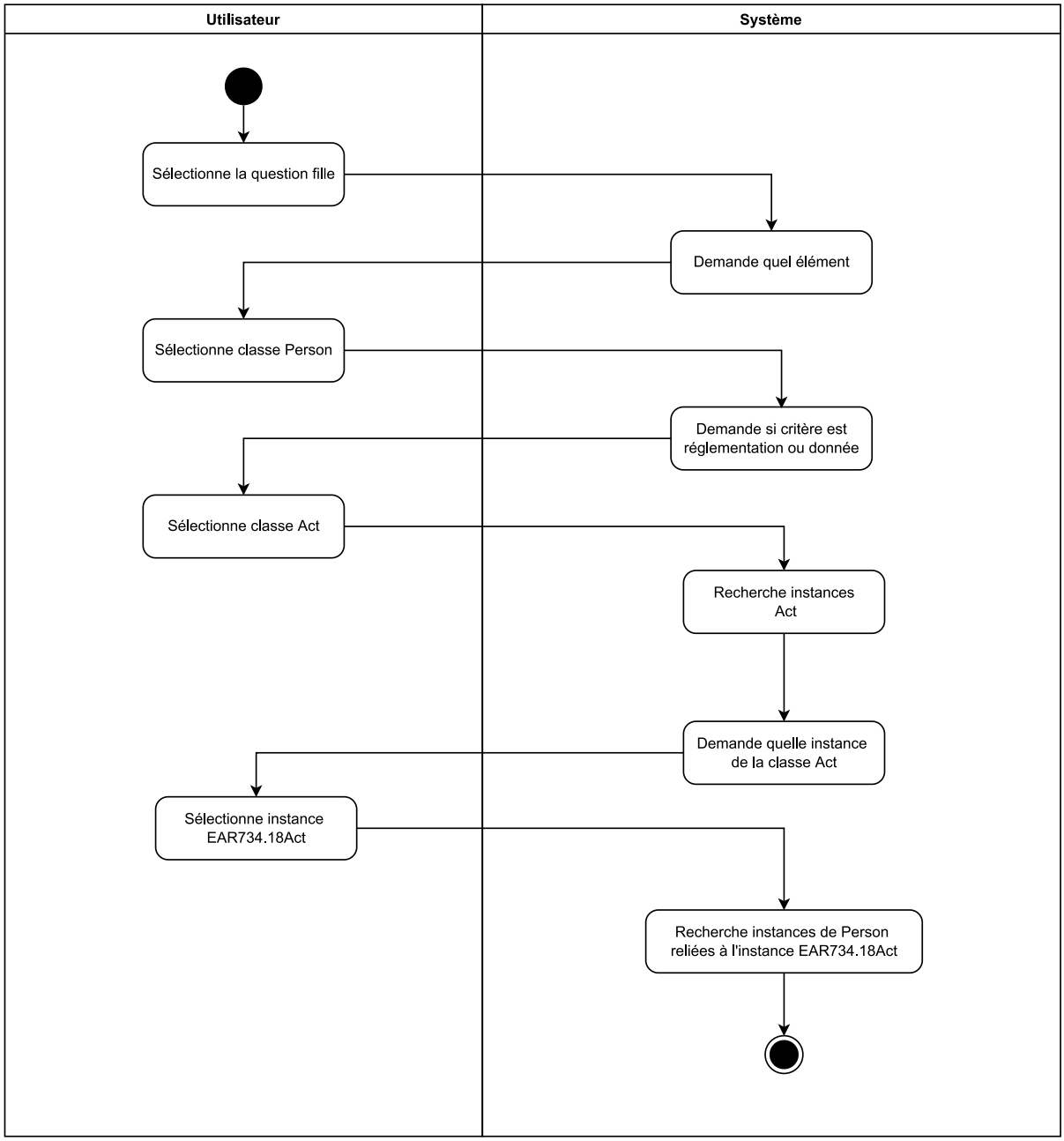


Figure 48 Exemple du processus de recherche

Dans ce cas précis l'outil recherche les instances de « Person » qui utilisent des « IT_System » qui effectuent des « Action » sur des « Technological_Data » qui sont gouvernées par des « Act ». Cela revient pour l'outil d'effectuer la recherche suivante :

Person ⊑ uses.**IT_System** ⊓ performs.**Action** ⊓ process.**Technological_Data**

Pour laquelle les **«** Technological_Data » sont :

Business_Data ⊑ isBusinessDataGovernedBy.**Act**

Personal_Data ⊑ isPersonalDataGovernedBy.**Act**

Conclusion

Le modèle pour l'identification du Data Regulation Risk (DRR) est une proposition originale basée sur une ontologie de 15 concepts de haut niveau.

L'ensemble de la proposition comprend les éléments permettant d'illustrer et identifier un DRR :

- Des ensembles de concepts permettent de représenter une organisation, sa sécurité de l'information, son contexte juridique et sa localisation ;
- Des ensembles de propriétés d'objet permettent d'illustrer les relations entre les concepts.

La mise en œuvre du modèle proposé dans ce chapitre doit permettre de valider notre proposition par :

- La réalisation d'une solution qui permet aux utilisateurs de gérer les informations nécessaires pour identifier, traiter et répondre à un DRR ;
- La réalisation d'une solution qui permet aux utilisateurs la consultation, recherche et création d'instances.

Chapitre 4 – Prototypage et expérimentation

Ce chapitre sera consacré à l'expérimentation et au prototypage de notre proposition de solution. Nous avons réalisé un prototype afin d'illustrer la faisabilité de l'implantation des travaux présentés dans cette thèse et de valider notre approche pour supporter la gestion du DRR.

Nous avons repris les éléments du troisième chapitre et testé notre prototype sur des sections de textes juridiques. Ce chapitre est organisé comme suit :

La première section présente les objectifs de notre prototype, les choix techniques et la description de la plateforme logicielle. La deuxième section présente le protocole de test et de retours d'expérience utilisateurs ainsi que les étapes nécessaires pour un passage à l'échelle et une mise en production du prototype dans des conditions réelles.

Section I. Gestion du prototypage

Afin de montrer la faisabilité des travaux présentés dans ce mémoire, nous avons réalisé un prototype reposant sur notre ontologie. Par le développement des fonctionnalités de recherche, consultation et instanciation, son objectif est de permettre aux utilisateurs d'accéder aux informations nécessaires dans le cadre de gestion du DRR.

1.1. Objectif et choix techniques

Nous souhaitons apporter le support nécessaire aux activités des individus impliqués dans la gestion du DRR qui sont :

- La direction des systèmes d'information ;
- La direction des risques et de la sécurité ;
- La direction juridique.

Le prototype a alors pour ambition d'être utilisé par différents experts pouvant être situés sur des sites ou pays distants. Les contraintes d'utilisations nous incombe de répondre aux besoins d'utilisations collaboratives et d'accès en ligne par la création d'une application web.

Pour répondre à cela, nous avons choisi le modèle MVC (Modèle – Vue – Contrôleur) qui est basé sur une architecture en trois couches entièrement autonomes : les données (le modèle), l'interface homme machine (la vue) et la logique de contrôle (le contrôleur).

Le prototype est principalement réalisé en Python[21]. Les raisons justifiant ce choix de langages sont multiples. Nous citons : logiciel libre (open source), large catalogue de bibliothèques disponibles, simplicité du code qui facilite le développement de prototype, présence d'API qui permettent la manipulation d'ontologies. Python est également le langage privilégié par notre entreprise terrain. Le choix d'un langage compatible avec l'environnement dans lequel il sera déployé et testé nous permet d'anticiper des besoins éventuels d'intégration que nous détaillons dans la deuxième section de ce chapitre.

[21] https://www.python.org/

Notre modèle, qui repose sur une ontologie, est construit en utilisant le logiciel Protégé Desktop[22] (Noy et al., 2000), un logiciel libre développé par l'université de Stanford supportant la création et l'édition d'ontologies.

L'ontologie développée à partir de Protégé correspond à notre proposition de solution présentée dans le chapitre trois. Celle-ci comprend l'ensemble des concepts, des propriétés d'objet et de données. Nous avons retenu le langage OWL pour la création de notre ontologie car celui-ci offre l'expressivité nécessaire tel que nous l'avons vu dans le deuxième chapitre de ce mémoire.

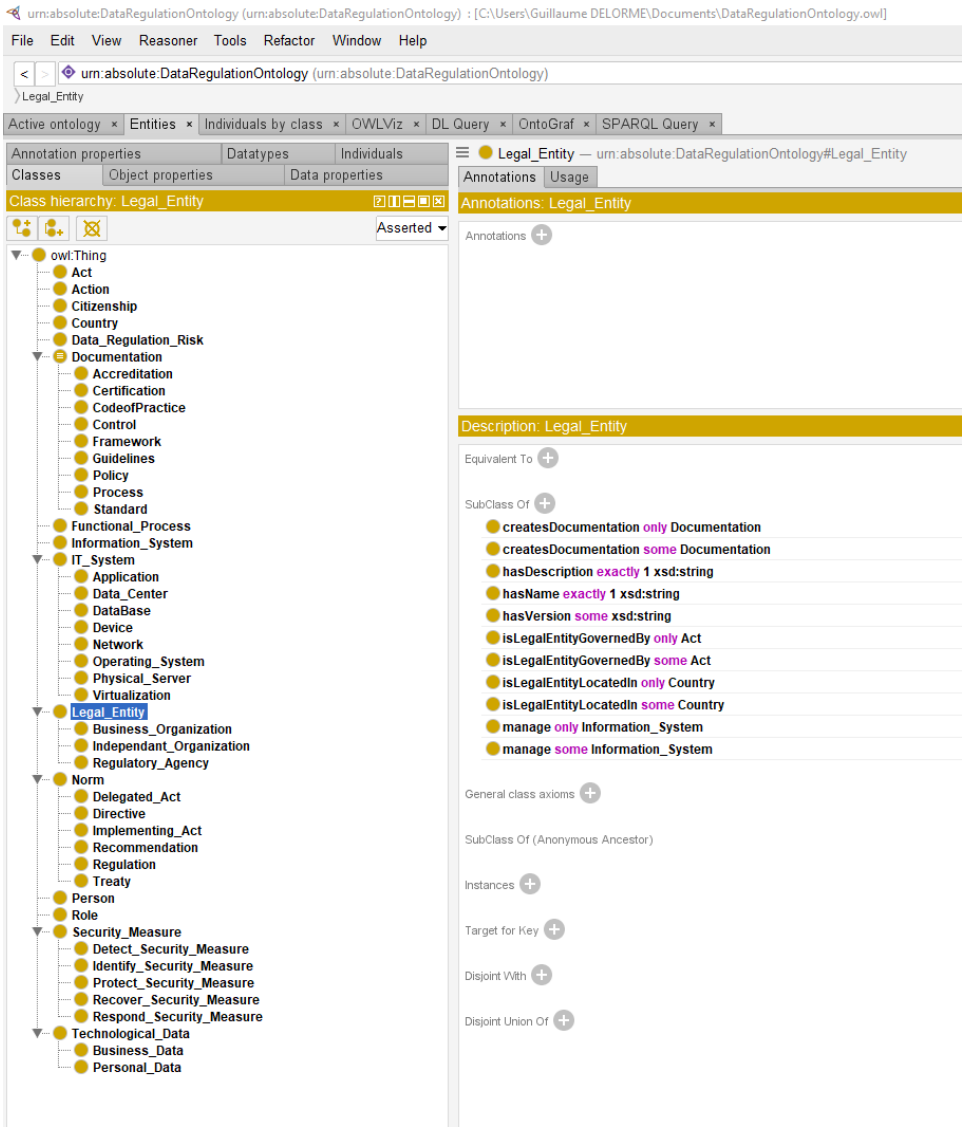


Figure 49 Interface de Protégé présentant les concepts

[22] https://protege.stanford.edu/

Afin de pouvoir manipuler l'ontologie développée sur Protégé par notre application développée en langage Python, nous utilisons le package Owlready2[23]. Développé par Lamy au laboratoire de recherche LIMICS de Paris (Lamy, 2017), celui-ci permet l'import, la création, la modification et la sauvegarde d'ontologies OWL2.0 en objet Python.

Owlready2 permet l'import d'ontologies sauvegardées en fichier .owl à partir d'une copie locale[24].

```
from owlready2 import *
class DataRegulationRepository:
    def __init__(self, app) -> None:
        file_path = app.config.get("OWL_FILE_PATH")
        self.onto = get_ontology(f'file://{file_path}').load()

    def get_all_classes(self):
        classes = self.onto.classes()
        return list(classes)

    def get_all_properties(self):
        properties = self.onto.object_properties()
        return list(properties)

    def get_subclass_of(self, subclass):
         data = self.onto.search(iri=f"*{subclass}*")
         print(list(data[0].get_properties()))
         classes = self.onto.search(subclass_of=data)
         return list(classes)
```

Figure 50 Exemple d'importation d'ontologies avec Owlready2

Une fois importés, les différents éléments (classes, instances et propriétés) de l'ontologie peuvent être manipulés (création, modification, sauvegarde, requête) comme s'ils

[23] https://owlready2.readthedocs.io/en/v0.37/intro.html

[24] Il est également possible d'importer des ontologies à partir de leur IRI ou d'un URL.

étaient des objets Python. L'ensembles des fonctionnalités d'Owlready2 sont documentées et librement accessibles sur le site du concepteur[25].

1.2. Présentation du prototype

L'architecture du système et les choix techniques étant définis, nous allons maintenant décrire les fonctionnalités de l'application que nous avons développé. Une fois sur l'application web, l'utilisateur peut accéder à deux interfaces correspondants aux fonctionnalités décrites dans le chapitre trois de ce mémoire :

- Recherche (assistée) d'instances existantes ;
- Consultation (non assistée) d'instances existantes ;

La page principale de l'application web permet uniquement d'accéder au différentes fonctionnalités par un bandeau. Nous avons créé les instances vues dans les exemples précédents afin de pouvoir illustrer les fonctionnalités.

Processus de recherche

Un utilisateur peut accéder au processus de recherche assistée en cliquant sur « Ask question ». Ceci le redirige vers une page dans laquelle il peut sélectionner les questions de recherche prédéfinies. Le prototype dispose d'une seule question de compétences : « Quels sont les éléments du système d'information qui sont impliqués dans le traitement d'une donnée réglementée ? ».

[25] https://owlready2.readthedocs.io/en/v0.37/index.html#

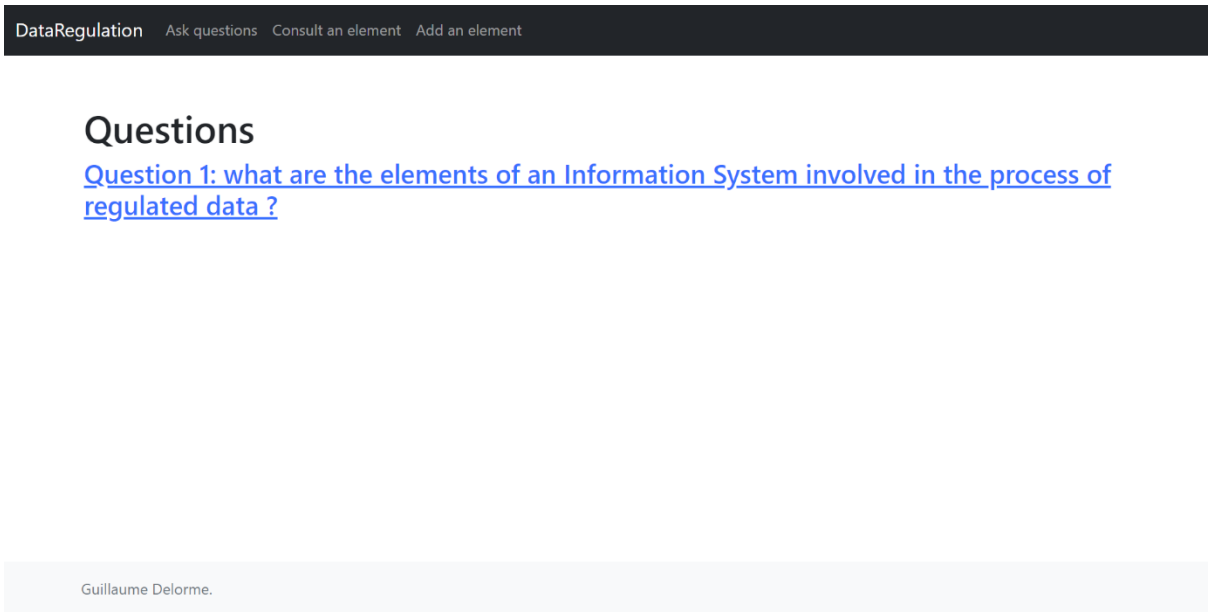


Figure 51 écran de sélection de la question de compétence

L'algorithme de sélection de requête SPARQL permet de sélectionner la requête correspondant à la question de compétence sélectionnée. Les deux premiers critères (« query1 » et « query2 ») permettent à l'algorithme de sélectionner la requête correspondant (« path »). Le troisième critère qui correspond à une instance sélectionnée par l'utilisateur permet de préciser l'élément nommé « lastcrit » dans chaque requête. L'extrait de code ci-dessous présente une partie de l'algorithme et deux requêtes.

La recherche assistée repose sur un principe de requêtes SPARQL prédéfinies. Les éléments sélectionnés par l'utilisateur permettent à l'algorithme de sélectionner la requête associée ainsi que de préciser quelle est la variable à utiliser pour effectuer la recherche.

```
def update_data(self):
        form = self

        query1 = form.data["query1"]
        query2 = form.data["query2"]
        query3 = form.data["query3"]

        query1Sf = self.repo.remove_prefix(query1)
        query2Sf = self.repo.remove_prefix(query2)

        if query1 is not None:
            form.query2.choices = [""] + self.repo.get_class_data_regulated()
            form.query2.render_kw={}

        if query2 is not None:
            form.query3.choices =  self.repo.get_instances_of_class(query2)
            form.query3.render_kw={}

        if query3 is not None:
            path = """""""
            if (query1Sf == PERSON and query2Sf == TECHNOLOGICAL_DATA):
                path = """
                    ?subject dro:isPersonInvolvedIn ?funcprocess.
                    ?itsystem dro:isITSystemInvolvedIn ?funcprocess.
                    ?itsystem dro:performs ?activite.
                    ?activite dro:process ?lastcrit.
                """

            if (query1Sf == IT_SYSTEM and query2Sf == TECHNOLOGICAL_DATA):
                path = """
                    ?subject dro:performs ?activite.
                    ?activite dro:process ?lastcrit.
                """
```

Figure 52 extrait de code de l'algorithme de sélection de requête SPARQL

Reprenant les étapes de la

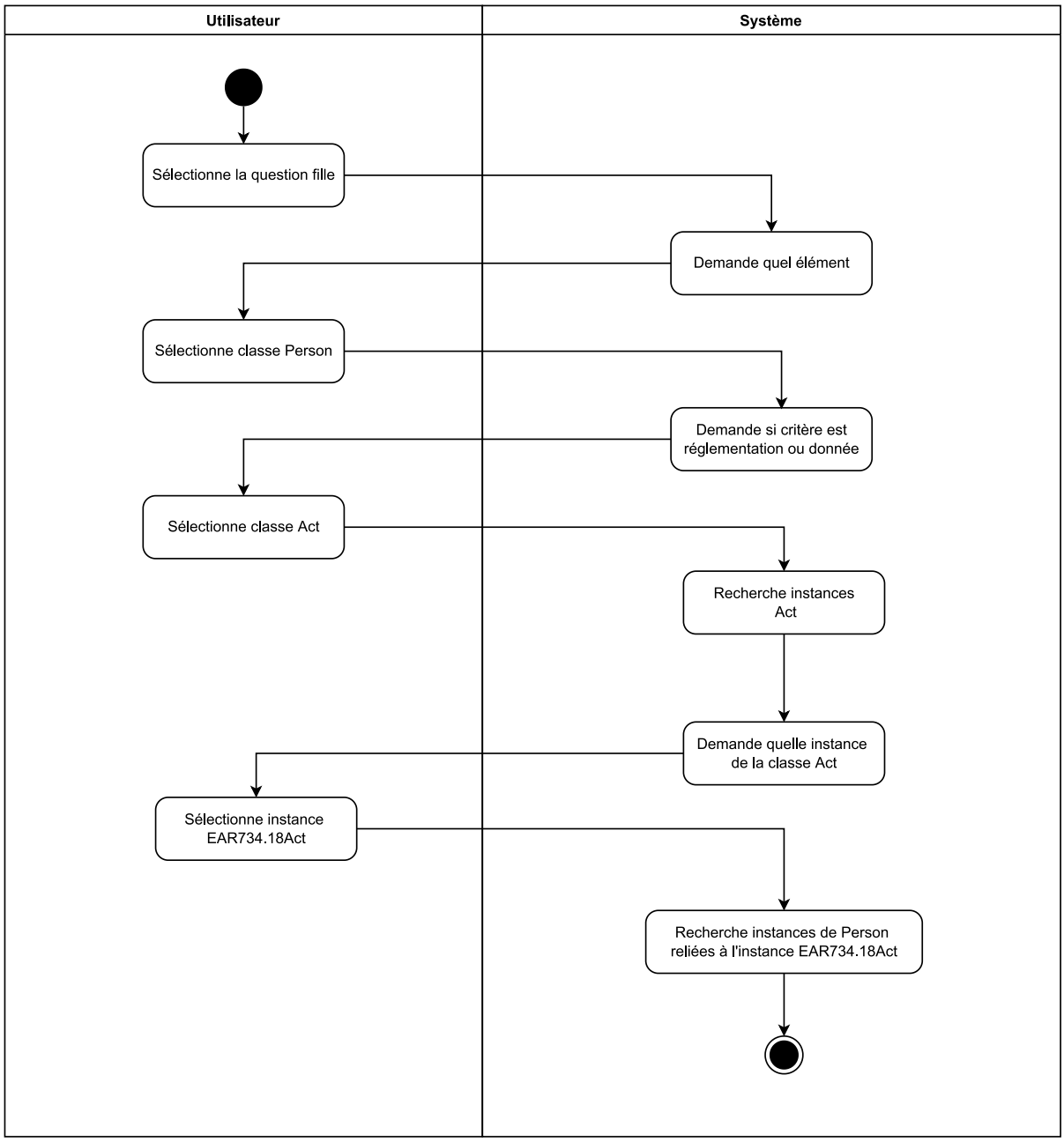


Figure 48 Exemple du processus de recherche, un utilisateur peut effectuer une recherche assistée pour identifier les individus impactés par la réglementation EAR et plus précisément l'Act 734.18 de l'EAR.

Une fois la classe « Person » sélectionnée par l'utilisateur comme premier élément de recherche, le système demande si le critère de recherche doit être une instance de la classe « Donnée » ou de la classe « Act ».

Une fois le critère précisé, l'outil recherche et présente l'ensemble des instances de « Norm ».

L’utilisateur sélectionne l’instance « EAR734.18Act ».

L’outil recherche alors l’ensemble des instances de la classe « Person » qui sont reliées à l’instance « EAR734.18Act ».

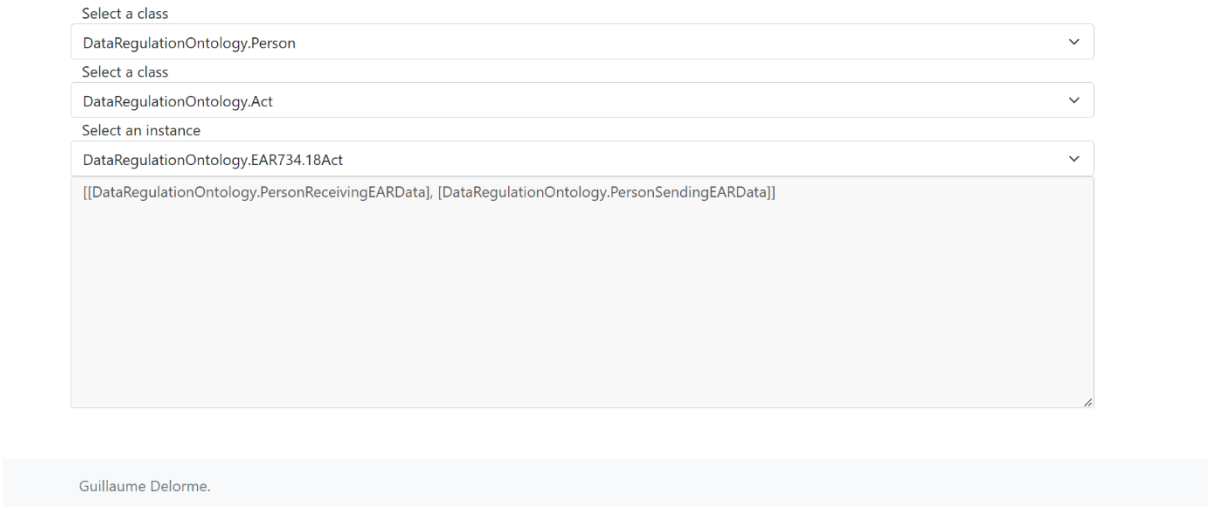


Figure 53 Interface de recherche

Cette recherche assistée repose alors sur la requête suivante :

```
PREFIX dro: <urn:absolute:DataRegulationOntology#>
        SELECT
        DISTINCT ?subject
        WHERE {
        {?subject dro:isPersonInvolvedIn ?funcprocess.
        ?itsystem dro:isITSystemInvolvedIn ?funcprocess.
        ?itsystem dro:performs ?activite.
        ?activite dro:process ?dataOne.
        ?dataOne dro:isBusinessDataGovernedBy ?lastcrit.
        ?lastcrit dro:hasName "EAR734.18Act"^^xsd:string }
UNION
        {?subject dro:isPersonInvolvedIn ?funcprocess.
        ?itsystem dro:isITSystemInvolvedIn ?funcprocess.
        ?itsystem dro:performs ?activite.
        ?activite dro:process ?dataOne.
        ?dataOne dro:isPersonalDataGovernedBy ?lastcrit.
        ?lastcrit dro:hasName "EAR734.18Act"^^xsd:string }}
```

Figure 54 exemple de requête SPARQL

Les résultats affichés sont :

```
[[DataRegulationOntology.PersonReceivingEARData], [DataRegulationOntology.PersonSendingEARData]]
```

Les résultats se lisent comme suit : [Ontologie utilisée.Instance].

Par exemple, le premier résultat correspond à l'instance « PersonReceivingEARData » de l'ontologie « DataRegulationOntology ».

Processus de consultation

En sélectionnant « Consult and element », un utilisateur peut également accéder à l'espace de consultation qui correspond à une recherche non assistée.

La recherche non assistée repose sur des fonctions permettant de requêter les instances d'une classe, les propriétés (propriétés de données) d'une instance et les relations (propriété d'objet) tel que présenté dans l'extrait de code ci-dessous.

```
def get_instances_of_class(self, targetClass):
    targetClass = self.remove_prefix(targetClass)
    classObj = self.onto[targetClass]
    return list(classObj.instances())

def get_instance_properties(self, targetInstance):
    targetInstance = self.remove_prefix(targetInstance)
    instance = self.onto[targetInstance]
    return list(instance.get_properties())

def get_relations_properties(self, targetProperties):
    targetProperties = self.remove_prefix(targetProperties)
    prop = self.onto[targetProperties]
    return list(prop.get_relations())
```

L'interface de consultation permet entre autre d'effectuer une recherche de classe parmi l'ensemble des classes et sous classes présentent dans l'ontologie.

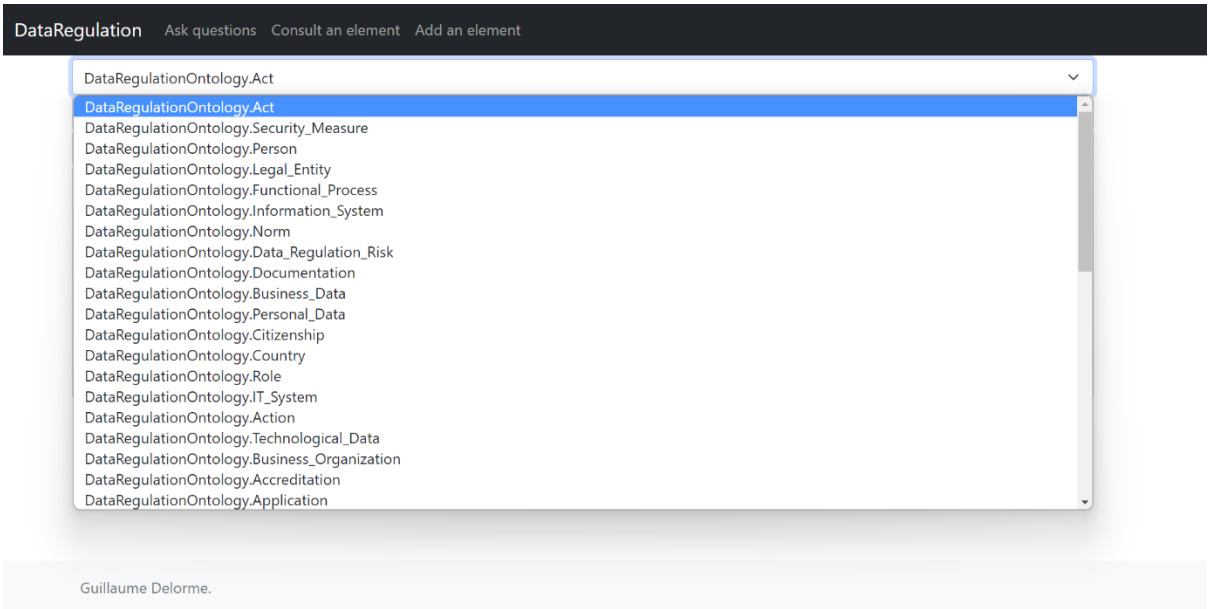


Figure 55 Interface de consultation

La Figure 56 Résultat d'une consultation montre une consultation simple permettant à un utilisateur de connaitre l'ensemble des mesures de sécurité qui sont appliquées aux données de classes « Business Data » et de l'instance « EARBusiness_Data ».

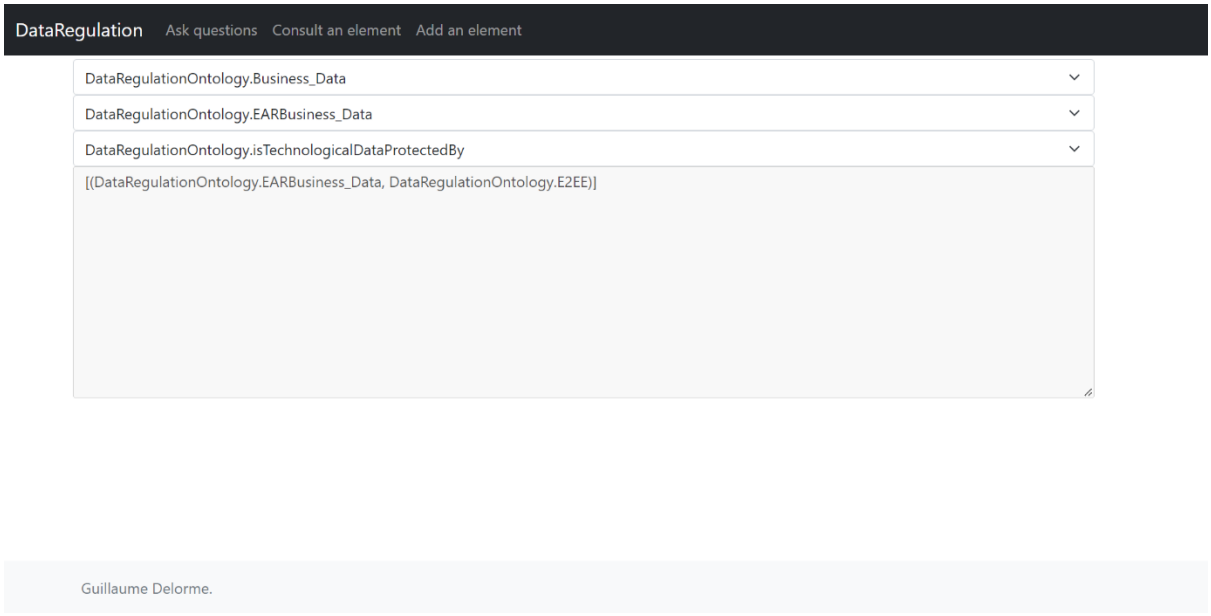


Figure 56 Résultat d'une consultation

La prochaine partie présente comment l'entreprise de référence peut suivre et mettre en place un protocole de validation du prototype et de mise en production de la solution.

Section II. Evaluation et mise en production

Afin de supporter la prise de décision de l'entreprise de référence quant à l'investissement dans le développement et le déploiement d'une solution pérenne, nous avons réalisé un protocole de validation du prototype et de mise en production de la solution.

2.1. Phase de validation du prototype

Nos travaux, et par extension notre prototype, s'inscrivent dans la volonté de l'entreprise de référence de mettre en œuvre les moyens nécessaires pour assurer sa conformité réglementaire. Ceux-ci faisant partie d'un plus large programme de conformité, nous devons nous assurer que notre proposition de solution non seulement réponde de manière efficace à un besoin métier existant mais aussi qu'elle soit acceptée par les utilisateurs.

La réponse à un besoin métier est l'origine et la motivation de ces travaux. Nous avons étudié son contexte dans le premier chapitre. Nous avons également présenté notre proposition de solution dans le troisième chapitre.

Nous devons maintenant nous assurer que notre proposition réponde aux besoins des utilisateurs, de leur adhésion ainsi que celle du management tout en limitant les risques et incertitudes liés à son déploiement. Afin de supporter la prise de décision de l'entreprise de référence quant à l'investissement dans le développement et le déploiement d'une solution pérenne, il est nécessaire de recueillir les retours d'expérience des utilisateurs et testeurs du prototype.

2.1.1. Préparation des retours d'expérience

La première étape de préparation de ce protocole de test et de retour d'expérience est la détermination des objectifs et de son périmètre. Pour cela, les retours d'expérience se focalisent sur les deux critères suivants:

- Quelle est la perception du service rendu par l'utilisation du prototype ?
- Quelle est la perception de la facilité d'utilisation du prototype ?

Ces critères permettent d'assurer que les utilisateurs appréhendent le prototype tel que nous l'avons conçu (service rendu) et garantissent une bonne expérience utilisateur (facilité d'utilisation).

La deuxième étape consiste en l'identification et l'implication des acteurs clés. Afin de disposer d'une vision exhaustive, nous avons besoin d'un panel représentatif d'utilisateurs. Nous avons précédemment identifier quels étaient les utilisateurs visés (la direction des systèmes d'information, la direction des risques et de la sécurité et la direction juridique). Afin de pouvoir présenter une évaluation pertinente de notre prototype, les testeurs devrons représentés à part égale ces différentes fonctions.

Enfin, l'identification des acteurs clés comprend également le comité de validation du projet. Celui-ci aura la charge de déterminer les Facteur Clef de Succès (FCS) du prototype et des tests. Si les FCS sont atteints, alors le comité pourra valider le projet de mise en production et de passage à l'échelle.

2.1.2. Mise en œuvre des retours d'expérience

Les retours d'expérience seront mis en œuvre via deux protocoles :

- Le protocole System Usabiity Scale (SUS) (Brooke, 1996) ;
- Le protocole User Experience Questionnaire (UEQ) (Schrepp et al., 2014).

Ces deux protocoles nécessitent que les utilisateurs aient le temps de tester l'ensemble des fonctionnalités du prototype. Pour cela, un calendrier de test doit être établi, validé et respecté par les différents testeurs. Les deux protocoles seront employés à la fin du calendrier de test.

Afin de d'évaluer la qualité du service rendu par notre proposition de solution, nous reposons sur le protocole System Usabiity Scale (SUS) développé par Brooke (Brooke, 1996). Ce questionnaire repose sur dix questions dont l'objectif est de déterminer le niveau de satisfaction des utilisateurs d'un service.

Celui-ci est une évaluation « à chaud » dont le score est compris entre 0 et 100. Utilisant une échelle de Likert (Likert, 1932), un testeur à le choix entre cinq réponses pour chaque question allant de « Pas du tout d'accord » à « Tout à fait d'accord » comme illustré dans la Figure 57 System Usability Scale.

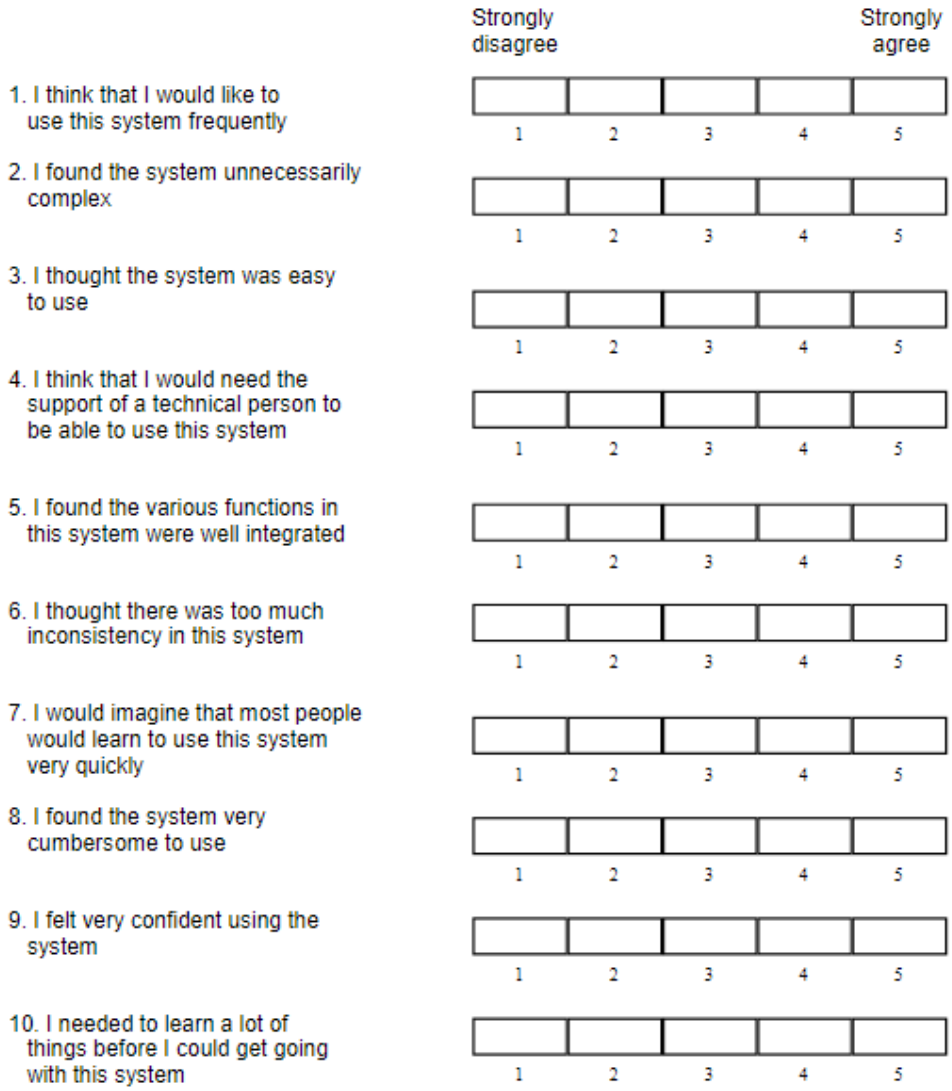


Figure 57 System Usability Scale

Selon le concepteur du SUS, le questionnaire doit être utilisé après une utilisation de la solution mais avant toutes conversations ou débriefing avec les testeurs.

Dans un second temps, nous reposons sur le protocole User Experience Questionnaire (UEQ) de Schrepp et al. (Schrepp et al., 2014) pour évaluer la facilité d'utilisation et l'expérience utilisateur. Disponible en libre accès, le questionnaire est également complété d'un outil d'analyse des retours d'expérience[26].

[26] https://www.ueq-online.org/

	1	2	3	4	5	6	7		
Agaçant	○	○	○	○	○	○	○	Agréable	1
Incompréhensible	○	○	○	○	○	○	○	Compréhensible	2
Moderne	○	○	○	○	○	○	○	Sans fantaisie	3
Appropriation simple	○	○	○	○	○	○	○	Appropriation compliquée	4
Apporte de la valeur	○	○	○	○	○	○	○	Peu de valeur ajoutée	5
Ennuyeux	○	○	○	○	○	○	○	Captivant	6
Inintéressant	○	○	○	○	○	○	○	Intéressant	7
Imprévisible	○	○	○	○	○	○	○	Prévisible	8
Rapide	○	○	○	○	○	○	○	Lent	9
Original	○	○	○	○	○	○	○	Conventionnel	10
Rigide	○	○	○	○	○	○	○	Facilitant	11
Bien	○	○	○	○	○	○	○	Médiocre	12
Compliqué	○	○	○	○	○	○	○	Simple	13
Repoussant	○	○	○	○	○	○	○	Attractif	14
Habituel	○	○	○	○	○	○	○	Avant-gardiste	15
Désagréable	○	○	○	○	○	○	○	Agréable	16
Sécurisant	○	○	○	○	○	○	○	Insécurisant	17
Stimulant	○	○	○	○	○	○	○	Soporifique	18
Répond aux attentes	○	○	○	○	○	○	○	Ne répond pas aux attentes	19
Inefficace	○	○	○	○	○	○	○	Efficace	20
Clair	○	○	○	○	○	○	○	Déroutant	21
Non pragmatique	○	○	○	○	○	○	○	Pragmatique	22
Sobre	○	○	○	○	○	○	○	Surchargé	23
Attrayant	○	○	○	○	○	○	○	Rébarbatif	24
Sympathique	○	○	○	○	○	○	○	Inamical	25
Conservateur	○	○	○	○	○	○	○	Innovant	26

Figure 58 User Experience Questionnaire

Utilisant également une échelle de Likert (Likert, 1932), ce questionnaire est constitué de 26 questions qui peuvent être répondues pendant ou après la phase de test. Celui-ci permet entre autres d'évaluer l'attractivité, la perception, l'efficacité, l'autonomie, la stimulation et l'originalité de notre solution.

Une fois les retours d'expérience analysés et présentés au comité de validation de l'entreprise de référence, celui-ci peut décider de la mise en production du prototype et son passage à l'échelle.

2.2. Phase de mise en production

Le passage à l'échelle et la mise en production comprend des points d'attention qui ont volontairement été écartés lors des phases de test et de prototypage.

Premièrement, les contraintes techniques de développement et d'intégration devront être étudiées. L'environnement et le système d'information de l'entreprise de référence reposent principalement sur des technologies Cloud. Le prototype devra alors est déployé sur un environnement de « Plateform as a Service » tel que Google Cloud Platfom[27]. Ceci nécessite de prévoir le développement de solutions techniques pour la gestion des bases de données externalisées, une gestion des identités et des accès, d'administration et gestion des performances etc.

Il est donc nécessaire de prévoir et planifier un projet de redéveloppement pour répondre aux contraintes techniques de l'entreprise de référence. La planification de ce projet devra être anticipée et valider avec les équipes responsables.

Deuxièmement, la gestion du déploiement auprès des utilisateurs. La gestion de mises en production consiste à mettre en service une unité de mise en production alors que la gestion des déploiements a pour objectif de transférer dans les environnements de production, tous les éléments qui composent un changement (matériels, logiciels, procédures, etc. (BAUD, 2020)). Selon l'objectif, le périmètre et les contraintes du déploiement, plusieurs types existent. Par exemple un déploiement « Big Bang » consiste en une seule opération vers tous les utilisateurs au même instant alors qu'un déploiement par phases s'effectue selon un plan de déploiement par phases successives.

Bien que notre prototype s'inscrive dans la volonté de l'entreprise de référence de mettre en œuvre les moyens nécessaires pour assurer sa conformité réglementaire, il n'est pas réaliste d'espérer que l'ensemble des utilisateurs soient disponibles dès la première phase de déploiement.

[27] Google Cloud Platform : https://console.cloud.google.com/getting-started?hl=fr&pli=1.

Conclusion

L’objectif de ce chapitre était double. Dans un premier temps, celui-ci permettait de présenter nos choix techniques pour l’expérimentation et le prototypage de notre proposition de solution. Nous avons réalisé un prototype principalement développé en Python reposant sur une ontologie développée grâce au logiciel Protégé.

Afin d’illustrer la faisabilité de l’implantation des travaux présentés dans cette thèse et de valider notre approche pour supporter la gestion du DRR, nous avons également détaillé et présenté les principales fonctionnalités de notre prototype.

Dans un second temps ce chapitre présente le protocole de test et de retours d’expérience utilisateurs ainsi que les étapes nécessaires pour un passage à l’échelle et une mise en production du prototype dans des conditions réelles.

Conclusion générale

Bilan

L'objectif de ces travaux était de proposer une solution de représentation du risque règlementaire multidisciplinaire afin de permettre sa gestion. Nous avons choisi d'aborder le problème de la gestion de la conformité des données réglementées en proposant un modèle permettant la représentation des modalités juridiques, des évolutions réglementaires ainsi que des impacts métiers et opérationnels. Celui-ci repose sur une ontologie multi-domaines.

L'originalité de notre recherche est de proposer un modèle permettant la représentation de manière intelligible et précise des contraintes et obligations réglementaires qu'une organisation doit respecter pour assurer sa conformité.

Ce choix a été guidé par l'état de l'art :

- Des méthodologies de gestion du risque notamment les propositions permettant la gestion de risque transverse, multidisciplinaire et multi-scénario;
- Des ontologies du domaine juridique ;
- Des ontologies du domaine des technologies de l'information et de la sécurité.

Du point de vue juridique, nous sommes partis des recherches de Van Kralingen (van Kralingen, 1997), Palmirani et al. (Palmirani et al., 2018) et Hoekstra et al. (Hoekstra et al., 2009) qui offrent des propositions de représentation du corpus et des modalités juridiques, du contexte des organisations et des individus ainsi que des jeux de documents d'une organisation.

Du point de vue IT et sécurité, nous sommes partis des recherches de Ekelhart et al. (Ekelhart et al., 2006) et de Fenz (Fenz & Ekelhart, 2009) qui permettent la représentation des mesures de sécurité, un système d'information et ses composants.

L'incrément que nous apportons par rapport à ces travaux est la création d'une ontologie composée d'un ensemble de 15 concepts utilisés pour représenter notre domaine, le risque lié à la réglementation des données ou Data Regulation Risk (DRR) (Delorme et al., 2021). Ces concepts sont répartis en quatre sous domaines :

- Le sous domaine de l'entreprise ;

- Le sous domaine de la sécurité ;
- Le sous domaine réglementaire ;
- Le sous domaine de la localisation.

Nous avons réalisé un prototype afin de valider la faisabilité de cette proposition. Celui-ci regroupe l'ontologie ainsi que les fonctionnalités principales de recherche, de consultation et d'instanciation qui permettent d'aider un utilisateur dans sa gestion du risque réglementaire.

Perspectives

Notre solution et prototype nous permettent de valider la faisabilité de cette proposition cependant le passage à l'échelle, la mise en production et l'adoption par les utilisateurs sont des challenges que nous devons anticiper. Par conséquent des possibilités de développement peuvent être envisagées.

Intégration avec un système d'information existant

L'entreprise de référence dispose de plusieurs dizaines de millier d'éléments composant son infrastructure IT. Il est alors impératif de prévoir une solution automatisée pour alimenter notre base de données. Le prototype repose actuellement sur une alimentation manuelle par un utilisateur qui n'est pas adaptée à une mise en production et un passage à l'échelle.

Il est également nécessaire de se poser la question de la mise à jour des instances présentes dans notre ontologie afin d'assurer la pérennisation de l'outil et la pertinence des résultats. Une infrastructure IT évolue dans le temps et les modifications associées peuvent impacter la conformité d'une organisation. Afin de pouvoir soutenir la prise de décision et répondre aux objectifs initiaux fixés, la mise à jour doit être prévue.

Gestion des utilisateurs

Il n'y a à ce jour aucune gestion des utilisateurs prévues par le prototype. Il en est de même pour la gestion des accès à la plateforme ou au sein de la plateforme. La réflexion quant à la criticité des données présentes dans l'outil doit être prise en compte en amont d'un passage à l'échelle.

Questions de compétences

Nous avons avancé dans la première section du Chapitre 3 qu'il n'était pas réaliste de prévoir, dès l'étape de conception du modèle, l'ensemble des questions auxquelles les différents utilisateurs vont chercher à obtenir les réponses. En effet, celles-ci vont dépendre du contexte de l'organisation tels que le champ législatif applicable, son système d'information mais également des connaissances de l'utilisateur ou encore de son activité professionnelle.

La création, le développement et l'intégration de nouvelles questions de compétences doit se faire conjointement avec les utilisateurs finaux afin d'en assurer la pertinence.

Intégration avec d'autres modèles existants

Nous avons limité notre modèle et notre prototype à notre domaine d'étude, à savoir :

- Comment représenter un risque règlementaire multidisciplinaire ?
- Comment permettre une meilleure gestion d'un tel risque ?

Dans une optique gestion transverse du risque et d'une approche multidisciplinaire et pluri-risques, il est nécessaire de se poser la question de l'intégration de modèles externes. Nous citons : modèle de gestion de risques financiers, de gestion de risques opérationnels, de risques et menaces cyber, etc.

Veille réglementaire

Notre dernier axe de développement concerne l'automatisation de la veille réglementaire, de la mise à jour et de la validation des éléments constituants notre base de connaissance réglementaire. Tout comme la mise à jour des composants informatiques de l'entreprise de référence permet de garantir la pertinence dans le temps de notre proposition de solution, la mise à jour et l'actualisation des modalités juridiques est primordial.

Table des matières

Appendix

Lien	SynSet référence	Définition SynSet	Commentaire
Govern	202511551	bring into conformity with rules or principles or usage; impose regulations	N/A
Have	202203362	have or possess, either in a concrete or an abstract sense	N/A
Located	302126430	situated in a particular spot or position	N/A
Belong	202756359	be a member, adherent, inhabitant, etc.	Besoin d'ajouter la précision que "membre de" signifie "employé de" dans notre modèle.
Involve	202677567	engage as a participant	N/A
Protect	201128193	shield from danger, injury, destruction, or damage	Besoin d'ajouter des précisions quant aux dommages dans le cadre du DRR.
Define	200947077	decide upon or fix definitely	N/A
Manage	202436349	be in charge of, act on, or dispose of	N/A
Own	202204692	have ownership or possession of	N/A
Compose	201626138	put together out of existing material	N/A
Impact	200137313	Have an effect upon	N/A
Create	201617192	make or cause to be or to become	N/A
Process	101023820	a particular course of action intended to achieve a result	Besoin d'ajouter des précisions avec la définition du RGPD pour être adapté au modèle.
Use	100947128	the act of using	N/A
Perform	201712704	carry out or perform an action	N/A

Appendix 1 Liste des relations

Concept	Synset	WordNet synsets definition	Commentaire
Legal_Entity	100001740	That which is perceived or known or inferred to have its own distinct existence (living or nonliving).	"Distinct existence" permet de considérer les personnes physiques ou morales. Besoin d'ajouter la précision la diversité des personnes morales afin d'être centré sur le sujet DRR.
Information_System	103164344	System consisting of the network of all communication channels used within an organization.	Besoin d'élargir la définition pour intégrer l'implication des individus et des intéractions (processus).
Functional_Process	101023820	A particular course of action intended to achieve a result	Besoin de préciser la nécessité d'avoir des éléments entrants pour obtenir un résultat.
IT_System	104377057	Instrumentality that combines interrelated interacting artifacts designed to work as a coherent entity	N/A
Action	100030358	Something that people do or cause to happen.	
Person	100007846	A human being	N/A
Role	100720565	The actions and activities assigned to or required or expected of a person or group	
Technological_Data	105816622	An item of factual information derived from measurement or research.	Besoin de préciser que la donnée est sous format numérique pour répondre à notre sujet.
Security_Measure	100823316	Measures taken as a precaution against theft or espionage or sabotage etc.	Besoin de préciser que les mesures de sécurité ont pour but de protéger la confidentialité, l'intégrité et la disponibilité de ses informations et de répondre à un ensemble d'exigences de sécurité définies
Documentation	106588326	Program listings or technical manuals describing the operation and use of programs.	Besoin de préciser que la documentation définissent, regroupent et garantissent la pertinence et les conditions d'applicabilité des Security Measures et ne se limitent pas à la description des opérations et utilisations de programmes.
Norm	106532330	Legal document setting forth rules governing a particular kind of activity	Besoin de préciser que nous limitons aux lois et textes juridiques régissant le traitement de données et/ou la gouvernance et les processus des technologies de l'information et communication et/ou les services et technologies de l'information.
Act	100805034	The act of controlling or directing according to rule	Besoin de préciser que l'action qui occure doit être régie par une loi.
Data_Regulation_Risk	114541852	A source of danger; a possibility of incurring loss or misfortune	N/A
Country	108544813	The territory occupied by a nation	N/A
Citizenship	113953467	The status of a citizen with rights and duties.	N/A

Appendix 2 Liste des classes de l'ontologie

Legal_Entity	hasName	IRI	hasVersion	hasDescription	isRefOrganization	
Information_System	hasName	IRI	hasVersion	hasDescription		
Functional_Process	hasName	IRI	hasVersion	hasDescription		
IT_System	hasName	IRI	hasVersion	hasDescription	isOwnedInternally	isManagedInternally
Action	hasName	IRI	hasVersion	hasDescription	hasActionType	
Person	hasName	IRI	hasVersion	hasDescription	isEmployeeInternal	
Role	hasName	IRI	hasVersion	hasDescription		
Technological_Data	hasName	IRI	hasVersion	hasDescription		
Security_Measure	hasName	IRI	hasVersion	hasDescription	hasSecurityMeasureNature	hasSecurityMeasurePurpose
Documentation	hasName	IRI	hasVersion	hasDescription	isDocumentationInternal	
Norm	hasName	IRI	hasVersion	hasDescription	hasHasNormSource	
Act	hasName	IRI	hasVersion	hasDescription	hasActType	
Data_Regulation_Risk	hasName	IRI	hasVersion	hasDescription		
Country	hasName	IRI	hasVersion	hasDescription		
Citizenship	hasName	IRI	hasVersion	hasDescription		

Appendix 3 Liste des propriétés de données par classe

Bibliographie

Abdullah, N. S., Indulska, M., & Sadiq, S. (2016). Compliance management ontology – a shared conceptualization for research and practice in compliance management. *Information Systems Frontiers*, *18*(5), 995–1020. https://doi.org/10.1007/s10796-016-9631-4

Agence Nationale de la Sécurité des Systèmes d'Information. (2010). *EBIOS MÉTHODE DE GESTION DES RISQUES*.

Agence Nationale de la Sécurité des Systèmes d'Information. (2018). *EBIOS Risk Manager*.

Agrafiotis, I., Nurse, J. R. C., Goldsmith, M., Creese, S., & Upton, D. (2018). A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate. *Journal of Cybersecurity*, *4*(1). https://doi.org/10.1093/cybsec/tyy006

Al-Ahmad, W., & Bassil Mohammad. (2012). Can a single security framework address information security risks adequately? *International Journal of Digital Information and Wireless Communications*, *2*(3), 222–230.

Alavi, M., & Leidner, D. E. (2007). Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, *25*(1), 107–136.

Amoroso, E. (2010). *Cyber Attacks: Protecting National Infrastructure* (P. Chester, Ed.; 1st ed., Vol. 1). Butterworth-Heinemann.

Anthony Cox, L. (2008). What's wrong with risk matrices? *Risk Analysis*, *28*(2), 497–512.

Asenov, E. (2015). Characteristics of Compliance Risk in Banking. *Economic Alternatives*, *4*(1), 20–28.

Aven, T. (2012). The risk concept—historical and recent development trends. *Reliability Engineering & System Safety*, *99*, 33–44. https://doi.org/10.1016/j.ress.2011.11.006

Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, *253*(1), 1–13. https://doi.org/10.1016/j.ejor.2015.12.023

Aven, T., & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, *12*(1), 1–11. https://doi.org/10.1080/13669870802488883

Avizienis, A., Laprie, J.-C., Randell, B., & Landwehr, C. (2004). Basic Concepts and Taxonomy of Dependable and Secure Computing. *IEEE Transactions on Dependable and Secure Computing*, *1*(1), 11–33.

Baader, F., & Hollunder, B. (1991). A terminological knowledge representation system with complete inference algorithms. *International Workshop on Processing Declarative Knowledge*, 67–86.

Baader, F., Horrocks, I., & Sattler, U. (2005). Description Logics for the Semantic Web. In Springer (Ed.), *Mechanizing mathematical reasoning* (pp. 228–248). Heidelberg.

Bagchi, K., & Udo, G. (2003). An Analysis of the Growth of Computer and Internet Security Breaches. *Communications of the Association for Information Systems*, *12*(1). https://doi.org/10.17705/1CAIS.01246

Baggini, J. (2002). David Hume: An Enquiry concerning Human Understanding (1748). In J. Baggini (Ed.), *Philosophy: Key Texts* (pp. 61–84). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-4039-1370-8_4

Baker, M. J. (2000). Writing a literature review. *The Marketing Review*, *1*(2), 219–247.

Barker, J. P. (2008). *Brokering under the international traffic in arms regulations* (pp. 181–197).

Barlette, Y., & Fomin, V. v. (2009). The adoption of information security management standards: A literature review. In *Cyber Security and Global Information Assurance: Threat Analysis and Response Solutions* (pp. 119–140). IGI Global. https://doi.org/10.4018/978-1-60566-326-5.ch006

Barth, J. R., Caprio Jr, & Levine, R. (2004). Bank regulation and supervision: what works best? *Journal of Financial Intermediation*, *13*(2), 205–248.

BAUD, J.-L. (2020). *ITIL® 4 - Comprendre la démarche et adopter les bonnes pratiques* (Vol. 2). Editions ENI.

Bench-Capon, T., Visser, P., & Jones, D. (1998). *Methodologies For Ontology Development*. *33*. https://www.researchgate.net/publication/2760967

Bier, V. M., & Lin, S. W. (2013). Should the Model for Risk-Informed Regulation be Game Theory Rather than Decision Theory? *Risk Analysis*. https://doi.org/10.1111/j.1539-6924.2012.01866.x

Bizer, C., & Heath, T. (2011). Special Issue on Linked Data. *International Journal on Semantic Web and Information Systems*, 205–227. http://linkeddata.org/docs/ijswis-special-issue

Bradley, N. (2002). *The XML companion*. Addison-Wesley Professional.

Broderick, J. S. (2006). ISMS, security standards and security regulations. *Information Security Technical Report*, *11*(1), 26–31. https://doi.org/10.1016/j.istr.2005.12.001

Brooke, J. (1996). SUS -- a quick and dirty usability scale. *Usability Evaluation in Industry*, *189*(194), 4–7.

Camp, L. J. (2002). Designing for Trust. *Workshop on Deception, Fraud and Trust in Agent Societies*, 15–29.

Canongia, C., & Mandarino, R. (2012). Cybersecurity: The New Challenge of the Information Society. In *Handbook of Research on Business Social Networking*. IGI Global. https://doi.org/10.4018/978-1-61350-168-9.ch009

Caralli, R., Stevens, J., Young, L., & Wilson, W. (2007). *Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process*. http://www.sei.cmu.edu/library/abstracts/reports/07tr012.cfm

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K. (2016). A review of cyber security risk assessment methods for SCADA systems. *Computers and Security*, *56*, 1–27. https://doi.org/10.1016/j.cose.2015.09.009

Churchman, C. W. (1971). *The design of inquiring systems basic concepts of systems and organization*.

Ciolan, I. M. (2014). Defining cybersecurity as the security issue of the twenty first century. A constructivist approach. *Revista de Administratie Publica Si Politici Sociale*, *12*(1).

Club de la sécurité de l'information français. (2007). *Mehari 2007 - Risk Analysis Guide*.

Collard, G., Ducroquet, S., Disson, E., & Talens, G. (2017). A definition of Information Security Classification in cybersecurity context. *Proceedings - International Conference on Research Challenges in Information Science*, 77–82. https://doi.org/10.1109/RCIS.2017.7956520

Collectif Dalloz. (2022). *Code de la cybersécurité 2022* (Dalloz, Ed.; Codes Dalloz, Vol. 1).

Collegio, D. G., & Alberto, C. (2001). Trust in Signs. In K. S. Cook (Ed.), *Trust in Society* (Russel Sage, Vol. 2, pp. 148–184). https://www.researchgate.net/publication/209409853

Collins, S., Genova, F., Harrower, N., Hodson, S., Jones, S., Laaksonen, L., Mietchen, D., Petrauskaitė, R., & Wittenburg, P. (2018). *Turning FAIR into reality: Final report and action plan from the European Commission expert group on FAIR data*.

Commission Nationale de l'Informatique et des Libertés. (2018). *Analyse d'impact relative à la protection des données*.

Craigen, D., Diakun-Thibault, N., & Purse, R. (2014). Technology Innovation Management Review Defining Cybersecurity. *Technology Innovation Management Review*, *4*(10). www.timreview.ca

Crook, J. A. (2009). National insecurity: ITAR and the technological impairment of US national space policy. *Technological Impairment of U.S. National Space Policy, 74 J. Air L. & Com*, *74*(1), 505–526. http://digitalrepository.smu.edu.https://scholar.smu.edu/jalc/vol74/iss2/7

Crozier, M., & Friedberg, E. (1977). *L'acteur et le système* (Editions du Seuil).

de Chalendar, G., Grau, B., & Ferret, O. (2000). Généralisation de graphes conceptuels. *Actes Du Congrès Reconnaissance Des Formes et Intelligence Artificielle (RFIA)*, 359–368.

Delorme, G., Talens, G., & Disson, E. (2022a). An Ontology For Data Regulation. *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*.

Delorme, G., Talens, G., & Disson, E. (2022b). Data Regulation Ontology. *Proceedings of the 34th International Conference on Software Engineering and Knowledge Engineering*, 503–507. https://doi.org/10.18293/SEKE2022-024

Delorme, G., Talens, G., Disson, E., Collard, G., & Gaget, E. (2021). On the Definition of Data Regulation Risk. In H. Hacid, F. Outay, H. Paik, A. Alloum, M. Petrocchi, M. R. Bouadjenek, A. Beheshti, X. Liu, & A. Maaradji (Eds.), *Service-Oriented Computing – ICSOC 2020 Workshops* (pp. 433–443). Springer International Publishing.

Douglas, M. (1990). Risk as a Forensic Resource. *Daedalus*, *119*(4), 1–16. http://www.jstor.org/stable/20025335

Ekelhart, A., Fenz, S., Klemen, M. D., & Weippl, E. R. (2006). Security Ontology: Simulating Threats to Corporate Assets. *International Conference on Information Systems Security*, 249–259. https://doi.org/10.1007/11961635_17

Eloff, M. M., & von Solms, S. H. (2000). Information security management: a hierarchical framework for various approaches. *Computers & Security* , *19*(3), 243–256.

EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1. (n.d.).

Fahey, L., & Prusak, L. (1998). The Eleven Deadliest Sins of Knowledge Management. *California Management Review*, *40*(3), 265–276.

Fensel, D., Horrocks, I., Decker, S., Harmelen, F., Erdmann, M., & Klein, M. (2000). OIL in a Nutshell. *International Conference on Knowledge Engineering and Knowledge Management*, 1–16.

Fenz, S., & Ekelhart, A. (2009). Formalizing Information Security Knowledge. *Proceedings of the 4th International Symposium on Information, Computer, and Communications Security*, 183–194.

Ferndndez, M., G6mez-P~rez, A., & Juristo, N. (1997). *METHONTOLOGY: From Ontological Art Towards Ontological Engineering*. www.aaai.org

Fischer, D. H. (1998). From Thesauri towards Ontologies? *Advances in Knowledge Organization*, *6*, 18–30.

Flowerday, S., & von Solms, R. (2006). Trust: An Element of Information Security. *IFIP International Information Security Conference*, 87–98.

Gambetta, D. (2000). Can We Trust Trust? *Trust: Making and Breaking Cooperative Relations*, *13*(1), 213–237. http://www.sociology.ox.ac.uk/papers/

Gaudemet, A. (2019). Qu'est-ce que la compliance ? *Commentaire*, *165*(1), 109. https://doi.org/10.3917/comm.165.0109

Gefen, D., Rao, V. S., & Tractinsky, N. (2003). The Conceptualization of Trust, Risk and Their Relationship in Electronic Commerce: The Need for Clarifications. *36th Hawaii International Conference on System Sciences (HICSS)*, 192–202. www.eweek.com/article/0,3658,

Gerber, M., & von Solms, R. (2005). Management of risk in the information age. *Computers and Security*, *24*(1), 16–30. https://doi.org/10.1016/j.cose.2004.11.002

Ghernaouti, S. (2016). *Cybersécurité sécurité informatique et réseaux* (Dunod, Ed.; 6th ed.).

Gros, F. (2013). Coopérer contre soi-même. In *Deals de justice* (pp. 173–187). Presses Universitaires de France. https://doi.org/10.3917/puf.gara.2013.01.0173

Gruber, T. R. (1993). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, *43*(5–6), 907–928.

Grüninger, M., & Fox, M. S. (1995). *Methodology for the Design and Evaluation of Ontologies*.

Guarino, N., & Gangemi, A. (1997). Understanding, building and using ontologies. *International Journal of Human-Computer Studies*, *46*(2–3), 293–310.

Guarino, N., & Welty, C. (2002). Evaluating ontological decisions with ontoclean. *Communications of the ACM*, *45*(2), 61–65. https://doi.org/10.1145/503124.503150

Guerra, G. A., Zizzo, D. J., Dutton, W. H., & Peltu, M. (2003). Economics of Trust in the Information Economy: Issues of Identity, Privacy and Security. In *SSRN*. https://doi.org/10.2139/ssrn.723201

Guizzardi, G. (2020). Ontology, ontologies and the "i" of fair. *Data Intelligence*, *2*(1–2), 181–191. https://doi.org/10.1162/dint_a_00040

Haimes, Y. Y. (2006). On the Definition of Vulnerabilities in Measuring Risks to Infrastructures. *Risk Analysis*, *26*(2). https://doi.org/10.1111/j.1539-6924.2006.00755.x

Haimes, Y. Y. (2009). On the definition of resilience in systems. *Risk Analysis: An International Journal* , *29*(4), 498–501. https://doi.org/10.1111/j.1539-6924.2009.01216.x

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), 1–121. https://doi.org/10.2200/S00334ED1V01Y201102WBE001

Hendre, A., & Joshi, K. P. (2015). A Semantic Approach to Cloud Security and Compliance. *2015 IEEE 8th International Conference on Cloud Computing*, 1081–1084.

Herzog, A., Shahmehri, N., & Duma, C. (2011). An Ontology of Information Security. *International Journal of Information Security and Privacy*, *1*(4), 1–23. https://doi.org/10.4018/jisp.2007100101

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

Hoekstra, R., Breuker, J., di Bello, M., & Boer, A. (2009). LKIF core: Principled ontology development for the legal domain. *Frontiers in Artificial Intelligence and Applications*, *188*(1), 21–52. https://doi.org/10.3233/978-1-58603-942-4-21

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, *4*(1), 1–23.

Holton, G. A. (2004). Defining Risk. *Financial Analysts Journal*, *60*(6), 19–25. https://doi.org/10.2469/faj.v60.n6.2669

Höne, K., & Eloff, J. H. P. (2002). Information security policy –what do international information securitystandards say? *Computers & Security* , *21*(5), 402–409.

Humphrey, J., & Hubert Schmitz. (1998). Trust and inter-firm relations in developing and transition economies. *The Journal of Development Studies*, *34*(4), 32–61.

International Organization for Standardization. (2018a). *ISO 31000:2018 Risk management — Guidelines*.

International Organization for Standardization. (2018b). *ISO/IEC 27005:2018 Technologies de l'information — Techniques de sécurité — Gestion des risques liés à la sécurité de l'information*.

Jakoubi, S., Tjoa, S., & Quirchmayr, G. (2007). Rope: A Methodology for Enabling the Risk-Aware Modelling and Simulation of Business Processes. *ECIS 2007*.

Jallow, A. K., Majeed, B., Vergidis, K., Tiwari, A., & Roy, R. (2007). Operational risk analysis in business processes. In *BT Technology Journal* (Vol. 25, Issue 1, pp. 168–177). https://doi.org/10.1007/s10550-007-0018-4

Janowicz, E. K., Hitzler, P., Adams, B., Kolas, D., & Ii, C. V. (2014). Five Stars of Linked Data Vocabulary Use. In *Semantic Web* (Vol. 0). IOS Press.

KALINOWSKI, G. (1971). UNE NOUVELLE BRANCHE DE LA LOGIQUE: LA LOGIQUE DEONTIQUE. Son histoire, ses formes, ses résultats. *Archives de Philosophie*, *34*(1), 3–36. http://www.jstor.org/stable/43033310

Kaplan', S., & Garrick2, B. J. (1981). On The Quantitative Definition of Risk. *Risk Analysis*, *I*, *No. I*, 11–27.

Karagiannis, D., Junginger, S., & Strobl, R. (1996). Introduction to Business Process Management Systems Concepts. In B. Scholz-Reiter & E. Stickel (Eds.), *Business Process Modelling* (pp. 81–106). Springer Berlin Heidelberg.

Kemmerer, R. A. (2003). Cybersecurity. *25th International Conference on Software Engineering, 2003. Proceedings.* https://doi.org/10.1109/ICSE.2003.1201257

Kenny, M., & Watson, F. (2004). Legal risk and the financier. *Network and Correspondent Banking Review*.

Klimburg, A. (2012). National cyber security framework manual. In *NATO Cooperative Cyber Defense Center of Excellence*.

Knight, F. H., & Kelley, A. M. (1964). *Risk, uncertainty and profit, reprints of economic classics*.

Kosseff, J. (2017). Defining Cybersecurity Law. *Iowa Law Review*, *103*(1), 985–1031.

Labuschagne, L., & Vorster, A. (2005). A framework for comparing different information security risk analysis methodologies. *Proceedings of the 2005 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries*, 95–103.

Lamy, J. B. (2017). Owlready: Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies. *Artificial Intelligence in Medicine*, *80*, 11–28. https://doi.org/10.1016/j.artmed.2017.07.002

Lassila, O., & Swick, R. R. (1998). *Resource description framework (RDF) model and syntax specification*.

le Querler, N. (1996). Typologie des modalités. *Presses Universitaires de Caen*.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22 140*, 55.

Luhmann, N., Barrell, R., Stehr, N., & Bechmann, G. (2017). *Risk: a sociological theory*. Routledge.

Lukasik, S. J. (2000). Protecting the global information commons. *Telecommunications Policy*, *24*(6–7), 519–531. https://doi.org/10.1016/S0308-5961(00)00038-0

Maglitta, J. (1996). Smarten up! *Computerworld*, *29*(23), 84–90.

Mahler, T. (2007). DEFINING LEGAL RISK. *Proceedings Of The Conference" Commercial Contracting For Strategic Advantage-Potentials And Prospects*, 10–31.

Maniraj, V., & Sivakumar, R. (2010). Ontology Languages – A Review. *International Journal of Computer Theory and Engineering*, *2*(6).

Mar, B. W. (1997). Back to Basics Again–A Scientific Definition of Systems Engineering. *INCOSE International Symposium*, *7*(1), 309–316.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Martin, C., Kadry, A., & Abu-Shady, G. (2014). Quantifying the financial impact of it security breaches on business processes. *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, 149–155. https://doi.org/10.1109/PST.2014.6890934

Mayer, R. C., Davis, J. H., & David Schoorman, F. (1995). An Integrative Model of Organizational Trust. *Source: The Academy of Management Review*, *20*(3), 709–734.

McGuinness, D. L., & van Harmelen, F. (2004). OWL Web Ontology Language Overview. *W3C Recommendation*, *10*(10). http://www.w3.org/TR/2003/PR-owl-features-20031215/

Mcqueen, R. J. (1998). *Four Views of Knowledge and Knowledge Management*.

Mellado, D., Enrique, L., Crespo, S., Fernández-Medina, E., Rebollo, O., & Sanchez, L. E. (2011). *Comparative Analysis of Information Security Governance Frameworks: A Public Sector Approach*.

Miller, D., & Shamsie, J. (1996). The Resource-Based View of the Firm in Two Environments: The Hollywood Film Studios from 1936 to 1965. *The Academy of Management Journal*, *39*(3), 519–543.

Minsky, M. (1974). *A framework for representing knowledge*.

Moisand, D., & Garnier de Labareyre. (2009). *Cobit Pour une meilleure gouvernance des systèmes d'information* (Eyrolles).

Montgomery, D. C., & Woodall, W. H. (2008). An overview of six sigma. In *International Statistical Review* (Vol. 76, Issue 3, pp. 329–346). https://doi.org/10.1111/j.1751-5823.2008.00061.x

Motik, B., Patel-Schneider, P., & Parisa, B. (2009). OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax. *W3C Recommendation*, *27*(65), 159.

Myagmar, S., Lee, A. J., & Yurcik, W. (2005). Threat Modeling as a Basis for Security Requirements. *Symposium on Requirements Engineering for Information Security (SREIS)*, 1–8.

Myerson, R. B. (1991). *Game theory: Analysis of conflict* (Vol. 66). The President and Fellows of Harvard College.

Nardi, D., & Brachman, R. J. (2003). An Introduction to Description Logics. *Description Logic Handbook*, *1*, 1–40.

National Institute of Standards and Technology. (2018a). *Framework for Improving Critical Infrastructure Cybersecurity*.

National Institute of Standards and Technology. (2018b). *Risk Management Framework for Information Systems and Organizations A System Life Cycle Approach for Security and Privacy*.

National Institute of Standards and Technology. (2020). *Security and Privacy Controls for Information Systems and Organizations*. https://doi.org/10.6028/NIST.SP.800-53r5

Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T., & Swartout, W. R. (1991). Enabling Technology for Knowledge Sharing. *AI Magazine*, *12*(3), 36.

Newell, S., Swan, J., Galliers, R., & Scarbrough, H. (1999). The intranet as a knowledge management tool? Creating new electronic fences. In Hershey (Ed.), *Proceedings of The Information Resources Management Association International Conference, Managing Information Technology Resources in Organizations in the Next Millennium*.

Nonaka, I., & Takeuchi, H. (1995). The Knowledge-Creating Company: How JapaneseCompanies Create the Dynamics of Innovation. *Oxford University Press*.

Noy, N., Fergerson, R. W., & Musen, M. A. (2000). The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility. *EKAW*.

Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., & Robaldo, L. (2018). PrOnto: Privacy Ontology for Legal Compliance. *Proc. 18th Eur. Conf. Digital Government (ECDG)* , 142–151.

Pandit, H. J., Fatema, K., O'sullivan, D., & Lewis, D. (2018). GDPRtEXT-GDPR as a Linked Data Resource. *European Semantic Web Conference*, 481–495.

Pearce, W. B. (1974). Trust in interpersonal communication. *Speech Monographs*, *41*(3), 236–244.

Peirce, C. S. (1978). *Écrits sur le signe* (Vol. 31). Seuil.

Penrose, E. T., & Penrose, E. (2009). *The Theory of the Growth of the Firm* (Vol. 4). Oxford University Press.

Piccoli, G. (2007). *Information systems for managers: texts and cases* (John Wiley & Sons Inc, Ed.).

Pinto, H. S., & Martins, J. P. (2004). Ontologies: How can They be Built? *Knowledge and Information Systems*, *6*(4), 441–464. https://doi.org/10.1007/s10115-003-0138-1

Ponzetto, S. P., & Strube, M. (2007). Deriving a Large Scale Taxonomy from Wikipedia. *Association for the Advancement of Artificial Intelligence*, *7*, 1440–1445.

Poveda-Villalón, M., Espinoza-Arias, P., Garijo, D., & Corcho, O. (2020). Coming to Terms with FAIR Ontologies. *International Conference on Knowledge Engineering and Knowledge Management*, 255–270.

Rahmouni, H. B., Solomonides, T., Mont, M. C., & Simon Shiu. (2010). Privacy compliance and enforcement on European healthgrids: An approach through ontology. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *368*(1926), 4057–4072. https://doi.org/10.1098/rsta.2010.0169

Renn, O. (2008). *Concepts of Risk: An Interdisciplinary Review*. Oekom-Verl.

Rose, A., & Liao, S. Y. (2005). Modeling regional economic resilience to disasters: A computable general equilibrium analysis of water service disruptions. *Journal of Regional Science*, *45*(1), 75–112. https://doi.org/10.1111/j.0022-4146.2005.00365.x

Saleh, M. S., & Alfantookh, A. (2011). A new comprehensive framework for enterprise information security risk management. *Applied Computing and Informatics*, *9*(2), 107–118. https://doi.org/10.1016/j.aci.2011.05.002

Schatz, D., Bashroush, R., & Wall, J. (2017). Towards a More Representative Definition of Cyber Security. *The Journal of Digital Forensics, Security and Law*, *12*(2). https://doi.org/10.15394/jdfsl.2017.1476

Schiavone, S., Garg, L., & Summers, K. (2014). Ontology of Information Security in Enterprises. *The Electronic Journal Information Systems Evaluation*, *17*, 71–087.

Schlarman, S. (2007). Selecting an IT control framework. *Information Systems Security*, *16*(3), 147–151. https://doi.org/10.1080/10658980701225440

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). *Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios*. https://doi.org/10.1007/978-3-319-07668-3_37

Shameli-Sendi, A., Aghababaei-Barzegar, R., & Cheriet, M. (2016). Taxonomy of information security risk assessment (ISRA). *Computers and Security*, *57*, 14–30. https://doi.org/10.1016/j.cose.2015.11.001

Sheikhpour, R., & Modiri, N. (2012). An Approach to Map COBIT Processes to ISO/IEC 27001 Information Security Management Controls. *International Journal of Security and Its Applications*, *6*(2).

Simons, A., Niehaves, B., Riemer, K., vom Brocke, J., Plattfaut, R., & Cleven, A. (2009). Reconstructing The Giant: On The Importance Of Rigour In Documenting The Literature Search Process. *ECIS 2009 PROCEEDINGS*.

Snyder, C., Wilson, L., McManus, D., Snyder, C. A., Todd Wilson, L., & Johnson McManus, D. (1998). Knowledge Management: A Proposed Process Model. *AMCIS*.

Sohrabi Safa, N., von Solms, R., & Furnell, S. (2016). Information security policy compliance model in organizations. *Computers and Security*, *56*, 1–13. https://doi.org/10.1016/j.cose.2015.10.006

Sowa, J. (1984). *Conceptual structures: information processing in mind and machine*. Addison-Wesley Longman Publishing Co.

Spender, J. C. (1996). Organizational Knowledge, Learning, and Memory: Three Concepts inSearch of a Theory. *Journal of Organizational Change Management*, *9*, 63–78.

Stage, L. (2002). Les modalités épistémique et déontique dans les énoncés au futur (simple et composé). In *Revue romane*.

Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers and Security*, *24*(2), 124–133. https://doi.org/10.1016/j.cose.2004.07.001

Stojanovic, L. (2004). Methods and tools for ontology evolution. In *PhD thesis, University of Karlsruhe*. https://www.researchgate.net/publication/35658911

Studer, R., Benjamins B'c, V. R., & Fensel, D. (1998). Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, *25*(1–2), 161–197.

Su, X., & Ilebrekke, L. (2002). A Comparative Study of Ontology Languages and Tools. *International Conference on Advanced Information Systems Engineering*, 761–765.

Syed Abdullah, N., Sadiq, S., & Indulska, M. (2012). A Compliance Management Ontology: Developing Shared Understanding through Models. *International Conference on Advanced Information Systems Engineering*, *7328*, 429–444.

Takahashi, T., & Kadobayashi, Y. (2014). Reference Ontology for Cybersecurity Operational Information. *Computer Journal*, *58*(10), 2297–2312. https://doi.org/10.1093/comjnl/bxu101

Tang, Z., & Pan, Y. (2015). Big data security management. *Handbook of Research on Trends and Future Directions in Big Data and Web Intelligence*, 53–66. https://doi.org/10.4018/978-1-4666-8505-5.ch003

Tixier J, Salvi, D. G., & Gaston D. (2002). Review of sixty two risk analysis methodologies of industrial plants. *Journal of Loss Prevention in the Process Industries*, *15*(4), 291–303.

Tjoa, S., Jakoubi, S., & Quirchmayr, G. (2008). Enhancing business impact analysis and risk assessment applying a risk-aware business process modeling and simulation

methodology. *ARES 2008 - 3rd International Conference on Availability, Security, and Reliability, Proceedings*, 179–186. https://doi.org/10.1109/ARES.2008.206

Tuomi, I. (2000). Data is More Than Knowledge : Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory. *Journal of Management Information Systems*, *16*(3), 107–121.

Ullman, R., & Ferrera, D. (1998). Crime on the Internet. *Boston Bar Journal*, *6*.

Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, Methods and Applications. *Knowledge Engineering Review*, *11*(2).

Uschold, M., & King, M. (1995). Towards a Methodology for Building Ontologies. *Workshop on Basic Ontological Issues in Knowledge Sharing Held in Conjunction with IJCAI*, 1–13.

van Heijst, G., Chreiber, H. S., & Ielinga, B. J. W. (1997). Using explicit ontologies in KBS development. *Int. J. Human-Computer Studies*, *45*, 183–292.

van Kralingen, R. (1997). A Conceptual Frame-based Ontology for the Law. *Proceedings of the First International Workshop on Legal Ontologies*, 6–17. https://pdfs.semanticscholar.org/1073/93363c1fb2493f5871cb3c5fe08f70810c1b.pdf

Vance, A., Siponen, M., & Pahnila, S. (2012). Motivating IS security compliance: Insights from Habit and Protection Motivation Theory. *Information and Management*, *49*(3–4), 190–198. https://doi.org/10.1016/j.im.2012.04.002

Vance, D. (1997). Information, Knowledge and Wisdom: The Epistemic Hierarchy andComputer-Based Information Systems. *Association for Information Systems Electronic Library (AISeL)* , 8–15.

Visser, P. R. S., & Bench-Capon, T. J. M. (1998). A Comparison of Four Ontologies for the Design of Legal Knowledge Systems. *Artificial Intelligence and Law*, *6*(1), 27–57.

von Solms, R., & van Niekerk, J. (2013). From information security to cyber security. *Computers and Security*, *38*, 97–102. https://doi.org/10.1016/j.cose.2013.04.004

Vorobiev, A., & Bekmamedova, N. (2010). *An Ontology-Driven Approach Applied to Information Security* (Vol. 42, Issue 1).

Wangen, G. (2007). Information Security Risk Assessment: A Method Comparison. *IEEE Computer Magazine Special Issue: Cyber Physical Systems and Security Risk Assessment*, *6*(1), 52–61.

Wangen, G., Hallstensen, C., & Snekkenes, E. (2018). A framework for estimating information security risk assessment method completeness: Core Unified Risk Framework, CURF. *International Journal of Information Security*, *17*(6), 681–699. https://doi.org/10.1007/s10207-017-0382-0

Webster, J., & Watson, R. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii–xxiii.

Wiener, N. (2014). *La Cybernétique. Information et régulation dans le vivant et la machine*. Média Diffusion.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R.,

… Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*. https://doi.org/10.1038/sdata.2016.18

Woods, D. D. (2017). Essential Characteristics of Resilience. In CRC Press. (Ed.), *Resilience engineering* (pp. 21–34).

Zimmermann, A., Lorenz, A., & Specht, M. (n.d.). *The Use of an Information Brokering Tool in an Electronic Museum The Use Of An Information Brokering Tool In An Electronic Museum Environment*.