HAL
open science

# Conception contextuelle des interactions entre un modèle de processus opérationnel et des modèles de processus supports

Pierre-Aymeric Masse

## To cite this version:

Pierre-Aymeric Masse. Conception contextuelle des interactions entre un modèle de processus opérationnel et des modèles de processus supports. Génie logiciel [cs.SE]. Ecole nationale supérieure Mines-Télécom Atlantique, 2019. Français. NNT : 2019IMTA0039 . tel-02483328

**HAL Id: tel-02483328**

**https://theses.hal.science/tel-02483328v1**

Submitted on 18 Feb 2020

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

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

# THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies*
*de l'Information et de la Communication*
Spécialité : *Informatique*

Par
**Pierre-Aymeric MASSE**

# Conception contextuelle des interactions entre un modèle de processus opérationnel et des modèles de processus supports

**Thèse présentée et soutenue à Brest, le 19/12/2019**
**Unité de recherche : Lab-STICC**
**Thèse N° : 2019IMTA0039**

**Rapporteurs avant soutenance** :

| | |
|---|---|
| Selmin NURCAN | Maître de Conférence HDR - Université Paris 1 Panthéon Sorbonne |
| Walid GAALOUL | Professeur - Télécom SudParis (ex INT) |

**Composition du Jury :**

| | | |
|---|---|---|
| Président : | Giuseppe BERIO | Professeur - UBS |
| Examinateurs : | Marie-Pierre GERVAIS | Professeur - Université Paris Nanterre |
| | Selmin NURCAN | Maître de Conférence HDR - Université Paris 1 Panthéon Sorbonne |
| | Walid GAALOUL | Professeur - Télécom SudParis (ex INT) |
| Dir. de thèse : | Patrick MEYER | Professeur - IMT Atlantique |
| Encadrants : | Nassim LAGA | Docteur - Orange |
| | Jacques SIMONIN | Directeur d'études - IMT Atlantique |

# Résumé

De nos jours, l'utilisation des techniques de gestion de processus au sein des entreprises permet une augmentation significative de l'efficacité des systèmes opérationnels. Ces techniques assistent les experts métiers lors de la modélisation des processus de l'entreprise, de leur mise en œuvre, de leur analyse et de leur amélioration. L'expert métier, responsable de la gestion et de la supervision des processus de son domaine d'activité, applique des techniques d'analyse afin de mieux comprendre l'exécution de processus et, le cas échéant, les améliorer. La mise en œuvre des processus métiers est un enjeu pour les entreprises dans un marché dynamique et compétitif. Cette mise en œuvre doit être facilitée par l'accessibilité, au sens sémantique, des traitements par l'expert métier. De plus, elle doit être enrichie par une relève rapide des anomalies lors du déroulement des processus. L'étude bibliographique a permis d'identifier l'importance du contexte d'exécution sur le déroulement d'un processus.

En effet, le contexte d'exécution d'un processus métier contient des informations permettant d'identifier et de comprendre les interactions entre lui-même et les autres processus. Le premier processus est déclenché suite à un besoin client (processus opérationnel) et les autres, suite à un besoin du processus opérationnel (processus supports). La satisfaction d'un besoin client dépend en partie d'une interaction efficace entre le processus opérationnel et les processus supports qui lui sont liés. Ces interactions se définissent au travers des données opérationnelles, manipulées par le processus opérationnel, et des données supports, manipulées par les processus supports.

Nos travaux se situent dans le cadre de l'Ingénierie dirigée par les Modèles. L'approche proposée dans cette thèse est fondée sur l'annotation des modèles de processus opérationnels ou supports. Cette annotation est réalisée à l'aide d'une ontologie définissant le domaine métier décrit par ces processus. Ces annotations sont ensuite exploitées afin de constituer un ensemble de données, dites données contextuelles, qui font le lien entre les données opérationnelles et les données supports. L'analyse des traces de l'exécution du processus et de ces données contextuelles permet de sélectionner les meilleures d'entre elles, au sens métier. Ainsi, un processus opérationnel pourra être enrichi par un ensemble de processus supports via les données contextuelles.

# Abstract

Nowadays, the use of business process management techniques within companies allows a significant improvement in the efficiency of the operational systems. These techniques assist business experts in modelling business processes, implementation, analytics, and enhancements. The business expert, responsible for the management and supervision of the processes in his area of activity, applies analytical techniques to better understand the process execution and, if necessary, to improve them. The implementation of business processes within the company is an issue for companies in a dynamic and competitive market. This implementation must be facilitated by the accessibility, in the semantic sense, of the treatments by the business expert. Moreover, during the execution of the process, it must be enriched by a quick fix of anomalies during the process execution. The literature review identified the importance of the execution context of a process.

Indeed, the execution context of a business process contains information to identify and understand the interactions between itself and other processes. The first process is triggered following a customer need (operationnal process) and the others, following the need of the operationnal process (support processes). The satisfaction of a customer need depends on an effective interaction between the operationnal process and the related support processes. These interactions are defined through operational data, manipulated by the operational process, and support data, manipulated by the support processes.

Our work is based on the framework of Model Driven Engineering. The approach proposed in this thesis is based on the annotation of operational and support process models. This annotation is performed with the help of an ontology defining the business domain described by these processes. These annotations are then exploited to constitute a set of data, called contextual data. The analysis of the traces of the execution of the operational process and of these contextual data makes it possible to select the best subset of contextual data, in the business sense. Thus, an operational process can be associated with a set of support processes via the contextual data.

# Publications

## Conférences et Ateliers internationales

- Pierre-Aymeric Masse, Nassim Laga, Oussama Mohammed Kherbouche, Jacques Simonin. An Approach Based on Ontology for Discovering Data Impacting the Execution of a Business Process. IEEE International Conference on Information Systems & Technology (CIST), pages 216-221, 2016.
- Nassim Laga, Oussama Mohammed Kherbouche, Pierre-Aymeric Masse. Communication-Based Business Process Task Detection - Application in the CRM Context. EDOC Workshops, pages 1-8, 2016.
- Pierre-Aymeric Masse, Nassim Laga, Jacques Simonin. A Contextual Data Selection Tool for an Enhanced Business Process Analysis. IEEE International Conference on e-Business Engineering (ICEBE), pages 1-8, 2016.
- Oussama Mohammed Kherbouche, Nassim Laga, Pierre-Aymeric Masse. Towards a better assessment of event logs quality. IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments, pages 1-8, 2016.
- Pierre-Aymeric Masse. Enrichment of Business Process Management with External Data. CAiSE (Doctoral Consortium), pages 47-55, 2015.

## Brevets

- Pierre-Aymeric Masse, Nassim Laga. Extraction d'informations expliquant le comportement d'un processus (2016).
- Nassim Laga, Pierre-Aymeric Masse, Sonia Laurent, Philippe legay. Mechanism for detecting important words in a text (2016).

# Table des matières

**Résumé**

**1 Introduction** **1**
1.1 Domaine de recherche et concepts associés . . . . . 1
1.2 Problématique de recherche . . . . . 2
1.3 Plan du document . . . . . 5

**2 État de l'art** **7**
2.1 Gestion de processus métiers . . . . . 7
2.1.1 Définition d'un processus métier . . . . . 8
2.1.2 Gestion de changement d'un processus métier . . . . . 10
2.2 Gestion de changement mono-processus . . . . . 11
2.2.1 Gestion de changement d'un processus fondé sur les données . . . . 11
2.2.2 Ingénierie des processus dirigée par les modèles . . . . . 19
2.3 Gestion de changement multi-processus . . . . . 25
2.3.1 Relations entre processus métier . . . . . 26
2.3.2 Schéma de mise en relation . . . . . 26
2.3.3 Conclusion . . . . . 27
2.4 Application de l'aide à la décision multicritère au sein du domaine du BPM 28
2.5 Conclusion . . . . . 30

**3 Introduction de la contribution et définitions** **33**

**4 Conception sémantique des données contextuelles d'un processus opérationnel et support** **43**
4.1 Données contextuelles d'un processus métier . . . . . 43
4.1.1 Annotation d'un processus métier avec une ontologie . . . . . 45
4.1.2 Sélection des données contextuelles à partir de l'annotation . . . . . 49
4.2 Partitionnement des traces des instances d'un processus opérationnel pour détecter des comportements . . . . . 54
4.2.1 Premier partitionnement des traces des instances d'exécution . . . . 55
4.2.2 Second partitionnement à l'aide des instances des données contextuelles . . . . . 60

4.3 Sélection et représentation des données contextuelles significatives par rapport à un comportement . . . . . 61
4.3.1 Sélection des données contextuelles associées à un comportement . . 61
4.3.2 Représentation de sous-ensembles de données contextuelles en un arbre de décision métier . . . . . 62
4.4 Conclusion . . . . . 63

**5 Évaluation métier des processus supports gérant chaque donnée contextuelle 65**
5.1 Sélection des processus supports gérant chaque donnée contextuelle . . . . 65
5.2 Évaluation des processus supports en fonction des préférences de l'expert métier . . . . . 66
5.2.1 Identification du problème . . . . . 68
5.2.2 Formulation du problème . . . . . 69
5.2.3 Modèle d'évaluation . . . . . 72
5.2.4 Expérimentation . . . . . 73
5.3 Conclusion . . . . . 78

**6 Développement de l'outil d'enrichissement d'un processus opérationnel par des processus supports 79**
6.1 Analyse fonctionnelle de l'outil . . . . . 79
6.1.1 Analyse fonctionnelle statique de l'outil fondée sur les transformations 81
6.1.2 Analyse fonctionnelle dynamique de l'outil fondée sur les transformations . . . . . 94
6.2 Architecture applicative de l'outil d'enrichissement . . . . . 101
6.2.1 Architecture technique de déploiement de l'outil . . . . . 101
6.2.2 Architecture applicative statique de l'outil . . . . . 102
6.2.3 Architecture applicative dynamique de l'outil . . . . . 106
6.3 Conclusion . . . . . 113

**7 Conclusion et Perspectives 115**
7.1 Conclusion . . . . . 115
7.2 Perspectives . . . . . 116
7.2.1 Traitement de grande volumétrie de données . . . . . 116
7.2.2 Automatisation des transformations . . . . . 116
7.2.3 Incorporer d'autres éléments tels que les objectifs . . . . . 117

**Bibliographie 119**

Chapitre 1

# Introduction

Ce chapitre introduit et explicite le domaine de recherche relatif à mes travaux. La première section présente le domaine de recherche de manière succincte et introduit les concepts qui lui sont associés. La deuxième section explicite la problématique de recherche ainsi que les questions de recherche sous-jacentes. La dernière partie de ce chapitre annonce le plan du document.

## 1.1 Domaine de recherche et concepts associés

Il est communément admis que la performance d'une entreprise dépend en partie de son organisation et de son fonctionnement. En effet, les entreprises évoluent dans des environnements très dynamiques et doivent s'y adapter le plus rapidement possible [105]. La mondialisation et l'évolution technologique continue sont les principaux moteurs de ces évolutions. L'adaptation des entreprises passe alors par la mise en place d'une stratégie adéquate qui a, pour conséquence immédiate, de modifier l'organisation de l'entreprise et les processus métiers sous-jacents [118].

Cette recherche d'adaptabilité n'est pas nouvelle. En effet dès 1880, l'ingénieur Frederick Winslow Taylor présenta une méthode afin d'améliorer la performance des entreprises en utilisant des principes scientifiques (Taylorisme [125]). Quelques années après, Henry Ford proposa un modèle d'organisation qui a permis d'améliorer la rentabilité des chaînes de construction dans l'industrie automobile. Il aura fallu attendre 1967 pour que le terme de « modélisation de processus métiers » soit utilisé pour nommer les méthodes décrivant le fonctionnement d'une entreprise. Ces méthodes de gestion se basaient principalement sur le principe de décomposition du travail en tâches, introduit par Adam Smith en 1776. Cependant, ces méthodes ne prenaient pas en compte la qualité et la diversité que peuvent avoir les produits fabriqués (couleur, options différentes, etc.). En 1983, Peter Drucker a repris certains principes du taylorisme en expliquant que la réussite du management d'une entreprise est basée sur cinq points [34] :

- fixer des objectifs ;
- mesurer leur réalisation ;
- organiser le travail ;

- motiver les salariés (au travers du salaire par exemple) ;
- former et communiquer avec eux.

Avec l'arrivée massive de l'informatique, les entreprises ont dû repenser leur processus métier. En 1993, Michel Hammer et James Champy ont publié le livre *Reengineering the Corporation* expliquant qu'il ne suffisait pas d'informatiser voire d'automatiser certaines tâches du processus pour intégrer les outils informatiques, mais qu'il était nécessaire de repenser les processus dans leur globalité. Depuis, le domaine de la gestion des processus métiers a évolué vers une approche incrémentale cherchant à améliorer continuellement les processus métier. Cette approche découle de l'amélioration continue telle que proposée par Demings [84].

Les techniques d'analyse des évènements d'un processus se regroupent au sein du domaine de la « fouille de processus » ou *process mining* en anglais. En 2012, Will Van Der Aalst explique dans *Process Mining Put Into Context* [127] l'intérêt du contexte d'exécution d'un processus métier, c'est-à-dire l'intérêt de prendre en compte durant son analyse l'ensemble des éléments constituant son environnement d'exécution. Toutefois, la majorité des travaux continue de se concentrer sur l'étude du modèle du processus, des évènements générés par ce processus et de ses performances sans prendre en compte dans quel environnement il se déroule.
C'est pourquoi, les travaux de cette thèse porteront sur la prise en compte de l'environnement d'un processus métier durant son analyse. Ils permettront d'aider l'entreprise, prise dans son contexte, à avoir des processus métier les mieux adaptés possible.
A ce jour, peu d'outils sont à disposition des entreprises pour la démarche d'analyse d'un processus. Ainsi, cette dernière reste en grande partie manuelle. Les employés en charge de l'analyse, que l'on nommera experts métier dans le reste de ce document, doivent identifier manuellement les facteurs qui induisent sur le comportement du processus. Il existe plusieurs comportements associés à un processus, certains peuvent être associés à des exécutions déviantes du modèle théorique entraînant potentiellement une sur-performance ou une sous-performance. Ces comportements peuvent être définis par un type de déroulement et par son contexte caractérisés par des données contextuelles.

La prise en compte du contexte dans lequel se déroule les processus métiers est donc une des clés de la réussite de l'entreprise.

## 1.2 Problématique de recherche

Le système opérationnel d'une entreprise est composé d'une multitude de processus métiers. Ces processus sont le plus souvent inter-dépendants. La bonne réalisation opéra-

tionnelle d'un des processus est donc liée à la bonne réalisation des autres processus en collaboration.

Mes travaux se concentrent sur la relation entre des processus de nature différente, c'est à dire entre des processus répondant à un besoin métier (processus opérationnel) et des processus en collaboration avec ce dernier (processus support). Cette étude a été réalisée afin de compléter la littérature qui s'est, dans un premier temps, focalisée sur la relation reliant deux processus de même nature (relation entre plusieurs processus opérationnels) [13].

Ces travaux permettront de répondre à la problématiques suivante : **comment, au travers du contexte que l'on représente par des données contextuelles, peut-on améliorer l'efficacité des processus d'une entreprise en améliorant les interactions entre processus opérationnel et processus supports ?**

Cette problématique induit des questions de recherche sur la manière de détecter et sélectionner les données contextuelles, sur l'exploitation de ces données et d'améliorer la relation que peut avoir un processus opérationnel avec des processus supports.
Les trois questions de recherche suivantes posent les barrières scientifiques à lever afin d'apporter une solution à cette problématique.

La première question de recherche (QR1) fait référence à la conception de l'ensemble des données contextuelles pouvant induire un changement sur l'exécution d'une instance d'un processus métier. La définition d'un contexte est un sujet étudié depuis une vingtaine d'années [22, 23, 132, 36]. Les travaux de Saidani et al. [113, 114] sont des exemples de l'usage du contexte dans le domaine du *Business Process Management* (BPM). Wieland et al. [133] proposent une méthodologie afin de rendre un modèle de processus sensible à son contexte. Notre approche consiste à créer une ontologie avant d'exécuter un processus pour enrichir le modèle du processus métier. Cette ontologie décrit, au travers de ses concepts, les données contextuelles et les activités associées à des processus opérationnels ou à des processus supports. Cet enrichissement va nous permettre de récupérer les informations du contexte directement ou indirectement associées à l'exécution du processus. Cette approche diffère d'autres techniques proposant de rechercher les données associées à l'exécution d'un processus dans les systèmes d'informations (SI) en exploitant les relations que peuvent avoir des données au sein du SI. **La question portera donc sur l'identification des données contextuelles**.

La deuxième question de recherche (QR2) fait référence au *Process context analysis* [7, 58, 128]. Il s'agit de l'exploitation des informations contenues dans le contexte afin d'approfondir la compréhension de l'exécution des processus aussi appelée analyse approfondie

des processus (*deep business analysis* en anglais) [91]. Cette analyse permet de détecter les variantes de processus [14, 59], ou encore d'identifier les facteurs justifiant la performance d'un processus (taux de réalisation d'objectifs) [7, 58]. Certaines méthodes proposent de s'appuyer sur ces informations afin de modifier le processus [91] pour le rendre plus performant. Cependant, la plupart de ces méthodes se basent sur un ensemble restreint de données subissant un fort biais de sélection ou nécessitant un travail en amont afin de définir les analyses à effectuer [59]. Ces méthodes sont limitées, c'est pourquoi notre approche nous permettra une meilleur identification des facteurs impactant l'exécution du processus par l'automatisation de l'exploitation des données précédemment sélectionnées (QR1), ainsi que du modèle du processus ou encore des évènements de celui-ci. **La question portera donc sur l'analyse des données contextuelles afin d'identifier les facteurs impactant le comportement d'un processus métier**.

La troisième question de recherche (QR3) fait référence au domaine de l'enrichissement de processus opérationnels par des processus supports en introduisant l'apport du contexte. Les méthodes du domaine du BPM utilisent principalement des données prédéfinies [53, 118], des instances de processus déjà réalisées [41] ou encore des règles métiers [96] voir des *patterns* d'optimisation [91] afin d'améliorer la définition d'un processus. Ces méthodes n'exploitent donc ni les préférences que peuvent avoir les experts métiers ni le contexte d'exécution. **La question portera donc sur la conception d'interaction automatique entre processus en exploitant les données contextuelles**. Par exemple, relier un processus opérationnel de livraison et un processus support de réparation de véhicule afin de pouvoir déclencher une intervention de façon automatique lors d'une défaillance. Relier ces deux types de processus permettra notamment d'avoir une vue consolidée des coûts complets du processus opérationnel.

Ma thèse présente une approche permettant l'apport du contexte dans le processus d'analyse des processus opérationnels, depuis sa modélisation jusqu'à un modèle de processus enrichi par des processus supports. Cette approche, basée sur les trois questions de recherche, est construite comme une succession d'étapes pouvant être utilisées indépendamment.

Mes travaux ont pour conséquence d'assister opérationnellement les experts métiers au sein de l'entreprise dans :

- La sélection des données à surveiller lors de l'exécution : l'objectif est d'aider les experts métiers à répertorier les données à faire surveiller par les équipes chargées de l'exploitation des services réalisant les processus opérationnels. Ces données permettront de détecter les déviations, ou encore les anomalies pouvant survenir durant

l'exécution d'un processus ;

- Leur compréhension des exécutions de processus : l'objectif est d'analyser les données associées au déroulement d'un processus afin d'améliorer la compréhension de ce dernier. Cette assistance permettrait d'accélérer la correction d'une anomalie. Par exemple, l'arrêt d'un service ou encore le retard d'une livraison ;

- La transformation des modèles des processus métiers en modèles dynamiques : l'objectif est d'aider les experts métiers à rendre dynamiques les modèles de processus en leur proposant des modifications pertinentes pour le métier et notamment en modélisant des liens entre processus métiers. Créer ces liens permet, durant l'exécution d'un processus opérationnel, de déclencher le bon processus support au bon moment et ainsi répondre à un besoin du processus opérationnel. Dans notre cas, nous nous focaliserons sur la réponse à une anomalie. Ces liens rendent possibles la création d'une vue consolidée de plusieurs processus permettant, par exemple, une meilleure estimation du coût d'exécution.

## 1.3 Plan du document

L'introduction de ce document présente les questions de recherche traitées par cette thèse ainsi que le contexte de recherche. Le reste du document est composé de cinq chapitres.
Le deuxième chapitre, État de l'Art, décrit les travaux existants concernant les points clés de la contribution.
Le troisième chapitre introduit les concepts associés à notre contribution.
Le quatrième détaille notre contribution dans la conception sémantique d'une relation entre un processus opérationnel et son contexte d'exécution.
Le cinquième chapitre aborde la conception sémantique entre le contexte d'exécution d'un processus opérationnel et des processus supports.
Le sixième chapitre présente en détail notre approche et le développement d'un outil d'enrichissement de processus opérationnels à l'aide de processus supports en se fondant sur les relations conçues dans les chapitres précédents.
La conclusion résume notre travail, et présente les perspectives futures.

Chapitre 2

# État de l'art

Ce chapitre est une revue des travaux de la littérature souhaitant répondre à notre problématique de recherche. La première partie se focalise sur une explication générale du domaine. Les deux parties suivantes se concentrent sur l'élément central de la problématique à savoir la modification des processus. Ces deux parties reviendront respectivement sur les méthodes de modification centrées sur un seul processus (mono-processus) et puis sur les méthodes de modification centrées sur plusieurs processus métiers (multi-processus). L'application de ces méthodes nécessite la prise de décision quant aux modifications à effectuer. C'est pourquoi ce chapitre se termine sur une revue de l'utilisation de techniques d'aide à la décision appliquées notre domaine.

## 2.1 Gestion de processus métiers

La gestion de processus métiers, *Business Process Management* (BPM) en anglais, est une discipline composée d'un ensemble de méthodes et d'outils. Cette discipline a pour objectif d'améliorer l'efficacité du travail au sein d'une organisation au travers de processus métier. Le BPM se caractérise par la modélisation, l'exécution, le contrôle, la mesure et l'optimisation des processus alignés aux objectifs de l'entreprise, aux systèmes, aux employés, aux clients et aux partenaires.

Les processus métiers aident les employés à formaliser un ensemble de séquences d'activités permettant d'améliorer leur efficacité en organisant leur travail, en limitant les erreurs et en améliorant le travail en équipe. Les processus décrivent précisément le travail à effectuer, les rôles et les acteurs. Ils permettent ainsi de réaliser une communication ciblée sur les bon acteurs, c'est-à-dire, la capacité d'informer d'un problème l'acteur concerné. Il sera alors en mesure de traiter l'information critique au moment opportun et ainsi améliorer le service rendu au client. Puisque les processus peuvent être décrits, ils peuvent être améliorés de manière continue. Les processus vont naturellement créer de l'information, au fur et à mesure de leur utilisation par une organisation, qui pourra être analysée et utilisée pour améliorer le processus. Ils sont de très bons supports pour capturer la connaissance métier puisque la valeur et les avantages de la surveillance d'un processus doivent rapidement permettre d'identifier les processus critiques ainsi que de permettre une « véritable » compréhension de ce qui est important pour la réussite de l'entreprise.

La mise en place d'une gestion de processus permet de mesurer les évolutions des processus avant, pendant et après un changement majeur (par exemple : une fusion, délocalisation, migration des systèmes, ...). Ils permettent d'analyser temporellement des anomalies et de justifier des succès opérationnels (par exemple, augmentation des ventes de téléphone mobile pour un opérateur).

Comme de nombreuses domaines de recherche, le BPM a pour vocation à fonctionner avec de nouvelles techniques simplifiant la démarche d'amélioration des processus, leurs usages et leurs compréhensions.

### 2.1.1 Définition d'un processus métier

Un processus métier, ou *Business Process* en anglais, est défini par le *Workflow Management Coalition* (WfMC). Il s'agit de « l'automatisation des procédures où les documents, les informations ou les tâches sont transmis entre les participants selon un ensemble défini de règles pour atteindre ou contribuer à un objectif fonctionnel global » [136]. En effet, un processus métier met en exergue une collection de tâches (ou activités) et de tâches organisées dans le temps qui, une fois achevées, permettront d'atteindre un objectif organisationnel ainsi qu'un résultat précis et mesurable. Autrement dit, il est considéré comme un ensemble de relations logiques entre un groupe d'activités incluant des interactions entre différents partenaires et participants. Il peut s'agir des acteurs humains ou encore d'autres processus métiers sous la forme d'échange d'informations et de données pour fournir une valeur ajoutée tangible aux clients et aux différents interlocuteurs d'une organisation [55].

#### 2.1.1.1 Types de processus métier

Un processus métier est défini comme étant une séquence d'activités amenant à une production métier. Les processus métiers sont au coeur des organisations et se déclinent dans tous les domaines, que ce soit dans le management, dans la gestion financière ou encore dans la gestion de l'excellence opérationnelle (ISO 9001, etc.). Il existe trois principaux types de processus métier : les processus opérationnels, parfois appelés *primary* en anglais, les processus supports, et les processus de management [130].

**Processus opérationnel** Les processus opérationnels, aussi nommés *operational processes* en anglais, sont des processus tels que définis par Hammer et Champy [54], ayant comme but de satisfaire le besoin d'un client externe. Ces processus opérationnels sont déclenchés suite à un besoin client et ont vocation à y répondre.

Ils nécessitent des processus internes à l'entreprise assurant leur fonctionnement. Ce sont les processus supports.

**Processus support** Les processus supports [130], aussi nommés *supporting processes* en anglais, ont pour objectif de supporter les processus métiers dans la réalisation de la demande par un client interne à l'entreprise. Ces derniers peuvent être déclenchés suite à un évènement pouvant être aléatoire, mais ne s'agissant pas d'une demande de client interne. Il s'agit par exemple des processus associés à la résolution de pannes.

**Processus de management** Le dernier type de processus est le processus de management [130] (*management processes*). Ils sont en charge de la configuration d'autres processus métier, de la définition de leurs objectifs, de leur planification, du contrôle des ressources et de la conduite de leur exécution. On citera en exemple, le processus de gestion d'entreprise, de définition de stratégie, de supervision de budget.

#### 2.1.1.2 Cycle de vie d'un processus métier

Le cycle de vie d'un processus métier a été défini au sein du manifeste de l'analyse de processus (*Process Mining*) [128]. Ce manifeste écrit par l'IEEE CIS Task Force on Process Mining est un document, signé par plus de 70 chercheurs, qui a vocation d'expliquer et promouvoir le domaine de recherche du *Process Mining*.
Ce cycle de vie, illustré par la Figure 2.1, est composé de sept phases :

- *diagnosis* : le *diagnosis* correspond à l'analyse du processus et de ses évènements une fois qu'un certain nombre d'exécutions du processus sont terminées. Cette phase est aussi l'occasion de l'utilisation des méthodes d'analyse de processus ;
- *(re) design* : la modélisation d'un processus métier consiste à représenter les méthodes de travail telles qu'elles sont mises en place dans l'organisation. Le *re-design* consiste à modifier le modèle d'un processus en fonction des observations issues de la phase de *diagnosis* ;
- *analysis* : l'analyse, contrairement au *diagnosis*, se focalise sur l'analyse du modèle de processus. Cette phase a pour objectif la détection d'anomalies au niveau de la modélisation telle que les boucles infinies (lorsqu'une suite d'activité dans certaines conditions s'exécute en boucle ne permettant pas de terminer l'exécution du processus), ou une impasse (lorsqu'une suite d'activité dans certaines conditions ne permet pas de terminer l'exécution du processus) ;
- *implementation* : la phase d'implémentation consiste au développement et à la mise en place des services permettant l'exécution du processus ;
- *(re) configuration* : la configuration des services réalisant le processus se déroule pendant la phase de *configuration* et peut être déclenchée suite à l'implémentation

ou la modification du processus (*re-design*) ;

- *execution* : la phase *execution* correspond à l'utilisation du processus. Chaque fois qu'un processus est exécuté, on parlera d'instance du processus ;
- *adjustement* : durant la phase *execution*, les processus sont supervisés et peuvent s'ajuster à l'utilisation.

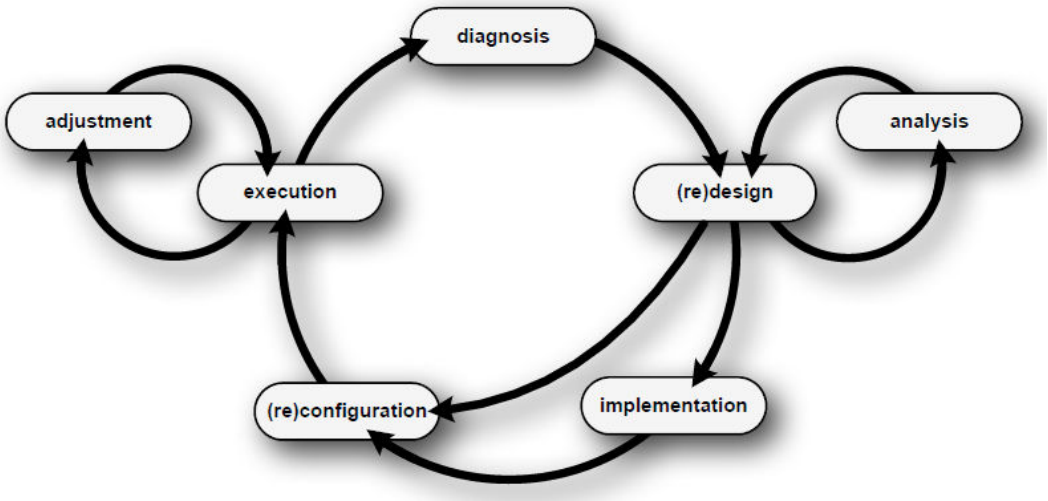


FIGURE 2.1 – Cycle de vie d'un processus métier.

Notre problématique de recherche fait donc référence au changement initié lors de la phase de *re design*. Ce changement est souvent initié par un souhait d'aligner le processus avec les nouveaux besoins de l'entreprise.

### 2.1.2 Gestion de changement d'un processus métier

Le changement des processus métiers, nommé *Business Process Change (BPC)* en anglais, est l'une des méthodes utilisées par les organisations pour s'adapter à un environnement en évolution rapide. Les changements accélérés au sein de leur environnement font pression pour qu'elles réagissent en remaniant continuellement leur stratégie, leur organisation et leurs processus [95]. En ne réagissant pas, les entreprises risquent de perdre au fil du temps leur avantage compétitif au profit d'entreprises plus flexibles ou plus innovantes. Cet environnement très dynamique a suscité un intérêt croissant pour la recherche visant à améliorer les processus opérationnels afin d'accroître la souplesse et d'améliorer le rendement [109]. Les projets de BPC sont souvent complexes et entachés d'incertitudes, de retards fréquents, voire d'échecs. Le BPC est une approche holistique qui comporte de nombreux risques organisationnels, technologiques, économiques et sociaux. Environ

60% à 80% des projets de BPC ont échoué [109, 68, 122]. L'un des principaux facteurs de réussite est la capacité des organisations à comprendre et à faire face aux tâches organisationnelles et économiques complexes que ces projets leur imposent. Le BPC est une approche de management [51, 67] qui s'applique sur tout type de changement et est définie comme une « initiative organisationnelle axée sur la stratégie visant à (re)concevoir les processus opérationnels afin d'obtenir des améliorations, incrémentales ou de rupture, significatives en matière de performance (qualité, réactivité, coût, flexibilité, satisfaction et autres mesures critiques des processus) par des changements dans la relation entre la direction, les informations, la technologie, la structure organisationnelle et le personnel » [51]. Étant donné que ces initiatives peuvent avoir une portée différente selon le degré de changement qui est encouragé dans chaque organisation, la définition du BPC implique l'intégration d'approches incrémentales, comme par exemple la gestion de la qualité totale (GQT ou TQM en anglais), ou de rupture, comme par exemple la réingénierie des processus opérationnels (*Business Process Reengineering*, BPR) [52, 115, 90].

Il est donc en partie question dans le BPC de modifier les processus afin de les aligner avec leur environnement et avec l'organisation de l'entreprise. C'est-à-dire, modifier le processus pour qu'il puisse le mieux possible répondre aux besoins de l'organisation dans son environnement d'exécution spécifique (législation, etc.). À partir de l'explication qui a été donnée ci-dessus, mes travaux appartiennent donc au BPC et utilisent une approche incrémentale. Cependant les changements peuvent impacter un ensemble de processus de manière individuelle ou de manière ensembliste. Nous allons donc maintenant nous intéresser au changement d'un seul processus.

## 2.2 Gestion de changement mono-processus

Le changement au sein d'un processus est un sujet très étudié dans la littérature [77, 98, 91, 106]. Elle intervient lorsqu'un besoin identifié de modifier un processus apparaît nécessaire sans toutefois nécessiter la modification d'autres processus. Cette modification d'un seul processus métier doit donc pouvoir répondre seule au besoin identifié. Ce changement consiste à aligner un processus avec la stratégie de l'entreprise, qui est transposée en objectifs métiers [96], avec son environnement d'exécution [98] ou avec l'organisation de l'entreprise [118].

### 2.2.1 Gestion de changement d'un processus fondé sur les données

Nous avons regroupé les méthodes de modification d'un processus dans trois différents groupes. Les groupes ont été définis par la diversité de type des données utilisées dans la modification du processus. Le premier groupe est composé des méthodes exploitant

l'exécution du processus et des données métiers. Ce sont les méthodes les plus anciennes, et sont utilisées soit pour modifier le processus après son exécution, ou soit directement pendant son exécution. Le deuxième groupe est composé des méthodes exploitant des données supplémentaires prédéfinies ou en relation avec les données métiers. Enfin le troisième groupe est composé de méthodes utilisant le contexte.

#### 2.2.1.1 Changements fondés sur les données d'exécution d'un processus et ses données métiers

Les processus sont conçus dans le but de guider les employés dans leurs activités. Cependant, il est communément admis qu'un processus trop rigide est contre-productif, car il ne permet pas de s'adapter aux caractéristiques des employés. De même, s'il est trop flexible, il ne permettra pas de gain de productivité ni de qualité. Les processus flexibles permettent ainsi de trouver un compromis en créant des portes logiques dans le processus afin d'offrir plusieurs parcours conditionnés par des « gardes » (c'est-à-dire, plusieurs scénarios de déroulement du processus). Ces gardes peuvent être activées en fonction de données ou encore en fonction de comment « l'instance se déroule ». La différence entre les processus configurables et les processus flexibles réside dans la manière de créer ces portes logiques.

Les processus configurables rationalisent la conception de modèles spécialisés pour un métier en fournissant une solution générique [61] aussi nommée modèles de références. Ces modèles fournissent un référentiel de modèles potentiellement pertinents pouvant être utilisés pour accélérer le processus de modélisation en s'adaptant à un scénario de déroulement du processus. En effet, la modélisation des processus métiers à partir de zéro devient une tâche fastidieuse. Pour utiliser le paradigme « Conception par Réutilisation », les processus configurables sont généralement « plug and play » après quelques ajustements en fonction des besoins individuels [106]. Un modèle de processus configurable doit être configuré en fonction des exigences spécifiques d'une organisation afin de dériver une variante du processus. Nous pouvons citer plusieurs langages permettant la définition de tel processus métier [5, 49, 140] : c-EPC, c-BPMN. Ce type de processus est souvent utilisé par des progiciels de gestion intégrée ou PGI (en anglais *Enterprise Resource Planning* ou ERP) tel que SAP.

Un processus flexible est un processus dont les portes logiques ont été définies durant sa définition. Chaque instance de ce type de processus va pouvoir s'ajuster en fonction des données afin de mieux s'exécuter. Cette adaptation fait référence à la phase d'*adjustement* définie dans le cycle de vie (Cf Figure 2.1). Talib et al. [124] proposent un système fondé sur un processus flexible. Leur système permet de trouver le meilleur employé disponible pour exécuter une certaine activité en fonction des résultats historiques de ce dernier,

c'est-à-dire que le système va dynamiquement exécuter une activité en fonction de donnée issue du système de gestion de ressource. En fonction de la disponibilité des employés et de leurs précédentes performances sur cette activité, le système proposé va essayer de choisir le meilleur employé. L'objectif est de répondre le mieux possible à la demande du client.

Selon Rosemann et Recker [107], la flexibilité des processus métiers est la capacité de générer des modifications de l'exécution d'un processus en réaction à un déclencheur externe. Cependant, Rosemann et al. [102] soulignent que, tout en étant capable de réagir au changement, il est particulièrement important de maintenir également la stabilité du processus métier. Le changement ne peut concerner que les parties d'un processus réellement affectées (c'est-à-dire modifier les parties affectées sans remplacer l'ensemble du processus [107]). Ce changement ne peut être obtenu que par une structure appropriée (ni trop simple ni trop complexe) avec le respect des problèmes et des tâches de l'entreprise [102].

L'application de la flexibilité au sein d'un processus peut apporter un gain de performance pour le processus [124]. Cependant, leur utilisation dans la pratique est toutefois souvent limitée. La sélection et l'application des modifications à opérer sont des tâches difficiles. Par conséquent, de nombreuses possibilités ne sont pas considérées comme de potentiellement amélioration. Utiliser des données supplémentaires (données métiers, ou encore données contextuelles) durant l'analyse permet de réduire cette difficulté.

#### 2.2.1.2 Changements fondés sur les données externes au processus

Lorsqu'il est question d'ajouter de nouvelles données durant l'analyse, une difficulté apparaît concernant la manière d'associer un processus à ces données [77]. Cette difficulté est aussi présente dans notre première question de recherche (Cf. QR1) qui met en exergue l'utilisation de plusieurs types de données afin d'enrichir l'analyse d'un processus métier. Les données peuvent être celles associées au processus ou encore celles associées aux perturbations survenues lors du déroulement de ce dernier.

Par exemple, Radeschutz et al. [98] proposent de mettre en relation les données métiers produites par un processus avec des données disponibles dans le SI afin de créer une vue consolidée des données nommée *deep business data description (dBDD)*. Cette vue consolidée est ensuite utilisée dans une analyse globale. Ils présentent le concept d'analyse de processus approfondie (en anglais *Deep Business Process analysis*) qui peut être définie comme l'appariement d'un processus avec des données supplémentaires pour réaliser une analyse globale de ce dernier. Au travers de leur publication, les auteurs montrent l'intérêt de créer une vue consolidée des principales sources de données métiers afin d'analyser et d'optimiser les processus de manière beaucoup plus complète. Ils proposent plusieurs méthodes d'appariement telles que : (1) une correspondance manuelle et statique des données

produites par le processus avec des données disponibles dans le SI, et (2) une correspondance semi-automatique basée sur l'utilisation de techniques sémantiques. L'expert métier annote manuellement les activités du processus ainsi que les données avec des concepts issus d'une ontologie couvrant le domaine métier. Le système cherche ensuite à associer les données entre elles au travers de l'ontologie. Une variante est aussi proposée, dans cette publication, utilisant des techniques de traitement du langage naturel (en anglais *Natural Language Processing, NLP*) pour annoter automatiquement les données.

Cette mise en relation d'une donnée métier avec d'autres données est très intéressante et permet de répondre partiellement à notre question de recherche. Il est cependant à noter que cette approche n'exploite pas la sémantique disponible dans le nom des activités produisant les données métiers. De plus, les auteurs de [98] ne décrivent pas de mécanisme permettant l'association automatique des données avec les concepts de l'ontologie.

Une approche similaire a été proposée par Niedermann et al. [91]. Les auteurs utilisent la mise en relation entre un processus et des données du SI afin de pouvoir modifier le processus et l'optimiser. Ils évoquent l'utilisation des évènements du processus issus de l'exécution ainsi que l'utilisation des données « opérationnelles » afin de sélectionner les modifications à apporter sur le modèle du processus. Les données opérationnelles sont les données associées aux processus tel que les ressources utilisées durant une exécution, les données métiers, etc. L'outil nommé « deep Business Optimization Platform » (dBOP) agrège les « informations pertinentes », les données opérationnelles, pour adapter le processus afin de maximiser une « fonction objectif » définie par l'expert métier. Cette fonction objectif est une fonction décrivant les objectifs que doit réaliser le processus en fonction de critères. La publication propose quatre critères tels que le coût, le temps d'exécution, la qualité et la flexibilité. En fonction des métriques calculées au sein de la plate-forme, de l'analyse du modèle de processus et de la fonction objectif, l'outil sélectionne des schémas d'optimisation à appliquer. Cette plate-forme est cependant limitée à l'utilisation des données du processus et de ses données métiers, ne permettant donc pas de prendre en compte des évènements perturbateurs caractérisés par des données contextuelles. De plus, l'expertise métier n'est pas exploitée et seule la fonction objective permet de guider les axes d'optimisations sans considérer de possible contraintes métiers. Cependant cette plate-forme montre la faisabilité d'une approche de bout en bout permettant d'améliorer un processus en utilisant des données métiers.

Nous pouvons voir que plusieurs approches ont déjà été proposées pour utiliser les données du processus afin de mieux comprendre son exécution et ses performances. Cependant, aucune d'entre elles ne répond entièrement à notre question de recherche. Toutefois, nous pouvons constater que la plupart des approches se fondent sur les données métiers comme vecteur de découverte d'autres « données pertinentes » [77, 98, 99].

#### 2.2.1.3 Changements fondés sur le contexte

Le contexte peut être défini comme « toute information pouvant être utilisée pour caractériser la situation des entités qui sont considérées comme pertinentes pour l'interaction entre un utilisateur et une application, y compris l'utilisateur et la requête elle-même » [30]. Le contexte est utilisé dans divers domaines. Par exemple, l'utilisation du contexte dans la santé est souvent associée au contexte environnemental d'un patient. Il est alors composé d'éléments sur le lieu de vie, le style de vie, etc. Ces informations sont alors utilisées durant l'analyse du patient permettant ainsi un meilleur diagnostique. Dans le domaine des moteurs de recherche, le contexte correspond à l'information supplémentaire contenue dans une requête. Par exemple, le contexte d'une recherche sera l'information qui permet de choisir si le mot clé « java » se réfère au café, au langage de programmation ou à l'île. Ces informations peuvent être le profil de l'utilisateur (ses précédentes recherches [28], etc.) ainsi que les autres mots de la requête [45], par exemple « java compilation » sera associé au langage de programmation « java ».

Dans notre approche nous utilisons délibérément une définition simplifiée du contexte afin de généraliser notre approche d'analyse. Cependant, des formulations plus détaillées peuvent être trouvées dans la littérature et peuvent être utilisées pour définir formellement et représenter le contexte. Par exemple, dans [27], les auteurs proposent une approche pour caractériser le contexte d'un processus dans un domaine donné au moyen de modèles conceptuels structurés en couches. Dans [29], les auteurs proposent un méta-modèle qui peut être utilisé pour définir sans ambiguïté des indicateurs de performance de processus qui se prêtent à une analyse automatisée. Dans [108], les auteurs présentent un cadre de référence afin d'identifier, documenter, comprendre et intégrer le contexte « pertinent » aux modèles de processus métiers. Ils distinguent quatre types de contextes, illustrés par la Figure 2.2, en fonction de leur proximité par rapport au processus : le contexte immédiat, le contexte interne, le contexte externe et le contexte environnemental.

De telles définitions peuvent être utilisées conjointement avec notre approche pour améliorer davantage l'analyse des processus.

Un système contextuel [23] est donc un système amélioré par l'ajout d'informations du contexte (informations contextuelles). Dans les SI, ces systèmes permettent des fonctionnalités dites « intelligentes » telles que la personnalisation des applications à l'utilisateur afin d'améliorer la qualité du service rendu. Dans le domaine des processus métiers, « toute information reflétant un changement de circonstances au cours de l'exécution d'un processus peut être considérée comme une information contextuelle » [113]. Les auteurs de [113] présentent une modélisation du contexte s'appuyant sur le « Context Related Knowledge » dans le cadre de la gestion dynamique des rôles d'un processus. Dans ces travaux, le contexte est divisé en quatre catégories mais peut être étendu au gré des besoins :

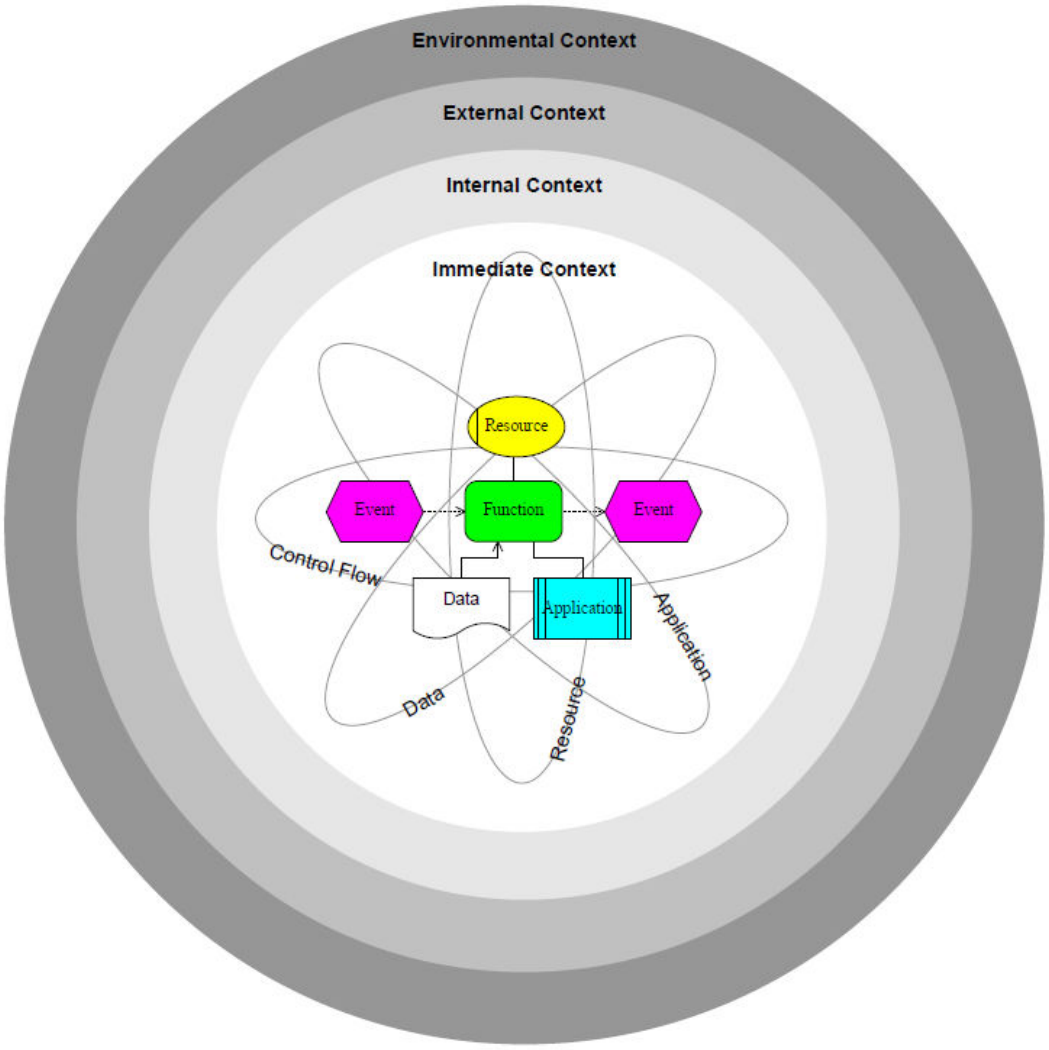


FIGURE 2.2 – Définition du contexte [108].

- Contexte lié à l'emplacement : représente les caractéristiques de l'emplacement. Par exemple, l'affectation d'un acteur à un rôle dépend de sa localisation ou encore dépend du domaine spécifique dans lequel l'acteur travaille ;
- Contexte temporel : reflète les caractéristiques liées au temps. Cela peut inclure le temps d'exécution, l'urgence, la durée du travail, la fréquence, etc ;
- Contexte lié aux ressources : décrit les propriétés des ressources matérielles et humaines. Concernant les ressources humaines cela peut être leur âge, leur sexe, leur taille ou encore leur motivation, leur performance annuelle, etc. Pour les ressources matérielles, ce contexte peut représenter la quantité disponible de ressources, leur coût de production, etc ;

- Contexte organisationnel : concerne les caractéristiques du lieu de travail, par exemple le type de structure organisationnelle (« hiérarchique », « transversale », ...), les aspects culturels et sociaux, etc.

Pour modéliser le contexte, les auteurs de [114] utilisent une ontologie. Ils se fondent sur leurs précédents travaux sur la connaissance relative au contexte,*Context Related Knowledge* CRK en anglais, afin de proposer un framework de modélisation du contexte *Context Meta-Model for Business Process Management* (CM4BPM). Ce framework distingue deux types de contexte :

- Le contexte indépendant du domaine d'activité et des processus, par exemple la disponibilité des acteurs, la durée d'exécution du processus, etc.

- Le contexte qui dépend du domaine ou du processus métier. Par exemple, dans un processus de gestion de prêt, le degré d'urgence d'une demande de prêt peut être considérée comme une information contextuelle pertinente.

Le contexte est divisée en deux ontologies reflétant la distinction des deux types de contexte. Une première ontologie générique représente les informations du contexte indépendantes aux domaines métier et peut ainsi être utilisée pour plusieurs processus. Une seconde ontologie spécifique à un domaine métier permet de capturer le détail des concepts généraux de domaines d'activité spécifiques. Par exemple, dans un processus du domaine académique, l'entité de localisation peut être caractérisée par l'attribut région. L'entité de temps peut être caractérisée par l'attribut semestre.

D'autres travaux se sont focalisés sur l'apport du contexte dans l'amélioration d'un processus. Certains ont proposé d'inclure le contexte dès la conception du processus métier [37]. L'inclusion permet ainsi d'aider l'expert métier en lui recommandant des noms d'activités pertinentes, voire l'ajout de sous-processus. Certains autres se fondent sur la définition et l'utilisation de données contextuelles afin de mieux comprendre les performances d'un processus ou encore afin de choisir quelles modifications peuvent être apportées à un processus pour l'améliorer. Hompes et al. [58] proposent une approche pour analyser les principaux indicateurs de performance du processus en tenant compte de son contexte. Ils ont introduit un framework qui aide à découvrir les différences significatives dans les résultats de performance et leurs causes. Les éléments d'un processus se voient attribuer des éléments de contexte découverts, ou calculés, à l'aide de fonctions de contexte. Les éléments contextuels d'un processus sont obtenus en appliquant une fonction de contexte à un élément du processus. Une fonction de contexte associe un élément de processus à un élément de contexte (valeur descriptive) qui vise à décrire le contexte de l'élément du processus. Par exemple, en prenant les instances d'activité comme élément et la ressource d'exécution comme contexte, nous pouvons voir comment différentes ressources

affectent les indicateurs clés de performance tels que le temps d'attente, la durée, etc. Cette technique permet d'analyser automatiquement l'effet de n'importe quel contexte sur la performance du processus aux travers de ces indicateurs. Ainsi, les causes des retards, des goulots d'étranglement et des écarts par rapport au processus, etc. peuvent être découverts. Ce framework se base sur des analyses statistiques afin de vérifier si un élément de contexte explique une différence significative de performance.

Bessai et al. [9] introduisent une méthode d'identification des faiblesses d'un modèle de processus et de déterminer les modifications à effectuer pour son amélioration. Dans un premier temps, cette méthode utilise le contexte du processus afin de découvrir sa nature. Bessai et al., se fondent sur quatre types de nature de processus : les processus de production, administratifs, collaboratifs et ad hoc [9].

Les processus de production impliquent des processus métiers répétitifs et déterministes, ce sont les processus les plus complexes des quatre, et ils permettent le bon fonctionnement de l'entreprise. Par exemple, le processus de demande de crédit pour une banque.

Les processus administratifs sont des processus bureaucratiques où les étapes à suivre sont bien établies et où il existe un ensemble de règles connues de tous les intervenants. Prenons l'exemple de l'inscription à une université, de la demande d'une carte grise, la demande d'obtention du permis de conduire. Il est à noter que la différence entre ces deux types de processus n'est parfois qu'une question de perspective.

Les processus collaboratif (PC) se caractérisent principalement par le nombre de participants impliqués, ainsi que le nombre et la nature des interactions entre eux.

Les processus ad hoc (AHP) sont des processus similaires au processus administratifs mais sont conçus pour traiter des exceptions ou des situations uniques et dépendent des utilisateurs impliqués.

Dans un second temps, la méthode aide un expert à sélectionner la modification à apporter et à modifier le modèle de processus. Les modifications choisies ne sont que celles qui sont pertinentes par rapport à la nature du processus dans un contexte particulier.

#### 2.2.1.4 Conclusion

Nous avons pu voir que la littérature était riche en méthodes permettant de prendre en compte des données du processus afin de l'améliorer. Ces données pouvant être les données manipulées par le processus, celles qu'il génère ou encore les données de contexte. En tant que tels, ces travaux nous permettraient de répondre partiellement à la problématique de cette thèse. En effet, ces travaux se focalisent sur un seul processus et non sur les interactions que peuvent avoir les processus entre eux en se basant sur le contexte. Notre approche s'inspire de ces travaux, particulièrement concernant l'utilisation du contexte au

sein d'un processus pour améliorer son analyse. Nos travaux vont plus loin en y incorporant la notion de dépendance entre processus.

La prise de décision des modifications des travaux présentés ne prend pas en compte l'expertise et l'expérience des experts métiers et ne se repose donc que sur l'analyse statistique des données. Cependant, il est relativement compliqué de pouvoir monitorer l'ensemble du contexte afin d'être capable de prendre une décision de modification tout en prenant en compte les préférences de l'expert, voire celle de l'entreprise. De plus, il est très difficile de modéliser dans le contexte les effets de compensation entre les éléments du contexte. Par exemple, si des travaux sur voies retardent des livraisons, le choix de remplacement de ces livraisons pourrait être conditionné au coût associé mais aussi à la politique d'entreprise qui souhaite minimiser l'usage du fret aérien. Dans ce cas, seul des méthodes issues du domaine de l'aide à la décision multicritère peuvent y répondre.

De plus, peu de méthodes permettent de partir d'un modèle de processus et de son contexte lors de son exécution afin d'appliquer des modifications pertinentes, c'est à dire exploiter l'ensemble de son cycle de vie. Les méthodes utilisées pour modifier le processus sont souvent laissées à l'appréciation de l'expert métier et ne l'accompagnent pas de bout en bout car cela devient très complexe. Pour corriger ce manque et réduire la complexité induite par notre approche, nous avons choisi d'utiliser l'ingénierie dirigée par les modèles pour accompagner l'expert dans la modification du processus.

### 2.2.2 Ingénierie des processus dirigée par les modèles

L'Ingénierie Dirigée par les Modèles (IdM) est une discipline du génie logiciel qui promeut les modèles en entités de première classe dans le développement logiciel [10]. L'IdM propose un nouveau paradigme par lequel il sera possible de lever le verrou de l'automatisation du développement logiciel. L'OMG (Object Management Group) et de nombreux chercheurs travaillent sur ce domaine afin d'enrichir les modèles qui sont utilisés dans la conception d'applications, à en définir de nouveaux, et à les adapter aux besoins des utilisateurs [31].

L'IdM permet de générer tout ou partie d'une application informatique à partir de modèles. Au sein de ce domaine, les modèles occupent une place prépondérante et doivent être suffisamment précis et riches d'information afin de pouvoir être interprétés ou transformés par des machines. Afin de créer un produit fini, il est nécessaire d'appliquer un ensemble de transformations à un ou plusieurs modèles où chaque transformation prend en entrée un ou plusieurs modèles et produisant un ou des modèles en sortie, jusqu'à l'obtention d'artefacts finaux. Ces transformations étant complexes, les chercheurs se sont intéressés à l'outillage permettant une simplification de la gestion des modèles ainsi qu'à la simplification de la gestion des transformations avec des langages dédiés (QVT Operationnel [72]). L'intérêt pour l'IdM a été fortement amplifié, en 2000, lorsque l'OMG a rendu publique

son initiative MDA (Model-Driven Architecture) [121] qui vise la définition d'un cadre normatif pour l'IdM. En plus de la génération automatique d'artefacts de développement, l'IdM permet notamment la validation/test des transformations effectuées [86] puisque l'automatisation de ces dernières peuvent aussi introduire un risque supplémentaire d'erreur à cause de transformations de modèles erronées, la vérification ou la traçabilité [3] des modèles résultant des transformations.

L'OMG est un consortium international de normalisation technologique sans but lucratif, à adhésion libre, fondé en 1989. Ils regroupent des industriels, des fournisseurs d'outils et des académiques dont l'objectif principal est d'établir des standards pour résoudre les problèmes d'interopérabilité entre les systèmes d'information. Ce consortium a notamment développé l'*Unified Modeling Language (UML)*, le standard *Meta-Object Facility (MOF)* et le *Model Driven Architecture (MDA)*, permettent de simplifier la conception, l'exécution et la maintenance de logiciels ou processus. Les standards de l'OMG sont centrés sur les notions de méta modèles et de méta méta modèles.

#### 2.2.2.1 *Meta-Object Facility (MOF)*

L'OMG a proposé un langage de définition de méta modèles, *Meta-Object Facility (MOF)*, afin de donner un cadre général aux utilisateurs de l'IdM. MOF est un méta formalisme c'est-à-dire un formalisme pour établir des langages de modélisation permettant eux-mêmes d'exprimer des modèles [31]. Ce méta méta modèle est le sommet de la pyramide représentant l'architecture à quatre niveaux de l'OMG (Cf. Figure 2.3). Ce dernier a la particularité de s'auto-décrire (méta circularité). L'illustration de la Figure 2.3 est composée de quatre niveaux. A gauche, ou couche M3, se trouve le méta méta modèle MOF. La couche M2, est le méta modèle en charge de la définition de la grammaire utilisée pour décrire les modèles M1 (UML, BPMN, etc.). La couche M1 est le modèle représentant le monde réel (couche M0).

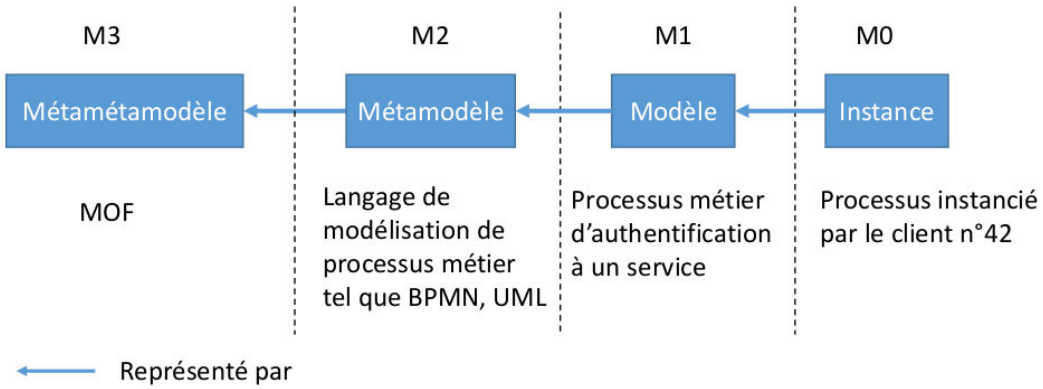


FIGURE 2.3 – Illustration de l'architecture à quatre niveaux de l'OMG.

#### 2.2.2.2 *Model Driven Architecture (MDA)*

Dans le cadre de l'IdM, l'accent est mis sur la conception et la gestion de modèles conceptuels. L'OMG a proposé en 2000 une architecture dirigée par les modèles, *Model Driven Architecture (MDA)*, proposant aux experts métiers ayant peu de connaissance informatique de pouvoir concevoir leur modèle indépendamment de l'implémentation et de la plateforme finale. L'idée principale derrière l'approche MDA est de décrire des éléments indépendants de la plate-forme pour ensuite les transformer en éléments spécifiques de la plate-forme contraints par la description de la plate-forme. MDA est une architecture structurée par plusieurs types de modèles [31] :

- modèles indépendants de l'informatisation appelés *CIM (Computational Independent Model)*. Ils modélisent les exigences d'un système afin d'aider à la compréhension du problème mais aussi de fixer un vocabulaire commun pour un domaine ;
- modèles indépendants des plates-formes appelés *PIM (Platform Independent Model)*. Ils ont pour but de décrire le savoir-faire ou la connaissance métier d'une organisation ;
- modèles décrivant une plate-forme appelés *PDM (Platform Dependant Model)*. Ils modélisent la plate-forme sur laquelle le système va être exécuté ;
- modèles générés à l'aide de transformations à partir des PIM en se basant sur les *PDM (Platform Dependant Model)* appelés *PSM (Platform Specific Model)*.

La Figure 2.4 ci-après donne une vue générale d'un processus MDA faisant apparaître les différents niveaux d'abstraction associés aux modèles.

#### 2.2.2.3 IdM et BPM

L'IdM est souvent utilisée dans le cadre du BPM afin de modéliser des processus et leurs interactions. L'auteur de [21] a travaillé sur la synchronisation de deux modèles de natures différentes car générés par deux outils différents. Le travail se place dans le cadre de la modélisation de pièces mécaniques mais la conceptualisation apportée par l'IdM permet de nous inspirer de ce travail. Le problème soulevé dans cette thèse est qu'il est difficile de travailler de manière collaborative lorsque les outils ne sont pas prévus pour. Cependant, de nos jours, de nombreux équipes de différentes entreprises doivent pouvoir travailler ensemble. C'est pourquoi Chettaoui [21] propose une méthode issue de l'IdM pour synchroniser deux modèles différents. Il propose une méthodologie de synchronisation où l'information décrite par un outil métier est un modèle qui est conforme à un méta-modèle partagé par tous. Les concepteurs utilisant les outils métiers hétérogènes pour modéliser

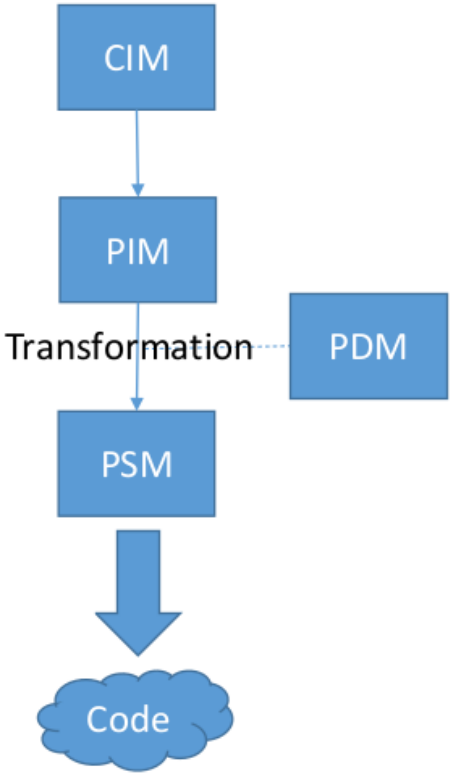


FIGURE 2.4 – Principes du Processus MDA.

leurs expertises contribuent à la conception des différentes vues d'un même produit. Les correspondances entre ces vues sont formalisées au travers de règles de transformations permettant de passer d'un modèle à un autre.

L'IdM est aussi utilisée pour la génération d'applications, ou encore pour générer des ontologies [43]. Faucher et al. [43] proposent une méthodologie de construction d'ontologies fondée sur l'annotation de modèles métier exprimés avec des diagrammes de classes UML. Cette méthodologie fournit la possibilité d'utiliser une partie des connaissances acquises sur un domaine métier pour construire une ontologie. Néanmoins le modèle UML comprend des différences sémantiques relativement importantes par rapport aux différents modèles d'ontologies. Pour résoudre ce problème, les auteurs proposent des profils UML constitués de stéréotypes et de valeurs étiquetées. Chaque élément du profil est associé à des règles de transformation visant à produire l'ontologie dans un format le plus indépendant possible. Ce format, nommé modèle « pivot », permet une navigation hiérarchique et associative entre concepts. Le passage du modèle « pivot » vers l'ontologie est fondé sur les correspondances entre les modèles de classe UML et les ontologies proposées dans la spécification *ODM (Ontology Definition Metamodel)* produite par l'OMG [25]. L'ontologie

générée sert de support à la navigation dans la base de connaissance. Leur contribution porte également sur la prise en compte des instances du modèle métier afin de peupler la base de connaissance. La transformation de modèles proposée s'applique ainsi à deux niveaux : au niveau modèle (M1), elle permet la génération de l'ontologie, au niveau des données (M0), elle permet d'utiliser les données du modèle métier pour construire la base de connaissance.

L'IdM est aussi utilisée afin de prendre en compte le contexte applicatif. Brossard et al. [17] soulignent qu'il existe de nombreuses solutions pour les usagers des transports en commun pour acquérir de l'information sur leur trajet. Cependant, malgré l'afflux toujours grandissant d'informations, il est plus intéressant pour une application de fournir une « information pertinente » afin de répondre simplement aux besoins de ces utilisateurs. Une information pertinente peut être définie comme toute information ayant un impact sur la prise de décision d'un utilisateur et/ou l'expérience utilisateur. Si la personnalisation de l'information à l'utilisateur permet de le faire jusqu'à un certain point, ce n'est pas suffisant puisque l'information peut aussi dépendre du contexte de l'utilisateur. Les auteurs proposent une solution, fondée sur l'IdM, permettant de gérer le contexte à l'intérieur des modèles conceptuels d'une application. Brossard et al. utilisent le contexte pour filtrer les informations fournies à l'utilisateur. Ils utilisent des arbres de décision lors de la phase de conception de l'application. Ces arbres de décision, qui ont été conçus pour être indépendants de la méthode de modélisation utilisée, séparent les tâches interactives et les tâches non interactives. Les arbres de décision sont conçus pour être manipulés par des non-informaticiens et pour ne fournir que les informations nécessaires à l'analyse du contexte de l'application. Ces arbres de décision permettent de générer des règles métiers permettant de modifier le modèle du processus métier sous-jacent à l'application afin de prendre compte le contexte utilisateur. Les auteurs soulignent que l'apport du contexte lors de la phase de modélisation permet une intégration plus aisée dans l'application et facilite son évolution dans le temps. Comme indiqué dans la publication, la proposition ne résout qu'une petite partie du problème de ré-ingénierie liée aux processus métier existants prenant en compte leur contexte.

Ce constat peut-être partagé avec les travaux de Popp et al. [96] qui proposent une utilisation de règles métier afin d'aligner le processus avec la stratégie de l'entreprise. Cette transformation est possible grâce à l'utilisation de l'IdM et la création de règles métier traduisant la stratégie de l'entreprise. Les auteurs de [96] proposent une méthodologie pour créer ces règles métiers de façon indépendante des processus afin qu'elles soient associables à des règles de transformations de modèles et ainsi de permettre l'automatisation de l'alignement d'un processus avec l'entreprise au travers de la transformation de modèles.

D'autres travaux ont souhaité prendre en compte le contexte dans les transformations de modèles. Simonin et al. [119] proposent d'enrichir le processus associé au MDA en y en incorporant le contexte afin de transformer un modèle PIM en un autre modèle PIM enrichi du contexte. Cette approche diffère grandement des approches usuelles en MDA puisque les transformations sont contraintes par un autre modèle, le contexte, et non simplement par un ensemble de règles de transformation. Dans cette approche, le *TCM (Transformation Contextual Model)* représente les données disponibles qu'un expert exploite pour influencer un processus. La méta modélisation du TCM doit prendre en compte la méthode de production des données correspondantes. Le méta modèle TCM décrit le mode de production sélectionné pour les données contextuelles. La Transformation Contextuelle (*CT*) permet de passer d'un modèle PIM en *PICM (Platform Independent Contextual Model)*. Cette transformation (Cf. Figure 2.5, partie de gauche) est un ensemble de règles, spécifiées par un expert métier, permettant d'enrichir le modèle PIM avec des données contextuelles à partir du modèle TCM. Le modèle PICM peut être ensuite utilisé au sein du processus usuel de MDA. Ces travaux proposent un modèle pour intégrer les données générées par des variables en tant que composants du système, c'est le contexte.

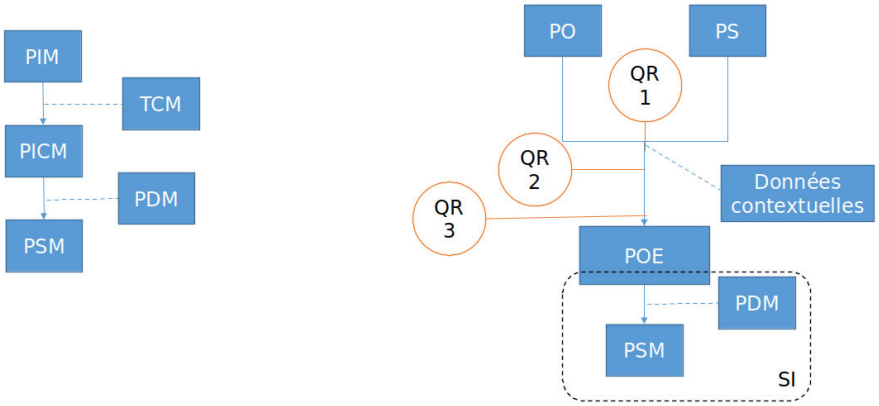


FIGURE 2.5 – Processus MDA incorporant le contexte à la transformation [119] à gauche, processus d'enrichissement d'un processus opérationnel à l'aide de processus supports et de données contextuelles à droite.

#### 2.2.2.4 Conclusion

L'IdM permet de dé complexifier un problème en le modélisant. A l'aide des outils associés, elle permet de transformer automatiquement les processus et ainsi d'aider l'expert métier.

Comme nous avons pu le voir, aucun des ces travaux ne permet de répondre entièrement à notre question, cependant chacun des ces travaux contribue partiellement à traiter notre problématique.

Notre approche se fonde sur l'IdM afin de profiter de ces propriétés permettant de répondre à notre problématique. Nous utilisons les transformations de l'IdM afin d'incorporer le contexte au sein même de la modification du processus. Par rapport à notre seconde question de recherche, consistant à permettre l'exploitation des informations contenues dans le contexte afin d'approfondir la compréhension de l'exécution des processus, l'état de l'art souligne l'intérêt du contexte dans l'analyse des processus. Cependant, le contexte est souvent exploité au niveau du SI. Dans notre contribution, le contexte sera exploité au niveau métier comme dans [119]. La Figure 2.5 montre bien la différence des deux approches et comment notre approche étend l'idée originelle afin de répondre à notre problématique. La partie de droite de cette figure montre la transformation d'un processus opérationnel (PO) et au moins un processus supports (PS) en un processus opérationnel enrichi (POE). Cette transformation est rendue possible au travers des réponses aux questions de recherche 1 et 2 et à l'aide d'une ontologie. Cette transformation va créer un nouveau processus associable à un modèle PCM à l'aide de la réponse à la question de recherche 3. Le point le plus important qu'exprime cette figure est que nous souhaitons rester au-dessus du SI, c'est-à-dire que ce travail chercher à mettre en relation les processus d'un point de vue métier et non technique. Ceci permet de laisser une grande latitude aux équipes techniques pour implémenter des services pouvant composer le nouveau processus enrichi.

## 2.3 Gestion de changement multi-processus

Nous avons pu voir l'intérêt de modifier un processus pour répondre à un besoin exprimé. Cependant, la solution générée par les méthodes mono-processus n'est pas toujours implémentable ni souhaitable par le métier. Dans les cas où modifier un seul processus n'est pas la bonne solution, il est intéressant de regarder si la modification de la relation qu'a un processus opérationnel avec d'autres processus ne permettrait pas de répondre aux besoins de l'organisation voire d'améliorer les performances du processus.

On retrouve majoritairement dans la littérature des publications concernant le lien que peuvent avoir les processus métiers avec l'organisation [118, 38] et avec le SI [110, 92]. Dans notre cas, nous nous intéressons plus particulièrement aux liens que peuvent avoir les

processus entre eux. En effet, c'est ce lien particulier qui permet de résoudre des cas métiers complexes. Il arrive donc qu'il soit nécessaire de modifier ces relations afin d'améliorer un processus. Cette problématique fait référence à notre troisième question de recherche.

### 2.3.1 Relations entre processus métier

Le premier constat auquel on aboutit lorsqu'on étudie les relations entre processus métiers provient de la définition même des processus. On peut, par exemple, revenir à la définition de processus support qui introduit une dépendance et donc une relation entre ce type de processus et les processus opérationnels. Il existe aussi des processus qui s'enchaînent les uns à la suite des autres, ou encore des processus qui s'exécutent en parallèle. En analysant ces relations, on remarque qu'elles se fondent sur l'échange de données métiers [81] ou encore sur l'échange de messages déclenchés par le processus [2]. Enfin on trouve dans le standard du langage de modélisation BPMN2.0 « les messages » permettant de modéliser des échanges d'informations entre processus.

Les processus collaboratifs [13] sont des exemples de processus métiers qui exploitent les évènements afin de synchroniser deux sous-processus. Ces sous-processus peuvent être réalisés par différentes organisations, ils sont inter-organisationnels. Les travaux présentés dans [2] utilisent les évènements pour lier des activités entre elles et ainsi composer un nouveau processus qui permettra de répondre aux besoins d'un client. Enfin, le cas le plus souvent rencontré dans la modélisation est celui où ce sont les données métiers qui sont utilisés pour créer une dépendance entre processus. Comme l'illustre les travaux de [81], cette relation est rarement modélisée. Cette relation existe lorsqu'un processus dépend d'une donnée d'un autre processus pour correctement s'exécuter ou encore s'adapter.

### 2.3.2 Schéma de mise en relation

Il existe plusieurs méthodes pour mettre en place une relation entre plusieurs processus. Les travaux que nous avons trouvés dans la littérature s'attachent à créer des liens que ce soit au niveau des activités [2, 135] ou au niveau des processus métiers.

L'orchestration de processus correspond à la mise en place d'un élément central, l'orchestrateur, qui coordonnera les autres processus. L'orchestrateur est responsable de l'invocation et de la combinaison des processus. La relation entre tous les processus participants est décrite par un seul critère d'évaluation (c'est à dire, un processus composite). L'orchestration comprend la gestion des transactions entre les différents processus. L'orchestration utilise une approche centralisée pour la composition des processus.

Il existe aussi une approche fondée sur la création d'un processus à partir de parties de plusieurs processus. Cette approche, similaire à l'orchestration, a le bénéfice de pouvoir

être utilisée dans les outils usuels de pilotage des processus. En effet, ces derniers ne sont que des processus gérant d'autres processus.

La chorégraphie entre processus est une approche où l'on synchronise des processus entre eux avec des points de synchronisation. Cette approche est notamment utilisée dans le cadre des processus inter-entreprises [13]. La chorégraphie utilise une approche décentralisée pour la composition des processus. Elle décrit les interactions entre de multiples processus, alors que l'orchestration représente le contrôle du point de vue d'une partie. Cela signifie qu'une chorégraphie diffère d'une orchestration quant à l'endroit où doit résider la logique qui contrôle les interactions entre les processus impliqués. Enfin la composition et la décomposition des processus en sous processus [135] permet de recréer dynamiquement de nouveaux processus répondant à de nouveaux besoins. On travaille à décomposer les processus pour les réutiliser et ainsi créer de nouveaux processus [134].

Ces méthodes permettent de répondre à différents cas de figures. En effet, la chorégraphie sera plus utilisée dans le cas où plusieurs organisations, ou acteurs, doivent travailler ensemble. Tandis que l'orchestration sera plutôt privilégié afin de réduire le coût d'implémentation ou afin de respecter des conventions du SI.

Au travers de notre QR3, nous cherchons à fusionner les processus entre eux en restant au niveau du processus. En effet, le fait de rester au niveau métier permet d'impliquer que les experts métier et ainsi intégrer précisément leurs besoins sans contraintes associées au SI. Pour ce faire, nous créons des portes logiques entre les processus. Cette approche peut être assimilée à de la composition de processus métier.

### 2.3.3 Conclusion

Les travaux présents dans la littérature ne font que rarement la distinction entre les types de processus métier (opérationnel, support, management). Cependant, ils ne font qu'exclusivement référence à une relation entre processus opérationnels. Nous nous sommes intéressés à la relation entre processus opérationnel et processus support qui est une relation que l'on retrouve souvent dans la réalité du fonctionnement de l'entreprise [56]. Cette relation peut être modélisée au niveau du modèle de processus mais aussi au niveau des données que ces processus doivent échanger pour fonctionner correctement [48].

Un point commun partagé par toutes les méthodes vu dans ce chapitre est la nécessité qu'un expert métier prenne la décision de réaliser une modification. Des chercheurs proposent des solutions spécifiques associées à des méthodes tel que Koschmider et al. [71], qui propose une méthode d'assistance utilisateur pour la recomposition d'un processus métier. C'est-à-dire aider un utilisateur à créer un nouveau processus à partir de bout de processus identifié afin de répondre à un nouveau besoin. Cependant, il existe des solutions plus génériques au travers de l'aide multicritère à la décision.

## 2.4 Application de l'aide à la décision multicritère au sein du domaine du BPM

La nécessité de laisser l'expert métier décider est une limitation courante dans les travaux présents de la littérature. Notre approche se démarque en proposant de l'assister dans ces décisions et de prendre en compte ses préférences. Notre troisième question de recherche (QR3) soulève ce besoin d'incorporer l'expert métier dans le processus de prise de décision. Cette décision permet d'optimiser le déroulement d'un processus opérationnel.

Une revue de la littérature sur la prise de décision et sa relation avec les processus opérationnels a montré une montée d'intérêt pour le traitement de la décision. Les recherches dans ce domaine sont axées sur la collecte, la modélisation et l'intégration de règles métier dans les processus métiers. Ce travail a permis de concevoir le langage Decision Model and Notation (DMN) qui est devenu un standard en matière de modélisation de la prise de décision dans les processus métiers. Plusieurs logiciels libres et propriétaires sont apparus afin de mettre en œuvre les normes BPMN et DMN en s'appuyant sur les langages BPEL (*Business Process Execution Language*, [4]) et FEEL (*Friendly Enough Expression Language*, [11]).

Nous pouvons remarquer l'absence de préférences métier conditionnant les critères de décision définis dans le DMN. En effet, les tables de décision ne traitent que de décisions prédéfinies prises à partir de critères connus et ne permettent pas de recourir au concept de pondération pour traiter la priorité des règles métier. Cette pondération peut venir des préférences d'un décideur, un expert métier par exemple, ou plus généralement de la stratégie de l'entreprise.

De plus, la prise de décision devient une tâche difficile dans l'un des cas suivants :

- Un grand nombre de possibilités de comparaison, dans ce cas, il s'agit d'une optimisation combinatoire ;
- Un nombre significatif de décideurs impliquant donc le recours aux compromis et à la théorie des jeux ;
- Plusieurs critères à prendre en compte amenant donc à la prise de décision fondée sur de multiples critères.

La prise de décision dans les processus métiers est généralement confrontée à la dernière difficulté, à savoir la prise de décision selon plusieurs critères.

L'aide multi-critère à la décision (AMCD) est une branche de la recherche opérationnelle permettant de choisir la meilleure solution ou la solution optimale parmi tout un ensemble de solutions en fonction de critères de préférence d'un décideur.

Dans le cadre de l'AMCD, nous identifions généralement au moins un décideur, nommé DM, qui est responsable de la décision à prendre. Il est confronté à plusieurs solutions possibles, que l'on appelle des alternatives de décision, qui sont évaluées sur un ensemble de critères ou de points de vue, parfois contradictoires. Le DM exprime certaines préférences relatives à ces alternatives et critères, qui sont généralement utilisés comme paramètres par les algorithmes d'AMCD afin de fournir une solution au problème de décision. Le type de résultats des algorithmes de l'AMCD, nommée problématique de décision, peut être de type « choix » (déterminer la « meilleure » alternative parmi toutes), « classer » (classer les alternatives du « meilleur » au « pire ») ou « trier » (assigner les alternatives aux classes prédéfinies et ordonnées). Pour illustrer ces concepts, présentons un court exemple. Un comité d'école est chargé d'attribuer un nombre fixe de bourses aux étudiants en fonction de leurs performances sur les sujets qui leur sont enseignés (par exemple, les mathématiques, l'informatique, la biologie, etc.). Dans ce cas, le comité d'école est le DM, les étudiants sont les alternatives de décision, tandis que les sujets représentent les critères. La problématique de décision, dans ce cas, est de classer tous les élèves du meilleur au pire (problème de classement) et de sélectionner les meilleurs étudiants en tant que bénéficiaires d'une bourse. Ce classement doit être effectué en fonction des préférences du comité d'école.

L'AMCD a été appliquée à de nombreux domaines, tels que la santé [131], la finance et les banques [44], la gestion de l'environnement [73], la planification urbaine en utilisant des systèmes d'information géographique [26], la robotique [123], la planification énergétique [44], la gestion des urgences nucléaires [94], la sélection des équipements, [57] etc. Le processus d'aide à la décision est souvent complexe, en fonction du domaine d'application et des préférences du décideur. En conséquence, de nombreux algorithmes d'aide multi-critères à la décision ont été développés au fil des ans ([50, 65]).

Comme mentionné, les décisions et les objets qui les concernent peuvent être très diverses. En fait, chacun de nous est confronté à une multitude de décisions chaque jour : quel itinéraire prendre pour se rendre au travail le matin, que choisir comme déjeuner, etc. De nombreux facteurs influencent ces décisions, telles que nos références, nos expériences antérieures, nos différentes contraintes, etc. Dans certains cas, il peut être difficile d'établir un équilibre entre ces facteurs. L'aide multi-critères à la décision sert de guide décisionnel dans la prise de décision lorsque plusieurs critères souvent contradictoires sont impliqués. Le processus commence généralement par définir le problème, les objectifs des DM et la manière dont la décision finale doit être prise [15]. L'un des aspects essentiels de l'aide multi-critères à la décision est que la décision finale n'a peut-être pas besoin d'être la meilleure possible, mais qu'elle doit être acceptable pour toutes les parties prenantes. Par conséquent, lorsque plusieurs DM sont impliqués, des conflits doivent être traités afin de parvenir à un consensus sur la décision finale. Le terme « cadre de décision », utilisé

par Tversky et Kahneman [126], corrobore le fait que les décideurs fondent souvent leur décision sur des jugements subjectifs. De plus, les problèmes de décision du monde réel sont devenus de plus en plus complexes.

En effet, la prise de décision au sein de processus métier est soumis à de multiples facteurs tels que, par exemple, les différents composants de la stratégie de l'entreprise ainsi que les préférences des décideurs conditionnés par leur expertise métier. C'est avec ce dernier facteur que l'on peut dire qu'une table de décision n'est pas suffisante pour retranscrire la complexité métier. Dans notre cas, la prise de décision qualifiant la gravité d'une anomalie survenue lors du déroulement d'un processus opérationnel ainsi que triant les solutions possibles est fondée sur ces deux facteurs (stratégie de l'entreprise et préférences de l'expert métier).

## 2.5 Conclusion

Suite à cette revue de la littérature, il est possible de reformuler notre contribution en se basant sur le travail de [119].

La contribution de cette thèse permet de transformer un processus opérationnel en un processus opérationnel enrichi dans lequel des liens ont été créés vers des processus supports. Cette solution permet d'améliorer la performance du processus opérationnel. Elle consiste à transformer un processus opérationnel et un processus support (PIM) grâce aux données contextuelles (TCM). Cette transformation va produire un processus opérationnel enrichi (PICM) qui sera transformé en service dans le système d'information (PCM), le PDM représentant la contrainte de la plate-forme technique. Cette formulation montre bien que l'on travaille au niveau métier (niveau du processus) et non du SI. Ceci nous différencie des travaux présentés dans ce chapitre qui ne proposent pas de solution pour améliorer un processus en prenant en compte à la fois les processus supports (parfois appelées ad hoc) et le contexte. Cette correspondance est illustrée par le tableau 2.1.

Table 2.1 – Tableau de correspondance entre l'approche proposée par [119] et notre approche.

| **Approche de** [119] | **Notre approche** |
|---|---|
| Platform Independant Model | Processus Opérationnel (PO) et Processus Supports (PS) |
| Transformation Contextual Model | Données Contextuelles |
| Platform Independent Contextual Model | Processus Opérationnel Enrichi (POE) |
| Platform Description Model | PDM |
| Platform Specific Model | PSM |

Partant de ces observations, nous avons conçu et développé un outil permettant l'enrichissement d'un processus opérationnel par des processus supports à l'aide des données contextuelles. Dans le chapitre suivant nous allons présenter les techniques sous-jacentes de cet outil.

Chapitre 3

# Introduction de la contribution et définitions

La problématique de cette thèse peut être reformulée comme étant : « peut-on enrichir (QR3) le modèle d'un processus opérationnel en le mettant en relation (QR2) avec des modèles de processus supports suivant le contexte d'exécution (QR1) » ?

Cette formulation introduit l'idée de comprendre le contexte d'exécution afin de pouvoir mettre en relation le modèle d'un processus opérationnel et ceux des processus supports qui pourraient lui être associés. Cette vue enrichie du processus opérationnel permet aux experts métiers d'améliorer le déroulement du processus opérationnel en faisant appel aux processus supports utiles.

Le processus métier (Cf. Définition 1) est utilisé au sein des organisations comme un support numérique permettant de décrire un « métier » par une séquence temporelle d'activités.

Nous choisissons de définir le processus métier en intégrant les données métiers associées à celui-ci.

**Définition 1** (Processus métier)**.** *Un processus métier est défini comme un quadruplet $P = (A, S, \gamma, DM)$ où $P$ est un ensemble d'objets qui peuvent être divisés en ensembles disjoints d'activités $A$, de données métiers $DM$, de gateways $\gamma$, de séquence $S$ entre deux activités où $S \subseteq A \times DM \times A$, ou entre une activité et une gateway. Cette définition est illustrée par la Figure 3.1.*

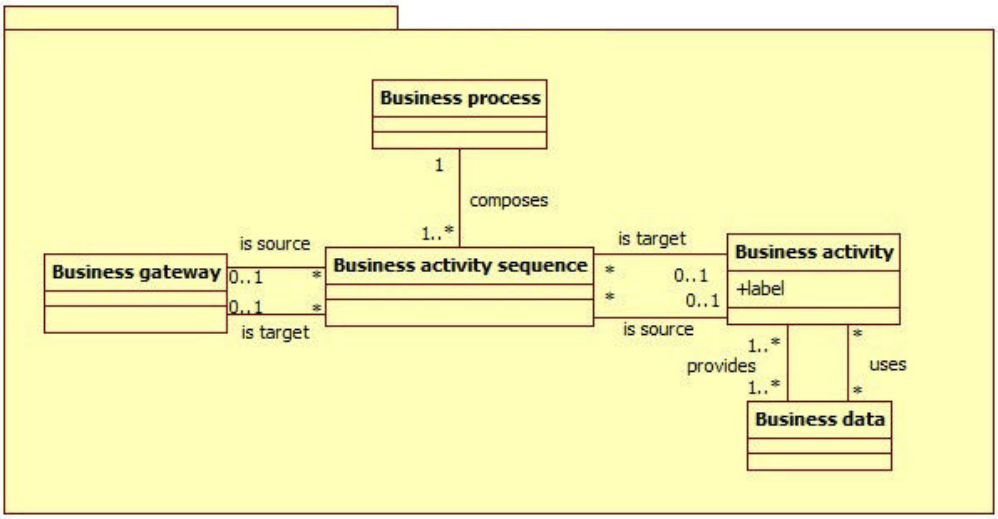


FIGURE 3.1 – Représentation à l'aide d'un méta modèle d'un processus métier. Ce méta-modèle est associé au niveau M1 du MOF défini dans le Section 6.1.1.6.

Les données métiers associées à un processus métier se déclinent en données opérationnelles lorsque celle-ci sont associées à un processus opérationnel et en données supports lorsque ces dernières sont associées à un processus support.

Afin d'illustrer la démarche mise en place pour répondre à la problématique de cette thèse, nous allons nous appuyer sur trois processus métiers dont un processus opérationnel et deux processus supports. Le processus opérationnel modélise l'accès à un service d'un opérateur téléphonique réalisé par trois entités que sont le client (*Customer*), le gestionnaire d'identité (*Identity manager*) et le gestionnaire du service (*Service manager*). Suite à la demande d'accès d'un client à un service (évènement « Wish to access to a service »), deux possibilités s'offrent au gestionnaire d'identité : (1) entrer les identifiants de client (« Entering crdentiels ») ou (2) mettre en service le terminal (« Turning on device »). Il est utile de noter que le service est connu en amont du processus. Une fois l'une de ces deux possibilités réalisée, le fournisseur de service peut créer une session (« Creating authentification session »). Une fois la session créée le client reçoit une notification de cette création (« service access autotized »).

L'un des processus supports est le processus de gestion d'incident affectant une ressource (terminal, livebox, etc.). Suite à la détection d'une anomalie par le fournisseur de ressource (évènement « Resource trouble »), le gestionnaire de ressource (« Resource manager ») crée un rapport d'anomalie affectant la ressource. Suite à cela, il relève l'anomalie « Correct and resolve resource trouble » et ainsi le fournisseur de ressource reçoit un acquittement (« Resource replaced or restored »).

L'autre processus support est le processus de gestion d'incident affectant un service. Suite à la détection d'une anomalie par le fournisseur de service (évènement « Service trouble »), le gestionnaire de service (« Service manager ») crée un rapport d'anomalie affectant le service. Suite à cela, il relève l'anomalie « Correct and resolve service trouble » et le fournisseur de service reçoit un acquittement (« Service restored »).

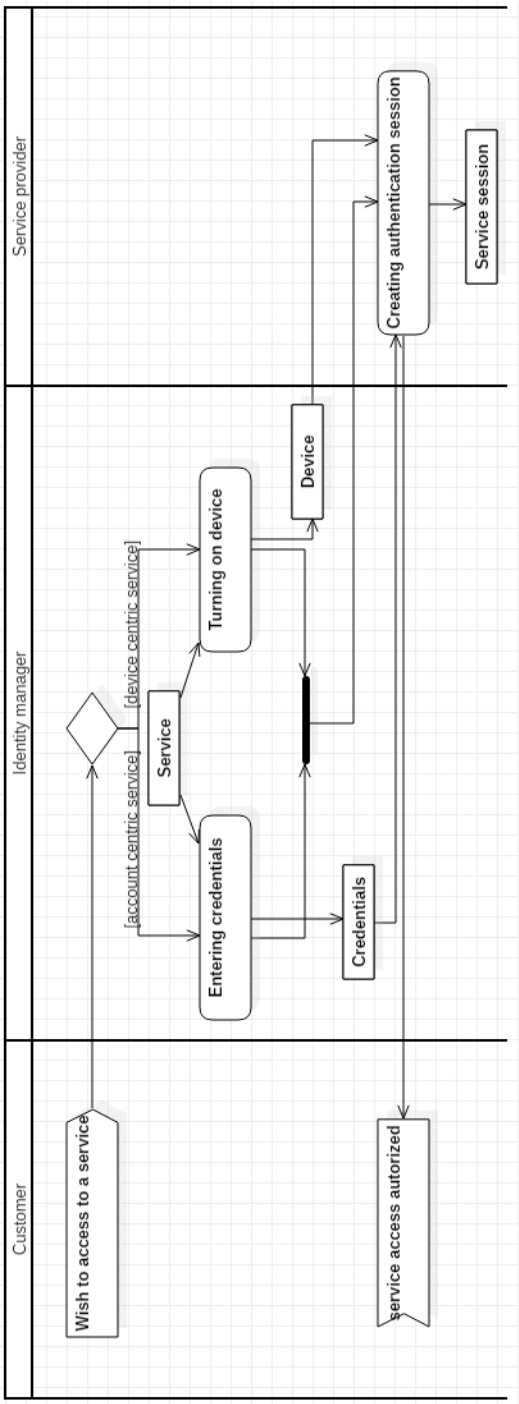


FIGURE 3.2 – Modèle d'un processus opérationnel d'accès à un service.

Nous utiliserons cet illustration afin d'expliciter les concepts de cette thèse et les appliquer dans un cas concret. Ces modèles sont inspirés de [40] écrit par E. Bertin et du

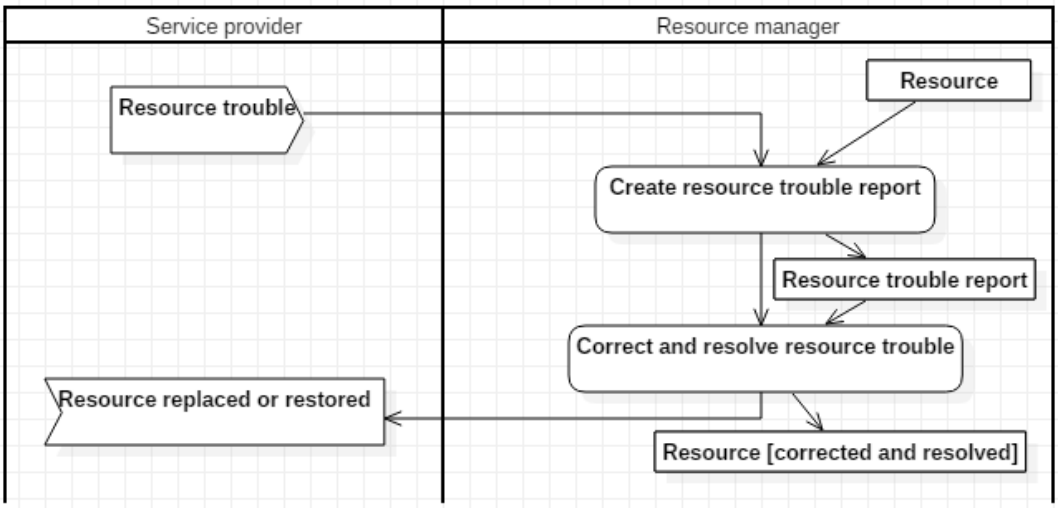


FIGURE 3.3 – Modèle d'un processus support de gestion d'incident d'une ressource.

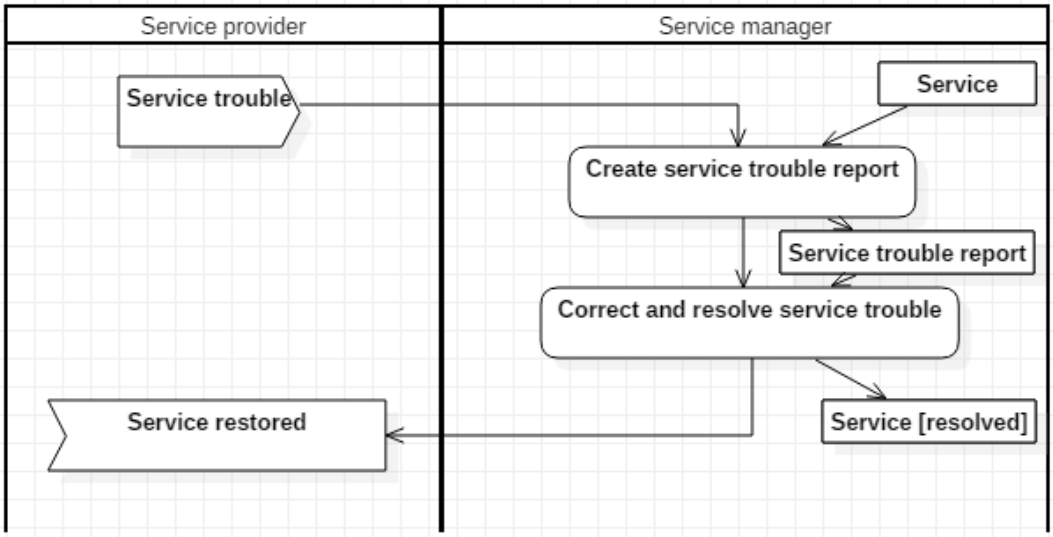


FIGURE 3.4 – Modèle d'un processus support de gestion d'incident d'un service.

framework eTom du TMForum [1].

En se référant au processus opérationnel de notre exemple (Cf. Figure 3.2) et à la définition 1, le processus est composé de plusieurs séquences d'activités. Ces séquences d'activités peuvent contenir une gateway (ou porte logique) comme, par exemple, dans la séquence « Wish to access to a service » vers « Turning on device ». Ces activités sont chacune associées à des données métiers (produites ou utilisées) telles que le type d'identification fondé sur un compte utilisateur ou sur un terminal.

1. https ://www.tmforum.org/business-process-framework/

Nous proposons une méta-modèlisation nécessaire à l'approche IdM que nous avons choisie. Le premier méta-modèle (Cf. Figure 3.1) cible la définition du processus métier. Les concepts sont illustrés ci-après à partir de la Figure 3.2 :

- Business process : Processus opérationnel d'accès à un service ;
- Business activity : « Turning on device », « Creating authentification session » ;
- Business activity sequence : Séquence ayant pour source « Turning on device » et cible « Creating authentification session » ;
- Business data : « Device » est produite par « Turning on device » et utilisé par « Creating authentification session » ;
- Gateway : Ayant pour cible « Turning on device » et succédant à l'évènement « Wish to access to a service ».

L'exécution d'un processus peut-être associé à une liste de logs tel que chaque log indique la date de début et la date de fin d'une activité. Ce sont des informations qui sont analysées habituellement afin de comprendre comment un processus s'est exécuté (*process mining*).

**Définition 2** (Logs)**.** *Les journaux d'évènements, ou fichiers de logs, sont des fichiers enregistrant de manière chronologique les évènements, aussi appelés logs, générés par des applications. Soit E l'ensemble des évènements $\epsilon$ contenus dans les logs. Usuellement les attributs suivants sont présents pour tous les évènements : activity($\epsilon$), time($\epsilon$), name($\epsilon$), id_instance($\epsilon$). Soit id_instance($\epsilon$), l'identifiant unique permettant de relier $\epsilon$ à une instance du processus P. Soit activity($\epsilon$), l'activité en lien avec $\epsilon$, soit time($\epsilon$), le timestamp d'$\epsilon$ enfin soit name($\epsilon$), le nom de $\epsilon$, $\forall name \in Name$. Un événement peut également contenir d'autres informations telles que le coût, la ressource, les données métiers manipulées, etc.*

**Définition 3** (Trace)**.** *Une trace est la retranscription d'une instance et d'une perspective d'un processus au sein des logs E. Soit $instance(P)_k$, $\forall k \in \mathbb{N}$ la kème instance de P.*

*La trace $\lambda_i$, $\forall i \in |instance(P)|$, est définie comme l'ensemble des $\epsilon$ où $trace(\epsilon) = i$. Posons $\Lambda$ l'ensemble des traces $\lambda$.*

**Définition 4** (Perspective)**.** *Une perspective est la projection d'un processus en fonction d'une ou plusieurs caractéristiques de celui-ci. Une projection peut se faire en se fondant sur les rôles liés au processus, sur les temps d'exécution des instances, sur l'enchaînement des activités, etc. Soit $\zeta_{time}(P)$, la projection de P par rapport au temps d'exécution. $\zeta_{time}(P)$ est une matrice de taille $|instance(P)| \times |A|$ composée des temps d'exécution des*

*instances et des activités correspondantes. En process mining, les données majoritairement utilisées sont les traces de processus et la perspective de l'enchaînement des activités est la plus communément utilisée, et est souvent nommée « chemin ».*

On peut utiliser les perspectives des activités d'un processus afin de calculer la similarité de deux instances. Cette similarité nous permet d'appliquer des algorithmes de partitionnement de traces afin de grouper les traces ayant une similarité forte.

**Définition 5** (Similarité d'une trace)**.** $M(E) = (E \times E) \rightarrow [0.0, 1.0]$ *désigne l'ensemble de toutes les matrices de similarité de traces possibles dans* $E$*.* $\forall(\lambda, \lambda') \in \Lambda, M(\lambda, \lambda')$ *désigne la similarité entre* $\lambda$ *et* $\lambda'$*. La similarité est calculée à l'aide d'une fonction qui transcrit une trace en un vecteur de longueur* $m$ *par rapport à une perspective* $p$*. Ce vecteur est noté* $\sigma_p(\lambda)$*.* $\forall i, j \in |instance(P)|, M(\lambda_i, \lambda_j) = \sigma_p(\lambda_i) - \sigma_p(\lambda_j)$*.*

Cette similarité entre traces permet de regrouper les traces similaires en se fondant sur les chemins afin d'avoir une représentation des « comportements » des utilisateurs du processus, c'est-à-dire, des différents types d'enchaînement d'activités.

**Définition 6** (Partitionnement de trace (*Trace clustering*))**.** *Soit* $E$ *l'ensemble des évènements du processus* $P$*. Un cluster de traces* $tc$*, aussi nommé groupe ou trace class, dans* $E$ *est un sous-ensemble de* $instance(P)$*.* $tc$ *est composé d'un ensemble de traces* $\lambda$ *similaires. Le partitionnement de traces* $TC \subseteq instance(P)$ *est un ensemble de clusters de traces dans* $E$*. Nous posons l'hypothèse que toutes les traces* $\lambda$ *font partie d'au moins un cluster, et qu'une trace ne peut pas être dans deux clusters différents tels que* $\cup\ tc = instance(P)$*.*

Un processus métier est amené à s'exécuter dans divers contextes. Un contexte peut être défini comme un ensemble d'éléments interagissant directement ou indirectement avec un processus métier (Cf. Définition 7) ; par exemple, le nombre de ressources mises en place pour l'exécution du processus, l'état de ces ressources ou encore le taux de disponibilité de celles-ci. Ces éléments influencent l'exécution du processus opérationnel en améliorant sa performance (diminution du taux d'échec, du temps d'exécution, etc.) ou au contraire en réduisant sa performance (diminution de la satisfaction client, augmentation du coût d'une exécution, etc.).

Le contexte d'un processus est une notion complexe, et fait référence à divers moments du cycle de vie du processus tels que le contexte d'exécution [133], de modélisation [113] ou encore d'analyse du processus [58]. Cette notion de contexte est aussi étudiée dans d'autres domaines de recherche, tel que dans la santé [129], les systèmes de recommandation, ou encore dans l'intelligence ambiante [47].

Dans le domaine des processus, le contexte est principalement utilisé pour améliorer la compréhension du déroulement d'un processus [113] et améliorer son analyse [77, 53].

Malheureusement, comme ces travaux l'ont souligné, une partie importante des informations manquantes cible ce contexte. Ce manque d'informations peut conduire à prendre des décisions avec des résultats limités, voire même contre-productifs.

L'avantage d'obtenir des informations sur le contexte est de permettre à l'expert métier de mieux comprendre le déroulement d'un processus métier. Le contexte peut être défini comme étant « toute information pouvant être utilisée pour caractériser la situation des entités (c'est-à-dire une personne, un lieu ou un objet) qui sont considérées comme pertinentes pour l'interaction entre un utilisateur et une application, y compris l'utilisateur et la demande elle-même » [30]. Cette définition est orientée sur l'intégration de la gestion du contexte au sein des applications du SI. Dans le cadre des processus métiers, les auteurs de [107] définissent le contexte comme étant « le minimum de variables contenant toutes les informations pertinentes qui ont une incidence sur la conception et l'exécution d'un processus métier ». D'autres auteurs [58] proposent de considérer le contexte comme un ensemble d'éléments obtenus à l'aide de « fonctions de contexte établissant une correspondance entre une instance de processus vers une étiquette de contexte (valeur descriptive) qui vise à décrire le contexte de l'entité ». On appelle une instance de processus son exécution. Il existe donc autant d'instances d'un processus qu'il y a eu d'exécutions du processus. Cette dernière définition permet de considérer le contexte comme un vecteur de données relatives à une instance de processus contrairement à la définition proposée dans [107] qui ne prend pas en compte l'influence de celui-ci sur le processus.

L'étude des données (nommées variables dans [107]) composant le contexte permet de mieux comprendre la relation qui existe entre le contexte d'exécution d'un processus et sa performance. La compréhension de cette relation sert aux experts métiers à améliorer le processus en alignant son modèle avec le modèle du contexte d'exécution. Comme le souligne [58], l'ajout de variables définissant le contexte permet un gain d'information. Tel qu'illustré par la Figure 3.5, ce gain permet de comprendre la corrélation entre les variables de contexte et la performance d'un processus. En effet, étudier seulement la distribution de la performance d'un processus fournit peu d'informations sur les raisons d'une sur-performance ou sous-performance. La Figure 3.5 montre l'intérêt d'étudier les performances en fonction du contexte pour fournir plus d'informations à un expert métier.

Un autre point digne d'intérêt à propos de ces éléments de contexte est leur impact sur le déroulement d'un processus. Prenons l'exemple des ressources (matérielles, humaines, etc.) associées à un processus, connaître leurs états, leur taux de disponibilité, leur temps de réponse durant l'exécution d'une instance du processus va permettre de comprendre les variations de performance qu'il rencontre. Ces informations sont traitées et gérées par des systèmes de monitoring accessibles au travers du SI de l'organisation. Nous trouvons donc un certain nombre de processus supports produisant et manipulant les données associées

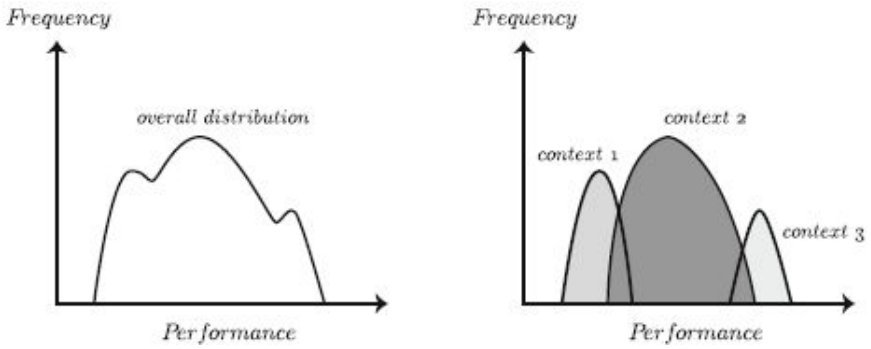


FIGURE 3.5 – Gain de l'ajout du contexte dans l'analyse de la performance [58].

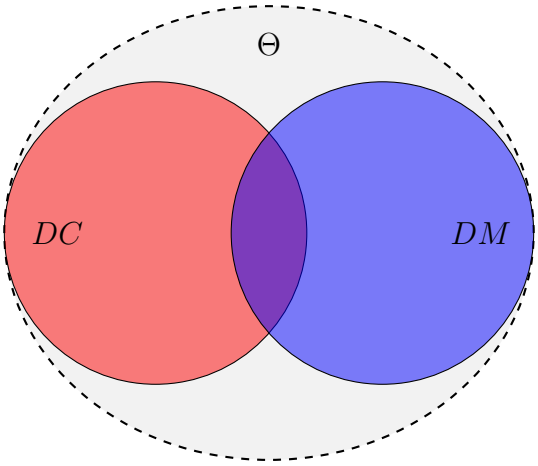


FIGURE 3.6 – Illustration du contexte $DC$, des données métiers $DM$ au sein de l'ensemble $\Theta$.

aux ressources (exemple : processus de relance de service après incident, etc.).

**Définition 7** (Données contextuelles)**.** *Soit $\Theta$ l'ensemble des données, $DC \subseteq \Theta$ l'ensemble de données contextuelles (aussi appelé variables de contexte, ou éléments de contexte), $DM \subseteq \Theta$ l'ensemble de données métiers. $P$ un processus composé de plusieurs activités $A$. On notera $DM_P \cup DC_P \subseteq \Theta_P$, les données relatives au processus $P$. $DC_P(A_i) = [DC_{P,A_i}(0), .., DC_{P,A_i}(j)]$ $\forall j \in |DC_P(A_i)|$ et $\forall i \in |A|$, un vecteur composé de données contextuelles relatives au processus $P$ et à l'activité $A_i$. Chaque donnée $DC$ fait référence au libellé d'une donnée contextuelle.*

*Soit $DC_{instance(P)_k}(A_i) = [DC_{instance(P)_k,A_i}(0), .., DC_{instance(P)_k,A_i}(j)]$ $\forall j \in |DC_{instance(P)_k}(A_i)|$ et $\forall i \in |A|$. Chaque instance de $DC$ fait référence à la valeur d'une donnée contextuelle. 3.6*

Au sein de l'ensemble des données contextuelles, certaines influencent l'exécution du processus métier.

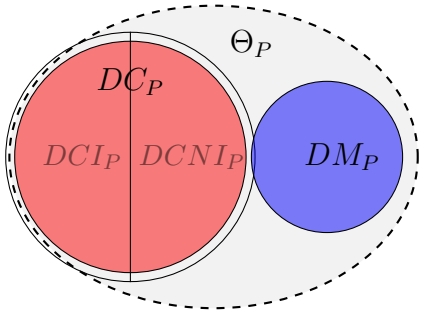


FIGURE 3.7 – Illustration des ensembles $DCI_P$ et $DCNI_P$ au sein de $\Theta_P$.

**Définition 8** (Données contextuelles influençant l'exécution d'un processus)**.** *Au sein de $DC$, l'ensemble des données contextuelles, il existe deux sous-ensembles $DCI$ et $DCNI$ tels que $DCI \cup DCNI = DC$ avec $DCI_P$ l'ensemble des données qui influencent l'exécution d'un processus $P$ et $DCNI_P$ l'ensemble des données qui n'influencent pas l'exécution d'un processus $P$. Soit une donnée $d \in DC$, $\theta_d$ est l'ensemble des valeurs possibles de $d$. $d \in DCI$ seulement s'il existe deux sous-ensembles $\alpha$ et $\beta \in \theta_d$ tels que $tc_i|\alpha \neq tc_j|\beta,\ \forall i \neq j,\ i, j \in |TC|$. Cette définition est illustrée par la Figure 3.7.*

La présentation de notre contribution est divisée en trois chapitres. Dans le premier chapitre, nous allons expliciter la relation existant entre un processus opérationnel, ou un processus support, et son contexte d'exécution. Dans le second chapitre, nous verrons comment le contexte nous permet de lier des processus supports à un processus opérationnel afin d'enrichir ce dernier. Enfin, dans le troisième chapitre, nous présentons le développement de l'outil d'enrichissement.

Chapitre 4

# Conception sémantique des données contextuelles d'un processus opérationnel et support

Dans ce chapitre, nous allons présenter une méthode permettant de réaliser la conception sémantique du contexte d'exécution relié à un processus opérationnel ou à un processus support. Le chapitre fait référence à la première question de recherche (QR1) sur les données associées aux processus et répond en partie à la seconde (QR2) sur l'utilisation du contexte dans l'analyse du processus opérationnel. Nous allons appliquer une méthodologie en cinq étapes (Cf. Figure 4.1) faisant référence au cycle de vie d'un processus métier. Les deux premières étapes concernent la modélisation du processus, l'étape suivante concerne son exécution alors que les deux dernières concernent l'analyse de cette exécution. Contrairement aux approches usuelles d'analyse de processus, notre approche utilise le modèle du processus comme source d'information.

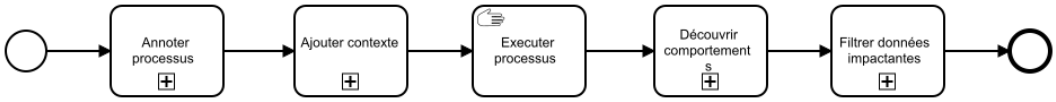


Figure 4.1 – Étapes de la prise en compte du contexte d'exécution à partir d'un modèle de processus.

## 4.1 Données contextuelles d'un processus métier

Dans cette section, nous proposons une démarche de conception des données contextuelles *DC* en s'appuyant sur l'expertise métier.

Notre approche consiste à associer des concepts décrit dans une ontologie, représentant un domaine métier, à un processus métier [78]. Cette association, aussi nommé annotation sémantique, cible une donnée métier ou une activité métier et un concept de l'ontologie. Ceci permet de définir des données contextuelles à partir de concepts de l'ontologie « proches » du concepts décrit dans l'annotation. Dans notre exemple, le concept ontologique « Device » est associé à l'activité de « Turning on device ». Lors de l'exécution, le

système récupère l'état du terminal, ou encore le niveau de batterie restant. Même si cette illustration semble simple, il est cependant nécessaire d'assister l'expert lors de l'annotation du processus à l'aide de l'ontologie afin de l'aider à trouver des concepts qui ne sont pas trivialement liés au processus.

Ce contexte d'exécution peut être explicité à l'aide de quatre types de contexte [113] :

- le contexte lié à ***l'organisation de l'entreprise***, les caractéristiques du lieu de travail, l'aspect social ;
- le contexte lié à ***l'emplacement***, la localisation de réalisation d'une activité ;
- le contexte lié au ***temps*** tel que le moment où une activité est terminée ;
- le contexte lié aux ***ressources*** reflétant les caractéristiques liées aux services, aux matériels utilisés et aux aspects humains tels que l'âge ou les compétences.

L'intérêt de cette définition du contexte est de permettre sa formalisation. Les auteurs de [114] ont étendu leur travail en proposant un méta-modèle de contexte limité à la modélisation d'un processus et sans prendre en compte sa spécificité (opérationnel, support, management). Ce méta-modèle nommé CM4BPM est fondé sur l'utilisation de deux ontologies. La première décrit les caractéristiques générales du contexte commun à tous les domaines d'activité.

La seconde ontologie est, quant à elle, spécifique à un domaine métier et décrit les caractéristiques de ce domaine.

Avec CM4BPM, les auteurs proposent une solution permettant d'appliquer un raisonnement sur le contexte en utilisant des règles d'inférence ou des règles définies par un expert métier permettant d'interpréter le contexte et d'en déduire d'autres informations contextuelles. Ces travaux de spécification du contexte de modélisation peuvent être appliqués dans le cadre de l'exécution d'un processus et ainsi permettre de l'éliciter.

De façon similaire, notre approche consiste à réaliser une analyse sémantique du modèle d'un processus métier afin d'en extraire les informations utiles permettant de sélectionner les données contextuelles, et ainsi, fournir un lien avec les processus supports. Cette méthode nécessite deux pré-requis. (1) Le premier pré-requis est la création d'une ontologie relative à un domaine métier [33], par exemple la logistique ou la relation client. Cette ontologie devra être réalisée de sorte qu'elle fasse ressortir les relations d'influence qui peuvent exister entre des concepts métiers et des concepts de données contextuelles. (2) Le second pré-requis est la mise à disposition d'interfaces permettant de récupérer les instances des données contextuelles décrites au sein de l'ontologie.

### 4.1.1 Annotation d'un processus métier avec une ontologie

Dans cette partie, nous décrivons une démarche d'annotation sémantique d'un processus métier puis celle de la conception des données contextuelles. L'annotation d'un processus se fonde sur l'utilisation d'une ontologie conçue par les experts métiers. Elle permet de représenter les liens d'influence entre d'une part les activités d'un processus et d'autre part le contexte de son déroulement. L'utilisation d'une ontologie permet d'évaluer la proximité des concepts d'un domaine métier. Plus simplement, on peut dire que l'« ontologie est aux données ce que la grammaire est au langage » (Wikipedia, 2018).

#### 4.1.1.1 Ontologie

Les ontologies [33, 79] sont la représentation d'une connaissance, par exemple, sous la forme d'un graphe (Cf. Définition 9). Elles sont utilisées dans de nombreux domaines tels que l'informatique, la bio-informatique, le web sémantique, etc. Dans le cadre de nos travaux, nous nous concentrons sur une ontologie relative à un domaine métier afin d'annoter le modèle de processus. Pour réaliser cette ontologie, nous nous reposons sur les quatre types de contexte vus précédemment afin de typer chaque concept de l'ontologie.

**Définition 9** (Ontologie)**.** *Soit $O$ une ontologie définie comme la paire $(C, R)$ où $C$ est un ensemble fini de nœuds représentant des concepts dans $O$, $R \subseteq C \times C$ est un ensemble fini d'arêtes représentant des relations entre les concepts de $C$. Nous pouvons définir le Path comme étant le chemin $C \times R$ reliant deux concepts $C$.*

Cette définition est modélisée par la Figure 4.2 et illustrée par notre exemple par la Figure 4.3.

La Figure 4.3 instantie le méta-modèle de la Figure 4.2. On y retrouve six « Ontology concepts » (Device, Synchronisation, Resource trouble report, Service, Timeout, Service trouble report) ainsi que les relations qui les relient, les « Ontology relation ».

Dans notre illustration, les relations entre les concepts permettent d'exprimer le fait que lorsqu'un timeout impacte un service, un ticket d'anomalie concernant le service y est associé. De même, lorsqu'un timeout ou un défaut de synchronisation affecte un terminal, un ticket d'anomalie y est associé.

L'annotation sémantique d'un processus avec des concepts d'une ontologie est une solution permettant d'enrichir le modèle d'un processus avec des informations complémentaires. C'est une tâche du domaine de la fouille de texte proche des méthodes de traitement automatique du langage naturel (TALN) [24]. Elle consiste à étiqueter les mots avec des liens qui pointent vers une description sémantique (concept de l'ontologie). Outre l'insertion de contenus qui permettent de compléter un texte (par exemple en identifiant

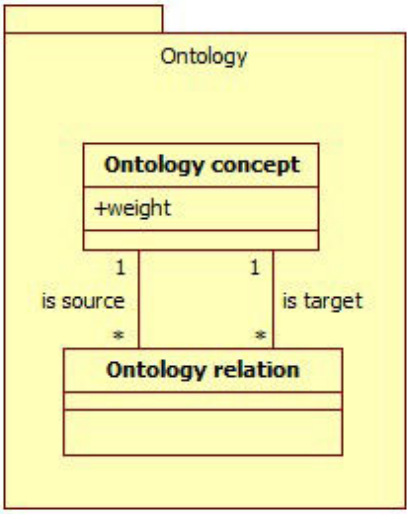


FIGURE 4.2 – Représentation à l'aide d'un méta-modèle d'une ontologie

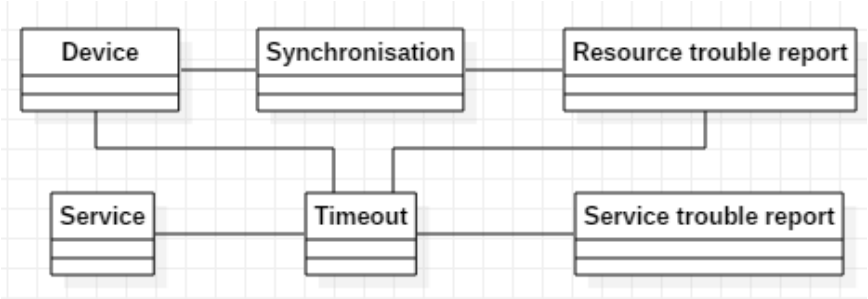


FIGURE 4.3 – Illustration d'une ontologie

automatiquement les noms de personnes et en fournissant leur date de naissance), l'annotation sémantique joue un rôle essentiel en désambiguïsation suite à la fourniture d'un identifiant unique représentant parfaitement l'identité du mot ou du groupe de mots annotés. Pour réaliser cette annotation, comme le montre la Figure 4.4, il faut analyser les informations contenues dans le modèle tel que les labels (libellés) d'activités et les données métiers manipulées par le processus. Nous appellerons cet algorithme « SA4BP ».

**Définition 10** (Annotation d'un processus métier à l'aide d'une ontologie)**.** *Soit $A \subset P$ un ensemble fini d'activités dans un processus, pour tous les concepts $c_i \in C$ avec $0 \leq i \leq n$. $n$ est le nombre de concepts présents dans l'ontologie. Nous définissons une fonction qui relie chaque activité à au moins un concept :* $\Gamma \subseteq A \times C$.

La relation entre un concept et une activité est construite à l'aide de l'annotation

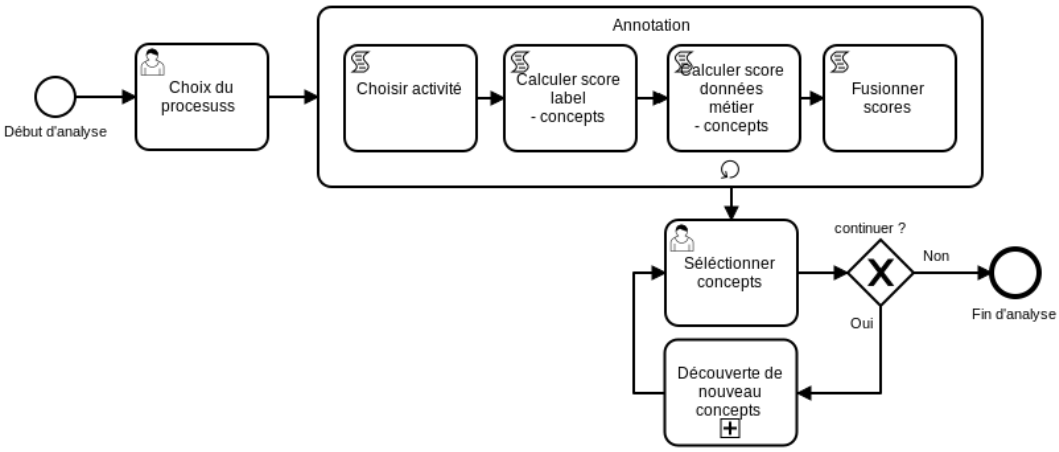


FIGURE 4.4 – Algorithme « SA4BP » d'annotation sémantique d'un processus métier.

sémantique et est illustrée dans la Figure 4.5 par la relation « maps with » entre « Business activity » et « Ontology concept ».

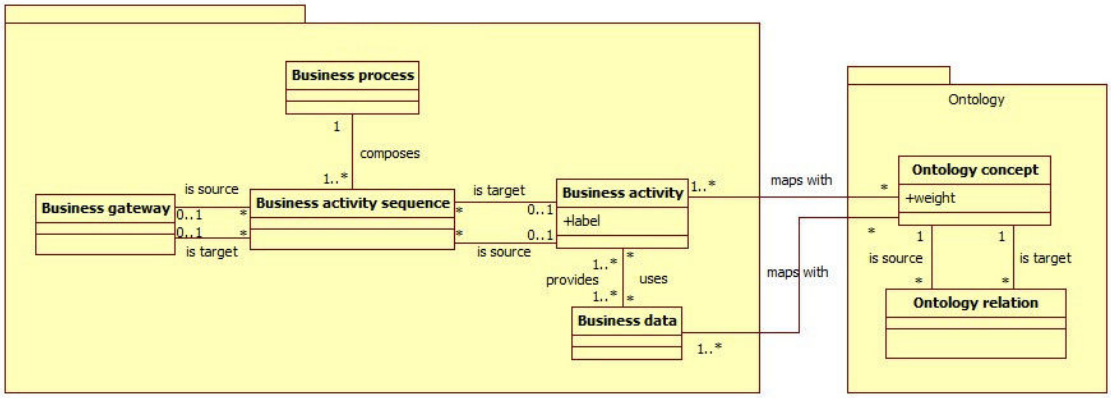


FIGURE 4.5 – Méta-modèle spécifiant la relation entre un processus et une ontologie.

La Figure 4.5 montre l'ensemble des relations entre un processus métier et une ontologie. Ces relations de type « maps with » associent les « Business data » et les « Business activity » par leur label avec un ou plusieurs « Ontology concept ». Un « Ontology concept » peut être mappé avec une infinité de « Business data » ainsi que de « Business activity ».

Dans notre illustration, l'activité « Turning on device » ainsi que la donnée métier « Device » sont alignés avec le concept « Device » de l'ontologie (Cf. Figure 4.3).

Nous proposons une distance permettant d'identifier les concepts de l'ontologie les plus proches d'une activité ou d'une donnée d'un processus. Cette distance est composée de la distance de Wu et Palmer et de celle de Levenshtein. Plus précisément, la distance de Wu et Palmer (Wup) [137] calcule la distance linguistique entre deux mots. Celle-ci est

représentée par le nombre minimum de nœuds qui relient deux mots au sein du dictionnaire WordNet [83], dictionnaire où les mots sont liés entre eux par des relations conceptuelles, sémantiques et lexicales. Quant à la distance de Levenshtein (Len), elle calcule la distance syntaxique entre une chaîne de caractères et les concepts de l'ontologie. La distance se fonde sur le nombre minimal d'opérations requises pour modifier une chaîne en une autre. Ces deux distances sont usuelles pour calculer la similarité entre les activités des processus métiers [32].

Soit $M$, l'ensemble contenant le label des activités d'un processus et les labels des données manipulées par celles-ci. $\forall m \in M$, et $\forall c \in C$, $0 \leq L(m,c) \leq 1$, la composante syntaxique du score est définie comme :

$$L(m,c) = \frac{Levenshtein(m,c)}{max(longueur(m), longueur(c))} \tag{4.1}$$

Soit $W(m,c)$, la composante linguistique du score, $\forall m \; et \; c, \; 0 \leq W(m,c) \leq 1$. Plus la valeur de $W$ est faible, plus $c$ et $m$ sont linguistiquement proches.

$$W(m,c) = 1 - Wup(m,c) \tag{4.2}$$

Nous assemblons la distance linguistique (Cf. Equation 4.1) et syntaxique (Cf. Equation 4.2) afin de calculer la distance de similarité $sim$. Ainsi, $\forall m \in M, et \; \forall c \in C, \; 0 \leq sim(m,c) \leq 2$ :

$$sim(m,c) = L(m,c) + W(m,c) \tag{4.3}$$

Ce choix d'assemblage, bien qu'arbitraire, a donné des résultats satisfaisants auprès des experts métiers. Nous avons reçu une validation de ces derniers sur la pertinence métier des distances à identifier les concepts de l'ontologie les plus proches d'une activité ou d'une donnée dans le cadre d'une relation client. C'est à partir de cette validation que la distance a été entérinée faute d'un ensemble de données de test suffisant.

Le concept $c$ est sélectionné lorsque $sim(m,c) > minScore$ avec $minScore$ le seuil minimum défini par l'expert métier. Le choix final de l'annotation est néanmoins laissé à cet expert métier afin de s'assurer de la cohérence métier du résultat.

Maintenant que nous avons expliqué l'annotation du processus, nous pouvons explorer l'ontologie afin de découvrir d'autres concepts pertinents pour le métier. Avant de présenter notre algorithme d'exploration « ContextExplorer », il est nécessaire de définir deux relations au sein de l'ontologie.

Concernant la première, nous souhaitons représenter au sein de l'ontologie une relation pouvant exister entre un processus métier et une donnée contextuelle. Pour ce faire, lors de la conception de l'ontologie, nous proposons une relation dite d'influence proche (Définition

11). Cette relation permet de relier deux concepts dont le premier est associé à une activité du processus métier et le second à une donnée contextuelle dans l'ontologie. Cette relation est définie comme suit :

**Définition 11** (Relation d'influence proche)**.** *Soit $c_i,\ c_j \in C$ avec $i, j \in |C|$, respectivement un concept d'une ontologie associé à une donnée contextuelle et un concept associé à une activité $A$. Nous disons que $c_i$ et $c_j$ sont associées par une relation d'influence proche notée $\bar{rel}(c_i, c_j)$ si et seulement si la donnée associée à $c_i$ influence l'exécution de l'activité $A$. Cette relation est nommée « wasInfluencedBy » lors de la conception de l'ontologie.*

La relation d'influence proche nous permet de définir la seconde relation dite séquentielle (Définition 12). Cette relation représente la retranscription des séquences entre deux activités en une relation entre les concepts associés à ces deux activités.

**Définition 12** (Relation séquentielle)**.** *Soit $a, b \in A$ (deux activités), $c_i, c_j \in C$ (deux concepts) de $O$ et $\Gamma : (a, c_i), (b, c_j)$ la fonction qui lie chaque activité à des concepts. On dit que $c_i$ et $c_j$ sont en relation séquentielle notée $rel(c_i, c_j)$ si et seulement si $b$ suit $a$ et qu'ils sont dans une relation causale (c'est-à-dire relation directe).*

### 4.1.2 Sélection des données contextuelles à partir de l'annotation

Notre objectif est maintenant d'explorer l'ontologie afin de savoir si de nouveaux concepts ne sont pas pertinents pour le métier.

Afin d'atteindre cet objectif, nous proposons un algorithme « ContextExplorer » d'exploration (Figure 4.6), à partir des concepts sélectionnés précédemment.

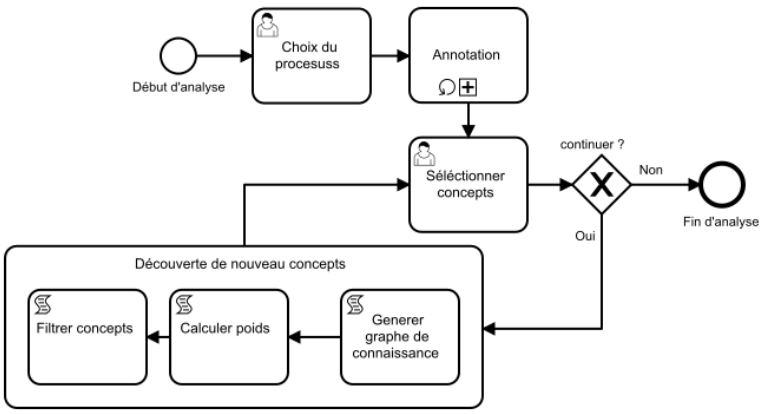


FIGURE 4.6 – Algorithme de découverte du contexte en se basant sur une ontologie.

Lorsque l'expert métier choisit de continuer l'exploration de l'ontologie, l'algorithme génère un graphe de connaissances à partir des relations d'influence proches et séquentielles de l'ontologie ainsi que des concepts précédemment sélectionnés. Il est ici question de graphe de connaissance car la modélisation réalisée au sein d'une ontologie correspond à une description conceptuelle d'une connaissance et spécifie uniquement la sémantique formelle des connaissances. Pour utiliser une ontologie à des fins d'exploration, il est nécessaire de réaliser une projection de cette description conceptuelle sur une description opérationnelle. Cette projection est appelée opérationnalisation [46]. Cela permet d'appliquer des méthodes de la recherche opérationnelle telles que le classement par popularité au sein d'un graphe. Le graphe de connaissance, résultat de l'opérationnalisation de l'ontologie, est un graphe dans lequel l'ensemble des relations entre nœuds sont bidirectionnelles. C'est une propriété bénéfique quant à la facilité de parcourir le graphe représentant l'ontologie.

L'expert métier obtient la liste des résultats à partir d'une pondération des concepts dans le graphe. Cette pondération est fondé sur les relations entre les concepts de l'ontologie. Il peut alors émettre un avis sur chaque concept en les sélectionnant ou en les rejetant. La pertinence métier du résultat est ainsi améliorée par l'apport de l'expertise métier de l'expert.

L'algorithme du *PageRank* proposé par [93] permet de calculer le poids d'un nœud dans un graphe à l'aide d'un classement par popularité. Par exemple, dans le contexte des moteurs de recherche, il permet de quantifier l'importance d'une page sur Internet. Le principe du *PageRank* consiste à poser l'hypothèse que lorsqu'un nœud est en relation avec un autre, on dit qu'il vote pour lui. Le poids associé à un nœud est déterminé en fonction des votes qui sont exprimés pour lui ainsi que du poids des nœuds diffusant ces votes. Plus le nombre de votes pour un nœud est élevé, plus le poids de ce nœud est élevé. De plus, le poids d'un nœud induit le poids du vote émis par ce nœud.

Soit $G = (N, E)$ un graphe orienté composé d'un ensemble de nœuds $N$ et d'arêtes $E$ où $E$ est un sous-ensemble de $N \times N$. $In(N_i)$ est l'ensemble des arêtes pointant vers $N_i$ : $In(N_i) = \{N_j \in N; \exists A \in E;\ A = (N_j, N_i)\}$. $Out(N_j)$ est l'ensemble des arêtes partant de $N_j$ : $Out(N_j) = \{N_k \in N;\ \exists A \in E;\ A = (N_j, N_k)\}$. Le poids PR du nœud $N_i$ est donné par l'équation :

$$PR(N_i) = (1 - d) + d * \sum_{N_j \in In(N_i)} \frac{PR(N_j)}{Out(N_j)} \tag{4.4}$$

où $d \in [0, 1]$.

L'algorithme du *PageRank* met en œuvre le « modèle du surfeur aléatoire », où un utilisateur simulé clique sur des liens au hasard avec une probabilité $d$ de passer à une page nouvelle. Cette probabilité $d$ joue le rôle d'un facteur d'amortissement. Elle permet

d'incorporer dans l'algorithme *PageRank* la probabilité de sauter d'un nœud à un autre nœud, aléatoirement, au sein du graphe. Ceci permet d'éviter de passer par les nœuds ne pouvant voter pour d'autres nœuds (bloquants). Sans ce facteur d'amortissement, la matrice des transitions entre les nœuds aurait comme propriété d'être réductible. Ceci signifie que le surfeur resterait bloqué sur un même nœud. Le facteur $d$ a été fixé à 0,85 par les auteurs de [16]. Les auteurs ont fixé ce facteur à partir d'une mesure de la probabilité qu'à un internaute de changer de signet dans son navigateur.

En partant de valeurs arbitraires attribuées à chaque nœud du graphe, l'algorithme *PageRank* itère jusqu'à convergence. C'est à dire, lorsque la variation des poids est évaluée en dessous d'un seuil donné. Ce seuil est usuellement défini comme étant une approximation de 0, soit $10^{-8}$ dans notre cas.

D'autres travaux sur le classement par popularité dans des graphes [139] prennent en compte des poids sur les relations entre nœuds [138]. Le calcul de ces poids peut être fondé sur l'analyse du contenu des nœuds, ici des pages web, en relation [103], ou sur le nombre de clics d'utilisateurs visitant les pages [19].

Les travaux de Matthew Richardson et de Pedro Domingos [103] portent sur un modèle plus évolué du surfeur aléatoire permettant de prendre en compte de nouvelles informations telles que les relations et le contenu de page web. Ce modèle se traduit par la formule du poids $Pg_q$ du noeud $N_j$ :

$$Pg_q(N_j) = (1-\beta)Pg'_q(N_j) + \beta \sum_{N_i \in In(N_j)} Pg_q(N_i)Pg_q(N_i \rightarrow N_j) \tag{4.5}$$

$Pg_q(N_i \rightarrow N_j)$ est la probabilité que le « surfeur » change du nœud $N_i$ vers la nœud $N_j$ et qu'il cherche un nœud à l'aide d'une requête q. $Pg'_q(N_j)$ spécifie le poids du noeud $N_j$ spécifique aux évènements représentant un saut d'un noeud vers $N_j$ en l'absence d'une relation entre ces deux noeuds. $Pg_q(N_j)$ est l'importance du nœud $N_j$ par rapport à la requête $q$ aussi appelé *query-dependant PageRank score*. Les auteurs posent $R_q(N_j)$ comme étant la mesure de la pertinence du nœud $j$ par rapport à la requête $q$ d'une page web.

On définit par conséquence $Pg'_q(N_j)$ comme étant :

$$Pg'_q(N_j) = \frac{R_q(N_j)}{\sum_{N_k \in N} R_q(N_k)} \tag{4.6}$$

avec $N$ l'ensemble des nœuds du graphe.

$Pg_q(N_i \rightarrow N_j)$ est défini par :

$$Pg_q(N_i \rightarrow N_j) = \frac{R_q(N_j)}{\sum_{N_k \in Out(N_i)} R_q(N_k)} \tag{4.7}$$

Comme le soulignent les auteurs, cette méthode permet au « surfeur » lorsqu'il change de page de favoriser les pages ayant une forte pertinence pour l'utilisateur vis-à-vis de sa requête $q$. La méthode appliquée ici est stochastique, apériodique et irréductible (en raison du facteur d'amortissement). Au vu de ces propriétés, la matrice des transitions entre les nœuds est donc une chaîne de Markov. Le théorème de Perron-Frobenius qui permet de garantir qu'une chaîne de Markov converge vers un point fixe unique est donc applicable [104].

En nous fondant sur cette variante de l'algorithme *PageRank*, nous avons introduit dans notre algorithme la détection des nœuds importants en prenant en compte l'état des nœuds. En effet, il est nécessaire dans notre approche de prendre en compte l'avis de l'expert métier dans la recherche des concepts les plus proches dans l'ontologie. La pertinence de la proximité d'un concept de l'ontologie pour l'expert est équivalente à l'état d'un nœud.

Les nœuds, ou concepts de l'ontologie, peuvent être dans un des trois états suivant :

- « sélectionné » : un nœud choisi est un nœud proposé et sélectionné par un expert métier comme concept pertinent. Par exemple, le concept « Device » est pertinent avec un processus d'authentification ;
- « rejeté » : un nœud refusé est un nœud proposé et refusé par l'expert pour sa non-pertinence. Par exemple, le concept « chat » n'a que peu de chance d'être pertinent pour un processus d'authentification ;
- « défaut » : état par défaut d'un nœud, c'est-à-dire lorsqu'il n'a pas encore été proposé à un expert métier.

L'algorithme basé sur *PageRank* que nous proposons, nommé « ContextExplorer », résulte de la pondération des nœuds permettant à l'expert métier de les classer ou de les rejeter. Le calcul de cette pondération doit satisfaire les axiomes suivants (Cf. axiomatisation de la mesure définie dans [116]) :

- Le poids d'un concept augmente (respectivement, diminue) si le poids des concepts directement relié à celui-ci (relation séquentielle ou relation d'influence proche) augmente (respectivement, diminue) ;
- Le poids d'un concept dépend du nombre de ses relations avec d'autres concepts et de l'état de ces concepts.

La modification de l'algorithme *PageRank* introduit deux constituants : une fonction de coût et une fonction de pénalité. Tout d'abord, la fonction de coût du noeud $N_j$, notée $jump(N_j)$ :

$$jump(N_j) = \begin{cases} \alpha \ si \ N_j \ est \ sélectionné \\ \beta \ si \ N_j \ est \ rejeté \\ \gamma \ si \ N_j \ a \ la \ valeur \ par \ défaut \end{cases} \tag{4.8}$$

avec $\alpha,\ \beta,\ \gamma \in \Re^+$ tels que $\alpha + \beta + \gamma = 1$. Le choix de la valeur d'$\alpha,\ \beta$ et $\gamma$ est laissé à l'appréciation de l'expert métier en fonction de la taille de l'ontologie ainsi que du nombre de relations entre concepts. Il va ainsi pouvoir réduire les parcours dans le graphe à partir de ses connaissances métiers.

La fonction de coût modifie la probabilité qu'a le surfeur aléatoire de sauter sur un nœud sélectionné.

Ensuite, la fonction de pénalité $pathToSelect(N_i, N_j)$ calcule la distance minimum permettant de relier le nœud $N_i$ au noeud $N_j$.

$$pathToSelect(N_i, N_j) = min_k\{Card(path_k(N_i, N_j)\} \tag{4.9}$$

où $path_k(N_i, N_j)$ est un ensemble d'arêtes du graphe composant un chemin de $N_i$ vers $N_j$.

Ces deux fonctions vont nous permettre d'adapter la Formule 4.4 de l'algorithme du *PageRank* à notre problématique. Le poids W d'un nœud est alors défini par :

$$W(N_i) = \frac{jump(N_i) * (1-d)}{Card(N)} + d * \sum_{N_j \in In(N_i)} \frac{W(N_j)}{Out(N_j)} * \frac{1}{pathToSelect(N_i, N_j)} \tag{4.10}$$

L'étape de calcul de l'algorihtme « ContextExplorer » est fondée sur cette équation donnant le poids W d'un noeud ou concept de l'ontologie. C'est une étape fondamentale de la définition du contexte d'exécution d'un processus métier, c'est à dire l'ensemble de ses données contextuelles, $DC_P$ (Cf. Figure 3.7, chapitre 3).

Ces étapes nous permettent d'intégrer le concept de données contextuelles (Cf. Figure 4.7) dans le méta-modèle d'un processus (Cf. Figure 4.5). Ce concept « Contextual data » est associé dans ce méta-modèle au concept d'activité d'un processus, « Business activity », et au concept de donnée métier, « Business data », par une association « maps with ». Cette association doit être cohérente avec les éléments de l'ontologie représenté par « Ontology concept » dans le méta-modèle.

Dans l'illustration, la donnée contextuelle « Noise » qui est alignée avec l'activité « Creating authentification session » et la donnée métier « Device ». Cette alignement est justifié par le concept « Synchronisation » de l'ontologie car le bruit affecte directement la synchronisation du terminal.

Les prochaines étapes consisteront à analyser les instances des données contextuelles ainsi que les traces de processus afin de déduire l'ensemble $DCI$ (Cf. Figure 3.7, chapitre

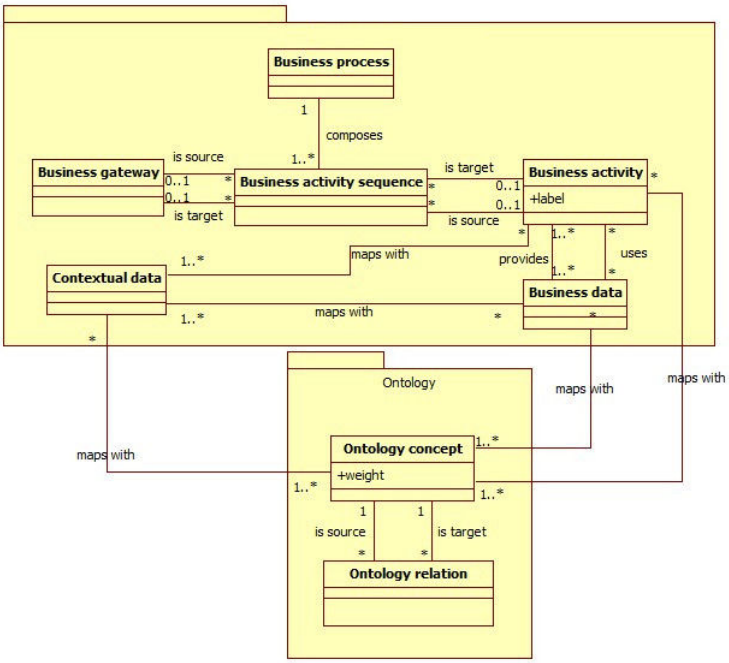


FIGURE 4.7 – Représentation des liens entre un processus métier et des données contextuelles.

3) des données contextuelles pertinentes.

## 4.2 Partitionnement des traces des instances d'un processus opérationnel pour détecter des comportements

L'analyse des traces générées par des instances d'un processus nous renseigne sur la manière dont un processus s'est déroulé, c'est-à-dire son comportement. Elle a pour but d'identifier chaque comportement par un regroupement des instances similaires dans un cluster (Définition 6). Dans [1], les clusters sont construits en deux étapes. La première est fondée sur l'analyse des traces et la seconde sur les KPI (*Key Performance Indicators*). Les KPI permettent de suivre l'évolution de l'atteinte des objectifs définis dans la stratégie de l'entreprise. Chaque objectif est associé à un ou plusieurs processus métiers. L'analyse de ces KPI complète ainsi l'analyse des traces de ces processus métiers.

Les auteurs de [1] proposent une méthode permettant de détecter les valeurs de KPIs propices au partitionnement des instances des processus associés. Pour chaque KPI, leur méthode découpe l'ensemble des instances en $k$ groupes selon la distribution des résultats de la KPI observée durant l'exécution du processus. Ils appliquent ensuite un algorithme de partitionnement résultant une détection des groupes d'instances similaires et en une

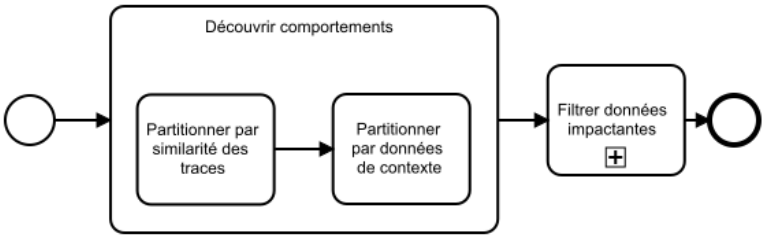


FIGURE 4.8 – Découverte des comportements associés au processus métier.

agrégation des groupes de KPI correspondants. Cette algorithme a pour avantage de découpler l'analyse des instances d'un processus et l'analyse des données liées, ici les KPI.

Dans le cadre de notre projet, l'étude des KPI est pertinente pour enrichir un processus opérationnel. Nous avons choisi comme critères de regroupement, complémentaires aux traces, les données contextuelles des processus métiers au lieu des KPIs. Ce choix a été motivé par une volonté de prioriser le déroulement du processus et non la stratégie de l'entreprise. Nous cherchons donc des sous-ensembles de données contextuelles, $DC$ (Cf. Figure 3.7), tels qu'une donnée contextuelle $dc_i$ soit associée à au moins un cluster d'instances de processus.

La Figure 4.8 montre les deux étapes que nous réalisons pour réaliser le partitionnement. Nous allons d'abord définir les clusters de trace $tc$ associés à un processus à l'aide de technique de fouille de données puis nous utiliserons les instances des données contextuelles afin de raffiner le résultat et définir $DCNI$ et $DCI$ (Cf. Définition 8).

### 4.2.1 Premier partitionnement des traces des instances d'exécution

Le partitionnement de données est généralement appliqué dans le cadre de la fouille de données comme une étape de pré analyse. En effet, l'analyse est dans ce cas fondée sur le regroupement de données similaire. Elle est utilisée pour améliorer le résultat final d'un algorithme en minimisant le bruit, c'est-à-dire en écartant les données aberrantes. Dans [1], le partitionnement est utilisé afin de regrouper des traces (Cf. Définition 3) similaires associées aux différentes exécutions d'un même processus (comportements).

Le fonctionnement du partitionnement de données, et a fortiori de données de traces, consiste généralement à utiliser une métrique (ou distance) permettant d'évaluer la similarité entre deux traces afin de pouvoir regrouper les traces les plus similaires. Pour réaliser ce partitionnement, deux points sont à prendre en compte : le choix de l'algorithme de partitionnement et le choix de la métrique permettant de regrouper les traces en fonction d'une ou plusieurs perspectives (Cf. Définition 4).

**Choix d'algorithme de partitionnement** Le choix d'un algorithme de partitionnement dépend du type de données manipulées et du besoin (pertinence métier des regroupements, performance temporelle du partitionnement). Nous avons sélectionné les algorithmes suivants : le *K-medoids* [85], le *Hierarchical Agglomerative Clustering* (HAC) [63] et le *Density-fonded Spatial Clustering of Applications with Noise* (DBSCAN) [64]. Ces algorithmes ont été choisis du fait de leur relative simplicité de fonctionnement permettant à un expert métier de bien les comprendre quand à leurs principes. En appliquant ces algorithmes, nous cherchons à la fois à construire des groupes de trace similaires, mais aussi à donner les motifs de ce regroupement compréhensible par un expert.

- *K-medoids* : C'est un dérivé d'une méthode très populaire en fouille de données, le k-means, exploitant la distance entre éléments. Il vise à partitionner $n$ observations en $k$ groupes dans lesquels chaque instance appartient à un groupe ayant la distance moyenne la plus faible avec lui. Le K-medoids est intéressant lorsque l'on a une idée du nombre approximatif de groupes contenu dans l'ensemble des instances. Une représentation usuelle consiste à utiliser le diagramme de Voronoï afin de visualiser les groupes. Ce diagramme est un outil illustrant les motifs de chaque regroupement ;

- *HAC* : Il se fonde sur le regroupement successive des deux groupes les plus ressemblants (ayant une faible distance entre eux). Lors de la première étape, l'algorithme considère chaque instance comme étant un groupe de taille un, et va chercher quels sont les deux groupes les plus proches pour les fusionner. L'étape suivante consiste à réitérer la recherche des deux groupes similaires pour les fusionner et ainsi de suite jusqu'à l'obtention d'un dendogramme complet. Le dendogramme a aussi pour avantage de fournir un résultat compréhensible pour un utilisateur final et notamment en expliquant visuellement la création des groupes dans l'analyse ;

- *DBSCAN* : C'est un algorithme de partitionnement qui fonde sa méthode de regroupement sur la mesure de la densité (nombre de traces) des groupes. Il trouve un nombre $k$ de groupes à partir de la distribution de densités estimées des instances. Il est l'un des algorithmes de partitionnement les plus courants. Il est très intéressant par sa capacité à découvrir le nombre de groupe et sa « résistance aux bruits ». Cependant contrairement à l'algorithme HAC, le DBSCAN ne propose pas une compréhension simple de la création des groupes.

Sur des données telles que des traces de processus, ses trois algorithmes ont des résultats similaires en performance [89, 64]. K-medoids est connu pour être capable d'analyser de grande quantité de données plus rapidement que les autres algorithmes, cependant il nécessite de connaître le nombre de clusters a priori. L'algorithme DBSCAN est bien plus

lent [64] sur de grandes quantités de données. Cet algorithme a le gros avantage dans le cadre de la découverte de cluster, de ne pas nécessiter de connaître a priori le nombre de cluster. L'algorithme HAC [89] est quant à lui l'algorithme qui nécessite le plus de temps de calcul des trois algorithmes. Il a une complexité $O(n^2)$ où $n$ est le nombre de données en entrée de l'algorithme. Comme DBSCAN, l'algorithme HAC ne nécessite pas de connaître les nombres de cluster puisqu'il est capable de le découvrir.

Notre besoin côté métier nécessite un algorithme non paramétré (nombre de cluster inconnu) afin d'être plus explicite. Il doit donc fournir à la fois un résultat pertinent et compréhensible. Seul l'algorithme HAC satisfait à ce besoin métier. L'algorithme HAC permet en effet de voir et connaître l'ensemble des étapes de création des clusters. On notera l'apport de la création d'un dendogramme (Cf. Figure 4.9) permettant de montrer visuellement à un expert métier à quel moment deux clusters ont fusionné. Ceci permet à un expert de mieux comprendre la création des clusters. Un autre avantage de cet algorithme est sa capacité à intégrer de nouvelles données, ici de nouvelles traces, sans avoir besoin de reprendre tout le calcul. On choisit donc l'algorithme HAC dans notre contexte métier même si le coût de temps de calcul est plus important.

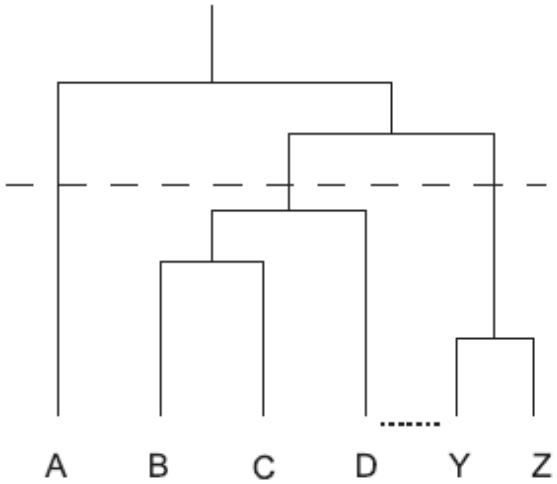


FIGURE 4.9 – Illustration d'un dendogramme.

Nous allons donc détailler l'algorithme HAC. Celui-ci possède plusieurs méthodes permettant de calculer la distance entre deux clusters. Les quatre méthodes de *linkage* (fusion de cluster) correspondant à notre besoin sont les suivants. Soit $d(x, y)$ la distance entre deux traces $x$, $y$ appartenant respectivement aux clusters de traces $tc_i$ et à $tc_j$ :

- *Complete linkage clustering* : cherche la distance maximale possible entre tous les

points appartenant à deux clusters différents $tc_i$ et $tc_j$.

$$D_{completelinkage}(tc_i, tc_j) = max\{\ d_{x \in tc_i, y \in tc_j}(x, y)\} \tag{4.11}$$

- *Single linkage clustering* : cherche la distance minimale possible entre tous les points appartenant à deux clusters différents $tc_i$ et $tc_j$.

$$D_{singlelinkage}(tc_i, tc_j) = min\{d_{x \in tc_i, y \in tc_j}(x, y)\} \tag{4.12}$$

- *Mean linkage clustering* : cherche toutes les distances par paires possibles pour les points appartenant à deux clusters différents $tc_i$ et $tc_j$, puis calcule la moyenne.

$$D_{meanlinkage}(tc_i, tc_j) = \frac{1}{|tc_i| \bullet |tc_j|} \sum_{x \in tc_i} \sum_{y \in tc_j} d(x, y) \tag{4.13}$$

- *Centroid linkage clustering* : cherche le centre de chaque cluster et calcule la distance entre les centroïdes de deux clusters, $tc_i$ et $tc_j$. Soit $Centroids_{tc_i}$ et $Centroids_{tc_j}$ respectivement les centroïdes de $tc_i$ et $tc_j$.

$$D_{centroidlinkage}(tc_i, tc_j) = ||centroids_{tc_i} - centroids_{tc_j}|| \tag{4.14}$$

Les méthodes ont été évaluées sur des traces provenant du « Business Processing Intelligence Challenge » (BPIC) 2015, ainsi que des traces générées par l'outil PLG2 [1] [18]. Nous n'avons pas remarqué une différence significative dans les résultats quant à l'utilisation de ces quatre méthodes. Du faite de l'utilisation de tout les couples possibles x,y, nous avons considéré la méthode *mean linkage* plus exhaustive.

Le fonctionnement de l'algorithme HAC peut se résumer en trois étapes. Chaque cluster est initialisé avec une trace.

1. Identifier les deux clusters les plus proches ;
2. Fusionner les clusters ;
3. Répéter l'étape 1et l'étape 2 jusqu'à ce que toutes les traces se trouvent dans un cluster unique.

L'algorithme s'arrête lorsque toutes les traces ont été regroupées. C'est à ce moment que l'expert métier va pouvoir, à l'aide du dendogramme, décider du niveau où la fusion doit s'arrêter. Ceci peut se faire de plusieurs façons [74]. Il est possible de choisir une

1. http ://plg.processmining.it/, visité le 20/04/2017

distance maximale entre deux clusters en fonction de la méthode de linkage choisi. D'autres solutions sont possibles ne nécessitant pas de définir un seuil [74]. Par exemple, il existe des méthodes fondées sur la distance intercluster. Cette méthode propose d'arrêter la fusion des clusters lorsque la variation de distance entre deux niveaux successifs est forte. D'autres méthodes proposent d'exploiter la variance totale intracluster ou encore la similarité d'un objet avec son propre cluster (cohésion) par rapport aux autres clusters (séparation).

Une autre solution est d'appliquer la méthode *Elbow* [69]. C'est une méthode d'interprétation et de validation de la cohérence dans le partitionnement conçu afin de trouver le nombre approprié de groupes dans un ensemble de données. Elle se fonde sur le calcul de la variance totale intracluster. L'analyse est centré sur l'emplacement visuelle d'un virage dans la courbe (Cf. Figure 4.10) qui représente la variance totale intracluster par rapport au nombre de cluster choisi. Cet emplacement est considéré comme un indicateur du nombre approprié de clusters.

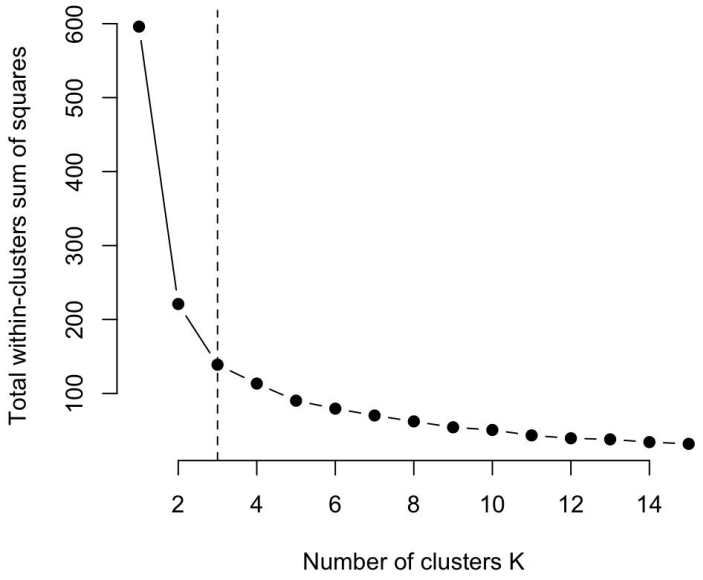


FIGURE 4.10 – Illustration d'une courbe issue de la méthode *Elbow*.

**Choix de la métrique de partitionnement** Le second point à prendre en compte est la métrique qui servira à calculer une distance de similarité.

Notre objectif est de regrouper les traces similaires d'un point de vue de la perspective associée à la séquence des activités qui les génèrent. Dans cette optique, nous nous sommes particulièrement intéressés à la métrique *Generic Edit Distance* (GED) proposée par [14].

En effet, cette métrique a une pertinence métier car elle se focalise que sur les séquences d'activités et leurs longueurs en terme d'activité.

Notre approche utilise cette métrique associé au partitionnement hiérarchique afin de découvrir des clusters. Ces clusters sont donc définis en fonction de la similarité des traces qui les composent. Comme le montre la Figure 3.5, l'usage du contexte est utile pour mieux comprendre le fonctionnement métier d'une instance. Nous allons maintenant expliquer comment incorporer la notion de contexte dans cette approche.

### 4.2.2 Second partitionnement à l'aide des instances des données contextuelles

Identifier les clusters de traces d'un processus est une information permettant de mieux comprendre quels sont les comportements d'un processus et, le cas échéant, identifier des améliorations à effec,tuer. Dans notre exemple, il existe plusieurs comportements tels que des instances qui se terminent avant la complétion du processus due à l'annulation d'un client, des instances qui se terminent avec beaucoup de retard ou encore des instances qui se déroulent normalement, etc. Mettre en relation les clusters précédemment découverts avec d'autres données telles que les données contextuelles permet d'identifier des comportements inadéquats par rapport à l'exécution théorique du processus ou identifier la raison d'une bonne ou mauvaise performance lors de l'exécution. Le but de cette partie est d'expliquer la méthode que nous proposons afin de raffiner les clusters en sous-ensemble en fonction d'une ou plusieurs données contextuelles.

En utilisant les clusters découverts, nous voulons étudier leur composition en fonction du contexte. Pour ce faire, nous proposons d'appliquer une seconde fois l'algorithme HAC avec une nouvelle métrique de distance. En effet, nous souhaitons diviser les clusters de traces par type de contexte pouvant ainsi résulter en plusieurs clusters différents ayant le même contexte, mais issues de différents clusters de traces. Pour ce partitionnement une distance entre les instances de données contextuelles est défini ci-après. Nous considérons l'espace vectoriel ayant pour chaque dimension une donnée contextuelle (la dimension de l'espace vectoriel est le nombre de donnée contextuelle). Nous proposons la distance $d$ entre $iedc_1$ et $iedc_2$ deux instances d'un ensemble de données contextuelles caractérisant une exécution d'un processus. $iedc$ est représenté par un vecteur de taille $n$ où chaque scalaire est l'instance d'une donnée contextuelle tel que $iedc_1 = (iedc_1 1, ..., iedc_n 1)$ et $iedc_2 = (iedc_{1,2}, ..., iedc_{n,2})$ où $iedc_{i,x}$ est une instance de $DC_i$.

$$d(iedc_1, iedc_2) = \sqrt{\frac{\alpha_1 \times (idc_{1,1} - idc_{1,2})^2 + .. + \alpha_n \times (idc_{n,1} - idc_{n,2})^2}{\sum_{k=1}^{n} \alpha_k}} \tag{4.15}$$

tel que $\alpha_k \in \Re+$ est défini par l'expert métier afin de compenser au mieux les différences d'échelle entre les instances de données.

Donc $n$ est le nombre de données contextuelle pris en compte. Les coefficients $\alpha_k$ sont déduits du choix de l'expert métier. En effet, l'expert métier choisit les données contextuelles à analyser et indique le cas échéant le coefficient de chaque donnée.

En conclusion, c'est à l'aide de ces deux partitionnement successif (Cf. 4.2.2) que nous obtenons dans chaque cluster de trace des clusters d'instance de données contextuelles associés. C'est ainsi que nous définissons un comportement par cluster de trace. La section suivante va nous permettre de sélectionner dans chaque cluster de trace les données contextuelles pertinentes.

## 4.3 Sélection et représentation des données contextuelles significatives par rapport à un comportement

L'objectif de cette section est de présenter notre approche permettant de séparer l'ensemble $DC$ en $DCI$ et $DCNI$. Nous allons dans un premier temps nous analyser le lien des instances des données contextuelles et les comportements. Ensuite, nous allons générer une représentation des instances de données contextuelles dans un arbre de décision afin d'aider l'expert métier dans sa compréhension des comportements.

### 4.3.1 Sélection des données contextuelles associées à un comportement

La sélectionner de données contextuelles les plus proches d'un comportement peuvent être traité par des méthodes issues de la fouille de donnée. Notre objectif est d'améliorer la pertinence métier de l'association entre des données contextuelles et un comportement. A cet effet, nous avons étudié plusieurs de ces méthodes de sélection de données afin d'en sélectionner une qui soit efficace pour le métier.

Nous distinguons deux objectifs à ces méthodes :

- **Interprétation** : Chercher à sélectionner toutes les données contextuelles qui sont fortement reliées au comportement même si les données contextuelles sont dépendantes entre elles.
- **Amélioration de la prédiction** : Chercher à sélectionner un sous-ensemble restreint de données contextuelles permettant la prédiction correcte du comportement (cluster de trace). Ceci permet notamment d'améliorer le temps d'apprentissage statistique des modèles prédictifs ainsi que de les simplifier.

Nous utilisons l'algorithme de régression logistique [39] auquel on applique la méthode Lasso (régularisation L1 [101], où les données contextuelles sont appelées variables)

justifiée par son impact sur la performance métier et sur le bon équilibre entre la suppression de données contextuelles redondantes et l'amélioration de la performance prédictive. L'ajout de la méthode de « stability selection » permet de rendre l'algorithme plus robuste [80]. Cette méthode de « stability selection » se fonde sur le sous-échantillonnage générer aléatoirement de données contextuelles et de leurs instances. Sur chaque sous-échantillon généré, la méthode calcule la capacité prédictive de chaque donnée contextuelle. Nous utilisons la méthode de « stability selection » afin de calculer la stabilité des modèles générés à l'aide de notre algorithme sur différents sous-ensembles de données contextuelles et de leurs instances.

L'idée sous-jacente à la sélection de stabilité est de déterminer les données contextuelles importantes dans chaque version échantillonnée en y appliquant une méthode de sélection de variables de Lasso. Les résultats sur chaque échantillon sont ensuite agrégés pour calculer un score de stabilité de chaque donnée contextuelle. Les données contextuelles peuvent ensuite être sélectionnées en choisissant un seuil approprié.

Nous séparons les données contextuelles restantes par rapport aux comportements. Le principe de la sélection de données contextuelles par rapport aux comportements (en anglais, *class-specific feature selection*) est d'extraire un sous-ensemble optimal de données contextuelles qui représentent chaque comportement [111]. A cet effet, nous proposons d'appliquer la stratégie du one-vs-all [60, 100]. Cette stratégie consiste à considérer les instances des données contextuelles associées à un comportement comme des observations positives, tandis que les autres comme observations négatives. Cette stratégie est notamment utilisée avec des algorithmes d'apprentissage afin de binariser (positif / négatif) des problèmes à $k$ groupes en les transformant en $k$ problèmes binaires.

Après cette binarisation, nous nous retrouvons avec $k$ problèmes binaires à traiter. Nous appliquons sur chacun des problèmes binaires l'algorithme de raffinement (régression logistique + lasso) des données contextuelles permettant de créer $k$ sous-ensembles de données contextuelles.

La dernière étape de notre approche consiste à représenter les sous-ensembles de données contextuelles obtenues en un arbre de décision afin d'assister l'expert métier dans sa compréhension de l'exécution du processus.

### 4.3.2 Représentation de sous-ensembles de données contextuelles en un arbre de décision métier

La compréhension de l'exécution du processus à l'aide de sous-ensemble de données contextuelles peut être représenter par un arbre de décision décrivant pour chaque comportement les instances de données contextuelles pertinentes.

Un arbre de décision [97] représente un ensemble de choix sous la forme graphique

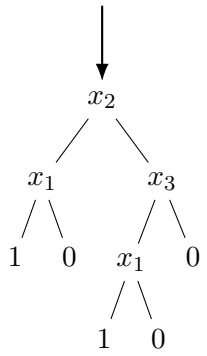


FIGURE 4.11 – Illustration d'un arbre de décision avec $x_i$ une garde.

d'un arbre (Cf. Figure 4.11). Les différentes décisions possibles sont situées aux extrémités des branches (les « feuilles » de l'arbre), et sont atteintes en fonction de décisions prises à chaque étape (« fourche »). L'arbre de décision est un outil utilisé dans des domaines variés tels que la sécurité, la fouille de données, la médecine, etc. Il a l'avantage d'être lisible et simple d'utilisation.

Dans notre cas, les feuilles de l'arbre sont associées à un comportement et les fourches sont définies par des gardes sur les instances de données contextuelles. Dans le cadre d'un processus de livraison de colis, lors de la livraison, $x_1$ = (distance de livraison $<$ 5km), $x_2$ = (météo == « Soleil »), $x_3$ = (nombre de produit à livré $>$ 3). L'instance 1 pour la feuille correspond au comportement « livraison dans les temps » et 0 correspond à un autre comportement.

Notre algorithme souligne la pertinence, dans l'arbre de décision, des données contextuelles par rapport au comportement. En effet, plus une donnée contextuelle apparaît dans une fourche haute dans l'arbre, plus elle est pertinente par rapport au comportement vu du métier (plus le score de « stability » calculé est haut). L'arbre est d'autant plus pertinent qu'il permet de faire le lien de façon visuelle entre une donnée contextuelle et un comportement donc avec un type de trace.

## 4.4 Conclusion

Dans ce chapitre, nous avons montré comment en partant de l'annotation d'un processus, nos algorithmes ont rendu possible la détection d'un ensemble de données contextuelles pertinentes vis-à-vis du déroulement d'un processus et du métier. Les données contextuelles sont une réponse aux questions de recherche QR1 en soulignant l'importance des données produites par les processus opérationnels et supports et QR2 où le contexte permet de mieux comprendre l'exécution d'un processus opérationnel. Cette méthode se découpe en quatre étapes.

Dans un premier temps nos algorithmes lient le processus à des données contextuelles au travers d'une ontologie (Algorithme « SA4BP »).

Dans un deuxième temps, l'algorithme permet de sélectionner d'autres concepts de l'ontologie en lien avec des données contextuelles (Algorithme « ContextExplorer »).

Dans un troisième temps, l'algorithme permet un partitionnement des traces d'un processus en comportement regroupant des instances de processus similaires.

Dans un dernière temps, nous raffinons les données contextuelles issues de l'annotation et de l'exploration de l'ontologie afin de créer un sous-ensemble fini de données contextuelles pertinentes pour chaque comportement et ainsi générer une liste de toutes les données contextuelles pertinentes pour un processus opérationnel.

Maintenant que nous avons les données contextuelles associées aux processus métiers et plus particulièrement les données contextuelles qui ont un impact sur le déroulement des processus opérationnels, nous souhaitons évaluer les processus supports afin de créer un sous-ensemble de processus supports permettant de réduire l'impact sur le métier des comportements dits anormaux.

Chapitre 5

# Évaluation métier des processus supports gérant chaque donnée contextuelle

Dans une organisation centrée sur les processus de l'entreprise, les processus supports sont au centre de la performance opérationnelle de l'entreprise. Les processus supports ont vocation à répondre aux besoins des processus opérationnels afin que ces derniers soient en mesure de répondre favorablement à la demande du client. Les processus opérationnels sont en relation directe avec les clients de l'entreprise. Les processus supports, directement reliés aux processus opérationnels, concourent à leur performance métier.

Dans ce chapitre, nous explicitons le lien entre les processus opérationnels et les processus supports à partir des données contextuelles. Le chapitre fait référence à la troisième question de recherche (QR3) en assistant l'expert métier afin de sélectionner les processus supports en se fondant sur ses préférences. Les algorithmes « SA4BP » et « ContextExplorer » explicités dans le chapitre précédant sont appliqués d'un coté aux processus opérationnels et de l'autre aux processus supports.

Le lien se fait donc à partir des données contextuelles associées à la fois à un processus opérationnel et à des processus supports.

L'objectif de ce chapitre est de proposer une méthode d'évaluation des processus supports ainsi sélectionnés par un expert métier à partir d'une aide multi-critère à la décision.

## 5.1 Sélection des processus supports gérant chaque donnée contextuelle

Un répertoire de processus [76] fournit un emplacement de référence centralisé où les employés de l'entreprise peuvent accéder aux informations de processus pour comprendre comment le métier devrait fonctionner. Les répertoires doivent être gérés en définissant des procédures pour s'assurer que tous les changements sont correctement contrôlés et validés.

Ces répertoires ont une forte valeur ajoutée puisqu'ils permettent, entre autres, l'historisation des mises à jour des processus, la recherche de processus, la gestion des droits d'accès et le travail collaboratif.

Dans le cas de la gestion d'anomalies, nous sommes à la recherche d'un ou des processus supports manipulant une donnée contextuelle liée à une anomalie. Une anomalie est un comportement divergent du comportement nominal d'un processus opérationnel. Pour sélectionner les processus support, il est nécessaire de parcourir l'ensemble du référentiel des processus supports à la recherche de processus support ayant leur première activité annotée par au moins une donnée contextuelle associé au processus opérationnel.

## 5.2 Évaluation des processus supports en fonction des préférences de l'expert métier

Pour chaque donnée contextuelle, plusieurs processus supports peuvent être sélectionnés. Il est nécessaire de les évaluer afin de maîtriser leurs impacts métiers sur le processus opérationnel associé. Nous utilisons l'AMCD qui permet de prendre en compte les préférences de l'expert métier afin de les évaluer.

Les étapes suivantes structurent un processus d'aide à la décision multi-critère : identifier le problème, le formuler, construire le modèle d'évaluation, puis aboutir à une recommandation finale [15, 12, 44].

Chacune de ces étapes contient des sous-étapes supplémentaires, qui dépendent des nombreux facteurs définissant un problème de décision. Plus précisément, l'étape de structuration du problème comprend des sous-étapes telles que l'identification des parties prenantes (ou des acteurs), l'identification du contexte du problème, les objectifs de la décision et ses propriétés respectives. La deuxième étape de la formulation du problème consiste à identifier les alternatives de décision et leurs critères, le type de problème de décision, ainsi que la gestion de plusieurs décideurs et de leurs différentes perspectives. La troisième étape implique le choix d'un modèle mathématique et son réglage de manière à ce qu'il reflète la perspective du décideur. En outre, une méthode de résolution doit également être sélectionnée afin de fournir une recommandation au problème de décision. Enfin, lors de la dernière étape, cette recommandation est présentée au décideur en lui proposant deux possibilités. Il peut valider la recommandation et ainsi demander des analyses supplémentaires ou re-visite les étapes précédentes afin d'affiner la recommandation. Soulignons que la structure du processus est non linéaire, complexe et itérative. Cela signifie que, à tout moment du processus, nous pouvons choisir différentes voies à suivre, dans certains cas, en revenant aux étapes précédentes.

L'étape de résolution du processus d'aide multi-critères à la décision précédemment décrite implique l'utilisation d'un algorithme d'agrégation multi-critère. Diverses approches algorithmiques ont été proposées dans la littérature [15, 66, 112].

Deux principales écoles méthodologiques peuvent être identifiées, à savoir les méthodes

de surclassement et les méthodes fondées sur la théorie de la valeur (multi-attribut).

**Surclassement** L'idée principale des méthodes de surclassement est de comparer deux alternatives quelconques en fonction de leurs évaluations basées sur l'ensemble des critères définis soit par le décideur métier (DM), soit dans notre cas par l'expert métier et selon une règle de majorité. Pour deux alternatives $x$ et $y$ de A, si pour le DM il y a suffisamment d'arguments en faveur de l'affirmation « $x$ est au moins aussi bon que $y$ », alors $x$ surclasse $y$ ($xSy$) [112]. ces arguments reposent sur des différences d'évaluation entre les différents critères sui ont été comparés à des seuils de discrimination fixés selon les préférences du DM. De plus, un poids est associé à chaque critère permettant de donner plus ou moins d'importance à ces arguments locaux dans la règle de majorité. Un indice de concordance permet d'agréger ces arguments *via* une somme pondérée pour obtenir un degré de crédibilité du surclassement. Trois situations de préférence peuvent être déduites de cette relation de surclassement. $x$ et $y$ sont considérés comme indifférents si $xSy$ et $ySx$. Ils sont considérés comme incomparables par rapport aux informations disponibles si aucun surclassement ne peut être confirmé entre eux (ni $xSy$ ni $ySx$). $x$ (resp. $y$) est strictement préféré à y (resp. $x$) si $xSy$ et non $ySx$ (resp. $ySx$ et non $xSy$). Comme cette relation de surclassement n'est pas nécessairement complète ou transitive, son exploitation en vue de l'élaboration d'une recommandation de décision est, en général, assez difficile. De nombreuses procédures d'exploitation ont été proposées dans la littérature pour résoudre les trois principaux types de problèmes de décision multi-critère mentionnés précédemment.

**Théorie de la valeur multi-attribut** Les méthodes fondées sur la théorie de la valeur multi-attribut visent à construire une représentation numérique des préférences du DM sur l'ensemble des alternatives $A$. Ces techniques cherchent plus formellement à modéliser les préférences du décideur, supposé être un pre-ordre représenté par la relation binaire $\succeq$ sur $A$, au moyen d'une fonction de valeur globale $U : A \to \mathbb{R}$ tel que $x \succeq y \Leftrightarrow U(x) \geq U(y),\ \forall x, y \in A$. La fonction de valeur globale $U$ peut être déterminée *via* de nombreuses méthodes comme l'on fait, par exemple, von Winterfeldt et Edwards [35] dans le contexte d'un modèle de fonction de valeur additive. Idéalement, ces méthodes devraient consister en une discussion avec le décideur en utilisant le champs sémantique de son expertise et en évitant les questions techniques liées au modèle utilisé. Notez que la relation de préférence induite par une telle fonction de valeur globale est nécessairement un pre-ordre complet. Cela qui signifie que seules deux situations de préférence peuvent se produire : soit $x$ et $y$ sont considérés comme indifférents (si $U(x) = U(y)$), soit $x$ (resp. $y$) est strictement préféré à $y$ (resp. $x$) si $U(x) > U(y)$ (resp. $U(y) > U(x)$). Une fois que la fonction de valeur globale a été correctement déterminée, son utilisation pour la recommandation de décision est généralement simple, toutes les alternatives étant devenues comparables.

Les principales différences entre ces deux écoles méthodologiques résident dans la manière dont les alternatives sont comparées et dans le type d'informations requises du décideur. En outre, les méthodes de substitution pourraient être préférables si les évaluations des options sur les critères sont principalement qualitatives et si le décideur souhaite inclure une certaine imprécision dans ses préférences dans le modèle. Les méthodes fondées sur la valeur peuvent être privilégiées si un comportement compensatoire du décideur doit être modélisé.

Nous allons dans cette partie appliquer cette méthodologie à un exemple de gestion d'anomalies.

### 5.2.1 Identification du problème

La problématique de l'évaluation est due aux multiples associations entre un processus support et une donnée contextuelle. Notre objectif global est de créer une relation entre un processus opérationnel et un processus support afin de déclencher ce dernier lors d'une anomalie. Il faut donc évaluer l'intérêt du processus support par rapport à l'anomalie. Tout en sachant que, les préférences peuvent varier d'un expert à l'autre et que pour un expert métier donné, la gravité de l'anomalie va également impacter ses préférences. En effet, si l'anomalie a un fort impact sur l'entreprise, par exemple en terme d'image, il sera acceptable de déclencher un processus support coûteux, alors que si l'anomalie impacte peu l'entreprise, les solutions à faible coût seront à privilégier.

Par rapport au vocabulaire introduit dans l'état de l'art (section 2.4), dans notre contexte, le décideur est l'expert métier. Il s'agira donc de modéliser ses préférences, et de les utiliser afin de paramétrer un algorithme d'aide multi-critère à la décision.

Dans notre problème, nous souhaitons évaluer les processus supports afin de recommander à l'expert métier un ensemble de solutions adaptées à son besoin. Une méthode de classement (ordre total) ou une méthode de choix ne nous semblent donc pas appropriées. En effet, les processus supports seraient classés du meilleur au plus mauvais, ou bien un processus serait désigné comme le meilleur processus support. Ceci laisse peu de marge à l'expert métier puisqu'en recommandant un classement à l'expert métier, ce dernier sera tenté de prendre systématiquement le premier.

Dans notre cas, la problématique de décision du tri nous semble donc plus intéressante puisqu'elle laisse ainsi plus de flexibilité à l'expert métier dans la sélection des processus supports en lui recommandant un ensemble de processus considérés comme les meilleurs. Il aura ainsi la possibilité de choisir parmi ces alternatives, le meilleur processus en fonction de l'anomalie tout en ayant connaissance de celles qui ne sont pas recommandées.

Pour terminer, les alternatives de décisions sont donc les processus supports.

### 5.2.2 Formulation du problème

Afin de représenter la variabilité des préférences de l'expert métier en fonction de la gravité de l'anomalie, l'évaluation des processus supports se fait en deux étapes.

La première étape se charge d'évaluer la gravité de l'anomalie en lui assignant une catégorie.

La seconde, a comme objectif, en fonction de la gravité de l'anomalie, de choisir un modèle des préférences de l'expert métier afin de catégoriser les processus supports comme étant pertinent ou non. Nous résumons, ce processus est représenté par la Figure 5.1.

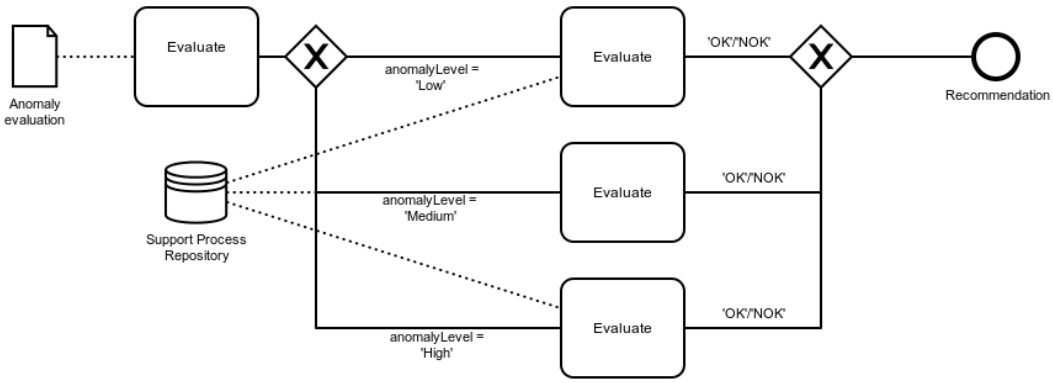


FIGURE 5.1 – Représentation du résultat du processus d'application de l'AMCD.

Les impacts se mesurent selon plusieurs critères qui sont définis en fonction du contexte et des préférences de l'expert métier. Ces critères, peuvent par exemple, être liés à des coûts d'exécution, à la satisfaction des clients ou des employés, à l'impact écologique, à l'image de l'entreprise, ou d'autres critères pouvant découler de la stratégie de l'entreprise.

Ces critères peuvent être de nature quantitative ou qualitative. Par exemple, le coût d'exécution est une valeur quantitative alors que l'image de l'entreprise auprès de ses clients est typiquement une valeur qualitative.

Comme pour chaque algorithme de cette thèse, nous nous intéressons principalement aux méthodes permettant une grande transparence de fonctionnement. En effet, il sera préférable d'avoir un résultat justifiable afin d'inciter l'expert métier à suivre la recommandation.

La nature hétérogène des critères et le besoin de justification sont des arguments en faveur de l'utilisation de méthodes de surclassement comme méthodes d'aide à la décision (voir section 2.4 de l'état de l'art).

Comme évoqué précédemment, notre problématique de décision est associée à une problématique de tri permettant ainsi de catégoriser d'abord les anomalies puis les processus supports.

Parmi l'ensemble des méthodes d'aide à la décision compatibles avec ces caractéristiques, nous avons choisi la méthode *Majority-Rule Sorting* (MR-Sort) [75, 120], qui est une version simplifiée du classique ElectreTRI [87, 88]. Cette méthode nous permet de construire une échelle d'évaluation qualitative globale pour l'évaluation des contributeurs, tout en présentant un modèle très lisible et opérationnel. Cette méthode permet d'affecter des alternatives de décision (ici les processus supports) à des catégories prédéfinies, en tenant compte des préférences du décideur (ici l'expert métier).

La méthode MR-Sort prend ainsi en compte les préférences du DM qui sont représentées par les paramètres préférentiels suivants :

- la pondération des critères (w), qui donne l'importance relative des critères ;
- le seuil de majorité ($\lambda$), qui indique le poids d'un ensemble de critères ;
- les limites de catégorie (b), qui sont utilisées pour la comparaison entre différentes alternatives et déduire l'affectation.

Considérons un ensemble fini d'alternatives $A$, un ensemble fini d'index de critères $J$ et un ensemble de limites de catégories séparant les profils $B = b_1, ..., b_{k-1}$. Chaque alternative et chaque limite de catégorie sont des vecteurs d'évaluations par rapport à tous les critères. L'évaluation par rapport au critère $j$ peut être vue comme une fonction $g_j : A \cup B \rightarrow R$, où $g_j(a)$ désigne l'évaluation de la variante $a \in A$ sur le critère $j$ et $g_j(b_h)$ désigne l'évaluation de la limite de catégorie $b_h, \forall h \in 1, ..., k-1$, sur le critère $j$. L'ensemble des limites de catégories est utilisé pour définir un ensemble de $k$ catégories $c_1, ..., c_k$, classées par leur désirabilité, où $c_1$ est considéré comme étant la pire catégorie et $c_k$ la meilleure. Chaque la catégorie $c_h$ est définie par sa limite supérieure, $b_h$, et sa limite inférieure, $b_{h-1}$, à l'exception des catégories les plus mauvaises et les meilleures, qui n'ont qu'une seule limite. Nous supposons que les performances sont telles qu'une valeur supérieure indique une meilleure performance. De plus, les performances des limites de catégorie ne sont pas décroissantes, à savoir $\forall j \in J$, $1 < h < k : g_j(b_{h-1}) \leq g_j(b_h)$. Une alternative $a$ est dite supérieure à la limite de catégorie $b_{h-1}$ si et seulement si il y a une coalition suffisante de critères appuyant l'affirmation « $a$ est au moins aussi bon que $b_{h-1}$ ». Une coalition de critères correspond à un sous-ensemble de critères qui « s'accordent » sur la manière dont une alternative se compare à une limite de catégorie, qu'elle soit au moins aussi bonne ou strictement pire. Pour mesurer cela, nous définissons pour chaque critère $j$ une fonction $C_j : A \times B \rightarrow 0, 1$ qui détermine si le critère j prend en charge cette affirmation ou non :

$$\forall j \in J, a \in A, 1 \leq h \leq k : C_j(a, b_{h-1}) = \begin{cases} 1, \text{if} g_j(a) \geq g_j(b_{h-1}), \\ 0, \text{autrement.} \end{cases} \tag{5.1}$$

Pour déterminer si une coalition de critères favorise ou non le surclassement, $\forall a \in A, 1 \leq h \leq k$, nous définissons d'abord la concordance globale comme étant :

$$C(a, b_{h-1}) = \sum_{j \in J} w_j C_j(a, b_{h-1}) \tag{5.2}$$

où $w_j$ est le poids du critère j.

Les poids sont définis de manière à être positifs ($w_j > 0, \forall j \in J$) et ont leur somme égale à un ($\sum_{j \in J} w_j = 1$). Cette concordance globale est ensuite comparée à un seuil de majorité $\lambda \in [0.5, 1]$ extrait des préférences du décideur avec les pondérations. Comme nous ne considérons aucune règle de veto ici, la relation de surclassement $S$ est alors définie comme suit :

$$aSb_{h-1} \Leftrightarrow C(a, b_{h-1}) \geq \lambda \tag{5.3}$$

Si $C(a, b_{h-1}) \leq \lambda$, la coalition de critères n'est pas suffisante, l'alternative ne dépasse pas la frontière $b_{h-1}$ et sera donc affectée à une catégorie inférieure à $c_h$. L'alternative a est attribuée à la catégorie la plus élevée qu'elle surclasse. Cette règle peut être écrite comme :

$$a \in c_h \Leftrightarrow aSb_{h-1} \textit{ and not } aSb_h \tag{5.4}$$

Pour que cette règle d'affectation soit utilisée pour les catégories de limites, deux profils fictifs $b_0$ et $b_k$ doivent être ajoutés à $B$, le premier contenant les évaluations les plus basses possible pour tous les critères, tandis que le second conserve les plus élevées possible.

Nous fournissons un exemple simple illustrant cette règle d'affectation (Figure 5.2) provenant de [6]. Nous considérons trois alternatives désignés par $TM_1$, $TM_2$ et $TM_3$, qui sont évalués en fonction de cinq critères avec des échelles ordinales correspondant à leurs niveaux de performance : bon ($g$), neutre ($n$) et mauvais ($b$). Ces échelles sont codées numériquement comme -1 pour $b$, 0 pour $n$ et 1 pour $g$. Nous souhaitons affecter ces alternatives à l'une des trois catégories suivantes : $c_1$ (mauvais), $c_2$ (moyen), $c_3$ (bon). Au sommet, nous observons les paramètres du modèle de tri. Le premier paramètre ($\lambda$) est le seuil de majorité, suivi du vecteur de pondération des critères ($w$) et de deux profils de séparation de catégories ($b_1$ et $b_2$). En dessous des paramètres, nous illustrons l'affectation de trois alternatives en utilisant la règle d'affectation de l'équation (5.4). Dans le cas des première et dernière catégories, nous avons ajouté deux limites fictives afin de simplifier l'exemple. La limite inférieure de la catégorie la plus défavorable ($b_0$) est considérée comme ayant les évaluations les plus basses possibles pour tous les critères (dans ce cas, $b$), alors que le profil supérieur de la meilleure catégorie ($b_3$) est considéré comme ayant les évaluations toujours meilleures que toutes les alternatives en $A$ sur chaque critère. $TM_1$ est au moins aussi bonne que $b_1$ sur 4 critères sur 5 et au moins égal à $b_2$ sur 3. Par conséquent,

$\lambda = \frac{3}{5} \quad w = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}) \quad b_1 = (n, n, n, n, n) \quad b_2 = (g, g, g, g, g)$

| $A$ | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ | $a_i \, S \, b_0$ | $a_i \, S \, b_1$ | $a_i \, S \, b_2$ | $a_i \, S \, b_3$ | assignment |
|---|---|---|---|---|---|---|---|---|---|---|
| $SC_1$ | $g$ | $g$ | $g$ | $n$ | $b$ | ✓ | ✓ | ✓ | × | $c_3$ (good) |
| $SC_2$ | $g$ | $n$ | $n$ | $n$ | $g$ | ✓ | ✓ | × | × | $c_2$ (medium) |
| $SC_3$ | $b$ | $b$ | $g$ | $b$ | $n$ | ✓ | × | × | × | $c_1$ (bad) |

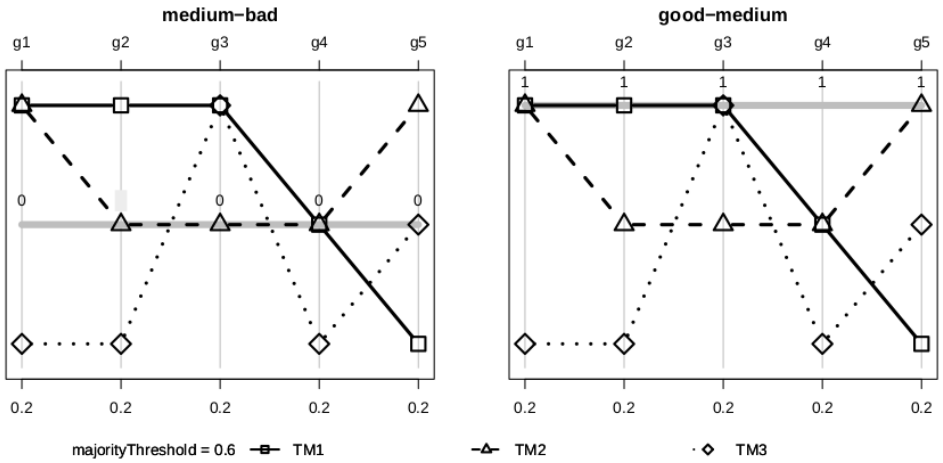


FIGURE 5.2 – Exemple illustrant la procédure d'affectation MR-Sort. [6]

$TM_1$ surpasse $b_1$ et $b_2$ et est donc classé dans la catégorie la plus élevée. La deuxième alternative est au moins aussi bonne que $b_1$ pour tous les critères, mais au moins aussi bon que $b_2$ pour deux critères seulement. Par conséquent, $TM_2$ dépasse, $b_1$ mais pas $b_2$ et est donc affecté à la deuxième catégorie. Enfin, la dernière alternative est au moins aussi bonne que $b_1$ pour deux critères seulement et est donc classée dans la catégorie la plus basse. L'illustration qui sert de support visuel contient deux graphiques de la relation entre les trois alternatives et les profils séparant chaque paire consécutive de catégories. Les profils de séparation sont tracés en noir avec des marqueurs ronds remplis, tandis que les alternatives sont colorées en fonction de leur affectation de catégorie et contiennent chacun un marqueur différent.

### 5.2.3 Modèle d'évaluation

Le modèle d'évaluation que nous proposons pour répondre aux enjeux décrits dans la section précédente est illustré dans la Figure 5.3. Il s'agit d'un processus d'aide à la décision en 2 grandes étapes. En entrée du processus d'aide à la décision, nous avons l'évaluation d'une anomalie en fonction de multiples critères. Cette évaluation est alors triée à l'aide d'un premier algorithme MR-Sort en trois catégories (« Haute », « Moyenne », « Faible »)

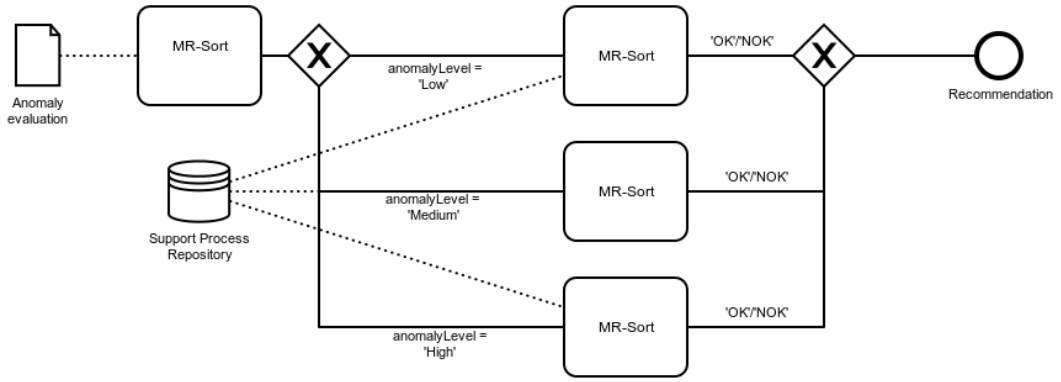


FIGURE 5.3 – Représentation du modèle d'évaluation.

qui représente l'impact global de l'anomalie sur le déroulement du processus opérationnel. En fonction du résultat de ce premier tri, une configuration des préférences de l'expert métier est choisie. Un deuxième algorithme MR-Sort est alors paramétré selon ces préférences pour trier les processus supports en deux catégories (« OK », « NOK »). Les métriques utilisées pour réaliser ce second tri sont la différence entre l'évaluation d'un processus support avec l'évaluation de l'anomalie qui lui est associé. Ceci permet alors d'évaluer l'intérêt de mettre en oeuvre au moins un processus support par rapport à l'anomalie.

Chaque algorithme MR-Sort du processus proposé doit être configuré selon les préférences de l'expert métier. Pour cela, 2 choix sont possibles : soit une configuration directe de l'algorithme en construisant les paramètres préférentiels en collaboration avec l'expert, soit en apprenant ces paramètres de manière indirecte à partir d'exemples d'affectation (comme suggéré par [75, 82]).

Afin de valider l'approche choisie, nous avons utilisé la première méthode en nous fondant sur notre double expertise « aide à la décision » et « processus métier ». Cependant, dans le cadre d'une mise en exploitation, il sera préférable de se tourner vers les méthodes d'apprentissages à partir d'exemples d'affectation. En effet, dans le cadre d'une entreprise, la détermination des paramètres de MR-Sort peut s'avérer compliquée car cela n'est pas du domaine de l'expert métier. De plus, les méthodes d'apprentissage permettent de prendre en compte les subtilités issues des exemples opérationnels et se fonder sur des décisions passées. Nous avons, dans le cadre de notre implémentation de notre outil permettant l'AMCD, implémenté les deux possibilitées de configuration (manuelle ou par apprentissage).

### 5.2.4 Expérimentation

Nous allons maintenant appliquer les notions vues ci-dessus au processus opérationnel d'accès à un service. Nous nous basons sur des données générées pour le processus opéra-

tionnel, les processus supports et leurs contextes.

Dans le cadre du pilotage de l'entreprise et du métier, la stratégie de l'entreprise oriente l'expert métier dans la définition des critères d'évaluation de l'impact d'une anomalie ou d'un processus métier sur l'entreprise. L'entreprise de notre exemple a trois grands axes stratégiques. Ces axes sont la satisfaction client, la satisfaction des employés à travailler dans l'entreprise et la gestion des coûts. Nous définissons ainsi cinq critères représentant ces trois axes et auxquels s'ajoutent le point de vue financier ainsi que l'image de marque de l'entreprise.

Pour résumer nous avons donc les cinq critères suivants :

- **Le coût financier** (C1) : Le coût financier associé à la mise en place et à l'exécution de la solution, ici des processus supports (ou d'une anomalie). Le critère est définit entre 0 et 100, avec 0 correspondant à un impact nul et 100 un impact très fort ;
- **L'impact écologique** (C2) : L'impact écologique se calcul à l'aide de trois sous-critères, la quantité de déchets générés par un processus support (ou par une anomalie), la quantité d'électricité consommée et la quantité de $C0_2$ rejetée. Le critère est définit entre 0 et 5, 0 étant un impact nul et 5 un impact très fort ;
- **La satisfaction client** (C3) : L'impact négatif d'un processus support (ou d'une anomalie) sur les clients de l'entreprise. Le critère est définit entre 0 et 5, 0 étant un impact nul et 5 un impact très fort ;
- **Le niveau de stress des employés** (C4) : Le niveau de stress induit par l'application d'un processus support (ou d'une anomalie). Le critère est définit entre 0 et 10, 0 étant un impact nul et 10 un impact très fort ;
- **L'image de marque de l'entreprise** (C5) : L'impact d'un processus support (ou d'une anomalie) sur la représentation de l'entreprise perçue par les clients. Le critère est définit entre 0 et 5, 0 étant un impact nul et 5 un impact très fort.

Comme on peut le voir dans la description des critères, ces derniers n'ont pas la même échelle d'évaluation.

Table 5.1 – Critères d'évaluation (variable criteriaMinMax pour la première ligne et criteriaWeights pour la seconde ligne en R).

| C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|
| min | min | min | min | min |
| 3 | 0 | 2 | 1 | 1 |

Comme le montre la Table 5.1, les critères ainsi définis sont ensuite intégrés à notre outil en définissant si le critère doit être minimisé ou maximisé. En effet, l'AMCD permet

de minimiser et/ou maximiser les critères, c'est-à-dire définir s'il l'on souhaite qu'un critère soit le meilleur possible lorsqu'il est le plus grand possible (maximiser) ou lorsqu'il est le plus petit possible (minimiser). Il est aussi possible de définir des poids pour pondérer les critères et ainsi prioriser certains. Dans notre exemple, l'entreprise souhaite avoir le plus faible coût financier possible, c'est pourquoi nous avons choisi ce critère comme le critère le plus important en lui assignant une note de 3.

Maintenant que les critères d'évaluation ont été définis, intéressons nous aux anomalies du processus opérationnel. Les anomalies survenues dans notre jeu de données sont :

- Un temps de réponse trop grand (A1) ;
- Un problème de synchronisation (A2).

A l'aide des données contextuelles, nous avons pu sélectionner trois processus supports permettant de résoudre les anomalies. Ces trois processus supports sont :

- Processus de « échange d'appareil » (PS1) : Ce processus permet à un utilisateur d'échanger d'appareil pour un fonctionnel ;
- Processus de « sollicitation du support utilisateur » (PS2) : Ce processus permet de déclencher la sollicitation du support utilisateur aux travers d'un appel ;
- Processus de « re-synchronisation » (PS3) : Ce processus permet de relancer une tentative de synchronisation de la ligne.

L'expert métier a eu l'occasion lors de la modélisation des processus supports de les évaluer en fonction des critères précédemment définis. Le résultat de cette évaluation est enregistré dans un fichier dans le format CSV, et est montré dans la Table 5.2.

Table 5.2 – Évaluation des processus supports en fonction de nos critères d'évaluations (variable « performanceTable » en R).

| ID | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| PS1 | 62 | 1.1 | 2 | 3 | 1 |
| PS2 | 25 | 0.1 | 1 | 3 | 1 |
| PS3 | 5 | 0 | 0 | 0 | 0 |

Nous avons donc précédemment explicité les anomalies survenues sur notre processus opérationnel. Nous allons maintenant évaluer ces anomalies en fonction des critères, comme nous l'avons fait avec les processus supports. Cette évaluation est réalisée lors de la découverte de l'anomalie par l'expert métier. Le résultat de cette évaluation est enregistré dans un fichier dans le format CSV, et est montré dans la Figure 5.3.

Table 5.3 – Évaluation des anomalies en fonction de nos critères d'évaluations (variable performanceTable en R).

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| A1 | 0 | 0 | 3.2 | 0 | 3 |
| A2 | 0 | 0 | 3.0 | 0 | 3 |

Afin de pouvoir prendre la décision si une anomalie a un impact fort, moyennement fort ou faible sur l'entreprise, il est nécessaire de définir un profil. Ce profil découle de la stratégie de l'entreprise et des préférences de l'expert métier qui est ici notre décideur. Ce profil est illustré par la Table 5.4

Table 5.4 – Profil des niveaux d'impact en fonction de nos critères (variable profiles en R).

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| High Impact | 50 | 2 | 2 | 2 | 2 |
| Medium Impact | 25 | 1 | 1 | 1 | 1 |
| Low Impact | NA | NA | NA | NA | NA |

Avant de lancer l'algorithme MRSort, il est nécessaire de trier les niveaux d'impact (code R) :

```
#Creation de la variable categoriesRank qui classe les niveaux
#d'impact les uns par rapport aux autres.
categoriesRank <- (1, 2, 3)
names(categoriesRank) <- ("High Impact","Medium Impact",
                          "Low Impact")
```

Nous avons maintenant l'ensemble des éléments nécessaires pour appliquer notre premier algorithme de tri, MRSort. L'ensemble des fichiers précédemment générés sont dans un premier temps lu, puis nous appliquons l'algorithme MRSort en se fondant sur le « profil » définit précédemment (code R) :

```
#Appel de la fonction MRSort du package R - MCDA
assignments <- MRSort(performanceTable, profiles, categoriesRank,
                      criteriaWeights, criteriaMinMax)
```

Nous retrouvons dans cette ligne de code R, la variable *performanceTable* qui fait référence à la Table 5.3, *profiles* fait référence quant à lui à la Table 5.4 enfin *criteriaWeights* et *criteriaMinMax* font référence à la Table 5.1.

En appliquant l'algorithme, les anomalies A1 et A2 ont reçu l'assignation « Medium ».

Après l'assignation des anomalies dans les catégories « High », « Medium », « Low », le résultat est sauvegardé dans un fichier CSV.

Il ne reste plus qu'à utiliser l'algorithme de tri sur les évaluations des processus supports afin de pouvoir ensuite prendre la décision d'appliquer ou non ces solutions, c'est-à-dire lancer l'exécution des processus supports. Pour réaliser ce tri, nous avons besoin de trois nouveaux profils correspondant aux trois niveaux d'impact ainsi que de calculer la différence des évaluations des processus supports par rapport à l'anomalie associée. Ces profils, comme le précédent, sont définis en se fondant sur la stratégie de l'entreprise et les préférences de l'expert métier. Chaque profil contient deux catégories, « OK » et « NOK ». Lorsqu'un processus support reçoit l'assignation « OK », cela signifie qu'il est recommandé d'appliquer ce processus support à l'anomalie. Lorsqu'un processus reçoit l'assignation « NOK », cela signifie qu'il est recommandé de ne pas appliquer ce processus support. Nous nommerons Delta C$x, \forall x \in 1..5$, la différence entre l'évaluation du critère C$x$ d'un processus support et l'évaluation du critère C$x$ pour l'anomalie associée à ce processus support.

Table 5.5 – Profil « High » en fonction de critères.

| | Delta C1 | Delta C2 | Delta C3 | Delta C4 | Delta C5 |
|---|---|---|---|---|---|
| NOK | 30 | 2.3 | 2 | 1 | 1.5 |
| OK | NA | NA | NA | NA | NA |

Table 5.6 – Profil « Medium » en fonction de critères.

| | Delta C1 | Delta C2 | Delta C3 | Delta C4 | Delta C5 |
|---|---|---|---|---|---|
| NOK | 20 | 0.5 | 0 | 1.5 | 0 |
| OK | NA | NA | NA | NA | NA |

Table 5.7 – Profil « Low » en fonction de critères.

| | Delta C1 | Delta C2 | Delta C3 | Delta C4 | Delta C5 |
|---|---|---|---|---|---|
| NOK | 10 | 0 | 0 | 0.5 | 0 |
| OK | NA | NA | NA | NA | NA |

En appliquant ces profils aux processus supports, les processus supports PS2 et PS3 sont classés comme étant « OK », et PS1 est classé dans la catégorie « NOK ».

La recommandation issue de notre algorithme est donc d'exécuter les processus supports PS2 et PS3 lorsque leurs anomalies respectives sont détectées. Concernant le processus support PS1, l'algorithme recommande de ne pas l'exécuter puisqu'il ne répond pas aux préférences des experts métiers.

## 5.3 Conclusion

Dans ce chapitre, nous avons proposé une méthode permettant la sélection d'un ensemble de processus supports associés à un processus opérationnel à l'aide de ses données contextuelles. Puis, nous avons proposé un modèle d'évaluation permettant d'évaluer chaque processus support en fonction des anomalies issues du partitionnement de traces. Cette évaluation recommande à l'expert métier les meilleurs processus supports à mettre en oeuvre. Ceci permet de répondre à notre troisième question de recherche (QR3).

Maintenant que nous avons pu voir comment sélectionner les processus supports, nous allons expliquer comment concevoir cette relation entre le modèle d'un processus opérationnel et le modèle d'un processus support dans le cadre de l'Ingénierie dirigée par les Modèles.

Chapitre 6

# Développement de l'outil d'enrichissement d'un processus opérationnel par des processus supports

Ce chapitre décrit le développement de l'outil d'enrichissement d'un processus opérationnel. L'enrichissement d'un processus opérationnel fait référence à la modélisation des données associées à un processus opérationnel ou à un processus support, à l'introduction du contexte via l'ontologie (QR1), l'enrichissement d'un processus opérationnel par un ou plusieurs processus supports à partir de ce contexte (QR2, QR3).

Le développement de l'outil est fondé sur la méthode Unified Process (UP) [62]. Cette méthode est itérative. Elle consiste à analyser les fonctionnalités de l'outil puis à architecturer applicativement l'outil, et enfin, l'implémenter, lors de chaque itération. Chaque itération est composée de l'analyse fonctionnelle, de l'architecture technique et applicative et de la programmation.

Plus précisément, nous mettons en oeuvre la méthode Enterprise Architecture for Unified Process (EA4UP) [117] afin de prendre en compte les réutilisations de solutions externes telles que les ontologies, les services web et les services opérationnels. Il s'agit d'une mise en œuvre opérationnelle de la méthode UP. Cette méthode comporte les phases d'analyse fonctionnelle de l'outil (statique et dynamique), d'architecture applicative de l'outil (statique et dynamique) contrainte par l'architecture technique et réalisant l'analyse fonctionnelle. La phase de vérification de cohérence de l'architecture applicative par rapport à l'analyse fonctionnelle complète la méthode. Nous utiliserons le langage UML pour représenter le résultat de chaque phase.

Une originalité de ce développement est de spécifier les fonctionnalités offertes à l'expert métier à partir de l'approche d'Ingénierie dirigée par les Modèles. En effet, nous associerons chaque fonctionnalité à une transformation de modèle.

## 6.1 Analyse fonctionnelle de l'outil

L'analyse est fondée sur cinq transformations élicitées dans les deux chapitres précédents (Cf. Figure 6.1).

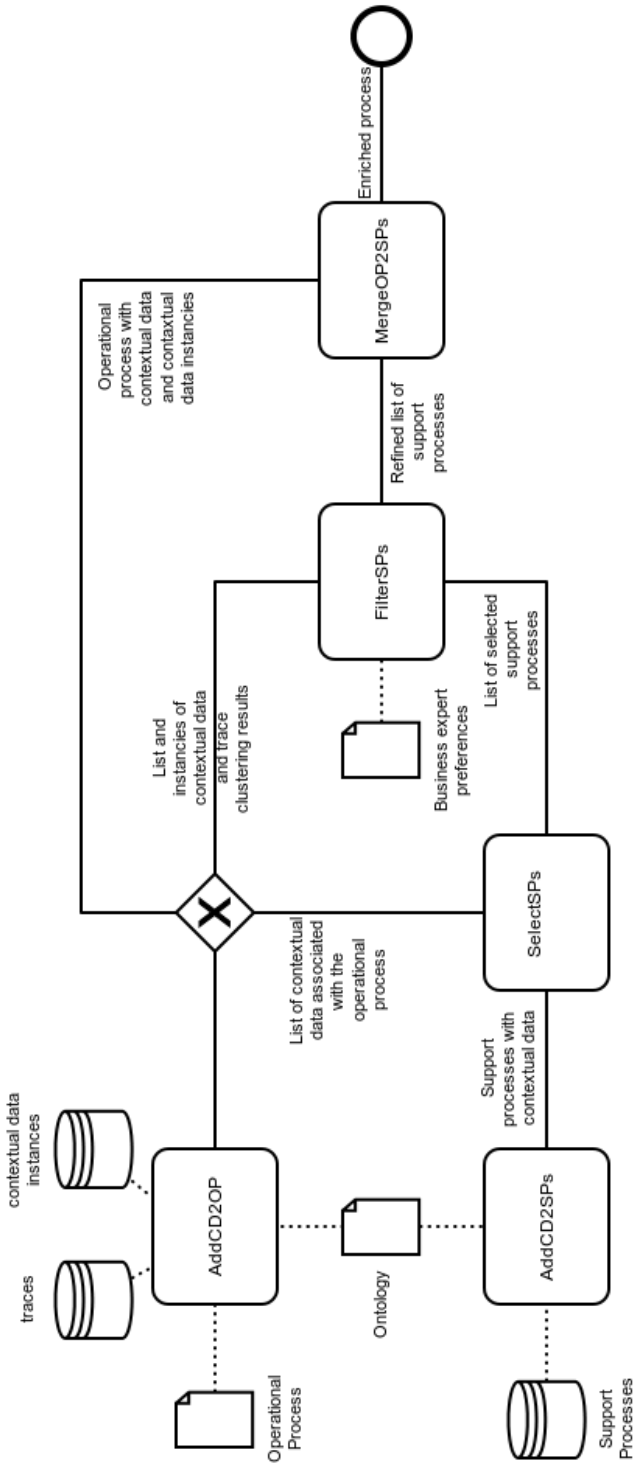


FIGURE 6.1 – Enchaînement des transformations transformant un processus opérationnel en un processus enrichi à l'aide de processus supports.

Les cinq transformations sont :

- **AddCD2OP** : Transformation permettant d'annoter un processus opérationnel à l'aide d'une ontologie et de sélectionner une liste de données contextuelles qui impactent l'exécution ce processus. Elle réalise les algorithmes « SA4BP » (Cf. Figure 4.4) et « ContextExplorer » (Cf. Figure 4.6) ;
- **AddCD2SPs** : Transformation permettant d'annoter des processus supports à l'aide d'une ontologie. Elle réalise l'algorithme « SA4BP » (Cf. Figure 4.4) ;
- **SelectSPs** : Transformation permettant de réduire la liste des processus supports à partir de la liste des données contextuelles du processus opérationnel et des processus supports annotés. Cette transformation fait référence à la Section 5.1 ;
- **FilterSPs** : Transformation permettant de raffiner la liste des processus supports fournie par la transformation « SelectSPs » à partir de l'analyse des traces du processus opérationnel et des préférences de l'expert métier. Elle est associée à l'utilisation de l'aide multi-critère à la décision explicitée Section 5.2 ;
- **MergeOP2SPs** : Transformation permettant de concevoir les interactions entre un processus opérationnel et la liste des processus supports raffinée à partir des instances de données contextuelles.

A partie de ces cinq transformations, nous allons spécifier les cas d'utilisation et les scénarios associés, puis en déduire les entités participantes à ces cas d'utilisations.

### 6.1.1 Analyse fonctionnelle statique de l'outil fondée sur les transformations

La spécification des cas d'utilisation de notre outil d'enrichissement et le diagramme des entités participantes à ces cas d'utilisation (diagramme de classe UML) composent le livrable d'analyse fonctionnelle statique.

#### 6.1.1.1 Cas d'utilisation de conception des données contextuelles associées à un modèle de processus opérationnel (UCAddCD2OP)

Le cas d'utilisation « UCAddCD2OP » représente la conception des données contextuelles associées à un modèle de processus opérationnel. Le rôle contributeur de ce cas d'utilisation est celui de l'expert métier qui le déclenche afin de prendre en compte le contexte d'un processus opérationnel.

La pré-condition de ce cas d'utilisation est un modèle de processus opérationnel modélisé en BPMN.

Afin de décrire le contexte d'un processus opérationnel défini par des séquences d'activité où chaque activité produit ou utilise une donnée opérationnelle et des points de décision, l'expert métier sélectionne les données contextuelles les plus proches de chaque donnée opérationnelle et activité opérationnelle. Cette proximité entre deux données, ou entre une activité et une donnée contextuelle se mesure entre les concepts correspondants définis dans une ontologie. Les relations entre concepts permettent de parcourir l'ontologie (Cf. Section 4.1.2). Les entités participantes sont, pour cette partie, les concepts du méta modèle d'un processus métier (Cf. Figure 4.7) spécialisé pour un processus opérationnel comme illustré par la Figure 6.2.

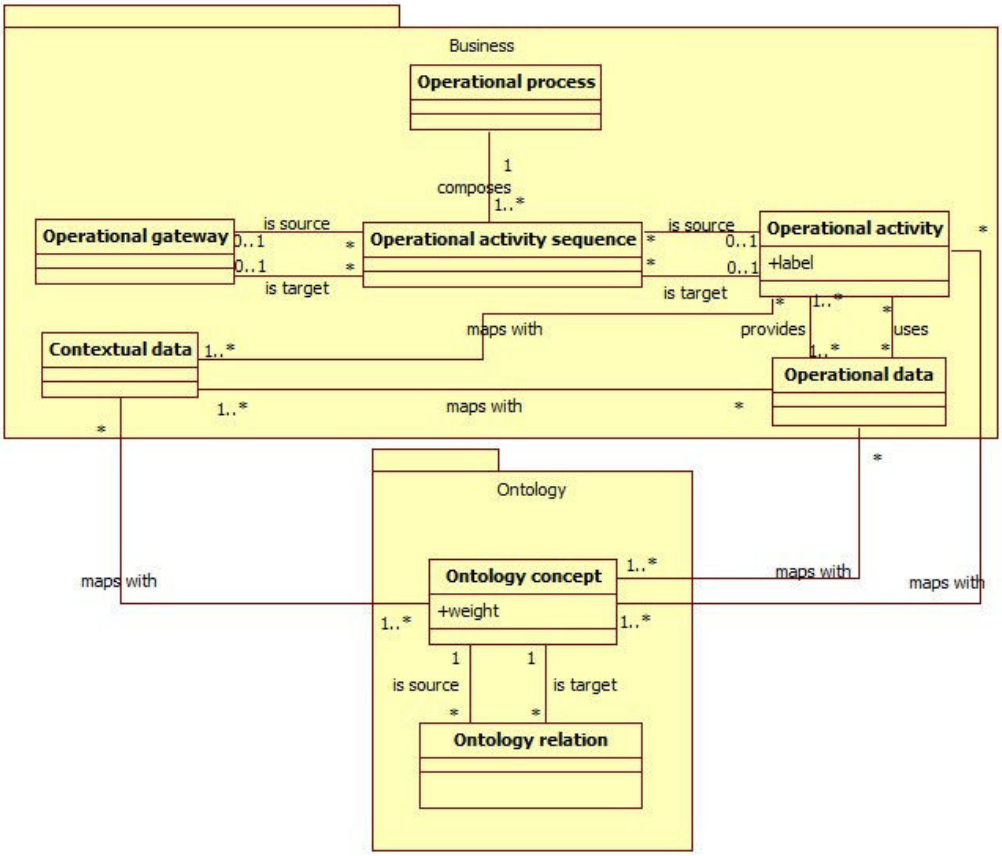


FIGURE 6.2 – Spécialisation du méta modèle d'un processus métier associé à une donnée contextuelle mappée par l'ontologie en un méta modèle processus opérationnel.

L'illustration ci-dessous est celle du méta-modèle du processus métier (Cf. Figure 4.7).

Si l'on se place dans le cas de l'authentification au travers d'une Livebox, nous pouvons instancier notre exemple et ainsi illustrer la Figure 6.2. Dans notre exemple, nous disposons d'un processus d'authentification pour l'usage d'un service. Pour s'authentifier, c'est à la

livebox d'annoncer l'identité de l'utilisateur. Cependant, plusieurs évènements peuvent perturber le bon déroulement de ce processus. L'expert métier a au préalable identifié trois données contextuelles que sont le « Livebox timeout instance » (Cf. Figure 6.3), le « Domain name system timeout instance » (Cf. Figure 6.4), le « synchronisation instance » (Cf. Figure 6.5). Chacune de ces données contextuelles est associée à un service permettant de récupérer la valeur courante et ces données possèdent toutes un seuil au-dessus duquel elle est considérée comme anormale.

L'instanciation des données opérationnelles associée à l'utilisation de la Livebox et du DNS est illustrée par les Figures 6.6 et 6.7.

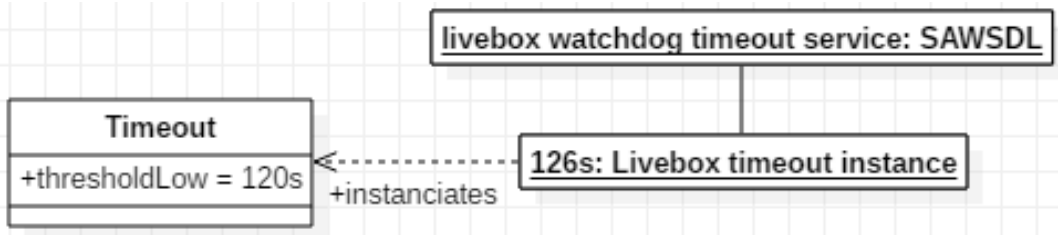


FIGURE 6.3 – Instanciation du timeout de la Livebox.

L'instanciation de la figure 6.3 signifie que suite au dépassement de 6sec lors de la synchronisation la livebox a déclenché un « timeout ». Le « livebox watchdog timeout service » permet de déclarer ce « timeout ».

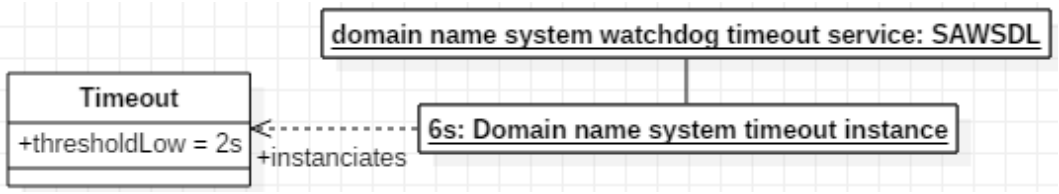


FIGURE 6.4 – Instanciation du timeout du DNS.

L'instanciation de la figure 6.4 signifie que suite au dépassement de 4sec lors de la requête DNS la livebox a déclenché un « timeout ». Le « domain name system watchdog timeout service » permet de déclarer ce « timeout ».

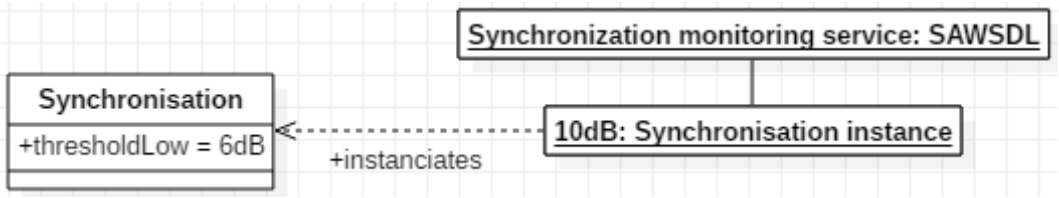


FIGURE 6.5 – Instanciation de la synchronisation.

L'instanciation de la figure 6.5 signifie que suite à un bruit dépassant le seuil de 6dB, la livebox a déclenché un défaut de synchronisation. Le « synchronization monitoring service » permet de déclarer ce défaut.

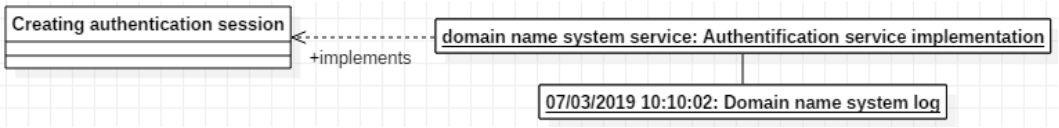


FIGURE 6.6 – Instanciation de l'activité « Creating authentification session » pour le DNS.

L'instanciation de la Figure 6.6 signifie que la trace « Domain name system log » est générée par l'activité « Creating authentification session ». Cette activité est implémentée, entre autres services, par le service « domain name system service ».

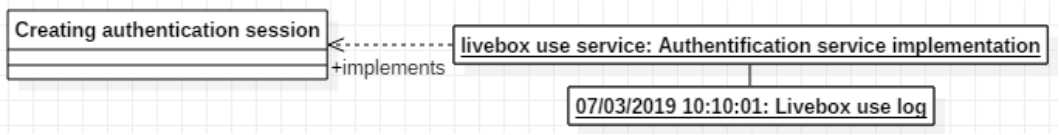


FIGURE 6.7 – Instanciation de l'activité « Creating authentification session » pour la Livebox.

L'instanciation de la Figure 6.7 signifie que la trace « Livebox use log » est générée par l'activité « Creating authentification session ». Cette activité est implémentée, entre autres, par le service « Livebox use service ».

Afin de réduire l'ensemble des données contextuelles à celles qui impactent le déroulement du processus, il est nécessaire de faire appel aux traces (« Operational logs ») issues des « Operationnal service » implémentant les activités du processus et des instances des données contextuelles préalablement sélectionnés par l'expert métier à l'aide de service externe aux processus (« SAWSDL »). Les entités participantes sont, pour cette partie, citées ci-dessus. Ces entités sont représentées dans la Figure 6.8.

Dans notre illustration, l'instance de la donnée contextuelle « 10dB » est instancié par la donnée contextuelle « Noise ». Le service opérationnel « Livebox use service » implémente l'activité opérationnelle « Creating authentification session ».

La contrainte du cas d'utilisation est représentée par les concepts du méta-modèle « Ontology » (Cf. Figure 4.2) ainsi que les traces du processus opérationnel et les instances des données contextuelles.

La post-condition de ce scénario est la relation, d'une donnée opérationnelle ou d'une activité opérationnelle, avec une ou plusieurs données contextuelles, contrainte par une ontologie.

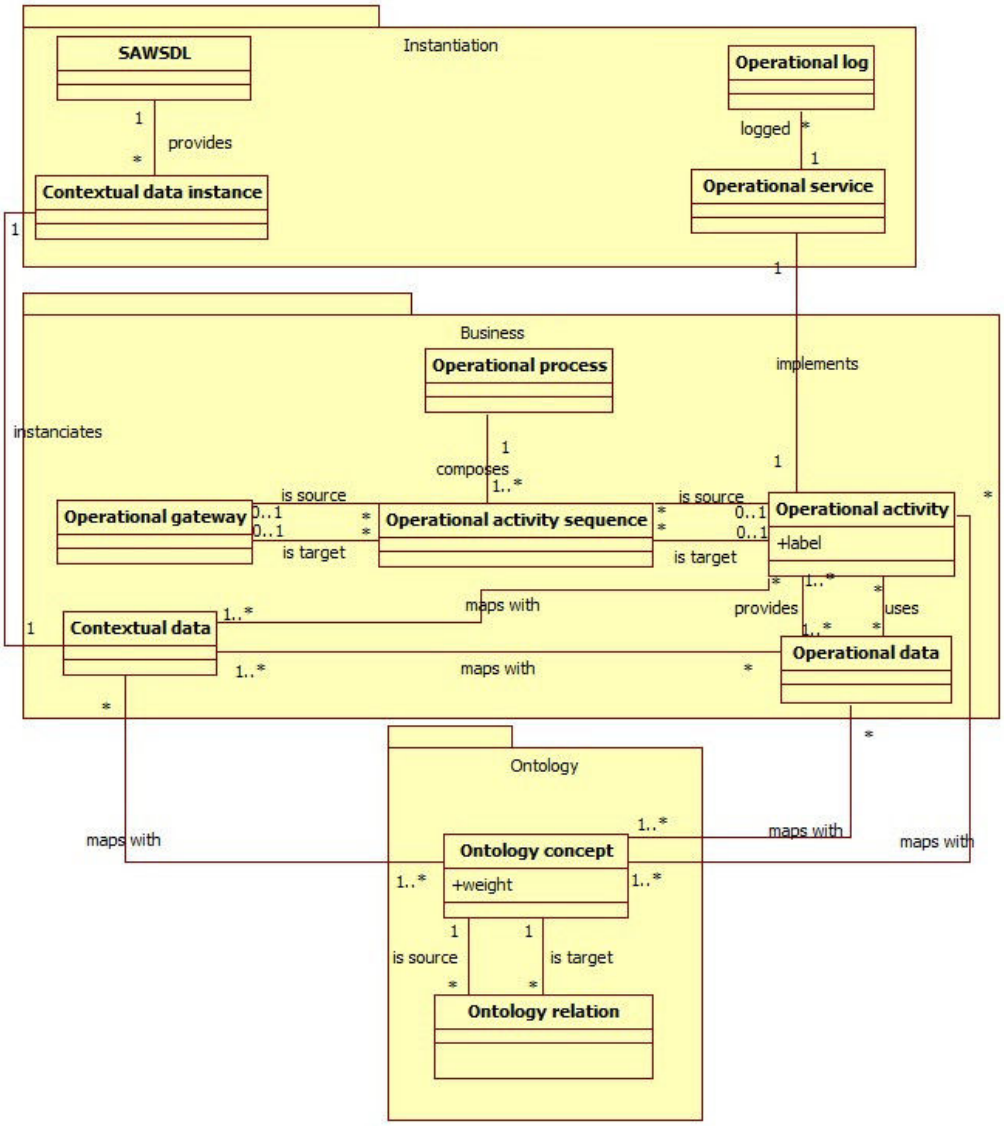


FIGURE 6.8 – Méta-modèle d'un processus opérationnel supportant la transformation « AddCD2OP ». Ce méta-modèle est associé au niveau M0, c'est-à-dire à l'instanciation du niveau M1 (Cf. Figure 3.1), du MOF défini dans le Section 6.1.1.6.

#### 6.1.1.2 Cas d'utilisation de conception des données contextuelles associées à un modèle de processus support (UCAddCD2SPs)

Nous appliquons la même spécification au cas d'utilisation « UCAddCD2SPs » qui représente la conception des données contextuelles associées à un modèle de processus support. Le rôle contributeur est celui de l'expert métier qui le déclenche afin de prendre en compte le contexte d'un processus support.

La pré-condition de ce cas d'utilisation est un modèle de processus support modélisé en BPMN.

Afin de décrire le contexte d'un processus support défini par des séquences d'activité où chaque activité produit ou utilise une donnée support et des points de décision, l'expert métier sélectionne les données contextuelles les plus proches de chaque donnée support et activité support. Cette proximité entre deux données, ou entre une activité et une donnée contextuelle se mesure entre les concepts correspondants définis dans une ontologie. Les relations entre concepts permettent de parcourir l'ontologie (Cf. Section 4.1.2). Les entités participantes de ce cas d'utilisation sont les concepts du méta-modèle représentant les liens entre un processus métier et des données contextuelles (Cf. Figure 4.7) spécialisés pour un processus support dans le méta-modèle (Cf. Figure 6.9).

Les concepts de ce méta-modèle sont illustrés de la manière suivante :

- Support process : « Gestion d'incident d'un service » (3.4) ;
- Support activity : « Create service trouble report », « Correct and resolve service trouble » ;
- Support activity sequence : Séquence ayant pour source « Create service trouble report » et cible « Correct and resolve service trouble » ;
- Support data : « Service trouble report » est produite par « Create service trouble report » et utilisé par « Correct and resolve service trouble » ;
- Support gateway : Il n'y a pas de gateway néanmoins on pourrait concevoir une gateway entre « Create service trouble report » et « Correct and resolve service trouble » indiquant qu'il existe des cas ou la correction d'anomalie n'est pas possible (décision superviseur) ;
- Contextual data : Le « Noise » est aligné avec le concept ontologique « Synchronisation » ainsi qu'avec l'activité support « Correct and resolve service trouble » et la donnée support « Service [resolved] ».

Le cas d'utilisation a donc pour entité participante les concepts du méta-modèle « Ontology » (Cf. Figure 4.2).

La post-condition de ce cas d'utilisation est la relation entre des données supports et des activités avec une ou plusieurs données contextuelles contraintes par une ontologie.

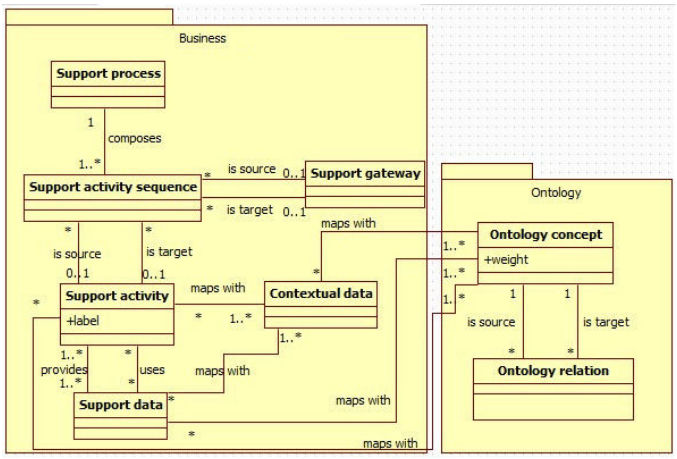


FIGURE 6.9 – Spécialisation du méta-modèle d'un processus métier associé à une donnée contextuelle alignée avec l'ontologie pour un processus support.

#### 6.1.1.3 Cas d'utilisation de création d'un ensemble de processus supports en lien avec un processus opérationnel et une donnée contextuelle (UCSelectSPs)

Ce cas d'utilisation permet de sélectionner les processus supports en liant avec un processus opérationnel via une donnée contextuelle partagée. Le rôle interagissant avec ce cas d'utilisation est celui de l'expert métier.

La pré-condition du cas d'utilisation est à la fois la relation d'une donnée/activité opérationnelle avec une ou plusieurs données contextuelles contrainte par une ontologie et la relation d'une donnée/activité support avec une ou plusieurs données contextuelles contraintes par une ontologie.

La relation entre la donnée opérationnelle et une donnée contextuelle offre un lien entre cette dernière et une activité du processus opérationnel. Il en est de même entre une activité du processus support et une donnée contextuelle. L'attente de l'expert métier est d'obtenir, pour chaque donnée contextuelle partagée par ces deux types de processus, une liste des processus supports associés à un processus opérationnel. Les entités participantes sont celles spécifiées pour les deux cas d'utilisation précédents.

La post-condition du cas d'utilisation est la liste des processus supports associés aux données contextuelles du processus opérationnel.

Dans notre illustration, les processus supports de gestion d'incident de service et de gestion d'incident de ressource sont associés avec le processus opérationnel d'accès à un service.

#### 6.1.1.4 Cas d'utilisation de création d'un ensemble de processus support raffiné à l'aide des préférences de l'expert métier (UCFilterSPs)

Il s'agit ici grâce à l'AMCD de prioriser, pour chaque processus opérationnel, les processus supports associés à une même donnée contextuelle. Cette priorisation est contrainte par rapport aux préférences de l'expert métier. Les niveaux d'anomalies associés aux données contextuelles sont des paramètres de la priorisation.

La pré-condition du cas d'utilisation est la liste des processus supports associés aux données contextuelles du processus opérationnel ainsi que les anomalies issues du scénario nominal « ScAddCD2OP »

Pour des besoins métiers, il y a besoin d'anticiper les cas d'interruption du processus. Pour ce faire, il est utile de hiérarchiser les anomalies du processus détectées par la garde associée aux différents comportements traçant le déroulement d'un processus opérationnel. Cette hiérarchisation est associée à différents seuils de gravité caractérisant le type de contexte (l'organisation, la localisation, le temporel et les ressources). Nous utiliserons l'algorithme AMCD pour obtenir le niveau des anomalies et prioriser les processus supports.

L'entité participante décrivant la donnée contextuelle est composée de trois attributs :

- type : cet attribut fait référence au type de contexte (Organisation, Localisation, Temporel, Ressource). ;
- thresholdHigh, thresholdLow, thresholdEqual : ces attributs correspondant à la valeur des seuils définissant l'impact d'une anomalie lors de l'exécution du processus. Le seuil peut-être un intervalle ou une valeur numérique ou un état de la donnée contextuelle. Par exemple, un temps de synchronisation trop long est qualifié trop long lorsque ce dernier dépasse un certain seuil ;
- anomalyLevel : cet attribut correspond au niveau d'importance de l'anomalie pour le métier.

Dans le cadre de notre exemple, ces instances de donnée contextuelle sont associées à des comportements anormaux que l'on qualifie d'anomalies.

La contrainte est composée des préférences de l'expert métier ainsi que des évaluations des anomalies et des processus supports (Cf. Section 2.4).

La post-condition du cas d'utilisation est la liste définie par les préférences de l'expert métier des processus supports associés aux données contextuelles d'un processus opérationnel.

Dans notre illustration, ce filtrage a conservé les deux processus supports sélectionnés précédemment.

#### 6.1.1.5 Cas d'utilisation de création d'un processus enrichi (UCMergeOP2SPs)

L'objectif de ce cas d'utilisation est de représenter dans un modèle un processus opérationnel enrichi par les processus supports filtrés précédemment.

La pré-condition du cas d'utilisation est l'instanciation des concepts du méta modèle et leurs relations (Cf. Figure 4.7) appliquées à un processus opérationnel et à un processus support.

La première règle d'enrichissement du processus opérationnel consiste à créer une séquence entre une activité du processus opérationnel et une activité du processus support conditionné par une instance d'une donnée conditionnelle.

**Règle 1** (Send.event.PM.to.PS) : On crée une instance de donnée contextuelle qui va déclencher un processus support via un traitement d'anomalie (acteur interne de gestion des incidents)

- Instancier une « Guard gateway » à partir des connaissances de l'expert métier issues des transformations précédentes ;
- Instancier une « Contextual activity sequence » qui va de la « Guard gateway » vers une activité du processus support (Cf. Figure 6.10).

La seconde règle d'enrichissement cible le retour du processus support vers le processus opérationnel.

**Règle 2** (Receipt.event.PM.from.PS) : La validation du traitement de l'anomalie (acteur interne de gestion des incidents) déclenche le retour vers une activité du processus opérationnel.

- Instancier une « Guard gateway » de validation de l'anomalie par l'acteur interne de gestion des incidents ;
- Instancier une « Contextual activity sequence » qui va de la « Guard gateway » vers une activité du processus opérationnel (Cf. Définition 13).

Nous définissons un processus opérationnel enrichi comme un modèle de processus pour lequel ses relations de dépendance sont explicitées et conditionnées par des gardes au sein du modèle.

**Définition 13** (Processus enrichi)**.** *Soit po un processus opérationnel et $PS$ un ensemble de processus de support noté ps, notons $PS(po)$ l'ensemble des processus de support en relation avec po, alors pe un processus enrichi est défini comme le triplet : (po, $PS(po)$, $\gamma$).*

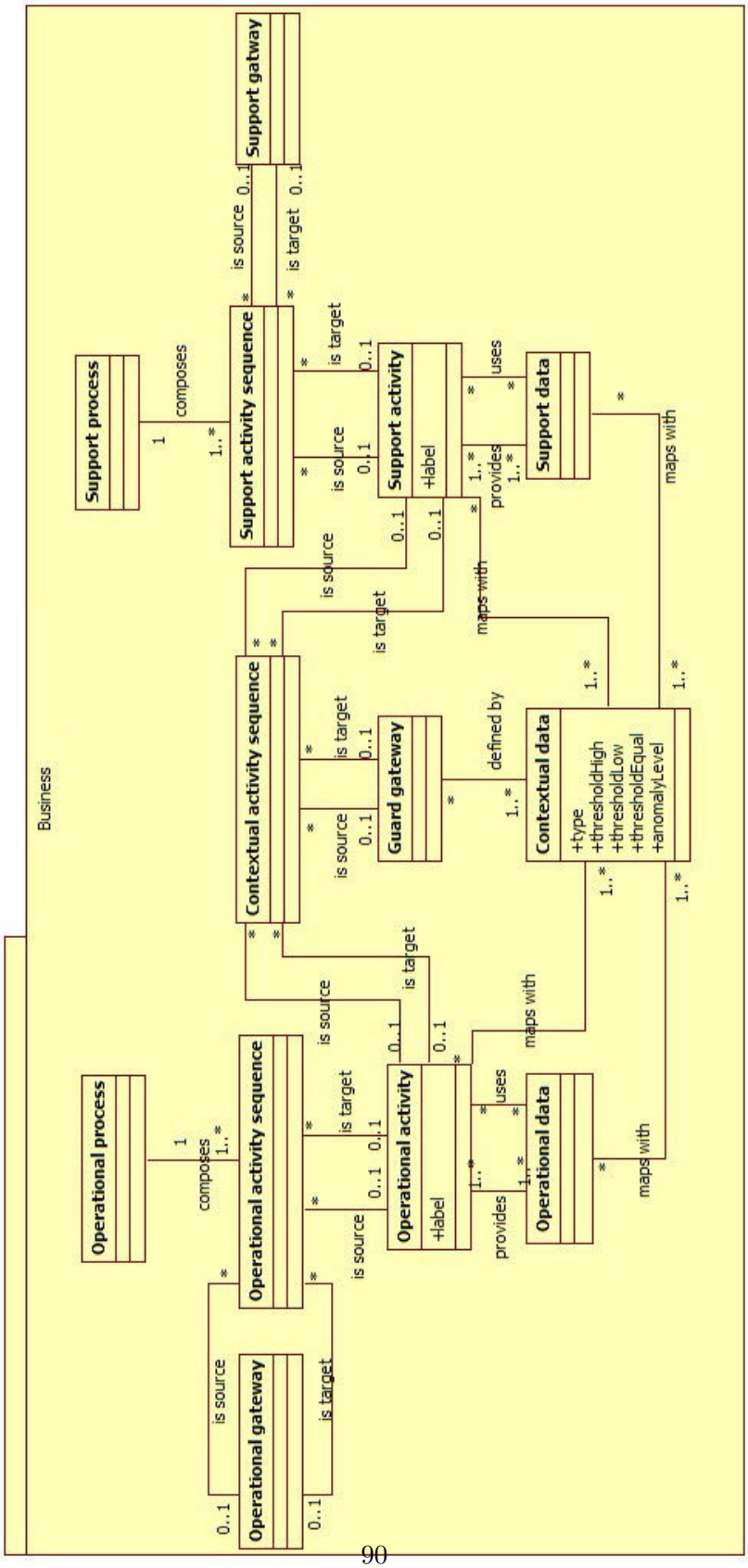




FIGURE 6.10 – Méta modèle d'un processus enrichi.

Ce méta modèle inclut les concepts du méta modèle du processus métier (Cf. Figure 4.7). Il est spécialisé pour un processus opérationnel et pour un processus support auxquels les concepts de garde et de séquence induite par ces gardes ont été ajoutés afin de modéliser l'enrichissement proposé.

Afin d'illustrer nous proposons dans la Figure 6.11 le résultat de l'enrichissement du processus opérationnel d'accès à un service par les deux processus supports de gestion des incidents associés soit à une ressource, soit à un service.

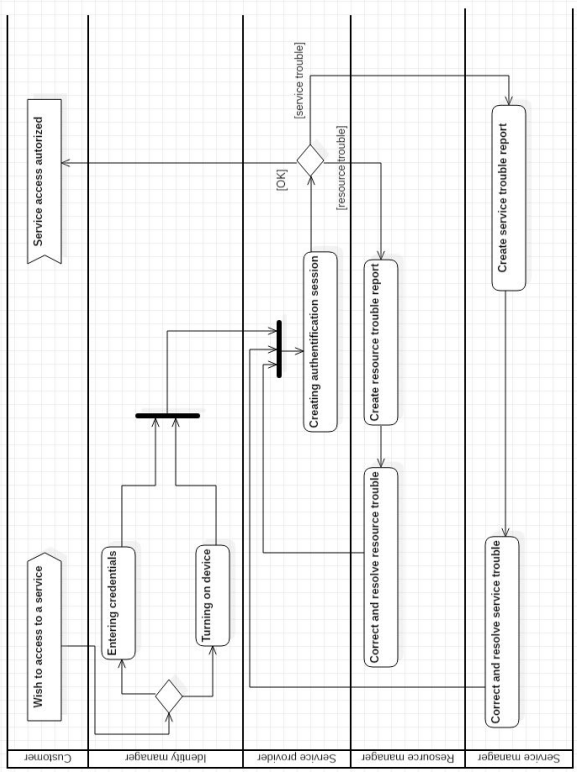


FIGURE 6.11 – Instanciation de notre illustration fondée sur les instanciations Figure 6.3, 6.4, 6.5, 6.6 et 6.7.

On retrouve donc ces deux processus reliés au processus opérationnel au travers d'une gateway. Cette gateway est donc issue des différentes transformations explicitées dans ce chapitre.

La « guard gateway » est illustrée par le point de décision entre l'activité « Creating authentification session », en amont, et les activités « Create resource trouve report » et « Create service trouble report », en aval. Elles forment ainsi deux séquences d'activités contextuelles (« Contextual activity sequence ») : de « Creating authentification session » vers « Create resource trouve report » et de « Creating authentification session » vers « Create service trouble report ». Cette « guard gateway » est définie pour la don-

née contextuelle « Noise » ayant pour seuil bas (attribut « thresholdLow ») de 6dB. Cette donnée contextuelle est alignée d'une part avec l'activité opérationnelle « Creating authentification session » et la donnée opérationnelle « Service », et d'autre part avec l'activité support « Create service trouble report » et la donnée support « Service[resolved] ».

La post-condition du cas d'utilisation est un ensemble de séquences d'activités contextuelles entre un processus opérationnel et un processus support associé à une garde sur une ou plusieurs données contextuelles. Les processus supports pris en compte sont ceux de la liste raffinée lors de la transformation « FilterSPs ».

En nous fondant sur l'analyse fonctionnelle qui reformule nos chapitres 4 et 5, nous pouvons en déduire le diagramme des entités participantes.

#### 6.1.1.6 Diagramme des entités participantes

Le diagramme des entités participantes (Cf. Figure 6.12) regroupe toutes les entités participantes des cinq cas d'utilisations ciblés par l'analyse fonctionnelle statique.

Nous avons empaqueté les entités participantes dans deux catégories suivant les niveaux M0 et M1 du standard Meta-Object Facility (MOF) proposé par l'OMG. Nous utilisons trois des infrastructures à quatre couches[1] proposées par ce standard afin de représenter notre modèle. Cette infrastructure à quatre couches vise à distinguer les quatre niveaux de modélisation d'un modèle. Il est composé par :

- M0 est le calque du « monde réel ». Dans notre cas, c'est l'instance composée d'au moins un événement ;
- M1 est la couche modèle. Pour nous, il est relatif au modèle de processus, par exemple le modèle de notre processus 3.2 ;
- M2 est la couche de méta-modèle telle que définie par Figure 6.12 qui montre les différentes entités participantes de notre architecture ;
- M3 est le méta-méta-modèle qui décrit la modélisation du méta-modèle M2. Dans notre travail, ce dernier est le méta-modèle de l'UML qui ne nous semble nécessaire de décrire.

Ces entités participantes vont être instanciées lors de l'analyse fonctionnelle dynamique pour chaque scénario nominal illustrant un cas d'utilisation.

1. https ://www.omg.org/mof/, visité le 11/10/2018

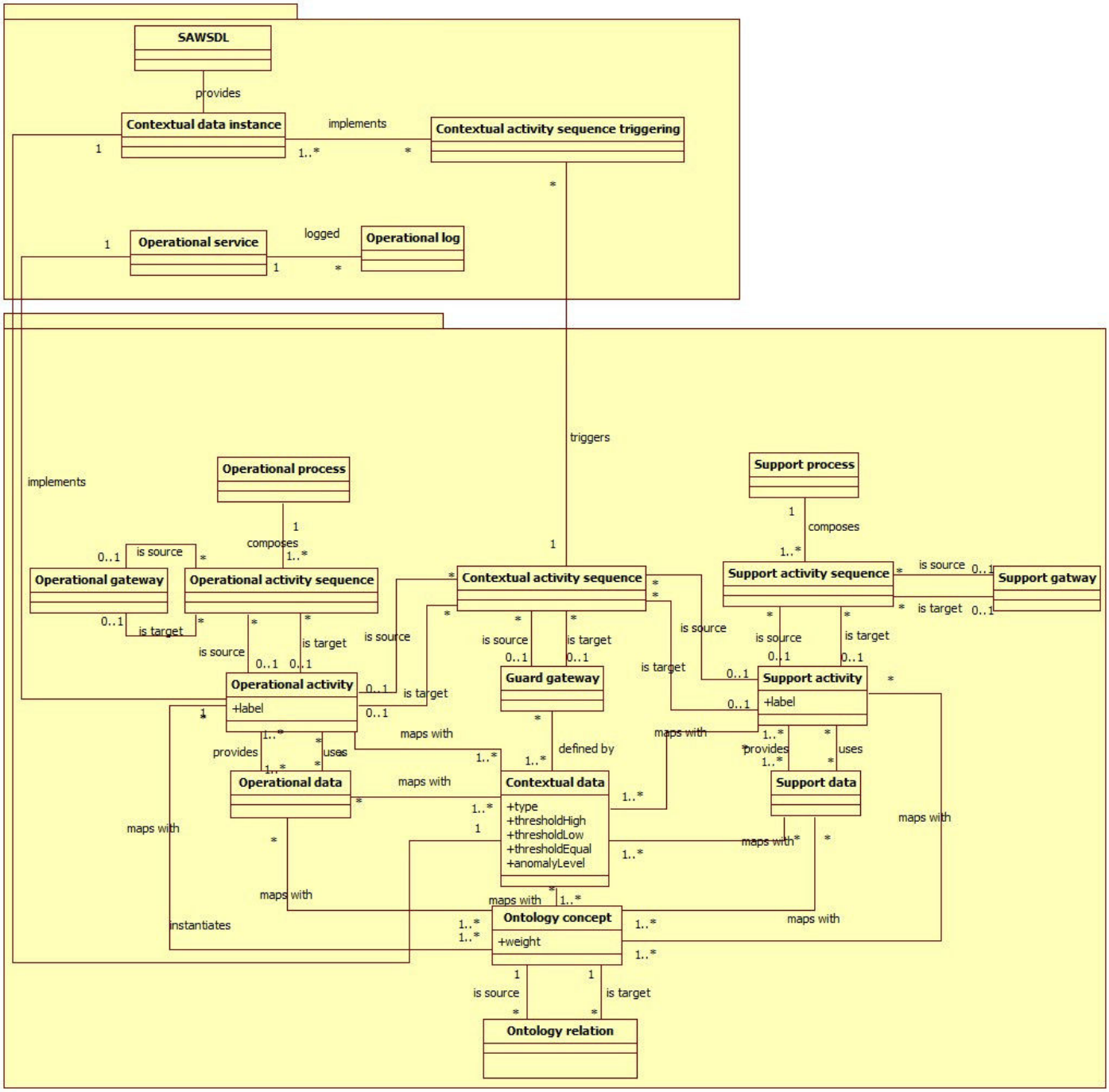


FIGURE 6.12 – Diagramme des entités participantes.

### 6.1.2 Analyse fonctionnelle dynamique de l'outil fondée sur les transformations

L'analyse fonctionnelle dynamique est caractérisée par les scénarios qui illustrent les cas d'utilisation spécifiés en analyse fonctionnelle statique. L'échange de messages entre entités participantes dans le diagramme de séquence est cohérent avec le diagramme des entités participantes (Cf. Figure 6.12). Cette cohérence signifie qu'un message envoyé de l'instance de l'entité E1 vers une instance de l'entité E2 instancie une association entre ces deux entités. Ce sera le cas dans tous les diagrammes de séquences d'analyses fonctionnelles suivants. Nous avons choisi de représenter cette instanciation à l'aide de diagramme de séquence UML.

#### 6.1.2.1 Scénario de conception des données contextuelles associées à un modèle de processus opérationnel (ScAddCD2OP)

Le scénario nominal « ScAddCD2OP » représente la conception des données contextuelles associées à un modèle de processus opérationnel. Le rôle du contributeur est celui de l'expert métier qui le déclenche afin de prendre en compte le contexte d'un processus opérationnel.

La pré-condition de ce scénario est un modèle de processus opérationnel modélisé en BPMN.

- Step1 : Lire le processus opérationnel en BPMN avec ses séquences, soit entre activités, soit entre une activité et une gateway, soit entre gateways ainsi qu'avec les données opérationnelles fournies ou utilisées par une activité opérationnelle ;
- Step2 : Chercher le concept de l'ontologie correspondant à une activité opérationnelle ou à une donnée opérationnelle ;
- Step3 : Concevoir la donnée contextuelle, mappée avec un concept de l'ontologie lui-même relié au concept opérationnel de l'ontologie par un chemin de relation de l'ontologie ;
- Step4 : Évaluer les traces du processus en récupérant les logs opérationnels à l'aide de services opérationnels ;
- Step5 : Récupérer les instances de données contextuelles afin de les exploiter avec le résultat de l'analyse des traces afin de raffiner les données contextuelles.

La post-condition de ce scénario est la relation d'une donnée/activité opérationnelle avec une ou plusieurs données contextuelles contrainte par une ontologie.

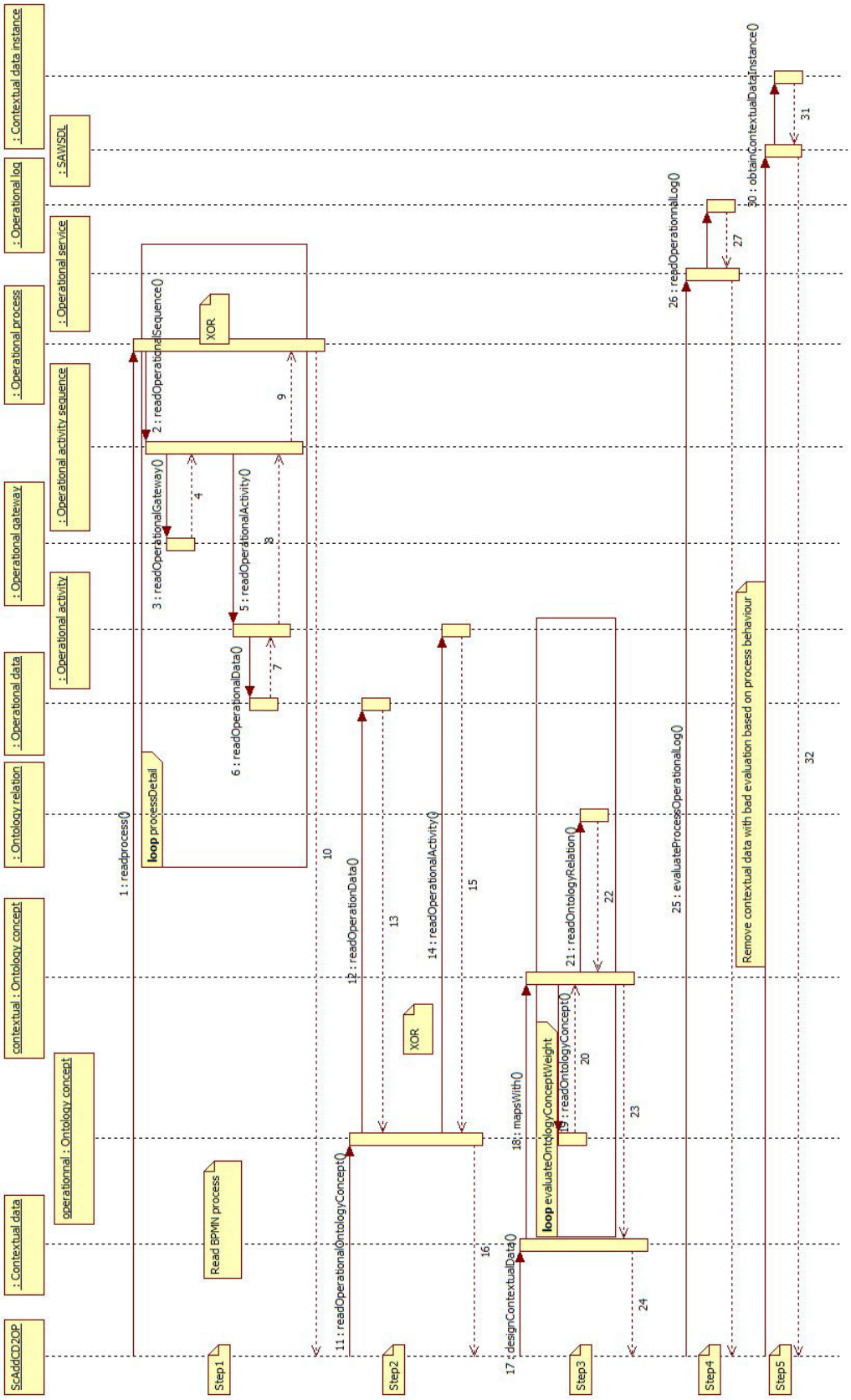




FIGURE 6.13 – Scénario de conception ScAddCD2OP.

#### 6.1.2.2 Scénario de conception des données contextuelles associées à un modèle de processus support (ScAddCD2SPs)

Le scénario nominal « ScAddCD2SPs » représente la conception des données contextuelles associées à un modèle de processus support. Le rôle du contributeur est celui de l'expert métier qui le déclenche afin de prendre en compte le contexte d'un processus support.

La pré-condition de ce scénario est un modèle de processus support modélisé en BPMN.

- Step1 : Lire le processus support en BPMN ainsi que les données supports fournies ou utilisées par chaque activité support ;
- Step2 : Chercher les concepts de l'ontologie correspondants à une activité support ou à une donnée support ;
- Step3 : Concevoir la donnée contextuelle, mappée avec un concept de l'ontologie lui-même relié au concept support de l'ontologie par un chemin de relation de l'ontologie.

La post-condition de ce scénario est la relation d'une donnée/activité support avec une ou plusieurs données supports contraintes par une ontologie.

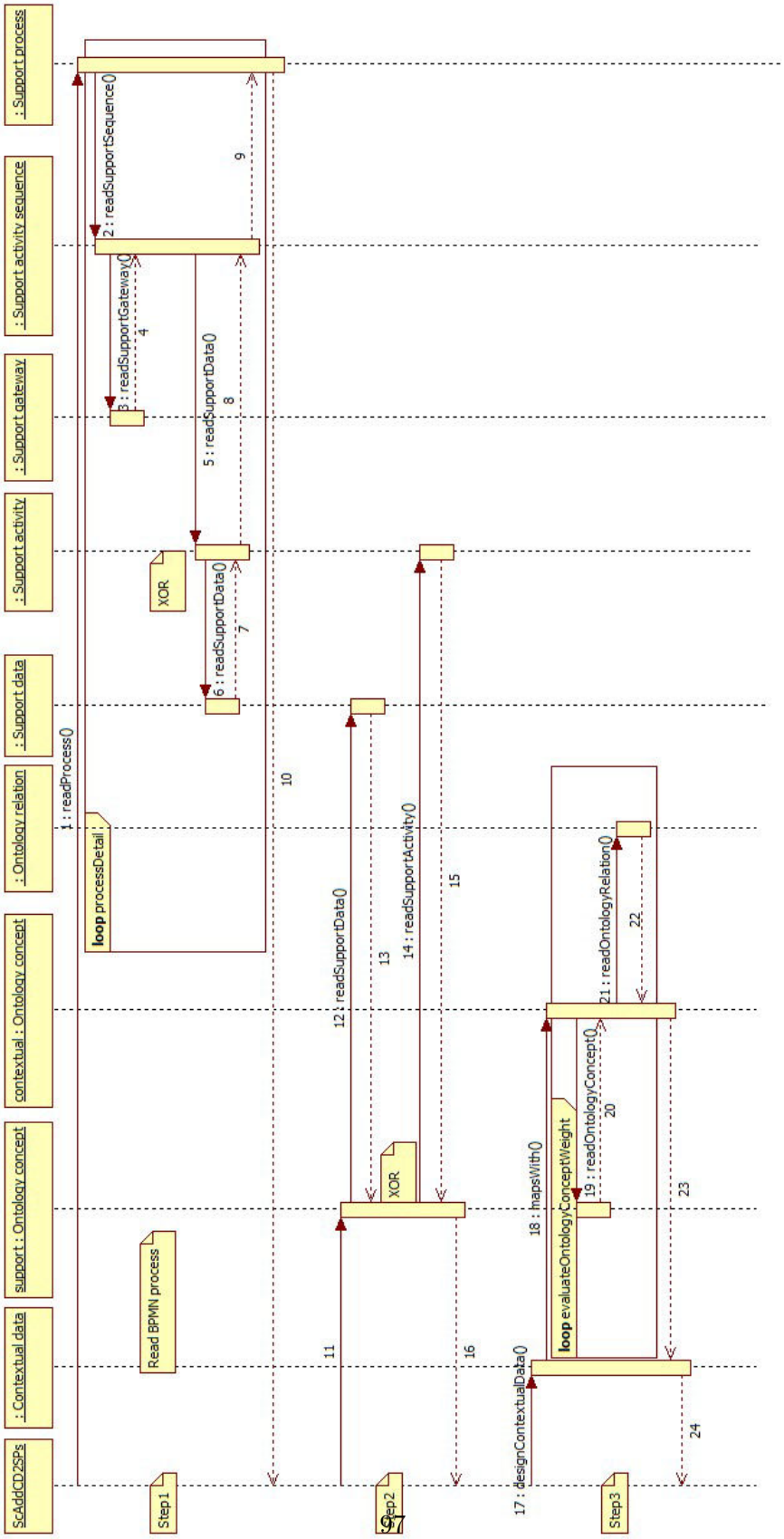


FIGURE 6.14 – Scénario de conception ScAddCD2SPs.

#### 6.1.2.3 Scénario de conception d'un ensemble de processus supports en lien avec un processus opérationnel et une donnée contextuelle (ScSelectSPs)

Ce scénario permet de sélectionner les processus supports en lien avec un processus opérationnel via une donnée contextuelle partagée. Le rôle interagissant avec ce cas d'utilisation est celui de l'expert métier.

La pré-condition du scénario est à la fois la relation d'une donnée opérationnelle avec une ou plusieurs données contextuelles contraintes par une ontologie et la relation d'une donnée support avec une ou plusieurs données contextuelles contraintes par une ontologie.

- Step1 : Collecte des données contextuelles associées à un processus opérationnel via la donnée opérationnelle ou l'activité opérationnelle (Cf. ScAddCD2OP) ;
- Step2 : Lister les processus supports dont, soit la donnée support utilisée par la première activité soit la première activité du processus support, sont associée à une donnée contextuelle collectée lors de l'étape précédente (Cf. ScAddCD2SP) ;
- Step3 : Mise en relation d'un processus opérationnelle et des processus supports listés dans l'étape précédente via les données contextuelles qui leur sont associées tel que décrit dans la section 5.1.

La post-condition du scénario est la liste des processus supports associés aux données contextuelles du processus opérationnel.

Ce scénario cible une seule entité et ne nécessite donc pas de représentation graphique.

#### 6.1.2.4 Scénario de conception d'un ensemble de processus de support raffiné à l'aide des préférences de l'expert métier (ScFilterSPs)

Il s'agit ici grâce à l'AMCD de prioriser les processus supports associés à un processus opérationnel via une donnée contextuelle. Cette priorisation est contrainte par les préférences de l'expert métier. Les niveaux d'anomalies associés aux données contextuelles sont des paramètres de la priorisation.

Pour des besoins métiers, il y a besoin d'anticiper les cas d'interruption du processus. Pour ce faire il est utile de hiérarchiser les anomalies du processus détectées dans les différents comportements traçant le déroulement d'un processus opérationnel. Cette hiérarchisation est associée à différents seuils de gravité caractérisant le type de contexte (l'organisation, la localisation, le temporel et les ressources). Nous utiliserons l'algorithme AMCD pour obtenir le descriptif des anomalies.

La pré-condition du cas d'utilisation est la liste des processus supports associés aux données contextuelles du processus opérationnel ainsi que les logs opérationnels qui tracent

les services implémentant les activités opérationnelles dans le SI. Ces logs composent les comportements sources de la priorisation.

- Step1 : Lister et évaluer les processus supports associés à un processus opérationnel via une donnée contextuelle (Cf. ScSelectSPs) ;
- Step2 : Lister et évaluer les anomalies (Cf. ScAddCD2OP) ;
- Step3 : Prioriser les processus supports par rapport aux anomalies.

La post-condition du cas d'utilisation est la liste ordonnée, par préférence de l'expert métier, des processus supports associés aux données contextuelles du processus opérationnel.

Ce scénario, du fait de se simplicité, ne nécessite pas de représentation graphique.

#### 6.1.2.5 Scénario de conception d'un processus enrichi (ScMergeOP2SPs)

L'objectif de ce scénario est de représenter l'enrichissement d'un processus opérationnel par des processus supports provenant de la liste ordonnée.

La pré-condition du cas d'utilisation est l'instanciation des concepts du méta-modèle 4.7 appliquées à un processus opérationnel et à un processus support.

- Step1 : Lire une donnée contextuelle associée à une activité opérationnelle et une activité support des processus supports listés dans ScSelectSPs ;
- Step2 : Conception d'une condition (« Guard gateway ») associée dans une séquence d'activité contextuelle entre l'activité opérationnelle et l'activité support déclenchée à partir d'une instance de la donnée contextuelle.

La post-condition du cas d'utilisation est un ensemble de séquences d'activités contextuelles entre un processus opérationnel et un processus support associé à une garde sur une ou plusieurs instances de données contextuelles. Les processus supports pris en compte sont ceux de la liste précédemment ordonnée. Ces ajouts permettent de former un processus enrichi tel que défini par le méta-modèle Figure 6.10.

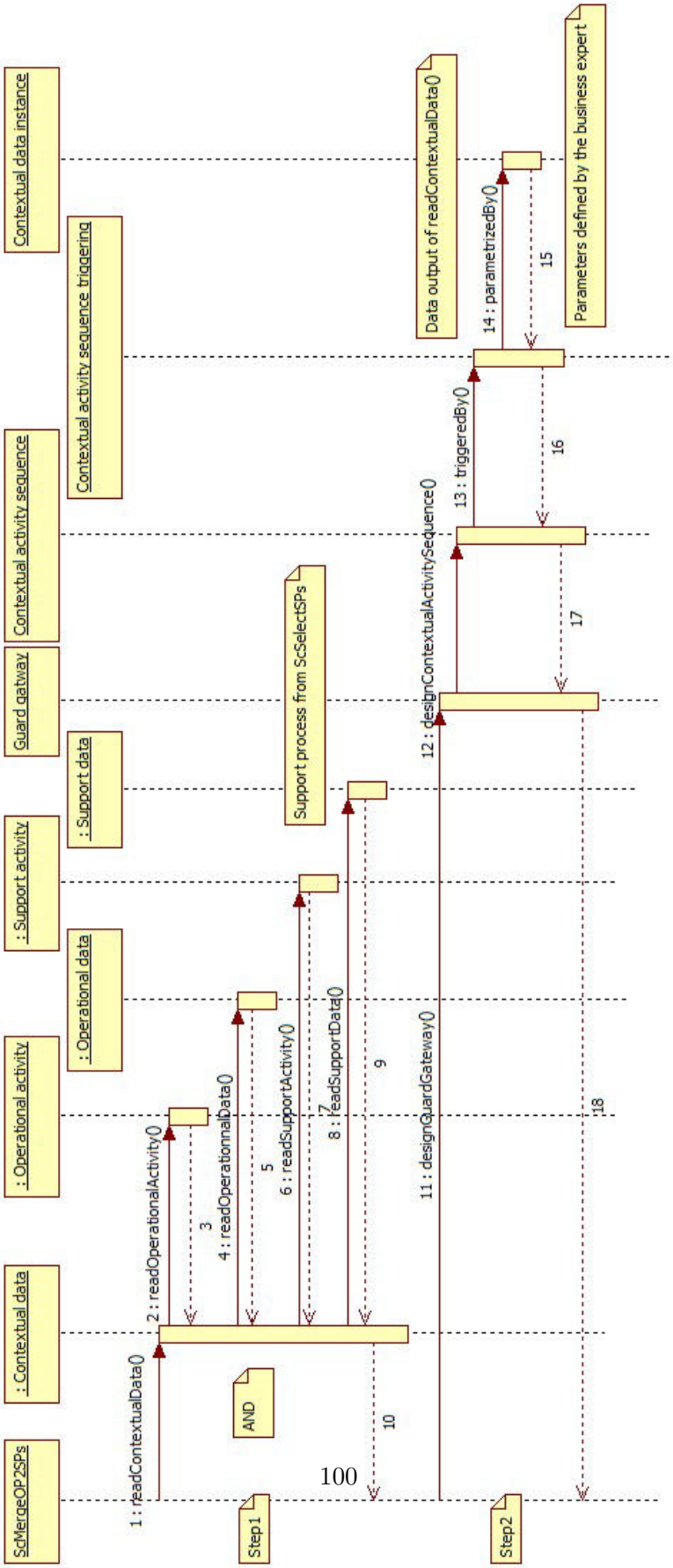


FIGURE 6.15 – Scénario de conception ScMergeOP2SPs.

L'analyse fonctionnelle dynamique valide ainsi le résultat de l'analyse fonctionnelle statique.

## 6.2 Architecture applicative de l'outil d'enrichissement

La phase d'architecture applicative cible, statiquement et dynamiquement, la réalisation de l'analyse fonctionnelle statique et dynamique contrainte par un environnement technique de l'outil. Cet environnement est décrit lors de la phase d'architecture technique où les éléments du déploiement de l'outil sont conçus.

### 6.2.1 Architecture technique de déploiement de l'outil

Les éléments du déploiement sont des nœuds d'exécution, représentant ici un environnement propre à l'IdM (modélisation et transformation de modèles), et les liens de communication qui les relies.

La Figure 6.16 représente nos quatre noeuds d'exécution et leurs liens.

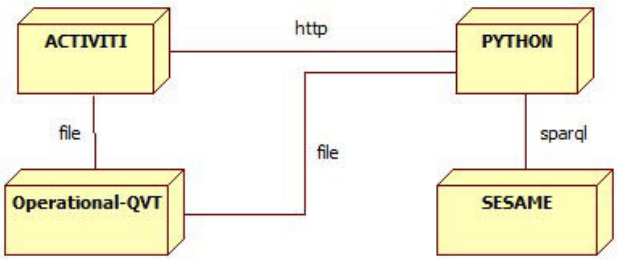


FIGURE 6.16 – Architecture technique de l'outil d'enrichissement.

En suivant les recommandations architecturales d'Orange, nous avons utilisé la suite open-source de gestion de processus « Activiti ».

Nous avons choisi d'implémenter les services d'analyse et de récupération en Python. Python est un langage haut niveau souvent utilisé dans le cadre de l'analyse de donnée. Il permet d'avoir une grande productivité de par sa simplicité et ses très nombreux modules matures.

Dans le domaine des outils de gestion d'ontologie, Sesame était la seule solution open source qui offrait une interface simple en SPARQL avec Python.

De plus, nous avons choisi d'implémenter les transformations d'enrichissement d'un processus opérationnel décrit dans la section précédente avec le langage Operational-QVT (Query, View, Transformation) préconisé par l'OMG. Au sein de la transformation « FilterSPs », implémentée en Operational-QVT, nous utilisons le langage statistique R pour

exploiter les algorithmes d'aide multi-critère à la décision. En effet, le langage R est le seul à avoir une implémentation à l'état de l'art de l'algorithme MR-Sort.

Cependant, afin de mettre en oeuvre une chaîne d'outillage classique, nous avons dû concevoir des transformations entre les méta modèles de ces outils et les méta modèles que nous utilisons. Cette transformation est nécessaire pour la modification des processus modélisés en BPMN vers un méta modèle BPMN implémenté par Activiti. En effet, le méta modèle implémenté au sein d'Activiti n'est pas strictement le même que celui proposé par l'OMG pour le BPMN, impliquant des incompatibilités lors des imports des modèles de processus.

Les quatre noeuds s'échangent en effet des informations via des protocoles. « PYTHON » et « ACTIVITI » utilisent le protocole HTTP[2] pour communiquer, « PYTHON » et « SESAME » utilisent le protocole Sparql [70], « PYTHON » et « Operational-QVT » s'échangent des fichiers (non automatiquement) ainsi qu'entre « Operational-QVT » et « ACTIVITI ».

### 6.2.2 Architecture applicative statique de l'outil

La phase d'architecture applicative statique cible ici la conception d'un modèle de composants applicatifs réalisant les cas d'utilisation spécifiés en analyse fonctionnelle et déployés sur les nœuds décrits en architecture technique.

#### 6.2.2.1 Composants applicatifs

L'architecture proposée est constituée de huit composants. Ces composants applicatifs sont conçus à partir des entités participantes spécifiées en analyse fonctionnelle statique.

La table 6.1 montre l'implément-ation des entités participantes, issues de l'analyse fonctionnelle, par les composants applicatifs.

2. IETF RFC 1945, https ://tools.ietf.org/h !ml/rfc1945

TABLE 6.1 – Implémentation des entités participatives au sein de composants applicatifs.

| Entité participante | Composant |
|---|---|
| Operational process<br>Operational activity sequence<br>Operational gateway<br>Operational activity<br>Operational data<br>Support process<br>Support activity sequence<br>Support gateway<br>Support activity<br>Support data | AC_PROCESS_DISPLAY |
| Contextual data | AC_SEMANTIC_ANNOTATION |
| Ontology concept (excepté weight)<br>Ontology relation | AC_ONTOLOGY_MANAGEMENT |
| Operational log | AC_LOG_ANALYSIS |
| Contextual activity sequence triggering<br>Contextual data instance | AC_CONTEXTUAL_DATA_<br>INSTANCE_ANALYSIS |
| SAWSDL<br>Operational service | AC_RUNNING |
| Contextual activity sequence<br>Guard gateway | AC_PROCESS_MERGE |
| Ontology concept - weight<br>Ontology relation | AC_SEMANTIC_ANALYSIS |

- AC_PROCESS_DISPLAY : Ce composant applicatif gère la lecture et l'affichage du processus. Il implémente donc l'ensemble des entités participantes relatives au méta modèle propre à un processus, les entités participantes étant : « Operational process », « Operational activity sequence », « Operational gateway », « Operational activity », « Operational data », « Support process », « Support activity sequence », Support gateway », « Support activity » et « Support data » ;

- AC_ONTOLOGY_MANAGEMENT : Ce composant applicatif gère l'ontologie et implémente les entités « Ontology concept » et « Ontology relation » ;

- AC_SEMANTIC_ANNOTATION : Ce composant applicatif gère l'annotation du processus en implémentant les entités « Contextual data » ;

- AC_SEMANTIC_ANALYSIS : Ce composant applicatif gère la recherche des concepts de l'ontologie pertinents pour l'annotation du processus. Cette recherche se fonde sur la notion de poids de l'ontologie (« Ontology concept - weight ») ainsi que sur la

notion de nombre de relation (« Ontology relation ») existante entre deux concepts de l'ontologie ;

- AC_RUNNING : Ce composant permet d'aller récupérer les données associées aux traces d'exécution d'un processus à l'aide des « Operational service » ainsi que des instances de données contextuelles à l'aide des « SAWSDL » ;
- AC_LOG_ANALYSIS : Ce composant est responsable de l'analyse des logs applicatifs (« Operational log ») ;
- AC_CONTEXTUAL_DATA_INSTANCE_ANALYSIS : Ce composant gère l'analyse des instances de données contextuelles (« Contextual data instance ») et l'instanciation des conditions de déclenchement des séquences vers les guards (« Contextual activity sequence triggering ») ;
- AC_PROCESS_MERGE : Ce composant applicatif gère la mise en relation d'un processus opérationnel avec un ou plusieurs processus supports. Pour cela, il instancie des séquences associées au contexte (« Contextual activity sequence ») et des portes conditionnées (« Guard gateway »).

Le diagramme de déploiement des composants (Figure 6.17) indique le nœud où est déployé chaque composant.

Chaque donnée applicative est définie à partir de ce tableau par un couple entité et composant applicatif. Le composant applicatif produit la donnée applicative qui réalise l'entité.

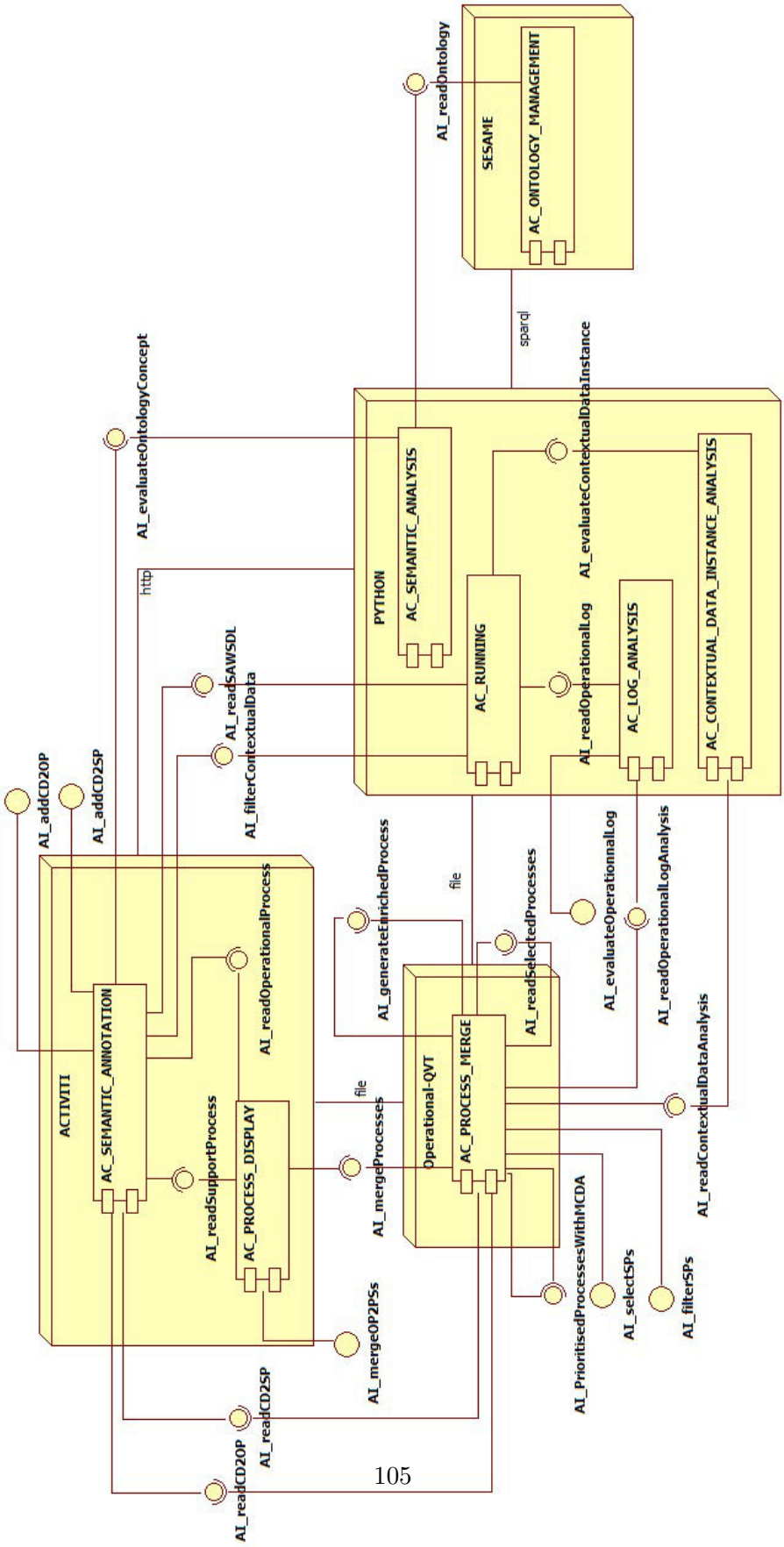




FIGURE 6.17 – Représentation des composants applicatifs et de leur interaction au sein d'un diagramme de déploiement

De plus, la Figure 6.17 illustre les liens entre les composants conçus précédemment ainsi que leur déploiement sur les liens de communication conçus en architecture technique. Les composants applicatifs conçus sont déployés sur ces noeuds d'exécution. Le premier noeud d'exécution, « ACTIVITI », prend en charge les composants « AC_PROCESS_DISPLAY » et
« AC_SEMANTIC_ANNOTATION ». Le second noeud, « PYTHON », supporte les composants « AC_SEMANTIC_ANALYSIS », « AC_RUNNING », « AC_LOG_ANALYSIS » et « AC_PROCESS_MERGE ». Le troisième noeud, « ONTOLOGY », s'occupe des composants liés à l'ontologie, « AC_ONTOLOGY_MANAGEMENT ». Enfin, le quatrième noeud supporte AC_PROCESS_MERGE. Le découplage de ces noeuds permet le partage des ressources entre différents projets. En effet, si le besoin de redimensionner les noeuds applicatifs se fait sentir, le découplage permet de ne modifier les ressources que du noeud ayant de faibles performances de calcul. Ce besoin de mise à l'échelle est principalement axé sur la nécessité de maintenir des temps de calculs acceptables par les utilisateurs.

Un lien est créé grâce à l'utilisation et à l'exposition d'interfaces applicatives (AI) fournies par un composant applicatif. L'utilisation de ces interfaces est justifiée ci-après par la réalisation de chaque scénario spécifiés en analyse fonctionnelle dynamique.

### 6.2.3 Architecture applicative dynamique de l'outil

Afin de valider notre proposition d'architecture, nous allons instancier les scénarios de conception définis lors de l'analyse fonctionnelle avec les composants applicatifs. Cette instanciation sera représentée à l'aide de diagramme UML. Chaque message échangé de type de requête d'une instance du composant applicatif AC1 vers une instance du composant applicatif AC2 est cohérent avec l'appel par le composant AC1 d'une interface du composant AC2 conçu dans le diagramme de composants applicatifs Figure 6.17. Chaque étape spécifiée dans les scénarios de l'analyse fonctionnelle dynamique est reportée dans les diagrammes de séquence applicatifs.

#### 6.2.3.1 Scénario de conception des données contextuelles associées à un modèle de processus opérationnel (ScAddCD2OP)

Le diagramme de séquence suivant représente l'architecture applicative dynamique du scénario ScAddCD2OP permettant d'enrichir sémantiquement le processus opérationnel avec des données contextuelles.

La stratégie d'architecture applicative choisie est de découpler les étapes Step1, Step2, Step3, l'étape Step4 et l'étape Step5. En effet, le premier ensemble d'étapes est associé au niveau M1 (Modèle) alors que les deux autres étapes sont associées au niveau M0 (Instance).

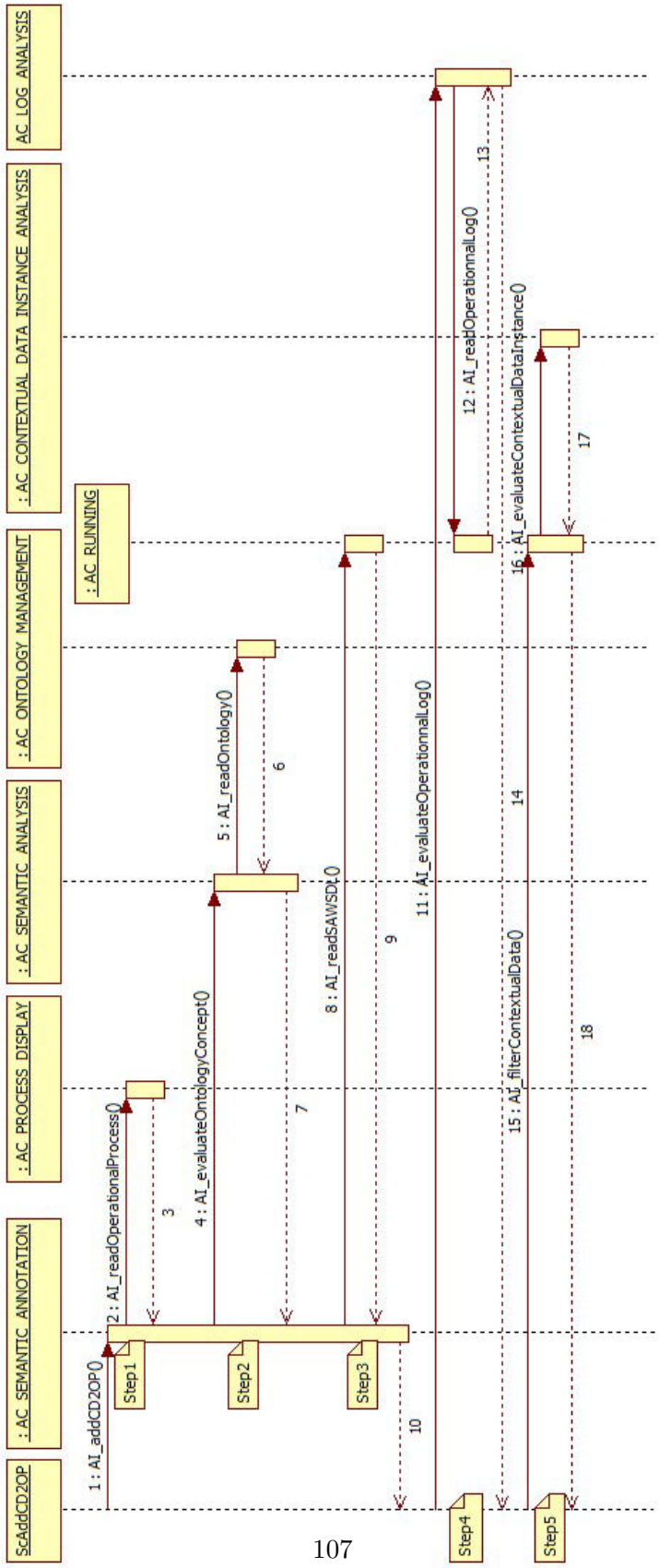




FIGURE 6.18 – Instanciation du scénario de conception ScAddCD2OP.

- Step1 : La lecture du modèle du processus opérationnel est réalisée par l'outil « Activiti ». Il permet d'interpréter un modèle BPMN afin de pouvoir le manipuler ;

- Step2 : La recherche des concepts de l'ontologie se fait par une requête SPARQL paramétrée par l'activité opérationnelle ou la donnée opérationnelle. L'implémentation des algorithmes « SA4BP » (Cf. Figure 4.4) et « ContextExplorer » (Cf. Figure 4.6) en Python dans « AC_SEMANTIC_ANALYSIS » permet de relier une donnée contextuelle avec une activité et/ou une donnée du processus opérationnelle. Ces algorithmes permettent d'explorer l'ontologie afin de découvrir des concepts proches de ceux précédemment identifiés ;

- Step3 : Nous associons ensuite les données contextuelles à des services internes ou externes à l'entreprise afin de pouvoir récupérer par la suite les instances de ces dernières à l'aide des SAWSDL ;

- Step4 : Une implémentation Python au sein du composant « AC_LOG_ANALYSIS » permet de récupérer les traces d'exécution des instances du processus opérationnel. Ces données sont ensuite utilisées afin de partitionner les logs afin de créer des clusters. Cette dernière utilise la bibliothèque logicielle PMLAB [20] afin de pouvoir lire le format XES ;

- Step5 : L'évaluation des instances des données contextes est réalisée par l'implémentation en Python de l'algorithme de « feature selection » tel que décrit dans la section 4.3.

#### 6.2.3.2 Scénario de conception des données contextuelles associées à un modèle de processus support (ScAddCD2SPs)

Le diagramme de séquence suivant représente l'architecture applicative dynamique du scénario ScAddCD2SPs permettant d'enrichir sémantiquement les processus supports avec des données contextuelles.

- Step1 : La lecture du modèle du processus support est réalisée par l'outil « Activiti ». Il permet d'interpréter un modèle BPMN afin de pouvoir le manipuler ;

- Step2 : La recherche des concepts de l'ontologie se fait par une requête SPARQL paramétrée par l'activité support ou la donnée support ;

- Step3 : L'implémentation des algorithmes « SA4BP » 4.4 et « ContextExplorer » 4.6 en Python dans « AC_SEMANTIC_ANALYSIS » permet de relier une donnée

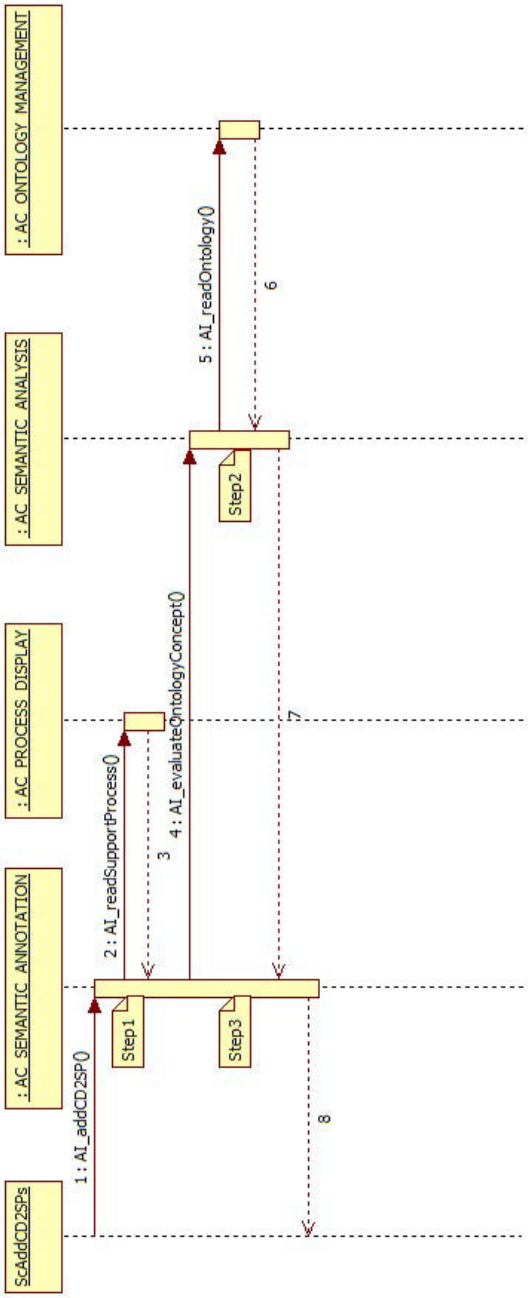


FIGURE 6.19 – Instanciation du scénario de conception ScAddCD2SPs.

contextuelle avec une activité et/ou une donnée du processus support. Ces algorithmes permettent d'explorer l'ontologie afin de découvrir des concepts proches de ceux identifiés dans l'étape Step2.

#### 6.2.3.3 Scénario de conception d'un ensemble de processus supports en lien avec un processus opérationnel et une donnée contextuelle (ScSelectSPs)

Le diagramme de séquence suivant représente l'architecture applicative dynamique du scénario ScSelectSPs permettant de collecter un ensemble de processus supports en relation avec un processus opérationnel.

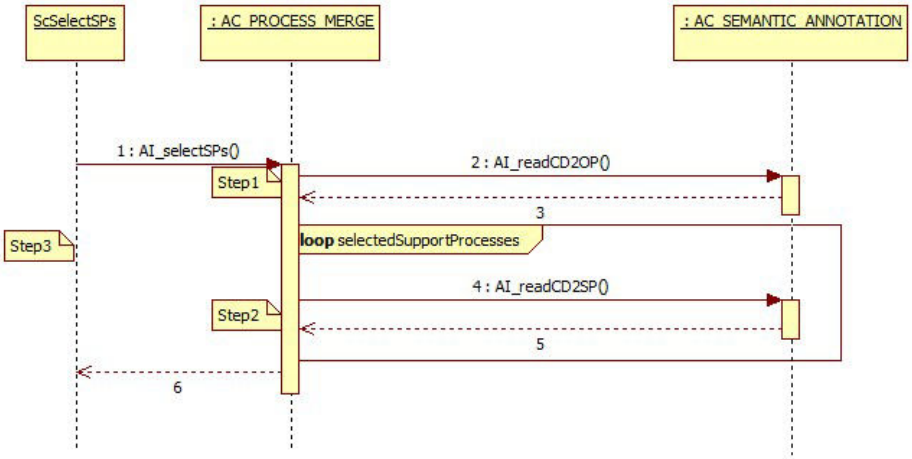


FIGURE 6.20 – Instanciation du scénario de conception ScSelectSPs.

- Step1 : La collecte des données contextuelles associées au processus opérationnel est implémentée en Python dans le composant « AC_PROCESS_MERGE » à l'aide des annotations sémantiques gérées dans « AC_SEMANTIC_ANNOTATION » ;
- Step2 : La collecte des données contextuelles associées aux processus supports est implémentée en Python dans le composant « AC_PROCESS_MERGE » à l'aide des annotations sémantiques gérées dans « AC_SEMANTIC_ANNOTATION » ;
- Step3 : La mise en relation des processus supports et du processus opérationnelles via les données contextuelles est réalisée par une transformation QVT dans « AC_PROCESS_MERGE » conformément au méta-modèle du processus opérationnel enrichi (Cf. Figure 6.10).

#### 6.2.3.4 Scénario de conception d'un ensemble de processus de support raffiné à l'aide des préférences de l'expert métier (ScFilterSPs)

Le diagramme de séquence suivant représente l'architecture applicative dynamique du scénario ScFilterSPs permettant de sélectionner un ensemble de processus supports en

relation avec un processus opérationnel.

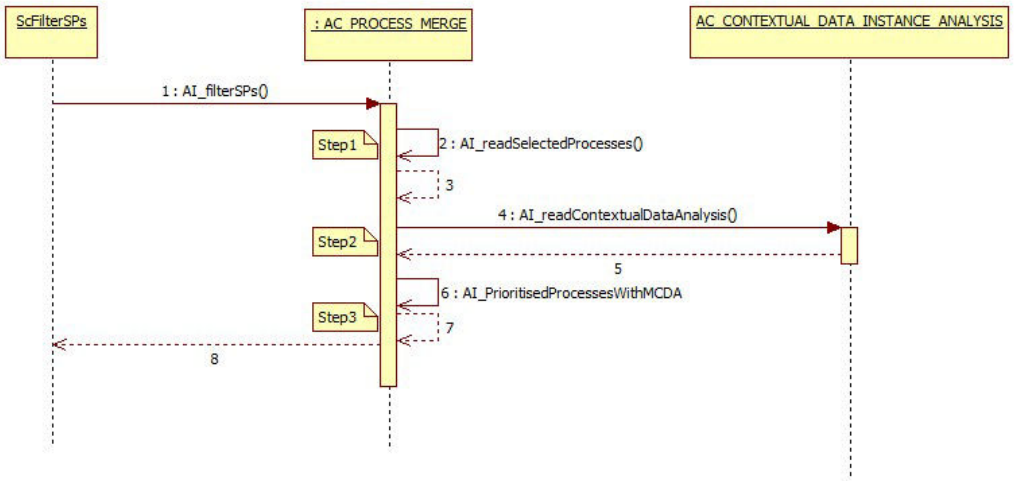


FIGURE 6.21 – Instanciation du scénario de conception ScFilterSPs.

- Step1 : Le composant « AC_PROCESS_MERGE » permet la lecture les processus supports précédemment collectés. Une implémentation du module de l'aide à la décision en R permet d'évaluer chaque instance de processus support ;
- Step2 : Une seconde implémentation du module de l'aide à la décision en R permet d'évaluer chaque instance des données contextuelles au sein du composant « AC_CONTEXTUAL_D-ATA_INSTANCE_ANALYSIS » ;
- Step3 : Une transformation QVT dans le composant « AC_PROCESS_MERGE » permet de filtrer les processus supports par rapport aux évaluations réalisées dans Step1 et Step2 à l'aide de méthode MCDA.

#### 6.2.3.5 Scénario de conception d'un processus enrichi (ScMergeOP2SPs)

Le diagramme de séquence suivant représente l'architecture applicative dynamique du scénario ScMergeOP2SPs permettant de générer un modèle de processus opérationnel enrichi.

Le composant « AC_PROCESS_DISPLAY », fourni par « ACTIVITI », permet d'afficher le résultat final, c'est-à-dire le processus opérationnel enrichi par les processus supports. C'est donc le composant « AC_PROCESS_MERGE » qui contient le service réalisant ScMergeOP2SPs.

- Step1 : Lire une donnée contextuelle associée à une activité opérationnelle et une activité support d'un des processus supports listés dans ScSelectSPs. La collecte

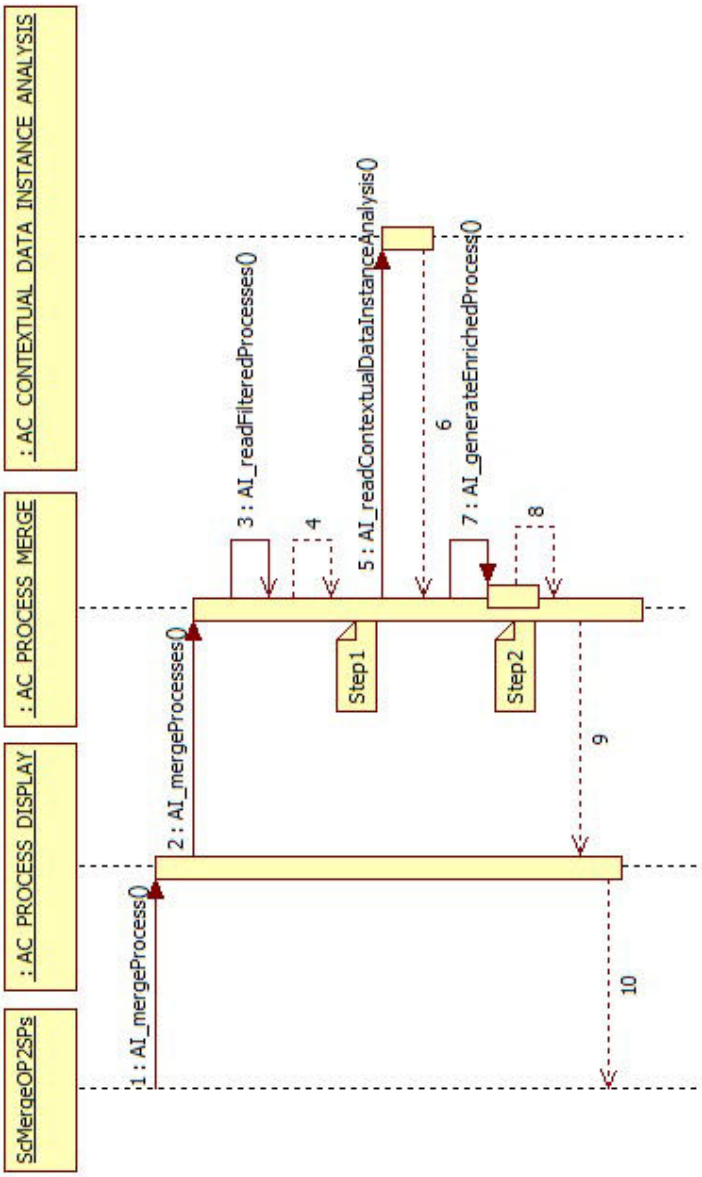


FIGURE 6.22 – Instanciation du scénario de conception MergeOP2PS.

des données, utiles pour réaliser l'enrichissement, est implémentée dans le composant « AC_PROCESS_MERGE » pour les données métiers (opérationnelles et supports) associées aux processus et dans le composant « AC_CONTEXTUAL_DATA_INSTANCE_ANALYSIS » pour les données contextuelles ;

- Step2 : Les données collectées dans Step1 sont transformées en un modèle enrichi conforme au méta-modèle du processus opérationnel enrichi (Cf. Figure 6.10). Cette

transformation QVT est réalisée dans le composant « AC_PROCESS_MERGE ».

Ces cinq diagrammes de séquence valident l'utilisation des interfaces des composants applicatifs proposés dans le diagramme de déploiement (Cf. Figure 6.17).

## 6.3 Conclusion

Nous avons développé un outil exploitant le contexte d'exécution d'un processus opérationnel comme pierre angulaire de sa transformation en un processus enrichi.

Ceci constitue la réponse outillée à nos trois questions de recherche.

Dans ce chapitre, nous avons reformulé au sein de l'analyse fonctionnelle nos travaux à l'aide de l'Ingénierie dirigée par les Modèles afin de déduire l'architecture de notre outil d'enrichissement. Cette architecture a ensuite été validée par la réalisation des cas d'utilisation. Cela consiste à réaliser les scénarios de conception en instanciant les composants applicatifs conçus statiquement. Nous avons ainsi pu vérifier que l'ensemble des composants applicatifs ainsi que leur relations sont utiles à tous les scénarios d'utilisation de l'outil.

Chapitre 7

# Conclusion et Perspectives

## 7.1 Conclusion

Ce manuscrit a présenté une approche permettant d'utiliser le contexte d'exécution d'un processus métier afin d'améliorer sa compréhension et potentiellement l'améliorer. Cette approche permet de formaliser le lien existant entre plusieurs processus métiers. Cette approche outillée s'articule autour de trois questions de recherche :

- Dans la perspective de définir le contexte d'exécution d'un processus métier, nous avons basé notre approche sur des techniques sémantiques afin d'enrichir le modèle d'un processus métier avec des annotations. Ces annotations permettent de relier les différentes activités avec un sous-ensemble de concepts d'une ontologie décrivant un domaine métier. Ce sous-ensemble représente le contexte d'exécution d'un processus. Cette approche est utilisable dans différents domaines à condition de construire une ontologie associée à ces domaines. Ce coût de construction d'une ontologie peut être réduit avec l'usage d'outil de génération d'ontologie basé sur les documents [8] (QR1) ;

- Afin d'exploiter ces données pour enrichir l'analyse des processus, nous nous sommes appuyés sur des techniques de partitionnement d'instances de processus. Cette analyse permet de découvrir des comportements similaires fondés sur la similarité des traces. Ensuite, le composant d'analyse du contexte nous permet d'extraire les données contextuelles en relation avec des groupes de comportements similaires (QR2) ;

- Nous proposons un outil accélérant la démarche de modification d'un modèle qui permet de mettre en relation plusieurs processus en se basant sur leur contexte respectif. Ce modèle prend également en compte les préférences du décideur qui est, dans notre cas, l'expert métier (QR3).

Cette modification du processus a pour but d'améliorer l'efficacité métier du processus et ainsi d'augmenter la satisfaction client. Elle permet également l'émergence de modèle de processus opérationnel dans lequel les relations avec des processus supports sont modélisées. Cette modification du processus rend possible l'automatisation et la communication entre un processus opérationnel et plusieurs processus supports. La modification

est réalisée à l'aide de techniques de transformation de modèle proposé par le domaine de l'Ingénierie dirigée par les modèles.

Tous ces travaux mettent l'humain au cœur de l'analyse et dans notre cas, l'expert métier. En effet, l'expert est en charge de la création de l'ontologie, mais aussi de la sélection des données contextuelles à prendre en compte lors de la modification du processus. Enfin, l'humain est au cœur du mécanisme de modification du modèle de processus en exprimant ses préférences sur les critères de décision et en choisissant les couples anomalie/processus support à instancier en se fondant sur la priorisation effectuée précédemment. L'utilisation des préférences de l'expert est possible grâce à l'aide multi-critère à la décision.

Cet outil a pour vocation à être utilisé par les experts métiers dès la conception des processus jusqu'à leur obsolescence.

## 7.2 Perspectives

### 7.2.1 Traitement de grande volumétrie de données

Durant la réalisation de nos travaux, nous nous sommes rendus compte de la grande différence d'utilisation qu'il pouvait y avoir entre les processus opérationnels de l'entreprise. En effet, la volumétrie associée aux processus opérationnels peut énormément varier. Si nous nous contentons de regarder le nombre d'instances, un processus tel que celui proposé dans l'illustration, sera instancié, par exemple, entre zéro et cents fois par jour. Alors qu'un processus tel que celui de la vente de téléphone mobile représente des volumétries devant être de l'ordre de cent fois celles de la pizzeria. Les méthodes que nous utilisons pour analyser ces traces, de par leur complexité algorithmique, vont mettre un temps bien plus long à fournir un résultat. Ce temps de réponse rendra l'utilisation de notre outil inefficace en entreprise. Il serait donc intéressant de travailler sur l'analyse des traces, mais aussi des données contextuelles afin de pouvoir l'utiliser à une échelle de volumétrie bien plus grande. Une approche souvent utilisée est la parallélisation des calculs comme appliqués par les auteurs de [42] dans le cadre de la découverte de processus à partir des logs applicatifs.

### 7.2.2 Automatisation des transformations

Nos travaux ont été orientés afin de valider l'hypothèse qu'il était possible de réaliser un outil exploitant les données contextuelles dans l'analyse d'un processus métier. Nous ne nous sommes donc pas attachés à minimiser les interactions avec l'expert métier. Cependant, il serait intéressant d'étudier comment nous pouvons réduire le nombre d'interactions entre les transformations et l'expert métier tout en gardant l'apport de son expertise. Une possibilité serait d'exploiter les exécutions précédentes du processus afin d'extraire

des informations utiles dans la sélection des processus supports et dans la définition des conditions des portes logiques.

### 7.2.3 Incorporer d'autres éléments tels que les objectifs

Un autre point qui nous semble intéressant et qui n'a pas été traité dans ce document est l'apport des objectifs d'un processus opérationnel dans la conception de ses relations avec les processus supports. Les objectifs sont l'instanciation de la stratégie de l'entreprise au niveau d'un processus ou d'une activité de processus. Les analyser permettraient de prendre en compte la stratégie de l'entreprise dans la définition des relations entre un processus opérationnel et des processus supports. Ainsi, en plus de pouvoir apporter la vision de la stratégie de l'entreprise au sein de l'analyse, cela permettrait d'utiliser les KPI associés aux objectifs afin de mieux raffiner les données contextuelles.

# Bibliographie


[1] Mari Abe and Michiharu Kudo. Analyzing business processes by automatically detecting kpi thresholds. In *IEEE International Conference on Services Computing (SCC)*, pages 187–194, 2016.

[2] Nancy Alexopoulou, Mara Nikolaidou, NY Chamodrakas, Y Chamodrakas, and D Martakos. Enabling on-the-fly business process composition through an event-based approach. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, pages 379–379. IEEE, 2008.

[3] Bastien Amar, Philippe Dhaussy, Hervé Leblanc, and Bernard Coulette. La description d'un modèle de traçabilité pour la mise en œuvre d'une technique de validation formelle de modèles. *Genie logiciel*, 85 :20, 2008.

[4] Tony Andrews, Francisco Curbera, Hitesh Dholakia, Yaron Goland, Johannes Klein, Frank Leymann, Kevin Liu, Dieter Roller, Doug Smith, Satish Thatte, et al. Business process execution language for web services, 2003.

[5] Nour Assy. *Automated support of the variability in configurable process models.* PhD thesis, Paris Saclay, 2015.

[6] Ann Barcomb, Nicolas Jullien, Patrick Meyer, and Alexandru-Liviu Olteanu. Integrating managerial preferences into the qualitative multi-criteria evaluation of team members. In *Multiple Criteria Decision Making and Aiding*, pages 95–143. Springer, 2019.

[7] Till Becker and Wacharawan Intoyoad. Context Aware Process Mining in Logistics. *Procedia CIRP*, 63 :557–562, 2017.

[8] Ivan Bedini. Automatic ontology generation : State of the art. *PRiSM Laboratory Technical Report. University of Versailles*, pages 1–15, 2007.

[9] Kahina Bessai, Bruno Claudepierre, Oumaima Saidani, and Selmin Nurcan. Context-aware business process evaluation and redesign. *Proceedings of BPMDS'08*, 335 :86–95, 2008.

[10] Jean Bézivin and Jean-Pierre Briot. Sur les principes de base de l'ingénierie des modèles. *L'OBJET*, 10(4) :145–157, 2004.

[11] Thierry Biard, Alexandre Le Mauff, Michel Bigand, and Jean-Pierre Bourey. Separation of decision modeling from business process modeling using new "decision model and notation"(dmn) for automating operational decision-making. In *Working Conference on Virtual Enterprises*, pages 489–496. Springer, 2015.

[12] Raymond Bisdorff, Luis C Dias, Patrick Meyer, Vincent Mousseau, and Marc Pirlot. *Evaluation and Decision Models with Multiple Criteria*. Springer, 2015.

[13] David Boddy, Douglas Macbeth, and Beverly Wagner. Implementing collaboration between organizations : an empirical study of supply chain partnering. *Journal of Management studies*, 37(7) :1003–1018, 2000.

[14] R. P. Jagadeesh Chandra Bose and Wil M.P. Van Der Aalst. Context aware trace clustering : Towards improving process mining results. In *SIAM International Conference on Data Mining*, pages 401–412, 2009.

[15] Denis Bouyssou, Thierry Marchant, Marc Pirlot, Alexis Tsoukias, and Philippe Vincke. *Evaluation and Decision Models with Multiple Criteria : Stepping stones for the analyst*. Springer Science & Business Media, 2006.

[16] Sergey Brin and Lawrence Page. The anatomy of a large scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30 :107–17, 1998.

[17] Arnaud Brossard, Mourad Abed, and Christophe Kolski. Taking context into account in conceptual models using a model driven engineering approach. *Information and Software Technology*, 53(12) :1349–1369, 2011.

[18] Andrea Burattin. Plg2 : Multiperspective process randomization with online and offline simulations. *Demo Track of the 14th International Conference on Business Process Management (BPM)*, pages 1–6, 2016.

[19] Zhou Cailan, Chen Kai, and Li Shasha. Improved PageRank Algorithm based on feedback of user clicks. In *Computer Science and Service System (CSSS), 2011 International Conference on*, pages 3949–3952. IEEE, 2011.

[20] Josep Carmona Vargas and Marc Solé. Pmlab : an scripting environment for process mining. In *Proceedings of the BPM Demo Sessions 2014 : Co-located with the 12th International Conference on Business Process Management (BPM 2014) Eindhoven, The Netherlands, September 10, 2014*, pages 16–20. CEUR-WS. org, 2014.

[21] Hanène Chettaoui. *Interopérabilité entre modèles hétérogènes en conception coopérative par des approches d'Ingénierie Dirigée par les Modèles*. PhD thesis, Institut National Polytechnique de Grenoble-INPG, 2008.

[22] Bachir Chihani, Emmanuel Bertin, and Noel Crespi. A graph-based context modeling approach. In *Smart Communications in Network Technologies (SaCoNeT)*, pages 1–6, 2013.

[23] Bachir Chihani, Emmanuel Bertin, Fabrice Jeanne, and Noel Crespi. Context-aware systems : a case study. *Digital Information and Communication Technology and Its Applications*, pages 718–732, 2012.

[24] Gobinda G. Chowdhury. Natural language processing. *Annual Review of Information Science and Technology*, pages 51–89, 2003.

[25] Robert Colomb, Kerry Raymond, Lewis Hart, Patrick Emery, Chris Welty, Guo Tong Xie, and Elisa Kendall. The object management group ontology definition metamodel. In *Ontologies for software engineering and software technology*, pages 217–247. Springer, 2006.

[26] João Coutinho-Rodrigues, Ana Simão, and Carlos Henggeler Antunes. A GIS-based multicriteria spatial decision support system for planning urban infrastructures. *Decision Support Systems*, pages 720–726, 2011.

[27] Talita da Cunha Mattos, Flávia Maria Santoro, Kate Revoredo, and Vanessa Tavares Nunes. A formal representation for context-aware business processes. *Computers in Industry*, 65(8) :1193–1214, 2014.

[28] Sampath Dechu and Pratyush Kumar. Search engine optimization techniques, April 25 2019. US Patent App. 15/790,393.

[29] Adela Del-RíO-Ortega, Manuel Resinas, Cristina Cabanillas, and Antonio Ruiz-Cortés. On the definition and design-time analysis of process performance indicators. *Information Systems*, 38(4) :470–490, 2013.

[30] Anind Dey, Gregory Abowd, and Daniel Salber. A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications. *Human-Computer Interaction*, pages 97–166, 2001.

[31] Samba Diaw, Redouane Lbath, and Bernard Coulette. État de l'art sur le développement logiciel basé sur les transformations de modèles. *Technique et Science Informatiques*, 29(4-5) :505–536, 2010.

[32] Remco Dijkman, Marlon Dumas, Boudewijn Van Dongen, Reina Krik, and Jan Mendling. Similarity of business process models : Metrics and evaluation. *Information Systems*, 36(2) :498–516, 2011.

[33] Dragan Djurić, Dragan Gašević, and Vladan Devedžić. Ontology modeling and MDA. *Journal of Object Technology*, pages 109–128, 2005.

[34] Peter F. Drucker. *Concept of the Corporation.* Transaction Publishers, 1993.

[35] Ward Edwards and Detloff Von Winterfeldt. Decision analysis and behavioral research. *Cambridge University Press*, pages 6–8, 1986.

[36] Dejene Ejigu, Marian Scuturici, and Lionel Brunie. An ontology-based approach to context modeling and reasoning in pervasive computing. *Pervasive Computing and Communications Workshops*, pages 14–19, 2007.

[37] Dejene Ejigu, Vasile-Marian Scuturici, and Lionel Brunie. An ontology-based approach to context modeling and reasoning in pervasive computing. In *PerComW'07*, pages 14–19, New York, United States, 2007.

[38] Mohamed Z Elbashir, Philip A Collier, and Michael J Davern. Measuring the effects of business intelligence systems : The relationship between business process and organizational performance. *International Journal of Accounting Information Systems*, 9(3) :135–153, 2008.

[39] Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 213–220. ACM, 2008.

[40] BERTIN Emmanuel and CRESPI Noël. *Architecture et gouvernance des services de communication.* Lavoisier, 2013.

[41] Elena V Epure, Charlotte Hug, Rebecca Deneckere, and Sjaak Brinkkemper. What shall i do next ? intention mining for flexible process enactment. In *Advanced Information Systems Engineering*, pages 473–487. Springer International Publishing, 2014.

[42] Joerg Evermann. Scalable Process Discovery using Map-Reduce. *IEEE Transactions on Services Computing*, PP(99) :1–16, 2014.

[43] Cyril Faucher, Frédéric Bertrand, and Jean-Yves Lafaye. Génération d'ontologie à partir d'un modèle métier uml annoté. *Revue des Nouvelles Technologies de l'Information*, 12 :65–84, 2008.

[44] JosÉ Figueira, Salvatore Greco, and Matthias Ehrogott. *Multiple Criteria Decision Analysis : State of the Art Surveys.* Springer New York, 2005.

[45] Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, and Eytan Ruppin. Placing search in context : The concept revisited. *ACM Transactions on information systems*, 20(1) :116–131, 2002.

[46] Frédéric Fürst, Michel Leclère, and Francky Trichet. Operationalizing domain ontologies : a method and a tool. In *Proceedings of the 16th European Conference on Artificial Intelligence*, pages 333–337. IOS Press, 2004.

[47] Hans-w Gellersen, Albrecht Schmidt, and Michael Beigl. Multi-Sensor Context-Awareness in Mobile Devices and Smart Artefacts. *Mobile Networks and Applications*, 7(5) :341–351, 2002.

[48] María Teresa Gómez-López, Antonia M Reina Quintero, Luisa Parody, José Miguel Pérez Álvarez, and Manfred Reichert. An architecture for querying business process, business process instances, and business data models. In *International Conference on Business Process Management*, pages 757–769. Springer, 2017.

[49] Florian Gottschalk, Wil M P Van Der Aalst, and Monique H. Jansen-Vullers. Configurable PROCESS MODELS - A foundational approach. In *Reference Modeling : Efficient Information Systems Design Through Reuse of Information Models*, pages 59–77. Physica-Verlag HD, 2007.

[50] Salvatore Greco, J Figueira, and M Ehrgott. *Multiple Criteria Decision Analysis - State of the Art Surveys*. Springer, 2016.

[51] V Grover and K Kettinger. Toward a theory of business process change. *Journal of Management Information Systems*, 12(1) :1–21, 1995.

[52] Varun Grover and M Lynne Markus. *Business process transformation*. ME Sharpe, 2008.

[53] Christoph Gröger, Holger Schwarz, and Bernhard Mitschang. Prescriptive analytics for recommendation-based business process optimization. In *Business Information Systems*, pages 25–37, 2014.

[54] Michael Hammer. Reengineering work : don't automate, obliterate. *Harvard business review*, 68(4) :104–112, 1990.

[55] Michael Hammer and James Champy. Business process reengineering. *London : Nicholas Brealey*, 444(10) :730–755, 1993.

[56] Paul Harmon. The scope and evolution of business process management. In *Handbook on business process management 1*, pages 37–81. Springer, 2010.

[57] Richard Edgar Hodgett. Comparison of multi-criteria decision-making methods for equipment selection. *The International Journal of Advanced Manufacturing Technology*, pages 1145–1157, 2016.

[58] Bart F. A. Hompes, Joos C. A. M. Buijs, and Wil M. P. van der Aalst. A Generic Framework for Context-Aware Process Performance Analysis. In *On the Move to Meaningful Internet Systems : OTM 2014 Conferences*, volume 8841, pages 300–317, Cham, 2016.

[59] Bart F A Hompes, Abderrahmane Maaradji, Marcello La Rosa, Marlon Dumas, Joos C A M Buijs, and Wil M P van der Aalst. Discovering causal factors explaining business process performance variation. *Advanced Information Systems Engineering*, pages 177–192, 2017.

[60] Px Huang and Rb Fisher. Individual feature selection in each One-versus-One classifier improves multi-class SVM performance. *Homepages.Inf.Ed.Ac.Uk*, 2015.

[61] IFIP-IFAC Task Force on Architectures for Enterprise Integration. GERAM : The Generalised Enterprise Reference Architecture and Methodology : Version 1.6.3 (Final). *Handbook on enterprise architecture*, pages 21–63, 2003.

[62] Ivar Jacobson, Grady Booch, and James Rumbaugh. *The Unified Software Development Process*, volume 16. Addison-Wesley Reading, 1999.

[63] Jae Yoon Jung, Joonsoo Bae, and Ling Liu. Hierarchical business process clustering. *Proceedings - 2008 IEEE International Conference on Services Computing, SCC 2008*, pages 613–616, 2008.

[64] Hari Krishna Kanagala and V V Jaya Rama Krishnaiah. A comparative study of k-means, dbscan and optics. In *Computer Communication and Informatics (ICCCI), 2016 International Conference on*, pages 1–6, 2016.

[65] R. L. Keeney, H. Raiffa, and D. W. Rajala. Decisions with Multiple Objectives : Preferences and Value Trade-Offs. *IEEE Transactions on Systems, Man, and Cybernetics*, pages 403–403, 1979.

[66] Ralph L Keeney and Howard Raiffa. Decision analysis with multiple conflicting objectives. *Wiley& Sons, New York*, 1976.

[67] William J Kettinger, James TC Teng, and Subashish Guha. Business process change : a study of methodologies, techniques, and tools. *MIS quarterly*, pages 55–80, 1997.

[68] Ralph L Kliem. Risk management for business process reengineering projects. *Information systems management*, 17(4) :71–73, 2000.

[69] Trupti M Kodinariya and Prashant R Makwana. Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies*, pages 2321–7782, 2013.

[70] Ilianna Kollia, Birte Glimm, and Ian Horrocks. SPARQL Query Answering over OWL Ontologies. In *CEUR Workshop Proceedings*, volume 796, pages 382–396, 2011.

[71] Agnes Koschmider and Emmanuel Blanchard. User assistance for business process model decomposition. In *Proceedings of the 1st IEEE International Conference on Research Challenges in Information Science*, pages 445–454, 2007.

[72] Ivan Kurtev. State of the art of qvt : A model transformation language standard. In *International Symposium on Applications of Graph Transformations with Industrial Relevance*, pages 377–393. Springer, 2007.

[73] Risto Lahdelma, Pekka Salminen, and Joonas Hokkanen. Using Multicriteria Methods in Environmental Planning and Management. *Environmental Management*, pages 595–605, 2014.

[74] Peter Langfelder, Bin Zhang, and Steve Horvath. Defining clusters from a hierarchical cluster tree : The Dynamic Tree Cut package for R. *Bioinformatics*, 24(5) :719–720, 2008.

[75] Agnes Leroy, Vincent Mousseau, and Marc Pirlot. Learning the parameters of a multiple criteria sorting method. In *International Conference on Algorithmic DecisionTheory*, pages 219–233, 2011.

[76] Zhilei Ma, Branimir Wetzstein, Darko Anicic, Stijn Heymans, and Frank Leymann. Semantic business process repository. *CEUR Workshop Proceedings*, 251 :55, 2007.

[77] Humaira Majeed, Ramzan Talib, Madeeha Aslam, and Hira Majeed. A Study on Matching of the Process Data and Operational Data for a Deep Business Analysis . Ontology Design for Possible Matching and Recommendations to Improve it. *International Journal of Computer Science and Mobile Computing*, 2014.

[78] Pierre-Aymeric Masse, Nassim Laga, Mohammed Oussama Kherbourche, and Jacques Simonin. An Approach Based on Ontology for Discovering Data Impacting the Execution of a Business Process. In *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*, pages 216–221, 2016.

[79] Deborah L McGuinness, Frank Van Harmelen, et al. Owl web ontology language overview. *W3C recommendation*, 10(10) :2004, 2004.

[80] Nicolai Meinshausen and Peter Bühlmann. Stability selection. *Journal of the Royal Statistical Society. Series B : Statistical Methodology*, 72(4) :417–473, 2010.

[81] Andreas Meyer, Luise Pufahl, Dirk Fahland, and Mathias Weske. Modeling and enacting complex data dependencies in business processes. In *Business process management*, pages 171–186. Springer, 2013.

[82] Patrick Meyer and Alexandru-Liviu Olteanu. Integrating large positive and negative performance differences into multicriteria majority-rule sorting models. *Computers & Operations Research*, 81 :216–230, 2017.

[83] George A. Miller. WordNet : A Lexical Database for English. pages 39–41, 1995.

[84] Ronald Moen and Clifford Norman. *Evolution of the PDCA cycle*. 2006.

[85] Andrew W. Moore. *K-means and Hierarchical Clustering*. 2001.

[86] Jean-Marie Mottu, Benoit Baudry, and Yves Le Traon. Test de transformation de modèles : Expression d'oracles. In *4ièmes Journées sur l'Ingénierie Dirigée par les Modèles*, 2008.

[87] V. Mousseau, R. Slowinski, and P. Zielniewicz. A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research*, pages 757–777, 2000.

[88] Vincent Mousseau, Roman Slowinski, and Piotr Zielniewicz. ELECTRE TRI 2.0 a. methodological guide and user's manual. *Document du LAMSADE*, pages 263–275, 1999.

[89] Daniel Müllner and others. fastcluster : Fast hierarchical, agglomerative clustering routines for R and Python. *Journal of Statistical Software*, 53(9) :1–18, 2013.

[90] Stephen CH Ng, Johnny M Rungtusanatham, Xiande Zhao, and TS Lee. Examining process management via the lens of exploitation and exploration : Reconceptualization and scale development. *International Journal of Production Economics*, 163 :1–15, 2015.

[91] Florian Niedermann and Holger Schwarz. Deep Business Optimization : Making Business Process Optimization Theory Work in Practice. In *Enterprise, Business-Process and Information Systems Modeling*, pages 88–102. Springer Berlin Heidelberg, 2011.

[92] Selmin Nurcan, Bruno Claudepierre, and Islem Gmati. Conceptual dependencies between two connected it domains : Business/is alignment and it governance. In *2008 Second International Conference on Research Challenges in Information Science*, pages 87–98. IEEE, 2008.

[93] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking : bringing order to the web. 1999.

[94] K. N. Papamichail and S. French. 25 years of mcda in nuclear emergency management. *IMA Journal of Management Mathematics*, pages 481–503, 2013.

[95] Stephen J Perkins and Raisa Arvinen-Muondo. *Organizational behaviour : People, process, work and human resource management.* Kogan Page Publishers, 2013.

[96] Roman Popp and Hermann Kaindl. Automated Adaptation of Business Process Models Through Model Transformations Specifying Business Rules. In *Ceur-Ws.Org*, volume 2, pages 65–72, 2014.

[97] J Ross Quinlan. *C4.5 : Programs for machine learning.* Elsevier, 2014.

[98] Sylvia Radeschütz, Bernhard Mitschang, and Frank Leymann. Matching of process data and operational data for a deep business analysis. In *Enterprise Interoperability III*, pages 171–182. Springer, 2008.

[99] Sylvia Radeschütz, Florian Niedermann, and Wolfgang Bischoff. Biaeditor : matching process and operational data for a business impact analysis. In *Proceedings of the 13th International Conference on Extending Database Technology*, pages 705–708. ACM, 2010.

[100] Sridhar Ramaswamy, Pablo Tamayo, Ryan Rifkin, Sayan Mukherjee, Chen-Hsiang Yeang, Michael Angelo, Christine Ladd, Michael Reich, Eva Latulippe, Jill P Mesirov, et al. Multiclass cancer diagnosis using tumor gene expression signatures. *Proceedings of the National Academy of Sciences*, pages 15149–15154, 2001.

[101] Pradeep Ravikumar, Martin J. Wainwright, and John D. Lafferty. High-dimensional Ising model selection using l1 -regularized logistic regression. *The Annals of Statistics*, 38(3) :1287–1319, 2010.

[102] Gil Regev, Ian F. Alexander, and Alain Wegmann. Modelling the regulative role of business processes with use and misuse cases. *Business Process Management Journal*, pages 695–708, 2005.

[103] Matthew Richardson and P. Domingos. The intelligent surfer : Probabilistic combination of link and content information in pagerank. *Advances in neural information processing systems*, 2(14) :1441–1448, 2002.

[104] François Robert. Théorèmes de perron-frobenius et stein-rosenberg booléens. *Linear Algebra and its applications*, 19(3) :237–250, 1978.

[105] John F Rockart, Michael J Earl, and Jeanne W Ross. Eight imperatives for the new it organization. *Inventing the Organizations of the 21st Century*, 297, 1996.

[106] M. Rosemann and Wil M.P. Van Der Aalst. A configurable reference modelling language. *Information Systems*, 32(1) :1–23, 2007.

[107] Michael Rosemann and Jan Recker. Context-aware process design : Exploring the extrinsic drivers for process flexibility. In *The 18th International Conference on Advanced Information Systems Engineering. Proceedings of Workshops and Doctoral Consortium*, pages 149–158, 2006.

[108] Michael Rosemann, Jan C Recker, Christian Flender, and Peter D Ansell. Understanding context-awareness in business process design. In Su Spencer and Adam Jenkins, editors, *17th Australasian Conference on Information Systems*. Australasian Association for Information Systems, 2006.

[109] Zuzana Rosenberg, Tobias Riasanow, and Helmut Krcmar. A system dynamics model for business process change projects. *Technische Universität München*, pages 1–27, 2015.

[110] Kiana Rostami, Robert Heinrich, Axel Busch, and Ralf Reussner. Architecture-based change impact analysis in information systems and business processes. In *2017 IEEE International Conference on Software Architecture (ICSA)*, pages 179–188. IEEE, 2017.

[111] Asim Roy, Patrick D. Mackin, and Somnath Mukhopadhyay. Methods for pattern selection, class-specific feature selection and classification for automated learning. *Neural Networks*, 41 :113–129, 2013.

[112] Bernard Roy. *Multicriteria Methodology for Decision Aiding*. Springer Science & Business Media, 1996.

[113] Oumaima Saidani and Selmin Nurcan. Towards context aware business process modelling. In *8th Workshop on Business Process Modeling, Development, and Support (BPMDS'07), CAiSE*, volume 7, page 1, 2007.

[114] Oumaima Saidani, Colette Rolland, and Selmin Nurcan. Towards a generic context model for BPM. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, pages 4120–4129, 2015.

[115] Suprateek Sarker, Saonee Sarker, and Anna Sidorova. Understanding business process change failure : An actor-network perspective. *Journal of management information systems*, 23(1) :51–86, 2006.

[116] Martin Shepperd and Darrel Ince. *Derivation and validation of software metrics.* Oxford University Press, Inc., 1993.

[117] Jacques Simonin, Francis Alizon, Jean Pierre Deschrevel, Yves Le Traon, Jean Marc Jezequel, and Bertrand Nicolas. EA4up : An enterprise architecture-assisted telecom service development method. In *IEEE International Enterprise Distributed Object Computing Workshop, (EDOC)*, pages 279–285, 2008.

[118] Jacques Simonin, Selmin Nurcan, and Judith Barrios. Evolution organisationnelle fondée sur la cohérence des relations entre acteurs avec les buts métiers. *Informatique des organisations et systèmes d'information et de décision*, pages 225–240, 2013.

[119] Jacques Simonin and John Puentes. Automatized integration of a contextual model into a process with data variability. *Computer Languages, Systems & Structures*, 54 :156–182, 2018.

[120] Olivier Sobrie, Vincent Mousseau, and Marc Pirlot. Learning a majority rule model from large sets of assignment examples. In *International Conference on Algorithmic DecisionTheory*, pages 336–350, 2013.

[121] Richard Soley et al. Model driven architecture. *OMG white paper*, 308(308) :5, 2000.

[122] Paul Strebel. Why do employees resist change ? *Harvard business review*, 74(3) :86, 1996.

[123] Patrick Taillandier and Serge Stinckwich. Using the promethee multi-criteria decision making method to define new exploration strategies for rescue robots. In *9th IEEE International Symposium on Safety, Security, and Rescue Robotics, SSRR 2011*, pages 321–326, 11 2011.

[124] Ramzan Talib, Bernhard Volz, and Stefan Jablonski. Agent assignment for process management : Agent performance evaluation framework. In *Proceedings - IEEE International Conference on Data Mining, ICDM*, pages 1005–1012, 2010.

[125] Frederick Winslow Taylor. *The principles of scientific management.* Harper, 1914.

[126] Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. *science*, pages 453–458, 1981.

[127] Wil MP Van Der Aalst and Schahram Dustdar. Process mining put into context. *IEEE Internet Computing*, 16(1) :82–86, 2012.

[128] Wil M.P. et al. Van Der Aalst. Process mining manifesto. *Lecture Notes in Business Information Processing*, pages 169–194, 2012.

[129] Upkar Varshney. Context-awareness in Healthcare. In *Pervasive Healthcare Computing*, pages 231–257, 2009.

[130] Mark von Rosing, Stephen White, Fred Cummins, and Henk de Man. Business Process Model and Notation—BPMN. In *The Complete Business Process Handbook*, pages 433–457, 2015.

[131] Philip Wahlster, Mireille Goetghebeur, Christine Kriza, Charlotte Niederländer, and Peter Kolominsky-Rabas. Balancing costs and benefits at different stages of medical innovation : a systematic review of Multi-criteria decision analysis (MCDA). *BMC Health Services Research*, 2015.

[132] X.H. Wang, Da Qing Zhang, Tao Gu, and H.K. Pung. Ontology based context modeling and reasoning using OWL. In *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*, pages 18–22, 2004.

[133] Matthias Wieland, Oliver Kopp, Daniela Nicklas, and Frank Leymann. Towards context-aware workflows. In *CAISE'07 Proceedings of the Workshop and Doctoral Consortium*, volume 2, pages 1–15, 2007.

[134] Piotr Wiśniewski. Decomposition of business process models into reusable subdiagrams. In *ITM Web of Conferences*, volume 15, page 01002. EDP Sciences, 2017.

[135] Piotr Wiśniewski, Krzysztof Kluza, and Antoni Ligęza. An approach to participatory business process modeling : Bpmn model generation using constraint programming and graph composition. *Applied Sciences*, 8(9) :1428, 2018.

[136] Workflow Management Coalition. The Workflow Reference Model. Technical report, 1999.

[137] Zhibiao Wu and Martha Palmer. Verb Semantics and Lexical Selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pages 133–138, 1994.

[138] Wenlei Xie, David Bindel, Alan Demers, and Johannes Gehrke. Edge-Weighted Personalized PageRank. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, pages 1325–1334. ACM Press, 2015.

[139] W. Xing and A. Ghorbani. Weighted PageRank algorithm. In *Proceedings. Second Annual Conference on Communication Networks and Services Research, 2004.*, pages 305–314. IEEE, 2004.

[140] Liangzhao Zeng, David Flaxer, Henry Chang, and Jun-Jang Jeng. Plm flow—dynamic business process composition and execution by rule inference. In *International Workshop on Technologies for E-Services*, pages 141–150, 2002.

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

**Titre :** Conception contextuelle des interactions entre un modèle de processus opérationnel et des modèles de processus supports

**Mots clés :** Gestion de processus métier, Données contextuelles, Analyse de processus

**Résumé :** De nos jours, l'utilisation des techniques de gestion de processus au sein des entreprises permet un gain significatif d'efficacité du système opérationnel. Ces techniques assistent les experts métier lors de la modélisation des processus de l'entreprise, de leur mise en œuvre, de leurs analyses et de leurs améliorations. Le contexte d'exécution d'un processus métier contient des informations permettant d'identifier et de comprendre les interactions entre lui-même et d'autres processus. Le premier processus est déclenché suite à un besoin client (processus opérationnel) et les autres, suite à un besoin du processus opérationnel (processus supports). La satisfaction d'un besoin client dépend d'une interaction efficace entre le processus opérationnel et les processus supports qui lui sont liés. Ces interactions se définissent au travers des données opérationnelles, manipulées par le processus opérationnel, et des données supports, manipulées par les processus supports.

Nos travaux se fondent sur l'Ingénierie dirigée par les Modèles. L'approche proposée dans cette thèse est fondée sur l'annotation des modèles de processus opérationnels ou supports. Cette annotation est réalisée avec l'aide d'une ontologie définissant le domaine métier décrit par ces processus. Ces annotations sont ensuite exploitées afin de constituer un ensemble de données, dites données contextuelles. L'analyse des traces de l'exécution du processus et de ces données contextuelles permet de sélectionner les meilleurs d'entre elles, au sens métier. Ainsi, un processus opérationnel pourra être associé à un ensemble de processus supports via les données contextuelles.

**Title :** Contextual design of interactions between an operational process model and a process model supports

**Keywords:** Business process management, Contextual data, Process analysis

**Abstract:** Nowadays, the use of business process management techniques within companies allows a significant improvement in the efficiency of the operational systems. These techniques assist business experts in modelling business processes, implementation, analytics, and enhancements. The execution context of a business process contains information to identify and understand the interactions between itself and other processes. The first process is triggered following a customer need (operational process) and the others, following the need of the operationnal process (support processes). The satisfaction of a customer need depends on an effective interaction between an operational process and the related support processes. These interactions are defined through operational data, manipulated by the operational process, and support data, manipulated by the support processes.

Our work is based on the framework of Model Driven Engineering. The approach proposed in this thesis is based on the annotation of operational or support process models. This annotation is performed with the help of an ontology defining the business domain described by these processes. These annotations are then exploited to constitute a set of data, called contextual data. The analysis of the traces of the execution of the operational process and of these contextual data makes it possible to select the best sub set of contextual data, in the business sense. Thus, an operational process can be associated with a set of support processes via the contextual data.