



HAL
open science

Développement d'un outil à base de connaissances pour l'aide à la décision dans le contexte de l'Industrie 4.0 : application au diagnostic des machines d'usinage à grande vitesse

Oussama Meski

► To cite this version:

Oussama Meski. Développement d'un outil à base de connaissances pour l'aide à la décision dans le contexte de l'Industrie 4.0 : application au diagnostic des machines d'usinage à grande vitesse. Intelligence artificielle [cs.AI]. Université de Nantes (UN), FRA., 2021. Français. NNT: . tel-03789795

HAL Id: tel-03789795

<https://theses.hal.science/tel-03789795>

Submitted on 27 Sep 2022

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'UNIVERSITE DE NANTES

ECOLE DOCTORALE N° 602

Sciences pour l'Ingénieur

Spécialité : Intelligence artificielle et informatique décisionnelle

Par

Oussama MESKI

« Développement d'un outil à base de connaissances pour l'aide à la décision dans le contexte de l'Industrie 4.0 : Application au diagnostic des machines d'usinage à grande vitesse »

Thèse présentée et soutenue à Nantes, le 08/01/2021

Unité de recherche : LS2N, UMR CNRS 6004

Rapporteurs avant soutenance :

Mme Virginie Goepf
M. Gregory Zacharewicz

Maitre de conférences des universités-HDR, INSA de Strasbourg
Professeur des universités, IMT - École des Mines d'Alès

Composition du Jury :

Président du jury : M. Nicolas Perry Professeur des universités, Arts et Métiers - Bordeaux
Examineur : M. Néjib Moalla Professeur des universités, Université Lumière Lyon

Dir. de thèse : M. Benoit Furet Professeur des universités, IUT de Nantes
Co-enc. de thèse : M. Farouk Belkadi Maître de conférences, Ecole Centrale de Nantes
M. Florent Laroche Maître de conférences-HDR, Ecole Centrale de Nantes

REMERCIEMENTS

Je tiens d'abord à adresser mes profonds remerciements aux membres de jury, Virginie Goepp, Gregory Zacharewicz, Néjib Moalla et Nicolas Perry, pour m'avoir fait l'honneur d'évaluer mes travaux de recherche.

Je tiens à remercier mon directeur de thèse, Benoit Furet, et mes encadrants, Farouk Belkadi et Florent Laroche, pour leurs précieuses directives, leurs idées scientifiques, et leur soutien continu. Qu'ils soient ici assurés de mon très grand respect et du plaisir que j'ai eu à travailler avec eux. Je les remercie également de m'avoir fait confiance en me laissant une grande autonomie.

Je tiens également à adresser mes remerciements à l'ensemble des partenaires du projet SmartEmma. Je remercie mes interlocuteurs issus des partenaires industriels du projet, Airbus, Mécachrome ou Europe Technologies, pour la qualité de nos échanges scientifiques. J'exprime également ma sincère reconnaissance à l'équipe de recherche du projet et notamment à Catherine Da Cunha, Mathieu Ritou, Asma, Zakaria et Zhiqiang.

J'exprime aussi ma sincère reconnaissance à tous mes collègues qui m'ont permis d'évoluer dans un lieu de travail à l'ambiance conviviale et particulièrement à Paul pour nos collaborations, son aide et ses conseils. À Loïc meilleur collègue de bureau, merci pour toutes nos discussions enrichissantes. Et je n'oublie pas l'équipe Ruwang, Qussay, Khaoula, Mourad, et Yassir, merci pour tous ces précieux moments passés ensemble.

Je remercie également mes parents, mon frère Aymen, et mes sœurs Omayma et Aya, ils sont ma source inépuisable de tendresse, de patience, de sacrifice, de sollicitude, et d'encouragements. Qu'ils trouvent dans ce travail le fruit de leur soutien et leurs sacrifices.

Je n'oublie pas enfin d'adresser mes sincères remerciements à tous mes amis qui étaient d'un grand soutien moral et d'une grande aide durant cette expérience difficile mais si enrichissante.

TABLE DES MATIERES

Remerciments.....	3
Table des matières.....	4
Table des Figures.....	8
Liste des abréviations.....	11
CHAPITRE 1 : Contexte et problématique.....	17
I.1. Introduction du chapitre.....	18
I.2. Contexte général de la recherche.....	18
I.2.1. L’industrie 4.0.....	18
I.2.2. Le recours au Big Data pour l’Industrie 4.0.....	20
I.2.3. La prise de decision dans le contexte d’Industrie 4.0.....	21
I.2.4. Les systèmes d’aide à la décision :.....	22
I.3. Le contexte de la thèse : l’usinage mécanique à grande vitesse.....	24
I.3.1. Les enjeux de l’usinage mécanique.....	24
I.3.2. La surveillance de l’usinage.....	26
I.3.3. Les problèmes de rupture de la chaine numérique :.....	27
I.3.4. Le projet SmartEmma.....	31
I.4. Positionnement et problématiques de recherche.....	34
I.5. Démarche scientifique.....	35
I.5.1. Analyse.....	36
I.5.2. Recueil.....	37
I.5.3. Modélisation.....	38
I.5.4. Implémentation.....	38
I.6. Conclusion du chapitre.....	38
CHAPITRE 2 : État de l’art – la gestion de la connaissance pour l’aide a la decision.....	39
II.1. Introduction du chapitre.....	40
II.2. De la donnée aux connaissances.....	40

II.3. Principaux types de connaissances	42
II.4. La gestion des connaissances.....	44
II.4.1. Le cycle de la gestion des connaissances.....	45
II.4.2. Le recueil de données et de connaissances	47
II.5. Modélisation des données et connaissances	49
II.5.1. Notion de Modèle et de modélisation	49
II.5.2. Modélisation des connaissances	50
II.5.2.1. Définitions	50
II.5.2.2. Structuration du modèle des connaissances.....	50
II.5.3. Les approches de modélisation	51
II.5.4. Exemples de modèles de connaissances	53
II.5.4.1. Le modèle d'information et de connaissance des moyens de la fabrication (MFIKM) 53	
II.5.4.2. Les modèles des normes et des standards	54
II.5.5. Les ontologies pour la modélisation et la gestion des connaissances	59
II.5.5.1. Les ontologies de références.....	59
II.5.5.2. Exemples d'ontologie pour la modélisation des connaissances.....	59
II.6. Les systèmes d'aide à la décision	64
II.6.1. Processus de SAD.....	65
II.6.2. Les Différentes catégories de SAD :.....	66
II.6.3. Les systèmes d'aide à la décision à base de connaissances (SADBC)	68
II.6.3.1. Le raisonnement à base des connaissances (KR).....	68
II.6.3.2. Les SADBC pour les industries de fabrication	69
II.6.3.3. Les systèmes de RBR	70
II.6.3.4. Les systèmes de RàPC	74
II.6.4. L'aide à la décision pour le diagnostic	76
II.6.4.1. Terminologie de la défaillance	77
II.6.4.2. Le Diagnostic des défaillances	78
II.7. Limitation et questions de recherches.....	85
II.8. Conclusion du chapitre	87

CHAPITRE 3 : Une approche d'aide à la décision basée sur des données et des connaissances	88
III.1. Introduction du chapitre	89
III.2. Rappel et prérequis sur les axes d'aide à la décision.....	89
III.3. Architecture de SAD à base de connaissances	91
III.3.1. Approche de distribution des bases de données et de connaissances	91
III.3.2. Le processus de diagnostic : une adaptation du Raisonnement à Partir de Cas.....	93
III.3.2.1. Construction et concepts.....	94
III.3.2.2. Fonctionnement	94
III.4. Métamodèle des connaissances	97
III.4.1. Structuration d'un « cas » de diagnostic :.....	101
III.4.2. Reporting et modèle de traçabilité :.....	106
III.5. Structuration des données et connaissances industrielles.....	109
III.5.1. Modélisation de la situation AS-IS.....	109
III.5.2. Méthodologie de développement des modèles.....	112
III.5.3. Classification et modélisation des données	113
III.5.3.1. Structuration des données Brutes	114
III.5.3.2. Structuration des données Smart :	118
III.5.3.3. Structuration des Bibliothèques : recours à la littérature et aux normes	122
III.5.3.4. Capitalisation des KPI	125
III.5.3.5. Construction des règles métiers.....	127
III.6. Conclusion du chapitre :.....	129
CHAPITRE 4 : Démonstrateur et besoins d'implémentation	130
IV.1. Introduction du chapitre	131
IV.2. Description du démonstrateur SmartEmma	131
IV.2.1. Définition de système multi-agents.....	131
IV.2.2. Architecture fonctionnelle de la solution multi-agents dans SmartEmma	133
IV.2.2.1. Liste des agents	134
IV.3. Le périmètre des implémentations	136

IV.3.1. Développement de la base de traçabilité	137
IV.3.2. Développement de la base de connaissances : recours aux ontologies	139
IV.3.2.1. Première couche	140
IV.3.2.2. Deuxième couche	143
IV.3.2.3. Troisième couche	144
IV.4. Aspects techniques et interface utilisateur (IU)	145
IV.5. Conclusion du chapitre.....	151
CHAPITRE 5 : Implémentation dans le cadre du Projet SmartEmma	152
V.1. Introduction du chapitre.....	153
V.2. Implémentation de l'approche dans le domaine de l'usinage.....	153
V.2.1. Principe de l'implémentation.....	153
V.3. Premier scénario de diagnostic	158
V.4. Deuxième scénario de diagnostic	164
V.4.1. Présentation du scénario	165
V.4.2. L'algorithme de diagnostic	166
V.5. Vers une généralité de l'approche proposée : Application au domaine de la restitution numérique archéologique	171
V.5.1. Étude de cas : une forge du XVIIIe siècle	173
V.5.1.1. Contexte historique.....	173
V.5.1.2. Une roue à aubes et son axe.....	175
V.5.1.3. Mise en œuvre de notre approche.....	176
V.6. Conclusion du chapitre	177
Conclusion et Perspectives	179
Liste des publications	184
Bibliographie.....	186
Annexes	203

TABLE DES FIGURES

Figure 0. Plan du manuscrit.....	16
Figure 1. Exemple de bris d’outil (à gauche) et de broutement (à droite) [Godreau, 2017]. ..	26
Figure 2. Les chaînes numériques des applications métiers et des systèmes d'information [Danjou, 2016]	28
Figure 3. Problème de rupture de la chaîne numérique en usinage [Godreau, 2017]	30
Figure 4. Structure globale du projet SmartEmma.....	31
Figure 5. Schéma de principe de l’EmmaTools	33
Figure 6. Démarche scientifique de développement	36
Figure 7. La pyramide DIKW	41
Figure 8. La méthodologie de KM de Grundstein.....	45
Figure 9. Les activités du cycle KM.....	45
Figure 10. Etapes du processus d’ECD [Boulila, 2012].....	48
Figure 11. Le modèle d'information et de connaissance des installations de fabrication (MFIKM) [Guerra-Zubiaga et young, 2006].....	54
Figure 12. Différentes catégories de normes catégorie 25	54
Figure 13. STEP-NC: Modèle de données pour MOCN intelligente [Hardwick et Loffredo,2007]	56
Figure 14. Diagramme UML du générateur du code STEP-NC	56
Figure 15. Modèle de produit dans le PLM [Cutting-Decelle et al., 2007].....	58
Figure 16. L’architecture du VFDM [Sacco et al., 2011]	60
Figure 17. Aperçu des principales classes et propriétés objets de l'ontologie [Lemaignan et al., 2006].....	61
Figure 18. Les concepts de base de l'ontologie du contexte [Giustozzi et al., 2018].....	62
Figure 19. Le processus de prise de décision [Shim et al., 2002]	66
Figure 20. L’évolution des SAD en fonction des domaines d’utilisation [Arnott et Pervan, 2008].....	67
Figure 21. Architecture d’un système à base de règles [Paquette, 2014].....	71
Figure 22. Structure du système d'aide à la décision basé sur les connaissances.....	72
Figure 23. Architecture du système de RàPC à base d’ontologie [Dendani et al., 2012]	75
Figure 24. Processus de RàPC à base d’ontologie [Adla, 2018].....	76

Figure 25. Les méthodes de diagnostic des défauts [Isermann, 2006]	79
Figure 26. Application de la DDD dans l'ingénierie [Katipamula et Brambley, 2005]	81
Figure 27. Schéma global de DDD basé sur la connaissance [Isermann, 1994 ; Isermann, 1997].....	82
Figure 28. Structure d'un SADBDC pour la détection et la prévention des défaillances	84
Figure 29. L'enchaînement des axes d'aide à la décision	90
Figure 30. Stratégie de structuration des bases de données et connaissances (1 ^{er} niveau : traçabilité).....	93
Figure 31. Principe général du système d'aide à la décision : processus de diagnostic.....	95
Figure 32. Séquencement d'un processus de diagnostic	96
Figure 33. Métamodèle de connaissances	100
Figure 34. Structure de la base de connaissances et interconnexion des couches.....	101
Figure 35. Métamodèle de structuration de cas.....	103
Figure 36. Dépendances entre les éléments de diagnostic	105
Figure 37. Métamodèle de traçabilité.....	108
Figure 38. Processus de réalisation (vue micro).....	111
Figure 39. Extrait des éléments de données brutes	116
Figure 40. Paramètres génériques	117
Figure 41. Diagramme de classe des données brutes « Emmatools ».....	117
Figure 42. Les données SMART « Modèle JPO »	119
Figure 43. Structuration des données JPO.....	119
Figure 44. Modèle de calcul des données Smart	120
Figure 45. Combinaison des modèles de données.....	121
Figure 46. Données Outils et conditions de coupe (à partir des normes).....	122
Figure 47. Modèle de structuration des bibliothèques.....	124
Figure 48. Exemple de connaissances (KPI + Seuil)	126
Figure 49. L'arborescence des mots-clés recherchés et leur relation [Monostori et al., 2016]	132
Figure 50. Architecture proposée pour le système multi-agents [Livrable Projet]	134
Figure 51. Liste des premiers agents implémentés dans le Framework SmartEmma, MAS vaut pour <i>Multi-Agent System</i> . [Livrable projet].....	135
Figure 52. Répartition des bases et communication avec le SAD.....	137
Figure 53. Structuration de la base de traçabilité implémentée dans phpMyAdmin.....	138
Figure 54. Exemple d'implémentation au niveau du logiciel « Protégé »	140

Figure 55. Extrait de la structuration de la base de connaissances.....	140
Figure 56. Un extrait du graphe ontologique de la base de connaissances	141
Figure 57. Aperçu hiérarchique de l'ontologie	142
Figure 58. Interface d'assertion des propriétés	143
Figure 59. Intégration du STEP-NC dans l'ontologie.....	144
Figure 60. Structuration de la troisième+ exemple d'instanciation.....	144
Figure 61. Technologies de développement.....	145
Figure 62. Séquencement de l'exécution d'un processus de diagnostic	149
Figure 63. Interface utilisateur pour la configuration de l'opération de diagnostic	150
Figure 64. Diagramme d'activité de l'algorithme de détection.....	159
Figure 65. Exemple de requête SPARQL.....	160
Figure 66. Extrait d'un fichier de données Smart	160
Figure 67. Instanciation de l'ontologie	161
Figure 68. Paramètres de définition d'un outil.....	162
Figure 69. L'instanciation automatique de l'ontologie.....	163
Figure 70. L'interface finale représentant le résultat du diagnostic.....	164
Figure 71. Processus de diagnostic d'une usure d'outil	166
Figure 72. Algorithme d'apprentissage des valeurs de Pmoy_min/Pmoy_max.....	167
Figure 73. Extrait du fichier « Summary ».....	168
Figure 74. Code de calculs des valeurs de référence.....	169
Figure 75. Forme et composants d'une boîte à moustache	170
Figure 76. Représentation des distributions des Pmoy en fonction des ID avec les boîtes à moustache.....	170
Figure 77. Structuration de la base de connaissances.....	173
Figure 78. Plan de la forge d'Arthez-d'Asson.....	174
Figure 79. Extrait de la structure de l'ontologie architecturale.....	175
Figure 80. Industrialisation du produit	206
Figure 81. Planification et préparation de la production	207
Figure 82. Homologation du code ISO	207
Figure 83. Fabrication du produit.....	208
Figure 84. Usinage	208
Figure 85. Contrôle du produit	209
Figure 86. As-Is avec l'EmmaTools.....	209
Figure 87. Le concept d'activité dans SADT	212

LISTE DES ABREVIATIONS

ACL: Agent Communication Language
BMM: Business Motivation Model
BPMN: Business Process Model and Notation
CIGEF: Club Informatique des Grandes Entreprises Françaises
CPS: Cyber Physical Systems
DL: Description Logic
DMN: Decision Model and Notation
ECBD: Extraction de Connaissances à partir des Bases de Données
ECD: Extraction de Connaissances des Données
ECD: Extraction des Connaissances à partir des Données
ERP: Enterprise Resource Planning
FAST: Function Analysis System Technique
FIPA: Foundation for Intelligent Physical Agents
IA: Intelligence Artificielle
TRE: Taux de Rendement Economique
ICAM: Integrated Computer Aided Manufacturing
IDEF: Integration DEFinition
IOT: Internet Of Things
ISO: International Organization for Standardization
KADS: Knowledge Analysis and Design System/Support
KDD: Knowledge Discovery in Databases
KM: Knowledge Management
KOD: Knowledge Oriented Design
KPI: Key Performance Indicator
KR: Knowledge Reasoning
MASK: Method for Analysis and Structuring Knowledge
MDA: Model Driven Architecture
MKSM: Methodology for Knowledge System Management
MOT: Modèle Organisationnel de Traitement
OCL: Object Constraint Language
OWL: Web Ontology Language

QQOQCP: Quoi, Qui, Où, Quand, Comment, Pourquoi

QVT: Query/View/Transformation

RDF: Resource Description Framework

SADT: Structured Analysis and Design Technique

SBC: Systèmes à base de connaissances

SQL: Structured Query Language

SQWRL: Semantic Query-Enhanced Web Rule Language

SWRL: Semantic Web Rule Language

SysML: Systems Modeling Language

TRS: Taux de Rendement Synthétique

UEML: Unified Enterprise Modelling Language

UML: Unified Modeling Language

VSM: Value stream mapping

XML: eXtensible Markup Language

KQML: Knowledge Query Meta Language model

Introduction Générale

Depuis son apparition, la quatrième révolution industrielle ne cesse de se propager dans la plupart des domaines d'activité. Sa mise en place présente un sujet de recherche majeur dans de nombreuses disciplines. Aujourd'hui, le point focal de la recherche autour de cette révolution est l'intégration des nouvelles techniques et technologies dans l'industrie. Il s'agit des systèmes cyber physiques, de l'intelligence artificielle, de l'internet des objets, de la réalité virtuelle, des cobots, du *cloud*, du *big data*, du *digital twin*, de la simulation ou de la fabrication additive : tout un ensemble d'outils rendus possible par l'évolution des technologies numériques [Saldivar et al., 2015 ; Julien et Martin, 2020].

Parmi les environnements les plus favorables à la migration vers l'industrie 4.0, figurent les systèmes de production. Les industries manufacturières ont pour objectif de fournir un produit optimal, fiable et conforme aux exigences des clients avec comme défi l'augmentation de la compétitivité et l'amélioration de la productivité tout en garantissant le triptyque « coût-délais-qualité ». C'est pour répondre à cet objectif que les industriels recherchent avec acharnement les meilleures techniques et outils d'amélioration continue de la production. Malgré les capacités des outils numériques à résoudre certaines problématiques industrielles, leur potentiel reste sous exploité voire inexploitable des fois, tant les données produites par les différentes machines et départements dépassent les capacités d'analyse de l'humain. Si l'expertise de l'ensemble des acteurs industriels est cruciale pour assurer un usage optimal des moyens disponibles, c'est la collaboration avec l'informatique qui permet de tirer parti de tout leur potentiel, ce que l'on nomme intelligence artificielle. Cette dernière consiste en la numérisation d'une partie de l'expertise humaine pour développer et automatiser un processus de proposition de solutions, afin de répondre à des problématiques complexes nécessitant des calculs et des efforts de raisonnement importants, hors de portée des capacités humaines.

Parmi les méthodes utilisées dans l'intelligence artificielle, les systèmes d'aide à la décision (SAD) servent essentiellement à analyser le contexte d'une situation problématique et à proposer des indicateurs de performance pour mieux comprendre les difficultés et, éventuellement, proposer des alternatives de solutions permettant de surmonter les obstacles pouvant affecter la productivité [Daihani, 2001]. Ces techniques sont très répandues depuis des années et leur utilité et leur complexité croissent avec l'apparition des nouvelles technologies

de l'information et de la communication (TIC). De plus, l'utilisation des nouvelles technologies embarquées, les capteurs et IoT, les énormes quantités de données générées chaque jour par les machines et autres ressources peuvent être profitable pour les nouvelles générations des SAD. Pour cela le point bloquant reste la gestion et l'interprétation de cette grande quantité de données. Coupler les bases de données avec des bases de connaissances industrielles incluant l'expertise humaine, l'expérience de résolution des problèmes passées et les règles métiers peut générer des résultats efficaces. Cependant, la diversité des solutions technologiques utilisées dans la chaîne numérique d'aide à la décision, depuis la collection jusqu'à la génération d'indicateur et de solutions d'aide à la décision, implique une hétérogénéité et une complexité des flux de données et de connaissances rendant leur réutilisation fortement contrainte. Par conséquent, il est primordial de suivre un processus de structuration et de gestion intégrée des données et connaissances industrielles, pour faciliter leur exploitation.

Ces travaux de recherche traitent l'exemple d'une industrie de production spécifique : l'usinage des pièces mécaniques aéronautiques. Ces pièces étant à forte valeur ajoutée et très coûteuses, le défi de productivité consiste à fabriquer des pièces conformes dès le premier coup. Il s'agit donc, implicitement, d'éviter l'occurrence de tous les phénomènes malveillants qui peuvent apparaître au moment de l'usinage. La solution adoptée est une surveillance de l'usinage et développer un SAD permettant de faire le *reporting* des indicateurs de performance et d'aider au diagnostic des causes de défaillances potentielles afin de les éviter dans les prochains ordres d'usinage. Cette thèse s'insère dans un projet de recherche ANR qui requiert une grande attention des industriels, qui sont de plus en plus conscients de l'importance de la données et connaissance pour développer leurs activités. Ce sujet s'inscrit en même temps en tête de la liste des enjeux de l'industrie 4.0. La proposition centrale est donc un système d'aide à la décision à base de connaissances avec une application particulière au diagnostic des machines d'usinage. Ce système repose sur une approche de modélisation et de gestion des connaissances industrielles permettant une meilleure exploitation des données collectées par les capteurs et leurs interprétations par des règles métier.

L'organisation du manuscrit

Ce manuscrit est structuré selon la méthodologie de recherche adoptée dans nos travaux, composée de trois parties principales : l'analyse, la modélisation et l'implémentation. Ainsi, ce rapport comporte deux premiers chapitres introductifs d'analyse, un troisième chapitre abordant le cadre conceptuel de l'approche proposée, un quatrième chapitre expliquant les aspects

d'implémentation et un dernier chapitre détaillant les scénarios exécutés pour valider l'approche proposée sur des cas industriels.

Le premier chapitre précise le contexte industriel de ces travaux de recherche, à savoir la prise de décision à base de connaissances au sein d'une approche de l'industrie 4.0. Dans un second temps, ce chapitre détaille le contexte applicatif dans lequel sont étudiées ces problématiques industrielles : le projet ANR SmartEmma dédié à l'industrie d'usinage mécanique aéronautique. Par la suite, La problématique et question de recherche de cette thèse sont expliqués.

Le deuxième chapitre présente une analyse de l'état de l'art et une étude bibliographique de l'ensemble des notions, concepts, approches et principes, relatifs à la prise de décision dans l'industrie. Cette étude débute par une définition des notions essentielles, de la donnée jusqu'aux compétences, en mettant l'accent sur les connaissances et leurs typologies. Les définitions des différents types de modélisation, les ontologies, les normes et standards seront également abordés. La deuxième partie de ce chapitre détaillera le fonctionnement des systèmes d'aide à la décision et leurs différents types en mettant l'accent sur les systèmes à base de connaissances. Le second chapitre permettra de positionner nos travaux par rapport aux avancées réalisées dans ce domaine et de poser les verrous à résoudre pour proposer une solution aux problématiques de recherche.

Le troisième chapitre débute par détailler les différentes étapes de construction de l'approche que nous proposons. Le Framework global du système d'aide à la décision ainsi que la méthode de gestion et d'intégration des données et des connaissances y sont présentées. Enfin, les méta-modèles développés pour la structuration des données et connaissances sont illustrés.

Dans le quatrième chapitre les méta-modèles proposés sont ainsi étendus sous forme d'ontologie pour fournir une base de connaissances qui inclue également les règles métier. Il décrit aussi la structuration globale du démonstrateur sous forme d'un système multi agents, ainsi que les différentes solutions technologiques nécessaires pour la mise en place du système d'aide à la décision.

Dans le cinquième et dernier chapitre, trois scénarios d'application industriels sont présentés pour évaluer la faisabilité de l'approche dans le cas du diagnostic. Le troisième scénario est dédié à démontrer la généricité de l'approche à travers une application dans le domaine de sauvegarde du patrimoine historique. La figure suivante synthétise la structuration de ce manuscrit de thèse.

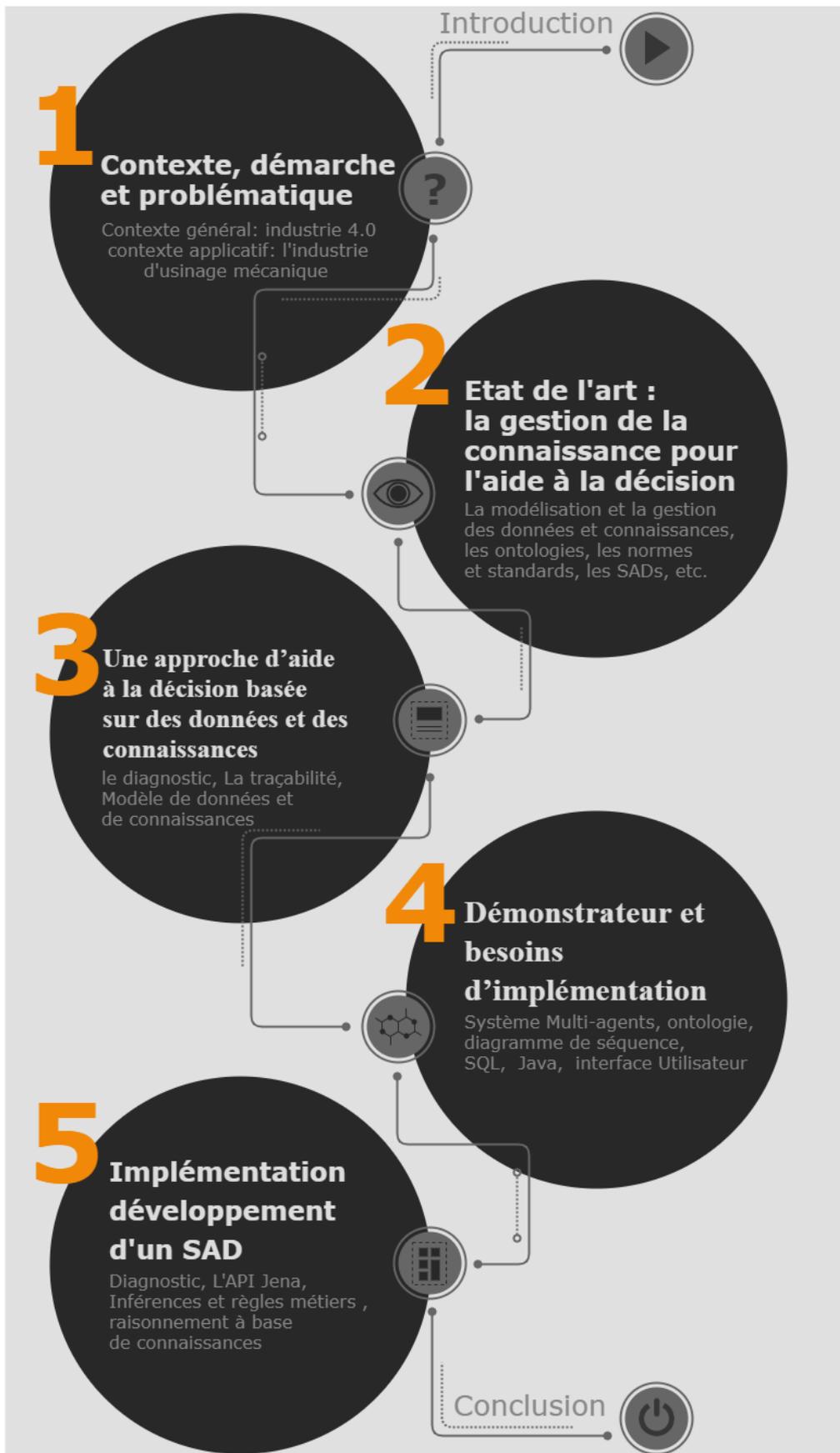


Figure 0. Plan du manuscrit

CHAPITRE 1 : CONTEXTE ET PROBLEMATIQUE

« Je vous demande...de garder le calme dans ce
CONTEXTE... nous devons avoir un esprit de
responsabilité »

Emmanuel Macron

le 16 mars 2020

Allocution au sujet de
l'épidémie de Coronavirus

I.1. INTRODUCTION DU CHAPITRE

L'objectif de ce chapitre est de définir, dans une première partie, le contexte industriel dans lequel ces travaux de recherche sont menés, ainsi que le contexte applicatif. Plus spécifiquement, ces travaux s'inscrivent dans les problématiques de la quatrième révolution industrielle et de la transition vers les industries 4.0. Le contexte d'application concerne les industries d'usinage mécanique dans secteur aéronautique. Particulièrement, le contexte applicatif, le projet ANR SmartEmma, est décrit dans la deuxième partie de ce chapitre, en précisant ses objectifs et le positionnement de cette thèse dans le Framework global.

Finalement, les problématiques et les verrous scientifiques sont identifiés, et la démarche scientifique de développement sera détaillée.

I.2. CONTEXTE GENERAL DE LA RECHERCHE

I.2.1. L'industrie 4.0

Grâce aux progrès réalisés dans les Technologies de l'Information et de la Communication (TIC), dans la digitalisation et l'Internet, le monde industriel a beaucoup évolué et se trouve face à une quatrième révolution industrielle. Celle-ci est marquée par l'apparition de la « fabrication intelligente » et des « usines intelligentes », dans laquelle les systèmes de production réels peuvent être intégrés avec une copie virtuelle dans le cyberspace par le biais des systèmes cyber physiques (ci-après CPS), de l'internet des objets (ci-après IoT) et du jumeau numérique (*digital twin*).

Le terme « Industrie 4.0 » a été inventé en 2013 par le gouvernement allemand en introduisant son plan d'action et ses stratégies concernant le développement du domaine des hautes technologies, qui encouragent la révolution numérique des industries [Blanchet, 2016]. Dans le même esprit, le terme « *Smart Industry* » décrit l'évolution des technologies industrielles à partir du microprocesseur, et les systèmes industriels embarqués jusqu'au CPS qui ont permis de relier intelligemment les différents départements de l'industrie et autres ressources industrielles, par Internet [Saldivar et al., 2015].

« L'usine virtuelle » repose sur l'emploi de technologies de l'ingénierie virtuelle (réalité virtuelle, augmentée ou mixte) dans l'objectif de réaliser des simulations avancées de processus de fabrication, de collaborer dans un environnement virtuel ou pour former les acteurs de l'usine à travers des techniques d'interactions avancées [Julien et Martin, 2020].

En France, plusieurs sociétés sont très impliquées dans le développement de solutions innovantes pour l'industrie 4.0 comme : la SNCF, Dassault Systèmes, Siemens, Schneider Electric, Bosch Rexroth, Fives, etc. [ADEME, 2017]. En parallèle, plusieurs startups naissent quotidiennement pour proposer de nouvelles solutions innovantes permettant l'optimisation des ressources et la performance globale de l'industrie en utilisant les nouvelles technologies, mais aussi en assurant de nouveaux services aux grandes entreprises industrielles.

Plusieurs organisations ont pris l'initiative de collaborer afin de joindre leurs compétences complémentaires, qui n'adressent pas forcément les mêmes types de problématiques mais qui partagent forcément un intérêt pour certains aspects de l'industrie du futur et de la modernisation des entreprises. Dans ce contexte, on peut citer l'exemple de l'Alliance Industrie du Futur qui a été créée en 2015 à l'initiative de 11 organisations professionnelles, établissements académiques et technologiques (elle comporte 35 membres aujourd'hui), ou de la *French Fab* qui, en s'inspirant de l'approche « *French Tech* », propose aux industries manufacturières de promouvoir et favoriser l'exploitation du numérique à tous les niveaux de l'entreprise [Site web French Fab].

L'implémentation des approches de l'industrie 4.0 doit assurer les principes suivants :

- L'Interopérabilité : la capacité de communication entre les différents éléments de l'usine (les niveaux décisionnels, les CPS, les systèmes d'information, l'effectif, les ressources, les produits intelligents, etc.) [Lu, 2017] ;
- La virtualisation : la capacité de fournir une copie virtuelle du système réel à travers la capitalisation des données et connaissances et la modélisation des processus industriels (*via* l'usage des *digital twin* par exemple) [Qi et Fao, 2018 ; Brettel et al., 2014] ;
- La décentralisation : la capacité de partager la prise de décision, en concevant des sous-processus autonomes dans les industries, à travers l'utilisation des CPS, l'IoT, etc. [Marques et al., 2017] ;
- La modularité et l'extensibilité : la flexibilité et la dynamique des solutions techniques utilisées, permettant de s'adapter aux besoins industriels, et aux variations de contexte [Gupta, 2019] ;
- Analyse en temps réel : la capacité d'analyser de grandes quantités de données (*Big Data*) qui permettent de contrôler et d'optimiser les processus, en facilitant tout

résultat et toute décision découlant du processus à tout moment et de façon immédiate [Silva Peres et al., 2018] ;

- La servicisation : la capacité de transférer les résultats générés par le système de production global sous forme de nouveaux services pour le client à travers des nouveaux modèles économiques [Lee et al., 2014]. Ce paradigme Produit-Service implique l'utilisation des nouvelles technologies pour mieux connecter le client avec la chaîne de valeur complète.

Afin de pouvoir mener à bien le principe d'analyse en temps réel, et en plus des systèmes cyber-physiques, du *cloud*, de la robotique et des différentes nouvelles technologies et évolutions dans le domaine de l'électronique numérique, la science des données est ainsi l'une des approches les plus indispensables dans l'industrie 4.0 [Uhlmann, 2017]. Le recours à l'internet des objets industriel (IIoT) et aux nouvelles technologies se développe à un rythme exponentiel et génère des quantités massives de données et connaissances industrielles. Ces dernières doivent être exploitées pour soutenir les objectifs stratégiques et opérationnels de l'entreprise. Par conséquent, il est indispensable d'adopter des technologies et les techniques de *big data* pour permettre l'analyse et l'exploitation des flux informationnels industriels [Al-Gumaei et al., 2018].

I.2.2. Le recours au Big Data pour l'Industrie 4.0

Le NIST (*National Institute of Standards and Technology*) définit les *big data* comme "des ensembles de données étendus - principalement en termes de volume, de variété, de vitesse et/ou de variabilité - qui nécessitent une architecture évolutive pour un stockage, une manipulation et une analyse efficaces" [NIST, 2015a]. La norme [ISO/IEC JTC 1, 2015] détaille ces caractéristiques des *big data* :

- Le volume : la quantité de données ;
- La variété : les nombreux types de sources et de formats différents ;
- La rapidité : le flux de données est continu et massif ;
- La variabilité : le changement continu dans le débit, la forme, la sémantique et la qualité des données ;

Dans le but de gérer ces ensembles de données, une plateforme *big data* est une solution informatique qui combine les caractéristiques et les capacités des applications et Framework

big data au sein d'une solution unique pour permettre aux entreprises de collecter, stocker, gérer, analyser et visualiser des volumes massifs de données et connaissances, avec beaucoup de rapidité, scalabilité et fiabilité [NIST, 2015b ; Site techopedia].

Les principaux défis des projets *big data* résident dans [NIST, 2015b] :

1. La capture et la capitalisation d'énormes quantités de données et connaissances provenant de sources multiples et hétérogènes avec une syntaxe et une sémantique différente ;
2. La préparation des données pour l'analyse en plusieurs étapes telles que la validation, le nettoyage, la transformation, l'indexation, l'agrégation et le stockage ;
3. Le choix de la technique de traitement appropriée qui correspond à la nature des données et aux exigences opérationnelles. Il peut s'agir d'un traitement par lots, où les données sont collectées au fil du temps puis envoyées pour traitement, ou d'un traitement en temps réel, où les données sont généralement produites de façon continue, en temps réel.

Malgré un recours massif aux technologies numériques, dans la stratégie de l'entreprise du futur la place de l'humain reste inestimable [Toro et al., 2015] en raison de sa capacité à manager, coordonner, à résoudre les problèmes [Brettel et al., 2015], à s'adapter avec les environnements de travail difficiles [Gorecky et al., 2017 ; Li et al., 2019] et à prendre des décisions [Stankovik, 2014]. Pour pérenniser cette ressource inestimable, la capitalisation des connaissances permet d'extraire l'expertise des différents acteurs humains afin de comprendre ces mécanismes de prise de décision et de les réutiliser devant des situations complexes. Il s'agit aussi de donner un outil fiable pour interpréter, selon le point de vue et les besoins de chaque expert métier, la grande quantité de données hétérogènes collectées à partir des machines et autres ressources dans l'atelier. Ceci conduit à une plus grande consolidation de la chaîne numérique en support à la chaîne de valeur globale de l'organisation [Kagermann et al., 2013]. Afin d'apporter une contribution face à ces enjeux, il serait important de comprendre quelles sont les mécanismes de prise de décision en entreprise.

I.2.3. La prise de décision dans le contexte d'Industrie 4.0

Certes, dans le contexte de cette révolution industrielle, l'évolution technologique joue un rôle mais il n'est pas toujours évident pour les décideurs de suivre cette transformation rapide. De plus, ils ne possèdent pas nécessairement l'expertise permettant de comprendre les technologies utilisées. Certaines situations, dépassent le périmètre des managers et experts,

peuvent en outre impliquer les opérateurs et autres experts opérationnels dans le processus de prise de décision. Une solution à cette multiplication des compétences techniques requises serait le renforcement continu des compétences des acteurs opérationnels et autres responsables pour être en phase avec la complexité croissante apportée par l'industrie 4.0 et les exigences de réactivité et de compétitivités.

Dès lors, que signifie dans ce contexte être un « bon » décideur ? Ce n'est pas un talent inné, mais bien quelque chose qui s'acquiert par des années d'exercice et par un entretien quotidien de son savoir. Une étude de Harvard a établi une liste de caractéristiques qui font qu'une décisions peut être qualifiée de bonne décision [Davenport, 2009] : elle doit résoudre un problème ; elle doit être cohérente avec les objectifs à atteindre ; elle est impartiale ; elle doit être facile à exécuter ; elle est souvent le résultat d'une collaboration et une tolérance des idée des autres ; elle doit être testée et validée suite à une analyse des avantages et des inconvénients ; elle doit être fondée sur un raisonnement solide et doit être transparente.

L'enjeu est donc de soutenir, par le biais d'un programme informatique, une prise de décision **efficace** conduisant à un résultat positif aussi bien sur le plan technique qu'organisationnel. On parle alors de systèmes d'aide à la décision.

I.2.4. Les systèmes d'aide à la décision :

Le système d'aide à la décision (DSS) est un outil efficace qui peut faciliter la prise de décision afin de résoudre le problème identifié. Les DSS intègrent de multiples fonctions telles que l'analyse, la modélisation, la prédiction, l'optimisation et le diagnostic, en combinant généralement des données et des modèles [Kulhavy, 2003].

Selon Bonczek (1980), un système d'aide à la décision (*Decision Support System, DSS*) est un système informatique composé, entre autres, d'un système linguistique, d'un système de connaissances et d'un système de traitement des problèmes qui interagissent les uns sur les autres [Daihani, 2001].

Il comprend généralement une interface utilisateur permettant de communiquer avec le décideur. Un système d'aide à la décision ne prend pas de décision, mais assiste le décideur humain en analysant les données et en présentant les informations traitées sous une forme qui lui est compréhensible. Les systèmes de décision sont généralement une combinaison d'ensembles de règles et de logique de décision (que l'on nomme « moteur de décision ou d'inférence ») qui s'appliquent sur des données particulières, traitées, organisées et accessibles.

Le moteur de décision peut agréger diverses données pour former des informations et connaissances utiles par le décideur, ou il peut rechercher les données pour trouver des modèles significatifs qui peuvent également être utiles [Crossland, 2016].

Dans le domaine industriel, nous distinguons trois catégories de décisions. Les décisions *stratégiques* permettent de représenter les orientations à long terme de l'entreprise. Par exemple, l'installation d'une nouvelle ligne de production ou l'investissement dans une nouvelle machine. Les décisions *tactiques* sont focalisées sur la planification de la production à moyen terme, comme une implémentation des orientations stratégiques. Enfin, le *niveau opérationnel* qui consiste à mettre en place les décisions prises au niveau tactique. Un exemple de décisions opérationnelles est l'ordonnancement et l'exécution à court terme des ordres de fabrication.

Les nouvelles générations de ces systèmes d'aide à la décision tirent plus d'avantage des technologies d'information et du *big data*. Cependant, compte tenu des attentes différentes des différents utilisateurs possibles de ces systèmes et de la variation des niveaux de compétence de ceux-ci, il est difficile d'imaginer une solution standard convenant à tous [Parasuraman et Riley, 1997]. Quel que soit le niveau hiérarchique, l'acceptation et l'utilisation de ce type de système requiert une première phase d'apprentissage. Par la suite, une phase intermédiaire d'application des nouveaux savoir-faire permet de valider les compétences, et enfin une dernière phase de perturbation permet de tester sa maîtrise des nouveaux outils et des moyens de la production [Mattsson et al., 2018]. Cependant, la mise en place de ces systèmes d'apprentissage et de test, nécessaires pour la construction de cette expertise, reste un défi difficile pour de nombreuses industries de production [Stentoft et al., 2019 ; Chengula et al., 2018].

Outre le travail sur les compétences des utilisateurs, une autre solution consiste à améliorer la disponibilité et l'accessibilité aux données tout au long du processus de fabrication, et garantir la continuité de la chaîne numérique entre les différents niveaux de management jusqu'au simple opérateur.

Par conséquent, les entreprises doivent revoir leurs capacités et stratégies organisationnelles afin d'identifier plus efficacement ses besoins en matière de données et informations, mais aussi les connaissances pertinentes [Li et al., 2019].

En effet, la multiplication des sources de données et la grande quantité de données collectées chaque jour ne sont pas sans poser des problèmes pour la prise de décision. Beaucoup de

problèmes d'incohérence et de difficultés d'accès à la « bonne information » sont souvent constatés. Pour remédier à cela, la traçabilité des processus de fabrication, la capitalisation des différents flux d'informations et leur analyse permettent de définir de nouveaux indicateurs agrégés qui simplifient l'interprétation des phénomènes réels. Dans un niveau supérieur, la capitalisation des connaissances des experts et leur intégration dans les systèmes d'aide à la décision peuvent aider les utilisateurs novices à mieux exploiter les données de contexte pour résoudre les différents problèmes. Cette question est au cœur de notre problématique de recherche comme il sera détaillé plus loin.

L'industrie d'usinage est l'une des industries qui travaillent sur l'évolution vers des industries du futur avec un usage de *big data*. Dans cette industrie, l'expertise métier pointue des opérateurs contraste avec un manque de numérisation historique de ce type d'atelier. Ces travaux de recherche traitent les thématiques de la gestion des connaissances comme support d'aide à la décision dans le domaine de l'usinage des pièces mécaniques de forte valeur ajoutée. La section suivante détaille les caractéristiques principales de ce domaine.

1.3. LE CONTEXTE DE LA THESE : L'USINAGE MECANIQUE A GRANDE VITESSE

1.3.1. Les enjeux de l'usinage mécanique

Le domaine de l'usinage a subi plusieurs améliorations depuis une trentaine d'années avec le passage de l'usinage conventionnel à l'usinage à grande vitesse. À l'époque de l'usinage conventionnel, l'automatisation des machines était très limitée et la présence d'un opérateur pour faire fonctionner la machine et la contrôler était primordiale. L'évolution informatique et l'apparition de l'électronique, ont permis de développer le domaine de l'usinage avec la création des Machines-Outils à Commande Numérique (MOCN) ainsi que la Fabrication Assistée par Ordinateur (FAO). La combinaison des MOCN et de la FAO est la base de ce que l'on nomme Usinage à Grande Vitesse (UGV).

L'usinage à grande vitesse est un procédé de fabrication qui prend en entrée une matière brute et génère en sortie une pièce définie à travers l'enlèvement de la matière, en utilisant des outils coupants spécifiques. Contrairement à l'usinage conventionnel, l'UGV travaille avec des vitesses de coupes plus élevées. Utiliser l'UGV c'est améliorer la qualité des pièces, augmenter la productivité et donc assurer un gain du temps de production et finalement favoriser la possibilité de faire des pièces de plus en plus complexes.

Quel que soit le type d'usinage, plus les pièces fabriquées sont complexes plus le risque de dégradation de l'état de la machine ou des outils est fréquent. Or, une telle dégradation implique une dégradation de la qualité de la pièce et dans la littérature, plusieurs travaux traitent des causes et des phénomènes qui peuvent affecter la qualité de l'usinage. Par exemple, de nombreux travaux s'intéressent à l'état de l'outil en traitent la détection ou le suivi de l'usure, la prédiction, et la détection des bris [Godreau, 2017 ; De Castelbajac, 2012].

L'étude de l'état de l'outil consiste à mesurer certaines de ses caractéristiques avant ou après l'opération d'usinage en utilisant des techniques d'analyse des signaux optiques ou laser. Si ces techniques permettent de contrôler l'outil après chaque usinage, elles restent limitées et ne peuvent pas renseigner sur l'évolution de l'usure pendant l'usinage. C'est pour remédier à cela que plusieurs recherches s'intéressent au contrôle en continu de l'usure à travers le calcul de plusieurs paramètres tels que les efforts ou le couple au niveau de différents composants : la broche, l'outil, le support-outil, ou même la pièce [Ritou et al., 2007 ; Kuljanic, 2009].

L'outil coupant n'est pas le seul à rentrer en compte dans la qualité finale de l'usinage et d'autres travaux s'intéressent à l'état de la broche ou de la MOCN dans son ensemble. Concernant la broche, les études réalisées s'appuient surtout sur l'étude de la fonction de rotation, par exemple l'étude de l'état des roulements. En effet, plus ces roulements se dégradent, plus les vibrations augmentent. L'analyse vibratoire est donc un moyen de déduire l'état de ces roulements [De Castelbajac, 2012]. Concernant la machine elle-même, d'autres travaux s'intéressent à la mesure de la puissance et du courant d'alimentation des axes d'avance, qui permettent le déplacement de la broche [Furet, 1994].

Toute dégradation de l'outil ou de la broche a comme conséquence l'apparition de défauts de surface sur la pièce usinée. La mesure et le contrôle de l'intégrité de la surface, en cherchant par exemple la présence de traces de broutement, représente aussi l'une des thématiques de recherche les plus prometteuses dans ce domaine [Ritou et al., 2018]. La figure 1 montre deux exemples réels de phénomènes, très fréquents en usinage, le « bris d'outil » et le « broutement ».



Figure 1. Exemple de bris d'outil (à gauche) et de broutement (à droite) [Godreau, 2017].

Dans ce travail de thèse, nous nous intéresserons à l'intégration des données et des connaissances liées à ces phénomènes, pour proposer de l'aide à la décision. Parmi ceux-ci, l'aide au diagnostic qui permet de trouver la cause de l'apparition et proposer des solutions pour limiter l'usure des outils et faire face à l'apparition de défauts de surface. En retour, cela permet d'assurer une bonne qualité des pièces usinées et une meilleure productivité. Les données et connaissances à intégrer pour permettre l'aide à la décision se basent principalement sur les procédés de surveillance d'usinage.

I.3.2. La surveillance de l'usinage

La surveillance de l'usinage consiste à contrôler le bon fonctionnement d'une ressource, ou le bon déroulement d'une étape d'un process. L'objectif est d'empêcher le passage à l'étape suivante dans le cas où un dysfonctionnement serait détecté. Il serait ainsi possible d'intervenir pour corriger le problème et passer à l'étape suivante.

Plusieurs paramètres peuvent être surveillés en usinage [Teti et al. 2010] et chaque type de surveillance garantit plusieurs avantages :

- Le contrôle des vibrations assure principalement une plus grande durée de vie de l'outil et de la broche ainsi qu'une meilleure qualité de production en toute sécurité ;
- Le contrôle de la puissance et de la force permet de réduire les temps morts dans le processus de production, d'assurer une configuration des machines différente de celle posant des problèmes ce qui favorise plus de sécurité et de respecter les durées de vies des ressources matérielles utilisées et principalement celles des outils ;
- Le contrôle du déplacement assure la prévention des rebuts de production ainsi que des différentes erreurs dues aux variations géométriques et dimensionnelles, la

réduction des coûts par rapport aux anciennes solutions de contrôle, et finalement une réduction de l'usure.

Ces exemples ne représentent pas, évidemment, toutes les possibilités de surveillance, mais confirment que l'objectif principal de la surveillance reste l'amélioration de la qualité d'usinage et des pièces produites.

I.3.3. Les problèmes de rupture de la chaîne numérique :

La chaîne numérique permet de gérer d'une façon cohérente tous les flux de données et informations circulant dans l'industrie, entre les différentes phases de la chaîne de valeur : de la conception jusqu'à la fabrication. Au fur et à mesure, plusieurs technologies informatiques sont mises en place en industrie pour assurer ou contrôler les différentes opérations de développement du produit, nous décrivons quelques exemples ici.

La chaîne numérique permet de créer un lien entre ces différentes phases pour garantir une fiabilité au moment de la transmission de l'information entre les différents départements d'une usine de fabrication. Dans la chaîne numérique spécifique au processus d'usinage de pièces mécaniques, nous pouvons classer les outils impliqués selon leur appartenance aux principales phases du processus : conception, industrialisation et production (figure 2).

La conception est la première phase du processus de réalisation d'une pièce mécanique. Cette étape commence au bureau d'études et consiste à concevoir la géométrie en trois dimensions d'une pièce en se basant sur les besoins et les exigences des clients. Elle utilise pour cela un logiciel de Conception Assistée par Ordinateur (CAO). Vient ensuite l'industrialisation, qui est la phase intermédiaire se passant au bureau des méthodes. Elle permet la génération de la gamme de fabrication (c'est-à-dire l'ensemble des étapes de la fabrication de la pièce mécanique), le choix des outils à utiliser pendant l'usinage, la détermination des trajectoires de ces outils, etc. Le service de programmation, en utilisant un logiciel de Fabrication Assistée par Ordinateur (FAO), va transformer les données fournies par le bureau d'étude en un programme spécifique, appelé code G (code ISO), qui contient les paramètres et les trajectoires à réaliser par la Machine-Outil à Commande Numérique. Ce code G sera ensuite transcrit en langage machine à travers un Post-Processeur. Il est aussi possible de réaliser une simulation, qui consiste à tester le code G et valider son fonctionnement de manière virtuelle pour assurer la conformité de la pièce usinée avec les exigences clients, ou détecter les éventuels problèmes de collision de l'outil.

La dernière phase est celle de production. C'est une étape planifiée par les logiciels de Gestion de Production Assistée par Ordinateur (GPAO). Cette étape commence quand la MOCN exécute le code ISO après avoir reçu l'ordre de fabrication et la gamme de fabrication. Une fois la pièce usinée, une première phase de contrôle est appliquée afin de valider sa conformité. En cas de défaut, la pièce passe au parachèvement afin de la rendre conforme aux exigences initiales. Un contrôle final est établi afin d'avoir une pièce finie prête à être livrée selon les spécifications du commanditaire.

La chaîne numérique contient les systèmes d'information dédiés à la gestion des données et de tous les processus et ressources de l'entreprise : le système de gestion des données techniques ou *Product Data Management* (PDM), le système de gestion des données de simulation ou *Simulation Data Management* (SDM), le système de gestion des processus de fabrication ou *Manufacturing Process Management* (MPM) et enfin le progiciel de gestion intégrée ou *Entreprise Ressource Planning* (ERP). Pour gérer les évolutions du produit tout au long de son cycle de vie, il est important de mettre en œuvre ces différentes solutions [Duigou, 2010]. Enfin, les systèmes de gestion du cycle de vie du produit (*Product Lifecycle Management* - PLM) permettent une stratégie d'entreprise pour la gestion intégrée de toutes les informations et processus liés aux produits au long de leur cycle de vie, de l'idée initiale jusqu'à la fin de vie [Saaksvuori, 2004]. Leur implémentation correspond à une intégration intelligente d'un ensemble de système d'information couvrant les différentes phases du cycle de vie du produit [Belkadi et al, 2015 ; Danjou, 2016]

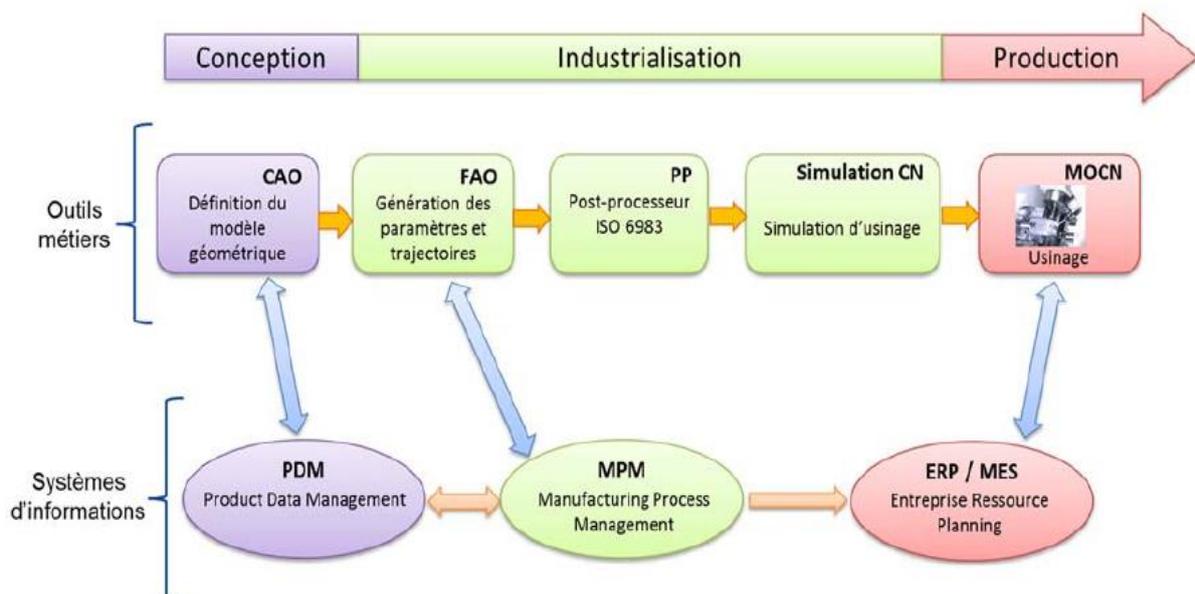


Figure 2. Les chaînes numériques des applications métiers et des systèmes d'information [Danjou, 2016]

Malgré les différents outils de gestion des flux d'informations existants, des problèmes sérieux de rupture de la chaîne numérique existent. C'est pour contrevenir à ces problèmes que des solutions ont été développées, solutions qui ont mené à l'apparition de la 4^{ème} révolution industrielle.

Parmi les phénomènes menant à une rupture de la chaîne numérique, le manque de traçabilité est un problème majeur dans la plupart des industries. Franck Cochoy et Gilbert De Terssac définissent la traçabilité comme étant l'opération qui consiste à affecter à un produit physique ou à une action de travail une ou plusieurs informations. Ces informations doivent permettre, le cas échéant, de suivre le produit et les actions associées « à la trace ». Elles permettent également de remonter dans le temps le processus et de rapporter les caractéristiques du produit aux différentes étapes/aux différents responsables de sa fabrication [Cochoy et De Terssac, 2000]. Il s'agit également d'identifier les différentes décisions prises dans ce processus.

Pour le contexte de ces travaux de recherche nous traitons la traçabilité comme : l'identification unique du produit (la pièce mécanique pour notre cas), l'analyse de la situation du produit et la génération des rapports, à chaque étape de sa fabrication et donc à chaque phase de la chaîne valeur dès l'acquisition de la matière première jusqu'à l'obtention d'un produit fini. Ainsi, une maîtrise du flux d'information industriel garantit une bonne traçabilité des conditions de mise en œuvre, des normes, des ressources et des procédures utilisées. La traçabilité est donc primordiale pour le suivi de la production, et assure automatiquement une qualité et une conformité des produits, une satisfaction de la clientèle et donc *in fine* une compétitivité de l'industrie.

Dans le cas particulier de l'industrie aéronautique qui nous occupe, le manque de traçabilité peut se présenter sous plusieurs formes. Par exemple, le parachèvement manuel : dans le cas d'une détection d'un problème de non-qualité par un technicien, celui-ci procède au parachèvement et son objectif est l'amélioration de la qualité de la pièce sans forcément décrire les étapes et les méthodes utilisées. Or, renseigner ces étapes et méthode c'est permettre de capitaliser l'expertise du technicien, mais également de conserver la traçabilité de l'ensemble des opérations subies par le produit. Dans le cas de changement du code ISO à la suite d'une non-conformité de la pièce, généralement les modifications apportées par le programmeur au niveau des instructions ne sont pas forcément tracées et capitalisées, ce manque peut causer une perte de temps.

En plus de la traçabilité, le manque de communication entre les différents services de l'entreprise est une cause importante de la rupture de la chaîne numérique. Aujourd'hui, la

plupart des industries souffrent d'un manque ou d'un défaut de communication entre les différents départements de la même industrie mais aussi entre plusieurs services du même département. Ce manque résulte d'un cloisonnement des données de chaque département. Dans les industries, il existe aussi un manque de communication entre le système d'information de l'entreprise et le réseau informatique de l'atelier. Inopportunément, ces problèmes s'aggravent beaucoup plus quand il s'agit d'une industrie dont les sites de production sont dispersés géographiquement.

La figure 3 illustre différents problèmes de rupture de la chaîne numérique industrielle.

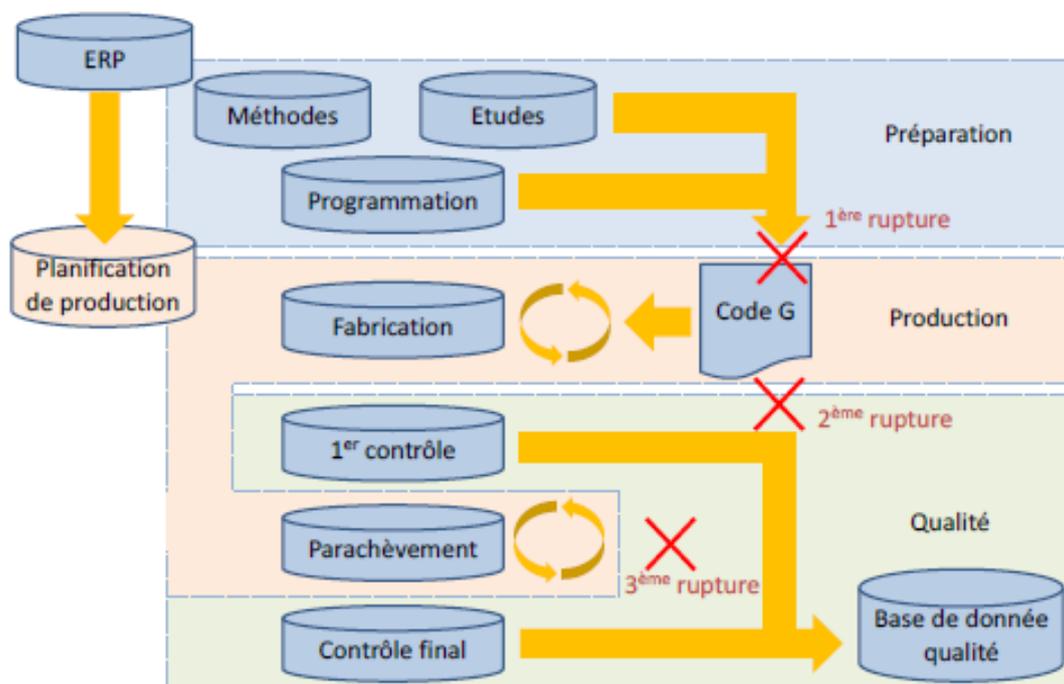


Figure 3. Problème de rupture de la chaîne numérique en usinage [Godreau, 2017]

Enfin, la diversité des moyens de production génère une variété de formats de données : celles-ci présentent par exemple plusieurs temporisations, liées à la différence des fréquences d'échantillonnage, d'acquisition, de traitement, etc. Il s'agit d'une incompatibilité, nuisant à l'interopérabilité entre les composants de la chaîne numérique. Qu'est-ce que l'interopérabilité ? La norme ISO 25010:2011, traitant la qualité logicielle, permet de mieux comprendre les notions de compatibilité et d'interopérabilité.

La Compatibilité est le degré auquel un produit, un système ou un composant peut échanger des informations avec d'autres produits, systèmes ou composants, et/ou exécuter ses fonctions

requis, tout en partageant le même environnement matériel ou logiciel. Cette caractéristique se compose des deux sous-caractéristiques suivantes :

- *La coexistence* : Degré auquel un produit peut remplir ses fonctions requises de manière efficace, tout en partageant un environnement commun et leurs ressources avec d'autres produits, sans impact négatif sur tout autre produit ;
- *L'interopérabilité* : Degré auquel deux ou plusieurs systèmes, produits ou composants peuvent échanger des informations et utiliser les informations qui ont été échangées.

C'est dans ce contexte, à savoir répondre à un ensemble de problématiques liées à la mise en place des approches de l'industrie 4.0 dans le domaine d'usinage aéronautique, que s'intègre le projet de recherche SmartEmma.

I.3.4. Le projet SmartEmma

L'objectif principal du projet SmartEmma est d'assurer une évolution dans le domaine de l'usinage à grande vitesse à travers le développement de machines-outils intelligentes et connectées. Ce projet s'intègre donc dans les enjeux de l'industrie 4.0, à travers la surveillance de l'usinage, la capitalisation des données et des connaissances, l'extraction des indicateurs de performances clés (*Key Performance Indicator*, KPI) et le pilotage de l'usinage afin de pouvoir prévoir des défaillances et prendre des décisions. La structure globale des différents axes de recherches du projet est représentée dans la figure 4.

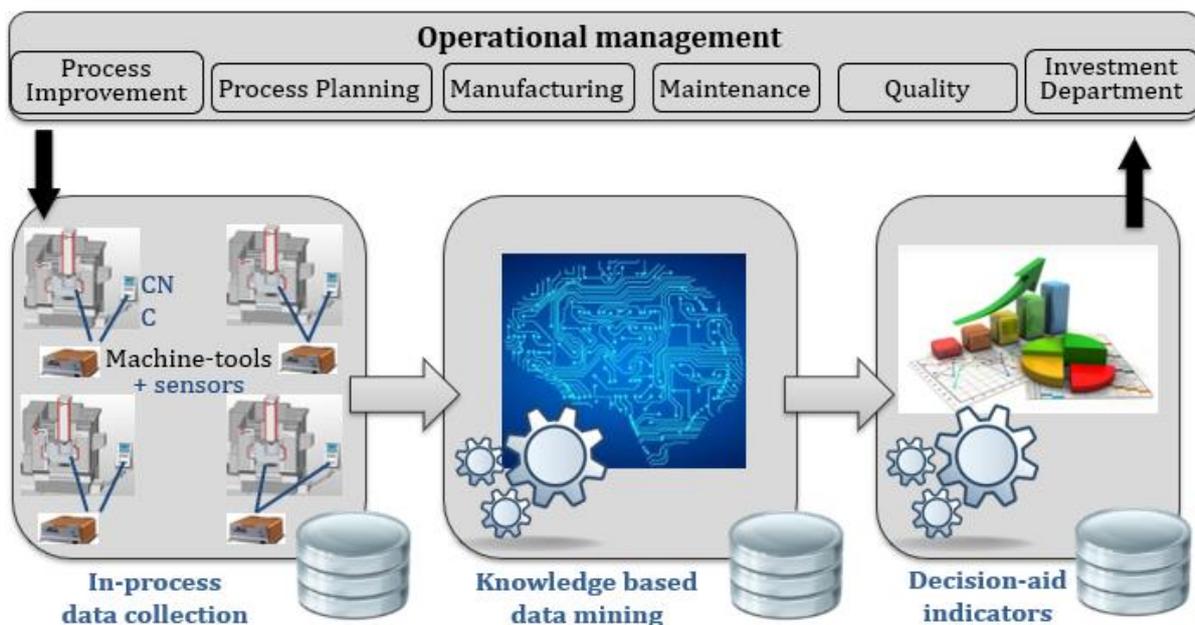


Figure 4. Structure globale du projet SmartEmma

SmartEmma est un projet collaboratif financé par l'ANR. Le consortium est composé de quatre partenaires : un laboratoire LS2N (Laboratoire des Sciences du Numérique de Nantes) avec deux équipes de recherche différentes et complémentaires (RoMas - *Robots and Machines for Manufacturing, Society and Services*) et IS3P (Ingénierie Système pour les Produits, Processus et Performance) et trois entreprises industrielles (Europe Technologies, en tant que fournisseur de solutions de capture et de traitement des données pour la surveillance de l'usinage), Airbus et Mécachrome (en tant qu'utilisateurs finaux dans l'industrie aéronautique). Le projet est approuvé par le pôle de compétitivité EMC2, dont tous les partenaires sont membres. Une brève description de ces entités est donnée en annexe.

Le processus d'aide à la décision dans le Framework SmartEmma, commence par une phase d'acquisition des données d'usinage issues de la MOCN. Ces signaux sont mesurés au plus près de la coupe à travers l'instrumentation de la MOCN et plus particulièrement de la broche par des capteurs capables de mesurer plusieurs types de données : puissances, températures, vitesses ou positions. Le Framework récupère aussi le contexte d'usinage issu de la Commande Numérique (CN) de la machine, tous ces données sont collectées dans des bases de données spécifiques. Par la suite, des méthodes de fouilles de données, d'intelligence artificielle et d'extraction des connaissances à partir des données sont appliqués pour identifier les indicateurs de performance et pour pouvoir récupérer un maximum de données et connaissances. Après cette phase de capitalisation, une méthode de modélisation adaptée et de gestion optimale des connaissances est nécessaire pour assurer l'interopérabilité.

Par ce processus, le Framework SmartEmma assure une bonne communication entre l'atelier et les différents bureaux et départements de l'entreprise et évite ainsi la rupture de la chaîne numérique. Deux types de pilotage sont identifiés, un premier type dit « ascendant » permettant, suite à une traçabilité, d'assurer une meilleure compréhension des processus de fabrication et le fonctionnement de l'atelier. Ce pilotage concerne des membres se trouvant au niveau supérieur des bureaux, et dans les services organisationnels. Un deuxième type dit « descendant », qui consiste principalement en des décisions prises par les bureaux, ou en des améliorations préventives, à destination de équipes d'usinage, afin d'assurer une bonne qualité et conditions d'usinage, et par conséquent, une bonne productivité et compétitivité.

Le Framework SmartEmma vient rajouter la couche « Smart » sur un équipement existant nommé EmmaTools. C'est un système de surveillance (figure 5) développé à l'origine par le LS2N durant un projet collaboratif et industrialisé par la société Europe Technologies. Ce système permet d'enregistrer à l'aide de plusieurs capteurs (vibrations, puissance moteur,

température, etc.) ce qui se passe sur la machine pendant l'usinage. Il est aussi exploité pour réaliser chaque jour une signature vibratoire de la broche afin de connaître son état et d'en déduire quotidiennement un niveau d'usure.

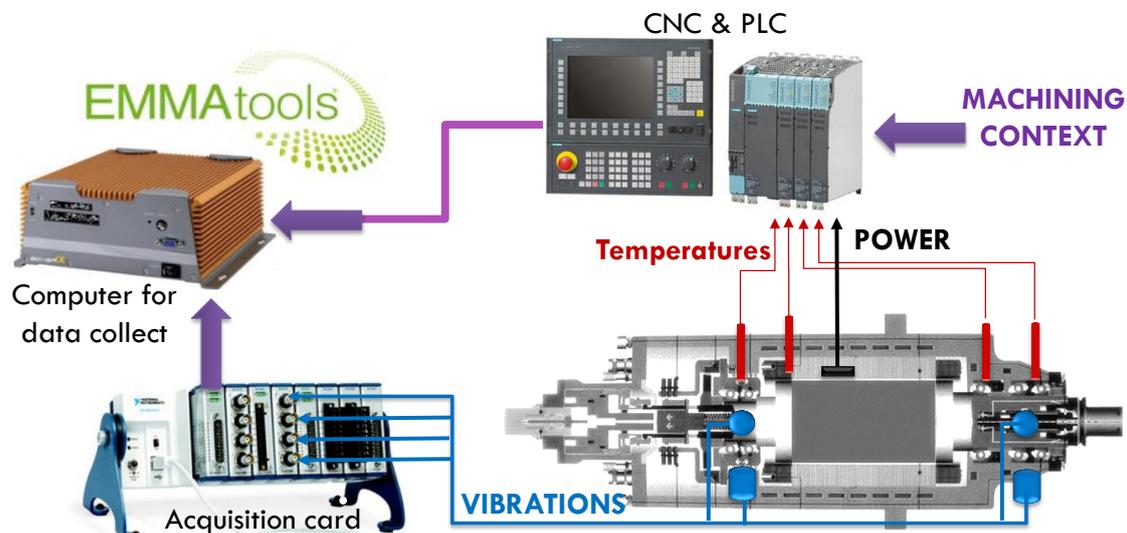


Figure 5. Schéma de principe de l'EmmaTools

La base de données EmmaTools est remplie à partir de deux sources d'information : la commande numérique et les capteurs installés sur les machines. La commande numérique fournit des informations sur le contexte ainsi que certaines données de capteurs internes au contrôleur. Les données de contexte correspondent aux informations telles que l'outil en broche, le programme en cours, le jour, etc. Les données des capteurs de la CN comprennent la vitesse de rotation de la broche, les différentes vitesses d'axe, la puissance instantanée de la broche, etc. Afin de compléter ces mesures, d'autres capteurs ont été intégrés dans la broche (Vibration et température).

L'objectif global de ce projet est de développer un système d'aide à la décision (SAD), participant à la transformation numérique industrielle, et permettant le suivi des phénomènes malveillants de l'usinage, affectant la productivité. L'originalité est de fournir pour l'ensemble des départements du management opérationnel un ensemble d'outils d'aide à la décision sur la base du même socle de connaissances, et répartis en quatre axes :

- *Reporting* et traçabilité d'information : à ce premier niveau d'aide à la décision, le système doit être capable de fournir les KPI et l'ensemble des informations nécessaires pour l'analyse des processus et la compréhension du bon fonctionnement afin de pouvoir établir un raisonnement et prendre des décisions ;

- Diagnostic et analyse : à ce deuxième niveau plus développé, le système doit se baser sur le premier axe de traçabilité pour pouvoir mettre en place une logique d'analyse automatique permettant d'établir des relations causes-effets. Les experts pourront par la suite détecter les comportements aberrants, distinguer les phénomènes indésirables, et surtout comprendre les causes d'apparition et les risques ;
- Prédiction : ce troisième axe permet d'anticiper l'occurrence d'un phénomène et d'une panne, tout en capitalisant un ensemble d'indicateurs et de contextes types jugé critique vis-à-vis de l'état de santé des différents équipements (MOCN, broche, outils coupants) ;
- Planification et réaction : ce dernier axe se base sur les trois précédents pour aller au-delà des limites de l'aide à la décision et pour assurer la prise de décision. Le système devient ainsi capable de réagir, essentiellement dans les situations opérationnelles critiques et les pannes techniques. Il devra être capable de déclencher des alertes et proposer des solutions ou des éventuelles corrections.

Les travaux de recherche, décrits dans ce manuscrit, s'intègrent essentiellement au niveau du deuxième axe d'aide à la décision « le diagnostic ». L'objectif ici est de développer un système d'aide à la décision à base de connaissances permettant l'analyse et la compréhension de la cause principale des phénomènes malveillants qui peuvent apparaître au moment de l'usinage.

I.4. POSITIONNEMENT ET PROBLEMATIQUES DE RECHERCHE

La revue des avancées technologiques dans le domaine de l'usinage montre que les solutions utilisées génèrent une grande masse de données très hétérogènes. Malgré les soupçons de l'impact de l'industrialisation sur la réduction de la place du travail humain, il est intéressant de constater que l'expert humain reste primordial pour améliorer la performance industrielle, en raison du niveau de spécialisation requis pour exercer ce type de tâche. Le défi est d'assister ces experts par des solutions qui assurent une facilité d'interprétation de la grande masse de données et la réutilisation de connaissances hétérogènes pour résoudre les problèmes rencontrés dans leur travail quotidien.

De ce fait, les problématiques d'intégration des données et connaissances sont au cœur des verrous scientifiques pour le déploiement des approches de l'industrie 4.0. Cependant, les problématiques concernant l'interopérabilité et la consistance des données sont également au cœur des enjeux pour la réduction des ruptures de la chaîne numérique, elle-même indissociable

de l'industrie 4.0. Un point commun entre ces deux problèmes est la réalisation de modèles robustes capables de supporter à la fois la capitalisation des connaissances et la collecte des données d'une façon consistante. Ces verrous sont la principale source de motivation pour cette thèse et ils en définissent la problématique scientifique.

Fort de cette analyse, nous pouvons formuler la question de recherche principale de la sorte :

Comment mettre les nouvelles générations de systèmes d'aide à la décision au profit de l'industrie du futur à travers une conciliation entre les données et les connaissances métier ?

Cette question de recherche fait appel à plusieurs problématiques scientifiques que nous résumons ci-dessous et auxquelles nous proposerons des réponses dans la suite de ce manuscrit :

- Comment peut-on aider les entreprises grâce à des systèmes d'aide à la décision exploitant au mieux les connaissances métier en les connectant aux différents types de données ?
- Quelles sont les connaissances industrielles disponibles et récupérables ? Et quelles sont celles indispensables pour le développement d'un système d'aide à la décision à base de connaissances ?
- Quelle stratégie adopter pour assurer l'intégration « données – connaissances » dans le système d'aide à la décision ?
- Comment assurer un raisonnement juste et efficace pour des perspectives d'aide à la décision ?
- Quelles sont les limites d'adaptabilité de la solution proposée pour d'autres domaines ?

I.5. DEMARCHE SCIENTIFIQUE

Pour résoudre les problématiques scientifiques citées ci-dessus, nous avons suivi une démarche en quatre phases : *l'analyse, le recueil, la modélisation et l'exploitation*. Cette démarche est représentée par la figure 6.

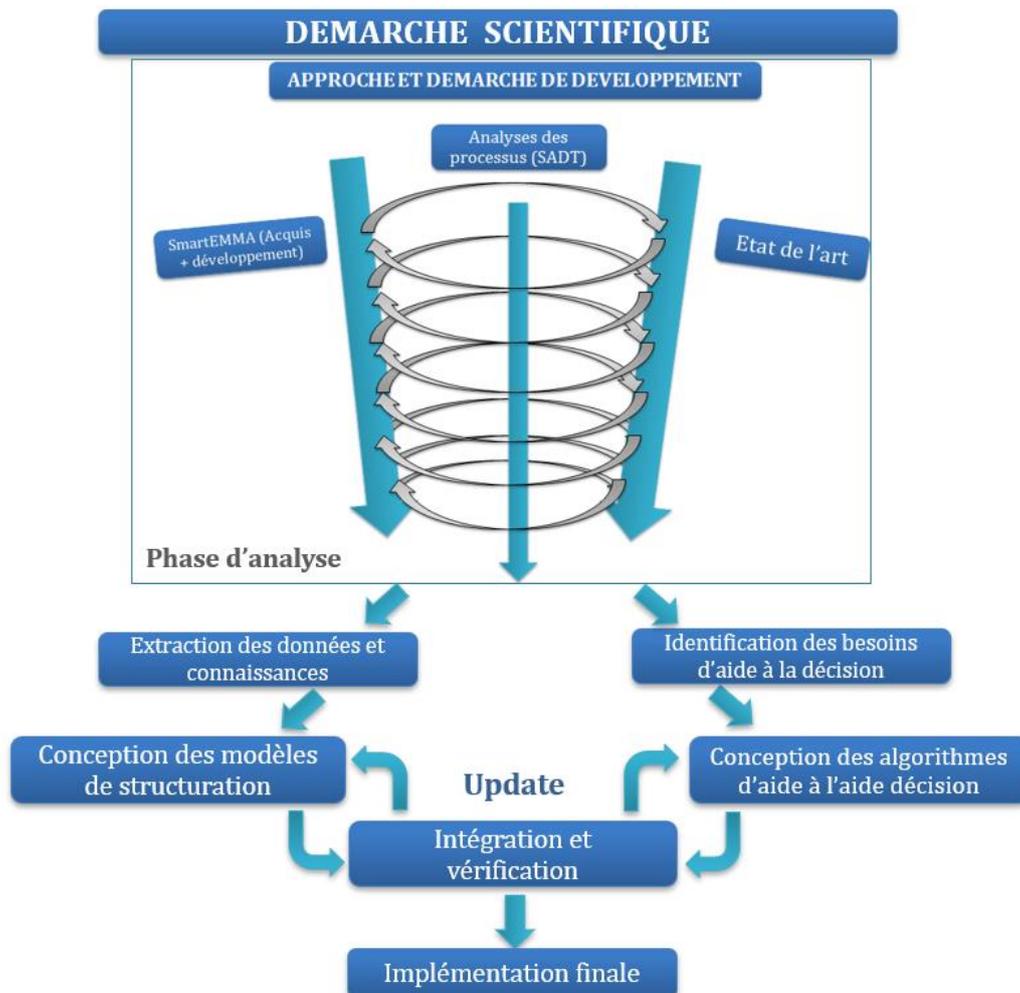


Figure 6. Démarche scientifique de développement

I.5.1. Analyse

La première phase consiste à l'analyse de 3 parties importantes :

- La première est l'analyse de la littérature et des thématiques de recherche similaire afin de distinguer les problématiques de recherche et les solutions existantes et ainsi de se positionner dans tout ce contexte pour mieux cerner les objectifs et pour essayer de trouver de nouvelles solutions ;
- La deuxième est l'analyse des pré-acquis du projet, des développements déjà réalisés au cours de la phase du lancement, des liens entre les différents acteurs du projet ainsi que de l'utilité et de la nécessité des solutions à proposer pour la mise en place du Framework global ;
- La troisième est l'analyse de la liste des besoins des partenaires industriels et la compréhension de leurs processus de fabrication pour identifier les flux de données circulant dans ces entreprises.

Cette première phase, nous a permis de collecter plusieurs types de données et connaissances. Par conséquent, cette phase de recueil a aidé à comprendre les flux d'informations industriels, la chaîne numérique, et à en extraire un premier niveau de connaissances. En parallèle, cette phase permet l'identification des besoins d'aide à la décision nécessaires pour l'implémentation des axes détaillés dans la partie 3.4.1.

I.5.2. Recueil

Dans la science de données, après l'étape de l'acquisition, il existe trois sous-étapes principales pour préparer les données et pouvoir les exploiter dans des algorithmes, ou voire même pour faire de l'extraction des connaissances dans les données l'ECD :

- La sélection : toutes les données ne sont pas toujours utiles pour répondre à une question ou pour expliquer un phénomène précis. Il est donc important de se limiter aux données qui portent une information sur le phénomène traité et d'éliminer le reste ;
- La préparation : pour des contraintes de fiabilité, il faut traiter les données pour éliminer les valeurs aberrantes et les bruits, et pour essayer de trouver et replacer les données manquantes ;
- La transformation des données, afin d'assurer une cohérence entre les données et pour pouvoir les exploiter dans des algorithmes. Cette étape est importante à cause de la diversification des sources d'acquisition et de temporisation des données.

Une fois que les données utiles sont bien structurées, une étape d'extraction de connaissances dans les données est primordiale. Cette extraction des connaissances se fait essentiellement à travers les processus d'analyse et de fouille de données. La mise en contexte des données, le raisonnement par situation et le regroupement des données selon des contraintes et des critères précis, permettent ainsi l'extraction de nouvelles connaissances.

L'enrichissement de cette première catégorie de connaissances s'avère également primordial. Cet enrichissement se fait à travers la capitalisation de l'expertise, les retours d'expériences, la littérature et l'analyse des anciens travaux de recherches dans des domaines similaires, les normes et les standards du domaine, les ontologies du domaine, etc. La capitalisation de ces différents types de connaissances hétérogènes permet de former une base de connaissances.

Cette base regroupe l'ensemble des connaissances hétérogènes du projet. Ces dernières sont différentes du point de vue des types, natures, sources, origines et formats. Cette variabilité

implique le recours à la modélisation afin de structurer et organiser ces données et connaissances et de créer des liens entre elles pour faciliter leur réutilisation et leur intégration

I.5.3. Modélisation

Une 3^{ème} phase de conception a été mise en place et plusieurs outils et méthodes de modélisation sont utilisés. Elle permet de structurer les bases de données et connaissances et de définir les algorithmes d'aide à la décision.

Suite au développement d'une première version des modèles et algorithmes, ceux-ci sont implémentés dans le système. Leur efficacité pour répondre aux besoins des utilisateurs est ainsi vérifiée, au fur et à mesure et en situation. La fiabilité des sous-systèmes d'aide à la décision étant fortement liée à la qualité des modèles de structuration de la base de connaissances, il est important d'établir un cycle d'amélioration et de mise à jour pour avoir finalement les algorithmes et les modèles les plus adéquats.

I.5.4. Implémentation

La dernière étape de notre démarche consiste en l'implémentation de l'approche d'aide à la décision à base de données et connaissances, dans un cas d'étude réel. Ceci permet de tester l'efficacité, et l'intérêt de la solution proposée. Ainsi, les bases de données et connaissances seront développées et implémentées avec des outils spécifiques. Le système de diagnostic sera aussi implémenté et intégré d'un Framework global d'aide à la décision testé qui est testé chez les partenaires industriels du projet.

I.6. CONCLUSION DU CHAPITRE

Ce chapitre a introduit le contexte industriel et scientifique de ces travaux de recherche ainsi que le cadre applicatif et le projet SmartEmma dans lequel cette thèse s'inscrit. Ces travaux étant au cœur des problématiques de la gestion des connaissances pour le développement des systèmes d'aide à la décision, plusieurs verrous scientifiques ont été identifiés et seront discutés dans la suite du manuscrit.

Enfin, la démarche scientifique que nous avons suivie tout au long de la réalisation de ces travaux a été détaillée. Les étapes d'analyse, de modélisation, et d'implémentation et exploitation des connaissances ainsi que la stratégie globale de gestion des connaissances font l'objet d'une description et d'un commentaire détaillé dans la suite de ce rapport de thèse.

CHAPITRE 2 : ÉTAT DE L'ART – LA GESTION DE LA CONNAISSANCE POUR L'AIDE A LA DECISION

**« Pour l'écrivain, la littérature est cette
parole qui dit jusqu'à la mort : je ne
commencerai pas à vivre avant de savoir
quel est le sens de la vie. »**

Roland Barthes

II.1. INTRODUCTION DU CHAPITRE

Ces travaux de recherche se basent sur la gestion de données et de connaissances pour supporter le développement d'un système d'aide à la décision (SAD) dans le contexte de l'industrie, pour la rendre plus intelligente, plus connectée et plus communicante : une industrie 4.0. L'analyse de la littérature permet de définir les différents concepts fondamentaux et de positionner notre approche par rapport à l'existant.

Ce chapitre décrit progressivement les concepts nécessaires à l'établissement du raisonnement de notre approche, en suivant la logique incrémentale de la méthodologie adoptée : analyse, modélisation, exploitation. Nous commençons par la description des notions de base, de la donnée jusqu'à la compétence, en mettant l'accent sur les typologies de connaissances. Après cela, les principes de gestion et de modélisation et les principaux exemples de modèles de connaissances sont détaillés. L'accent est mis sur les ontologies comme solutions robustes pour la modélisation sémantique. Ce chapitre sera clôturé par une étude approfondie des systèmes d'aide à la décision, leurs processus et les catégories existantes et l'on détaillera quelques exemples de travaux qui ont développé des SAD pour le diagnostic.

II.2. DE LA DONNEE AUX CONNAISSANCES

La quatrième révolution industrielle a favorisé l'émergence de nouvelles technologies dans le domaine industriel, ce qui a généré une profusion de données et de connaissances. Ces nouvelles ressources utilisées sont si différentes et complexes qu'elles impliquent une grande variabilité de types, de fréquences et de formats des flux d'informations qu'elles produisent. Pour suivre la distinction entre les données et les connaissances, nous nous appuyons dans cette partie sur le modèle DIKW (*Data, Information, Knowledge, Wisdom*, en français Données, Information, Connaissance, Sagesse/compétence), l'un des plus utilisés dans le domaine de la science de l'information et de la gestion des connaissances.

Ce modèle est généralement représenté par une pyramide (figure 7) de quatre niveaux fortement liés (le dernier niveau ne peut pas être atteint sans passer par le processus de transformation des autres niveaux).

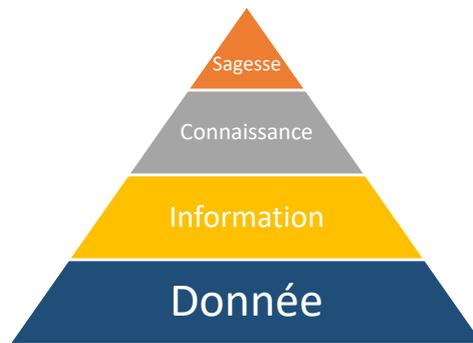


Figure 7. La pyramide DIKW

Dans la littérature, plusieurs travaux se sont basés sur ce modèle pour expliquer le processus de passage de la donnée aux connaissances [Ackoff, 1989 ; Zins, 2007 ; Rowley, 2007 ; Fricke, 2009 ; Ermine *et al.*, 2012 ; Baskarda et Koronios, 2013 ; Allen, 2017]. Plusieurs définitions ont été introduites dans la littérature pour définir les concepts de donnée et de connaissance, elles sont plus ou moins pertinentes en fonction du domaine d'étude. Parmi celles-ci, nous retenons celles qui correspondent à nos problématiques et qui font consensus dans la littérature.

La donnée est le reflet symbolique de nombres, quantités, grandeurs ou faits. Lorsque [Jashapara, 2005 ; Choo, 2006 ; Rowley, 2007] dressent un état de l'art sur le modèle DIKW, ils s'attardent sur les travaux de [Weggeman, 1997] traitant de la notion de données. Ces chercheurs ont travaillé sur l'extension du modèle DIKW en ajoutant une nouvelle catégorie de concepts qui est à l'origine de la donnée : le signal. Il est considéré comme la représentation directe de la réalité perçue. Le signal peut être traité par les sens humains, ou encore acquis par des capteurs spécifiques pour générer la donnée.

L'information est une donnée placée dans un contexte [J.L.C. Kemp, 1999]. Elle est aussi définie comme étant la donnée pertinente, utilisable, significative, utile ou traitée [Rowley, 2007]. [Ackoff, 1989] affirme que le passage de la donnée à l'information doit se construire en se posant les questions suivantes : « qui », « quoi », « où », « quand », ou « combien ».

La connaissance est l'information qui prend un (ou plusieurs) sens dans un (ou plusieurs) contexte(s) à travers les acteurs de l'entreprise [Ermine, 2000a]. Selon [Murray, 1996], la connaissance répond au « quoi » tout comme l'information, mais répond également au « pourquoi » et au « comment ». [Prax, 2000] pense que « la connaissance résulte d'une acquisition d'information et d'une action, elle est à la fois mémoire et processus de construction d'une représentation ».

La dénomination ou la notion du dernier niveau de la pyramide varie beaucoup en fonction de la discipline. On parle ainsi de *sagesse* en philosophie du numérique, ce qui sort du périmètre de ces travaux. D'autres distinguent plutôt la notion de compétence qui se rapproche de notre contexte de recherche. Nous traitons plutôt des notions d'intelligence, de savoir-faire, de savoir-agir et d'expertise : ces concepts se basent tous sur la connaissance. Par exemple l'intelligence est acquise grâce à l'accumulation de connaissances et à la sagacité nécessaire permettant de prendre les bonnes décisions. [Le Boterf 2000] précise que la compétence est synonyme de « savoir agir », qui signifie « la capacité de répondre à des prescriptions ouvertes dans des situations professionnelles plus larges et plus complexes ». Cet auteur définit également le savoir-faire, qui est relatif à la capacité d'effectuer une opération prescrite dans une organisation plutôt taylorienne. Dans ce projet de recherche, les connaissances sont aussi importantes que les données, et afin d'assurer une classification optimale des connaissances, il est important de détailler les principaux types de connaissance.

II.3. PRINCIPAUX TYPES DE CONNAISSANCES

Une analyse de la littérature permet de distinguer plusieurs catégories de connaissances en fonction de leurs formats, disponibilités, structures, facilités d'extraction, etc. Deux catégories, considérées comme les plus génériques, regroupent majoritairement l'ensemble des connaissances à traiter dans ces travaux de recherche. Ce sont les connaissances implicites et explicites.

Les connaissances implicites (*ou tacites*) sont celles qui n'ont pas de représentation claire mais qui transparaissent par les actions des membres de l'entreprise [Bar, 2006]. Ces connaissances sont spécifiques à une personne et leur formalisation est compliquée par le fait que la construction des connaissances à partir d'échanges verbaux dépend fortement du système d'interprétation cognitif propre à la personne en face. L'explication d'un scénario ou d'un incident observé en usinage par un opérateur peut s'intégrer parfaitement dans la catégorie des connaissances implicites. La traçabilité de ce type de connaissances permet de les transformer en connaissances *explicites*.

Les connaissances explicites sont, contrairement au premier type, formalisées et capitalisées. Elles sont donc stockées dans des archives de papiers, de documents, des bibliothèques, des fichiers et dossiers sur ordinateurs et des bases de données. Elles sont transmissibles et communiquées à travers des outils de formalisation ou des supports technologiques. Dans cette

catégorie s'intègrent toutes les données de la surveillance de l'usinage issues des capteurs, les documents et les fiches techniques de l'atelier, les fiches qualité, etc.

Plusieurs autres types de connaissances peuvent exister mais sont généralement considérés comme des sous-catégories de celles présentées ci-dessus. Par exemple, les règles métiers formalisées après des entretiens avec des experts métiers font partie des connaissances *structurées* [Chourabi, 2009]. D'ailleurs, ces mêmes règles métiers sont souvent appliquées sur un ensemble de connaissances en entrée pour pouvoir en générer d'autres en suivant un raisonnement spécifique. Ces nouvelles connaissances en sortie sont appelées *programmables* [Novins et Armstrong, 1998]. Le tableau ci-dessous, fait par [Dhouieb, 2016], résume les travaux de classification des principaux types de connaissances.

Tableau 1. Typologies des connaissances [Dhouieb, 2016]

Classification	Description	Auteur(s)
Connaissances implicites	Connaissances personnelles créées à partir des expériences et de l'intuition. Elles sont difficiles à spécifier et à communiquer.	[Nonaka et Takeuchi, 1995]
Connaissances explicites	Elles peuvent être codifiées formalisées et transmissibles.	
Connaissances Locales	C'est sont les connaissances qui s'appliquent seulement à un ensemble de conditions limitées.	[Novins et Armstrong, 1998]
Connaissances globales	Connaissances largement applicables	
Connaissances programmables	Les connaissances basées sur les règles.	
Connaissances uniques	Elles sont sensibles au contexte d'une situation spécifique.	
Connaissances pratiques	Connaissances dont l'action associée consiste à accomplir une transformation dans le monde matériel et physique.	[Lenne, 2009]
Connaissances théoriques	Connaissances dont l'action associée consiste à produire une explicitation dans un code de communication.	
Connaissances procédurales	Représentent comment un problème est résolu. Elle indique comment réaliser une tâche.	[Chourabi, 2009]
Connaissances déclaratives	Décrit ce qui est connu du problème à travers des énoncés ou des rapports.	
Connaissances heuristiques	C'est sont les connaissances empiriques acquises par un expert au fil de son expérience passée.	
Connaissances structurées	Décrivent un modèle mental de l'expert sous forme de structure.	
Connaissances stratégiques	Coût et effet élevés	[Xu, 2010]
Connaissances d'application	Coût faible, effet élevé.	
Connaissances basiques	Coût et effet faibles.	
Connaissances à abandonner	Coût élevé, effet faible.	

Face à cette variété de connaissances, il est important de mettre en place des moyens et des techniques permettant de faciliter les processus d'organisation des connaissances existantes et de production de nouvelles. Pour répondre à ces problématiques, la gestion des connaissances est la meilleure solution : l'ingénierie de la connaissance vient alors supporter la tâche de production de la connaissance. C'est une méthode particulièrement utile pour l'externalisation, c'est-à-dire la conversion de connaissances tacites en connaissances explicites. L'importance des connaissances tacites est aujourd'hui largement reconnue dans l'ingénierie et la gestion des connaissances [Schreiber et al., 2001].

II.4. LA GESTION DES CONNAISSANCES

Aujourd'hui, le défi ne se résume pas au recours et à l'utilisation d'un système d'information optimal et à des ressources technologiques développées : le vrai défi est de pouvoir adapter ces ressources pour faciliter la création des connaissances est surtout leur partage entre les différents acteurs de l'entreprise [Gunasekaran et E.W.T. Ngai, 2007b]. La gestion des connaissances permet de mettre en valeur les capacités de la connaissance à s'intégrer dans les différents niveaux décisionnels de l'entreprise : stratégique, tactique et opérationnel [Lalouette, 2013].

[Chow et al., 2005] définissent la gestion des connaissances comme l'ensemble des opérations et activités qui permettent la création et le stockage des connaissances pour pouvoir les distribuer et les appliquer dans les organisations. Toujours dans ce même contexte, les recherches continuent pour définir des fonctions de base du processus de gestion des connaissances. Plusieurs chercheurs associent la gestion à la capitalisation des connaissances. Par exemple, Prax précise que la gestion des connaissances est un processus de capitalisation et de création des savoirs, assurant la diffusion entre tous les acteurs d'une organisation. Elle représente aussi une approche de management des idées, des expériences, des pratiques, etc. [Prax, 2000]. Grundstein insiste également sur le fait que la capitalisation représente une vraie problématique pour toute activité, tandis qu'une phase de gestion des connaissances est cruciale après la capitalisation [Grundstein, 2000]. Enfin, nous pensons que la gestion des connaissances peut être définie comme :

Une étape centrale mettant en évidence l'ensemble des outils, méthodes et technologies nécessaires pour suivre le cycle de vie des connaissances dans n'importe quelle organisation ou système. Elle permet de gérer la capitalisation, la modélisation, la communication, la réutilisation et la génération des connaissances.

II.4.1. Le cycle de la gestion des connaissances

La méthode de gestion des connaissances de [Grundstein, 2009] est très bien adaptée à nos travaux. Sa méthode présente un cycle composé de 4 étapes et 1 processus support comme montré sur la figure 8 :

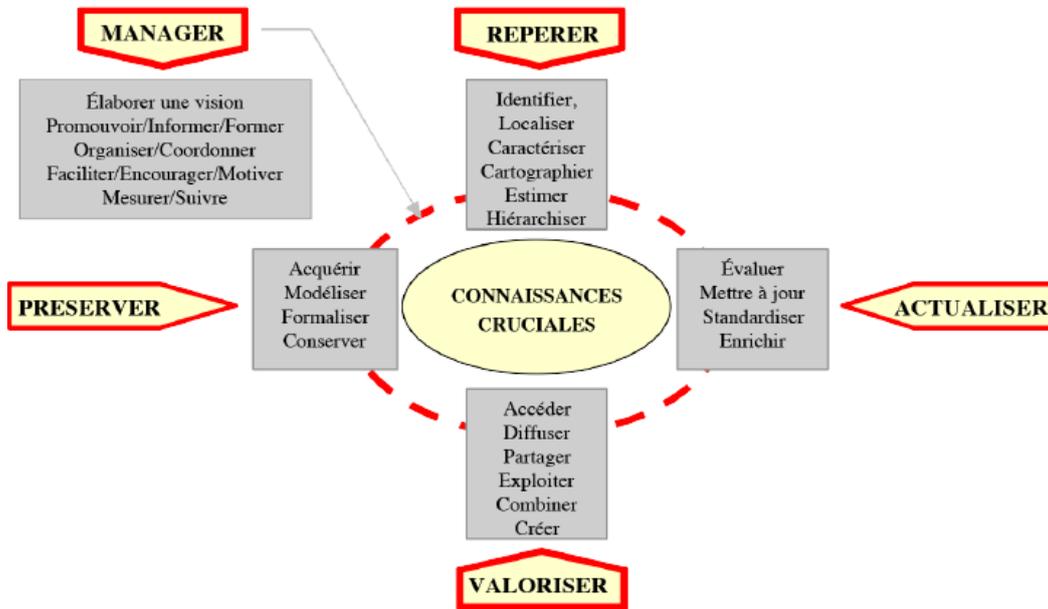


Figure 8. La méthodologie de KM de Grundstein

1. Repérer, c'est identifier et constituer une base de connaissances ;
2. Préserver, c'est structurer, modéliser et garder le plus de sens possible, afin de faciliter les interprétations ultérieures ;
3. Valoriser, c'est montrer, apporter et adapter à l'utilisateur ;
4. Actualiser, c'est maintenir et enrichir les connaissances, les mettre à jour et régénérer des nouvelles ;
5. Manager, pour mettre en place la démarche d'accompagnement au changement.

[Schreiber et al., 2001] admettent, suite à l'étude de la littérature, que la plupart des recherches traitant le cycle de la gestion des connaissances distinguent les 7 activités de base représentées dans la figure 9.



Figure 9. Les activités du cycle KM

- Identifier les connaissances existantes en interne et en externe ;
- Planifier les connaissances qui seront nécessaires à l'avenir ;
- Acquérir et/ou développer les connaissances nécessaires ;
- Distribuer les connaissances là où elles sont nécessaires ;
- Favoriser l'application des connaissances dans les processus opérationnels ;
- Contrôler la qualité des connaissances et les maintenir ;
- Disposer de la connaissance lorsqu'elle n'est plus nécessaire.

La différence entre ces deux cycles est que les premiers travaux de Grundstein structurent la gestion des connaissances autour des notions de connaissances vulnérables et cruciales, permettant de les repérer en fonction d'un jeu de critères (Rareté, Accessibilité, Coûts d'acquisition, Délais d'acquisition et Impact sur la stratégie, Impact sur le marché de l'entreprise, Impact sur la pérennité de l'entreprise) ; ce qui n'est pas le cas du cycle présenté par [Schreiber et al.,2001] qui s'intéressent dans un premier temps à toute les connaissances disponibles, et distinguent dans un second temps les connaissances nécessaires pour l'avenir.

La gestion des connaissances est très importante dans plusieurs domaines et particulièrement dans la production et la fabrication. D'ailleurs, [Berawi et Woodhead, 2005] pensent que l'amélioration de la stratégie de gestion de production se fait à travers le développement d'une bonne stratégie de gestion des connaissances. En effet, l'analyse des données et des connaissances disponibles permet de distinguer leur hétérogénéité due à la variabilité des sources, types, natures, fréquences, etc. Ainsi, un effort d'adaptabilité et de structuration est nécessaire pour pouvoir agréger, traiter et extraire des connaissances à partir de ces flux d'information [Farhoomand et Drury, 2002].

D'un point de vue pratique, le cycle de vie des données et des connaissances dans les entreprises et les industries commence toujours par l'étape de recueil et de capitalisation pour pouvoir, par la suite, les implémenter dans des systèmes de raisonnement à base de données et de connaissances : les systèmes experts, les systèmes d'aide à la décision, etc. [Olson et Wu, 2017b] traitent les processus de gestion des connaissances au service des systèmes d'aide à la décision et distinguent les 4 fonctions de base suivantes : identification, acquisition, stockage et récupération. L'étape d'acquisition ou de recueil est très importante puisqu'elle facilite la récupération [Olson, 2018]. Les techniques de recueils sont détaillées ci-après.

II.4.2. Le recueil de données et de connaissances

L'étude bibliographique permet de distinguer les cinq familles de méthodes ou outils de recueil de données et de connaissances les plus fréquents, et qui sont essentiellement : l'enregistrement, les observations, les questionnaires, les entretiens directs et l'analyse de données secondaires (ECD) [Leffondré et al., 2017]. Les deux premières techniques permettent surtout d'extraire les connaissances explicites contrairement aux autres qui extraient les connaissances implicites.

La technique basique d'enregistrement est très utilisée dans le contexte industriel réel avec le recours au *monitoring* qui permet de suivre les paramètres, les signaux et les grandeurs physiques utilisées dans le système de production, grâce aux capteurs logiques, analogiques et numériques.

L'observation est une technique très utilisée dans les domaines de sociologie et biologie pour assurer les études qualitatives. Elle permet de se focaliser sur le comportement et la situation pour expliquer un phénomène de façon très fidèle et fiable.

La méthode de questionnaire est utilisée lorsque le sujet de l'étude est bien maîtrisé, pour fournir des résultats quantifiés ou pour valider les résultats. Les questions peuvent être fermées avec un choix limité de réponses, ou ouvertes avec une entière liberté de réponse.

La meilleure méthode pour le recueil de l'expertise est l'entretien. Le processus de collecte de ces données impose une relation d'interaction et de confiance entre le cognicien d'une part, celui qui procède à l'externalisation de connaissances implicites, et l'expert du domaine d'autre part. Cette externalisation doit être complète dans le sens où le cognicien doit avoir la capacité et les outils nécessaires pour identifier la connaissance et la collecter dans un premier temps et pour la modéliser et la faire valider dans un second temps [Caulier, 1999]. L'expert est le spécialiste qui maîtrise le métier et qui possède le savoir-faire permettant de résoudre les différentes problématiques observées ou apparues. L'implication de cet acteur est très importante pour la réussite du développement du système expert. Il doit être motivé et disponible pour collaborer et appuyer la fiabilité du système [Bonnet et al., 1986] [Chaudhri et al., 2004]. Suite à cette collaboration, le cognicien doit identifier les connaissances nécessaires pour faciliter la prise de décision ou pour analyser un problème, comprendre un phénomène et proposer une solution. Dans ce contexte [Houriez, 1994] distingue quatre catégories de connaissances : les connaissances du domaine relatives aux différents concepts manipulés ; les connaissances sur les opérations et les inférences appliquées sur la première catégorie ; les connaissances sur les méthodes utilisées pour résoudre un problème ; les connaissances de contrôle.

En outre, une technique très efficace pour la capitalisation des connaissances à travers l'analyse des données est l'utilisation du processus d'extraction des connaissances à partir des données ou « ECD ». Il permet d'extraire le maximum de connaissances implicites à partir des données collectées et ceci en ayant recours à plusieurs techniques telles que les méthodes de *data mining*. D'après [Fayyad et al., 1996a], l'ECD représente : « un processus non trivial d'identification de nouveaux modèles valides, utiles, et compréhensibles à partir des bases de données ». À un niveau abstrait, l'ECD concerne le développement de méthodes, modèles et techniques pour donner un sens aux données. L'objectif de ce processus est la cartographie des données de bas niveau, souvent appelées brutes (qui sont généralement trop volumineuses pour être comprises et exploitées facilement), sous d'autres formes qui pourraient être plus compactes (par exemple, un rapport court), plus abstraites (une description approximative ou le modèle du processus qui a généré les données), ou plus utiles (un modèle prédictif pour l'estimation de la valeur de cas futurs) [Fayyad et al., 1996b]. En 2012, Boulila a réutilisé les travaux de [Fayyad et al., 1996a] pour proposer les étapes du processus d'ECD représentées par la figure suivante :

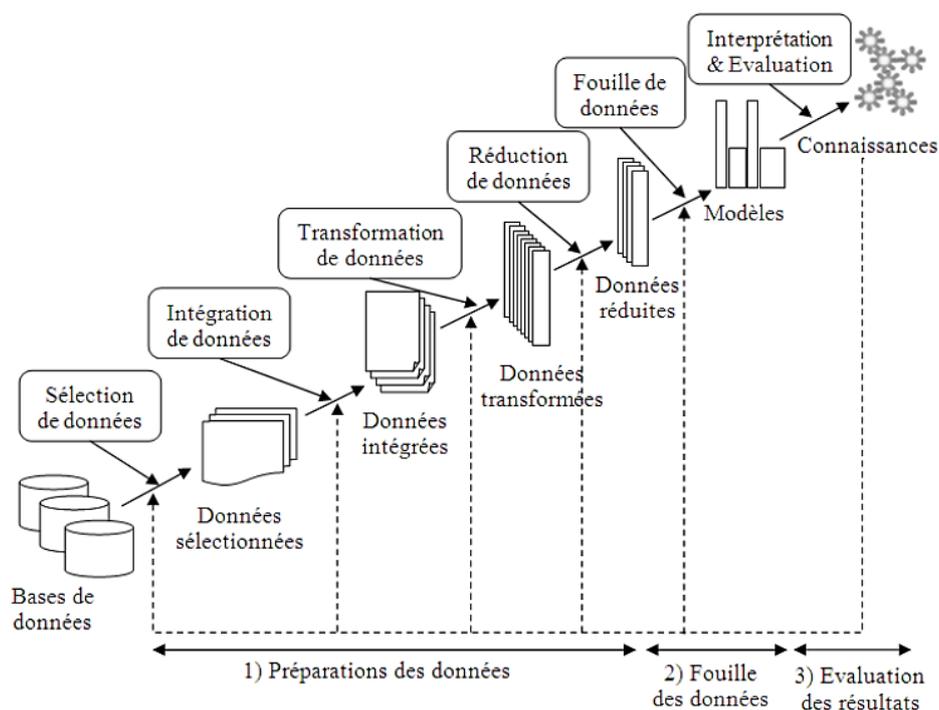


Figure 10. Étapes du processus d'ECD [Boulila, 2012]

Boulila a repris les travaux initiaux avec les différentes sous-étapes formant le processus et a défini trois étapes génériques qui sont :

- La préparation des données, qui définit le besoin initial et les objectifs attendus par le processus d'ECD, ce qui permettra de distinguer les types de données à utiliser ;

- La fouille de données, aussi connue comme exploration de données, *data mining*, prospection de données, ou encore forage de données, cette étape permet de proposer des modèles et des structurations intéressants à partir des données, selon des critères spécifiques et en utilisant des algorithmes divers issus de plusieurs domaines tels que l'IA, ou les statistiques. Il existe pour cela plusieurs méthodes, classées en trois grandes familles : les méthodes de classification et de structuration, les méthodes de visualisation et de description et enfin les méthodes d'explication et de prédiction [Zemirline, 2008 ; Zighed et Rakotomalala, 2000] ;
- L'évaluation des résultats est la dernière étape, qui clôture la boucle, permettant d'analyser les résultats obtenus et de sélectionner les connaissances à exploiter selon le besoin de l'utilisateur. Cette validation se fait à l'aide de deux méthodes : Les méthodes objectives utilisant les techniques statistiques et les méthodes subjectives utilisant les avis des experts métiers [Silberschatz et Tuzhilin, 1995]. C'est cette deuxième méthode que nous avons choisie dans le cadre de nos travaux de recherche.

Une fois le processus de recueil de données entamé, il est important de passer à l'étape de modélisation de données et de connaissances. Ces aspects sont traités dans la partie suivante.

II.5. MODELISATION DES DONNEES ET CONNAISSANCES

II.5.1. Notion de Modèle et de modélisation

La modélisation est utilisée dans plusieurs domaines tels que les sciences cognitives, l'ingénierie, les sciences de gestion, l'informatique, etc. Un modèle est généralement utilisé pour représenter et décrire de façon formelle et simplifiée un système complexe selon plusieurs points de vue correspondant à des intentions métiers différentes. [Candlot, 2006] s'est inspiré des différents critères à valider par un modèle (qui sont principalement la compréhensibilité, la maniabilité, l'expressivité et l'utilité) et des définitions citées dans [Labrousse 2004 ; @Gdt ; Harani, 1997 ; Rumbaugh, 1995 ; Vernadat, 1994] pour proposer la définition de modèle suivante : « Un modèle est un consensus sur une représentation partielle de la réalité ». De même, [Charlet, 2000] a défini la modélisation comme étant « une activité qui consiste à transformer ces descriptions textuelles en descriptions formelles, pour permettre la mise en œuvre *in fine* des connaissances décrites dans les textes ». D'ailleurs, la séparation entre les niveaux de données et de connaissances permet de distinguer les modèles de connaissances.

II.5.2. Modélisation des connaissances

II.5.2.1. Définitions

La définition de la modélisation proposée par [Paquette, 2002] est la plus adaptée à ce contexte d'étude, elle définit la modélisation comme étant :

Une activité de représentation des connaissances relatives à un domaine donné qui vise à identifier et structurer les connaissances en une représentation schématique pour les rendre visibles, manipulables, compréhensibles et communicables.

La modélisation des connaissances doit mettre à la disposition des experts la méthode la plus adéquate pour pouvoir exploiter au mieux l'ensemble des connaissances [Demoly et al., 2010], il est ainsi nécessaire de proposer une représentation bien structurée et riche [Belkadi et al., 2012]. Il s'agit ainsi d'une modélisation sémantique qui permet de rajouter du sens aux relations et aux objets, généralement, à travers les métadonnées. Le terme modélisation des connaissances est à l'origine un terme issu de l'intelligence artificielle. Il est utilisé pour déterminer la liste des connaissances à introduire dans un système informatique pour qu'il devienne « intelligent ».

« Un système dit intelligent est un système capable de faire des inférences (donc de générer de nouvelles connaissances) à partir d'une base de connaissances intégrée dans le système. » [Ammar Khodja, 2007]. Les ingénieurs en informatique capitalisent les données à travers les outils d'extraction des connaissances dans les données. Ils font ainsi l'extraction de connaissances qui seront par la suite formalisées, rendues compréhensibles et surtout exploitables. [Stojanovic, 2004] distingue deux types de modélisation des connaissances :

- **La modélisation descriptive** permettant de développer un modèle abstrait qui regroupe l'ensemble des connaissances du domaine ;
- **La modélisation opératoire** permettant de développer un modèle abstrait qui regroupe l'ensemble des connaissances de raisonnement nécessaires pour la résolution des problèmes et la prise de décision.

II.5.2.2. Structuration du modèle des connaissances

Après une analyse bibliographique des travaux de recherche sur la structuration des modèles de connaissances, nous avons particulièrement remarqué l'étude menée par [Schreiber et al., 2000]. Au vu de l'intérêt de cette méthode et de son adéquation avec notre problématique de recherche, l'approche que nous proposons est inspirée de cette structuration. Pour ce groupe de chercheurs, un modèle de connaissances comporte trois parties, chacune d'entre elles englobant

un groupe connexe de structures de connaissances. Chaque partie représente une catégorie de connaissances.

- La première catégorie est appelée *connaissances du domaine*. Cette catégorie distingue les connaissances spécifiques au domaine, nécessaires pour le développement d'un SBC. Par exemple, les connaissances du domaine d'une application de diagnostic utilisable dans une industrie de fabrication peuvent contenir des définitions de produits, de process, de ressources, d'anomalies, de causes possibles, ainsi que des relations entre ces types ;
- La deuxième partie du modèle de connaissance contient *les connaissances d'inférence*. Elles décrivent les étapes fondamentales d'inférence que nous voulons faire en utilisant les connaissances du domaine. Les inférences sont perçues comme étant les éléments constitutifs du système de raisonnement. Dans cette catégorie de connaissances nous pouvons placer les règles métiers formalisées suite à l'externalisation du savoir-faire des experts techniques, qui permettent d'analyser une situation et de prendre des décisions. Elles peuvent être une cause d'apparition d'une anomalie dans le contexte d'une application de diagnostic ;
- La troisième catégorie de connaissances contient *les connaissances de tâches*. Elle définit quel(s) objectif(s) sont à assurer par le SBC, et comment ces objectifs peuvent être atteints par une décomposition en sous-tâches et en inférences. Par exemple, une application de diagnostic des anomalies industrielles pourrait avoir pour tâche principale le diagnostic, pouvant être réalisé par une séquence de sous tâches de récupération de données, d'analyse et de prise de décision.

II.5.3. Les approches de modélisation

Afin de mieux maîtriser la modélisation des connaissances il est nécessaire d'étudier les approches méthodologiques pour ce type de modélisation. Nous pouvons classer ces approches en trois catégories : descendantes (modélisations préalables des connaissances) coopératives (modélisation des connaissances peu formalisées) ascendante (l'analyse d'informations pour construire des nouvelles connaissances).

Dans la littérature, une approche permet de structurer et clarifier la visibilité des modèles de données produit, il est possible de citer les approches suivantes :

- L'approche Produit, Processus, Ressources ou PPR qui a été traitée dans plusieurs travaux. [Cutting-Decelle et al., 2007 ; Agyapong-Kodua, 2014] confirment qu'une bonne modélisation des données ou des connaissances dans ce domaine doit tourner autour de ces trois axes. Dans ce contexte, les travaux de [Ramis Ferrer et al., 2016] se basent sur le recours à différentes ontologies de domaine et démontre comment elles peuvent être couplées en mappant des éléments de l'ontologie grâce à une approche basée sur des règles. La mise en œuvre de cette approche permet de fusionner des ontologies modulaires et donc de créer un lien entre les concepts et les données couverts par différents domaines qui sont fréquemment décrits avec des normes non liées. La cartographie des données est réalisée à l'aide d'un modèle ontologique qui permet non seulement la description des lignes d'assemblage, mais aussi la représentation des processus que les composants peuvent effectuer.
- L'approche *Function, Behaviour, Structure – Product Process Resource External Effects* « FBS-PPRE » (fonction comportement structure – produit processus ressource contexte) [Labrousse, 2004] étend le modèle FBS (*Function Behaviour Structure*) proposé dans les travaux de [Gero, 1990]. Les fonctions décrivent de manière abstraite le but d'un objet. Les fonctions opérationnelles sont formulées indépendamment de toute solution. Au contraire, les fonctions techniques dépendent des choix techniques. Le comportement décrit l'aspect dynamique d'un objet. Il comprend un ensemble de règles et un graphique d'états séquentiels représentant la transformation d'un objet stimulé au cours d'un processus. La structure définit les éléments qui font partie de l'objet. Elle définit également les attributs de ces parties. Le modèle FBS-PPRE permet des améliorations dans trois domaines principaux : l'exhaustivité de la modélisation, la gestion de la dynamique des objets et l'unification conceptuelle.
- L'approche *People Process Product* qui a été utilisée pour la première fois par Motorola dans son développement de *Six Sigma*. C'est une approche très importante puisqu'elle met en valeur la place de l'humain dans l'industrie, domaine dans lequel l'implication de l'être humain reste incontestable [Ericson et Larsson, 2009]. Cette dernière approche est assez importante surtout face aux problématiques actuelles de garantie de la place humaine au moment de la propagation de la quatrième révolution industrielle. L'usine du futur doit permettre la collaboration technologique et humaine collaborative pour répondre aux enjeux industriels actuels. [Westkämper, 2014] discute l'importance de la prise en compte de la dimension humaine dans l'usine du futur : « *Les usines sont le lieu de travail des humains. Les humains génèrent et dirigent les usines. Les humains sont la ressource la plus flexible*

du système de fabrication [...] mais leurs perceptions cognitives, physiques et tactiles sont limitées [...] Les humains peuvent utiliser l'intelligence informatique pour apprendre et surmonter l'"oubli" et profiter des connaissances informatiques. Les êtres humains sont capables d'apprendre, d'oublier, de communiquer et de coopérer dans l'environnement numérique [...] Leur motivation et leur créativité ainsi que leur expérience sont les moteurs de l'innovation, de l'adaptation et de l'optimisation dans le secteur manufacturier. »

Plusieurs modèles ont été développés pour la structuration des connaissances des entreprises, et plus spécifiquement celles des industries de fabrication. Dans la partie suivante quelques exemples de modèles sont présentés.

II.5.4. Exemples de modèles de connaissances

II.5.4.1. Le modèle d'information et de connaissance des moyens de la fabrication (MFIKM)

Ce modèle permet le stockage, l'accès et la gestion des informations et des connaissances liées aux installations de la fabrication. L'accent a été mis sur la connaissance des processus et des ressources pour soutenir les décisions de planification des processus. Ce travail explique également la pertinence de l'utilisation de différentes catégories de connaissances. L'ensemble des connaissances en matière de moyens de la fabrication est organisé dans une superclasse « *facility knowledge* » qui est divisée sous forme de connaissance des processus et des ressources. Cette superclasse peut inclure différentes connaissances de fabrication telles que le broyage, rodage, moulage, etc. Toutefois, les deux principales classes considérées pour la connaissance des processus sont la connaissance du fraisage et la connaissance du tournage [Guerra-Zubiaga et young, 2006]. Le modèle global du MFIKM est représenté par la figure 11.

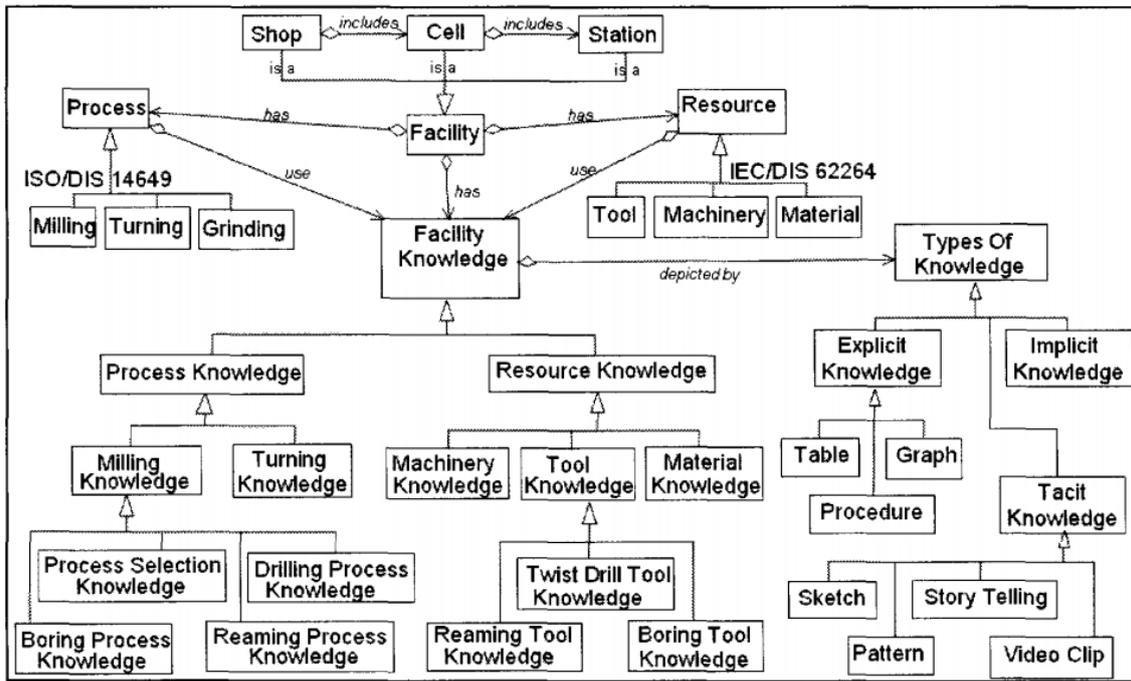


Figure 11. Le modèle d'information et de connaissance des installations de fabrication (MFIKM)
[Guerra-Zubiaga et young, 2006]

Il existe d'autres catégories de modèles de connaissances qui sont standardisés et relatifs à des normes spécifiques. Une analyse bibliographique a été établie pour la distinction des normes les plus adéquates et utiles pour le contexte de ces travaux de recherche.

II.5.4.2. Les modèles des normes et des standards

Nous nous sommes basés sur le catalogue des normes ISO, et nous avons distingué la norme 25 qui regroupe un ensemble de règles et définitions qui s'intéressent à normaliser l'environnement des techniques de la fabrication, représentée par la figure 12.

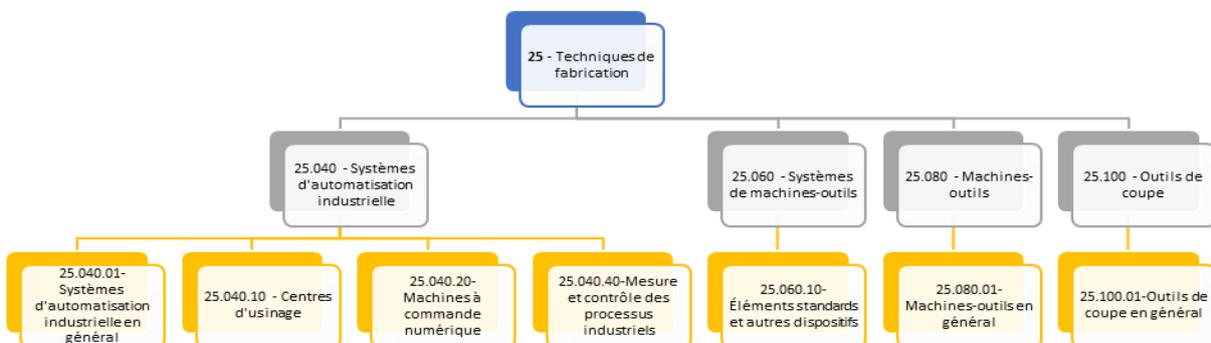


Figure 12. Différentes catégories de normes catégorie 25

Dans cette catégorie, il y a plusieurs sous-catégories. Les plus adaptées pour l'exploitation dans ces travaux de recherche, sont principalement : 25.040 - Systèmes d'automatisation industrielle, 25.060 - Systèmes de machines-outils, 25.080 - Machines-outils, 25.100 - Outils de coupe. Après l'analyse approfondie de ces différentes sous-catégories, quelques normes ont été jugées utiles pour une intégration dans nos travaux de recherche afin d'assurer la généricité des modèles proposés. Deux modèles ont été utilisés dans le système global, ceux des normes STEP-NC et MANDATE.

STEP-NC

STEP-NC (ISO 14649) (*STEP compliant Numerical Control*) a été élaboré et publié par l'ISO en juin 2006. Il remplace le G-Code (ISO 6893) et améliore le STEP (ISO 10303) (*Standard for the Exchange of Product model data*, Standard pour l'échange de données de produit). C'est un standard d'échange de données pour la programmation de commande numérique. Il permet surtout de faire la communication entre les différentes parties de la chaîne numérique CAO-CFAO-CN. L'intérêt de ce standard est de regrouper les propriétés du STEP en plus des informations de haut niveau concernant la fabrication du produit. Ce standard a été développé dans le but de trouver des solutions aux manquements de la chaîne numérique actuelle en ce qui concerne la fabrication sur machine-outil à commande numérique (MOCN). Pour détailler encore plus les caractéristiques du STEP-NC, des modèles de données ont été développés afin de structurer les différentes informations et données nécessaires à la génération du code.

Le fichier STEP-NC est, de plus, transportable d'une machine à l'autre sans nécessité d'adaptation car il comporte majoritairement des informations génériques traitables par tous les interpréteurs. La figure 13 représente une vue simplifiée du modèle de donnée standardisé STEP-NC pour les MOCN intelligentes

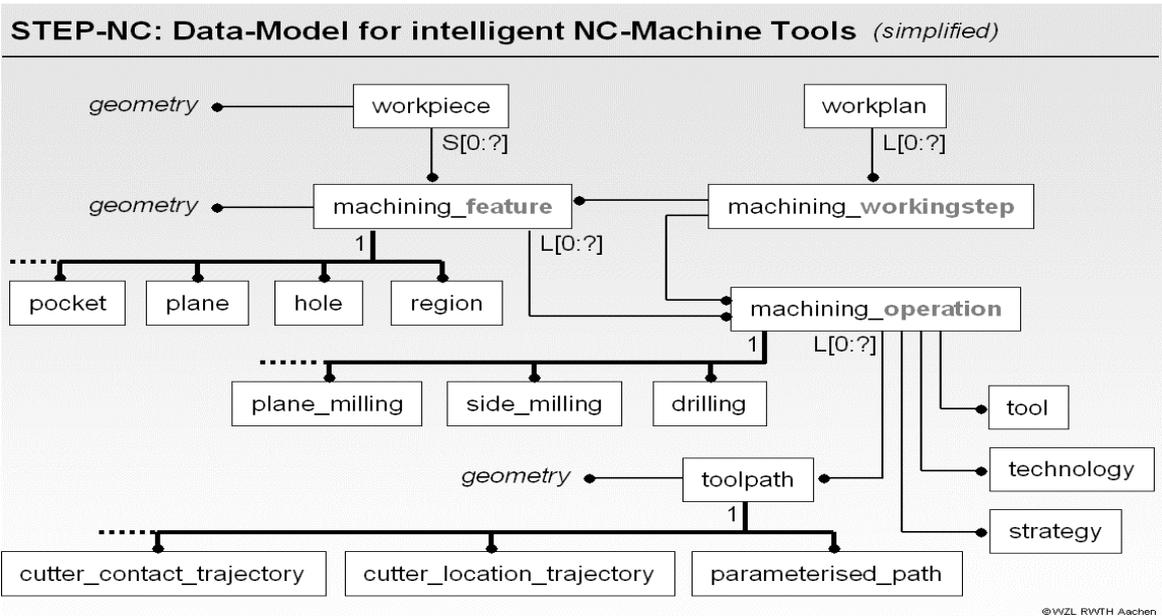


Figure 13. STEP-NC: Modèle de données pour MOCN intelligente [Hardwick et Loffredo,2007]

Dans la figure 14, le diagramme UML du générateur de code STEP-NC est présenté.

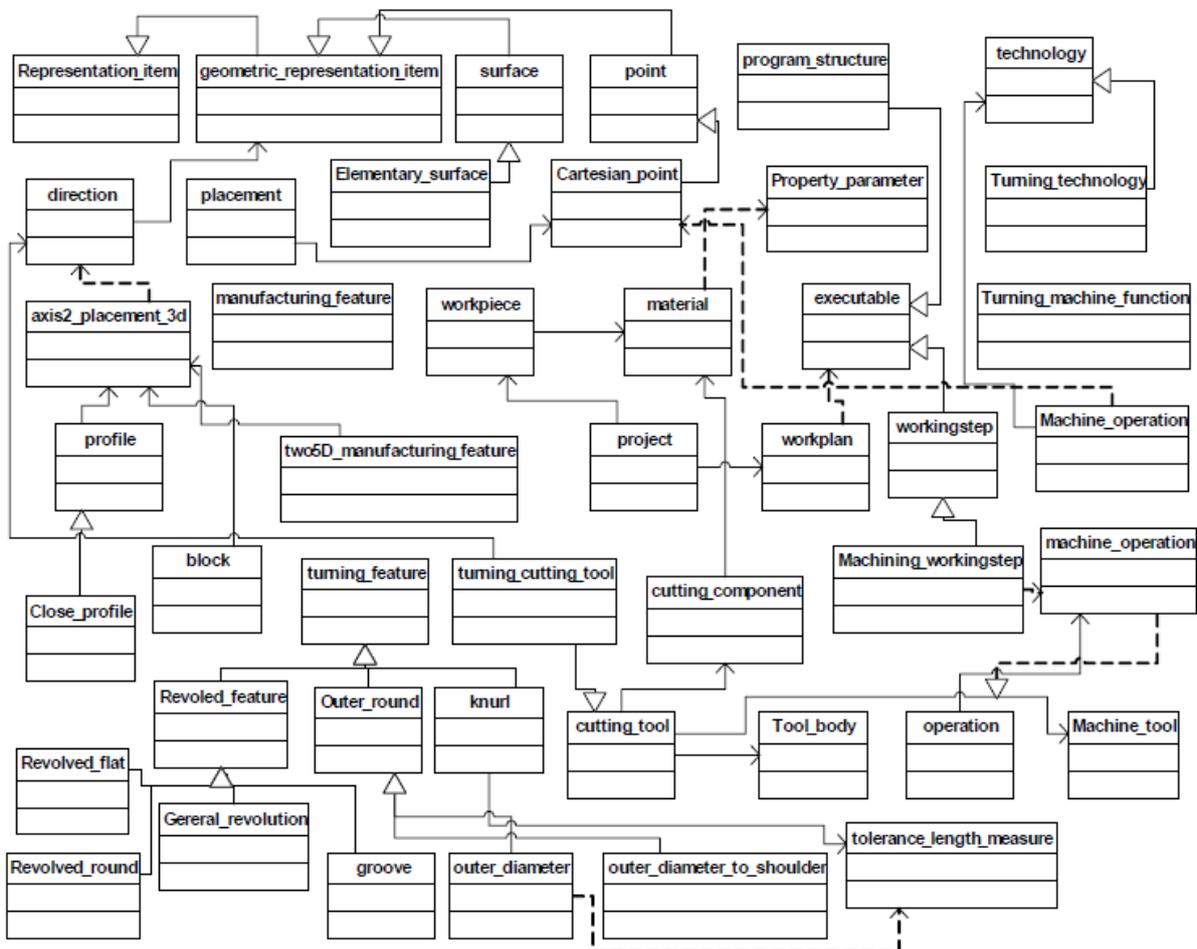


Figure 14. Diagramme UML du générateur du code STEP-NC

Par son format de données enrichies, le standard STEP-NC permet de réunir, en plus de la description complète de la géométrie de la pièce sous formes d'entités, toutes les informations relatives à la stratégie ainsi que les paramètres d'usinage en y associant également les informations sur les outils. De plus, le standard STEP-NC, par sa couverture des différentes informations permet une propagation des informations sur toutes les étapes de la chaîne numérique [Danjou, 2016].

Le problème du protocole STEP-NC est qu'il ne permet pas d'assurer la communication et les échanges avec les systèmes d'information de l'entreprise, et la capitalisation des connaissances. Dans la littérature, plusieurs sont les méthodes et les outils de capitalisation et de structuration des connaissances. En étudiant les travaux orientés vers la modélisation de données produit et processus de fabrication dans les industries de production, on constate qu'une majorité se met d'accord sur l'importance et l'utilité des modèles ontologiques pour la capitalisation et les échanges d'informations. Comme réponse à cette problématique [Danjou, 2016] a développé l'ontologie OntoSTEP-NC et le processus *Closed-Loop Manufacturing*. Le modèle OntoSTEP-NC est un support d'échanges de données au sein des chaînes numériques d'industrialisation. Le processus *Closed-Loop Manufacturing* est un cadre de capitalisation des connaissances et savoir-faire de fabrication, en vue d'une intégration au niveau de l'industrialisation.

MANDATE

La norme ISO 15531-1:2004 donne un aperçu général de l'ensemble de la norme ISO 15531 (MANDATE). Elle précise son champ d'application et fournit un certain nombre de définitions de base sur lesquelles l'ensemble de la norme est construit conformément à la « théorie générale des systèmes » et aux concepts définis dans le dictionnaire APICS. Ses annexes informatives fournissent une description des relations entre le MANDATE et d'autres normes (en particulier les normes ISO/TC 184), ainsi qu'une clarification des concepts de *capabilité* et *capacité* tels qu'ils sont utilisés dans le MANDATE et d'autres normes qui font référence explicitement ou implicitement à la théorie des systèmes.

MANDATE aborde la modélisation des données de gestion de la fabrication telles que :

- Données sur la gestion des ressources (modèle de ressources) ;
- Fonctions liées au temps (Modèle du temps) ;
- Données de gestion des flux dans la fabrication (modèle de gestion des flux).

MANDATE, en association avec les normes STEP, PLIB et d'autres normes SC 4 (ou non SC 4), peut être utilisé dans toute application logicielle qui traite des informations relatives à la gestion de la fabrication telles que les données de gestion des ressources et des flux. À ce titre, la norme vise à faciliter les échanges d'information entre les applications logicielles telles que ERP, les logiciels de gestion de la fabrication, les logiciels de gestion de la maintenance, les logiciels de devis, etc. MANDATE a été écrit en EXPRESS. Au cours des phases de développement de la norme MANDATE, la compatibilité de la norme avec la norme ISO 10303 (STEP) a fait l'objet d'une analyse approfondie [ISO 2004].

L'un des travaux les plus proches de notre sujet sont ceux de [Cutting-Decelle et al. 2007] qui ont développé des modèles de données et des ontologies et ont traité l'interopérabilité dans les systèmes de production, les scénarios de fabrication, les modèles de ressources, les modèles du temps, etc. la figure suivante capitalise les normes qui décrivent quelques composants de la chaîne de production durant le cycle de vie du produit.

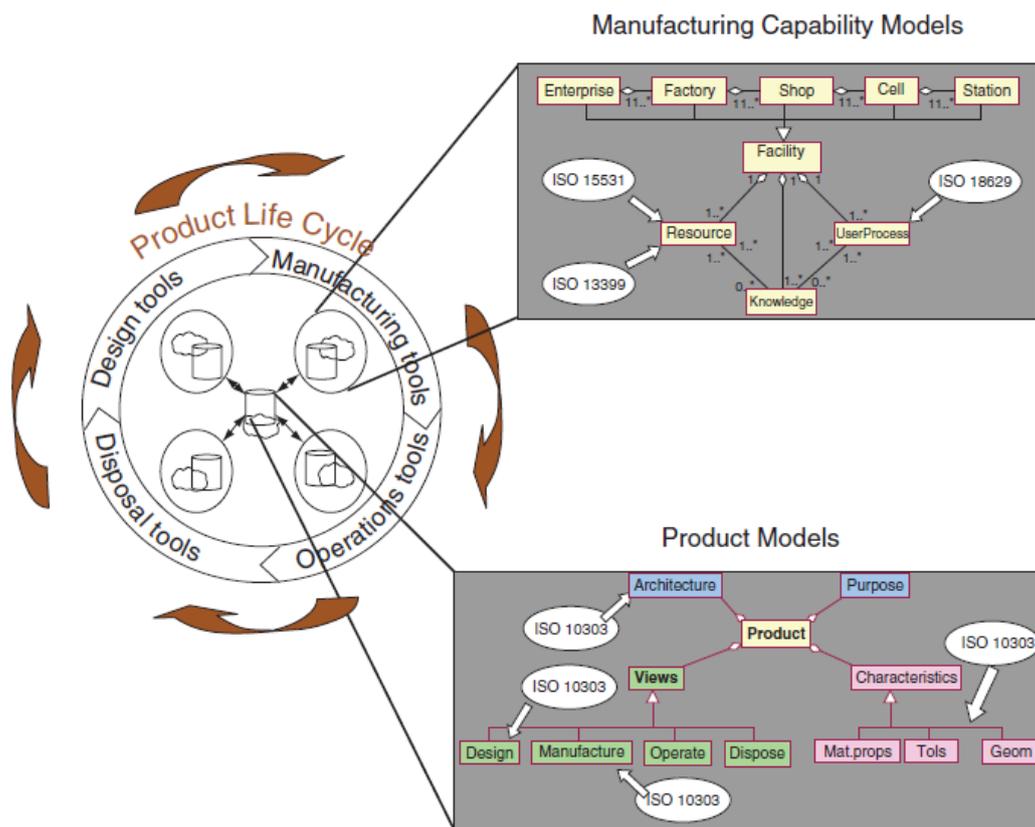


Figure 15. Modèle de produit dans le PLM [Cutting-Decelle et al., 2007]

II.5.5. Les ontologies pour la modélisation et la gestion des connaissances

Une ontologie est une technique de représentation et d'organisation des connaissances d'un domaine d'application précis en utilisant des mécanismes de raisonnement et d'inférence [Octaviani, 2015 ; Munir, 2017] (définitions en annexe 3).

II.5.5.1. Les ontologies de références

Dans un domaine spécifique, le plus simple est le recours à l'utilisation des ontologies de références [Annamalai et al., 2011 ; Borsato 2014 ; Song et al., 2013], considérées comme des ontologies de domaine d'un niveau très générique, elles sont souvent le résultat d'une fusion entre plusieurs ontologies spécifiques [Palmer et al., 2018].

Dans la littérature, il existe plusieurs exemples de ces ontologies de référence, par exemple : DOLCE [Masolo et al., 2004 ; Sanfilippo & Borgo, 2015] SUMO [Oberle et al., 2007] ou BFO [Arp et al., 2015].

D'autres ont eu recours à l'utilisation des normes pour le développement des ontologies, par exemple : OntoSTEP [Krima et al., 2009] en se basant sur la norme STEP (*Standard for the Exchange of Product model data*), ou encore « l'ontologie PSL » basée sur le langage de spécification des processus. PSL a été conçue pour faciliter l'échange correct et complet d'informations sur les processus entre les systèmes de fabrication, tels que l'ordonnancement, la modélisation des processus, la planification des processus, la planification de la production, la simulation, la gestion de projet, le flux de travail et la réingénierie des processus d'entreprise [Schlenoff et al., 2000 ; Gruninger and Kopena 2005].

Il existe plusieurs fondations dont l'objectif est de centraliser la recherche sur le développement d'ontologies dans chaque domaine spécifique. Dans le domaine de la fabrication numérique, l'*Industrial Ontologies Foundry* (IOF) tente de créer un ensemble d'ontologies de référence centrales et à accès libre qui couvrent l'ensemble du domaine [Kulvatunyou et al. 2018].

L'analyse de la littérature permet de distinguer plusieurs travaux qui utilisent les ontologies pour la modélisation des connaissances dans le domaine de la fabrication et la production de façon générale. Quelques exemples sont détaillés dans la partie suivante.

II.5.5.2. Exemples d'ontologie pour la modélisation des connaissances

Modèle de données de l'usine virtuelle

Les avantages d'une modélisation des données basée sur une ontologie ont été exploités pour le développement du modèle de données de l'usine virtuelle (*Virtual Factory Data Model* :

VFDM) dans le cadre du *Virtual Factory Framework* (VFF) [Efthymiou et al., 2015], qui a été introduit pour fournir un « environnement virtualisé collaboratif ». L'objectif de cet environnement était de faciliter le partage des connaissances et des ressources de l'usine et de soutenir les phases de planification de l'usine. Au cœur du VFF se trouve le *Virtual Factory Manager* (VFM) [Sacco et al., 2011], qui orchestre les modules de l'usine virtuelle, ou *Virtual Factory*, et régit le référentiel partagé qui stocke les données et connaissances définies dans la VFDM. Le VFDM vise à formaliser et à intégrer les concepts de produit, de processus, de ressource de production et de bâtiment tels qu'ils sont traités par les outils numériques qui soutiennent les phases du cycle de vie de l'usine. Les données de ce cycle de vie sont liées à de multiples domaines [Terkaj et al., 2012], à savoir : l'usine, le processus, le produit, les ressources de production, le système de production, la construction (la structure physique réelle de l'usine). De plus, le VFDM a été conçu pour exploiter les normes techniques existantes et étend leurs définitions pour représenter les caractéristiques d'un système de fabrication en terme de produits à fabriquer, de processus de fabrication qu'ils doivent subir et de ressources habilitées à exploiter les différentes opérations de fabrication. La version actuelle de la VFDM est principalement basée sur les normes IFC (*Industry Foundation Classes*) et STEP-NC qui ont été traduites en un ensemble d'ontologies [Sacco et al., 2011]. L'architecture du VFDM est représenté par la figure 16.

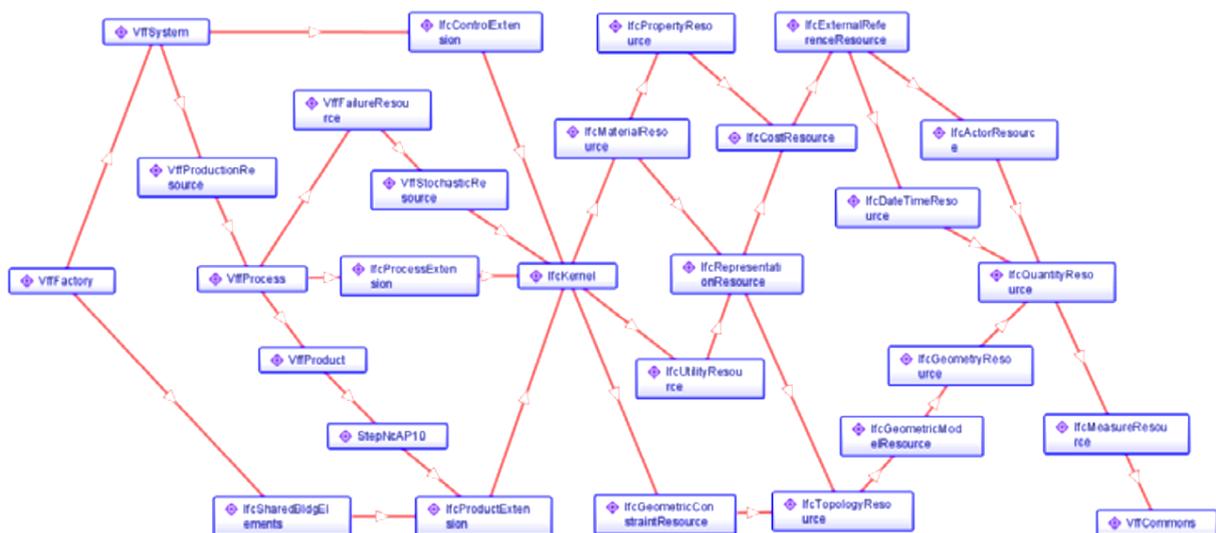


Figure 16. L'architecture du VFDM [Sacco et al., 2011]

Le VFDM est un modèle très pertinent pour créer un modèle de données de référence basé sur une ontologie car il montre comment une norme existante hors du domaine de la fabrication peut être adaptée pour une application dans le domaine de la fabrication. Toutefois, le modèle

ne peut pas être appliqué directement car d'une part il n'est pas accessible au public, et d'autre part il concerne des cas d'utilisation particuliers.

MASON (Manufacturing's Semantics ONtology)

L'ontologie MASON propose une architecture et des outils pour l'évaluation automatique des coûts de fabrication en s'appuyant sur les entités, les opérations, les ressources et les échanges d'informations haut-niveau. Deux applications de cette ontologie sont exposées : l'estimation automatique des coûts et un système multi agent sémantique orienté fabrication [Lemaignan et al., 2006]. La proposition s'appuie sur trois concepts principaux : les entités, les opérations et les ressources. Elle est inspirée de la décomposition de la fabrication de [Martin et D'Acunto, 2003] en produit, processus et ressource.

Les entités regroupent tous les concepts relatifs aux différentes entités mises en place pour spécifier un produit : entité administrative, entité d'assemblage, etc. Les opérations concernent la description du processus. Elles couvrent tous les processus liés à la fabrication dans une large mesure : les opérations de fabrication, y compris les opérations d'usinage, du contrôle et d'assemblage, les opérations humaines, logistiques, et de lancement. Les ressources sont classées en trois catégories : matérielles, humaines et géographiques. La figure 17 montre les principaux sous-concepts qui héritent des concepts principaux ainsi que les principales relations entre ces concepts.

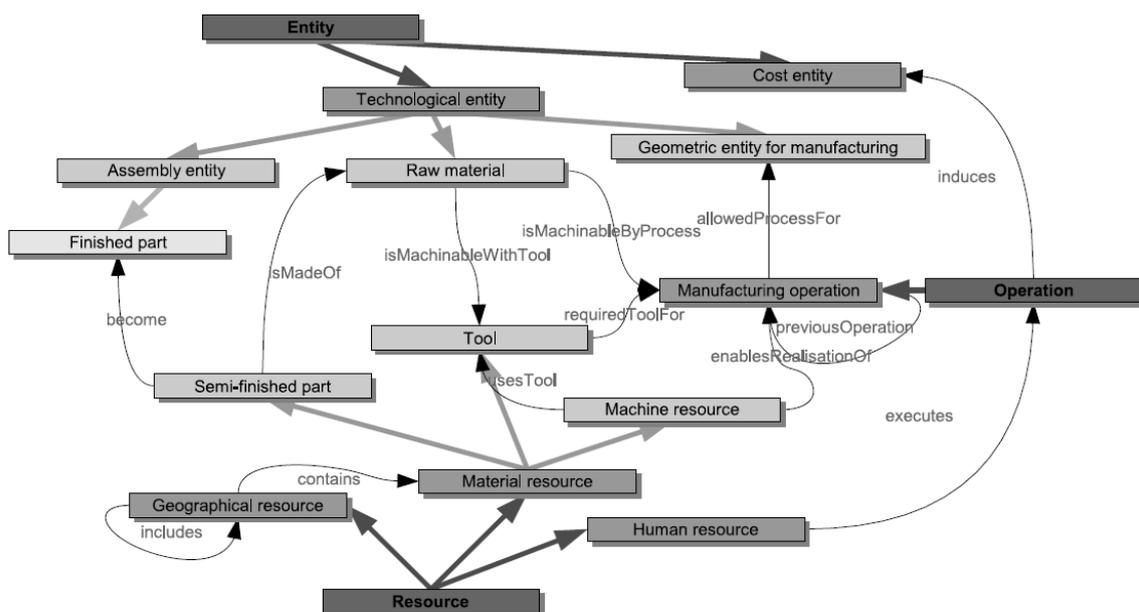


Figure 17. Aperçu des principales classes et propriétés objets de l'ontologie

[Lemaignan et al., 2006]

Il existe d'autres approches intéressantes telles que *Onto-pdm (ONTOlogy for product data management)*. C'est un modèle ontologique qui permet de définir une organisation pour assurer l'interopérabilité des systèmes d'information en favorisant l'échange d'informations liées à la production [Panetto et al., 2012]. De même, *PRONTO (PROduct ONTOlogy)* est une approche ontologique permettant de définir un modèle global qui puisse renseigner toutes les informations produites pour tous les acteurs du processus. L'objectif est d'enrichir les données produites pour connaître les incompatibilités d'un produit dans un système [Giménez et al., 2008].

Si de nombreuses ontologies produites pour l'échange et la capitalisation des informations de production existent, il n'existe pas beaucoup de modèles ontologiques qui présentent une granularité suffisante pour supporter les informations spécifiques à l'usinage. Il paraît ainsi évident de poser la question suivante « Comment choisir la meilleure façon de structuration des données pour l'interopérabilité des chaînes numériques d'industrialisation ? » [Danjou, 2016].

Après l'analyse des solutions ontologiques proposées dans la littérature, et compte tenu de leur intérêt pour assurer l'interopérabilité et la communication des connaissances entre les services de l'industrie, il convient de noter que la plupart de ces travaux sont orientés vers une communication partielle. Elle est assurée uniquement entre les services directement impliqués dans la réalisation du produit : essentiellement entre l'atelier et les services de conception ou de fabrication. Cela implique un manque de communication entre les différents services de l'industrie, qui sont étroitement ou indirectement liés à la réalisation d'un produit, ainsi qu'à la communication entre les différents systèmes d'information de l'entreprise.

En outre, les solutions proposées ne gèrent pas toutes les connaissances hétérogènes provenant de sources et de formats différents. Un autre inconvénient est le manque de généralité des solutions identifiées. Cela implique une réutilisation difficile dans d'autres domaines.

Cette partie de revue de la littérature, à propos des méthodes de structuration des données et connaissances, soulève dans un premier temps la question de recherche suivante :

- *Comment développer la meilleure approche de gestion des connaissances capable de structurer la variété des données et des connaissances hétérogènes ?*

Dans un second temps, une fois que les données et connaissances sont bien structurées et capitalisées dans des bases spécifiques la question suivante se pose :

- *Dans quelles catégories d'application ces bases sont-elles efficaces et comment peut-on les intégrer dans des systèmes industriels réel ? Et comment elles sont exploitées pour la réutilisation et la génération de nouvelles connaissances ?*

L'objectif global de ces travaux est de développer un système d'aide à la décision à base de connaissances. Une analyse bibliographique a été établie pour définir cette catégorie de SAD et trouver quelques réponses, dans la littérature, à ces questions de recherches.

II.6. LES SYSTEMES D'AIDE A LA DECISION

Un SAD est un système d'information interactif, flexible, adaptable et spécifiquement développé pour fournir aux décideurs des informations pertinentes qui les aideront à prendre les meilleures décisions. La coopération homme-machine est importante dans l'aide à la décision. Elle assiste le décideur humain et garanti donc le partage ou la répartition des tâches entre l'homme et la machine [Milot et Lemoine, 1998].

Les SAD sont installés pour améliorer la productivité. Les systèmes de fabrication traditionnels se transforment en environnements de fabrication fondés sur la connaissance, basés sur la numérisation, les réseaux informatiques, l'intelligence artificielle et la réaction rapide tout au long de la chaîne de production. Les recherches autour du développement des solutions technologiques capables de supporter l'aide à la prise de décision pour résoudre les problèmes ont commencé à partir des années 1970 [Gorry et Scott Morton, 1971]. La mise en place d'un SAD consiste au développement et à l'implémentation de systèmes utilisant les technologies de l'information et de la communication pour assurer la prise de décision et la résolution des problèmes [Shim et al., 2002].

Les SAD sont très utiles pour améliorer et faciliter le management dans les différents niveaux de décision dans l'industrie : le niveau stratégique qui s'intéresse aux actions planifiées à long terme pour l'amélioration de la productivité, le niveau tactique qui gère les processus et les ressources et le niveau opérationnel pour le contrôle et la gestion des opérations spécifiques. [Klement et al., 2017]

Selon [Marakas, 2003], les composantes d'un SAD peuvent généralement être classées en cinq parties distinctes :

- Un système de gestion de base de données et la base de données associée : qui stocke, organise, trie et renvoie les données pertinentes pour un contexte particulier de prise de décision ;

- Un système de gestion de base de modèles et la base de modèles associée : qui a un rôle similaire au système de gestion de base de données, sauf qu'il organise, trie et stocke les modèles quantitatifs de l'organisation ;
- Le moteur d'inférence et la base de connaissances : qui effectue les tâches relatives à la reconnaissance des problèmes et à la génération de solutions finales ou intermédiaires, ainsi que les fonctions relatives à la gestion du processus de résolution des problèmes ;
- Une interface utilisateur : qui représente un élément clé dans les fonctions du système global ;
- Un utilisateur : qui constitue une partie intégrale du processus de résolution des problèmes.

Il existe trois types de SAD [Haettenschwiler, 2001] :

- Le passif qui permet l'implémentation du processus de prise de décision mais ne permet pas de proposer des solutions de résolution ;
- Le SAD actif, qui contrairement au premier permet la génération de solution ;
- Le Coopératif qui permet de faire intervenir le décideur dans la proposition de solution finale, celui-ci pouvant valider, modifier ou redéfinir la décision proposée par le système. Ce dernier prend en considération les modifications établies par le décideur avant la validation finale.

II.6.1. Processus de SAD

[Drucker, 1956 ; Simon,1965 ; Archer, 1980] définissent 3 phases primordiales pour le processus de prise de décision. La première est la phase d'intelligence qui s'occupe essentiellement de la définition de la problématique.

La deuxième intermédiaire pour la structuration et la définition des liens entre les problèmes et les solutions proposé est la phase de conception et modélisation. Et enfin, une dernière phase, la phase du choix, permet la prise de décision et la proposition de solution finale validée. [Mintzberg et al., 1976] définissent ces trois phases comme l'identification, le développement et la sélection. En se basant sur ces 3 phases, plusieurs autres travaux continuent jusqu'à aujourd'hui à travailler sur la définition des étapes du processus d'aide à la décision. Les recherches les plus citées restent celles de [Shim et al., 2002] qui définissent les 7 étapes représentées par la figure 19.

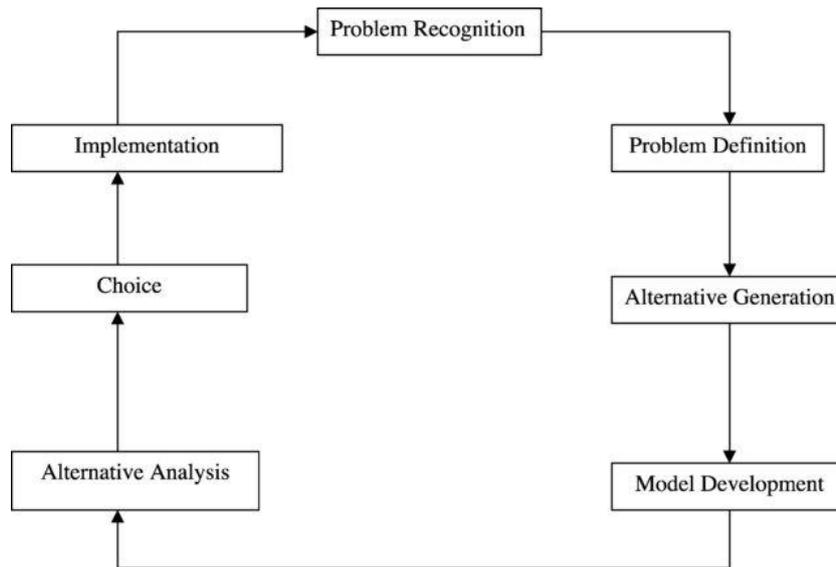


Figure 19. Le processus de prise de décision [Shim et al., 2002]

Pour [Shim et al., 2002], les premières étapes traitent l'identification du problème. Face à chaque problème, logiquement il existe une ou plusieurs méthodes d'élimination ou d'évitement de l'apparition. Pour proposer ces différentes alternatives il est important de développer des modèles bien précis qui définissent les relations entre les problèmes et les solutions proposées ainsi que l'ensemble des paramètres complémentaires et les règles métiers. Ainsi, ceci facilite la phase d'analyse alternative : il suffit de choisir la solution ou la décision adéquate pour pouvoir l'implémenter dans la dernière étape.

II.6.2. Les Différentes catégories de SAD :

Avec les évolutions technologiques et le développement des recherches sur les SAD, ainsi que de leurs domaines d'application, cette discipline a beaucoup évolué et plusieurs approches et sous-domaine ont apparues au fur et à mesure. [Arnott et Pervan, 2005] traitent l'histoire des SAD et définissent les principaux sous-systèmes d'aide à la décision spécifiques et leurs évolutions au cours du temps. Ils sont représentés dans la figure 20 :

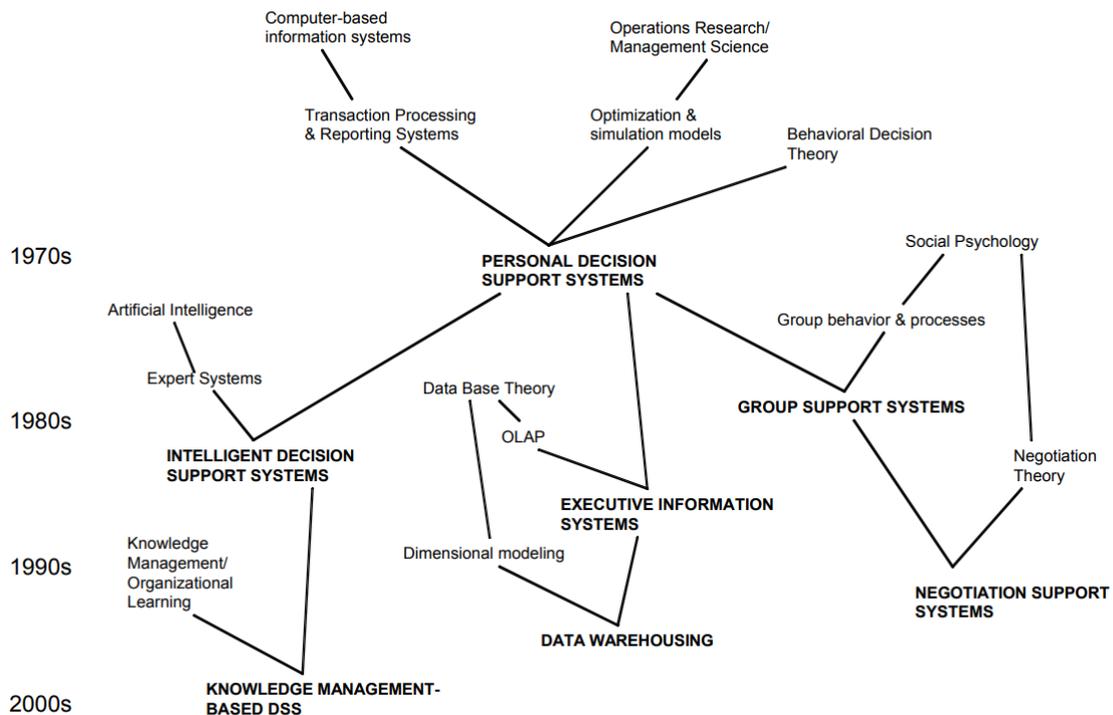


Figure 20. L'évolution des SAD en fonction des domaines d'utilisation [Arnott et Pervan, 2008]

Parmi ces SAD, il est possible de différencier :

- *Les SAD personnels* : comme indique leur nom sont souvent des systèmes de décision personnels développés pour un seul manager ou un groupe de manager indépendants ;
- *Les SAD du groupe* : sont développés pour améliorer le travail et la collaboration entre un groupe de travail. L'objectif principal de ce système est de favoriser une prise de décision basée sur l'ensemble des suggestions proposées par les différents membres du groupe ;
- *Les SAD de négociation* : ce type de système utilise les technologies informatiques pour faciliter la négociation. Ils existent deux catégories de systèmes de négociation : les systèmes orientés problèmes et ceux orientés *process* ;
- *Les SAD intelligents* : cette catégorie de système introduit l'utilisation de l'intelligence artificielle dans l'aide à la décision. Ils sont classés en deux générations : les systèmes à base de règles et la deuxième génération regroupe les algorithmes génétiques, la logique floue, les réseaux de neurones ;
- *Les SAD basés sur la gestion des connaissances* : ces systèmes permettent la prise de décision en se basant sur la capitalisation des connaissances et le raisonnement à base de connaissances. La manipulation, le transfert et la réutilisation des connaissances, pour un objectif de prise de décision permet la création de nouvelle connaissance.

Ces travaux de recherches s'intègrent, essentiellement, dans cette dernière catégorie. Ainsi, une analyse bibliographique sur ces systèmes a été établie.

II.6.3. Les systèmes d'aide à la décision à base de connaissances (SADBC)

Afin de comprendre l'utilité et le fonctionnement des SADBC, il est important d'étudier les principes de raisonnement et particulièrement ceux à base de connaissances.

II.6.3.1. Le raisonnement à base des connaissances (KR)

Historiquement, la science de formalisation du raisonnement a commencé en Grèce antique, lorsque le philosophe Aristote a proposé le syllogisme qui est à la base du raisonnement déductif moderne. Du simple calcul, qui construit les concepts de fonction et d'application et définit les ordinateurs, aux diverses plates-formes numériques intelligentes, la notion de raisonnement est indispensable à tout système de calcul. Nos travaux de recherche s'intéressant à l'étude des concepts de base du KR, nous présentons dans cette partie les différentes définitions synthétisées à partir de la littérature.

[Zhang et Zhang,1992] ont souligné que le raisonnement est le processus d'analyse, de synthèse et de prise de décisions sur diverses notions, depuis la collecte des faits existants, la découverte des interconnexions entre les objets, jusqu'au développement de nouvelles connaissances. En bref, le raisonnement est le processus qui consiste à tirer des conclusions à partir de faits existants en appliquant des règles. Les travaux sur ces thématiques ont commencé dès les années 1960, dans les domaines de la logique (la logique du premier ordre et la logique des prédicats) et de l'ingénierie de la connaissance (des réseaux sémantiques et des modèles graphiques de connaissances). À l'époque, plusieurs chercheurs ont proposé des méthodes pour améliorer la rigidité du processus de raisonnement, telles que le raisonnement flou [Zadeh, 1965] et le raisonnement non monotone [McCarthy, 1980].

Les recherches continuent et d'autres définitions apparaissent. [Tari, 2013] a défini le concept de KR comme le mécanisme qui permet d'acquérir de nouvelles connaissances basées sur les faits et les règles logiques existants. En général, le KR est le processus qui consiste à utiliser des connaissances connues pour en déduire de nouvelles [Chen et al., 2020].

Avec l'apparition des nouvelles technologies et la croissance des flux de données circulant sur Internet, les méthodes traditionnelles de gestion des bases de connaissances (KB) deviennent limitées et ne permettent pas d'assurer une exploitation maximale des connaissances capitalisées. C'est pourquoi les méthodes de KR restent toujours parmi les sujets de recherche

les plus actifs dans le domaine des sciences cognitives et précisément en gestion et ingénierie des connaissances. Elles spécialement trop utilisée pour le développement des systèmes d'aide à la décision. Les SADBC sont utilisés dans plusieurs domaines et principalement celui de la production, y compris la conception, l'approvisionnement et la logistique, la planification et l'ordonnancement, la fabrication et l'optimisation des processus.

II.6.3.2. Les SADBC pour les industries de fabrication

Ces systèmes sont très efficaces pour les industries manufacturières pour plusieurs raisons : Premièrement, les activités manufacturières sont très importantes dans tous les domaines pour fournir aux clients des produits et des services de haute qualité et qui répondent à leurs exigences, dans les délais et à des coûts compétitifs. Aussi, aujourd'hui une énorme quantité de connaissances hétérogènes et de catégories différentes, est disponible en industrie. Par conséquent, il y a une nécessité de fournir les meilleurs outils et pratiques permettant la capitalisation, l'apprentissage, le partage et la réutilisation des connaissances pour supporter le travail des experts et pour participer à l'amélioration continue des processus industriels. Finalement, avec les aspects de digitalisation et les enjeux de l'industrie 4.0. La prise de décision devient de plus en plus complexe. Plusieurs types de décisions stratégiques et opérationnels sont à prendre dans une industrie de fabrication, à savoir le choix des fournisseurs, la conception des produits, le changement d'un processus, l'achat de matériels, la maintenance des produits, l'investissement, etc. Pour prendre une bonne décision, il faut pouvoir compter sur l'appui des connaissances de toute la chaîne de valeur. De nombreux travaux de recherche ont adopté les techniques de systèmes d'aide à la décision à base de connaissances pour améliorer la productivité. [Harding et al., 2006] présente une étude sur les systèmes de fabrication intelligents basé sur l'extraction des connaissances à partir des données. [Zhang et al., 2011] ont développé un système d'aide à la décision multi-rôle basé sur plusieurs catégories de connaissances pour l'optimisation des processus de fabrication.

En générale, les SADBC utilisent les différentes connaissances disponibles pour établir un raisonnement et générer de nouvelles connaissances sous forme de décision. Nous distinguons plusieurs technologies de raisonnement ou d'inférence, par exemple [Liu et Zaraté, 2014] différencient : les systèmes de raisonnement à base de règles RBR, les systèmes de raisonnement à partir de cas RàPC, les systèmes de raisonnement narratif, et algorithmes génétiques. [Ben Rabah et al., 2017] quant à eux distinguent trois méthodes de diagnostic industriel : le diagnostic basé sur le RBR, le diagnostic basé sur les modèles et le diagnostic basé sur le RàPC. Toutes ces techniques sont souvent combinées avec les solutions ontologiques

pour la gestion des connaissances. Ces méthodes seront expliquées dans cette partie et des exemples de la combinaison avec les ontologies seront justifiées avec des exemples pratiques tirés de la littérature.

II.6.3.3. Les systèmes de RBR

Définitions

L'intérêt récent pour l'intelligence artificielle et les systèmes experts a attiré l'attention sur la représentation des connaissances fondée sur des règles. Il faut mettre davantage l'accent sur l'interaction homme-machine, qui permet à l'utilisateur d'acquérir les connaissances nécessaires à la prise de décision. Dès les premières utilisations de ces systèmes dans des applications pratiques, il a été démontré que la représentation des connaissances basée sur des règles peut conduire à des techniques plus efficaces d'analyse et d'aide à la décision [Bohanec et al., 1983; Efstathiou et Rajkovic, 1979]. Les RBR permettent de représenter des connaissances heuristiques, profondes ou stratégiques, en utilisant la logique des propositions, des prédicats ou d'ordre 2. Cette technologie a été utilisée dans la mise en œuvre de plusieurs systèmes [Armengol, 2011 ; Wakulicz-Déjà et al., 2011 ; Kong et al., 2012 ; [Zhang, 2011 ; Rajkovic et al., 1987] définissent les règles comme une méthode naturelle de communication. Elles sont utilisées pour la représentation des connaissances permettant de faciliter l'interaction homme-machine et homme-homme en termes d'acquisition, d'apprentissage, de vérification, et de traitement des connaissances.

Plusieurs avantages justifient l'utilité des systèmes RBR. Premièrement, ils représentent les systèmes à base de connaissances les plus utilisés dans le développement des applications et des systèmes d'aide à la décision. De plus, plusieurs éditeurs de règles sont disponibles, adaptés aux variétés des besoins et des technologies utilisées. Toute l'efficacité des RBR réside dans le module d'inférence intégrée dans ce système permettant d'établir le raisonnement à prendre des décisions. Selon [Paquette, 2014] un système de RBR est composé par les éléments représentés dans la figure suivante :

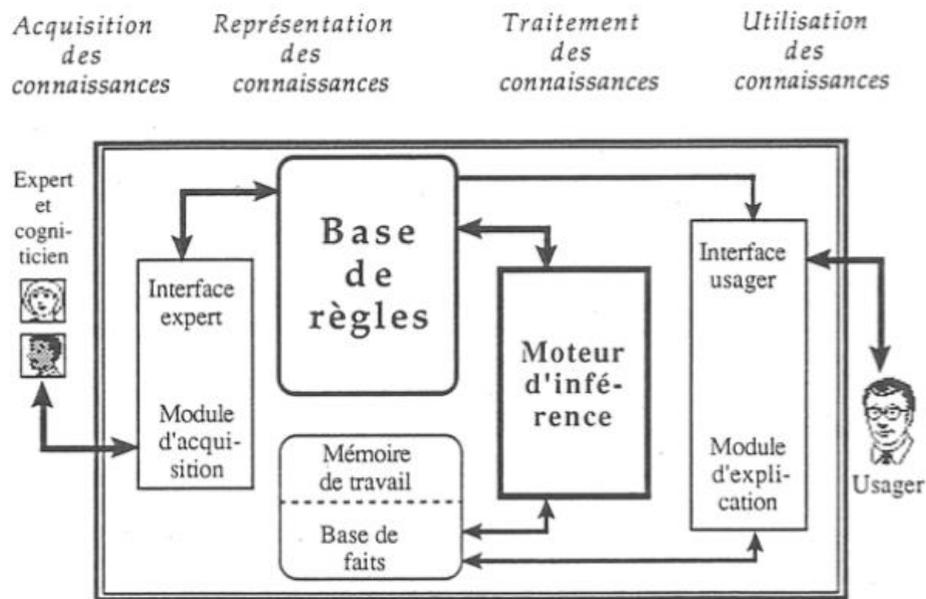


Figure 21. Architecture d'un système à base de règles [Paquette, 2014]

Les composants de base sont ainsi les connaissances relatives à un domaine spécifique et les faits qui représentent les données spécifiques relatives à un problème donné. Le deuxième composant primordial est la base de règles qui représente et formalise le savoir-faire des experts métiers. Le moteur d'inférence permet de faire des déductions et proposer des décisions en se basant sur la base de connaissances. La mémoire de travail représente quant à elle un emplacement intermédiaire permettant de stocker des déductions du moteur d'inférence les conditions nécessaires pour prendre des décisions, les messages échangées, etc. Enfin, les interfaces peuvent être exploitées en entrée, pour configurer le système et distinguer le besoin de l'opérateur, ou encore pour apprendre le système et intégrer des nouvelles connaissances et règles de la part de l'expert. Elles communiquent aussi en sortie pour afficher les résultats de l'analyse. La représentation des connaissances sous forme de règles est une méthode de formalisation assez simple, prenant souvent la forme d'une expression conditionnelle « Si – alors », et plusieurs méthodes de représentation existent.

Les règles logiques de prédicats du premier ordre

Cette méthode utilise les propositions comme unité de base de raisonnement. Les propositions contenant des individus et des prédictions. Les individus qui peuvent exister indépendamment correspondent à des objets entités dans la base de connaissances. Ils peuvent être une entité concrète ou un concept abstrait. Le prédicat est utilisé pour décrire la nature et les éléments de l'individu. Par exemple, la modélisation des relations interpersonnelles peut être réalisée en

utilisant une logique de prédicats du premier ordre. On considère alors les relations comme des prédicats, les personnages comme des variables, et on utilise des opérateurs logiques pour exprimer les relations interpersonnelles, en fixant la logique et les contraintes du raisonnement relationnel pour effectuer un premier type de KR simple. Par exemple, Si « Père » est le père de « Enfant1 » et si « Enfant 1 » et « Enfant 2 » sont frères alors « Père » est le père de « Enfant2 ». Cette règle se formalise de la manière suivante :

$$(Pere, EstPèreDe, Enfant1) \wedge (Enfant1, EstFrèreDe, Enfant2) \Rightarrow (Pere, EstPèreDe, Enfant2)$$

Plusieurs travaux de recherche ont été développés autour de ces thématiques : [Schoenmackers et al., 2010 ; Galárraga et al., 2013 ; Demeester et al., 2016]. Parmi les plus intéressants, les travaux de [Ozbayrak et Bell, 2003] utilisent un SAD basé sur des règles pour gérer les composants et les outils d'une chaîne de production. Ils proposent trois modèles basés sur la connaissance pour faciliter le processus de prise de décision : un système expert de planification de la production, un système d'aide à la décision basé sur la connaissance pour la gestion des outils et un système de diagnostic des erreurs de gestion des outils. L'ensemble du système a été créé de manière hiérarchique et comprend plus de 400 règles. La figure 22 présente la structure du SAD à base de règles pour la gestion des pièces et des outils dans les systèmes de fabrication flexibles

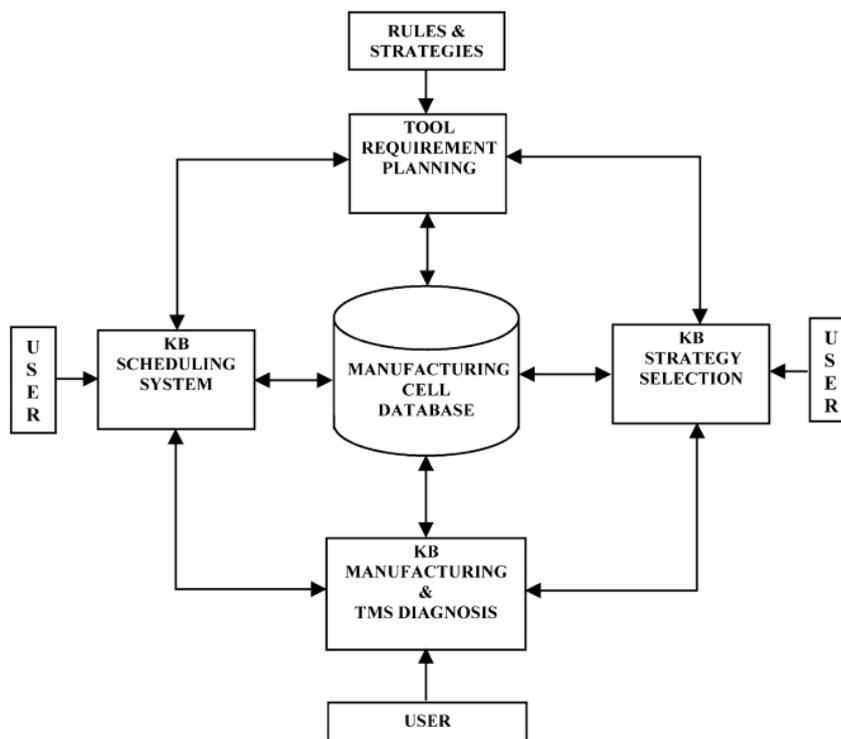


Figure 22. Structure du système d'aide à la décision basé sur les connaissances

Ces méthodes ont été aussi adaptées au profit du web sémantique et de l'inférence. Certains chercheurs proposent un langage tractable : la logique de description (*Description Logic*, DL).

La logique de description

La logique de description est un fondement essentiel de la recherche sur les ontologies, développée sur la base de la logique propositionnelle et de la logique des prédicats du premier ordre. L'objectif de la logique de description est d'équilibrer le pouvoir de représentation et la complexité du raisonnement. Elle peut fournir une sémantique bien définie et des outils de raisonnement puissants pour les graphes de connaissances et satisfaire les besoins de la construction, de l'intégration et de l'évolution de l'ontologie. Il s'agit donc d'un langage ontologique idéal [Berners-lee et al., 2001].

Ceci nous emmène à étudier une deuxième sous-catégorie : le RBR basé sur la combinaison entre les ontologies et les règles. Une ontologie (définie en annexes) effectue le raisonnement pour inférer une conclusion logique en utilisant un ensemble de faits et d'axiomes affirmés. Le modèle d'ontologie basé sur OWL-DL est composé d'un ensemble de triplets, chaque triplet étant représenté par (a,R,b) ou $R(a,b)$, où a et b désignent deux individus liés à la propriété R . Ainsi, le processus de raisonnement des relations implicites peut être décrit par le modèle mathématique défini par l'équation :

$$P_1(a_1, b_1) \wedge P_2(a_2, b_2) \wedge \dots \wedge P_N(a_N, b_N) \rightarrow \exists x, y R_1(x, y)$$

$$P_i \in FactsBase ; R \in InferrBase ; a_i, b_i, x, y \in OntIndividuls$$

Dans l'équation, $R_1(x,y)$ est le résultat d'un raisonnement. *OntIndividuls* représente l'ensemble des individus. *FactsBase* est un ensemble de tous les individus et propriétés explicites, tandis que *InferrBase* désigne l'ensemble des propriétés qui ont été déduites. La première moitié de l'équation (la combinaison des P_i) représente une contrainte. Dans chaque contrainte, toutes les prémisses doivent être vraies pour garantir le résultat correct de l'inférence.

En détails, le mécanisme de raisonnement peut être représenté par les étapes suivantes [Qiao et al., 2017] :

- Étape 1 : instancier tous les individus explicites et les propriétés entre eux, puisqu'ils constituent les faits pour le raisonnement ;
- Étape 2 : définir les règles du web sémantique sous forme d'équations de contraintes, qui peuvent être interprétées avec le SWRL. Ce langage peut être reconnu par un moteur de raisonnement basé sur des règles comme Pellet [Sirin et al., 2007]. Sur la base des

règles définies dans le SWRL, le moteur peut déduire les relations implicites entre les individus, créant ainsi de nouveaux triplets pour générer les nouvelles connaissances inférées telles que définies dans l'équation. Par exemple, une règle dans SWRL définit que s'il existe une relation entre deux individus Ind1 et Ind2, et une autre relation entre les individus Ind2 et Ind3, un nouveau triplet sera formé en déduisant une relation entre Ind1 et Ind3 ;

- Etape 3 : Interroger le résultat en définissant une fonction objective comme indiqué dans l'équation. Ensuite le langage SQWRL pourrait être utilisé pour interroger les triplets qui satisfont à la fonction objective.

Ainsi, il est clair que le raisonnement avec les connaissances permet de générer et de capitaliser de nouvelles connaissances. La capitalisation de données et de connaissances ne permet pas forcément une exploitation directe de celles-ci, il est crucial d'en faire une modélisation pour pouvoir conceptualiser le système à base de connaissances envisagé. Cette modélisation permettra de distinguer les concepts et objets de base du système et créer les différentes relations internes ce qui facilitera par la suite le raisonnement et la génération de nouvelles connaissances.

II.6.3.4. Les systèmes de RàPC

La technologie de raisonnement à partir de cas est une méthodologie de résolution de problèmes, qui s'appuie sur des cas passés et similaires pour trouver des solutions à de nouveaux problèmes. Elle permet de tracer les leçons passées et de mettre à disposition les meilleures pratiques sous forme de cas [Ben Rabah et al., 2017].

Le cas peut être défini comme un élément de connaissance contextualisé représentant une expérience qui enseigne une leçon fondamentale pour atteindre les objectifs du raisonneur [Kolodner, 1992]. Selon [Napoli et al., 1996], un cas est un ensemble formé par le problème et sa solution, et il peut être structuré de différentes manières. Les anciens cas, stockés dans la base de cas, sont appelés cas sources, tandis que chaque problème résolu est appelé cas cible. Le cycle du RàPC se compose de quatre étapes essentielles [Aamodt et Plaza, 1994] :

- **Retrouver** les cas sources les plus similaires au problème traité. La comparaison se fait sur la base de plusieurs mesures de similarité, la plus répandue dans la littérature est le calcul de la distance entre les deux cas source et cible ;
- **Réutiliser** la solution la plus satisfaisante et adaptée ;

- **Réviser** la solution proposée, en utilisant des méthodes d'évaluation spécifiques et l'adapter au besoin en utilisant des règles d'adaptation pour proposer la solution nécessaire pour résoudre le problème ;
- **Mémoriser** la nouvelle solution dans la base de règle pour des futures utilisations ;

De nombreux travaux de recherche utilisent le RàPC pour résoudre des problèmes dans les systèmes industriels. Une analyse bibliographique a permis d'étudier certains systèmes qui visent à résoudre des problèmes de diagnostic, de maintenance ou de reconfiguration tout en représentant la phase de diagnostic dans leurs systèmes. L'efficacité d'un système de RàPC est fortement liée à la qualité de représentation du cas. En effet, elle peut influencer toutes les autres phases du raisonnement. Plusieurs travaux de recherche appuient la représentation de cas avec l'intégration des ontologies dans le système.

[Dendani et al., 2012] proposent une approche qui comprend l'utilisation d'ontologies pour construire des modèles de connaissance générale du domaine. Bien que dans un système de RàPC la principale source de connaissances soit l'ensemble expériences précédentes, ces travaux proposent des applications intégrées qui combinent des connaissances spécifiques à chaque cas avec des modèles des connaissances du domaine. Les recours à la capitalisation et à la modélisation des connaissances du domaine permettent aux processus de raisonnement d'obtenir des résultats plus précis. La figure 23 présente l'architecture proposée, composée de trois composantes fonctionnelles : l'ontologie de domaine, l'application RàPC et le médiateur.

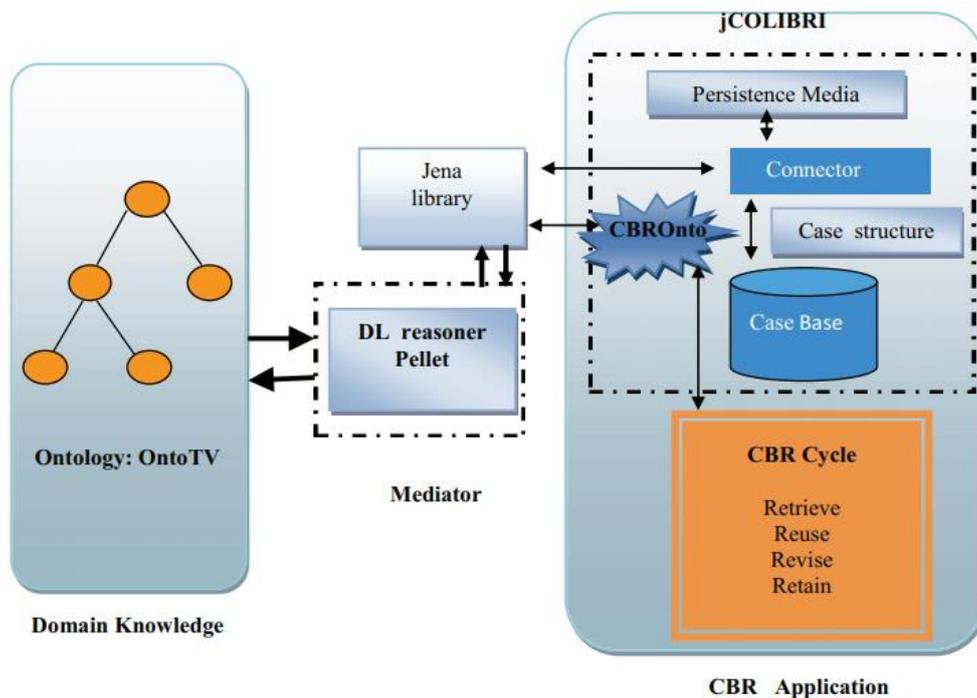


Figure 23. Architecture du système de RàPC à base d'ontologie [Dendani et al., 2012]

[Adla, 2018] propose un système de RàPC qui est composé par deux parties principales. La première est standard, le système utilisant une base de cas pour retrouver des cas similaires au problème à résoudre. La deuxième partie est utilisée lorsque le processus de récupération échoue ou que les cas récupérés ne sont pas satisfaisants pour le décideur, le système utilise des ontologies. Il utilise les relations sémantiques entre les concepts pour trouver d'autres solutions au problème. En utilisant l'ontologie, le système dérive des décisions plus spécifiques ou plus générales que celles initialement récupérées par le système. Il peut également définir la solution relative à l'équipement en visualisant l'élément concerné.

De même, d'autres descripteurs de cas peuvent être utilisés comme entrées dans l'ontologie pour agrandir ou réduire l'espace de la solution. Lorsqu'une solution est retenue puis testée et validée, elle est stockée dans la base de cas en tant que nouveau cas (avec tous ses descripteurs). Le processus de raisonnement est représenté dans la figure 24.

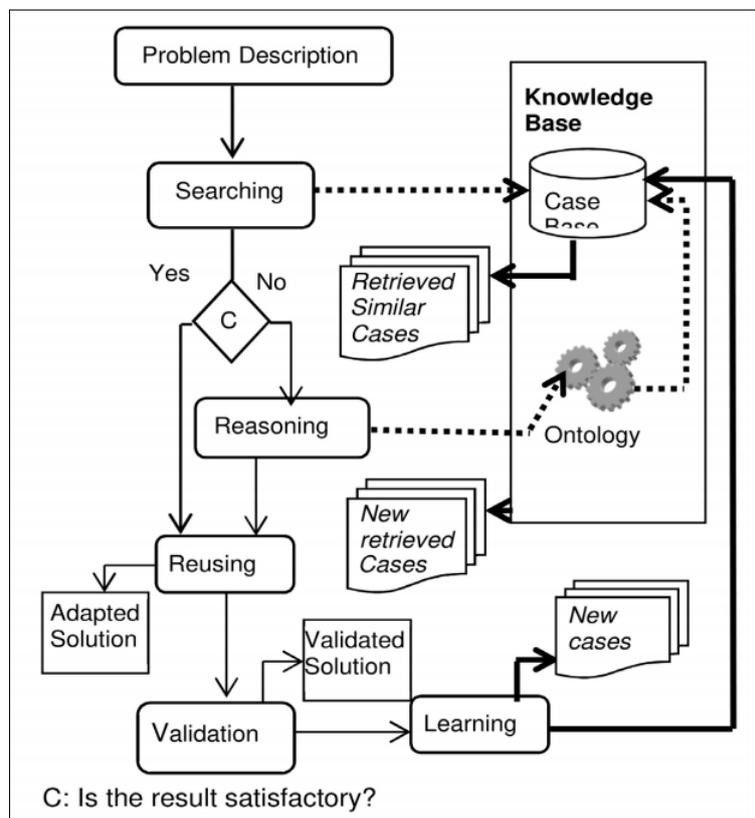


Figure 24. Processus de RàPC à base d'ontologie [Adla, 2018]

II.6.4. L'aide à la décision pour le diagnostic

Plusieurs aspects caractérisent tout système de production, essentiellement la fiabilité, la disponibilité, la maintenabilité et la sécurité. Ces contraintes doivent être prises en compte car elles sont pertinentes pour assurer l'engagement et la responsabilité des fabricants et en

contrepartie la satisfaction des clients. Plusieurs techniques ont été développées pour répondre aux exigences de sécurité et de fiabilité, issues essentiellement des disciplines de recherche traitant de l'intégrité du système et de la sûreté de fonctionnement du système. Parmi ces techniques figure le diagnostic des défaillances. Cette solution s'intègre dans le cadre de développement de systèmes d'aide à la décision pour supporter le travail des industriels dans l'analyse et la compréhension des phénomènes malveillants qui peuvent apparaître pendant la fabrication. La section suivante se focalise sur les systèmes d'aide à la décision pour le diagnostic. Avant de décrire ces solutions il est important de comprendre les distinctions terminologiques et les notions de défaillances qui peuvent apparaître dans la fabrication.

II.6.4.1. Terminologie de la défaillance

Vu la variété des domaines industriels, les terminologies relatives à ces problèmes sont très variables. Le comité technique de l'IFAC est néanmoins parvenu à des définitions et des terminologies standardisées. On distingue ainsi les notions suivantes :

- Un défaut est une déviation non autorisée d'au moins une propriété caractéristique de son état acceptable, habituel et standard. L'écart non autorisé est la différence entre la valeur du défaut et le seuil de dépassement d'une zone de tolérance pour sa valeur habituelle. Un défaut est un état dans le système ;
- Une défaillance est une interruption permanente de la capacité d'un système à exécuter une fonction requise dans des conditions de fonctionnement spécifiées. Une défaillance est un événement, qui résulte d'une ou de plusieurs défauts, survenant généralement après le début de l'exploitation ou par une sollicitation croissante du système ;
- Un dysfonctionnement est une irrégularité intermittente dans l'accomplissement de la fonction souhaitée d'un système. Un dysfonctionnement est une interruption temporaire, est considéré comme un événement, résulte d'une ou plusieurs défaillances.

Les défauts présentent généralement un comportement caractéristique pour les différents composants. Ils peuvent être distingués par leur forme, leur comportement temporel ou leur étendue. La forme peut être systématique ou aléatoire. Le comportement temporel peut être décrit par un comportement permanent, transitoire, intermittent, par le bruit ou la dérive. L'étendue des défauts est soit locale, soit globale et comprend la taille.

Le matériel électronique présente des défauts dits systématiques lorsqu'ils proviennent d'erreurs de spécification ou de conception. Une fois en fonctionnement, les défauts des composants matériels sont le plus souvent aléatoires, avec toutes sortes de comportements temporels. Les

défauts ou erreurs dans les logiciels (bugs) sont généralement systématiques, par exemple à cause d'une mauvaise spécification, d'un mauvais codage, d'une mauvaise logique, de débordements de calculs, etc. Ils ne sont généralement pas aléatoires contrairement aux défauts du matériel. Les défaillances des systèmes mécaniques peuvent être classées selon les mécanismes de défaillance suivants : distorsion (flambage, déformation), fatigue et fracture (fatigue cyclique, fatigue thermique), usure (abrasive, adhésive, cavitation) ou corrosion (galvanique, chimique, biologique), voir par exemple [Morris et al. 1995 ; Brown 2007].

L'amélioration de la productivité et la satisfaction de la clientèle sont les premières préoccupations de n'importe quel système industriel. Or, l'occurrence des défaillances au moment de la fabrication peut freiner le rendement de l'industrie. Par conséquent, le recours à l'analyse et la compréhension de l'origine des défaillances est une technique très répandue dans ce domaine et ceci à travers la mise en place des processus de diagnostic.

II.6.4.2. Le Diagnostic des défaillances

Il existe plusieurs définitions de la fonction de diagnostic. Par exemple, [Combacau et al., 2000] a défini le diagnostic comme trois entités :

- L'emplacement : déterminant le sous-système responsable de la défaillance ;
- L'identification : précisant les causes de cette défaillance ;
- L'explication : qui justifie les conclusions du diagnostic.

D'autres travaux, comme [Darkhovski and Staroswiecki, 2003 ; Phillippot et al., 2006], considèrent la détection comme une partie à part entière du processus, il s'agit des systèmes de détection et de diagnostic des défaillances (ou défaut) ou « DDD ». Les auteurs le définissent comme un processus de décision impliquant trois entités :

- La détection qui permet de détecter toute déviation du comportement du processus par rapport à son comportement normal et d'alerter en cas de défaillance ;
- La localisation qui permet de spécifier l'origine de la défaillance (zone et composant affecté) ;
- L'identification qui permet de déterminer l'heure d'apparition de la défaillance, sa durée et sa gravité.

Les méthodes de diagnostic des défaillances

Les méthodes de diagnostic des défaillances visent à trouver les causes profondes des défaillances qui apparaissent dans un système. Pour minimiser les retards de production dûs

aux impacts des défauts, le diagnostic doit être rapide et efficace. Diverses méthodes ont été développées pour soutenir le diagnostic des défauts. La première classification est celle de [Abele and Weyrich, 2017] :

- Les méthodes basées sur des modèles qui comparent le comportement du système avec un système modélisé et peuvent évaluer les erreurs dues aux écarts entre les deux ;
- Les méthodes basées sur les signaux qui analysent les données des capteurs de surveillance pour trouver des symptômes de défaut dans les courbes de signaux anormaux ;
- Les méthodes basées sur les défauts ou sur les historiques qui comparent l'état actuel du système avec des échecs dans le passé pour empêcher les défaillances de se produire à nouveau.

Les méthodes basées sur les signaux et les défauts deviennent de plus en plus utilisées face aux problématiques du *big data* et grâce aux nouvelles infrastructures informatiques disponibles, qui permettent une intégration facile de ces méthodes [Patwardhan et al., 2016]. D'autres études plus détaillées ont été établies. [Isermann, 2006] a par exemple étudié les méthodes de diagnostic des défaillances et a proposé deux catégories globales : les méthodes de classification et d'inférence. La figure 25 énumère la majorité des méthodes détaillées dans la littérature.

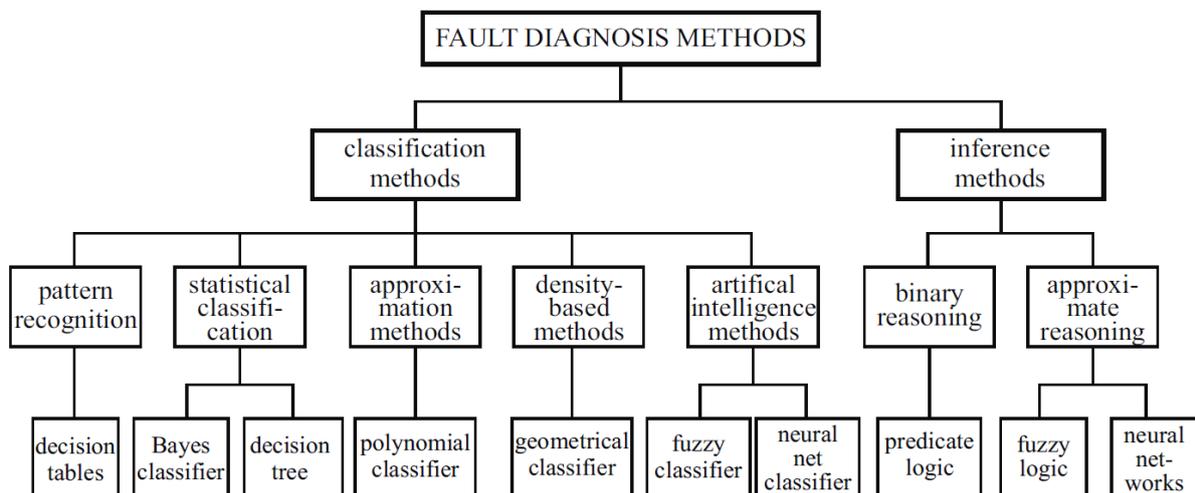


Figure 25. Les méthodes de diagnostic des défauts [Isermann, 2006]

Les problématiques de ces travaux de recherche se situent essentiellement au niveau de la deuxième catégorie, celle des méthodes d'inférence. Plusieurs techniques sont utilisées pour implémenter ces méthodes dans des SAD.

Techniques de développement des SAD pour le DDD

Plusieurs méthodes sont utilisées pour implémenter des SAD pour le diagnostic. Elles correspondent, essentiellement, aux différentes catégories détaillées précédemment dans la partie SADB. Les mêmes peuvent donc être orientées pour la réalisation d'un diagnostic :

- Le diagnostic basé sur un modèle est une technique de résolution de problèmes basée sur une description du système réel à diagnostiquer (un modèle). Il est utilisé pour détecter et localiser les défaillances multiples. Cependant, la création d'un modèle est un processus pointu et long, aussi un changement dans le système implique-t-il une mise à jour du modèle. La méthode n'a en effet pas de capacité d'apprentissage automatique [Darkhovski and Staroswiecki, 2003 ; Phillippot et al., 2006] ;
- Le diagnostic basé sur les ontologies a pour objectif de représenter les connaissances sous une forme interprétable par une machine. Le partage et la réutilisation des connaissances ainsi que l'interopérabilité sémantique des systèmes hétérogènes sont les avantages de cette technique. Pour modéliser un domaine spécifique par le biais d'une ontologie, il faut analyser les entités pertinentes. L'intégration des règles d'adaptation aux ontologies conduit à une base de connaissances et l'affirmation d'un moteur d'inférence produit un système expert [Uddin et al., 2011] ;
- Le diagnostic basé sur des règles nécessite la traduction de l'opération de diagnostic en expressions booléennes. Il est très adapté et facile à appliquer dans les systèmes industriels. Cependant, des efforts d'extraction du savoir-faire des experts métier et de formalisation de leur expertise et leurs connaissances internes sous forme règles sont nécessaires. En outre, c'est une méthode qui n'a pas de capacité d'apprentissage (elle ne peut pas gérer un échec non prévu dans sa base de règles) [Saddem et Phillippot, 2014] ;
- Le diagnostic à base d'un RàPC est une méthode utilisée pour résoudre de nouveaux problèmes de diagnostic en utilisant des expériences antérieures stockées dans une base de cas. L'avantage de cette technique par rapport aux autres est sa capacité d'apprentissage et de réintégration [Xiong et al., 2012].

[Katipamula et Brambley, 2005] précisent que l'objectif premier d'un système DDD est la détection précoce des défauts et le diagnostic de leurs causes, ce qui permet de corriger les défauts avant que des dommages supplémentaires ne soient causés au système ou que le service ne soit interrompu. Pour ce faire, on surveille en permanence le fonctionnement d'un système, on utilise DDD pour la détection et le diagnostic des conditions anormales et les défauts qui y sont associés puis on évalue l'importance des défauts détectés et on décide de la manière d'y

répondre. Une opération typique de l'utilisation d'un processus de DDD est le suivi de la fabrication et/ou de la maintenance. Comme montre la figure 26, d'un point de vue pratique, ces systèmes automatisés possèdent quatre fonctions distinctes. La première étape consiste à surveiller le système ou le dispositif physique et détecter toute situation anormale (ou problème). Cette étape est généralement appelée « détection des défauts ». Lorsqu'une situation anormale est détectée, le diagnostic de panne est utilisé pour analyser et déterminer ses causes. Ces deux étapes constituent le processus DDD. Après cela, l'évaluation de la défaillance permet d'évaluer l'ampleur et l'importance de l'impact sur la performance du système (en termes de consommation d'énergie, de coût, de disponibilité ou d'effets sur d'autres indicateurs de performance). Sur la base de l'évaluation de la défaillance, une décision est alors prise pour spécifier la manière de réagir (par exemple, en prenant une mesure corrective ou même en n'en prenant aucune).

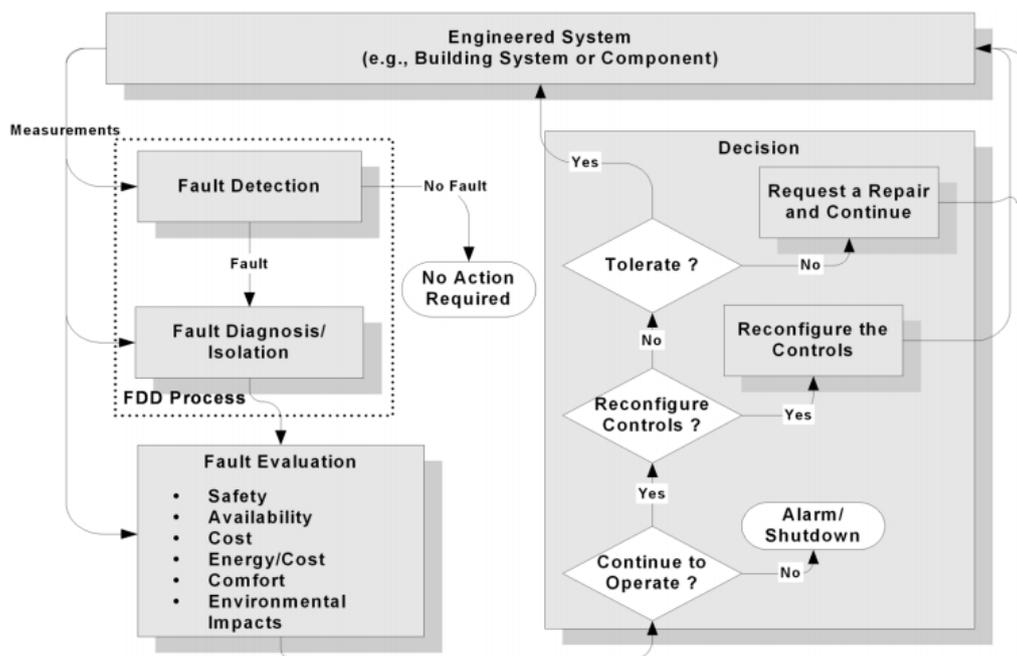


Figure 26. Application de la DDD dans l'ingénierie [Katipamula et Brambley, 2005]

La détection et le diagnostic des défauts, en général, sont basés sur des variables mesurées par des instruments, et des variables et des événements observés par des opérateurs humains. Le traitement automatique des variables mesurées pour la détection des défauts nécessite une connaissance analytique du processus [Struss et al., 1996], et l'évaluation des variables observées nécessite une connaissance experte humaine, appelée connaissance heuristique. La détection et le diagnostic des défauts peuvent donc être considérés dans le cadre d'une approche

basée sur la connaissance [Rasmussen,1993]. La figure 27 montre un schéma global d'un système DDD à base de connaissances [Isermann,2006].

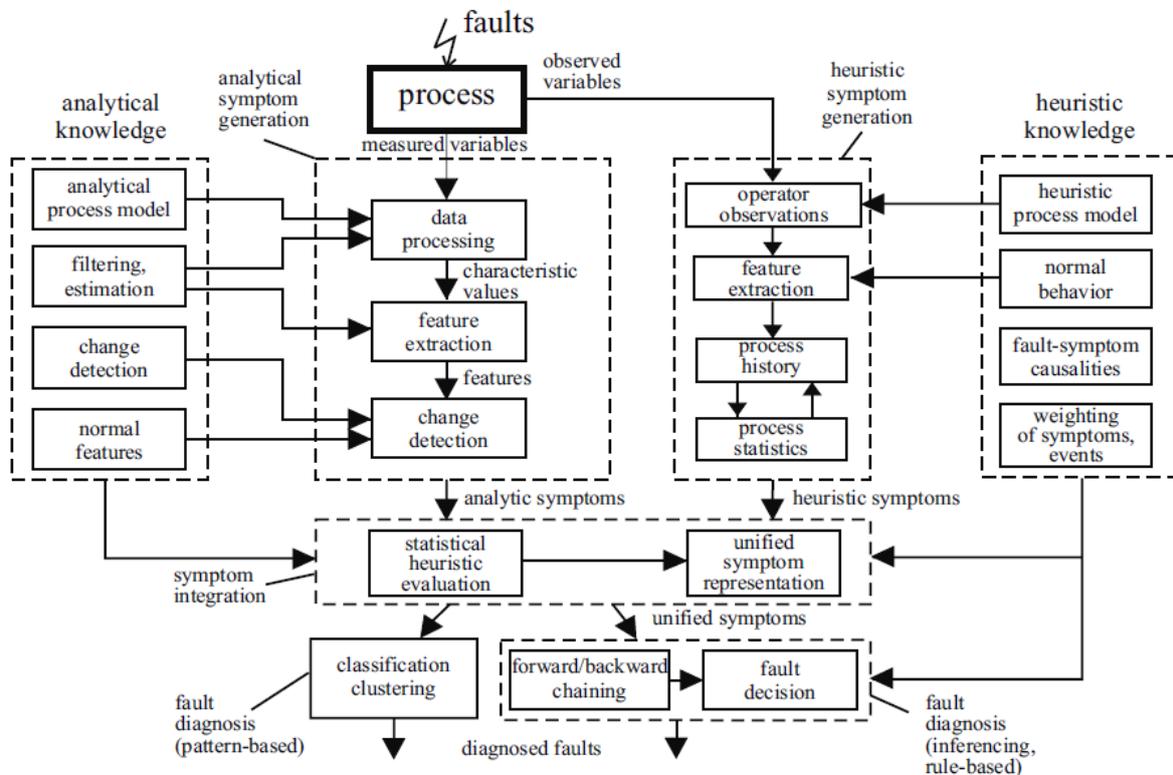


Figure 27. Schéma global de DDD basé sur la connaissance [Isermann, 1994 ; Isermann, 1997]

La troisième étape d'évaluation de la défaillance ou du défaut permet de vérifier plusieurs propriétés relatives au bon fonctionnement du système. Les contraintes les plus critiques, surtout dans un système industriel, sont la fiabilité et la sécurité du système. Pour étudier l'effet des défauts sur ces propriétés une série de méthodes d'analyse ont été mises au point, telles que la technique d'analyse de fiabilité, les méthodes d'analyse avec l'arbre des événements et l'arbre de défaillance, la technique d'analyse et de classification des risques et la méthode d'analyse des modes de défaillance et de leurs effets [IEC 60812]. Cette dernière a été très inspirante pour la construction de notre approche détaillée dans le chapitre suivant.

La méthode d'analyse des modes de défaillance et de leurs effets « AMDE »

AMDE (FMEA en anglais) est une méthode formalisée permettant de prendre en compte tous les composants, leurs fonctions, les modes de défaillance et les causes des défaillances du système. La méthode a été développée en 1960 et utilisée par la NASA pour le projet Apollo [NASA, 1999], puis pour l'aérospatiale et les centrales nucléaires [Reichow, 1991]. Elle est

aujourd'hui une méthode standard, également utilisée dans l'industrie aéronautique [Johnson, 2010], automobile et médicale [Rezaie et al, 2007 ; Rezaie et al, 2008].

Elle consiste d'abord en l'énumération de tous les composants, de leurs modes de fonctionnement et de leurs modes de défaillance. Elle examine ensuite les causes possibles de chaque mode de défaillance et décrit leurs effets sur le module considéré et pour l'ensemble du système. Elle énumère ensuite les contre-mesures [Ebrahimipour et al. 2010]. Habituellement, seules les défaillances simples sont prises en compte. L'utilisation de fiches de travail AMDE constitue une méthode formalisée qui permet de détecter les points faibles de la conception au cours des premières étapes et du développement dans les dernières étapes. La force de l'AMDE réside dans son exhaustivité. Cependant, elle peut entraîner une procédure très longue.

L'analyse des modes de défaillance, des effets et de la criticité AMDEC est une extension de l'AMDE. L'importance de la défaillance de chaque composant est prise en compte en considérant sa probabilité d'occurrence et ses effets sur le système, c'est-à-dire en exprimant son risque sur le fonctionnement. Un indice de priorité du risque IPR peut être calculé pour montrer la criticité des différentes défaillances [Donier et al., 2017] :

$$IPR = F \times G \times D$$

Où {F : fréquence d'apparition de la défaillance ; G : gravité de la défaillance ; D : détectabilité de la défaillance}. Ces éléments sont souvent notés entre 1 et 10.

Il convient de noter que des défaillances peuvent être introduites après tout changement et mise à jour d'un produit ou d'un processus. Ainsi, l'AMDEC peut avoir besoin d'être examiné (et mis à jour) lorsque tout changement est apporté au système.

L'AMDEC étant un outil pour aider à assurer la fiabilité et la sécurité, il accorde certains avantages pour la gestion de projet. Cette technique est souvent utilisée dans le cadre d'un travail d'équipe, pour échanger des idées et acquérir des connaissances. L'objectif de son utilisation est de développer des systèmes de prévention vis-à-vis des problèmes, d'améliorer les essais et les développements, et d'assurer la satisfaction des clients. Plusieurs travaux de recherche ont utilisé cette méthode pour le développement de SADBC.

[Nazakat et Hoing, 2018] propose une approche pour détecter et prévenir les défaillances à base de connaissances. Elle comporte deux parties : la détection des défaillances et la prévention. Ils ont utilisé une architecture 5C pour prévoir et détecter la défaillance. Cette architecture fournit une ligne directrice systématique pour développer et déployer un CPS pour les applications de

fabrication intelligente [Lee et al., 2015 ; Wu et al., 2020]. Lorsque la défaillance est prédite ou détectée, les actions recommandées pour chaque défaillance sont déclenchées pour prévenir en fonction de la gravité du défaut. Dans le niveau cybernétique de l'architecture 5C, également appelé niveau d'auto-comparaison, les informations actuelles qui viennent d'être acquises par le capteur sont comparées à des cas similaires ou aux données historiques de ce capteur afin de prédire les performances du capteur et de prévoir sa fiabilité. Dans ce même niveau, des algorithmes analytiques complexes sont utilisés pour comparer les données. Il est ainsi possible de prévoir ou d'identifier les défaillances. Une fois que la défaillance est identifiée, le Framework, proposé pour la prévention et l'atténuation des défaillances, utilise la technique d'analyse des risques très répandue, l'AMDEC. Elle permet d'identifier les modes de défaillance afin d'évaluer le risque associé et par la suite identifier et mettre en œuvre des actions correctives afin de répondre aux préoccupations critiques. La figure 28 présente la structure du SADBC développé.

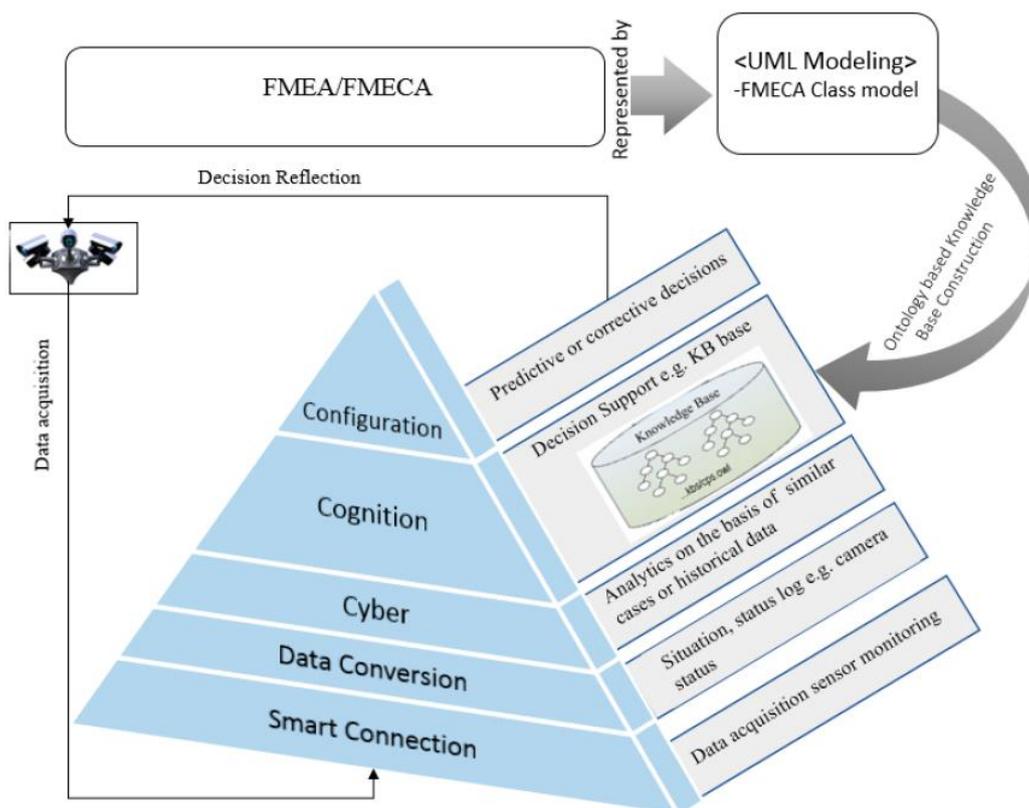


Figure 28. Structure d'un SADBC pour la détection et la prévention des défaillances

Dans la plupart des cas, la détection des défauts est généralement plus facile que l'évaluation des impacts découlant du défaut, qui à son tour est généralement plus facile que le diagnostic de la cause du défaut. Nous avons vu dans ce chapitre que plusieurs utilisent des techniques

différentes pour le diagnostic et surtout le recours aux systèmes d'aide à la décision à base de connaissances. En examinant les technologies et les méthodes mises en place pour le développement des SADBC, certains défis et limites ont été observés et plusieurs problématiques et questions de recherche ont été relevées.

II.7. LIMITATION ET QUESTIONS DE RECHERCHES

L'analyse bibliographique présentée dans ce chapitre est essentiellement liée au contexte de ces travaux de recherche, la plupart des solutions présentées sont ainsi sélectionnées pour leur utilité et efficacité pour la construction de notre approche. Il est évident que la liste n'est pas exhaustive et que plusieurs autres solutions intéressantes sont à explorer. Toutefois, cette étude bibliographique a montré un fort recours aux solutions ontologiques. Bien que cette technique représente un bon moyen de collecte et de structuration des connaissances, la construction d'une ontologie de taille moyenne dans le contexte de développement d'un SAD reste une tâche de longue haleine, qui nécessite une expertise et une maîtrise du domaine d'étude. Le défi réside dans la capacité d'acquisition d'un maximum de connaissances du domaine, tout en spécifiant les différentes relations sémantiques entre les concepts. Pour relever ce défi, plusieurs travaux de recherche ont travaillé sur l'émergence des techniques de l'apprentissage ontologique [Easton et al., 2011 ; Elnagar et al., 2020 ; Missikoff et al., 2002 ; Maedche et Staab, 2001]. L'apprentissage, la génération ou l'extraction de l'ontologie correspond à la création automatique ou semi-automatique d'ontologies, à travers l'extraction des concepts de base d'un domaine donnée, ainsi que les différentes relations entre ces concepts. Cette extraction se fait à partir de plusieurs sources de données mais essentiellement du texte en langage naturel. L'apprentissage ontologique comprend aussi la formalisation des connaissances avec un langage formel de développement. La construction manuelle d'ontologies représente une tâche assez complexe et longue, il y a ainsi une grande motivation à automatiser le processus de génération d'ontologies [Maedche et Staab, 2004]. Cette technique d'apprentissage, malgré son efficacité, n'est pas encore assez mature pour être adoptée directement. Les travaux de recherche continuent à travailler sur ces thématiques d'apprentissage pour proposer l'approche la plus fiable.

En contrepartie, les technologies de raisonnement individuel, telles que le raisonnement à base de règles et le raisonnement à partir des cas, ont mûri et ont été testées dans des applications réelles. Bien qu'ayant prouvé leur efficacité, ces technologies présentent encore plusieurs limites. Par exemple, pour un raisonnement à base de règles, plus la base augmente, plus le

temps d'inférence et de calcul augmente, on parle alors d'un problème d'extensibilité, ce qui freine les performances du système [Wakulicz-Déjà et al., 2011]. Toutefois, cet inconvénient peut être corrigé par une utilisation combinée du raisonnement fondé sur des règles et de la technologie de regroupement, c'est-à-dire le regroupement de règles similaires pour former des groupes de règles distincts, permettent de réduire considérablement le temps nécessaire à l'inférence. Outre la rapidité, la précision et la fiabilité ont été un enjeu important pour la plupart des technologies de raisonnement. Par exemple, la formalisation des règles métiers et l'externalisation du savoir-faire est une tâche très sensible, à cause de la forte dépendance du développement de ce système à la réactivité et la disponibilité de l'expert métier. De plus, une fois formalisées, ces règles devraient être validées par les experts pour vérifier la fiabilité de la phase de formalisation de ses connaissances internes, et la conformité de l'explicitation de ces nouvelles connaissances. La nécessité de vérifier les connaissances devient encore plus critique dans les domaines sensibles qui imposent beaucoup de normes à respecter comme l'aéronautique. Parfois une seule connaissance incorrecte ou imprécise pourrait entraîner une recommandation dangereuse ou erronée qui pourrait à son tour causer des défaillances. Ceci nous emmène vers une troisième limite pour les technologies de raisonnement est celle de l'étude de l'incertitude des connaissances utilisées, y compris les règles métiers et les cas tracés. Plusieurs travaux de recherche proposent de combiner ces technologies aux techniques de la logique floue et aux réseaux artificiels, pour améliorer les performances du raisonnement et réduire l'incertitude [Yang et al., 2011].

Enfin, les connaissances ne sont jamais statiques et une capitalisation des connaissances du domaine n'est jamais exhaustive. Avec l'émergence des nouvelles technologies, de nouvelles connaissances sont générées quotidiennement dans tous les domaines. Les SADBC à leurs tours génèrent des nouvelles connaissances suite au processus d'inférence. Ainsi est-il primordial de prendre en considération ces aspects d'évolution et de réintégration de nouvelles connaissances dans cette catégorie de système [Huang et al., 2011].

Face aux diverses limitations que présentent les différentes méthodes représentées plus haut, nous sommes convaincus qu'une collaboration entre les données et les connaissances permettra l'interprétation des données et facilitera la prise de décision dans le cadre de développement des SAD. Plusieurs questions de recherche ont été soulevées, la principale étant :

Comment développer un système d'aide à la décision efficace en ayant recours à une conciliation des données et des connaissances industrielles ?

Cette problématique générale intègre plusieurs autres questions de recherches spécifiques.

- *Comment recueillir et capitaliser les données et les connaissances ?*
- *Quels sont les moyens nécessaires pour mettre au service de l'utilisateur des bases de données et de connaissances exploitables dans des applications réelles ?*
- *Quelle est l'utilité de la modélisation ? Et quelles sont les méthodes et langages de modélisation les plus utiles pour le contexte de ces travaux de recherche ?*
- *Quelle sont les caractéristiques d'une architecture globale de système d'aide à la décision permettant l'intégration des données et connaissances ?*
- *Comment tirer bénéfice des approches existantes de gestion des données et connaissances pour supporter un système d'aide au diagnostic ?*

Au-delà de la conceptualisation :

- *Quelle architecture logicielle choisir et quelles sont les technologies disponibles pour l'implémentation du SAD ?*
- *Comment valider l'approche scientifique à travers l'application dans un cas d'étude réel ? et quelles sont les limites de son efficacité ?*
- *Comment assurer la généricité de l'approche proposée pour assurer son adaptabilité à d'autres domaines d'application ?*

II.8. CONCLUSION DU CHAPITRE

Ce chapitre, à travers une étude bibliographique approfondie, aborde, définit et synthétise l'ensemble des notions et éléments nécessaires au développement d'un système d'aide à la décision à base des connaissances. Parmi ceux-ci, les concepts de structuration et gestion ainsi que les méthodes de développement des SADBC ont particulièrement été développés. L'étude de la littérature a permis la distinction de plusieurs travaux intéressants, développés dans le même contexte que nos recherches ou dans des domaines connexes. Ils ont été très inspirants pour construire nos réflexions autour de ces problématiques et proposer notre solution de recherche.

CHAPITRE 3 : UNE APPROCHE D'AIDE A LA DECISION BASEE SUR DES DONNEES ET DES CONNAISSANCES

“Information is not knowledge”

Albert Einstein

“Knowledge is not wisdom

Wisdom is not truth”

Frank Zappa

III.1. INTRODUCTION DU CHAPITRE

Ce chapitre détaille la proposition scientifique élaborée dans le cadre de ces travaux de recherche. Il propose ainsi une méthode d'intégration des données et des connaissances pour le développement des SAD. L'architecture globale du système, les métamodèles permettant l'implémentation des axes d'aide à la décision orientés traçabilité et diagnostic, ainsi que les modèles de données et de connaissances développés seront tous détaillés dans ce chapitre. Celui-ci débute par la présentation des méta-modèles génériques puis de la structuration de la base de connaissances globale. Il sera clôturé par les détails des différentes opérations de classification et modélisation des données et connaissances industrielle.

III.2. RAPPEL ET PREREQUIS SUR LES AXES D'AIDE A LA DECISION

Pour assurer une aide à la décision, le système que nous concevons doit utiliser plusieurs solutions technologiques et techniques de raisonnement. Dans ces travaux de recherche, nous avons choisi les technologies des SAD à base de connaissances (SADBC) dont le fonctionnement est basé sur une logique d'apprentissage pour la réaction. Deux axes sont importants pour le développement et l'implémentation de ces systèmes. Le premier impose l'analyse des données de surveillance et la traçabilité des caractéristiques des phénomènes malveillants. Le deuxième axe consiste à déterminer la cause de l'occurrence des défaillances et proposer des éventuelles solutions de correction.

Le fonctionnement des systèmes SADBC peut être plus complexe et intégrer un aspect prédictif permettant l'anticipation des futures occurrences de phénomènes (troisième axe) et proposant les décisions nécessaires pour empêcher leurs apparitions (quatrième axe).

Quatre axes d'aide à la décision ont donc été identifiés dans ce projet, ils ont été détaillés dans le premier chapitre :

- *Reporting et traçabilité des informations ;*
- *Diagnostic et analyse ;*
- *Prédiction ;*
- *Planification et réaction.*

Ces axes sont fortement liés et complémentaires : il est par exemple indispensable de passer par les opérations de *Reporting* pour pouvoir établir un processus de diagnostic. Remarquons ici la possibilité de déclencher les opérations de *Reporting* ou de Traçabilité à la

demande, périodiquement, ou dès la détection d'un phénomène malveillant. La succession du *Reporting* et du processus de diagnostic permet d'établir une première analyse, nécessaire pour mieux comprendre les causes du phénomène, ses conséquences, et le relier à son contexte d'apparition. Au bout de cette première phase un ensemble de caractéristiques décrivant le phénomène indésirable est stocké sous forme de rapport dans la base de traçabilité et facilite ainsi la mise en place du deuxième axe le diagnostic. Tant pour la prédiction que la réaction, un diagnostic préalable est primordial pour implémenter et exécuter ces tâches. La figure 29 présente le positionnement des axes d'aide à la décision et leur utilité par rapport à l'enchaînement des étapes du processus de surveillance et de suivi des phénomènes malveillants qui peuvent apparaître au moment de l'usinage mécanique. Le système développé étant de type SBC, nous retrouvons ainsi les bases de données et connaissances au centre de ces axes, un élément primordial pour réussir les développements.

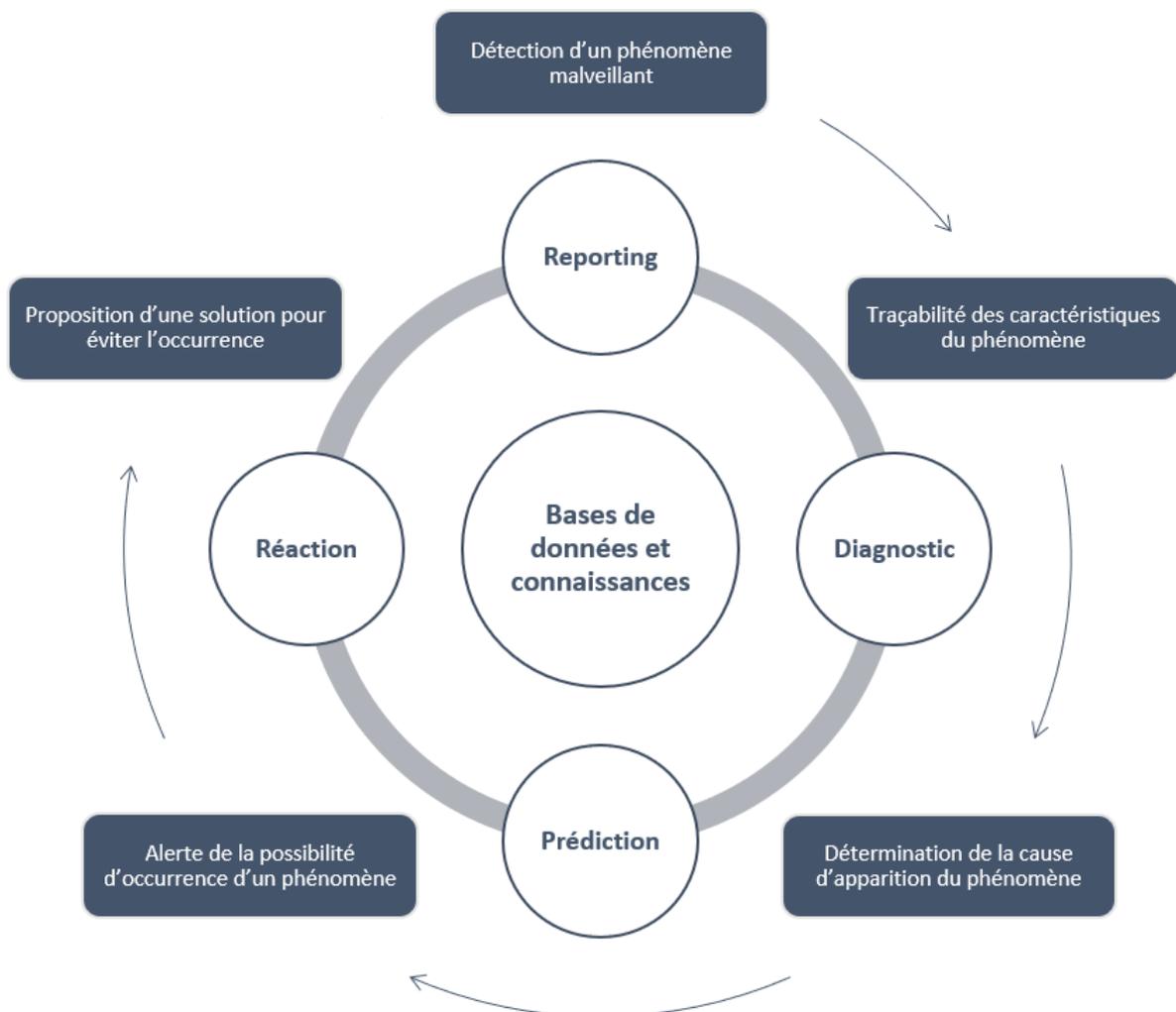


Figure 29. L'enchaînement des axes d'aide à la décision

L'implémentation de ces axes d'aide à la décision est dépendante de la qualité des bases de données et de connaissances mises en place, d'où l'importance de l'approche de modélisation pour la réussite de la mise en place du Framework global.

III.3. ARCHITECTURE DE SAD A BASE DE CONNAISSANCES

Nous allons détailler ici notre approche permettant la structuration et la conceptualisation des systèmes d'aide à la décision. Étant dédiée, dans un premier temps, aux industries de fabrication mécanique, notre approche se focalise, dans un premier temps, sur l'étude des processus liés aux deux premiers axes d'aide à la décision (le *Reporting* et le diagnostic). Comme déjà évoqué, la mise en place des axes de traçabilité et de diagnostic doit favoriser l'implémentation des systèmes de prédiction et de prise de décision. L'originalité de nos travaux se manifeste surtout par les moyens et méthodes adoptés pour faciliter le processus de diagnostic, tous ces aspects seront détaillés dans cette partie. Nous devons avant cela distinguer les frontières, définies dans ce projet, définissant les séparations des différents niveaux de données et de connaissances. Elles permettent de différencier les bases mises en place pour faciliter la capitalisation et la gestion des données et des connaissances.

III.3.1. Approche de distribution des bases de données et de connaissances

La connaissance est la richesse de toute discipline. Il est ainsi indispensable de bien la structurer pour pouvoir la réutiliser et en générer de nouvelles. Les méthodologies de gestion de connaissances sont généralement mises en place dans un objectif spécifique et des cas d'applications uniques dans un domaine précis. La réutilisation de ces approches est souvent contraignante en raison du manque de généricité et de la limitation des moyens d'adaptabilité.

Au-delà de la structuration, ces travaux de recherche visent à développer une stratégie complète de gestion des données et connaissances industrielles. Elle consiste à interconnecter plusieurs types de bases de données et connaissances distinctes et hétérogènes.

Grâce à une approche d'agrégation à plusieurs niveaux [Ritou *et al.*, 2019], le processus de capitalisation des données commence dès l'étape de surveillance qui permet l'acquisition des signaux des capteurs internes et externes instrumentés de la MOCN. Par la suite, une première étape d'agrégation s'applique sur ces signaux et alimente la base de données brutes. À son tour, cette base subit une deuxième étape d'agrégation pour générer des données intermédiaires regroupées dans la base de données Smart. Enfin, une troisième étape d'agrégation et d'analyse de données permet l'extraction de certains KPI qui sont stockés dans la base de traçabilité.

En parallèle, une base de connaissances à plusieurs niveaux contient les modèles formels utilisés dans le processus de multi-agrégation ainsi que les bibliothèques décrivant les caractéristiques des ressources utilisées, du produit développé et des processus adaptés. D'un point de vue technique, cette base est développée en ayant recours aux solutions ontologiques. Elle contient ainsi la combinaison sémantique contextuelle entre les indicateurs d'aide à la décision (composés de KPI et de données smart).

La figure 30 propose un Framework intégrant les différents entrepôts de données numériques. Ce schéma représente, à droite, les bases de données brutes et smart qui contiennent les enregistrements et les données agrégées ainsi que les instances de données. Elles sont structurées suivant des modèles spécifiques qui définissent les relations entre les paramètres et les mesures capitalisés. Ces modèles de structuration, images de la conceptualisation des bases de données, sont eux-mêmes des connaissances et par conséquent nous les retrouvons dans la base de connaissances, figurée à gauche dans le schéma. Dans cette dernière, d'autres formes de connaissances tels que les modèles permettant l'interprétation et le calcul des valeurs des KPI sont également présents. Ils sont, par la suite, stockés dans la base de traçabilité.

Cette stratégie de structuration en plusieurs bases est utile pour plusieurs raisons. Elle permet tout d'abord d'éviter la redondance des données et des concepts. Elle réduit également l'hétérogénéité grâce à l'utilisation de modèles cohérents et communs. Enfin, elle garantit la généralité et favorise en retour l'adaptabilité de la solution proposée à divers types d'applications et disciplines de recherche.

Le modèle en quatre bases facilite la mise en œuvre et l'implémentation de la solution technologique grâce à une approche modulaire. Il peut aussi soutenir les performances de calcul du système expert. Ces caractéristiques aident à résoudre la complexité de l'accès aux données et solutionnent la problématique de la manipulation des *big data*.

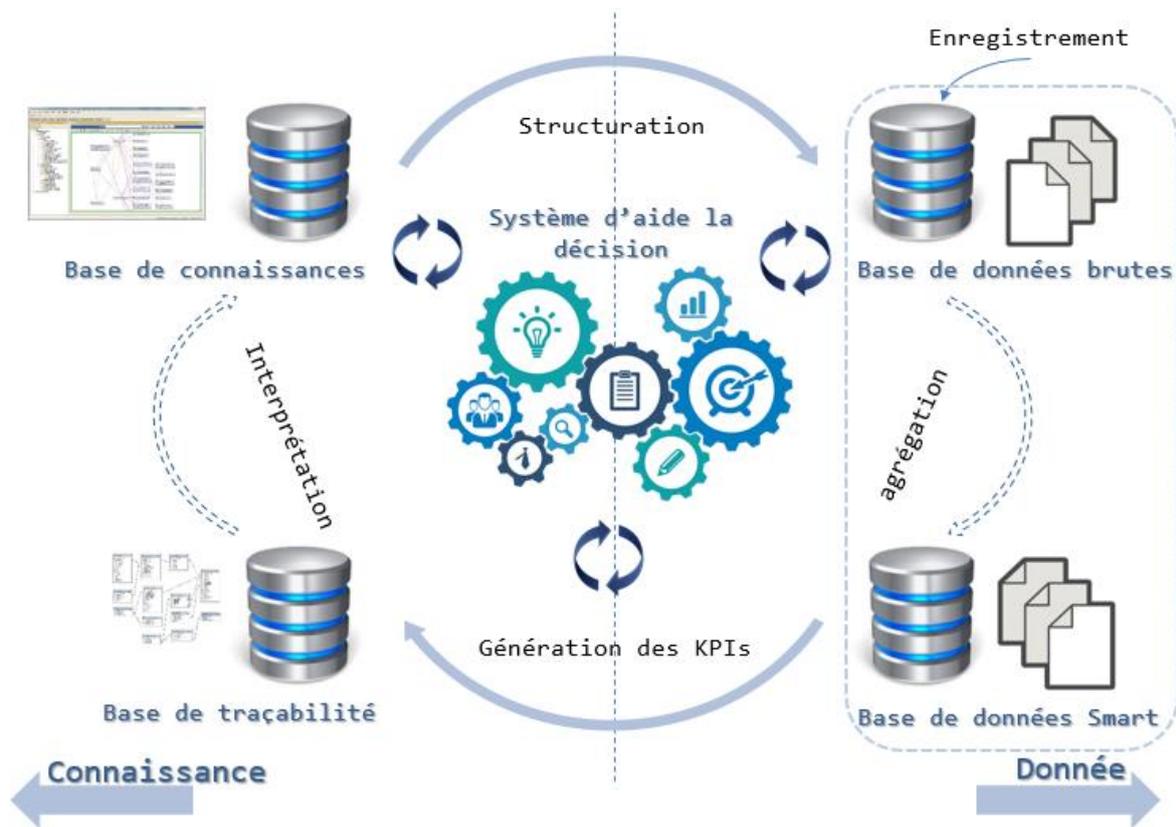


Figure 30. Stratégie de structuration des bases de données et connaissances
(1^{er} niveau : traçabilité)

Ce Framework assure la cohérence sémantique entre les différents modèles de données pour permettre l'interopérabilité entre les différentes bases. La cohérence sémantique est établie lorsque chaque concept partage une même signification entre les différents modèles, même s'il est représenté par des formulations différentes. Les ontologies hiérarchiques sont alors de précieuses alliées pour fournir une représentation conceptuelle commune et éviter les redondances ou l'insignifiance des concepts, depuis le niveau générique jusqu'aux domaines d'activités spécifiques.

Ce premier axe de traçabilité est indispensable à la mise en place des processus de diagnostic qui permettent de générer de nouvelles catégories de données et de connaissances. Le principe de ce deuxième axe d'aide à la décision est détaillé dans la partie suivante.

III.3.2. Le processus de diagnostic : une adaptation du Raisonnement à Partir de Cas

Le diagnostic permet d'analyser les défaillances des ressources utilisées et les phénomènes malveillants apparus au cours du processus d'usinage. Le scénario de diagnostic est basé sur

une caractérisation du contexte de travail pour déterminer les causes de ces phénomènes et proposer la solution appropriée sur la base de l'expérience. À terme, cela permettra d'améliorer la qualité, de réduire le délai du diagnostic classique, et de renforcer la fiabilité du service qualité. Le diagnostic vise également à proposer une solution d'amélioration après occurrence, ou à assurer une opération de prévention permettant d'alléger le travail du service de maintenance en évitant la répétition de phénomènes indésirables similaires à l'avenir.

III.3.2.1. Construction et concepts

Au niveau conceptuel, l'approche de diagnostic proposée est basée sur la méthode AMDEC (présenté dans le chapitre 2) afin de mettre en place les principaux critères et paramètres représentant les défaillances, les causes et les effets. Le recours à cette méthode a permis d'introduire de nouvelles connaissances en termes de règles et de procédures dans l'ontologie, afin de soutenir le processus de l'aide à la décision dans la phase du diagnostic.

Le système d'aide à la décision proposé repose quant à lui sur les principes généraux des systèmes de Raisonnement à Partir de Cas. La notion de cas a été adoptée et les mesures de similarité ont été adaptées à notre Framework par un système d'inférences à base de règles, implémenté dans l'ontologie. La première phase de développement a débuté en définissant la structure d'un cas et les éléments nécessaires pour sa construction, éléments qui sont stockés dans la base de connaissances. Une deuxième phase de construction de la base de règles métier a ensuite été établie à travers des entretiens avec les experts du métier.

III.3.2.2. Fonctionnement

Chaque opération de diagnostic permet de récupérer des éléments de contexte, d'exécuter une règle métier spécifique, de déterminer une cause et ainsi de structurer un nouveau cas. Ceci permettra de déclencher, par la suite, une étape d'apprentissage et de capitalisation des cas, pour générer une nouvelle catégorie de base : la base de cas. Une étape de traçabilité est nécessaire à chaque exécution d'un scénario de diagnostic. L'exécution de plusieurs tests permet, au fur et à mesure des exécutions, à la base de cas de devenir de plus en plus fiable avec un ensemble représentatif et valide de cas. Il devient ainsi possible de lancer un RàPC pour faire de la prédiction et de la réaction suite à la détection d'une situation s'étant déjà produite dans la base de cas.

Ces travaux de recherche ne décrivent pas le système de RàPC et l'implémentation de ce dernier est une perspective de ces travaux, pour la mise en place des troisième et quatrième axe d'aide à la décision. Par contre, nos recherches définissent un système de raisonnement à base de règles

métiers. Les différentes étapes du processus, ainsi que la répartition des différentes bases de données et de connaissances, sont résumées dans la figure ci-dessous.

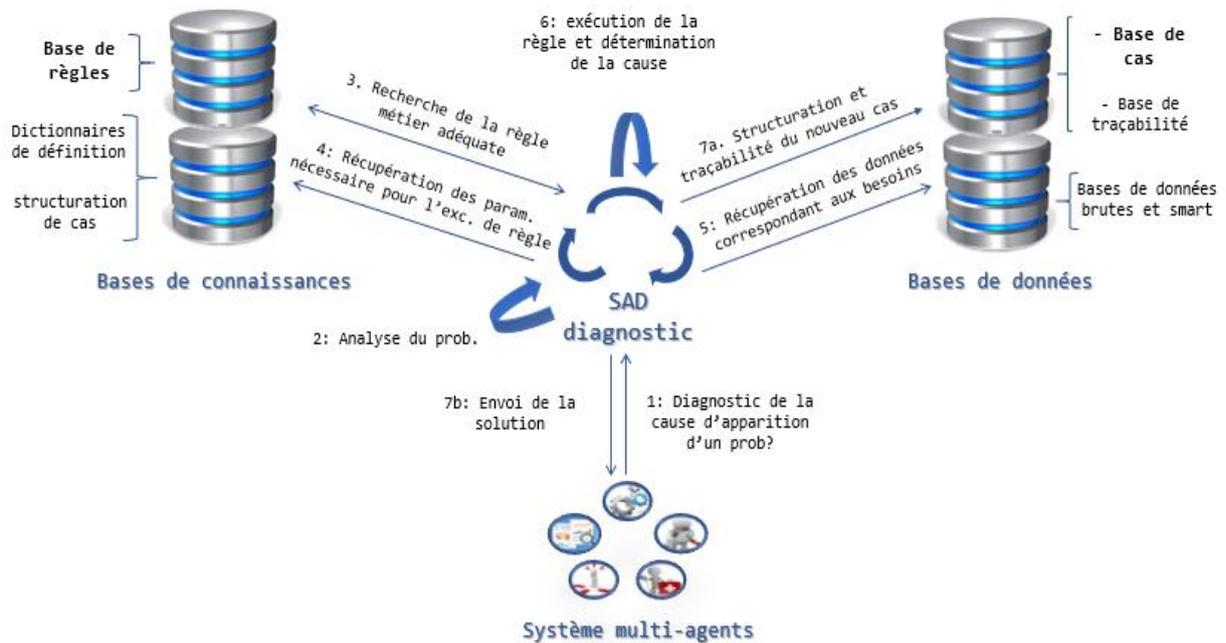


Figure 31. Principe général du système d'aide à la décision : processus de diagnostic

L'objectif de diagnostic étant de détecter les phénomènes anormaux, il est nécessaire d'étudier la notion d'anormalité et de définir les limites de classification. Nous avons eu recours aux techniques d'analyse de données et d'apprentissage automatique pour la définition de critères de détection relatifs à plusieurs phénomènes. La définition et le modèle de calcul de ces critères représentent une nouvelle forme de connaissance stockée dans la base globale. De plus, ces paramètres de détection ont été complétés par d'autres connaissances : les seuils de détection. Ces derniers sont calculés par des méthodes statistiques et représentent la valeur limite qu'un critère de détection ne doit pas dépasser. Le dépassement d'un de ces seuils implique l'occurrence d'un phénomène malveillant. Le diagramme d'activité représenté par la figure 32 détaille le processus d'exécution d'un scénario de diagnostic de la phase d'apprentissage jusqu'à l'instanciation de nouveaux cas. Cet algorithme représente la méthode standard d'analyse et de détermination de l'origine des phénomènes d'usinage pour la plupart des scénarios¹.

¹ Il existe toutefois quelques scénarios qui sont relativement simples et dont la cause, évidente, est directement déterminée, ils seront détaillés au chapitre 5.

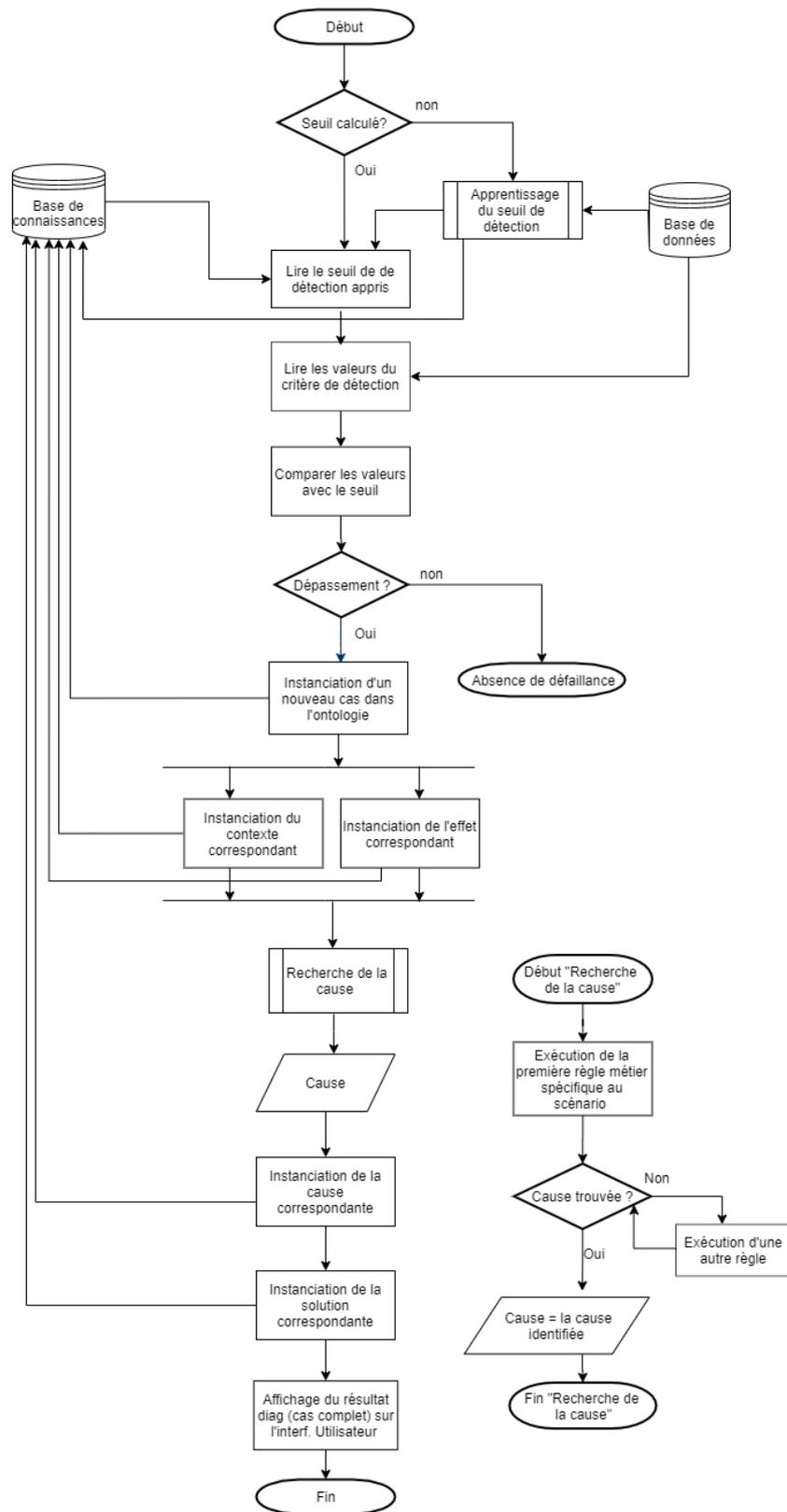


Figure 32. Séquencement d'un processus de diagnostic

Les bases de données et de connaissances sont en collaboration continue. L'exécution des scénarios commence par la communication avec la base de connaissances pour récupérer la

structure nécessaire, puis par l'interrogation, dans un second temps, des bases de données pour récupérer les données correspondantes. Le système exécute, par la suite, un traitement spécifique et instancie des individus dans la base de connaissances. Pour le diagnostic d'événements ces individus sont représentés par des cas et ceux-ci sont structurés avec les objets "Contexte, Cause, Effet, Décision". Ces éléments seront détaillés par la suite.

La notion de cas représente un élément fondamental dans notre approche puisqu'il regroupe les différents éléments de l'opération de diagnostic. La base de connaissances étant de structuration multicouche, ce concept de cas, représentant la troisième couche superficielle d'aide à la décision, permet en outre la connexion avec les autres couches intermédiaires. Le cas stocke les résultats de diagnostic de défaillances comme des interprétations complémentaires liées à un contexte opérationnel. Des connaissances de ce type pourraient par exemple faciliter le prochain processus d'exploration des données en indiquant celles qui causent les effets les plus critiques. Par conséquent, les coûts de traitement sont réduits et la boucle de gestion des données et des connaissances est fermée.

Pour résumer, la spécificité de notre approche repose sur l'intégration des données et des connaissances. Les modèles de structuration des bases de données sont intégrés dans la base de connaissances. Cette dernière est, à son tour, structurée par un méta-modèle global. Le système d'aide à la décision sollicite en continu les deux catégories de bases pour récupérer les définitions de paramètres et les modèles de calculs dans la base de connaissances, et extraire les données correspondantes dans les bases de données. Ceci favorise la généricité et l'adaptabilité de la solution proposée.

Les étapes de modélisation et d'implémentation des différentes bases ainsi que le choix des technologies sont justifiées plus loin dans ce manuscrit. Nous détaillons maintenant les développements réalisés pour définir l'architecture nécessaire à la mise en place des axes d'aide à la décision.

III.4. METAMODELE DES CONNAISSANCES

Notre approche de modélisation des connaissances s'est construite après les étapes d'analyse et de collecte de ces connaissances. Les flux ainsi capitalisés sont souvent relatifs à des objets spécifiques du système industriel global. Notre objectif étant de développer une solution générique, nous avons repéré des familles de concepts, communes à toutes les industries de fabrication, sur lesquelles se base notre solution. La deuxième contrainte est la capacité pour le

modèle proposé de supporter la mise en place des deux premiers axes d'aides à la décision de ce projet. Le fonctionnement des processus de traçabilité et diagnostic dépend en effet fortement de la qualité de structuration des bases de données et de connaissances. Ainsi est-il primordial de mettre en place des entrepôts de données et de connaissances suffisamment complets et, surtout, faciles à interroger. Face à ces critères, une architecture multicouche a été adoptée pour la structuration de la base de connaissances.

La base de connaissances est donc composée de trois principaux groupes de couches : le premier regroupe les couches de modèle de données et de librairies, le second comprend uniquement une couche de standards et le troisième regroupe les connaissances d'aide à la décision.

Pour une structuration formelle du premier groupe de couches de la base de connaissances, une phase d'étude de la littérature a été établie pour analyser les techniques de la modélisation dans les domaines de la production, de la fabrication et d'autres disciplines connexes. Il a été distingué que plusieurs travaux ont recours à l'utilisation des approches de modélisation orientées produits². Après la phase d'analyse et de synthèse de l'état de l'art, et en fonction du contexte du projet, la construction de notre approche a été établie à travers la fusion des 3 approches jugées les plus utiles :

- Approche Produit, Process, Ressources [A. F. Cutting-Decelle, 2007] [Borja, 2016] [Agyapong-Kodua, 2014] ;
- Approche Produit, Process, Ressources, Contexte [Labrousse, 2004] ;
- Approche Personne, Produit, Process [Ericson & Larsson, 2009].

Le choix de ces trois approches s'est basé principalement sur l'intérêt des différents concepts et objets de base qui les composent.

Notre réflexion sur la structuration de l'approche s'est construit comme suit : le cœur de tout système industriel de production est forcément le **produit** fabriqué à travers un **processus** spécifique prédéfini et validé par une **personne** occupant un rôle métier précis dans l'entreprise. Cet expert métier met en place, contrôle et gère les **ressources** nécessaires pour la fabrication, et valide dans un second temps la conformité du produit final à livrer. Le fonctionnement de tout ce système est caractérisé par un **contexte** global regroupant l'ensemble des caractéristiques et des connaissances disponibles.

² On pourra se référer à notre état de l'art pour connaître les détails de ces approches de modélisation orientées produit.

La fusion des différentes approches constitue notre métamodèle : CR3P (Context, Ressources, Product, Process, People) représenté sur la figure 33. Il est composé de cinq packages, le premier package regroupant l'ensemble des connaissances et caractéristiques du « Produit ». Les étapes et les spécifications du « Processus » sont décrites dans un deuxième. Un troisième package regroupe toutes les informations décrivant l'état et les descriptions des « Ressources » utilisées pour assurer la production.

Étant conscients de l'importance de l'humain dans ce domaine et dans le contexte des industries du futur de façon générale, nous le considérons comme l'un des piliers de la transition numérique industrielle. Dans notre approche, l'intelligence et l'expertise humaine participent en amont à la constitution de la base des règles métiers nécessaires à la prise de décision, ou en assurant le développement et la supervision des *process*. Il est également présent en aval puisqu'il est le bénéficiaire principal de la mise en place du Framework global d'aide à la décision. Pour valoriser la place de l'humain, le quatrième package « Personne » a été intégré dans le métamodèle. Enfin, le package « Contexte » regroupe les connaissances spécifiques et nécessaires au système d'aide à la décision. Il permet la description d'un contexte spécifique de travail avec tous les détails et les états des composants du système et des facteurs internes et externes. Par exemple, il peut contenir la description des scénarios opérationnels en termes de référence outil, caractéristiques matériaux, état des broches, etc. La figure 33 représente le métamodèle générique structurant les éléments fondamentaux autour desquels sont capitalisées les connaissances.

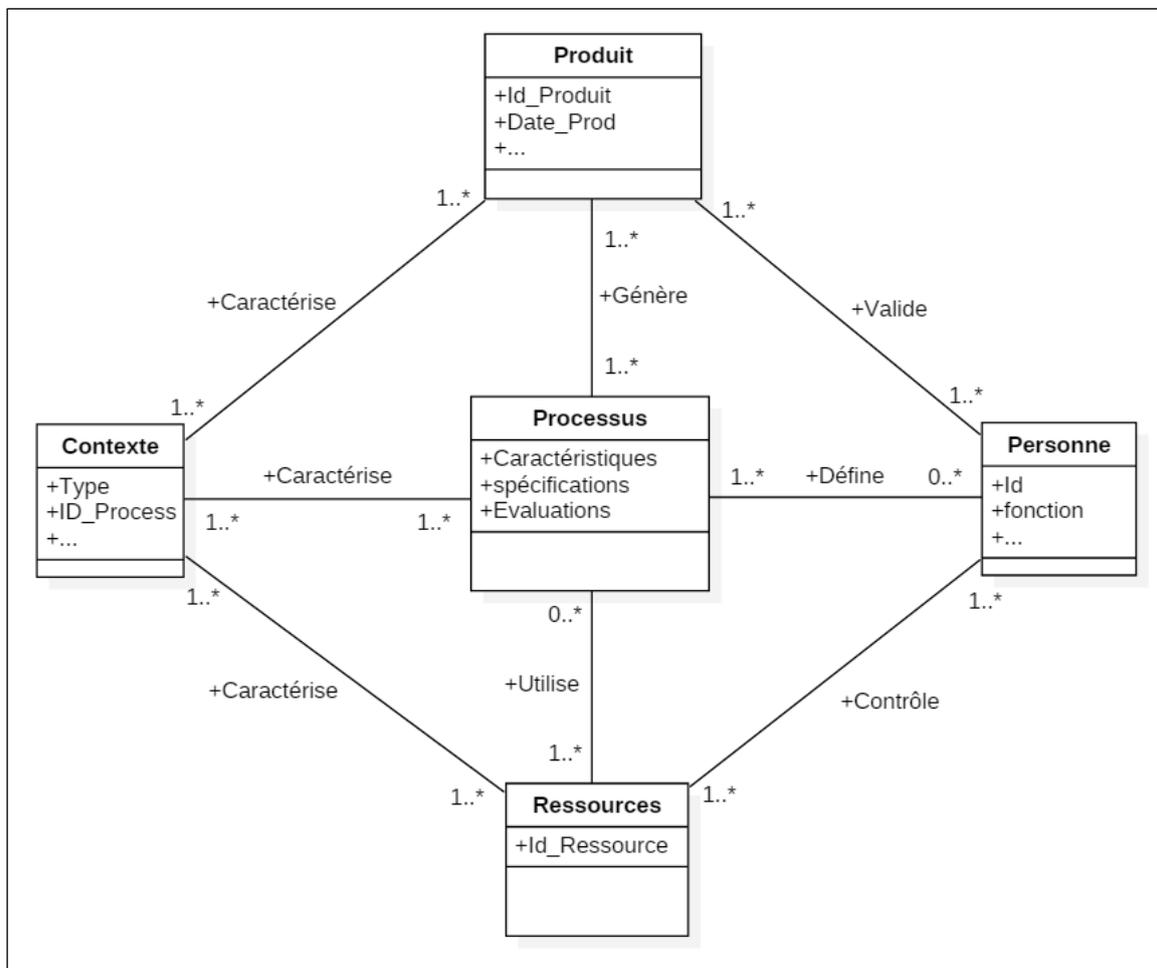


Figure 33. Métamodèle de connaissances

L'approche CR3P supporte la construction des couches de la bases de connaissances. Le premier groupe de couches regroupe un maximum de concepts relatifs au domaine de l'usage mécaniques, précisément les caractéristiques des ressources, des processus et des produits. Ces concepts prennent la forme de bibliothèques qui peuvent être utilisées par la suite pour effectuer les interprétations, le raisonnement et les opérations d'inférence. Les couches du premier groupe prennent donc la forme d'une ontologie de domaine et regroupe également, les modèles de structuration des bases de données. Pour plus de généralité, le deuxième groupe, composé d'une seule couche de normes et de standards, s'ajoute au premier pour assurer l'adaptabilité de la solution dans d'autres application. Le troisième groupe de couches de la base permet de gérer les connaissances décisionnelles, en entrée pour une lecture et réutilisation ou en sortie pour une écriture et ainsi une réintégration de nouvelles connaissances après traitement. Les couches de ce groupe sont structurées sous forme d'un cas composé par le quadruplet « Contexte-Cause-Effet-Décision ». La figure 34 représente la structuration en couche de la base de connaissances.

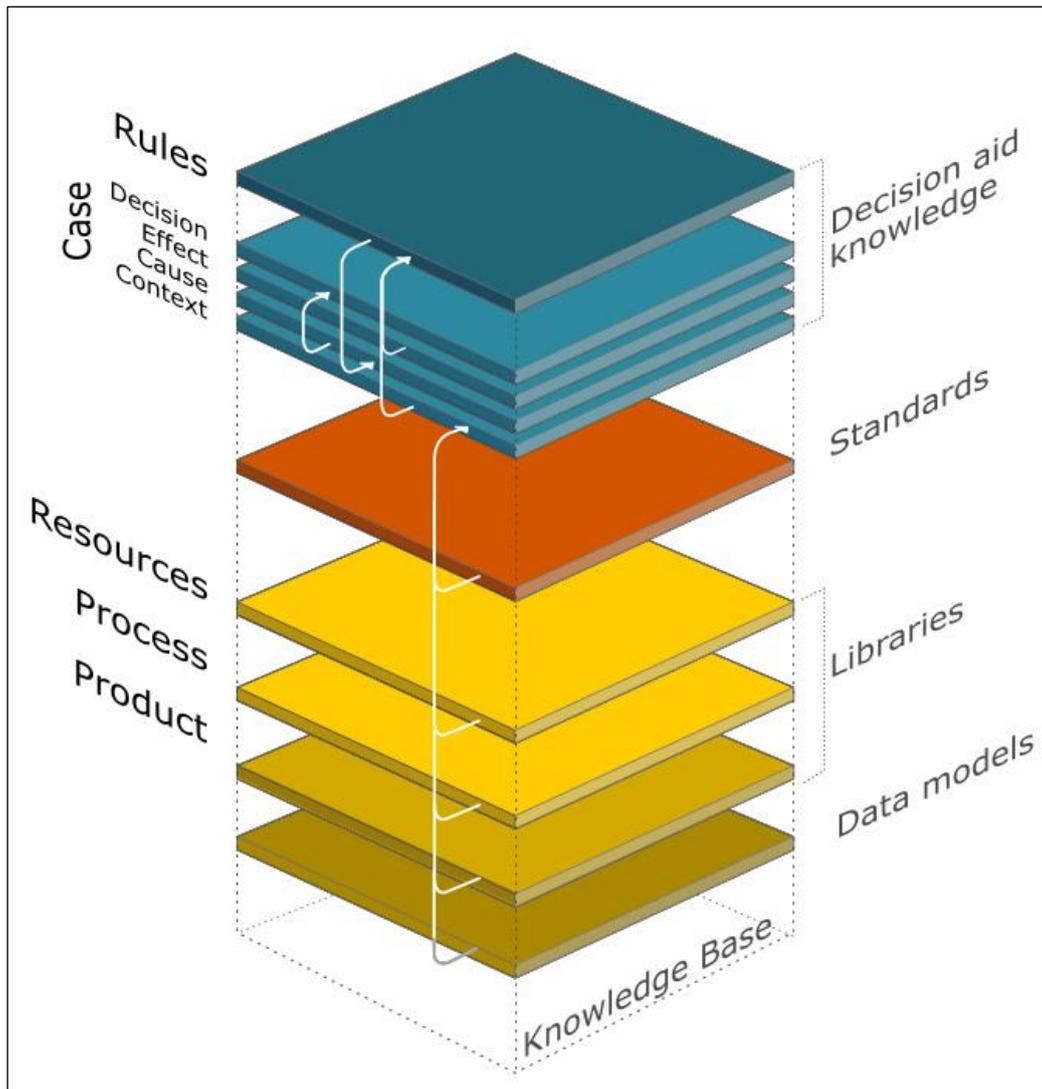


Figure 34. Structure de la base de connaissances et interconnexion des couches

La figure ci-dessus montre la forte interconnexion entre les trois couches constituant l'ontologie. Par exemple, l'exécution d'un scénario d'aide à la décision commence par l'analyse et l'identification des éléments du contexte, caractérisant le cas traité. Les éléments de contexte sont instanciés sur la base des bibliothèques et des connaissances standardisées. La dernière couche, nécessaire pour appuyer le processus décisionnel, est une des points forts de notre proposition scientifique. Sa structuration sous forme de cas, adoptée des systèmes RàPC, supporte notre objectif d'intégration données/connaissances. Ainsi les détails de structuration des cas et le métamodèle de diagnostic sont traités dans la partie suivante.

III.4.1. Structuration d'un « cas » de diagnostic :

Dans le cadre de ces travaux de recherche, un cas est selon nous la brique essentielle au développement de l'axe diagnostique du système d'aide à la décision.

Un SAD est conçu pour aider une ressource (humaine, matérielle ou logicielle) à répondre à une problématique donnée, automatiser une tâche, ou proposer une solution. La réflexion que nous avons construite en combinant l'étude bibliographique approfondie et le contexte industriel de ces travaux de recherche, sur ce sujet de SAD est la suivante :

Les résultats ou la décision prise en sortie d'un processus d'aide à la décision impliquent la compréhension de la relation existante entre un effet observé et une cause stimulante, et ce dans un contexte particulier.

Pour construire cette relation, les *règles métiers* forment une base très importante et représentent l'élément clé de la mise en place de la méthodologie. Elles formalisent l'expertise et l'intelligence humaine nécessaire à la distinction des éléments et à la logique qui sous-tendent une prise de décision.

Un cas de diagnostic doit donc supporter les objets de bases distingués ci-dessus (décision, effet, cause et contexte). Il doit, ainsi, décrire l'effet détecté avec des critères spécifiques, spécifier une ou plusieurs règles précises pour comprendre l'origine de l'apparition (la cause) et pour proposer la décision la plus fiable et adéquate. Il est en outre toujours important de relier l'exécution à un contexte spécifique, décrivant le statut du système et l'ensemble des facteurs internes et externes qui caractérisent la situation globale au moment de l'occurrence du problème. Dans le cadre de ces travaux, nous nous intéressons à la capitalisation de plusieurs paramètres relatifs aux ressources utilisées, au processus de fabrication, au produit final et à l'humain, qui peuvent être l'origine du problème ou la source de la solution et la décision. Plus ces éléments de contexte sont précis, plus l'opération de diagnostic est simplifiée, plus la décision finale prise (une solution d'amélioration dans le cas de ces travaux) est crédible et efficace. Ces éléments de contexte, en plus des critères de détection, sont des formes de connaissances : ce sont des paramètres définis dans la base de connaissances relatifs à des instances dans la base de données.

L'intégration entre ces données et ces connaissances se fait par le biais de la classe « EC- Caractéristiques » qui est connectée directement aux différentes sources de données. L'exécution d'un processus de diagnostic peut en effet avoir recours à l'utilisation des données brutes, Smart, ou KPI. De plus, cette classe est reliée à l'élément Point de traçabilité, relatif au métamodèle de traçabilité³, qui regroupe les rapports générés suite à l'exécution d'un processus de traçabilité. Ces rapports sont constitués d'un ensemble de données filtrées relatives à un

³ Le fonctionnement de ce Point de traçabilité est détaillé au §4.2

phénomène ou à une période d'acquisition spécifique. Ceci explique la forte interconnexion entre les deux axes d'aide à la décision. La figure 35, propose une conceptualisation permettant de connecter les concepts étudiés dans un métamodèle commun.

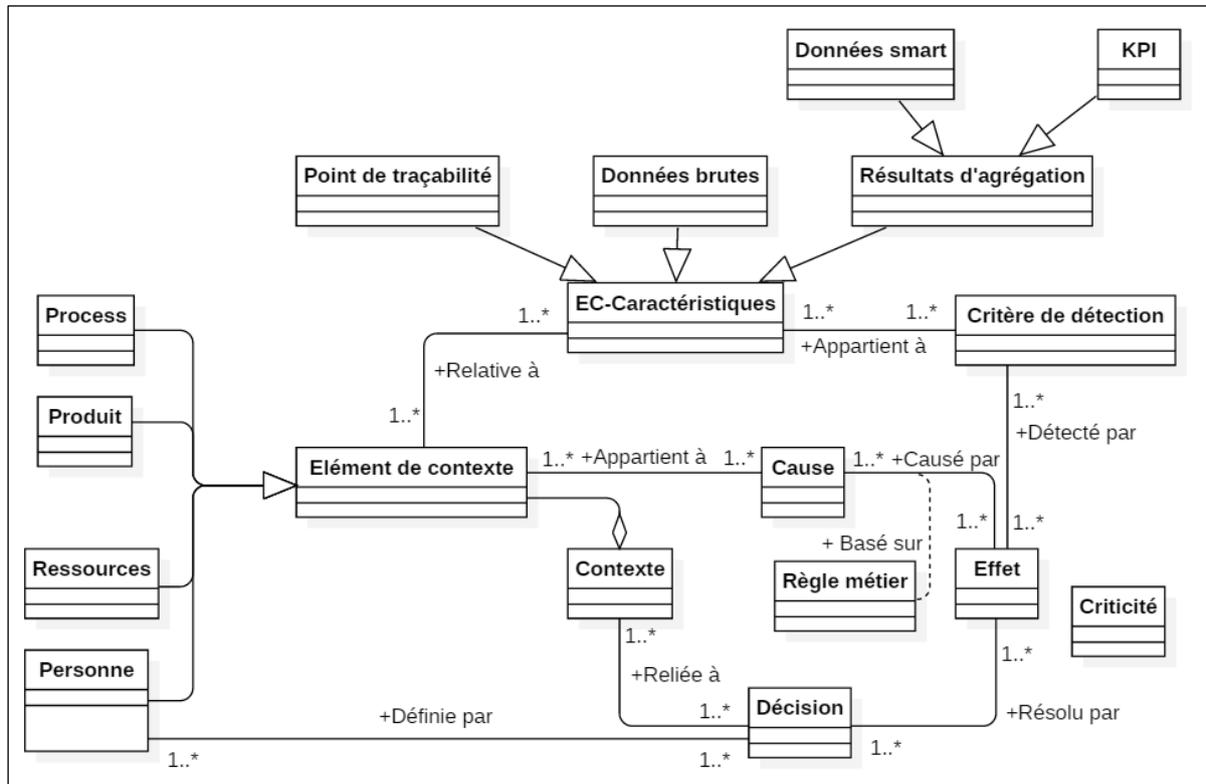


Figure 35. Métamodèle de structuration de cas

Notre approche est ainsi construite autour des objets structurant la troisième couche d'aide à la décision, qui sont les suivants :

- Le **CONTEXTE** fait appel aux éléments nécessaires à la caractérisation de l'environnement de fabrication, à savoir : le nom et la référence de l'outil, le nom du programme et du sous-programme, le jour, etc. Des caractéristiques supplémentaires sur les ressources d'usinage sont également accessibles dans la bibliothèque de ressources (première couche de la base de connaissances). Ces éléments contextuels dépendent fortement du cas spécifique à diagnostiquer et peuvent varier en fonction des besoins de décisions associées et du destinataire du rapport de diagnostic.

- L'**EFFET** représente les caractéristiques et la description de la défaillance à diagnostiquer, dans ce projet c'est le phénomène malveillant qui peut apparaître. L'instanciation de cet élément peut être automatique suite à une détection directe du problème ou bien indirecte suite à une phase d'analyse approfondie préalable. Plusieurs options caractérisent cet objet "effet", à

savoir : un effet peut être l'origine d'un deuxième effet, soit de façon plus générique, soit plus détaillée. Aussi, un effet peut être un élément de contexte d'un autre effet.

- La **CAUSE** est l'élément clé de l'approche de diagnostic à base de connaissance pour identifier l'origine d'événements ou de problèmes malveillants pendant l'usinage d'une pièce mécanique. Pour ce faire, la solution a consisté à formaliser l'intelligence humaine et les meilleures pratiques pour intégrer de l'intelligence artificielle dans les machines sous forme de règles métiers. Dans ce projet, ces règles sont définies dans le but de restituer l'interprétation d'experts sur les incidents passés. Le raisonnement sur l'incident actuel peut ainsi être fait sur la base de son lien avec une situation passée similaire. De nouvelles connaissances sont ensuite générées et instanciées au niveau du cas créé dans l'ontologie. L'exécution des règles de gestion dépend fortement des éléments contextuels et de l'effet observé pour générer la cause la plus fiable, d'où l'importance d'affiner le choix des bons éléments décrivant le contexte opérationnel.

- La **DÉCISION** : Une fois qu'une nouvelle situation indésirable est détectée, diagnostiquée et que ses causes sont identifiées, les derniers éléments à capitaliser sont les solutions proposées et toutes les décisions prises pour résoudre le problème. La boucle du cadre d'aide à la décision basé sur la connaissance est alors bouclée. La notion de décision (ou solution) comprend toutes les actions au-delà du processus de diagnostic qui visent à supprimer les conséquences observées (réparation) ou à anticiper/neutraliser les causes du problème (prévention). En d'autres termes, il s'agit de toutes les règles, actions et procédures qui contribuent à réinitialiser/maintenir le système de fabrication dans ses conditions normales de fonctionnement. Étant donné la spécificité de ce contexte de recherche, il est convenu que la décision est représentée par une solution, mais elle peut prendre des formes différentes dans d'autres cas d'application, en fonction de l'objectif et du besoin de l'utilisateur. Le tableau suivant détaille un exemple de cas, spécifiant le contenu ces différents éléments

Tableau 2. Exemple des éléments d'un cas de diagnostic « Bris d'outil »

Le contexte	L'effet	La cause	La Solution
Jour Programme Id_Outil	Bris d'outil	- Problème de conditions de coupe « CC » (Si les CC réel sont différents des CC de référence)	- Changer l'outil (maintenance) - Examiner les CC

Type d'opération		- Une durée de vie d'outil « DVP » mal définie (Si outil souvent cassé vers la fin	- Revoir la DVP de l'outil (Méthodes)
....			

Dans la stratégie de modélisation sous forme de cas définie ci-dessus, ces objets globaux sont fortement liés et peuvent dépendre les uns des autres. Plus précisément, un premier phénomène **A** (*i.e.* un premier effet "Effet1" avec une cause d'apparition "Cause1") peut être considéré comme l'origine ou la cause d'apparition d'un second phénomène **B** (*i.e.* un second effet "Effet2" avec une cause "Cause2"), comme le montre la figure 36. Cela implique l'exécution d'un second processus de diagnostic. Plus formellement :

$$\ll \text{Cause1} \rightarrow \text{Effet1} = \text{Cause2} \rightarrow \text{Effet2} \gg$$

Dans ce cas, le contexte peut être aussi bien identique pour les deux scénarios ou spécifique à chacun des deux. Plus formellement :

$$\exists \text{ContextGlob} \text{ tel que } \ll \text{Cause1} \rightarrow \text{Effet1} = \text{Cause2} \rightarrow \text{Effet2} \gg$$

$$\text{Sachant que } \{ \text{Cause1}, \text{Effet1}, \text{Cause2}, \text{Effet2} \} \in \text{ContextGlob}$$

$$\text{Ou : } \exists (\text{Context1}; \text{Context2}) \in \text{ContextGlob} \text{ Tel que } \ll \text{Cause1} \rightarrow \text{Effet1} = \text{Cause2} \rightarrow \text{Effet2} \gg$$

$$\text{Sachant que } \{ \text{Cause1}, \text{Effet1} \} \in \text{Context1} \ \& \ \{ \text{Cause2}, \text{Effet2} \} \in \text{Context2}$$

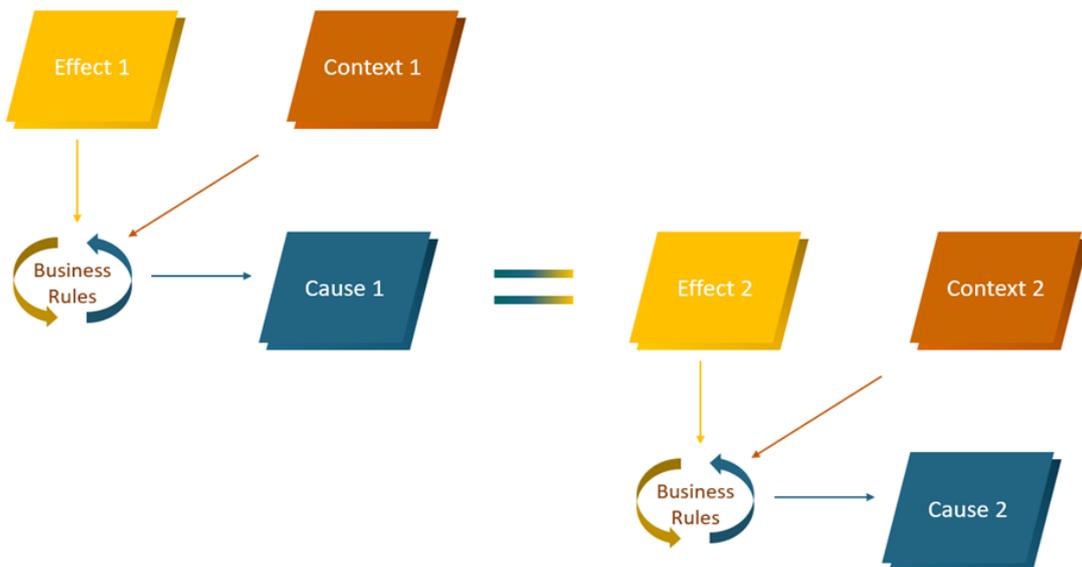


Figure 36. Dépendances entre les éléments de diagnostic

De point de vue technique, ces quatre objets de base ont été définis comme des classes pour la structuration de ces opérations de diagnostic au niveau de l'ontologie. En outre, la liste des relations qui lient ces classes aux attributs qui les identifient ont été aussi précisées. Au fur et à mesure que les analyses s'accumulent, ces classes seront instanciées sous forme de groupes d'individus correspondants. Cette instanciation s'appuie sur des bases de données industrielles préalablement structurées (Brutes, Smart et Traçabilité) : ceci met en évidence l'utilité de la stratégie de gestion des connaissances, qui permet l'interaction avec les données pour générer de nouvelles connaissances. Ces groupes d'individus sont reliés à chaque fois dans un nouveau cas, ce qui signifie qu'une base de cas sera construite au fur et à mesure de la réalisation des diagnostics. Le choix de cette structuration sous forme de cas n'est pas arbitraire mais résulte de notre stratégie de gestion. Cette structure garantira plusieurs perspectives de développement pour des futures travaux de recherche. Éventuellement, elle permettra la mise en place du troisième axe d'aide à la décision : la prédiction.

L'exécution d'un processus de diagnostic est une opération assez complexe et pointue. Il est important d'assurer un bon raisonnement pour offrir des résultats le plus fiables possible. Pour appuyer le fonctionnement de cet axe d'aide à la décision, et réduire les opérations de calculs et d'interopération, l'axe de traçabilité s'est montré très efficace. La partie suivante détaillera le métamodèle de structuration pour le *Reporting* et la traçabilité.

III.4.2. Reporting et modèle de traçabilité :

Aujourd'hui, plusieurs études montrent que plus que 85 % des données stockées dans le monde sont soit redondantes, soit obscures, soit inutiles, soit obsolètes. Le coût causé par la surcharge de données est estimé en 2020 à presque 3000 milliards d'euros [Le Rapport Mondial Databerg, 2016]. Ainsi, pour que le processus de traçabilité soit optimal, il est important qu'il ne s'exécute que suite à l'apparition d'un événement malveillant au moment de l'usinage. Ceci permettra d'éviter de surcharger la base de traçabilité en ne stockant que les données utiles à une traçabilité dite « sur événement ». Toutefois, suite à plusieurs entretiens avec les experts métiers, nous estimons qu'il est aussi important de prendre en considération la possibilité d'exécution d'un processus de traçabilité par une personne à un moment donné, traçabilité dite « sur demande », ou de façon périodique, traçabilité dite « périodique ». Ces analyses ont confirmé qu'il est indispensable de définir et d'étudier les concepts de modes et de fréquences d'exécution pour l'élaboration d'un processus de traçabilité. Par conséquent, nous distinguons les traçabilités sur demande, sur événement, ou périodique.

D'autre part, chaque rapport traité ou trace générée à la suite de l'exécution d'un traitement est destiné à une personne précise occupant un rôle métier spécifique dans un département de l'entreprise. Il est ainsi important de prendre en charge la ressource humaine à laquelle se destinent ces rapports comme un concept de base dans le méta-modèle global de structuration. Un troisième concept aussi important est celui du « Contexte d'utilisation ». Il est très important de relier chaque opération de traçabilité à son contexte spécifique. Cela facilitera surtout les opérations d'analyse futures et permettra, parfois, de comprendre l'origine du phénomène indésirables et donc diagnostiquer la cause d'apparition et ce qui fera de l'objet « contexte » le point centralisé et commun entre les deux premiers axes d'aide à la décision. Il est possible de distinguer plusieurs types de contextes ou de facteurs qui influencent le fonctionnement d'un système. Par exemple, le contexte environnemental regroupe les paramètres physiques de temps, d'espace, de températures, de bruits, etc. ; Le contexte technique peut être relié aux configurations du système, aux moyens technologies et ressources utilisées, aux réseaux, et aux fiabilité et stabilité du système, etc. ; Le contexte organisationnel qui est directement lié aux exigences et méthodologies de travail, aux décisions organisationnelles, et aux opérations de management et de contrôle. La capitalisation du contexte et la distinction du rôle métier spécifique, font partie des métadonnées du rapport de traçabilité. Plus encore, ils représentent l'Objectif de traçabilité. Par la suite, pour commencer la capitalisation des statistiques sur l'état des ressources utilisées, le système multi agent va récupérer des données hétérogènes (brutes, traitées, KPI, etc.) et les stocker dans le même rapport (la même table de la base de données) que les données de contexte, sous forme d'un « Point de traçabilité ». Le cumul de ces points forme des bases de données spécifiques, appelées « base de traçabilité ».

Pour conclure, la Base de traçabilité repose ainsi sur la notion de « Point de traçabilité » comme concept de base (implémentant des enregistrements dans la base de données). C'est cette classe Point de traçabilité qui permet de connecter le processus de traçabilité à celui de diagnostic. Nous proposons la définition de point de traçabilité suivante :

Un point de traçabilité regroupe plusieurs données tracées dans un objectif ou un besoin bien précis sur une période précise et en lien avec une référence pièce / processus unique. Les points de traçabilité seront exploités pour fournir les rapports de synthèse à destination des acteurs métier occupant un rôle spécifique dans un département de l'entreprise.

Afin de savoir comment structurer cette base de traçabilité, plusieurs scénarios ont été définis précisant les besoins en termes de données d'entrées et de sorties, les départements concernés par ces données ainsi que le mode et la fréquence de collecte et de restitution. Le déroulement

Pour conclure, l'approche détaillé ci-dessus est basé sur un raisonnement à base de connaissances. Il est ainsi important de bien structurer et capitaliser les flux d'informations industrielles pour faciliter leurs réutilisations. La partie suivante détaille l'analyse et la modélisation des données et des connaissances.

III.5. STRUCTURATION DES DONNEES ET CONNAISSANCES INDUSTRIELLES

Notre démarche de développement, comme expliqué dans le premier chapitre, est basée sur une approche d'analyse pour faciliter la modélisation et la collecte des données et des connaissances. Une première phase d'analyse de la situation telle qu'elle (*as-is*) chez les partenaires industriels du projet et une confrontation avec la littérature ont été établies afin de comprendre au mieux les flux industriels disponibles et afin de pouvoir développer des modèles efficaces et fiables. Conscient du fait que les modèles (en terme de structuration d'entités) sont en soi une forme de connaissance, nous avons choisi de regrouper dans la base de connaissances : les modèles de structuration des bases de données dédiées à l'enregistrement de mesures réelles caractérisant tout le processus de fabrication, les modèles descriptifs des bases de connaissances destinés à stocker les différentes règles métier et d'inférences, les métadonnées caractéristiques des machines et produits, et autres informations génériques telles que les normes. Dans cette catégorie, les seuils de comportement de la machine obtenus par apprentissage sont également pris en compte. Tous ces aspects d'analyse et de modélisation, en commençant par les techniques les plus simples jusqu'aux plus développées, seront expliqués dans les parties suivantes.

III.5.1. Modélisation de la situation AS-IS

La réussite de la modélisation repose sur la maturité de la phase d'analyse préalable. Cette étape est primordiale pour analyser les processus opérationnels, énumérer les ressources utilisées mais aussi pour identifier et comprendre les flux matériels et informationnels circulant entre les différentes phases d'un processus industriel.

La phase d'analyse a conduit dans un premier temps à l'analyse de la « *AS-IS situation* » qui décrit le système industriel tel qu'il existe et tel qu'il fonctionne actuellement. Selon les besoins et le contexte étudié, il existe plusieurs méthodes pour faire ces opérations d'analyse :

- Les diagrammes de processus ;

- Les diagrammes de cas d'utilisation ;
- Les diagrammes d'activité ;
- Les diagrammes FAST (*Function Analysis System Technique*) ;
- Les tableaux temps-procédure ;
- La méthode QQQQCP (Quoi, Qui, Où, Quand, Comment, Pourquoi), etc.

Pour respecter la contrainte de simplicité des modèles et pour qu'ils soient compréhensibles par l'ensemble des acteurs du projet, une première analyse de la littérature a permis de distinguer une technique basique permettant de faciliter la communication avec les partenaires industriels ainsi que l'analyse et la description des flux. Cette technique, c'est la méthode américaine IDEF0 qui découle de SADT.

Cette méthode permet de faire une analyse des activités de chaque service pour un industriel donné et donc des différentes étapes de la réalisation d'une pièce. Plusieurs avantages justifient le choix de cette méthode. Facile à utiliser et d'un fonctionnement compréhensible par tous, elle assure :

- La possibilité de faire des analyses fiables, de qualités, complètes et précises ;
- La facilité de communication entre les utilisateurs ;
- Une expression des besoins facilitée pour le client ;
- Un suivi de l'avancement et un contrôle facilité pour la direction ;
- L'amélioration du travail en équipe, à travers la répartition des tâches et la coordination des efforts.

Suite à l'entretien avec des spécialistes et des experts métiers, une première version des modèles de processus a été développée. Par la suite, une étude de la littérature sur les processus industriels de la fabrication mécanique a permis de les enrichir et de développer des modèles génériques regroupant un maximum de connaissances industrielles. Un exemple de modèle processus est présenté dans la figure 38. Les versions intermédiaires détaillées et validées de ces diagrammes sont représentées en annexes (Annexe 2).

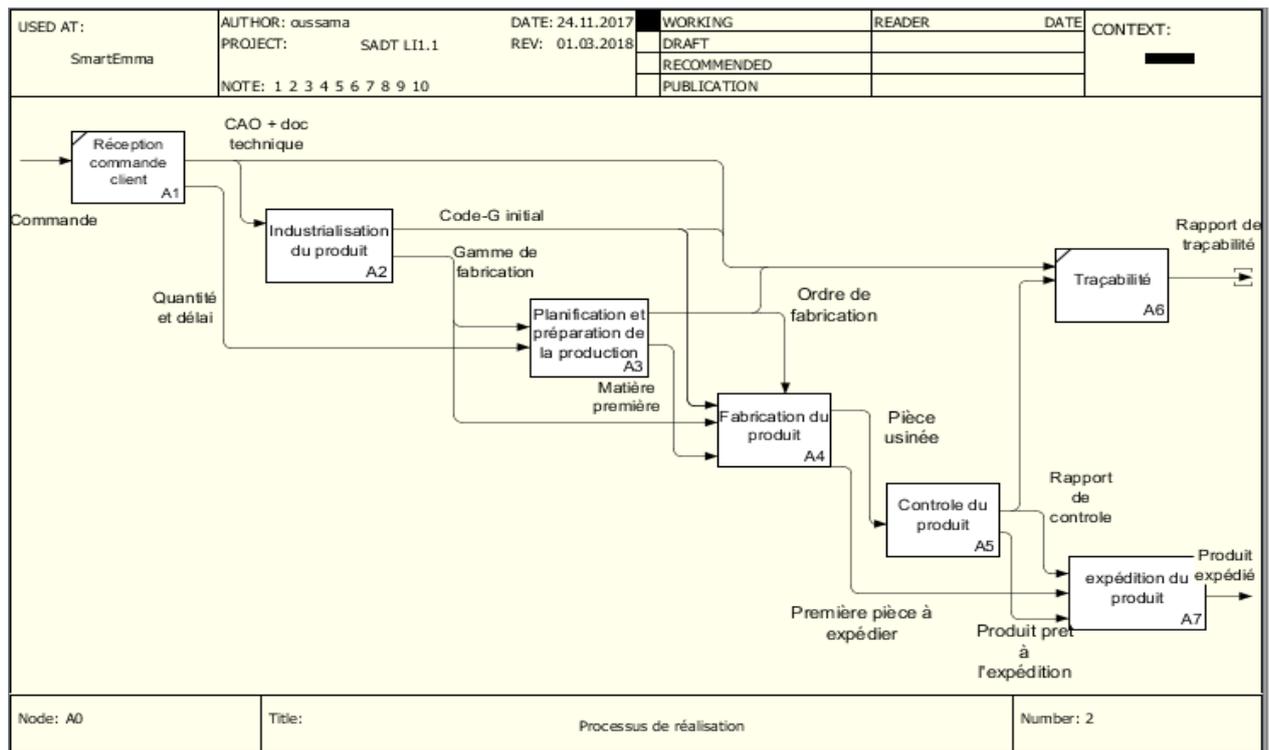


Figure 38. Processus de réalisation (vue micro)

Ce digramme décrit les différentes étapes nécessaires pour réaliser et livrer une pièce à un client. Il montre les flux d'informations générés à la sortie de chaque étape en allant de la commande du client jusqu'à l'expédition, tout en passant par la partie industrialisation au niveau du bureau des méthodes, la planification et la préparation de la production, et finalement la fabrication et le contrôle du produit.

Cette première phase d'analyse de la situation AS-IS a permis de mieux comprendre le système industriel étudié tout en analysant les différents éléments qui le composent. Ces diagrammes sont efficaces parce qu'ils permettent de distinguer les données qui concernent les ressources, les contraintes et les exigences liées au *process* ainsi que les entrées et sorties décrivant les différents états d'un produit au fur et à mesure de l'avancement de la fabrication. Ceci a permis de distinguer plusieurs catégories de données et de définir, surtout, les concepts de base composant le système global à savoir : les ressources, le *process* et le produit.

Une fois la phase d'analyse des processus terminée sur la base de la littérature et de la réalité industrielle, la deuxième phase consistait à mieux identifier et classifier les différents types de données en vue de les intégrer dans les modèles.

III.5.2. Méthodologie de développement des modèles

Notre objectif est de fournir un cadre de modélisation suffisamment général et robuste pour prendre en charge les différentes formes de données et de connaissances industrielles requises pour le développement du Framework global. Cette partie explique la méthode suivie pour développer ces modèles.

Ces travaux de recherche présentent un exemple concret d'implémentation d'un processus d'extraction des connaissances à partir des données, permettant d'affiner la méthodologie de développement tout en séparant les deux phases de modélisation de données et de modélisation des connaissances.

La première phase du processus d'ECD⁴ est la préparation des données qui est composée de quatre étapes : la sélection, le nettoyage, la transformation, et la réduction.

La préparation permet de réduire la taille et d'agrèger les énormes quantités de données collectées, pour garder uniquement celles qui sont jugées nécessaires afin d'entamer la deuxième phase du processus, la fouille de données, et d'assurer ainsi l'extraction des connaissances.

Différentes sous étapes permettent de générer plusieurs catégories de données de plus en plus significatives au fur et à mesure de l'application des opérations d'agrégation. Les données générées à chaque niveau sont nécessaires et doivent être prises en compte au niveau des modèles développés. En effet, chaque catégorie de donnée est utile pour une ou plusieurs personnes ou encore pour le développement d'une partie du Framework global.

Les opérations d'analyse et d'extraction sont cruciales pour bien étudier les données et les connaissances. Elles permettent de comprendre les flux disponibles pour assurer une modélisation optimale, mais avant cela, pour assurer une capitalisation efficace et complète des données.

L'objectif de ces premiers modèles est de supporter la structuration de données avant d'être exploités pour le développement de la plateforme d'aide à la décision.

Comme évoqué dans notre état de l'art⁵, la modélisation de données permet d'obtenir une représentation graphique simple et compréhensibles des données industrielles disponibles. Cette description, relationnelle dans le cas de ce projet de recherche, permet d'établir les liens

⁴ On pourra se reporter au premier chapitre pour l'explication de l'ECD.

⁵ Détaillé dans le chapitre 2 partie 4.2

entre les différents composants d'un système industriel donné, ce qui permet de renseigner les possibilités de communication, d'héritage ou de collaboration entre les flux de données.

Pour le développement des modèles de données, le langage UML a été choisi, et ceci pour plusieurs raisons. Principalement, UML est considéré comme le langage la plus utilisée pour l'implémentation de l'approche objet. Le principe de cette démarche est d'assurer une modélisation incrémentale (permettant une flexibilité de mises à jour) et itérative (permettant de passer d'une première version simple à une version finale complète). Cette approche, dépendante fortement des besoins clients, doit décrire l'ensemble des informations en plus des actions à établir dans la même entité. L'adaptabilité du langage UML, avec sa multitude de diagramme mise à disposition et spécifiquement le diagramme de classe pour cette dernière contrainte, est donc un atout pour cela. Autre avantage, UML offre la possibilité de génération de plusieurs scénarios de test et de composants logiciels de façon automatique.

UML est tellement répandu qu'il favorise une collaboration entre les différents acteurs métiers sur la base d'un seul document, de façon aussi bien statique que dynamique. En plus de ses aspects structurels, ce langage permet d'assurer des modèles de très haut niveau, abstraits et modulaires.

De fait, le recours à l'utilisation d'UML s'avère incontournable dans ce projet. Le développement des modèles a suivi la même démarche incrémentale que nous avons suivie pour le développement de la plateforme. Ainsi, la première version a été développée et enrichies tout au long du développement du projet. Nous détaillons maintenant les étapes de classification et modélisation des données.

III.5.3. Classification et modélisation des données

L'analyse de la littérature permet de distinguer plusieurs techniques de collecte de données : observation directe, mesure, enregistrement, entretiens, questionnaires, analyse par inspections régulières, génération de données auxiliaires, etc.

La plupart de ces techniques ont été mises en place dans ce projet pour assurer la capitalisation des données et connaissances. Par exemple, la mesure et l'enregistrement ont été utilisés pour récupérer des bases de données de surveillance de l'usinage, à travers l'installation des appareils Emmatools au niveau des MOCN, et l'instrumentation des machines par l'ensemble des capteurs nécessaires pour l'acquisition des signaux.

Les techniques d'analyse par inspection et génération de nouvelles données avec le recours aux opérations de fouilles, ont quant à elles permis de transformer les données brutes de surveillance vers une autre catégorie de données plus fine appelée Smart.

Le développement de ces travaux de recherche a eu recours à un autre aspect de collecte important, celui de l'entretien avec les experts pour pouvoir développer une base de règles métiers nécessaire au développement des scénarios de diagnostic (ces aspects seront traités dans le chapitre suivant).

III.5.3.1. Structuration des données Brutes

Selon *Le grand dictionnaire terminologique*, les données brutes sont recueillies sur un sujet précis à travers les observations ou les mesures, et qui n'ont pas subi un traitement ou une manipulation :

En télédétection, les données brutes proviennent d'un signal issu d'un capteur. Elles sont mises sous une forme soit analogique, soit numérique, et rendues disponibles en vue de leur traitement. Les données brutes n'ont été ni validées ni triées. Elles s'opposent ainsi aux données statistiques qui, elles, ont été traitées [Site Web GDT , 1999].

Dans nos travaux, les données brutes correspondent essentiellement aux bases de données récupérées directement auprès des partenaires industriels du projet. Il est possible de distinguer deux catégories : Les données de surveillance acquises à travers le système Emmatools et les autres types de données stockées sur des supports variés (les catalogue d'outils, la liste des programmes et des sous-programmes, les conditions de coupe, les fiches qualité, les ordres de fabrication, etc.). Dans Emmatools, les données des capteurs sont acquises à l'aide d'une carte NI (*National Instrument*) échantillonnée à 25 kHz.

La mise en place de ce système d'acquisition génère des énormes quantités d'informations, Ainsi un premier niveau d'agrégation et de traitement a été adopté pour calculer des critères de mesure (V_{rms} , V_{pic} , A_{rms} , A_{pic})⁶ sur une période de 0,1 s puis stockés dans une base de données de surveillance, nommée base Emmatools.

⁶ A_{rms} : représente la moyenne quadratique du signal d'accélération brut sur une période donnée. Ce critère est très sensible aux vibrations hautes fréquences (par rapport à la fréquence de rotation de la broche) ce qui favorise une bonne détection du broutement.

A_{pic} : permet essentiellement de détecter un choc. Il est calculé de la même façon que A_{rms} en ne gardant que la valeur max de l'acquisition sur la période donnée.

V_{rms} : représente le critère le plus utilisé dans le domaine de la surveillance de machines tournantes. Son calcul se fait sur trois étapes. Premièrement, il faut appliquer un filtre passe bande (50 - 3500 Hz) au signal afin de supprimer les effets qui ne seraient pas liés à l'usinage (roulement broche ou modes vibratoires de la machine). Ensuite, il faut réaliser une intégration numérique pour transformer l'accélération en vitesse. Et la dernière étape consiste à appliquer un opérateur de moyenne quadratique pour obtenir la valeur finalement stockée. Ce critère est surtout très sensible aux basses fréquences,

Le spectre fréquentiel est calculé sur chaque accéléromètre (par algorithme FFT, *Fast Fourier Transform*) sur une fenêtre glissante de 0,3 s (pour améliorer la fiabilité). Comme la taille de ce spectre est très grande, seuls les trois à cinq plus grandes composantes en amplitude et leurs fréquences respectives sont collectées.

Par conséquent, il est possible de distinguer deux premiers niveaux : *les données réelles* qui représentent le signal original collecté par les capteurs et *les données brutes* qui sont stockées dans la base EmmaTools après traitement. Cependant nous considérons seulement la deuxième catégorie dans le modèle de données, en plus des paramètres utilisés pour le traitement du signal réel et les différentes opérations d'agrégations.

Dans ce même contexte, la première phase d'analyse de la situation *AS-IS* incluait aussi l'étude des bases de données de surveillance, récupérées directement chez nos partenaires industriels, pour pouvoir énumérer et comprendre l'utilité des données disponibles. Un aperçu de ces mesurandes est représenté par la figure 39 :

- La première colonne du tableau comporte les paramètres, récupérés chez le premier industriel (données de contexte, données temporelles, puissances, fréquences, positions, vitesses, etc.) ;
- La deuxième colonne du tableau définit les unités de chaque valeur mesurée ;
- La troisième colonne détaille les sources d'acquisition des données, sachant qu'il y a principalement trois sources : les données de contexte de la commande numérique, les données des capteurs internes de la MOCN, les données acquises à travers les capteurs additionnels implémentés dans la machine ;
- Une quatrième colonne définit chaque paramètre ;
- Finalement, une dernière colonne de *mapping* permet d'associer le nom correspondant du paramètre équivalent, se trouvant dans la base de données de surveillance chez le deuxième industriel. Cette colonne a été ajoutée, puisque la version initiale du logiciel d'acquisition ne fournissait pas un fichier avec les mêmes noms de paramètres pour les différents partenaires industriels.

Au vu de la différence entre les références et versions des MOCN implémentées chez les différents industriels, les types d'opérations d'usinages exécutés, les outils utilisés, ainsi

il est souvent utilisé pour la détection du balourd ou pour la distinction des premières harmoniques qui sont créées par le passage des dents dans la matière.

Vpic : Pareil que Vrms, les trois étapes de calcul sont utilisées, la seule différence est au niveau de la dernière étape où juste la valeur maximale est stockée.

que des nombreuses autres contraintes d'utilisation, plusieurs paramètres peuvent être absents dans un contexte et présents dans un autre.

Données Industriel 1	Unité	Source	Description	Equivalent Industriel 2
tempstotal	min	Données CN (contexte)	Temps total d'utilisation de la broche	tspT
tempsbroche	min	Données CN (contexte)	Temps broche en rotation	tps B
date	#	Données CN (contexte)	Date	date
prog1	#	Données CN (contexte)	Identifiant programme principal	idpp
idbloc	#	Données CN (contexte)	Identifiant bloc du programme	id pc
jour	#	Données CN (contexte)	Le jour	jour
outilref	#	Données CN (contexte)	Identifiant de l'outil	id outil
xavarms	m/s ²	Données capteurs additionnels	Xav	Arms
yavarms	m/s ²	Données capteurs additionnels	Yar	Arms
xavapic	m/s ²	Données capteurs additionnels	Xav	Apic
yavapic	m/s ²	Données capteurs additionnels	Yar	Apic
xavvrms	mm/s	Données capteurs additionnels	Xav	Vrms
yavvrms	mm/s	Données capteurs additionnels	Yar	Vrms
xavvpic	mm/s	Données capteurs additionnels	Xav	Vpic
yavvpic	mm/s	Données capteurs additionnels	Yar	Vpic
px	mm	Données CN (capteurs)	X	Pos
py	mm	Données CN (capteurs)	Y	Pos
pz	mm	Données CN (capteurs)	Z	Pos
pa	*	Données CN (capteurs)	A	Pos
pb	*	Données CN (capteurs)	C	Pos
n	tr/min	Données CN (capteurs)	Vitesse de rotation de la broche	N
p	Kw	Données CN (capteurs)	Puissance consommée par la broche	P
fft3pts_harmo_freqN_xav	Hz	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_harmo_ampN_xav	m/s ²	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_harmo_freqN_yar	Hz	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_harmo_ampN_yar	m/s ²	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_nharmo_freqN_xav	Hz	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_nharmo_ampN_xav	m/s ²	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_nharmo_freqN_yar	Hz	Données capteurs additionnels	Spectre fréquentielle	FFT
fft3pts_nharmo_ampN_yar	m/s ²	Données capteurs additionnels	Spectre fréquentielle	FFT
chatter_ar		Données capteurs additionnels	Valeur broutement	Pas chez Indus2
chatter_av		Données capteurs additionnels	Valeur broutement	Pas chez Indus2

Figure 39. Extrait des éléments de données brutes

Il est évident que ces incompatibilités et différences entre les données brutes issues des différents partenaires industriels sont sources d'inefficacités et peuvent retarder ou freiner les projets collaboratifs entre les industries, ou l'évaluation statistique des performances, etc. L'importance de cette première phase d'analyse se manifeste donc dès à présent : en plus de son utilité pour la modélisation, cette énumération a présenté un premier support de discussion avec les industriels, autour des éventuelles solutions à mettre en place. Afin d'uniformiser l'extraction des données chez les différents industriels, des paramètres génériques ont été mis en place. La figure 40 présente un aperçu l'ensemble de ces paramètres génériques.

III.5.3.2. Structuration des données Smart :

La deuxième catégorie de données concerne les données Smart. Celles-ci sont souvent générées à travers l'exécution d'un traitement, une interprétation, ou une agrégation, sur les données brutes à des fins d'amélioration, de simplification ou de visualisation de données. Parmi ces traitements, l'agrégation, adoptée dans ce projet, consiste à regrouper plusieurs éléments et paramètres avec une logique précise pour réduire la taille d'un fichier initial de données brutes et en générer un nouveau plus simple à exploiter. Plusieurs classifications des méthodes d'agrégation existent : stricte/semi stricte, par loi de probabilité ou par mini modèle d'apprentissage, ou par approche. Cette dernière permet de classer les méthodes d'agrégation comme suit : approche d'agrégation multicritères, approche d'agrégation de listes (d'ordonnements, de sur-classement, d'apprentissage d'ordonnements), approche d'agrégation de pertinence multidimensionnelle, ou l'approche d'agrégation de valeurs. L'agrégation de valeur englobe les moyennes arithmétiques et les mécanismes de combinaison linéaire classiques. Ainsi, les sommes, les variances, les moyennes, les régressions linéaires, les écarts types, les occurrences représentent tous des formes de valeurs agrégées [Moulahi, 2015].

Pour exécuter des opérations d'agrégation, un modèle est défini en fonction du domaine d'étude et des caractéristiques des données manipulées. Il permet d'identifier les paramètres intervenant dans les calculs : dans notre projet, les données Smart sont calculées suivant le modèle JPO (Jour, Programme, Outil) [Godreau, 2017]. Le principe de cette agrégation consiste à regrouper l'ensemble de données brutes (des données temporelles, des puissances, des vitesses, etc.) suivant les appels d'outils coupants dans un programme d'usinage précis exécuté durant une journée de travail précise.

En se basant sur les données agrégées suivant le modèle JPO, et en ayant recours à l'utilisation de plusieurs seuils mesurés au préalable avec méthodes de *data mining*, il est possible de calculer d'autres données Smart. Celles-ci permettent de renseigner les phénomènes malveillants pouvant apparaître en usinage : par exemple, le temps passé en dépassement de seuil du balourd ou de broutement, l'aire de dépassement de seuil, etc. Ces données sont utiles au calcul des indicateurs de performance clés, pour fournir des statistiques sur l'état de ressources utilisées, pour faire de la « Traçabilité », etc. La figure 42 regroupe les données agrégées suivant le modèle JPO.

Data JPO		
Données	Unité	Description
Jour	#	Numéro du jour de l'usinage
programme	#	Identifiant programme utilisé
outil	#	Identifiant outil utilisé
temps_total	s	temps total d'usinage
temps_arret_complet	s	temps total d'arrêt complet de la MOCN
temps_hors_matiere	s	temps total de l'outil hors matière
temps_en_matiere	s	temps total de l'outil en matière
temps_ebauche	s	temps passé en ébauche
temps_finition	s	temps passé en finition
temps_G1_sans_usi	s	temps d'avance de travail
Pmoy_complet	kw	Puissance moyenne consommée
Pmoy_usi	Kw	Puissance moyenne consommée en usinage
Vfmoy_usi	mm/min	vitesse d'avance moyenne en usinage
Nmoy_usi	tr/min	vitesse de rotation moyenne en usinage
Pmoy_rota_vider	Kw	Puissance moyne pour broche en rotation à vide
balourd_moy	m/s ²	Balourd moyen de l'outil calculé hors usinage
CO_balourd	mm ²	aire de dépassement de seuil
T_balourd	s	temps passé en dépassement de seuil
CO_broutement	mm ²	aire de dépassement de seuil
T_brout	s	temps passé en dépassement de seuil
CO_arms	mm ²	aire de dépassement de seuil
T_arms	s	temps passé en dépassement de seuil
CO_vrms	mm ²	aire de dépassement de seuil
T_vrms	s	temps passé en dépassement de seuil
Qcop	cm ³ /min	Débit copeaux
Vol_retire	L ou cm ³	Volume Retiré

Figure 42. Les données SMART « Modèle JPO »

La structuration formelle de ces données JPO, en utilisant le langage UML, est représentée dans la figure suivante. Nous avons établi une classification en fonction des types des paramètres agrégés.

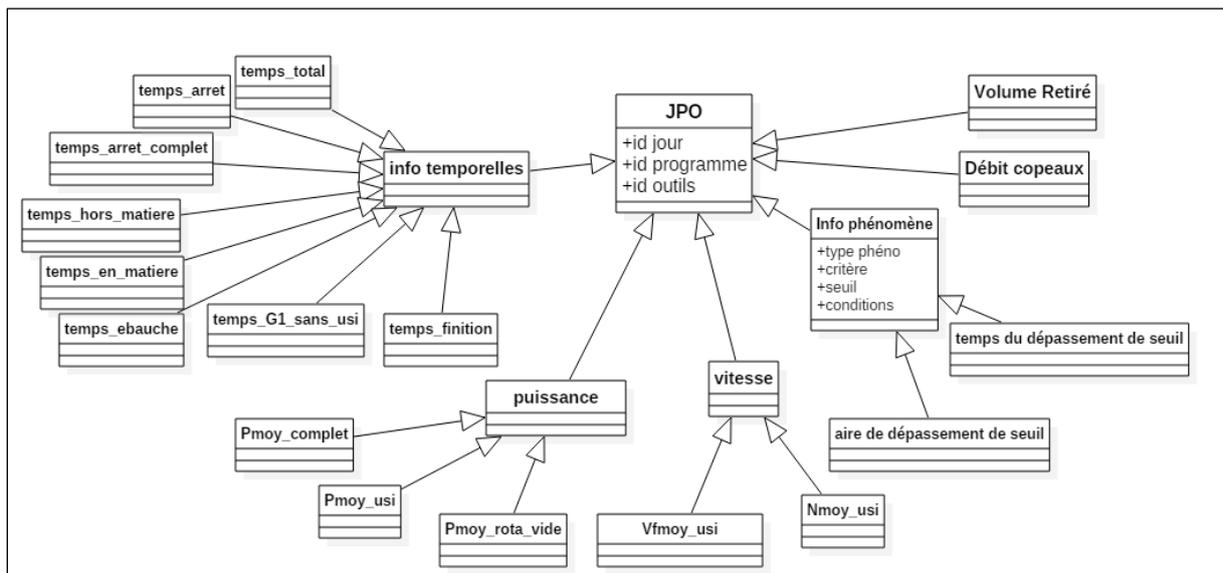


Figure 43. Structuration des données JPO

Les opérations d'agrégation et les méthodes de fouilles et d'analyse de données, exécutées pour l'obtention des données Smart, sont regroupées dans la catégorie « Smart_algorithme » et

classés dans la classe ressource logicielle. La génération des *Smart data* favorise l'enchaînement des opérations de calcul des KPI en utilisant les « KPI_algorithme ». La figure 44 représente le modèle de génération des données smart.

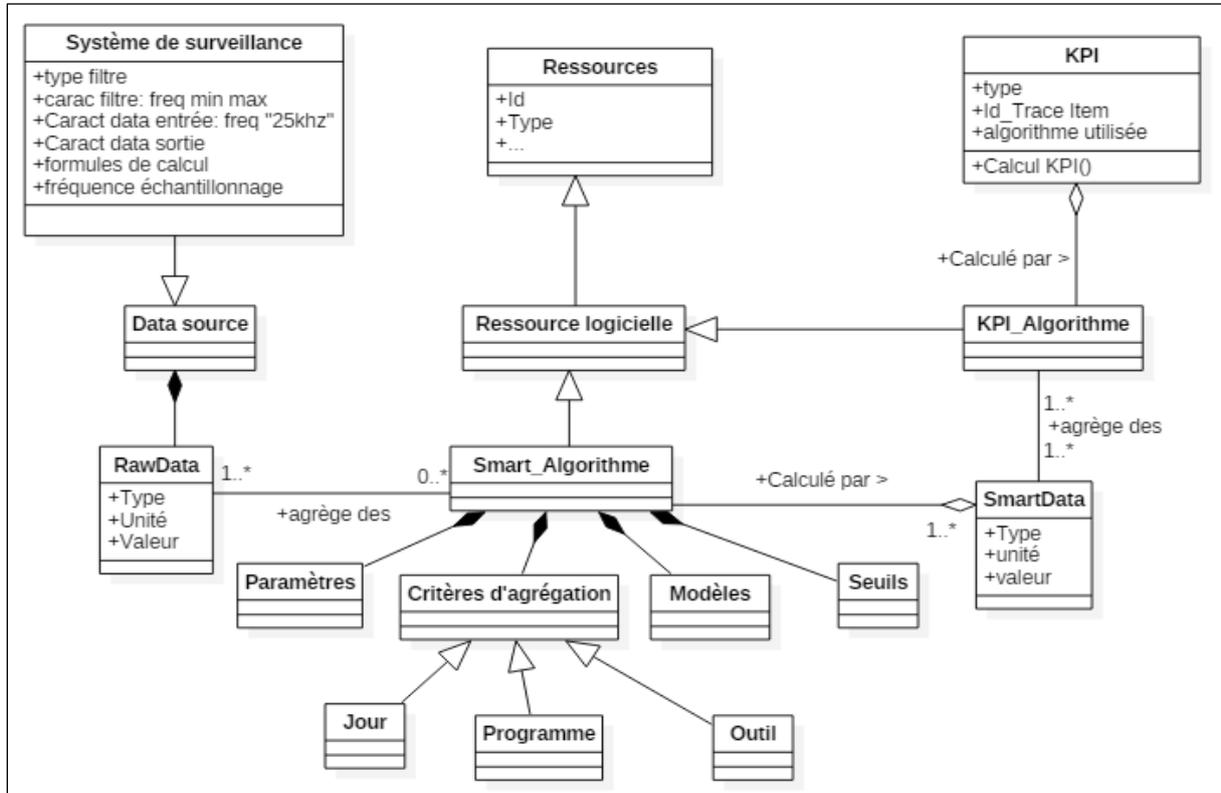


Figure 44. Modèle de calcul des données Smart

Cette classification et cette structuration des différentes données disponibles dans le cadre de ces travaux de recherche permet de définir les sous-modèles de données qui, ensemble, forment la première couche de la base de connaissances. La figure 45 représente l'interconnexion entre les sous-modèles de structuration des données.

III.5.3.3. Structuration des Bibliothèques : recours à la littérature et aux normes

Après avoir structuré des données industrielles et intégré des modèles dans la base de connaissances, et pour assurer la généricité des modèles développés, les premières bases collectées ont été confrontées et enrichies par des données complémentaires issues de l'étude de la littérature et par les normes et les standards. Plusieurs normes ont été étudiées dans un objectif de capitalisation de données, essentiellement issues du catalogue des normes ISO. La famille 25 de ce catalogue regroupe un ensemble de règles et définitions qui normalisent l'environnement des techniques de la fabrication⁷. La catégorie 25.080 des normes ISO a été très utile pour étendre la liste initiale des conditions de coupe, définie à travers l'étude des premières catégories de données décrites ci-dessus. De même, la catégorie 25.100 « Outils de coupe » et précisément la sous-catégorie 25.100.01 regroupent les différentes informations relatives aux outils, aux plaquettes, aux attachements, etc.

La **Erreur ! Source du renvoi introuvable.** présente un extrait des conditions de coupe et des caractéristiques d'outils coupants issus des normes ISO et capitalisés pour le développement des modèles de données.

Conditions de coupe		Outils			L'outil référencé
Ae engagement radial	mm	Caractéristiques	Unité	Description	Airbus
Angle de coupe	degré	Source (Airbus, mécachrome)			T 10006
Angle de direction d'arête Kr	degré	Référence de l'outil			Ventilateur LANG TECHNIK.de
Avance par tour f	mm/tr	type d'outil	#	Type de l'outil: fraise/foret; monobloc	
avance par dent fz	mm/dent	Type matière	#	Voir les taxo matières usinées	
Vf vitesse d'avance	mm/min	Diamètre	mm		330
D diamètre de l'outil	mm	longueur (Longeur plan de jauge)	mm		255
diametre taraud	mm	longueur utile (de sortie)	mm		229
diamètre foret	mm	Longueur de coupe	mm		165
Coefl spéc Kc	N/mm²	Nombre de dents	#		#
Durée de vie	min	Nombre de lèvres	#		#
Débit copeaux Q	cm³/min	Angle hélice	degré		#
État du surface Ra	µm	Angle de pointe	degré	Angle mesuré dans Pr entre l'arête de	#
Longueur arête de coupe	mm	Angle de dépouille	degré	C'est l'angle qui évite que l'outil ne talc	#
Longueur percée avant dété des bec	m	Angle attaque	degré	Il s'agit de l'angle entre l'arête de coup	#
Longueur usinée	m	Durée de vie	#	Durée de vie de l'outil	3939
lubrification	#	Revetement	#		#
N vitesse de rotation de l'outil	tr/min	Géométrie de pointe	#		#
Z Nombre de dents	#				
Nombre de passes	#				
Nombre de trou	#				
P pas de balayage en tréflage	mm				
pas	mm				
Profondeur de passe ap	mm				
Profondeur du trou	mm				
Profondeur tronçonnage P	mm				
Puissance de coupe absorbée Pc	KW				
Pénétration	#				
temps cycle	s				
temps usinage pièce	s				
Usure en dépouille Vb	mm				
Vitesses de coupe (Vc)	m/min				
Z nombre de dents de la fraise	#				

ISO 1832:2017

Plaquette

symboles

forme

dépouille normale

classe de tolérance

fixation et/ ou brise copeaux:

grandeur

épaisseur

configuration de la pointe

symboles facultatifs pour les plaquettes amovibles

condition d'arete

exécution de la pointe et l'appli de la plaquette (direction d'

Symboles supplémentaires pour les plaquettes brasées

condition d'arete

type d'arretes brasées ou monobloc

nombre de points brasés

longueur de l'arete brasée

Figure 46. Données Outils et conditions de coupe (à partir des normes)

⁷ Voir pour cela la partie Norme et Standards du chapitre précédent.

Pour la structuration formelle des bibliothèques de connaissances, nous nous sommes basés sur le méta-modèle global (Figure 33) et avons intégré le maximum de connaissances à partir de la littérature et des normes et standards (parmi lesquels les normes STEP-NC et MANDATE détaillés dans le chapitre 2). Ces connaissances standardisées sont essentiellement concentrées sur les concepts « Produit, Processus, Ressources » ce qui correspond aux deux premières couches de la base globale de connaissances (figure 34). La figure 47 présente un premier niveau de détail du modèle structurant ces bibliothèques.

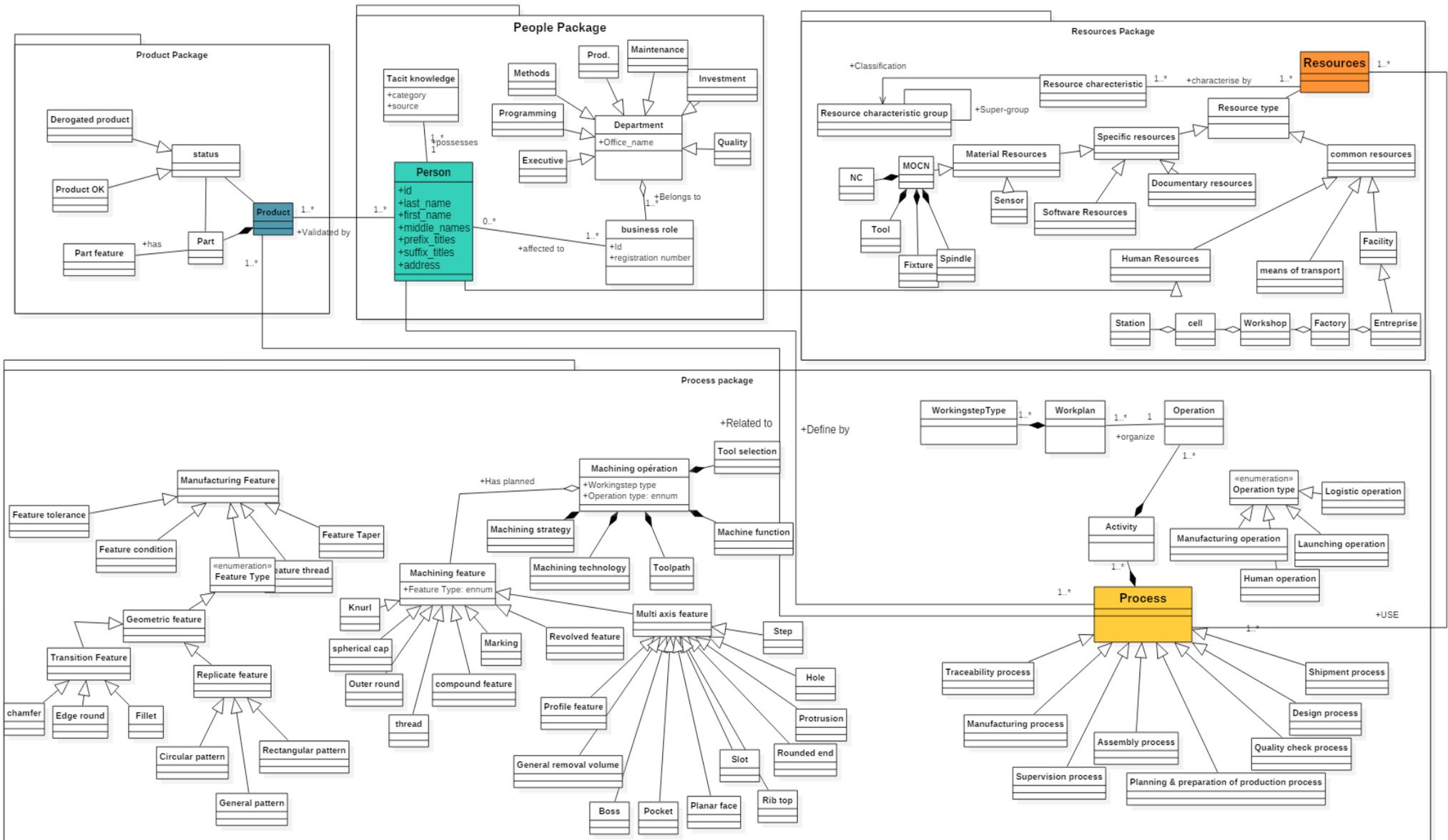


Figure 47. Modèle de structuration des bibliothèques

III.5.3.4. Capitalisation des KPI

Les connaissances industrielles peuvent avoir plusieurs formes⁸ et une étape d'analyse a permis de dresser une première liste de connaissances explicites (à savoir les statistiques, les critères et conditions de travail, les contraintes à respecter, les connaissances relatives directement aux produits, etc.). D'autres connaissances restent toutefois implicites, mais plusieurs techniques existent permettant l'externalisation, c'est à dire le passage de l'implicite à l'explicite. Ces moyens dépendent fortement des catégories et des types de connaissances à extraire. Parmi eux figure l'entretien avec des experts pour extraire des règles métiers. En revanche, pour extraire des connaissances statistiques, le recours à un processus d'ECD, et aux technologies du *data mining* sont plus efficaces. Au sein des connaissances statistiques, plusieurs connaissances restent cruciales à toute entreprise, comme par exemple les informations temporelles qui renseignent le rendement de l'entreprise (temps net, temps utile, temps requis, temps brut, etc.). Dans le domaine du pilotage industriel et du management de la production, plusieurs indicateurs sont ainsi calculés, tels que le TRE (un indicateur d'engagement des moyens de production), ou le TRS (un indicateur de productivité). Ce type de connaissances fait partie des indicateurs clés de performance permettant de mesurer l'efficacité des procédés par rapport à des objectifs précis. Les KPI sont essentiellement utilisés dans un objectif global commun : le management et l'amélioration continue [Meski et al, 2019b].

Pour des objectifs de suivi et surveillance de l'activité industrielle et d'amélioration, nous avons calculé plusieurs KPI pour faire de l'aide à la décision. Ils sont, essentiellement, destinés à répondre aux besoins managériaux spécifiques des différents départements d'une entreprise d'usinage. Un indicateur d'aide à la décision est défini comme une instanciation contextuelle d'un ensemble de KPI et de données Smart [Ritou et al. 2019]. La définition des KPI et la manière de sélectionner et d'agréger les données smart sont déterminées sur la base des connaissances industrielles telles que les règles métier et les meilleures pratiques. Dans ce projet plusieurs critères de détection des phénomènes malveillants et leurs seuils associés ont été mis en place.

Parmi les KPI relatifs aux critères de détection figurent la durée et l'intensité du dépassement de seuil. Un opérateur de criticité a été défini, et son idée générale est de mesurer le temps passé au-dessus d'un seuil fixé, et la hauteur du dépassement, chaque dixième de seconde.

⁸ Ces formes ont fait l'objet d'une description détaillée au chapitre 2.

Soit $X_i = \{x_i(k), k = 0..n\}$ la i -ième série temporelle représentant un critère de surveillance. L'opérateur de criticité sera alors défini par :

$$CO[X_i > T_{i,j}] = \sum_{k=tb}^{tf} Y(k) \text{ Avec } \begin{cases} Y(k) = x_i - T_{i,j} & \text{si } x_i > T_{i,j} \\ Y(k) = 0 & \text{si } x_i < T_{i,j} \end{cases}$$

Avec $T_{i,j}$ le seuil de criticité n°j appliqué au critère i , tb et tf les temps de début et de fin de chaque période.

Les seuils de détection et les modèles définissant les KPI et les données Smart constituent aussi de nouvelles connaissances formelles. La figure 48 présente un exemple de seuils utilisés dans le cadre de ces travaux de recherche et de KPI calculés avec les moyens d'analyse de données.

Seuils critiques identifiés par modélisations statistiques		
Critère	Seuils identifiés	unités
Usi	10	m/s ²
Broutement	20	m/s ²
Bris	7,25	m/s ²
Vrms	7	m/s ²
Arms	40	m/s ²

KPI	Description
FBPFO FBPFI	fréquences caractéristiques de défauts des roulements de la broche. Ces fréquences sont définies à partir d'un simple modèle cinématique dépendant des caractéristiques géométriques des roulements ([76, 77]). Ces fréquences correspondent à des défaillances sur les différentes parties des roulements comme, la piste extérieur (fBPFO) ou la piste intérieur (fBPFI) pour les plus importantes. Celles-ci peuvent aussi être modulées par la fréquence de rotation de la broche (fh1).
Fb	L'ensemble des fréquences liées au broutement, peut être décomposé en différents sous ensembles. Le premier issu du fondamental du broutement (fb1), fréquence principale du broutement, celle dont la fréquence est la plus basse, avec un ensemble de fréquences (fb1i) apparaissant suite à une modulation du broutement par la fréquence fondamentale du passage des dents (fc1). Puis d'autres ensembles correspondant aux harmoniques du broutement et leurs modulations (fb j i) toujours par rapport à fc1. Avec j représentant la j ème harmonique du broutement.

Figure 48. Exemple de connaissances (KPI + Seuil)

Pour faire le lien avec la première partie de ce chapitre, les indicateurs d'aide à la décision représentés ici constituent essentiellement les éléments de base des rapports de traçabilité. De ce fait, la définition d'un indicateur d'aide à la décision est obtenue en fonction de quatre paramètres d'instanciation, qui sont :

- L'objectif de la décision pour lequel les indicateurs d'aide à la décision sont définis ;
- Les éléments de contextes relatifs aux caractéristiques de machine, de pièce, de programme, d'outil ou d'autres entités ;

- Le mode d'instanciation qui déclenche le calcul des indicateurs d'aide à la décision. Trois modes sont considérés : périodique, sur demande et sur événement ;
- Le décideur qui utilise l'indicateur d'aide à la décision résultant selon un besoin métier spécifique.

L'humain possède un rôle très important dans ce système et peut aussi bien être l'utilisateur final du Framework d'aide la décision (c'est le cas des opérateurs et employés de l'atelier de fabrication ou des bureaux supports) que la source de connaissance et d'intelligence nécessaires pour alimenter le système (c'est le cas des décideurs ou des expert métiers).

III.5.3.5. Construction des règles métiers

Les règles métier sont généralement classées dans la catégorie des connaissances explicites. D'une part, ces règles permettent de définir et préciser les contraintes d'un processus particulier ; D'autre part, elles permettent le contrôle des processus industriels et la prise de décision sur la base des critères clairement définies, non ambigus et compréhensibles par tous. La structuration de ces bases de règles métier commence par une phase d'acquisition basée sur l'expertise et l'expérience des acteurs métier. À la fin de cette phase, un maximum d'expertise est récupéré. Les informations sont ensuite traitées et qualifiées pour éliminer la redondance et vérifier la cohérence. La troisième phase consiste à traduire ces règles dans un langage formel, numérique, que le moteur de raisonnement peut comprendre et utiliser [Meski et al, 2019b].

Les règles métiers sont très utilisées dans le domaine industriel et permettent de réduire les risques ou détecter les problèmes. Elles sont facilement intégrées dans les systèmes d'aide à la décision et permettent de gagner du temps et de réduire les coûts lors du développement d'une solution. Les règles métier sont autant utiles dans la gestion de projet, que dans le domaine de l'informatique, de la communication, de l'interopérabilité, de la sécurité, de l'étude des risques, ou de la perception de l'environnement. Par ailleurs, elles sont très utilisées pour l'exploitation des connaissances stockées, pour la génération de nouvelles connaissances et pour l'enrichissement des règles métiers préalablement définies.

Dans nos travaux, la prise en compte de l'interprétation de l'expert lors de l'intégration des règles métiers dans le modèle de connaissance nécessite une étape intermédiaire. Celle-ci consiste en une première présentation des résultats aux experts sur des cas spécifiques d'analyse et de diagnostic de phénomènes malveillants.

Pour mieux illustrer l'utilisation des règles métiers dans le cadre de ces travaux, prenons un exemple. Pour déterminer la cause d'apparition d'un bris d'outil (exemple de scénario de

diagnostic présenté table 1), une règle très simple peut tester si l'état de l'outil est l'origine du problème :

« L'outil est la cause du bris lorsque les conditions de coupe mesurées sont différentes des conditions de coupe de référence. »

Dans le domaine de l'usinage, le bureau des méthodes définit toujours des conditions de coupe de référence qui doivent être prises en considération au moment de la préparation du code ISO. Pendant la phase de fabrication, en utilisant les systèmes de surveillance, il est possible de mesurer les valeurs des conditions de coupe réelles. Parfois, en raison de la dégradation de l'état de l'outil, une variation entre les valeurs réelles et les valeurs de références apparaît. Dans ce cas, l'état de l'outil est la cause potentielle d'un problème de bris. Si tel n'était pas le cas, l'étude de l'état de la pièce, des caractéristiques du programme utilisé ou de l'état des autres ressources serait nécessaire. Pendant la phase d'implémentation cette règle sera formalisée en utilisant un langage spécifique qui pourra ainsi exécuter le raisonnement et assurer la comparaison.

Le recours aux règles métier est fortement relié à notre approche d'intégration des données avec les connaissances. Par exemple, la règle présentée ici s'appuie sur la base de connaissances pour y récupérer les valeurs des conditions de coupes de référence qui sont stockées dans les bibliothèques de la première couche. Par la suite, l'exécution de la règle nécessite de rechercher les conditions de coupes réelles dans les bases de données brutes ou Smart. Le résultat de l'exécution permet d'instancier les différents éléments d'un cas de diagnostic dans la base de connaissances.

Le développement des modèles de connaissances a permis de mettre en évidence l'importance des relations sémantiques entre les concepts manipulés et le fort recours aux subsomptions. La capitalisation d'un grand nombre de connaissances hétérogènes du domaine industriel génère une représentation, pour les concepts formant les modèles, en treillis ou en. D'autre part, le besoin de faire des inférences, de développer et d'exécuter des règles métiers et de réintégrer des nouvelles connaissances contraint le choix des méthodes d'implémentation de la base de connaissances. Finalement, en raison de ces contraintes et parce que la base de connaissances est développée pour être utilisée dans un système d'aide à la décision, une analyse bibliographique des moyens et techniques de structuration et de développement des bases de connaissances s'est avérée nécessaire. Les solutions ontologiques se sont révélées essentielles au contexte de ces travaux de recherche.

III.6. CONCLUSION DU CHAPITRE :

Ce chapitre a détaillé l'approche scientifique proposée. L'architecture mise en place pour l'implémentation des axes d'aide à la décision a été expliquée et la méthode de classification et modélisation des données a été présentée. Suite à la structuration de données et de connaissances industrielles, et afin de valider l'utilité des modèles développés, il est important de les implémenter dans des cas d'études réels. Plusieurs scénarios d'aide à la décision, essentiellement pour le diagnostic et ayant recours à l'intégration données/connaissances seront ainsi présentés dans les chapitres suivants.

CHAPITRE 4 : DEMONSTRATEUR ET BESOINS D'IMPLEMENTATION

**"On demontre qu'une chose
est consequence d'une autre.
pour cela, on construit la
consequence avec
l'hypothese."**

De Edmond Goblot

IV.1. INTRODUCTION DU CHAPITRE

Ce chapitre décrit le démonstrateur, développé dans le cadre du projet SmartEmma sous forme d'un système multi-agents « SMA ». L'implémentation des deux premiers axes d'aide à la décision sera également détaillée. Ce projet étant collaboratif, nous définirons la contribution de ces travaux de recherches à l'avancement de SmartEmma dans les développements. Enfin, les besoins technologiques, nécessaires pour la construction de l'agent de diagnostic et l'implémentation des scénarios d'aide à la décision, seront spécifiés.

IV.2. DESCRIPTION DU DEMONSTRATEUR SMARTEMMMA

Rappelons que ces travaux de recherche sont réalisés dans la cadre du projet ANR SmartEmma qui a pour objectif de développer un Framework d'aide à la décision à base de connaissances. Plusieurs acteurs académiques et industriels collaborent pour concevoir et développer les différents modules qui composent ce système. La stratégie d'aide à la décision, basée sur les quatre axes identifiés dans le projet et dont la méthodologie a été présentée dans le chapitre précédent, doit être complétée par une stratégie de pilotage des données. Après une analyse des technologies de développement pour ce SAD, la solution choisie pour piloter les données et connaissances et pour gérer les différents composants de la plateforme SmartEmma est basée sur la technologie des systèmes multi-agents. Ce système de pilotage permet de résoudre les problèmes d'interopérabilité et de modularité face à l'hétérogénéité des solutions technologiques utilisées et la diversité des axes d'aide à la décision.

IV.2.1. Définition de système multi-agents

Les travaux de [Monostori et al., 2016] ont traité l'implémentation et les axes de développement des Systèmes Cyber-Physiques (ci-après CPS, *Cyber-Physical System*) dans le contexte manufacturier vis-à-vis de l'informatique, les technologies de l'information et de la communication, et les technologies des sciences manufacturières. Dans l'objectif est de comprendre les techniques actuelles de développement et d'implémentation d'un système CPS, une recherche bibliographique statistique des mots clés liés à ces systèmes a été faite. La représentation graphique de relations liant les mots clés étudiés est représentée dans la figure 49, ci-dessous. Cette représentation montre que les systèmes multi-agents sont l'une des techniques les plus utilisées pour la conception et les développements des CPS, tout comme l'IOT, les ontologies, etc.

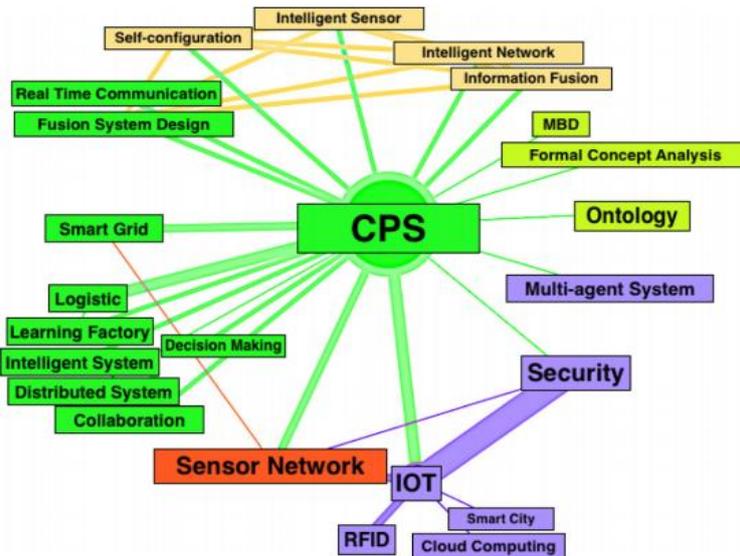


Figure 49. L'arborescence des mots-clés recherchés et leur relation [Monostori et al., 2016]

La littérature propose différentes définitions d'un agent et pour ne pas lister toutes celles-ci, nous reprenons celle de [Norvig, 1995] qui se rapproche de notre contexte :

An agent is defined as anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors. It has three main characteristics: autonomy, sociable (communication), reactivity and proactivity.

Un agent peut être une entité physique ou virtuelle. Il peut s'agir d'un système expert, d'un opérateur humain, d'une machine ou d'un algorithme par exemple. Dans une démarche de coopération intelligente entre l'opérateur humain et la machine, nous nous intéressons plus particulièrement aux agents informatiques virtuels définis par des algorithmes d'aide à la décision.

Qu'est-ce qu'un agent informatique virtuel ? Un agent informatique peut être considéré comme un objet (dans un système orienté-objet). Cependant, alors qu'un objet composé d'attributs et de méthodes exécute automatiquement une méthode lorsque cette dernière est appelée, un agent est autonome. L'agent a un aperçu partiel de son environnement, des objectifs qui lui sont propres et peut décider d'exécuter ou non un traitement (une fonction/méthode) selon sa propre stratégie. En plus de cette autonomie de prise de décision, un agent intelligent devrait posséder les compétences suivantes [Weiß, 2001] :

- Réactivité : il doit s'adapter aux changements ;
- Proactivité : il possède des processus internes qui fonctionnent même en absence de sollicitations externes ;
- Sociabilité : capable de coopérer, communiquer et négocier.

Pour faire simple, un agent informatique perçoit son environnement, interprète ces perceptions, puis exécute des actions pour atteindre les objectifs pour lesquels il a été conçu. Un SMA est représenté par un système d'agents distribués avec une considération stratégique de communication et coopération entre ses entités (agents).

Un premier démonstrateur de système multi-agents a été développé dans le cadre du projet SmartEmma. La construction de ce système d'aide à la décision est faite à travers une approche ascendante (*bottom-up*) agile. Il s'agit de définir et d'implémenter des fonctions élémentaires sur la base de scénarios opérationnels dont la complexité est augmentée progressivement. L'objectif de cette approche est de valider et de consolider les idées proposées au fur et à mesure avec les partenaires industriels pour assurer l'efficacité et la robustesse de la plateforme à moyen et long terme. La description de l'architecture de ce système sera détaillée dans la partie suivante.

IV.2.2. Architecture fonctionnelle de la solution multi-agents dans SmartEmma

L'analyse bibliographique (voir chap. 2) a montré un réel intérêt pour l'utilisation de la technique des systèmes multi-agents, qui a justifié le choix technologique du développement des différents modules de notre Framework global.

La solution que nous proposons étant focalisée principalement sur le pilotage et l'aide à la décision, elle a le rôle principal de percevoir l'environnement (l'équipement ou le logiciel) dans chacune des couches qui composent l'architecture, de comprendre les requêtes reçues des équipements et des autres agents, puis de réagir suite à leur interprétation. Nous proposons dans la figure 50, ci-après, l'architecture globale de pilotage de SmartEmma. Cette solution a pour objectif premier de résoudre les problèmes d'interopérabilité entre les différents éléments qui se trouvent dans une entreprise d'usinage. Elle est également capable de s'adapter à l'évolution des équipements/logiciels de l'entreprise. Cette architecture s'articule autour de trois couches comme le montre la figure 50 :

- La première regroupe les équipements physiques tels que la machine-outil et la commande numérique avec les capteurs installés. L'équipement principal de cette couche est le système de surveillance « EmmaTools » qui récolte les signaux réels, les agrège et les stocke dans la base de données brutes « BD EmmaTools » ;
- La seconde couche est composée de la base de données EmmaTools qui est en charge de l'interconnexion avec la couche supérieure. Cette couche regroupe les applications métiers hétérogènes, telles que les outils de conception et de simulation, les données du

système d'information, les règles métiers et les algorithmes smart, et surtout la base de connaissances globale et centralisée ;

- Enfin, toutes ces données et applications sont pilotées par un système multi-agents permettant l'interconnexion entre eux et qui forment la troisième couche. Le système multi-agents n'est pas utilisé pour faire du calcul exhaustif, mais uniquement pour le pilotage du processus d'aide à la décision ou de prise de décision. Il permet de récupérer l'information pertinente auprès de l'entité responsable et de lancer l'algorithme/outil de traitement opportun.

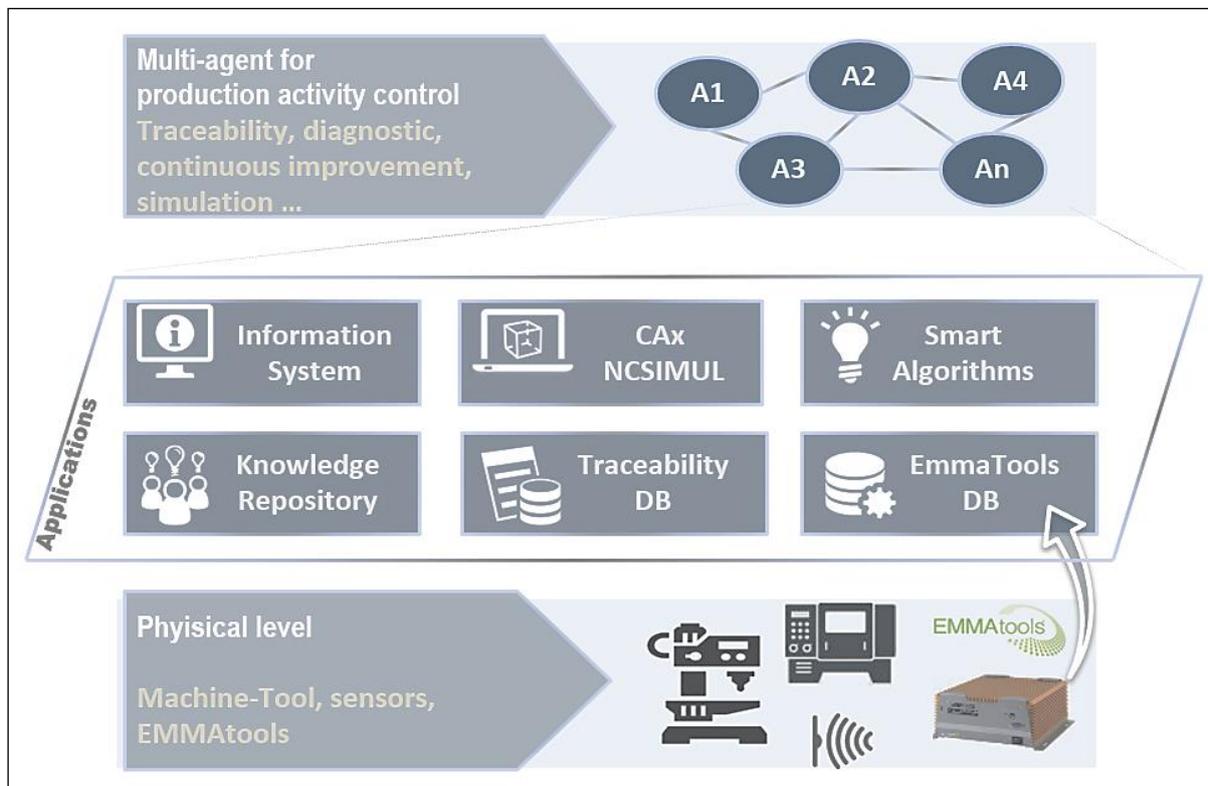


Figure 50. Architecture proposée pour le système multi-agents [Livrable Projet]

IV.2.2.1. Liste des agents

Comme montre la figure 51, plusieurs agents ont été développés au fil du temps. En raison de la spécificité et des contraintes de développement de ce système, le fonctionnement des agents s'appuie de façon directe sur la base de connaissances globale. La liste des agents est, évidemment, extensible et peut être complétée au fur et à mesure en fonction des besoins des utilisateurs. Parmi les agents déjà implémentés dans ce système, on distingue :

- Agent de communication : gère l'interface pour paramétrer l'analyse et envoie les paramètres à l'agent d'aide à la décision. Il permet aussi d'afficher les résultats

graphiques dans l'interface utilisateur du démonstrateur et d'envoyer les résultats à une personne spécifique ou une équipe par mail ;

- Agent d'aide à la décision : décide du scénario à exécuter et transmet les noms des scripts et arguments à utiliser à l'agent de calcul ;
- Agent de calcul « *Computing* » : lance tous les algorithmes relatifs à un traitement statistique, les algorithmes d'agrégation et d'analyse de données, de calculs statistiques, de calcul de KPI, etc. ;
- Agent de base de données « DB » : assure la communication avec les bases de données. Il permet d'interroger les bases en lecture pour la récupération de données et en écriture pour enregistrer les résultats de calcul. L'agent BD notifie également l'agent de communication de la disponibilité des données ;
- Agent de base de connaissances « KB » : assure la communication, en lecture et en écriture, avec la base de connaissances et la base de règles métiers.

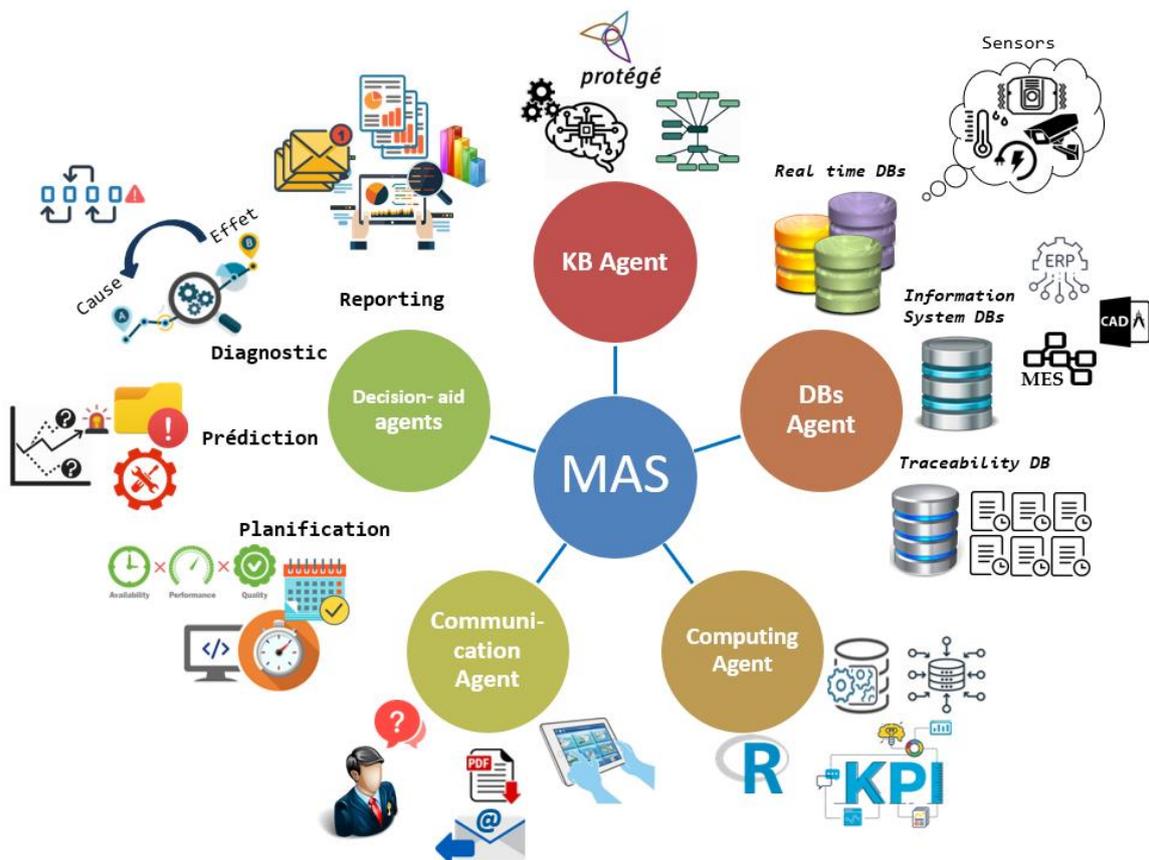


Figure 51. Liste des premiers agents implémentés dans le Framework SmartEmma, MAS vaut pour *Multi-Agent System*. [Livrable projet]

Parmi la liste des agents d'aide à la décision, intéressons-nous à l'agent de diagnostic. Il permet d'analyser les phénomènes malveillants qui peuvent apparaître en usinage et de déterminer la

cause d'appariation suivant un raisonnement à base de règles et le recours à une conciliation des données avec les connaissances métier. Cet agent est en collaboration continue avec les autres agents décrits plus haut. Il est surtout, comme évoqué dans le chapitre précédent, en communication continue avec la base de connaissances.

IV.3. LE PERIMETRE DES IMPLEMENTATIONS

Pour résumer, du point de vue de l'implémentation technique et du développement du démonstrateur, nos travaux de recherche se focalisent sur deux contributions majeures :

- L'implémentation de l'approche conceptuelle détaillée dans le chapitre 3, et ceci se manifeste à travers la mise en place des bases de données et de connaissances permettant le développement des deux premiers axes d'aide à la décision « le diagnostic et la traçabilité » ;
- L'implémentation de l'agent de diagnostic et l'exécution de quelques scénarios de test (détaillée dans le chapitre 5).

L'implémentation est très contrainte. Il faut en effet d'une part considérer le fait que le diagnostic est basé, en partie, sur le processus de traçabilité, et d'autre part il faut faciliter l'enregistrement, l'accès et l'interrogation des bases de données et de connaissances. Ainsi, pour réussir la mise en place de ce système, il est crucial de choisir les langages et outils d'implémentation les plus efficaces pour structurer les données et les connaissances dans les différentes bases. De nombreuses solutions technologiques de développement existent et leurs utilités varient en fonction des besoins d'aide à la décision et des contraintes techniques, telles que la technologie d'implémentation des autres modules ou le type de système de raisonnement. Dans ce cadre contraint, la technique « *file-based* » est utilisée pour la structuration des données smart sous forme de fichier CSV (*Comma-Separated Values*). Pour la base de traçabilité, une base de données relationnelle SQL est recommandée. La base SQL prend en charge le lien sémantique entre toutes les données traitées et permet une navigation et une interrogation plus faciles, en connexion avec la base de connaissances. Quant à cette dernière, elle est implémentée à l'aide du langage de représentation des ontologies OWL (*Web Ontology Language*). Les approches de modélisation ontologique présentent en effet l'avantage de fournir un raisonnement sémantique en permettant la réutilisation des connaissances existantes, l'inférence et la déduction de nouvelles connaissances, et la simplicité de la mise à jour de la base de connaissances. Comme montre la figure 52, le recours à ces technologies de développement facilite l'interaction en lecture et en écriture avec les bases.



Figure 52. Répartition des bases et communication avec le SAD.

La séparation nette entre la base de connaissance et la base de traçabilité, ainsi que le recours à l'utilisation d'une base SQL permet de ne pas surcharger la base globale. Toutefois, vu l'importance des rapports de traçabilité pour le développement du deuxième axe de l'aide à la décision, la liste des paramètres tracés est stockée dans la base de connaissance, au niveau de l'ontologie. Ceci permettra de mettre en valeur l'aspect collaboratif entre les données et les connaissances tout en définissant une frontière entre ces deux niveaux d'abstraction.

IV.3.1. Développement de la base de traçabilité

Cette première base a été développée en utilisant le langage informatique SQL qui permet d'ajouter, de rechercher, de modifier et de supprimer des données et de définir des contraintes spécifiques dans des bases de données relationnelles. Pour la création des tables de la base de nous avons utilisé l'application phpMyAdmin. C'est un outil logiciel gratuit écrit en PHP, destiné à gérer l'administration de MySQL sur le Web. PhpMyAdmin supporte un large choix d'opérations sous MySQL et MariaDB. Les opérations fréquemment utilisées (gestion des bases de données, des tables, des colonnes, des relations, des index, des utilisateurs, des autorisations, etc.) peuvent être effectuées via l'interface utilisateur, tout en ayant la possibilité d'exécuter directement n'importe quelle instruction SQL [Site web phpMyAdmin].

La stratégie de développement du Framework global étant ascendante, la création de la base de données s'est effectuée en même temps que l'exécution des scénarios de test initiaux. Au fur et à mesure du déroulement de ces scénarios, les tables de la base de traçabilité se sont développées. L'objectif principal de cette implémentation consistait à suivre et valider l'efficacité du modèle de structuration de la base de traçabilité, tout en respectant les besoins industriels. L'exécution des scénarios de test a permis de mettre à jour le modèle et de rajouter les concepts complémentaires permettant progressivement d'obtenir une version complète et fiable, telle que définie dans la figure 37. La figure 53 montre l'implémentation informatique du modèle conceptuel global.

phpMyAdmin

Serveur: 127.0.0.1 » Base de données: smartemma

Structure SQL Rechercher Requête Exporter Importer Opérations Privilèges Procédures stockées Événements Déclencheurs Suivi Concepteur Colonnes centrales

Sans titre *

Nouvelle base de données smartemma

- Nouvelle table
- business_role
- department
- hard_resource
- jpo_traceability
- kpi
- mode_traceability
- person
- piece
- process
- program
- raw_data
- ref_attribute
- semma_algorithms
- smart_data
- Structure
- traceability_point
- sys

smartemma process

- id_process : varchar(30)
- type : varchar(50)
- id_of : varchar(30)
- characteristics : varchar(100)

smartemma program

- id_prog : varchar(30)
- nom : varchar(30)

smartemma piece

- serial_num : varchar(30)
- type : varchar(30)
- l_of : varchar(30)

smartemma hard_resource

- id_resource : int(11)
- nom_resource : int(11)

smartemma spindle

- number : int(11)
- code : varchar(100)
- date_first_used : date
- date_last_used : date

smartemma ref_attribute

- id_process : varchar(30)
- id_program : varchar(30)
- id_piece : varchar(30)
- id_MR : varchar(30)
- id_SR : char(30)
- id_traca : varchar(30)

smartemma business_role

- id_role : varchar(30)
- function : varchar(30)
- id_depart : varchar(30)
- id_traca : varchar(30)

smartemma person

- id : varchar(30)
- nom : varchar(30)
- prénom : varchar(30)
- id_dep : varchar(30)

smartemma department

- id_department : varchar(30)
- name : varchar(30)

smartemma traceability_point

- id_pt_traca : varchar(30)
- period_from : varchar(30)
- period_to : varchar(30)

smartemma mode_traceability

- id-modetra : varchar(30)
- demand : tinyint(1)
- event : tinyint(1)
- periodic : tinyint(1)
- id_traca : varchar(30)

smartemma raw_data

- Name : varchar(100)
- unit : varchar(100)
- value : varchar(100)
- id_traca : varchar(30)

smartemma smart_data

- Name : varchar(100)
- Type : varchar(100)
- Value : varchar(100)
- id_traca : varchar(30)
- id_algo : varchar(30)

smartemma semma_algorithms

- id_algo : varchar(30)
- name_algo : varchar(50)
- type_algo : varchar(50)

smartemma jpo_traceability

- period_from : int(11)
- period_to : int(11)
- type : varchar(100)
- key : int(11)
- value1 : double
- value2 : double
- value3 : double
- spindle_number : smallint(5) unsigned

smartemma kpi

- id : varchar(30)
- type : varchar(50)
- valeur : varchar(100)
- id_algo : varchar(30)
- id_traca : varchar(30)

Console de requêtes SQL

Faire Ctrl+Entrée pour exécuter la requête

```

>SELECT * FROM `jpo_traceability`
>SELECT * FROM `kpi`
>SELECT * FROM `semma_algorithms`

```

127.0.0.1:8080/eds-modules/phpmyadmin470x180702135334/tbl_structure.php?server=1&db=smartemma&table=...

Figure 53. Structuration de la base de traçabilité implémentée dans phpMyAdmin

Pour la représentation des connaissances, nous avons choisi une solution basée sur les ontologies. Elle sera traitée dans la partie suivante.

IV.3.2. Développement de la base de connaissances : recours aux ontologies

En se basant sur l'étape d'analyse, les modèles de connaissances ont été développés et enrichis avec un haut niveau de granularité et de détails en utilisant le paradigme de la modélisation ontologique. En effet, dans le contexte de ce travail, le cadre de modélisation présente plusieurs spécificités et exigences à prendre en considération, à savoir : la grande quantité de connaissances hétérogènes, le besoin d'intégrer toutes les relations et contraintes nécessaires pour exécuter des requêtes interdépendantes, la possibilité de développer des règles métier nécessaires pour les opérations d'inférences, pour le développement d'un raisonnement et pour assurer une aide à la décision. L'utilisation d'ontologies est une solution idéale pour répondre aux exigences ci-dessus.

Ce choix est motivé par plusieurs raisons, tout d'abord, les ontologies sont très efficaces pour la capitalisation d'un maximum de connaissances d'un domaine donné. Elles favorisent la réutilisation des connaissances stockées et la réintégration de nouvelles connaissances et ceci grâce aux techniques d'interrogation et de communication avec son contenu. Les moyens ontologiques sont aussi favorables à la manipulation de connaissances hétérogènes. Enfin, les ontologies assurent l'interopérabilité technique et sémantique et permettent ainsi d'éviter les problématiques de rupture de la chaîne numérique.

L'ontologie proposée est structurée autour de plusieurs couches (correspondant à la structuration présentée dans la figure 34), formalisées en fonction de la spécificité, de la nature et de l'origine des connaissances acquises. Pour construire l'ontologie, nous avons eu recours au logiciel « Protégé » qui permet de définir les « Classes » représentant les différents objets de la base. Il permet aussi de déclarer les « Propriétés d'objet », qui représentent les relations entre deux ou plusieurs concepts, et « les propriétés de données », qui définissent la liste des attributs de chaque classe. La figure 54 montre la formalisation d'une simple relation de type « Un professeur enseigne une matière » dans le logiciel « Protégé ».

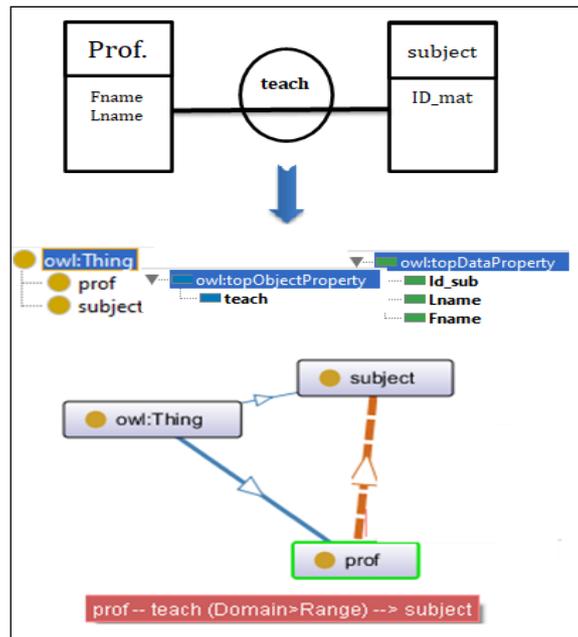


Figure 54. Exemple d'implémentation au niveau du logiciel « Protégé »

Ce logiciel fournit également des options permettant de préciser des contraintes, des exigences et des restrictions. En outre, Protégé met en place un ensemble de fonctionnalités plus avancées permettant d'interroger l'ontologie pour lire des données et des connaissances, ainsi qu'une interface pour développer des règles d'inférences permettant de générer de nouvelles connaissances et les intégrer dans l'ontologie.

IV.3.2.1. Première couche

La figure 55 montre un aperçu de classes de la 1ère couche de la base de connaissances.

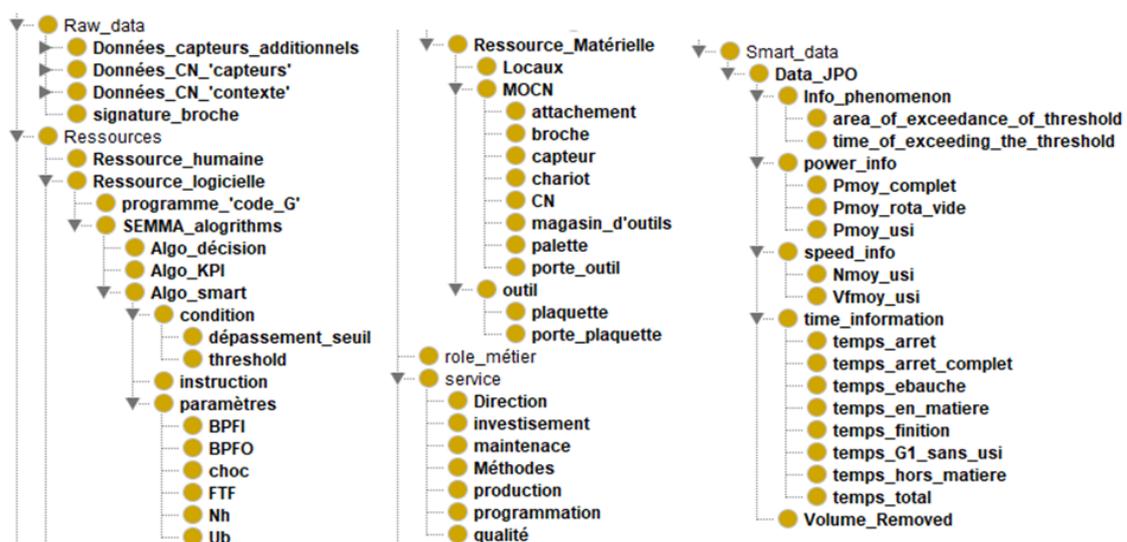


Figure 55. Extrait de la structuration de la base de connaissances

Le graphe ontologique de cette structuration est représenté dans la figure 56.

Un aperçu à petite échelle de la hiérarchisation est représenté par la figure 57 qui montre le chemin ontologique de la racine jusqu'au concept « Outil coupant » qui possède des sous éléments {plaquette ; porte plaquette} et des instances {Outil_22 ; Outil_26} appelées aussi individus dans le jargon de Protégé. La distinction des couleurs au niveau des flèches de liaisons est due aux différents types de relations entre les éléments : une relation entre classe et sous-classe (lien bleu) et ainsi différenciée d'une relation entre classe et individu (lien violet). Il est important de signaler que les relations ontologiques dans ces modèles ne sont pas toutes de type « Père-fils », mais peuvent être de plusieurs autres types. Dans cet extrait, la classe « outil » possède un premier lien de parenté avec la classe « Ressources Matérielles », et aussi une connexion avec la classe « Contexte » qui vient chercher des caractéristiques d'outils pour des fins d'aide à la décision.

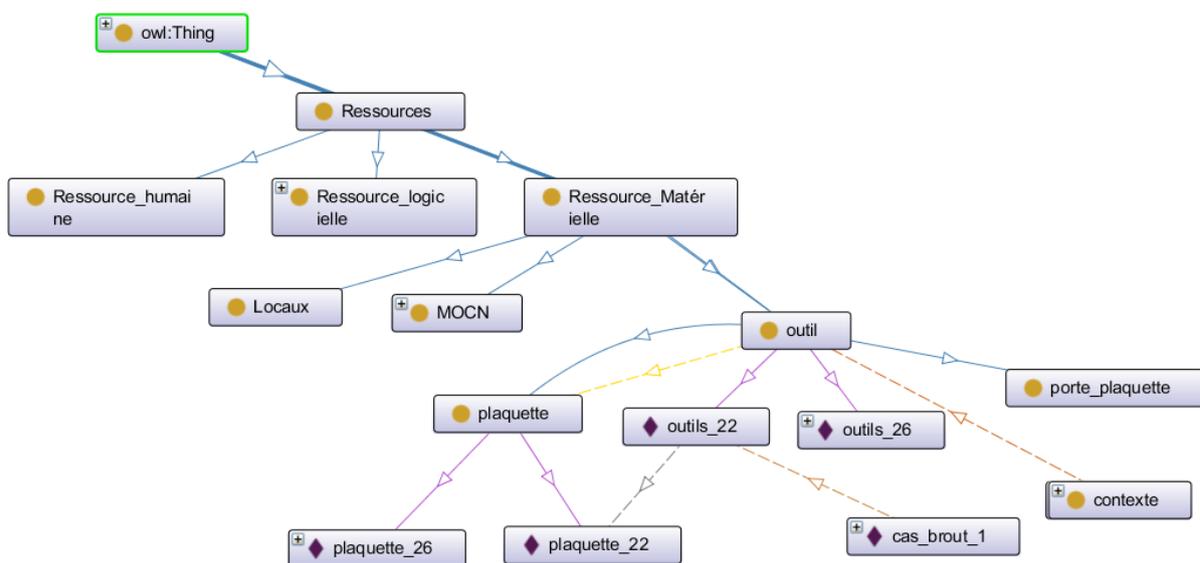


Figure 57. Aperçu hiérarchique de l'ontologie

Une autre vue du logiciel Protégé est présentée dans la figure 59. Dans cette interface d'assertion de propriétés, on peut définir les attributs associés à un individu spécifique et les relations qui le relie à d'autres individus. Dans cet exemple, on observe les attributs et caractéristiques de l'individu « Outil-22 » définis à travers les propriétés des données « Data property », et fixant la liste des caractéristiques et identifiants de l'outil ainsi que la liste de ses conditions de coupes (dont les valeurs ont été floutées pour des raisons de confidentialité). La propriété d'objet « Object property », définit la relation qui associe l'individu « plaquette_22 » à « l'outil 22 ».

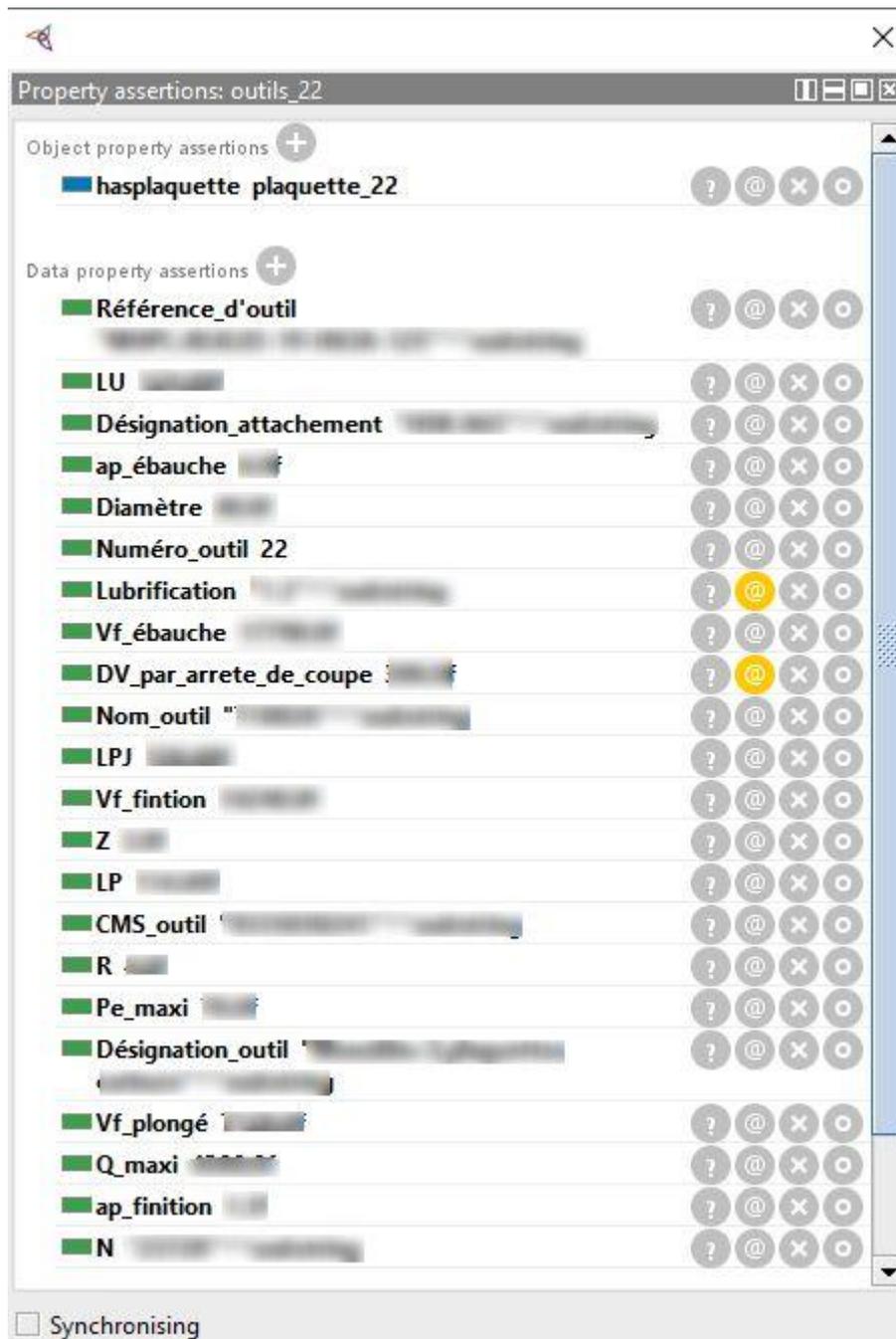


Figure 58. Interface d'assertion des propriétés

IV.3.2.2. Deuxième couche

Parallèlement à cette première couche, et afin d'assurer la généralité de la base de connaissances, plusieurs normes et standards ont été analysés et les plus appropriés ont été pris en compte pour structurer la deuxième couche du modèle ontologiques. Parmi ces normes, la partie 10 de la norme STEP-NC⁹ est centrale (14649-10 : Modèle de données pour les

⁹ Pour rappel, STEP-NC est une norme d'échange de données pour la programmation des commandes numériques et pour la communication entre les différentes parties de la chaîne numérique : conception, industrialisation et fabrication.

contrôleurs numériques informatisés — Données des procédés généraux). Elle comporte tous les paramètres, noms et concepts liés aux opérations, les caractéristiques de fabrication ainsi que l'ensemble des relations qui les regroupent. L'intégration de cette norme dans la base de connaissances globale consiste à transformer les diagrammes EXPRESS-G en classes, relations et attributs d'ontologie. La transformation de cette norme en ontologie a permis de développer plus d'un millier d'axiomes (figure 59).

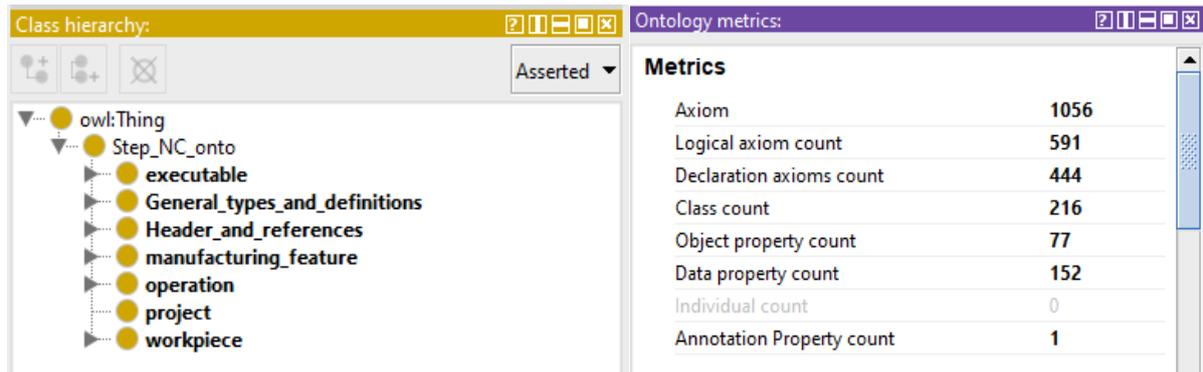


Figure 59. Intégration du STEP-NC dans l'ontologie

Cette structuration en couche permet la flexibilité et l'adaptabilité de la base de connaissance pour une réutilisation dans d'autres applications connexes, il suffit de remplacer une couche par une autre. Ainsi de nouveaux usages dans d'autres domaines peuvent être envisageables. L'intégration de cette base de connaissance dans le système global permet donc de faciliter et d'assurer la communication entre les différents niveaux de l'entreprise, des bureaux jusqu'à l'atelier en passant par le système d'information.

IV.3.2.3. Troisième couche

Pour répondre à la problématique de l'implémentation et de l'intégration de l'aide à la décision et suivant la structuration globale de la base de connaissances représentée dans le chapitre 3, la troisième couche a été ajoutée pour assurer la structuration des cas suite à chaque opération de diagnostic. Cette troisième couche a été implémentée sous forme du quadruplet « Contexte, Cause, Effet, Solution » comme montre la figure 60.

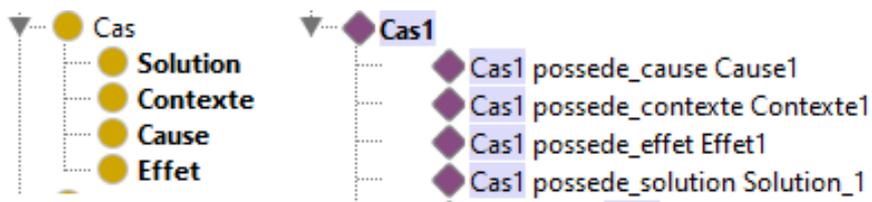


Figure 60. Structuration de la troisième+ exemple d'instanciation

L'efficacité de la solution proposée a été testée et implémentée dans le Framework global d'aide à la décision à travers plusieurs scénarios de diagnostic, testés dans le cadre du développement de l'agent de diagnostic. Ces scénarios seront détaillés dans le chapitre suivant, tandis que les besoins techniques et les technologies utilisées pour le développement de ces scénarios sont détaillés dans la partie suivante.

IV.4. ASPECTS TECHNIQUES ET INTERFACE UTILISATEUR (IU)

Pour le développement du SAD pour le diagnostic, plusieurs technologies et langages de programmation ont été utilisés. Ils sont représentés dans la figure 61.

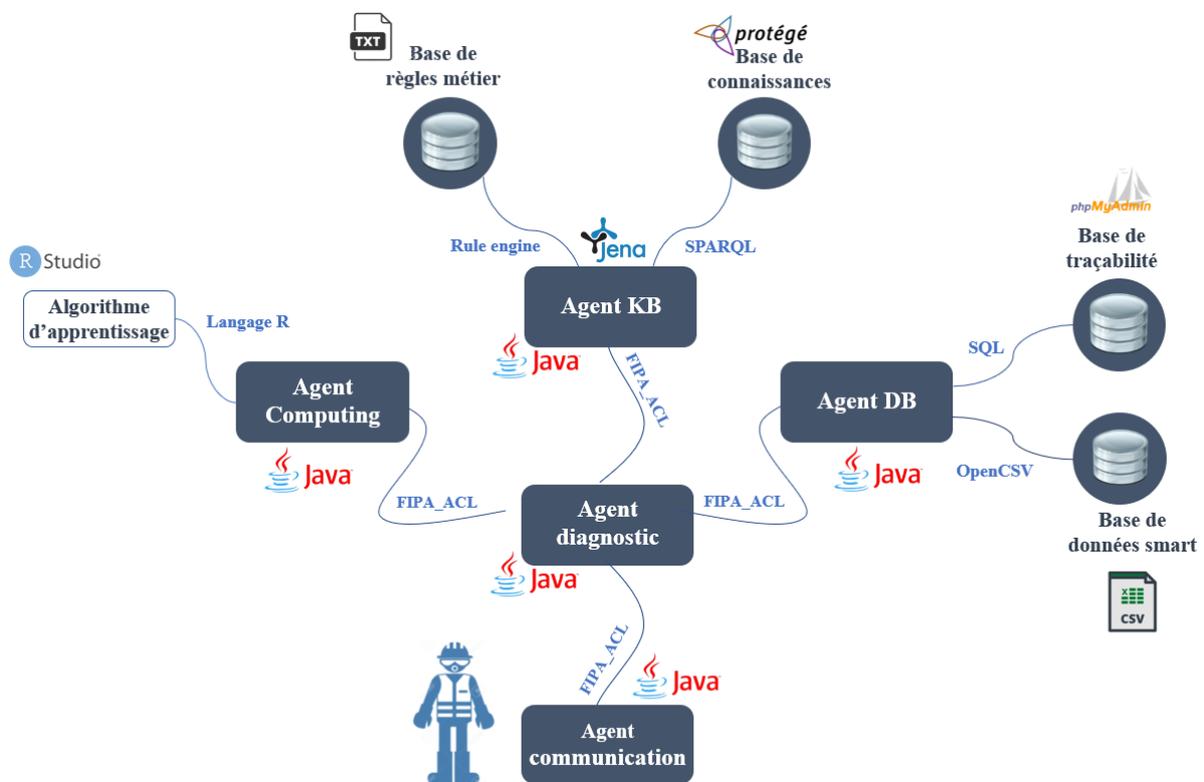


Figure 61. Technologies de développement

Pour la mise en place du système multi-agents global, le recours au langage « Java » a permis de développer un premier ensemble d'agents. Pour assurer une communication entre ces différents agents, nous avons eu recours à plusieurs normes et protocoles de communication. Les deux normes les plus utilisées pour définir l'encodage et gérer le transfert de messages entre agents sont les ACL FIPA « *Agent Communication Language* » proposé par la *Foundation for Intelligent Physical Agents* et KQML « *Knowledge Query Meta Language model* » [Soon et al. 2019]. Les deux langages partagent les mêmes concepts de base et les mêmes principes observés. Un message KQML et un message FIPA-ACL sont ainsi syntaxiquement identiques.

En revanche, les deux langages diffèrent principalement dans le détail de leurs cadres sémantiques :

- Les conditions préalables de faisabilité et les effets rationnels pour le FIPA-ACL ;
- Les conditions préalables, les conditions postérieures et les conditions de réalisation du KQML ;

Ils diffèrent également dans le choix et la définition des modalités qu'ils utilisent, c'est-à-dire dans le langage utilisé pour décrire les états des différents agents. Ils n'utilisent pas non plus la même technique de gestion des agents (inscription dans le système, localisation, appartenance à un groupe, etc.) [Poslad, 2007]. Bien que le KQML soit pertinent pour le transfert de messages entre agents, son exploitation directe dans la construction d'un système de coopération est très inefficace [Cost et al. 2000], ce qui explique le choix du FIPA ACL que nous avons utilisé.

L'agent diagnostic communique essentiellement avec trois agents : l'Agent KB, l'Agent DB et l'Agent Computing. Dans le cas de l'Agent KB, l'enjeu technologique était ici de trouver l'API la plus adaptée pour communiquer avec l'ontologie développée en OWL. La solution retenue est d'utiliser l'API *Jena*. C'est un Framework Java permettant de construire des applications web sémantiques et qui fournit un environnement de programmation pour :

- RDF : un modèle graphique pour décrire les ressources du web et leurs métadonnées, afin que ces descriptions puissent être traitées automatiquement. Développé par le W3C, le RDF est le langage de base du web sémantique ;
- OWL : un langage de représentation des connaissances construit sur le modèle de données RDF. Il fournit les moyens de définir des ontologies web structurées. C'est lui que nous avons utilisé pour le développement de notre ontologie ;
- SPARQL : un protocole qui permet de rechercher, d'ajouter, de modifier ou de supprimer des entités disponibles dans les ontologies RDF. Il a été utilisé pour la lecture des seuils de détection, appris au préalable, à partir de l'ontologie ;

Le sous-système d'inférence de Jena, appelé « Rule engine », est conçu pour permettre à une série de moteurs d'inférence ou de raisonneurs d'être connectés à Jena. Ces moteurs sont utilisés pour dériver des assertions RDF supplémentaires, qui sont issues de certaines RDF de base ainsi que de toute information ontologique optionnelle, et des axiomes et règles associés au raisonneur. Dans le cadre de ces travaux, il a été utilisé pour la formalisation et l'exécution des

règles métiers nécessaires à la détermination de la cause d'apparition d'un phénomène malveillant en usinage.

Pour l'exécution d'un processus de diagnostic, la communication avec l'agent de base de connaissances permet la récupération des connaissances nécessaires à savoir : les seuils appris, les règles métiers, les structurations et tous les éléments nécessaires pour l'opération de diagnostic.

La communication de l'agent de diagnostic avec l'Agent DB permet de récupérer les valeurs correspondantes au cas étudié et d'assurer l'instanciation des données. Cet agent permet d'accéder aux bases de données de surveillance, essentiellement Smart, à travers l'API Java « Opencsv ». C'est une bibliothèque d'analyseur CSV (*Comma-Separated Values*) d'utilisation aisée pour Java. Opencsv prend en charge toutes les opérations de base sur les fichiers CSV : lire un nombre arbitraire de valeurs par ligne, ignorer les virgules dans les éléments cités, configurer les séparateurs et les guillemets configurables, etc. L'Agent BD assure également la communication avec la base de traçabilité en utilisant le langage de requête structurée « SQL ».

Enfin, une troisième communication de l'agent de diagnostic avec l'Agent Computing est envisageable lorsqu'un apprentissage préalable des valeurs de références est nécessaire. Pour le développement des scénarios d'essai (détaillés dans le chapitre suivant), nous avons eu recours au langage R pour effectuer les opérations d'apprentissage automatique de quelques seuils de détection. R est, à la fois, un langage de programmation (de script) interprété, et un logiciel open source destiné aux calculs statistiques, l'analyse de données et à la science des données. Il est basé sur la notion de vecteur, ce qui le rend pratique pour les calculs mathématiques, et n'impose pas le typage des variables, ce qui permet de gagner en temps de développement.

Les différentes étapes du processus de diagnostic et la communication entre les différents composants et agents du système sont détaillées dans le diagramme de séquence en figure 62.

Une opération de diagnostic commence quand l'utilisateur introduit une demande au niveau de l'UI (*User Interface*, interface utilisateur) qui est gérée par l'« Agent de Communication ». Immédiatement ce dernier envoie la configuration et les paramètres spécifiés par l'utilisateur à l'agent de diagnostic. Deux alternatives se présentent et la première est très simple : si le phénomène existe déjà dans la base de traçabilité, l'Agent Diagnostic sollicite l'Agent BD pour récupérer les caractéristiques et le contexte du phénomène.

La deuxième alternative a lieu lorsqu'une phase de détection est nécessaire. L'Agent Diagnostic demande alors à l'Agent KB de récupérer les valeurs des seuils de détection de l'ontologie. S'il est nécessaire de déterminer les valeurs de références pour établir un raisonnement, une phase d'apprentissage automatique est enclenchée en communiquant avec l'Agent Computing. D'autre part, une autre sollicitation de l'Agent BD permet de récupérer les valeurs des paramètres de détection et, en les comparant avec les seuils, de détecter le phénomène et de le caractériser. À l'issue de cette phase, il est possible de débiter le raisonnement à base de règles.

L'agent de diagnostic doit alors choisir la règle métier adéquate en fonction de la demande du client, du contexte et des contraintes spécifiques. Le choix d'une règle appropriée est crucial dans la détermination de la cause du phénomène. Finalement, une étape importante est exécutée, celle de la communication, en écriture, avec l'Agent KB. Cette étape permet d'instancier le nouveau cas dans l'ontologie avec ses différents éléments de structuration : son contexte, son effet, sa cause et éventuellement la solution proposée. Par la suite, une simple interrogation de l'ontologie permet de retourner ces résultats à l'Agent de Communication et de les afficher à l'utilisateur. L'enchaînement de ces étapes sur des cas d'étude réels est détaillé dans le chapitre suivant.

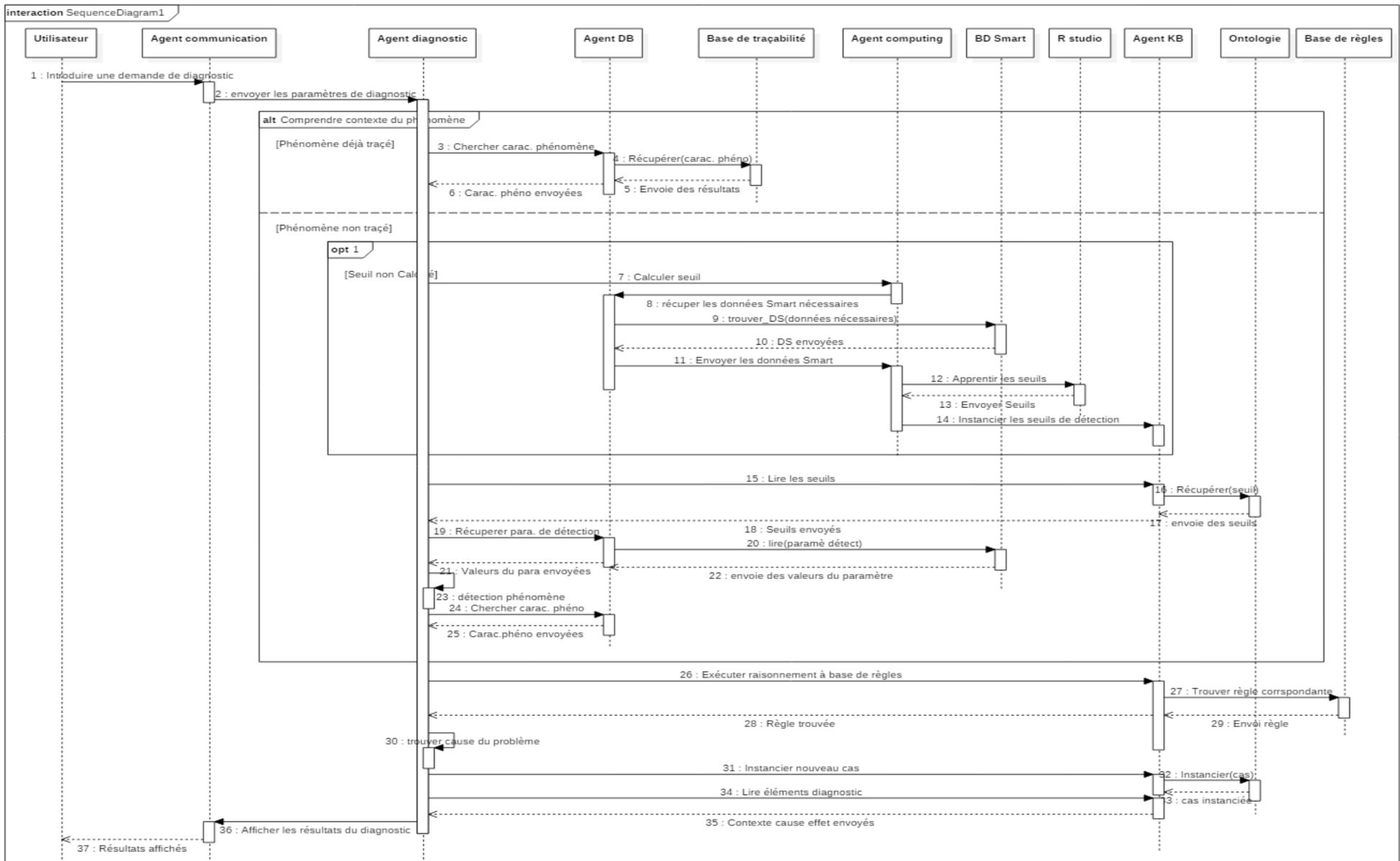


Figure 62. Séquencement de l'exécution d'un processus de diagnostic
149

Les IU de configuration gérée par l'agent de communication sont représentées dans la figure 63. Sur la première partie, on retrouve quatre boutons, les trois gris permettant d'accéder aux fonctionnalités de *Reporting* par détection d'événements, à la demande ou par planification. Pour le cas d'utilisation développé ci-dessus, l'utilisateur doit choisir le quatrième, "Diagnostic des phénomènes". Une deuxième interface apparaît, dans laquelle l'utilisateur peut choisir dans une liste les éléments du contexte (ligne de production, num. machine, période, type de phénomène à diagnostiquer, etc.). Dans ce scénario, l'utilisateur choisit de diagnostiquer la cause des phénomènes de broutement qui apparaissent au niveau des pièces usinées avec la broche n°1 de la ligne de production n°1 pour les 2 premières semaines de l'année 2020.



Figure 63. Interface utilisateur pour la configuration de l'opération de diagnostic

IV.5. CONCLUSION DU CHAPITRE

Ce chapitre a défini le cadre global d'implémentation en détaillant la structure générale du système d'aide à la décision. Les différentes contraintes et caractéristiques techniques ont été présentés et le déroulement du raisonnement a été expliqué. Pour assurer la validation de ces développements, il faut passer à l'étape de l'implémentation. Il est ainsi important d'établir des scénarios de test sur des cas d'études réels. Ceci permettra une amélioration des premiers développements grâce aux retours et de faire approuver la version finale par les experts. Le chapitre suivant va ainsi traiter plusieurs cas d'études réels, dans des domaines différents, permettant l'exécution du scénario de diagnostic et de valider les résultats.

CHAPITRE 5 : IMPLEMENTATION DANS LE CADRE DU PROJET SMARTEMMA

**“Un des tests de
l’autorité est d’identifier
un problème avant qu’il
ne devienne une
urgence.”**

Arnold H. Glasgow

V.1. INTRODUCTION DU CHAPITRE

Ce chapitre détaille les scénarios développés pour la validation de l'approche proposée. Ces développements se concentrent sur le deuxième axe d'aide à la décision, l'axe diagnostic, dans le domaine de l'usinage mécanique aéronautique. Le principe de l'implémentation choisie sera expliqué et les méthodes et technologies utilisées présentées. À la fin de ce chapitre, nous allons détailler la méthode d'adaptation de notre approche scientifique dans un autre domaine très différent, celui de la restitution numérique archéologique.

V.2. IMPLEMENTATION DE L'APPROCHE DANS LE DOMAINE DE L'USINAGE

V.2.1. Principe de l'implémentation

Pour commencer l'implémentation et tester l'efficacité des modèles développés, le point de départ est d'analyser les fichiers de données brutes et Smart. Une phase d'analyse de données et d'extraction des connaissances, en ayant recours à des algorithmes de *data mining* et d'apprentissage automatique (*Machine Learning*), est ainsi établie au préalable [Wang et al. 2019]. Des publications antérieures expliquent ces étapes de préparation nécessaires à l'établissement d'un processus de diagnostic [Ritou et al. 2019].

Les opérations d'analyse de données ainsi menées ont pu détecter un ensemble de phénomènes très fréquents en usinage, principalement : le broutement, l'usure, le bris des outils et la collision. Cette détection se fait en utilisant des seuils calculés grâce aux méthodes d'apprentissage automatique sur les données agrégées à partir des données brutes. Par exemple, pour détecter un broutement, il faut commencer par calculer le paramètre N_h , une donnée agrégée, qui est « la somme des amplitudes des cinq contributions non harmoniques dominantes du spectre des vibrations » [Godreau et al. 2019]. Par la suite, un seuil de détection des valeurs de ce paramètre N_h , a été appris avec les opérations de *data mining*. Un dépassement de ce seuil permettra de confirmer l'apparition du phénomène de broutement.

De même, un bris d'outil peut être détecté à travers un dépassement du seuil défini pour le paramètre U_b . Le calcul s'est basé sur le suivi du balourd de l'outil. Le critère U_b est défini comme « l'amplitude, dans le spectre fréquentiel du signal d'accélération, de la fréquence de rotation de la broche, pour éviter les perturbations liées à la coupe, le critère U_b est calculé uniquement lorsque l'outil est hors matière » [Godreau et al. 2019].

Plusieurs autres aspects peuvent être à l'origine d'un phénomène anormal, tels que des variations importantes de puissance (qui sont de bons indicateurs d'une collision) ou les variations des potentiomètres de vitesse de rotation et d'avancement des machines-outils (qui permettent d'avoir une idée des interventions humaines normales ou anormales et qui, à leur tour, peuvent retarder la productivité).

Après cette première phase de détection, et pour entamer l'étape de diagnostic et de compréhension de la cause de l'apparition, il était essentiel de formaliser les règles métier nécessaires pour établir des algorithmes de raisonnement.

Comme expliqué précédemment, la solution retenue consiste à s'appuyer sur l'expertise et le savoir-faire humain pour pouvoir distinguer les éventuelles causes possibles derrière chaque événement anormal. La capitalisation se fait à travers l'organisation d'entretiens avec des techniciens et des experts industriels, permettant de définir une première liste de règles.

La phase de formalisation commence par définir explicitement la liste des règles métiers. Le tableau 1 en présente quelques-unes. Sur chaque ligne, on trouve les paramètres nécessaires pour détecter un effet et la règle appliquée pour déterminer sa cause d'apparition.

Tableau 3. Liste des scénarios de diagnostic possibles

Les critères de détection	L'effet	La cause	La règle
$A_{rms} + P_{usi} > \text{seuil}$ Augmente en continu	Usure outil	Vitesse coupe trop élevée DVP (si fréquent entre outil)	Augmentation progressive d'énergie par couple outil/pièce Variation de puissance moyenne par rapport aux valeurs apprises
$U_{b\ nusi} + P_{usi} + V_{rms} > \text{seuil}$	Bris d'outil	Problème de cond. de coupe (N, Vf, Ap, Ae) Usure anormal outil Mauvaise DVP ou aléatoire	Si c'est récurrent : Il faut comparer avec les conditions de coupe de référence Si souvent bris en fin de vie

Nh > seuil A_{rms} > seuil	Broutement	Outil ou pièce Mauvaise programmation Mauvais bridage pièce	Il faut étudier la récurrence du phénomène pour savoir l'origine (outil, pièce)
Puissance variable par rapport à la valeur de Réf	Pièce non conforme	Montage de pièce incliné (prod grande série) Cause erreur de jauge (puissance élevée, différente)	Si op1 c'est moins grave et si op2 grave Correcteur en Z de l'outil : Temps utilisation outil plus long + variation de N
Potard (Vf ou N) ≠ 100	Problème potard	Manip opérateur : Approche manuel (programme démo) Défaut/Erreur usinage	Il faut vérifier le nom programme et le mode Pendant programme ok : pbl !
T_Vrms	Vibration excessive	Mauvaise CC/contexte - programmation (pièce spécifique) Problème Interface HSK - - Broche usée	Si y a toujours des sur-vibration avec un outil dans un contexte
Pusi	Pusi=0 ou Pas de vibration	Remonter une pièce finie déjà usinée (coté brut n'est pas usiné)	-----
BPFO	Broche usée	Gros bris (Co_Ub) ou gros broutement (Co_Nh)	Delta BPFO > seuil
Variation des puissances	Choc	Erreur de programme Erreur de jauge outil ou de palpage Mauvaise manip opérateur (mode manu) Problème réglage magasin outils	Vérification des noms de programme et sous-programme (Exp : CHANG_OUT.SPF pour le changement d'outil) Programme avec temps d'usinage Nul

Vol	Pas de matière usinée	Brut manquant ou pièce finie oubliée Bris outils (foret ou petite fraise carbure)	Vol < petite valeur Vibration = 0 => brut manquant Vibration faible => pièce réusinée
dT_usi, dT_nusi	Variation des durées de programme	Problème entre deux exécutions d'un programme (Arrêt Urgence, bris, RàZ) Intervention manuelle (M0), Manip Potard Vf Automatisation (Attente de palette, changement d'outil)	Il faut analyser les récurrences Code alarme « arrêt urgence » mode = automatique + en usinage / machine bouge => Problème grave de type collision bris
T	Problème d'échauffement	Augmentation de la Temp (serrage broche)	Vérifier Smart_data=Température
Q_lub, N, Vf	Mauvaise qualité métallurgique	Anomalie de la pompe	Débit de lubrification en dessous de la valeur de référence
N, Vf, Pot_av_rot	Arrêt de production	Pause déclarée /non déclarée Incident de travail Erreur manip opérateur (possible si variation potentiomètre) Panne machine	N=Vf=0 Pot av rot? Pot=0? Code alarme + mode maintenance

Pour donner suite à ce premier niveau de formalisation et de *mapping*, il est possible de commencer le développement de l'agent diagnostic. Ce dernier commence par une interface utilisateur permettant de définir le contexte de l'analyse (ligne de production, numéro de machine/broche, période d'acquisition, etc.) et de choisir l'opération de diagnostic à effectuer

(diagnostic de tous les phénomènes apparus ou d'un phénomène précis). Ces opérations peuvent dépendre du profil de l'utilisateur ou de l'importance et de la gravité des phénomènes étudiés. Dans ce second cas, il est important de définir un critère ou une note de criticité pour chaque anomalie. Cette notation est mesurée en se basant sur le tableau suivant.

Tableau 4. Critère de calcul de la criticité

Grille de cotation de la gravité de la défaillance G	
Critère	Cotation
Défaut mineur sans dégradation des performances	1
Défaut réparable sur site sans dégradation des performances	2
Défaut réparable sur site avec altération des performances	3
Défaut non réparable	4
Grille de cotation de la fréquence d'apparition F	
Critère	Cotation
Très faible	1
Faible	2
Fréquent	3
Très fréquent	4
Grille de cotation de la non-détection D	
Critère	Cotation
Détection préventive efficace (mesure, seuil optimisé, action)	1
Détection préventive (mesure, seuil non optimisé, action)	2
Détection préventive inexploité	3
Non détecté	4

Le fonctionnement de l'agent de diagnostic dépend fortement de sa capacité à accéder aux bases de données et de connaissances : tout le développement du système d'aide à la décision est basé sur la communication permanente avec ces bases. C'est dans ce contexte que nous avons pu tester l'efficacité de l'approche basée sur l'intégration des données et des connaissances. À la première étape, l'agent de diagnostic accède à la base de connaissances pour récupérer une liste de critères, de conditions, de seuils, etc. Puis, en accédant aux bases de données, il récupère les données correspondantes à ces paramètres pour exécuter le raisonnement spécifique requis. Enfin, il instancie tous les cas au niveau de l'ontologie avec respectivement le contexte, l'effet, la cause et la décision de chacun de ces cas.

V.3. PREMIER SCENARIO DE DIAGNOSTIC

Pour mieux illustrer ces aspects, nous détaillons un premier scénario spécifique concernant le diagnostic d'une rupture d'outil. Sur la base de la liste précédente (tableau 1), le scénario décrit dans le tableau 3, sera traité

Tableau 5. Les éléments du premier scénario de diagnostic : Bris d'outil

Les critères de détection	L'effet	La cause	La règle
Ub > seuil	Bris d'outil	Problème de conditions de coupe Une durée de vie d'outil mal définie	Comparer avec les conditions de coupe de référence Outil souvent cassé à la fin de sa vie

Le diagramme suivant (représenté par la figure 64) représente l'algorithme de détection et montre comment l'analyse d'un scénario particulier enrichit la base de connaissances. L'instanciation progressive des cas améliore le raisonnement et la fiabilité du système d'aide à la décision.

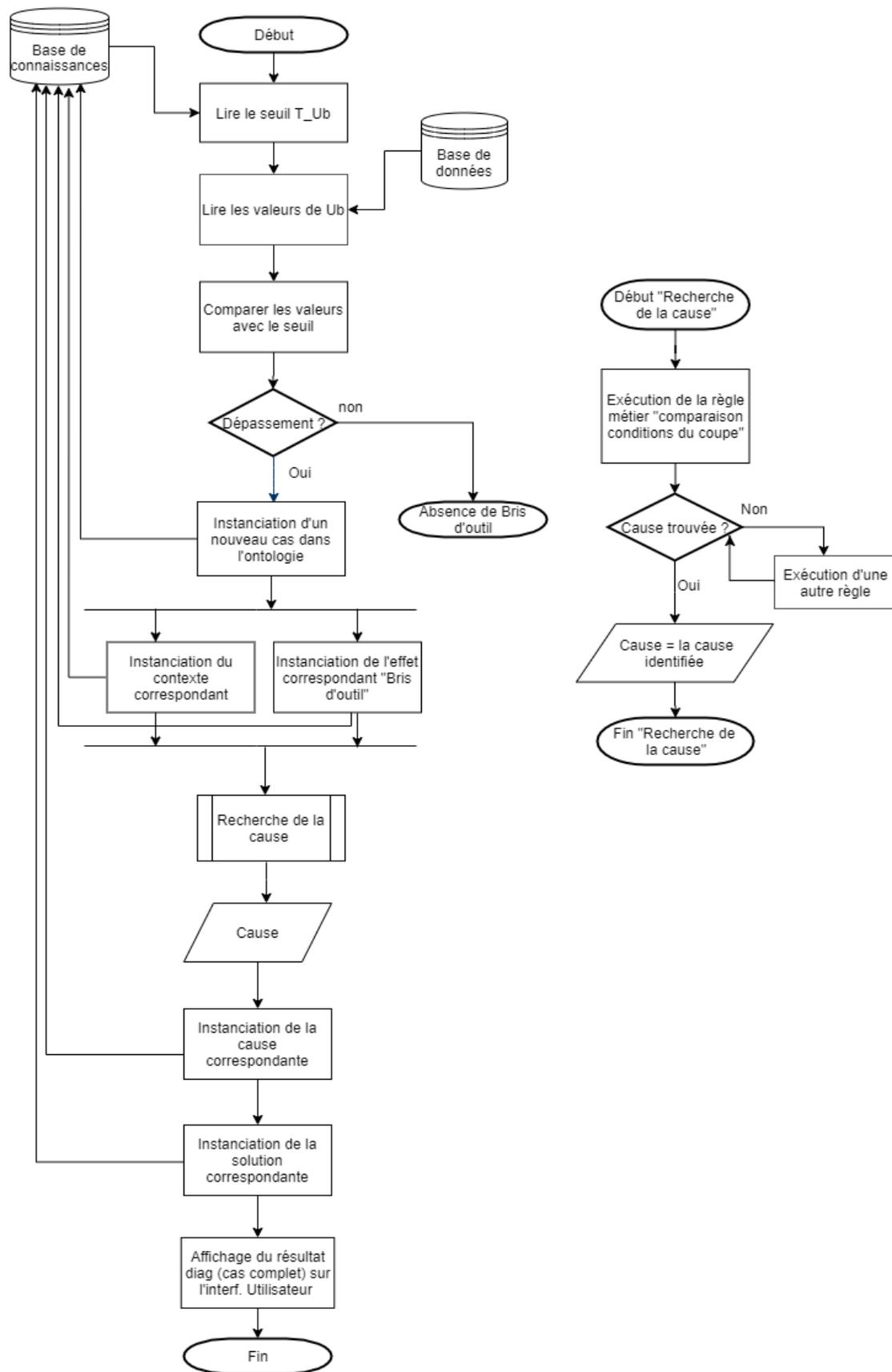


Figure 64. Diagramme d'activité de l'algorithme de détection.

Dans le cas spécifique d'une défaillance d'outil, la méthode consiste à vérifier le dépassement du seuil défini pour le critère de détection Ub pour lancer le cas échant un diagnostic. Pour ce faire, un seuil, calculé à l'aide d'algorithmes d'apprentissage automatique par des spécialistes d'analyse de données, a été instancié au niveau de la base de connaissances. C'est un bon exemple expliquant l'interaction entre les données et les connaissances. L'interrogation de l'ontologie pour la recherche de données se fait à l'aide du langage d'interrogation SPARQL. La figure ci-dessous montre un exemple de requête utilisée dans ce démonstrateur

```
queryString1 = "PREFIX ou:<http://www.semanticweb.org/pc/ontologies/2018/9/untitled-ontology-20#> "
+ "SELECT (?x as ?seuil) "
+ "where { ?y ou:Has_thresholdUb ?x."
+ " }";
```

Figure 65. Exemple de requête SPARQL

Une fois la valeur seuil récupérée, les bases de données de surveillance doivent être analysées en comparant les variations de Ub avec celle-ci. Un dépassement signifie l'apparition d'une défaillance de l'outil et déclenche le processus de diagnostic. Actuellement, les fichiers de surveillance bruts, et les fichiers Smart se présentent sous forme d'un fichier CSV. La figure 66 montre un extrait d'un fichier de données smart.

Jour	ModeCN	numOF	progP	progC	ID_Outil	OutilPF	temps_total	temps_arret	temps_arret	temps_Nvar	temps_G1_si	temps_G0_si
1	2	13	23	1	35							
1	2	13	23	7	35							
1	2	13	23	11	35							
1	2	13	23	9	35							
1	2	13	23	9	7							
1	2	13	23	10	7							
1	2	13	23	7	7							
1	2	13	23	11	7							
1	2	13	23	9	7							
1	2	13	23	10	7							
1	2	13	23	7	7							
1	2	13	23	11	7							
1	2	13	23	9	7							

Figure 66. Extrait d'un fichier de données Smart

Suite à la détection d'un bris d'outil, il convient d'instancier un nouveau cas de diagnostic dans l'ontologie développée avec le logiciel « Protégé ». Ainsi, il faut commencer avec la création d'un nouvel individu appelé « Case_tool_failure » et les individus qui vont permettre d'instancier la Cause, le Contexte et l'Effet. La figure 67 montre la liste des individus créés pour ce cas de diagnostic.

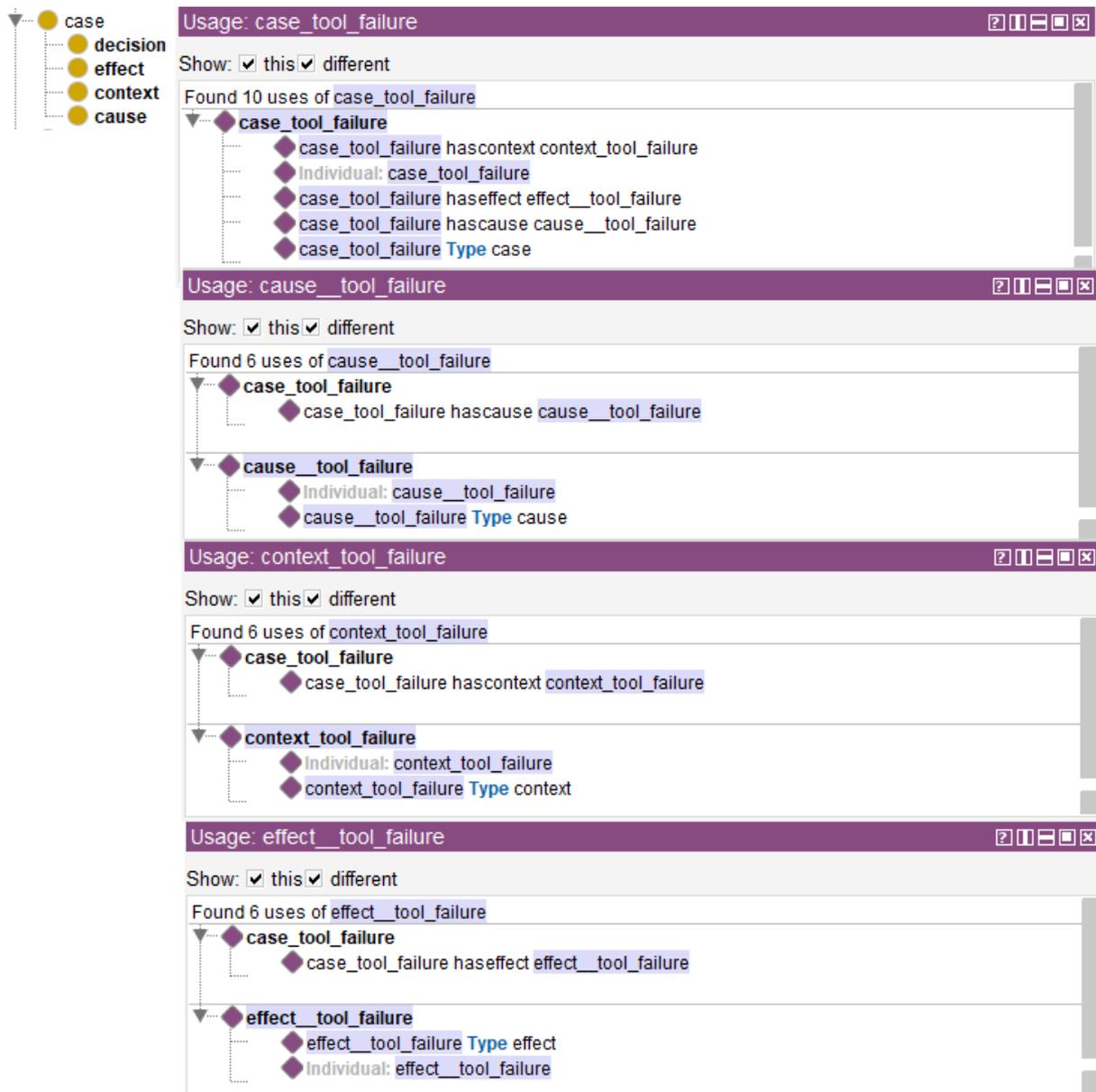


Figure 67. Instanciation de l'ontologie

Par la suite, une phase de mapping est nécessaire pour faire le lien entre ces éléments à l'aide des propriétés des objets (les *Object Properties*). Et enfin, en utilisant les propriétés des données (les *Data Properties*), les attributs de chaque individu sont ajoutés. Afin d'ajouter les éléments de contexte (dans cet exemple, il s'agit de l'instanciation de l'outil, du programme et du jour), il est nécessaire de se positionner sur la ligne du fichier Smart contenant un dépassement de seuil du critère Ub et de récupérer tous les paramètres nécessaires.

L'étape suivante, la recherche de la cause, est le cœur du processus de diagnostic. Elle est présentée dans le diagramme d'activités (voir figure 64) comme un processus à part entière.

Concernant ce scénario, une première règle métier a été formalisée suite aux entretiens avec les partenaires industriels du projet :

"L'outil est la cause du bris lorsque les conditions de coupe mesurées sont différentes des conditions de coupe de référence".

Ici, les conditions de coupe mesurées sont celles récupérées à l'aide des outils de surveillance et les conditions de coupe de référence sont prédéfinies par le bureau des méthodes. Au préalable, nous créons des fiches d'outils dans l'ontologie en nous basant sur les fiches outil fournies par les partenaires industriels. Techniquement, cela consiste à créer des individus « Outil_X » dans l'ontologie relatifs à la classe *outil* du dictionnaire de définition des ressources. Il faut compléter les valeurs de la liste des propriétés avec les conditions de coupe et les identifiants correspondant aux outils utilisés en usinage. La figure 68 présente les paramètres de définition possibles pour les fiches d'outils.

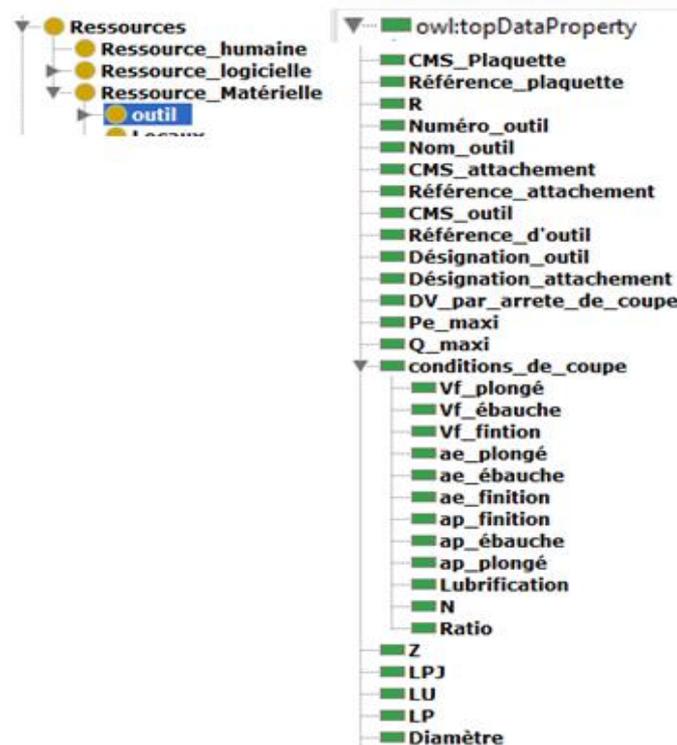


Figure 68. Paramètres de définition d'un outil

La même procédure d'identification se répète pour les programmes et les différentes ressources disponibles. Cette étape capitalise les paramètres industriels de référence dans la catégorie les bibliothèques qui constituent la première couche de la structuration de la base de connaissances.

Cet exemple, nous l'espérons, rend plus intelligible l'utilité et l'importance de cette catégorie de connaissances.

Pour implémenter la règle métier ci-dessus dans l'ontologie, une deuxième étape de formalisation est nécessaire afin de la traduire depuis le langage naturel vers le langage de règles formalisé spécifique de l'API JENA. Une fois traduit, le message prend la forme ci-dessous :

```
@prefix ns: <SpecificPrefix>.  
[rule1: (?a ns:hascontext ?b)(?a ns:hascause ?c)(?b ns:hastool ?x)(?x ns:hasVf ?d)(?x  
ns:hasVFRReal ?w) notEqual(?d,?w)  
→ (?c ns:hascausedirect 'tool problem: non-compliant cutting condition')]
```

L'exécution de cette règle métier implique la comparaison entre les deux catégories de conditions de coupe. En cas d'inégalité, l'individu « Case_tool_failure » est instancié et la cause est déterminée. La figure ci-dessous montre un cas où la cause déterminé est un « *Tool problem: non-compliant cutting condition* » (problème d'outil : condition de coupe non conforme)

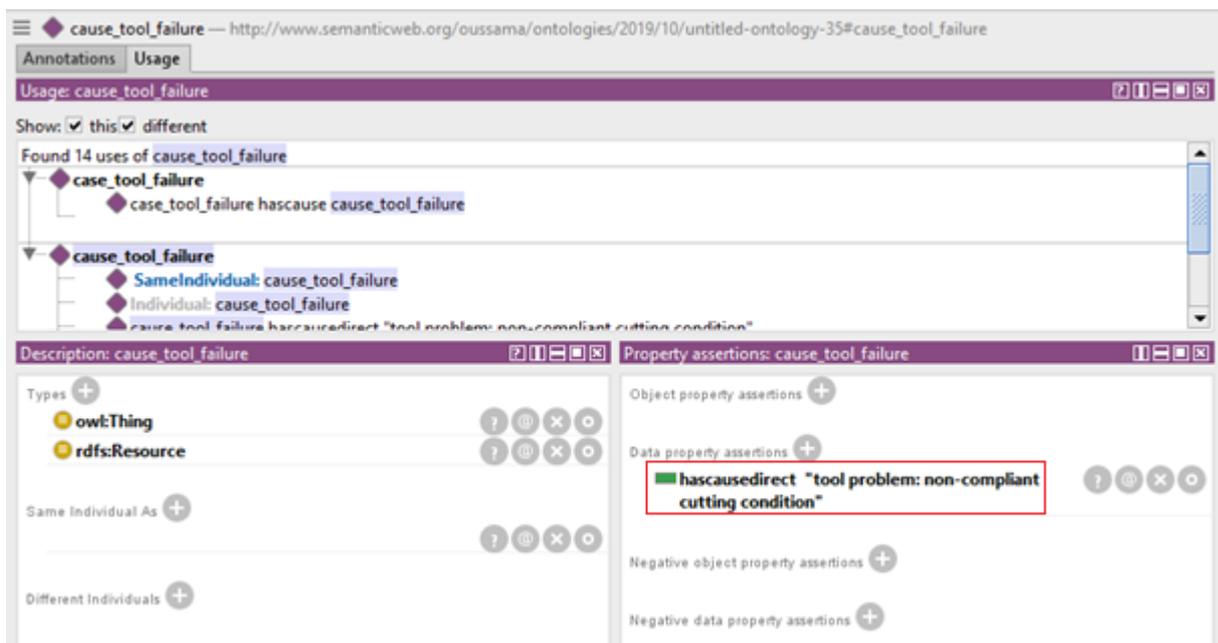


Figure 69. L'instanciation automatique de l'ontologie

Après cette phase d'instanciation de l'ontologie, le résultat de l'ensemble de l'opération de diagnostic est affiché au niveau de l'UI illustrée à la figure 70.

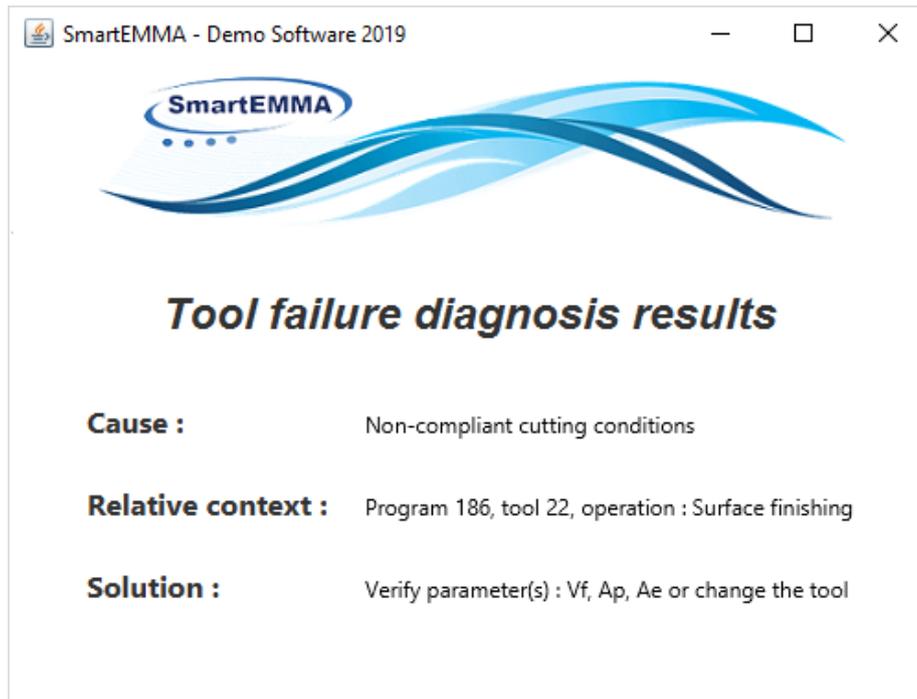


Figure 70. L'interface finale représentant le résultat du diagnostic

Ce scénario présente l'exécution d'une seule règle métier ; d'autres scénarios nécessitent de multiples règles et itérations pour trouver l'origine d'un seul problème. Par conséquent, la détermination d'une première cause directe peut conduire à un nouveau processus de diagnostic pour la détermination d'autres causes indirectes.

D'autres scénarios, nécessitent quant à eux l'exécution d'une phase d'apprentissage au préalable, pour la détermination de plusieurs valeurs significatives : des seuils, des intervalles de détections, des données aberrantes, etc. Le deuxième scénario présenté dans la partie suivante détaille notre recours à la fouille de données pour le calcul de plusieurs paramètres de référence.

V.4. DEUXIEME SCENARIO DE DIAGNOSTIC

Parmi la liste des scénarios de diagnostic implémentés dans le démonstrateur, cette section détaille le scénario de détermination de la cause potentielle de l'usure prématurée d'outils. Les éléments décrivant le scénario, nécessaires à la construction du raisonnement sont présentés dans le tableau 4.

Tableau 6. Les éléments du deuxième scénario de diagnostic : usure d'outil

Les critères de détection	L'effet	La cause	Règles
$A_{rms} + P_{usi} >$ seuil Augmente en continu	Usure outil	Cause vitesse coupe trop élevée DVP (si fréquent entre outil)	Augmentation d'énergie par couple outil pièce Progressive Variation de puissance moyenne par rapport à la valeur apprise (catégorie de pièce ou programme) valeur supérieure à la valeur apprise Quand la valeur de la puissance est bcp trop faible que la valeur apprise c'est possible que ça soit un problème de Programmation non productive (flexible)

V.4.1. Présentation du scénario

L'implémentation de ce scénario permet de détecter l'usure de l'outil coupant à travers le contrôle des puissances moyennes. D'après les experts métier, si la valeur de la puissance moyenne dépasse une valeur maximale donnée (par apprentissage automatique, par exemple), il est possible que la cause potentielle soit relative à une vitesse de coupe trop élevée. Dans le cas contraire quand il s'agit d'une vitesse de coupe trop faible par rapport à la valeur de la puissance minimale donnée, la cause possible est une programmation initiale non conforme de la vitesse de coupe par le bureau des méthodes. Pour l'implémentation du scénario dans le système, le diagramme d'activité de la Figure 71 explique les étapes nécessaires au déroulement du processus de diagnostic, et à l'instanciation d'un nouveau cas dans l'otologie en se basant sur un raisonnement à base de règles métier.

V.4.2. L'algorithme de diagnostic

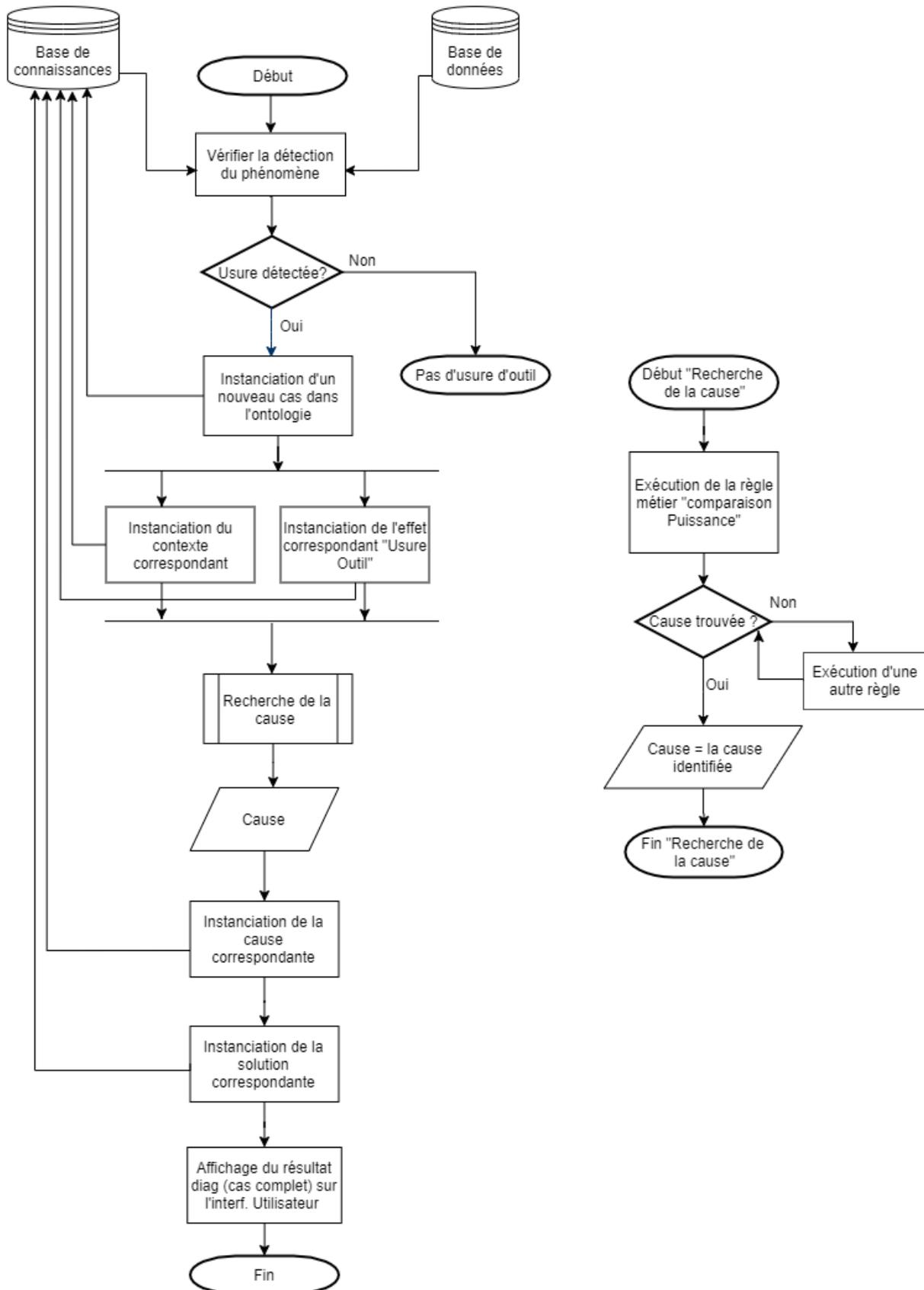


Figure 71. Processus de diagnostic d'une usure d'outil

La meilleure façon pour exécuter ce scénario est de se baser sur les rapports de traçabilité pour vérifier si un phénomène d'usure a été détecté ou non.

La spécificité de ce deuxième scénario est le recours aux outils d'analyse de données et d'apprentissage automatique pour déterminer les valeurs de puissance moyenne de référence sur lesquelles est basé le raisonnement. Les calculs se basent sur les données Smart et/ou Brutes (en fonction des besoins). Concernant l'exemple que nous traitons, vu la quantité importante de données manipulées et afin de faciliter la démonstration, il est possible de relier l'usure d'outil à un phénomène de bris d'outil ou à l'occurrence d'un broutement. Ainsi, nous nous limitons ici à l'étude des lignes de données qui détectent un phénomène de broutement. Par la suite, nous analysons les valeurs des puissances moyennes. Cet aspect de sélection illustre mieux la distinction entre une cause directe et une cause indirecte et donc un déclenchement d'un deuxième scénario suite à l'exécution d'un premier. L'algorithme d'apprentissage des valeurs de référence de la puissance moyenne est représenté par le diagramme suivant (figure 72).

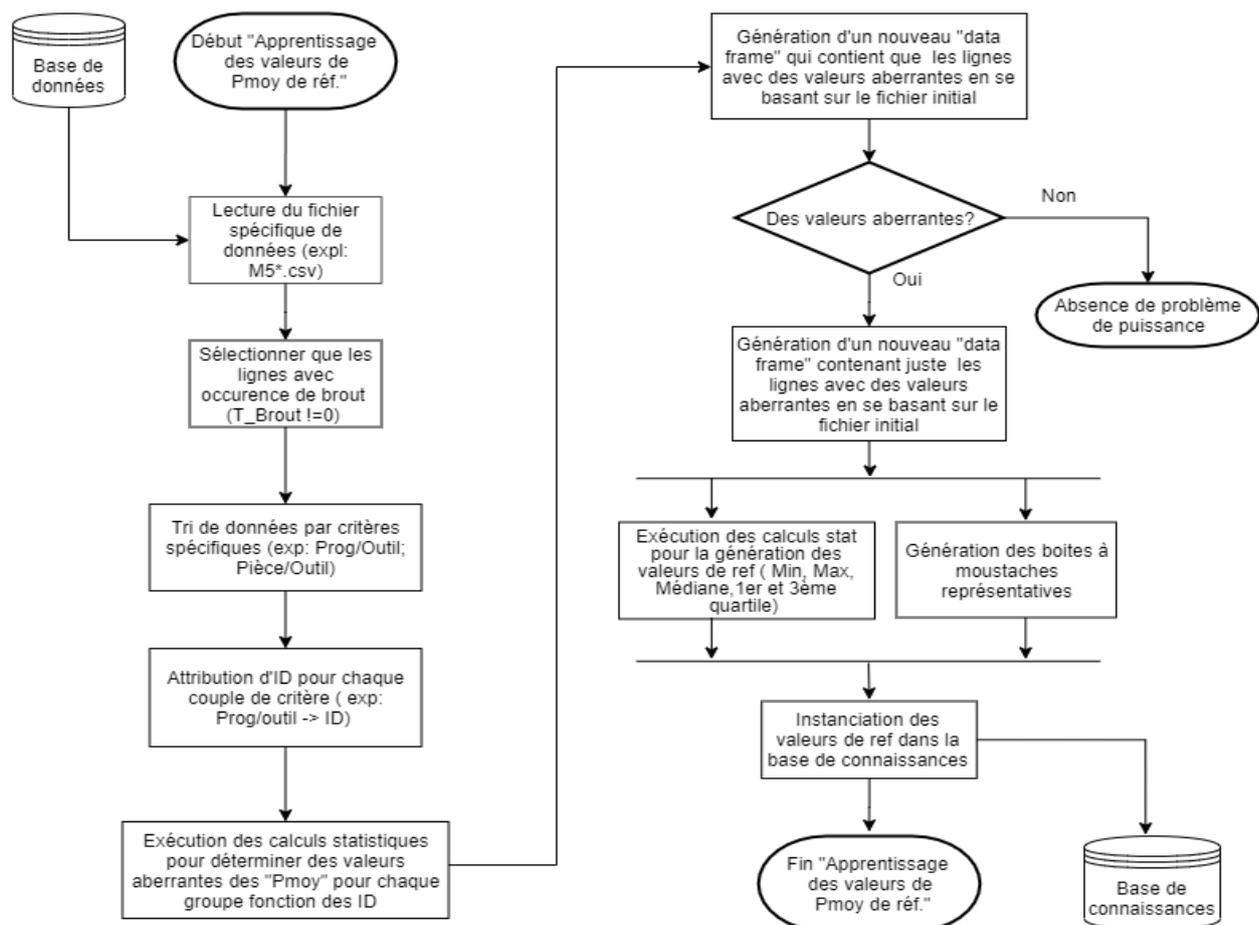


Figure 72. Algorithme d'apprentissage des valeurs de Pmoy_min/Pmoy_max

Une petite analyse de cet algorithme permet de distinguer deux phases principales pour l'extraction des connaissances qui sont : la préparation des données et l'analyse de données. La

première phase permet de structurer les données et de les identifier par couple Outil/Programme pour étudier l'occurrence et la récursivité du phénomène et contextualiser le raisonnement, ou par couple outil/pièce.

Dans le domaine des sciences des données, il est fréquent d'éliminer les données aberrantes dès la phase de préparation, ce qui n'est pas le cas dans ce système. Puisque l'objectif principal est d'étudier les problématiques d'usinage, plus spécifiquement les problèmes d'usure d'outil dans cet exemple, les données aberrantes sont justement celles qui portent les informations utiles à ce scénario. Toutefois, l'intérêt ne se limite pas à la capitalisation de ces données, mais surtout à la détermination des valeurs limites à partir desquelles les données sont jugées aberrantes, d'où l'importance de la deuxième phase de l'algorithme, l'analyse de données. La deuxième phase permet, en utilisant des fonctions prédéfinies et des calculs statistiques, de préciser les seuils de détermination des valeurs aberrantes. Autrement dit, il s'agit de calculer l'intervalle $[P_{moy_min} ; P_{moy_max}]$ et d'autres valeurs significatives comme la médiane, le premier quartile et le troisième quartile. Ces valeurs seront directement instanciées dans l'ontologie et vont être utilisées par la suite pour la détermination de la cause de l'usure. Ces données sont aussi stockées sous la forme d'un fichier de synthèse et peuvent être conservées dans la base de traçabilité pour de futures utilisations. Un aperçu de ce fichier est représenté par la figure 73.

	ID	progC	ID_Outil	Min	Max	Median	Lower_Quartile	Upper_Quartile	Outliers
1	102	254	22	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
2	102	254	22	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
3	106	259	23	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
4	107	259	24	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
5	112	262	24	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
6	112	262	24	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
7	115	273	22	3,101107	3,730343	3,407243	3,42307343	3,032024247	11,14300323
8	116	274	22	14,54957	14,78217	14,59338	14,55249131	14,69873798	8,747729302
9	116	274	22	14,54957	14,78217	14,59338	14,55249131	14,69873798	15,13603801
10	121	288	24	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
11	132	344	13	5,00000	5,00000	5,00000	5,00000	5,00000	5,00000
12	15	79	46	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000
13	16	81	13	4,00000	4,00000	4,00000	4,00000	4,00000	4,00000
14	16	81	13	4,00000	4,00000	4,00000	4,00000	4,00000	4,00000
15	16	81	13	4,273302	3,300000	4,300000	4,307027300	3,237127300	2,274200078

Figure 73. Extrait du fichier « Summary »

Par exemple, l'exécution de la première partie de l'algorithme permet d'associer l'ID = 116 au couple Programme/Outil « 274/22 ». Pour ce couple, le programme a retrouvé deux valeurs aberrantes de P_{moy} , valeurs qui se situent de part et d'autre des seuils critiques de détection.

Cet algorithme a été interprété et testé en utilisant le langage de programmation R. La figure 74 montre un aperçu du code développé. Lors de l'exécution et au fur à mesure du séquençement de l'algorithme, les différents *data frames* sont générés (la partie droite de la figure 74). Un *data frame* est un type utilisé dans le langage R pour la manipulation de données et peut être interprété comme une liste de vecteurs de dimensions identiques, ou comme une matrice. Il utilise les concepts « individus » ou « observations » pour distinguer les lignes et les variables pour les colonnes. Ceci explique les différentes descriptions quantitatives des *data frames* représentés ci-dessous. Nous remarquons que le nombre d'observations varie en diminuant entre les deux phases de l'algorithme suite à l'exécution des opérations d'agrégation et de sélection.

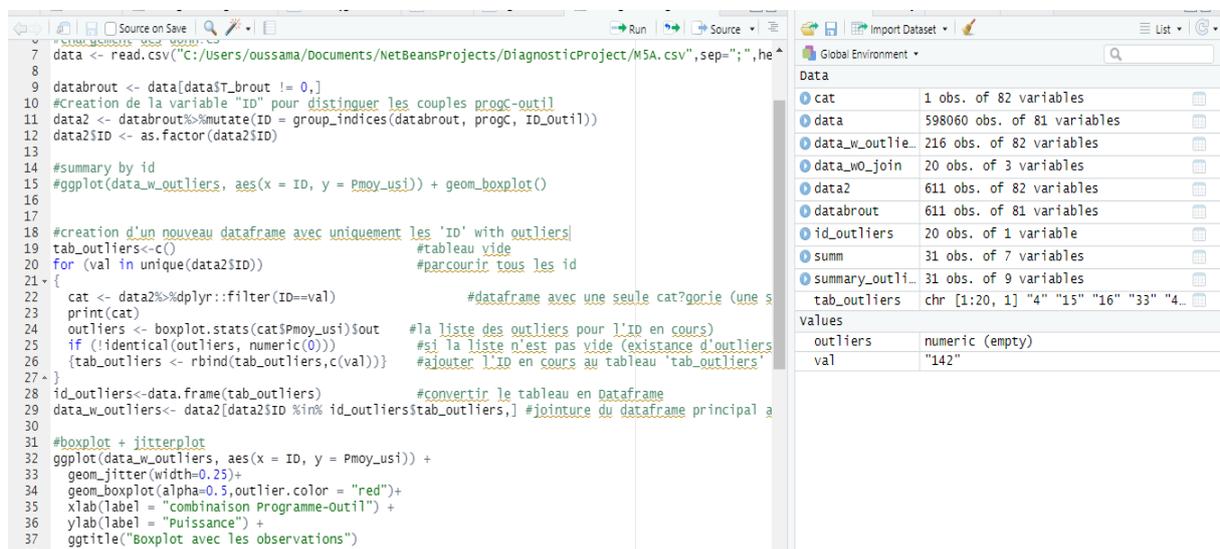


Figure 74. Code de calculs des valeurs de référence

Pour plus de visibilité, et suite à une étude des méthodes de représentation graphique et d'analyse de données, nous avons choisi les boîtes à moustache (*boxplot*, en anglais) pour représenter la distribution des puissances moyennes par couple d'outil/programme. Cette représentation, appelée aussi boîte de Tukey, du nom de son inventeur John Tukey, permet la distribution des observations en se basant sur des paramètres spécifiques qui sont : la médiane, le premier quartile, le troisième quartile, une valeur minimale et une valeur maximale qui figurent les extrémités des moustaches et qui sont calculées comme le produit de la différence entre le troisième et le premier quartile par 1,5. Les valeurs qui dépassent ces valeurs maximales et minimales représentent les données aberrantes. La figure 75 présente la structure d'une boîte à moustache.

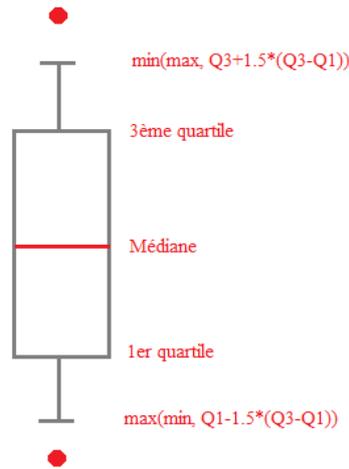


Figure 75. Forme et composants d'une boîte à moustache

Cette technique a été appliquée dans cet algorithme et a été testée sur les données réelles des partenaires industriels. Les distributions de puissances moyennes en fonction des combinaisons outil/programme, dans le cas de valeurs aberrantes, sont représentées par la figure 76.

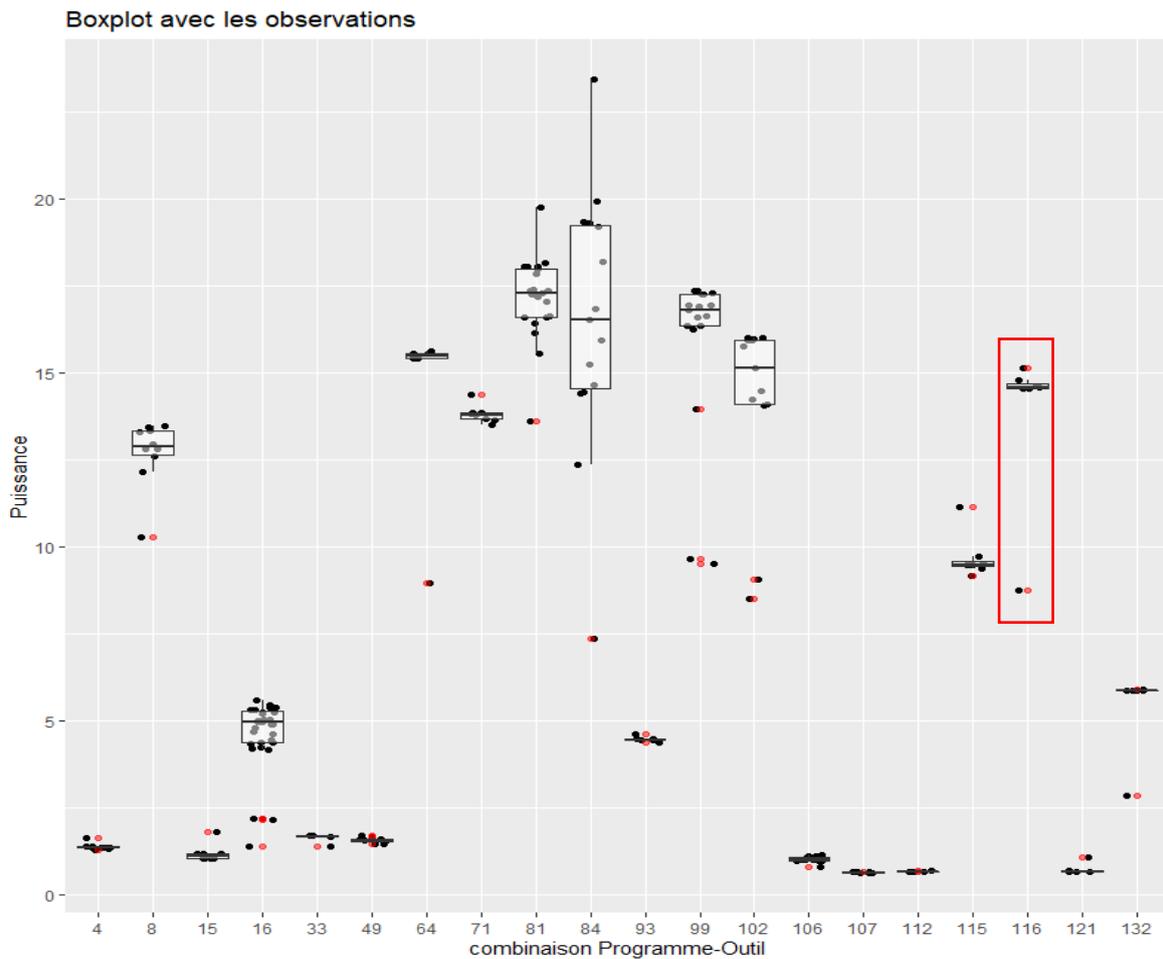


Figure 76. Représentation des distributions des Pmoy en fonction des ID avec les boîtes à moustache

Pour étudier la correspondance du fichier de données représenté dans la figure 73, et le diagramme en boîte, nous reprenons le même exemple du couple Programme/Outil = 274/22 identifié par l’ID 116. Et suite à une simple comparaison, il est immédiatement visible que les deux valeurs aberrantes {8,74 ; 15,13} extraites du tableau de données correspondent aux valeurs des point aberrants sur le diagramme.

Une fois les calculs statistiques réalisés, les nouvelles connaissances extraites sontinstanciées dans l’ontologie pour faciliter l’exécution des règles métiers permettant la détermination des causes d’apparition d’une usure d’outil.

@prefix ns: <*SpecificPrefix*>.

[rule: (?a ns:hascontext ?b)(?a ns:hascause ?c)(?b ns:hastool ?t)(?b ns:hasprogram ?p)]

Lire les éléments de contexte

[CombWithOut(?wo)(?wo ns:hastool ?tRef)(?wo ns:hasprogram ?pRef) equal(?t,?tRef)equal(?p, ?pRef)]

Détecter si la combinaison Outil programme possède des *outliers* (valeur aberrante)

(?b ns:hasPmoy ?pmoy)(?wo ns:hasPmoyMin ?Pmoymin) lessThan(?pmoy, ?Pmoymin)

Tester si la valeur de la puissance moyenne est inférieure à la valeur de référence Pmoy_min

→ (?c ns:hascausedirect ' Puissance faible: possibilité d’un problème de programmation non productive ')]

Instancier la cause dans l’ontologie

L’exécution de ces scénarios et du raisonnement à base de règles a permis de valider l’approche développée dans un premier temps dans le domaine d’étude de ce projet : l’industrie d’usinage mécanique aéronautique. L’objectif initial de ces travaux n’est pas de se limiter à une solution à usage unique, spécifique à une application précise ou à un domaine précis. Par conséquent, la généralité de l’approche proposée a été étudié et elle sera présenté dans le reste de ce chapitre.

V.5. VERS UNE GENERICITE DE L’APPROCHE PROPOSEE : APPLICATION AU DOMAINE DE LA RESTITUTION NUMERIQUE ARCHEOLOGIQUE

Dans le domaine de la gestion et de l’ingénierie des connaissances, la généralité reste un objectif principal pour la plupart des scientifiques et des experts de cette discipline. Dans le cadre de ces travaux de recherche, la notion de généralité concerne l’adaptabilité de l’approche complète de gestion et de modélisation des connaissances proposée, y compris des moyens

d'implémentation utilisés pour le développement d'un système d'aide à la décision associé, pour une réutilisation dans un autre domaine.

L'idée d'adapter notre solution dans un autre domaine est de permettre de tester son efficacité et de valider les développements. Ainsi, nous avons choisi de traiter un cas d'adaptabilité entre le domaine de la fabrication et de la production industrielle et celui de restitution numérique archéologique.

Le point fort de notre solution de modélisation est sa structuration en couches. Les deux premières couches regroupant l'ensemble des connaissances du domaine, elles permettent d'assurer l'aspect dynamique de la solution proposée avec une souplesse de modification.

Au vu de la conception modulaire de notre système, son adaptation à un autre domaine passe uniquement par le changement de l'ontologie de référence (première et deuxième couches) et des règles métiers associées au domaine, tout en gardant la même structuration pour la couche d'aide à la décision. Dans le contexte archéologique, trouver une ontologie de référence à proprement parler est difficile, mais s'appuyer sur des ontologies de description de l'architecture et de ses désordres est possible [Pérouse, 2011]. Les règles métier associées à la restitution archéologiques, en revanche, sont identifiées et définies au sein d'une discipline nommée archéologie du bâti. Elles lient l'observation d'un fait archéologique, l'effet, et de son contexte à une cause probable permettant la restitution d'un état antérieur, aujourd'hui disparu, ce qui constitue la décision. La figure suivante explique le principe de cette opération d'adaptabilité. Tout commence après l'intégration des ontologies de référence du domaine. Elles permettent une description fine du contexte et des effets observés (1). Ensuite, l'exécution des règles métier permet la détermination d'une cause probable (2). Cette cause entraîne enfin la prise d'une décision (3).

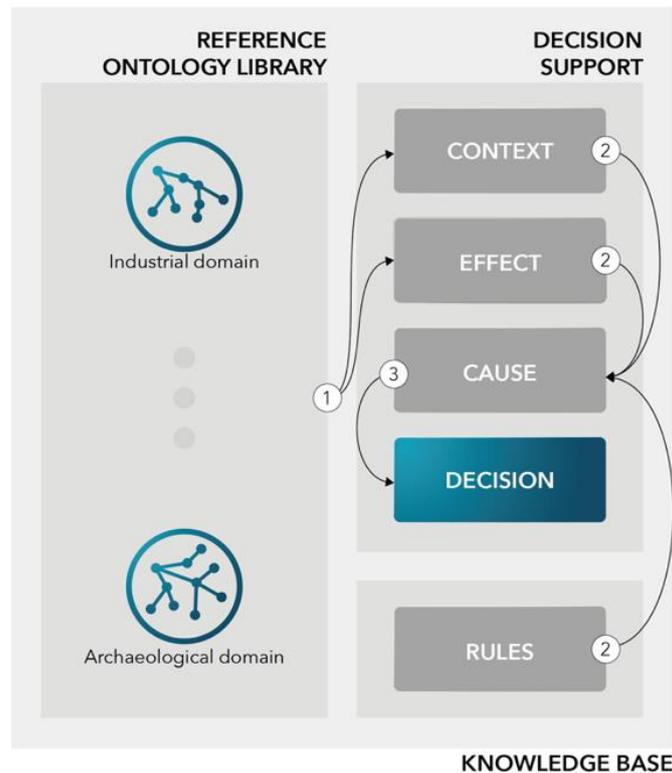


Figure 77. Structuration de la base de connaissances

Afin de tester l'efficacité de la solution de structuration et de cette stratégie d'adaptation, un cas d'étude spécifique a été établi.

V.5.1. Étude de cas : une forge du XVIII^e siècle

V.5.1.1. Contexte historique

Les forges d'Arthez-d'Asson sont situées dans les Pyrénées, dans la sud-ouest de la France. Les montagnes environnantes fournissent à la fois les ressources en bois et en fer qui ont permis le développement, dès la fin du Moyen Âge, d'une tradition de production de fer et de forgeage qui subsiste encore aujourd'hui. Les forges en question ici, dont l'état qui nous intéresse remonte au XVIII^e siècle, sont d'un type particulier, dit « à la Catalane » : une chute d'eau dans un tuyau percé latéralement provoque l'aspiration d'oxygène par effet Venturi et son injection sous pression dans le foyer [Pérouse, 2011].

Le site des forges dans la commune d'Arthez-d'Asson a été abandonné au XIX^e siècle lorsque le développement des hauts-fourneaux a rendu la technologie locale non compétitive [Larrouy, 1989]. Les bâtiments qui composent le site industriel sont donc restés inutilisés dans un environnement particulièrement rude au milieu de la Forêt des Pyrénées. S'il reste aujourd'hui

certaines murs de pierre, partiellement détruits, plus aucune installation mécanique, faites majoritairement de bois, ne subsiste.

La municipalité a donc mandaté une partie de l'équipe IS3P du LS2N (Florent Laroche et Paul François) pour participer à la conception du dossier patrimonial permettant la mise en valeur de ce site industriel par le biais d'une numérisation tridimensionnelle et d'une restitution numérique. Un relevé photogrammétrique a été entrepris du bâtiment de la forge afin d'obtenir un modèle 3D du bâtiment existant. Ce relevé a été couplé à une étude archéologique préliminaire des bâtiments du site. Sur la base de ces informations, nos collègues ont réalisé une restitution du complexe de la forge à partir d'une étude fonctionnelle et architecturale [Laroche et al., 2020]. Quelques documents et témoignages d'époque éclairent la compréhension du site, notamment les planches de l'*Encyclopédie* d'Alembert [Diderot, 1751-1772] qui représentent un certain nombre de mécanismes utilisés dans les forges du XVIII^e siècle. Néanmoins, la plupart des restitutions sont basées sur des considérations d'archéologie du bâti.

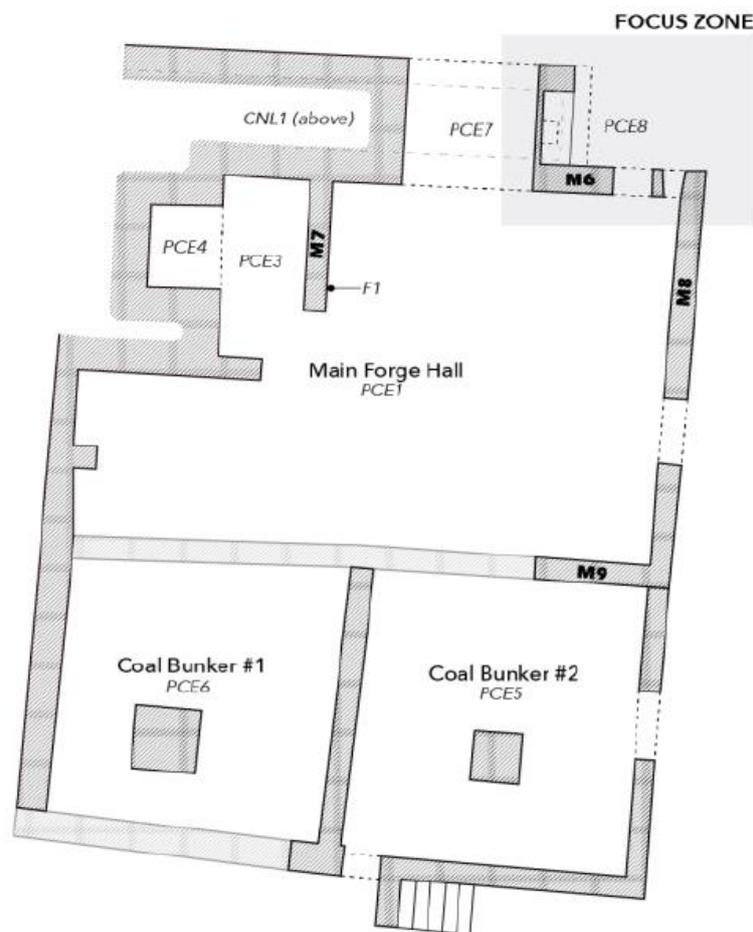


Figure 78. Plan de la forge d'Arthez-d'Asson.

V.5.1.2. Une roue à aubes et son axe

Pour les besoins de ces travaux, nous nous concentrerons simplement sur la zone dite PCE8 (voir plan du site, figure 78) qui abritait une chute d'eau amenée par un canal (CNL1, ci-dessus) permettant la rotation d'une roue à aubes dont l'axe traversait le mur nord de la forge (mur M6).

Nous cherchons non seulement à restituer la roue à aubes (notamment son diamètre, important pour l'interprétation mécanique des forges) mais aussi à interpréter les éventuels désordres de la construction résultant d'éventuelles modifications apportées au bâti.

Tous les faits que nous allons observer ont un contexte commun : le contexte historique, culturel, spatial et technique des forges et qui inclut notamment le savoir-faire industriel régional [Laroche, 2017]. Ce contexte limite et influence les causes probables des effets observés, ainsi que les décisions qui en découlent : ce qui pourrait être possible dans un contexte socioculturel donné pourrait ne pas l'être dans un autre. Physiquement, le contexte commun est l'espace dit PCE8.

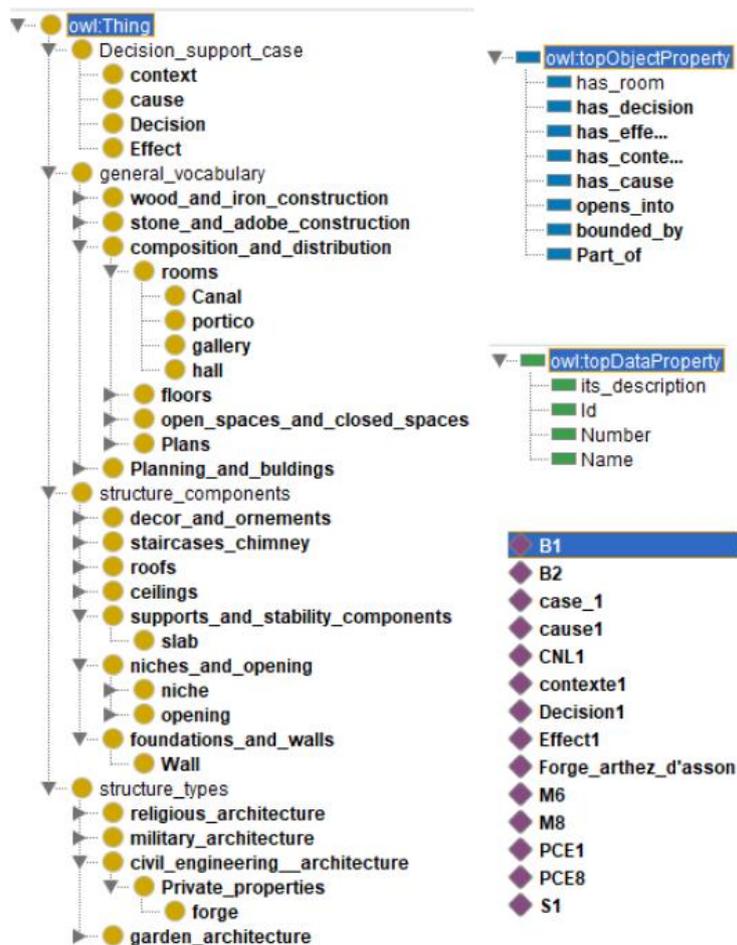


Figure 79. Extrait de la structure de l'ontologie architecturale

V.5.1.3. Mise en œuvre de notre approche

Afin de valider la méthode globale (figure 77), nous avons d'abord cherché une ontologie pour représenter l'architecture du bâtiment, sans considérations archéologiques. À partir de l'ouvrage de référence *Architecture, description et vocabulaire méthodique* [Pérouse 2011], nous avons construit l'ontologie correspondante. Ensuite, nous avons enregistré toutes les décisions prises au cours du processus d'interprétation des ruines afin de concevoir un ensemble de règles relatives à l'archéologie du bâti des Forges d'Arthez-d'Asson. Par exemple, dans le PCE8, la première observation que nous avons faite est la présence de joints ouverts et de pierres érodées dans la partie inférieure du mur M6 : c'est cet effet que nous étudions. L'étude du contexte montre qu'il existe une ouverture (B1) reliant l'espace du PCE8 au canal CNL1, permettant à l'eau de ce dernier de s'écouler dans la pièce. Dans ces conditions, on peut déduire, et c'est notre règle, que la cause de cette érosion est la présence d'eau en mouvement. La décision de restitution est donc la présence d'un équipement hydraulique dans PCE8.

La formalisation de cette règle à l'aide du langage SWRL donne :

```
Decision_support_case(?x)^has_effect(?x,?y)^      its_description(?y,      ?L)      ^
swrlb:stringEqualIgnoreCase(?L, "wall with open joints and eroded stones") ^ has_cause(?x,
?g)^ has_context(?x, ?m)^ has_desision(?x,?d)^ rooms(?r)^ opening(?n)^ has_room(?m, ?r)
^ bounded_by(?r, ?n)^ Canal(?p)^ bounded_by(?p, ?n) -> its_description(?g, "frequently
moving water environment"), its_description(?d, "restitution of a hydraulic machine")
```

De même, à partir de l'observation de la présence d'une seconde ouverture (B2) dans la paroi M6 séparant PCE8 de PCE1 et de la décision précédente, on peut faire l'hypothèse que l'ouverture étudiée correspond au passage de l'arbre de la roue à aubes, constituant ainsi une seconde règle. Cette hypothèse est soutenue par le fait qu'aucune autre ouverture menant au PCE8 ne permet la transmission d'une force mécanique.

Les règles peuvent également être appliquées à la composition de la paroi elle-même. Par exemple, la structure du mur M6 montre un changement dans la composition du mur (effet), mais la règle de l'archéologie du bâti implique que la cause en est une reconstruction partielle du mur, impliquant la restitution d'un état antérieur (décision). Cela peut en outre être considéré à son tour comme un effet. Et dans le contexte de la présence d'eau et de sol instable qui caractérise le PCE8, une nouvelle règle nous permet de déduire que la cause de cette

reconstruction est l'affaissement probable du mur dû à l'usure prématurée des pierres et du substrat.

L'adaptation de la méthodologie aux problèmes et aux défis de l'archéologie du bâtiment est une tâche profondément interdisciplinaire. Par conséquent, elle pose des problèmes communs à tous les travaux de ce type et en particulier la prise en compte des contraintes liées au domaine de l'archéologie. Pour ce faire, il a fallu reconstruire une ontologie adaptée au bâtiment étudié, ce qui représente un investissement de temps important. Néanmoins, l'intérêt de la méthodologie développée est la possibilité d'utiliser des ontologies existantes (de référence), ce qui réduit considérablement le temps nécessaire pour adapter la méthodologie à de nouveaux domaines. Par exemple, nous aurions pu utiliser l'ontologie CIDOC-CRM (*Conceptual Reference Model*), qui est particulièrement adaptée à la description de tout objet patrimonial mais qui s'est avérée trop importante pour le simple travail de validation que nous voulions entreprendre.

Le travail avec des experts en sciences humaines a également permis de mettre en évidence des pistes de développement de notre méthodologie. En effet, un concept clé de la restitution archéologique est la notion d'hypothèse : toute cause identifiée est toujours hypothétique et la décision prise doit pouvoir refléter cette incertitude. Afin de l'intégrer dans notre méthodologie, il serait nécessaire de définir des facteurs d'incertitude (basés sur des observations, des mesures, le manque de connaissances, etc.) qui peuvent être utilisés pour créer des indicateurs clés de performance (KPI) pour la reconstruction numérique.

De retour dans le domaine industriel, la prise en compte de l'incertitude dans l'aide à la décision industrielle pourrait donner au système expert une plus grande flexibilité pour proposer des solutions.

V.6. CONCLUSION DU CHAPITRE

Ce chapitre valide l'efficacité de l'approche proposée pour le développement des axes d'aide à la décision avec les scénarios détaillés. Nous avons aussi démontré la généralité de notre méthodologie d'aide à la décision et son adaptabilité à des domaines aussi éloignés de l'industrie que la reconstruction archéologique. L'ajout d'une couche supérieure d'aide à la décision structurée sous la forme du quadruplet "Cause-Contexte-Effet-Décision" s'est révélé particulièrement bien adapté dans ce domaine, confirmant ainsi nos choix de conception. En effet, cette couche supplémentaire permet une description détaillée du processus de restitution

et de l'arbre de décision associé, ce qui reste une question importante dans la modélisation 3D pour l'histoire et l'archéologie.

Cette approche peut assurer la mise en œuvre de plusieurs catégories d'aide à la décision telles que le diagnostic et la prédiction et permettrait à terme de prendre des décisions soit en amont lors de la phase de conception du produit, soit en aval lors de la phase de fabrication (en interagissant directement avec les machines-outils). Notre proposition montre que la gestion des connaissances et l'ingénierie sont autant capables de répondre aux problèmes de l'industrie 4.0 que de permettre l'aide à la restitution numérique de l'industrie du passé.

CONCLUSION ET PERSPECTIVES

La question de recherche principale de ces travaux est de trouver un moyen de mettre les nouvelles générations de systèmes d'aide à la décision au profit de l'industrie du futur à travers une conciliation entre données et connaissances métier, dans le contexte de l'industrie d'usinage mécanique. Pour répondre à cette question nous avons proposé une approche de conception et de développement d'un SAD à base de connaissances. Celui-ci permet de faciliter les opérations de traçabilité des données et connaissances relatives à un phénomène malveillant pouvant apparaître au moment de l'usinage. Il permet également d'effectuer le diagnostic nécessaire à l'identification de la cause de l'apparition du phénomène.

Pour la construction de l'approche, une méthode de développement a été adoptée. Composée de trois phases, elle permet dans un premier temps d'analyser les différentes sources de données et de connaissances. À ce niveau, une première problématique apparaît : comment gérer, structurer et capitaliser de façon efficace les flux de connaissances industrielles disponibles. Par conséquent, la deuxième phase de la méthodologie propose une stratégie de gestion de ces flux. La pertinence et l'originalité de nos travaux se manifestent dans la séparation entre les deux niveaux de données et de connaissances. Pour cela, plusieurs bases ont été structurées et implémentées. Elles sont fortement liées et communiquent continuellement, ce qui place les connaissances au service des données et vice versa. L'utilité de cette intégration a été démontrée dans la troisième phase de la méthode, consistant en l'implémentation des modèles et à mettre en place le SAD.

Un point clé de notre approche réside dans la structuration en couches de la base de connaissances qui sépare les couches inférieures, regroupant l'ensemble des connaissances métiers génériques, de la couche supérieure d'aide à la décision. Cette structuration a été à partir du modèle générique que nous proposons, le CR3P (Contexte, Ressources, Produit, Processus, Personne). Il regroupe les familles de concepts de base qui structurent toutes les données et connaissances industrielle. Nous estimons que ce modèle peut être réimplémenté et utilisé dans d'autres application appartenant au domaine de la production.

Le recours aux solutions ontologiques est apparu comme un moyen efficace de mise en place d'une base de connaissances globale. Suivant la structuration en couche de celle-ci, la base a été implémentée sous forme d'une ontologie du domaine de la fabrication mécanique formant

une première couche. Une deuxième couche regroupe les standards qui assurent la généralité de la solution proposée, favorisant sa réutilisation dans d'autres applications dans le domaine de l'usinage mécanique. La troisième couche comprend les connaissances décisionnelles structurées sous forme de cas. Cette modularité garantit à son tour l'adaptabilité dans plusieurs autres domaines de recherche en s'appuyant sur la possibilité de changer la première couche avec une autre ontologie de référence issue d'un autre domaine.

Finalement, en collaboration avec les experts métiers industriels partenaires du projet, une base de règles métiers a été mise en place. Ces règles représentent une autre forme de connaissances de haut niveau sur lesquelles se base le Système d'Aide à la Décision à Base de Cas pour assurer le raisonnement. L'architecture proposée pour le développement de ce SAD a été mise à jour et validée par l'exécution de plusieurs scénarios d'essai dans deux domaines différents. Si l'approche scientifique répond au besoin distingués initialement, de multiples perspectives pourraient être envisagées pour la poursuite de ces travaux.

Perspectives

Le fonctionnement du SAD étant basé sur des règles métiers, d'autres sessions d'entretien avec des experts industriels sont envisageables. Ceci permettrait d'améliorer et d'enrichir la base des règles métiers, afin de pouvoir mettre en œuvre d'autres scénarios de diagnostic. Surtout, cela autoriserait à capitaliser l'expertise et le savoir-faire des spécialistes, une ressource incontournable dans toute industrie, malheureusement rarement préservée. Le recours à l'intelligence artificielle représente aussi une solution efficace et surtout durable, permettant la sauvegarde d'une partie des savoir-faire des employés.

D'autres aspects pourraient être complétés dans notre proposition scientifique. Par exemple, le processus de sélection de règles pour déterminer la cause principale du problème est actuellement basé sur l'effet détecté et son contexte spécifique pour exécuter un raisonnement, choisir une règle et déterminer l'origine de l'effet. Malgré l'efficacité de l'approche et sa validation dans des scénarios réels, nous avons défini un critère de criticité qui permettrait d'étudier la priorité de phénomènes en fonction de leur fréquence d'occurrence et de leur gravité. Un deuxième indice pourrait être défini, relatif au nombre d'utilisation d'une règle pour un phénomène donné. La traçabilité de ces deux nouveaux KPI aiderait à affiner et faciliter le choix de la règle métier adéquate.

La structuration de la base de connaissances proposée dans ces travaux favorise plusieurs perspectives de développement dans ce projet. Premièrement, la structuration en quadruplet

"Contexte-Cause-Effet-Décision" regroupés sous forme de cas permettrait d'établir un système de raisonnement à partir de cas. Il supporterait et appuierait le travail du système à base de règles métier et permettrait la mise en place d'autres axes d'aide à la décision. Nous avons en effet laissé la possibilité de faire de la traçabilité des différents cas diagnostiqués au fur et à mesure de l'occurrence d'un phénomène malveillant et de l'exécution d'un scénario de diagnostic. La capitalisation de ces cas diagnostiqués et de tous les éléments décrivant le contexte de travail permettrait de faciliter le processus d'identification des cas et le calcul des coefficients et mesures de similarité. L'intégration du raisonnement à partir de cas dans le projet améliorerait, sans doute, la fonction de proposition de solutions et de décision. Il permettrait également de mettre en place la prévision de la possibilité d'occurrence d'un phénomène et la proposition d'actions préventives.

Une autre perspective de ce travail consiste à valider l'adaptabilité et la généricité de la méthodologie développée dans d'autres domaines d'application. L'expérience de validation de notre approche dans le domaine de la restitution numérique archéologique, pour comprendre l'origine des caractéristiques d'un site donné et pour proposer des décisions de restitution, était constructive et a permis de traiter un autre aspect de la solution qui va au-delà du diagnostic pour proposer des solutions. Cela mène à une architecture d'un système d'aide à la décision plus générique, dans laquelle la décision peut prendre plusieurs formes et dépend du domaine d'application et du contexte d'utilisation.

Au-delà du périmètre de ces travaux, d'autres problématiques restent aujourd'hui sujettes à études et recherches scientifiques. L'extraction et l'exploitation des connaissances sont dans l'intérêt de la plupart des industries et des entreprises dans tous les domaines. Conscient de l'importance de cette discipline, nous pensons que l'enjeu principal est de trouver un moyen pour mettre en place les techniques et méthodes plaçant la gestion des connaissances au cœur de la stratégie d'entreprise.

Une autre perspective de développement consiste en la proposition d'une stratégie de gestion des connaissances qui supporte non seulement la mise en place des systèmes d'aide à la décision, mais qui permette également d'étudier la fiabilité et la précision des décisions proposées.

D'autre part, grâce à l'efficacité et à l'utilité des systèmes d'aide à la décision, ceux-ci sont de plus en plus utilisés dans le contexte industriel réel. Toutefois, le fort recours à ces systèmes commence à empiéter sur la place de l'homme dans de nombreux secteurs. Il est ainsi

primordial de continuer à travailler sur la valorisation du rôle de l'humain dans ce type de système et l'importance de sa participation dans le processus de prise de décision. Il faut donc penser à proposer des solutions privilégiant l'expertise humaine face aux enjeux de la mise en place de l'industrie 4.0.

Limites

La mise en œuvre de la méthodologie de gestion des connaissances industrielles proposée dans cette thèse et l'expérimentation de plusieurs scénarios d'aide à la décision ont montré la faisabilité et l'intérêt de cette approche. Toutefois, cette approche présente certaines limites, notamment concernant la disponibilité et l'accessibilité des connaissances industrielles. Un tel outil doit être mis en œuvre dans un contexte industriel favorable qui offre les ressources et les moyens nécessaires, et qui encourage la collaboration entre le système numérique et les experts.

Ces travaux de recherche adoptent le raisonnement à base de règles métier pour assurer le processus de diagnostic. Cette technique, malgré son efficacité, demande un effort important d'extraction des connaissances implicites des experts métier (le savoir-faire), et beaucoup de précision au moment du passage du langage naturel à un langage formel informatisé. L'interprétation de l'information est relative au modèle cognitif de chaque personne, il est ainsi probable que la règle comprise et formalisée par le cognitivien, développeur du système d'aide à la décision, ne soit pas parfaitement conforme à l'expertise réelle. La fiabilité de la formalisation des règles est une contrainte indispensable pour la réussite de l'approche. Une règle impliquant une décision pouvant avoir des conséquences relativement importantes sur tout le système industriel, il est nécessaire qu'elle soit bien étudiée et validée par les experts avant son implémentation finale.

L'approche proposée dans ces travaux de recherche a été appliquée dans le cas d'étude industriel de diagnostic des défaillances des pièces mécaniques aéronautique. La liste des phénomènes possibles traités et analysés avec les partenaires industriels, est assez longue. Cependant, tous les problèmes imaginables n'apparaissent pas chaque jour dans le domaine industriel et parfois une surveillance supplémentaire et une analyse approfondie des processus sont nécessaires pour détecter davantage de défaillances. Par conséquent, la validation risque d'être complexe parce que tous les scénarios et toutes les règles métiers développés et capitalisés dans la base de règles ne peuvent pas être testés avant que la défaillance à laquelle chacun d'eux répond n'apparaisse.

Cet aspect souligne l'importance de l'homme dans le contexte de la numérisation et de la quatrième révolution industrielle. La nécessité d'une première phase d'entretiens directs avec les experts des entreprises souligne ce rôle crucial. Le défi consiste donc, pour le futur, à promouvoir l'importance de la collaboration entre le système et les utilisateurs qui restent maîtres et contrôleurs des opérations : l'objectif initial d'un système d'aide à la décision.

LISTE DES PUBLICATIONS

Journaux internationaux

1. Oussama Meski, Farouk Belkadi, Florent Laroche, Mathieu Ritou & Benoit Furet (2020) A generic knowledge management approach towards the development of a decision support system, International Journal of Production Research, DOI: 10.1080/00207543.2020.1821930
2. Oussama Meski, Farouk Belkadi, Florent Laroche, Asma Ladj, Benoit Furet, Integrated Data and Knowledge Management as Key Factor for Industry 4.0, IEEE Engineering Management Review, Year: 2019, Volume: 47, Issue: 4, Magazine Article, Publisher: IEEE
3. Oussama Meski, Paul François, Florent Laroche. An in virtuo system linking data corpus to 3D virtual model for industry 4.0. 2020, International Journal of Virtual Reality, Vol.20 No.1
4. Asma Ladj, Zhiqiang Wang, Oussama Meski, Farouk Belkadi, Mathieu Ritou, et Catherine Da Cunha. A knowledge-based Digital Shadow for machining industry in a Digital Twin perspective. Journal of Manufacturing Systems, Elsevier, In press, (10.1016/j.jmsy.2020.07.018).
5. Zakaria Yahouni, Asma Ladj, Farouk Belkadi, Oussama Meski, Mathieu Ritou. A smart reporting framework as an application of a multi-agent system in the machining industry. International Journal of Computer Integrated Manufacturing.

Conférences internationales

6. Oussama Meski, Benoit Furet, Farouk Belkadi, Florent Laroche. Towards a knowledge structuring framework for decision making within industry 4.0 paradigm. Manufacturing Modelling, Management and Control - 9th MIM 2019, Aug 2019, Berlin, Germany.
7. Oussama Meski, Farouk Belkadi, Benoit Furet, Florent Laroche. Towards a knowledge-based framework for digital chain monitoring within the industry 4.0 paradigm. CIRP DESIGN, May 2019, Póvoa de Varzim, Portugal. pp.118-123, (10.1016/j.procir.2019.04.250).

8. Oussama Meski, Paul François, Florent Laroche. An in virtuo system linking data corpus to 3D virtual model for industry 4.0. ConVRgence, Apr 2020, Laval, France.

Conférences nationales

9. Oussama Meski, Florent Laroche, Farouk Belkadi, Benoit Furet. La structuration des connaissances au service de l'industrie 4.0 : Le cas du projet " SmartEmma ". 16ème colloque national S-mart/AIP-PRIMECA, Apr 2019, les Karelis, France.
10. Oussama Meski, Farouk Belkadi, Florent Laroche, Benoit Furet. La gestion des connaissances au cœur de la continuité numérique pour l'industrie 4.0: Le cas du projet «SmartEmma». EGC, Jan 2019, METZ, France.
11. Oussama Meski, Farouk Belkadi, Florent Laroche, Benoit Furet, Mathieu Ritou. La structuration des connaissances pour l'aide à la décision dans le contexte de l'industrie 4.0. 24ème Congrès Français de Mécanique (CFM 2019), Aug 2019, Brest, France.

Journal national

12. Oussama Meski, Florent Laroche, Farouk Belkadi, et Benoît Furet. 2019. « La gestion des connaissances au cœur de la continuité numérique pour l'industrie 4.0 : Le cas du projet « SmartEmma » ». 2019. Revue des Nouvelles Technologies de l'Information vol.RNTI-E-35. Éditions RNTI, 381–382.

BIBLIOGRAPHIE

- Aamodt, Agnar, et Enric Plaza. 2001. « Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches ». *AI Communications* 7: 39-59.
- Abele, Sebastian, et Michael Weyrich. 2017. « Decision Support for Joint Test and Diagnosis of Production Systems Based on a Concept of Shared Knowledge ». *IFAC-PapersOnLine* 50(1): 15227-32.
- Ackoff, R.L. 1989. « From Data to Wisdom ». *Journal of Applied Systems Analysis* 16: 3–9.
- ADEME, Deloitte Développement Durable, G.-S.C.O.P. Deloitte In Extenso, et S.A.T.I.E. 2017. *Impacts du numérique au sein de l'industrie, au regard de la transition énergétique et écologique*.
- Adla, A. 2018. « Intelligent Semantic Case Based Reasoning System for Fault Diagnosis ». *Journal of Digital Information Management (JDIM)* 16 Number 2.
- Agyapong-Kodua, K., Csaba Haraszko, et István Németh. 2014. « Recipe-Based Integrated Semantic Product, Process, Resource (PPR) Digital Modelling Methodology ». *Procedia CIRP* 17: 112-17.
- Al-Gumaei, K. 2018. « A Survey of Internet of Things and Big Data Integrated Solutions for Industrie 4.0 ». In *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA, Turin, 1417–1424,*
- Allen, G.Donald. 2017. *Hierarchy of Knowledge – from Data to Wisdom*.
- Ammar-Khodja, Samar. 2007. « Processus d'aide à la spécification et à la vérification d'application d'ingénierie à base de connaissances expertes ». These de doctorat. Nantes. <http://www.theses.fr/2007NANT2159> (16 novembre 2020).
- Annamalai, G. et al. 2011. « An Ontology for Product-Service Systems ». In *Decision Engineering Report Series*, éd. R. Roy et Y. Xu. Cranfield: Cranfield University.
- Archer, E.R. 1980. « How to Make a Business Decision ». *Management Review* 2: 54–61.
- Armengol, E. 2011. « Classification of Melanomas in Situ Using Knowledge Discovery with Explained Case-Based Reasoning ». *Artificial Intelligence in Medicine* 51: 93–105.
- Arnott, D., et G. Pervan. 2005. « A Critical Analysis of Decision Support Systems Research ». *Journal of Information Technology* 20(2): 67–87.
- . 2008. « Eight Key Issues for the Decision Support Systems Discipline ». *Decision Support Systems* 44(3): 657–672.
- Arp, Robert, Barry Smith, et Andrew D. Spear. 2015. *Building Ontologies with Basic Formal Ontology*. MIT Press.

- Bar, M. et al. 2006. « Top-down facilitation of visual recognition ». *Proc Natl Acad Sci U S A* 103(2): 449–454.
- Baskarada, S., Koronios A. Data, et Knowledge Information. 2013. « Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and Its Quality Dimension ». *Australasian Journal of Information Systems* 18(1): 1–20.
- Belkadi, F., A. Bernard, et F. Laroche. 2015. *Knowledge Based and PLM Facilities for Sustainability Perspective in Manufacturing: A Global Approach*. Procedia CIRP.
- Belkadi, Farouk et al. 2012. « A Meta-Modelling Framework for Knowledge Consistency in Collaborative Design ». *Annual Reviews in Control* 36(2): 346-58.
- Berawi, M.A., et R.M. Woodhead. 2005. « Application of Knowledge Management in Production Management ». *Human Factors Ergon. Manuf* 15(3): 249–257.
- Berners-Lee, T., J. Hendler, et O. Lassila. 2001. « The Semantic Web ». *Scientific American* 284: 34–43.
- Bichindaritz, I., et S. Montani. 2011. « Advances in Case-Based Reasoning in Health Sciences ». *Artificial Intelligence in Medicine* 51: 75–79.
- Blanchet, M. 2016. « Industrie 4.0 Nouvelle donne industrielle, nouveau modèle économique ». *Outre-Terre* 46(1): 62–85.
- Bohanec, M., I. Bratko, et V. Rajkovic. 1983. « An Expert System for Decision Making ». In *Sol (Ed.), Processes and Tools for Decision Support*, éd. H.G. Amsterdam: North-Holland.
- Bonnet, A., J.P. Haton, et J.M. Truong-Ngoc. 1986. *Systèmes experts : vers la maîtrise technique*. InterEditions.
- Boulila, Wadii. 2012. *Extraction de connaissances spatio-temporelles incertaines pour la prédiction de changements en imagerie satellitale. Traitement des images [eess.IV]*. Télécom Bretagne, Université de Rennes 1.
- Brettel, Malte, Niklas Friederichsen, Michael Keller, et Marius Rosenberg. 2014. « How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective (Version 9997144) ». *International Journal of Mechanical, Industrial and Aerospace Sciences* 7(0(1)).
- Brown, Sam. 2007. « Forensic Engineering: Reduction of Risk and Improving Technology (for All Things Great and Small) ». *Engineering Failure Analysis* 14(ue 6). <https://doi.org/10.1016/j.engfailanal.2006.11.065>.
- Candlot, A. 2006. *Principes d'assistance à la maîtrise d'ouvrage pour la modélisation et l'intégration d'expertise, Thèse de doctorat*. Ecole Centrale de Nantes et Université de Nantes.
- Carthy, J.Mc. 1980. « Circumscription : A Form of Non Monotonic Reasoning ». *Artificial Intelligence* 13: 27–39,.

- Catalano, Chiara et al. 2009. « A Product Design Ontology for Enhancing Shape Processing in Design Workflows ». *Journal of Intelligent Manufacturing* 20: 553-67.
- Caulier, Patrice. 1999. *Concepts du modèle de l'expertise CommonKADS pour la modélisation au niveau connaissances*.
- Charlet, J., M. Zacklad, et G.Kassel D. Bourigault. 2000. *Ingénierie des connaissances : évolutions récentes et nouveaux défis*, Édition Eyrolles. Paris.
- Chaudhri, Vinay et al. 2004. *Graph-Based Acquisition of Expressive Knowledge*.
- Chen, Xiaojun, Shengbin Jia, et Yang Xiang. 2020. « A Review: Knowledge Reasoning over Knowledge Graph ». *Expert Systems with Applications* 141: 112948.
- Chengula, Z. et al. 2018. « State of Industry 4.0 across Six French Companies: A Pilot Study ». In *IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Stuttgart.
- Choo, C.W. 2006. *The Knowing Organization : How Organisations Use Information to Construct Meaning, Create Knowledge, and Make Decision*. Oxford: Oxford University Press.
- Chow, H.K.H., K.L. Choy, W.B. Lee, et F.T.S. Chan. 2005. « Design of a Knowledge-Based Logistics Strategy System ». *Expert Syst. Applic* 29: 272–290.
- Cochoy, Franck, et Gilbert Terssac. 2000. « Au-delà de la traçabilité : la mappabilité. Deux notions connexes mais distinctes pour penser les normes de management ».
- Combacau, M. et al. 2000. « Supervision and Monitoring of Production Systems ». *IFAC Proceedings Volumes* 33(17): 849-54.
- Cortes, U. et al. 2001. « Knowledge management in environmental decision support systems ». *AI Communication* 14: 3–12.
- Cost, R. Scott et al. 2000. « Using Colored Petri Nets for Conversation Modeling ». In *Issues in Agent Communication*, Lecture Notes in Computer Science, éd. Frank Dignum et Mark Greaves. Berlin, Heidelberg: Springer, 178-92. https://doi.org/10.1007/10722777_12 (16 novembre 2020).
- Crossland, Martin D. 2015. « Decision Support Systems ». In *Encyclopedia of GIS*, éd. Shashi Shekhar, Hui Xiong, et Xun Zhou. Cham: Springer International Publishing, 1-1. https://doi.org/10.1007/978-3-319-23519-6_269-2 (16 novembre 2020).
- Cutting-Decelle, A.F. et al. 2007. *15 Iso 15531 Mandate: A Product-Process-Resource Based Approach for Managing Modularity in Production Management Concurr Eng*.
- Daihani. 2001. *Sensitivity Test Application For MADM Models Using SAW and TOPSIS Methods*. Yogyakarta.
- Danjou, Christophe. 2015. « Ingénierie de la chaîne numérique d'industrialisation : proposition d'un modèle d'interopérabilité pour la conception-fabrication intégrées ». phdthesis. Université de Technologie de Compiègne. <https://tel.archives-ouvertes.fr/tel-01376520>

(16 novembre 2020).

- Darkhovski, B., et M. Staroswiecki. 2003. « A game-theoretic approach to decision in FDI ». *IEEE Transactions on Automatic Control* 48(5): 853-58.
- Davenport, T., et L. Prusak. 1998. *Working Knowledge: How Organizations Manage What They Know*. Boston: Harvard Business Review Press.
- Davenport, Thomas H. 2009. « Make Better Decisions ». *Harvard Business Review*. <https://hbr.org/2009/11/make-better-decisions-2> (16 novembre 2020).
- DE casteljajac, C. 2012. « Surveillance avancée et amélioration du procédé d'UGV ». *Thèse de doctorat de l'université de Nantes*.
- Demeester, Thomas, Tim Rocktäschel, et Sebastian Riedel. 2016. *Regularizing Relation Representations by First-order Implications*.
- Demoly, F. et al. 2010. « Multiple Viewpoint Modelling Framework Enabling Integrated Product–Process Design ». *International Journal on Interactive Design and Manufacturing (IJIDeM)* 4(4): 269–280.
- Dendani, N., Khadir M, et S. Guessoum. 2012. « Hybrid approach for fault diagnosis based on CBR and ontology: using jCOLIBRI framework Int ». In *Conf. on Complex Systems IEEE*, , 1–8.
- Dendani, Nadjette, et Rayene Allouani. 2019. « Rules-based decision support system and domain ontology for diabetes diagnosis ». In *30es Journées Francophones d'Ingénierie des Connaissances, IC 2019, AFIA*, 224–230 –02329609.
- Dey, A., et G. Abowd. *Towards a better understanding of context and context-awareness. Proceedings of the CHI 2000 Workshop on TheWhat, Who, Where, When and How of Context Awareness2000;4:1–6*.
- Dhouieb, M. A. 2016. *Structuration multi-échelle de la connaissance in-extenso d'entreprise*.
- Diderot, D., Le Rond D'Alembert, et J. 1751. *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*. Paris.
- Donier, L. et al. 2017. « Analyse de risques par la méthode AMDEC de la préparation des doses à administrer par un automate de formes orales sèches ». *Le Pharmacien Hospitalier et Clinicien* 52(1): e42-43.
- Drucker, P.F. 1956. « How to Make A Business Decision, Nation's Business ». *Archer, E.R:* 54–61.
- Easton, J.M., J.R. Davis, et C. Roberts. 2011. « Ontology Engineering: The “What's”, “Why's” and “How's” of Data Exchange ». *International Journal of Decision Support Systems Technology* 3(1): 40–53.
- Ebrahimipour, V., K. Rezaie, et S. Shokravi. 2010. « An Ontology Approach to Support FMEA Studies ». *Expert Systems with Applications* 37(1): 671-77.

- Efstathiou, J., et V. Rajkovic. 1979. « Multiattribute Decision Making Using a Fuzzy Heuristic Approach ». *IEEE Trans. Syst., Man Cyber SMC-9*: 326–333.
- Efthymiou, K., K. Sipsas, D. Mourtzis, et G. Chryssolouris. 2015. « On Knowledge Reuse for Manufacturing Systems Design and Planning: A Semantic Technology Approach ». *CIRP Journal of Manufacturing Science and Technology* 8: 1–11.
- Elnagar, Samaa, Victoria Yoon, et Manoj Thomas. 2020. *An Automatic Ontology Generation Framework with An Organizational Perspective*.
- Ericson, Åsa, et Tobias Larsson. 2009. *People, Product and Process Perspectives on Product/Service-System Development*. 10.1007/978-1-84882-909-1_11.
- Ermine, Jean-Louis 2000a. Les systèmes de connaissances. Hermes Science Publication, pp.144.
- Ermine, J.L. 2000b. *La gestion des connaissances, un levier stratégique pour les entreprises, Actes IC2000, Ingénierie des connaissances*. Toulouse.
- Ermine, Jean-Louis, Mahmoud Moradi, et Stéphane Brunel. 2012. 16 *Management international/International Management/Gestión Internacional*.
- Farhoomand, A.F., et D.H. Drury. 2002. « Managerial Information Overload ». *Communications of the ACM* 45(10): 127.
- Fayyad, U.M., G. Piatetsky-Shapiro, et P. Smyth. 1996a. « From Data Mining to Knowledge Discovery: An Overview ». In *Advances in Knowledge Discovery and Data Mining*, éd. U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, et R. Uthurusamy. AAAI Press, 1–34.
- . 1996b. « The KDD Process for Extracting Useful Knowledge from Volumes of Data ». *Communications of the ACM* 39(11): 27–34.
- Flores-Saldivar, Alfredo et al. 2015. *Industry 4.0 with Cyber-Physical Integration: A Design and Manufacture Perspective*.
- FouFou, S. et al. 2005. « A Core Product Model for PLM with an Illustrative XML Implementation ». In *International Conference on Product Lifecycle Management*.
- Franz, Miltner, Vogel Tobias, et Hemmje Matthias. 2014. « Towards Knowledge Based Process Planning Support for CAPP-4-SMEs: Problem Description, Relevant State of the Art and Proposed Approach" ». In *International Manufacturing Science and Engineering Conference (MSEC)*.
- Fricke, M. « The Knowledge Pyramid: A Critique of the DIKW Hierarchy ». *Journal of Information Science* 2009(;35(2):131-42).
- Furet, B. 1994. « La surveillance automatique de l'usinage en fraisage par l'analyse du courant de broche ».
- Galarraga, Luis Antonio, Christina Teflioudi, Katja Hose, et Fabian Suchanek. 2013. « Amie: Association Rule Mining under Incomplete Evidence in Ontological Knowledge Bases ». In *Proceedings of the 22nd International Conference on World Wide Web*,

- International World Wide Web Conferences Steering Committee, 413–422.
- Gernhardt, Benjamin et al. 2016. *Knowledge-Based Production Planning Within the Reference Planning Process Supporting Manufacturing Change Management*. 10.1115/MSEC2016-8658.
- Gero, J.S. 1990. « Design Prototypes: A Knowledge Representation Schema for Design ». *AI Magazine* 11(4): 26–36.
- Giménez, Diego, Marcela Vegetti, Horacio Leone, et Gabriela Henning. 2008. « PProduct ONTOlogy: Defining product-related concepts for logistics planning activities ». *Computers in Industry* 59: 231-41.
- Giustozzi, Franco, Julien Saunier, et Cecilia Zanni-Merk. 2018. 126 *Context Modeling for Industry 4.0: an Ontology-Based Proposal*, *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2018.08.001>.
- Godreau, Victor. 2017. « Extraction des connaissances à partir des données de la surveillance de l'usage ». These de doctorat. Nantes. <http://www.theses.fr/2017NANT4104> (16 novembre 2020).
- Gómez-Pérez, A. et Benjamins, et R. 1999. « Overview of Knowledge Sharing and Reuse Components: Ontologies and Problem-Solving Methods ». In *CEUR Workshop Proceedings. IJCAI and the Scandinavian AI Societies*,.
- Gorecky, D., M. Khamis, et K. Mura. 2017. « Introduction and Establishment of Virtual Training in the Factory of the Future ». *International Journal of Computer Integrated Manufacturing* 30(1): 182–90.
- Gorry, G.A., et M.S.Scott Morton. 1971a. *A framework for management information*.
- . 1971b. « A Framework for Management Information Systems ». *Sloan Management Review* 13(1).
- Gruber, T.R. 1993. « Toward Principles for the Design of Ontologies Used for Knowledge Sharing ». In *International Workshop on Ontology*, Padova.
- Grundstein, M. 2000. « Repérer et mettre en valeur les connaissances cruciales pour l'entreprise ». In *inProceedings of Actes du 10ème Congrès International de l'AFAV*, Paris. novembre.
- . 2009. « GAMETH®: A Constructivist and Learning Approach to Identify and Locate Crucial Knowledge ». *International Journal of Knowledge and Learning* 5: 289–305.
- Gruninger, M., et J.B. Kopena. 2005. « Planning and the Process Specification Language ». In *Proceedings of WS2 ICAPS 2005*, Monterey, CA, 22–29.
- Grüninger, Michael, et Joongu Lee. 2002. « Ontology ». *Communications of the ACM* 45: 39.
- Guerra-Zubiaga, David A., et Robert I.M. Young. 2006. « A Manufacturing Model to Enable Knowledge Maintenance in Decision Support Systems ». *Journal of Manufacturing Systems* 25(ue 2): 122–136,.

- Gunasekaran, et E.W.T. Ngai. 2007. « Knowledge Management in 21st Century Manufacturing ». *International Journal of Production Research* 45(11): 2391–2418,.
- Gupta, P. 2019. « Modularity Enablers: A Tool for Industry 4.0 ». *Life Cycle Reliab Saf Eng* 8: 157–163.
- Haettenschwiler, P. 2001. *Neues anwenderfreundliches konzept der entscheidungsunterstützung.* "Zurich, vdf Hochschulverlag AG.
- Harani, Y. 1997. « Une Approche Multi-Modèles pour la Capitalisation des Connaissances dans le Domaine de la Conception », Thèse de doctorat, Institut National Polytechnique de Grenoble.
- Harding, J. A., M. Shahbaz, Srinivas Srinivas, et A. Kusiak. 2006. « Data Mining in Manufacturing: A Review American Society of Mechanical Engineers (ASME) ». *Journal of Manufacturing Science and Engineering* 128(4): 969-76.
- Hardwick, Martin, et David Loffredo. 2007. *STEP-NC: Smart Data for Smart Machining.*
- Houriez, B. 1994. *Acquisition de connaissances pour l'aide à la conduite et à la supervision de procédés industriels. Habilitation à diriger des recherches.* Université de Valenciennes et du Hainaut-Cambrésis.
- Huang, Haoming, Michel Pasquier, et Chai Quek. 2011. « Decision support system based on hierarchical co-evolutionary fuzzy approach: A case study in detecting gamma ray signals ». *Expert Syst. Appl.* 38: 10719-29.
- IEC 60812. 1985. *Analysis techniques for system reliability procedure for failure mode and effects analysis (FMEA).* International Electrotechnical Commission.
- Isermann, R. 1994. « Integration of Fault Detection and Diagnosis Methods ». *IFAC Proceedings Volumes* 27(5): 575-90.
- . 1997. « Supervision, Fault-Detection and Fault-Diagnosis Methods — An Introduction ». *Control Engineering Practice* 5(5): 639-52.
- . 2006. *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance.* Berlin Heidelberg: Springer-Verlag. <https://www.springer.com/gp/book/9783540241126> (16 novembre 2020).
- ISO/IEC JTC 1. 2015. « Information Technology, Big data: Preliminary report ». *ISO/IEC.*
- Jashapara, A. 2005. *Knowledge Management : An Integrated Approach.* Harlow: FT Prentice Hall.
- Johnson, M. 2010. « Using process FMEA in an Aeronautical Engineering Technology capstone course ».
- Julien, Nathalie, et Eric Martin. 2020. *Le jumeau numérique : De l'intelligence artificielle à l'industrie agile.* Dunod.
- Kagermann, H., W. Wahlster, et J. Helbig. 2013. *Recommendations for Implementing the*

- Strategic Initiative INDUSTRIE 4.0. Acatech. National Academy of Science and Engineering.*
- Kant, J., et S. Sridharan. 1998. « The Values of Using Scheduling Information in Planning Material Requirements ». *Deci. Sci* 29(2): 479–497.
- Katipamula, Srinivas, et Michael R. Brambley. 2005. « ReviewArticle: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I ». *HVAC&R Research* 11(1): 3–25.
- Kemp, J.L.C. 1999. *Fractal organising of knowledge intensive organisations.*
- Klement, Nathalie, Cristóvão Silva, et Olivier Gibaru. 2017. « A Generic Decision Support Tool to Planning and Assignment Problems: Industrial Application & Industry 4.0 ». *Procedia Manufacturing* 11: 1684-91.
- Kolodner J., L. 1992. « An Introduction to Case-Based Reasoning ». *Artificial Intelligence Review* 6(1): 3–34.
- Kompridis, N. 2000. « So We Need Something Else for Reason to Mean ». *International Journal of Philosophical Studies* 8(3): 271–295.
- Kong, G. et al. 2012. « A Belief Rule-Based Decision Support Systems for Clinical Risk Assessment ». *European Journal of Operational Research* 219: 564–573.
- Krima, Sylvere et al. 2009. « OntoSTEP: OWL-DL Ontology for STEP ». In *Proceedings of the International Conference on Product Lifecycle Management PLM'09.*
- Kulhavy, R. 2003. « A developer's perspective of a decision support system ». *IEEE Control Systems Magazine* 23(6): 40–49,.
- Kuljanic, E., G. Totis, et Marco Sortino. 2009. « Development of an Intelligent Multisensor Chatter Detection System in Milling ». *Mechanical Systems and Signal Processing* 23: 1704–1718.
- Kulvatunyou, B. et al. 2018. « The Industrial Ontologies Foundry Proof-of Concept Project: IFIP WG 5.7 ». In *International Conference, APMS 2018, Seoul, Korea, August 26-30, 2018, Proceedings, Part II.*
- Labrousse, M. 2004. « Proposition d'un Modèle Conceptuel Unifié pour la Gestion Dynamique des Connaissances d'entreprise, Thèse de doctorat ». *École Centrale de Nantes et l'Université de Nantes.*
- Lalouette, C. 2013. *Gestion des connaissances et fiabilité organisationnelle : état de l'art et illustration dans l'aéronautique. Number 2013-01 of the Cahiers de la Sécurité Industrielle, Foundation for an Industrial Safety Culture.* Toulouse, France.
- Laroche, F. 2017. *KLM for Heritage, Thèse d'Habilitation à Diriger des Recherches.* Université de Nantes.
- Laroche, Florent, Virginie Rosato, et Paul François. 2020. « Patrimoine Industriel Numérique 3D : Reconstruction de la forge d'Arthez d'Asson ». *CILAC, Révéler le Patrimoine*

- Industriel* (74-75): 114–121.
- Larrouy, N. 1989. *Arthez d'Asson, Fer et Charbon, Éditions Marrimpouey*.
- Lau, H.C.W. et al. 2005. « A knowledge-based system to support procurement decision ». *J. Knowl. Mgmt* 9(1) : 87–100.
- Le Boterf G. 2000. "Construire les compétences individuelles et collectives", éd. d'organisation, Paris
- Le Duigou, J. 2010. « Cadre de modélisation pour les systèmes PLM en entreprise étendue : Application aux PME mécaniciennes ». *Thèse de doctorat de l'École Centrale de Nantes*.
- Lee, Jay, et Hung-An Kao. 2014. « Shanhu Yang, Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment ». *Procedia CIRP* 16: 3–8,.
- Lee, J.; Bagheri, B.; Kao, H.A. 2015. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manuf. Lett.* 3, 18–23
- Leffondré, Karen, Marthe-Aline Jutand, Marianne Savès, et Valérie Kiewsky. 2017. « Enseigner le recueil de données : étude de cas en épidémiologie ». *Statistiques et enseignement* 8(2).
- Lemaignan, S., A. Siadat, J.Y. Dantan, et Semenenko A. Mason. 2006. « A Proposal for an Ontology of Manufacturing Domain ». In *Distributed Intelligent Systems: Collective Intelligence and Its Applications, IEEE Workshop*, , 195–200.
- Li, Dan, Anna Landström, Åsa Fast-Berglund, et Peter Almström. 2019. 84 *Human-Centred Dissemination of Data, Information and Knowledge in Industry 4.0, Procedia CIRP*. <https://doi.org/10.1016/j.procir.2019.04.261>.
- Lin, C., H.-C. Hung, J.-Y. Wu, et B. Lin. 2002. « A Knowledge Management Architecture in Collaborative Supply Chain ». *J. Comput. Inform. Syst* 42(5): 83–94.
- Liu, Shaofeng, et Pascale Zaraté. 2014. *Knowledge Based Decision Support Systems: A Survey on Technologies and Application Domains*. 10.1007/978-3-319-07179-4_7.
- Lu, Yang. 2017. « Industry 4.0: A Survey on Technologies, Applications and Open Research Issues ». *Journal of Industrial Information Integration* 6: 1–10,.
- Maedche, A., et S. Staab. 2001. « Ontology Learning for the Semantic Web ». *IEEE Intelligent Systems* 16(2): 72–79,.
- Maedche, Alexander, et Steffen Staab. 2004. « Ontology Learning ». In *Handbook on Ontologies*, International Handbooks on Information Systems, éd. Steffen Staab et Rudi Studer. Berlin, Heidelberg: Springer, 173-90. https://doi.org/10.1007/978-3-540-24750-0_9 (16 novembre 2020).
- Marakas, George M. 2003. *Decision Support Systems in the 21st Century*. Upper Saddle River, NJ: Prentice Hall. <http://books.google.com/books?id=8OIOAAAAMAAJ> (16 novembre 2020).

- Marques, Maria, Carlos Agostinho, Gregory Zacharewicz, et Ricardo Jardim-Gonçalves. 2017. « Decentralized Decision Support for Intelligent Manufacturing in Industry 4.0 ». *JAISE* 9: 299–313 10 3233 –170436.
- Martin, P., et A. D'Acunto. 2003. « Design of a Production System: An Application of Integration Product-Process ». *Int. J. Computer Integrated Manufacturing* 16(7-8): 509–516.
- Masolo, C. et al. 2004. « Social Roles and Their Descriptions ». In *KR*, , 267–277.
- Matta, N., et J.L Ermine. 2001. *Gestion des connaissances et conception mécanique, Journée AIP-Priméca « Dynamique des connaissances en conception : acquisition, capitalisation et réutilisation »*. Grenoble, France.
- Mattsson, S., Å. Fast-Berglund, D. Li, et P. Thorvald. 2018. « Forming a Cognitive Automation Strategy for Operator 4.0 in Complex Assembly ». *Computers & Industrial Engineering*.
- McGuinness, D.L., et F.V. Harmelen. 2004. « OWL Web Ontology Language Overview [Online ». Available from. <http://www.w3>.
- McLean, C. 2005. *Shop Data Model and Interface Specification*. Gaithersburg, Maryland: National Institute of Standards and Technology. Technical report NISTIR 7198.
- M.E.D.E.F. 2002. *L'influence des contextes nationaux sur le management par les compétences*", *Cahiers du MEDEF, Objectif compétences, des pratiques européennes innovantes*. France.
- Meski, O., Laroche, F. Belkadi, F. et B. Furet. *La structuration des connaissances au service de l'industrie 4.0 : Le cas du projet " SmartEmmma "*. 16ème colloque national Smart/AIPPRIMECA, Apr 2019b, les Karelis. France.
- Millot, Patrick, et Marie-Pierre Pacaux-Lemoine. 1998. 1 *An attempt for generic concepts toward human-machine cooperation*.
- Mintzberg, H., Raisinghani D., et Théorêt A. 1976. « The Structure of Unstructured Decision Processes" ». *Administrative Science Quarterly* 21: 246–275.
- Missikoff, M., R. Navigli, et P. Velardi. 2002. « Integrated approach to Web ontology learning and engineering ». *Computer* 35(11): 60–63,.
- Monostori, L. 2016. « Cyber-Physical Systems in Manufacturing ». *CIRP Annals - Manufacturing Technology* 65: 621–641.
- Montclos, J.-M. 2011. *Architecture, description et vocabulaire méthodiques, Éditions du Patrimoine, Centre des Monuments Nationaux*. Paris.
- Morris, Seymour, Rome Laboratory (Griffiss Air Force Base, N.Y.), et Reliability Analysis Center (U.S.). 1995. *Reliability Toolkit: Commercial Practices Edition: A Practical Guide for Commercial Products and Military Systems Under Acquisition Reform*. Reliability Analysis Center.

- Moulahi, Bilel. 2015. « Définition et évaluation de modèles d'agrégation pour l'estimation de la pertinence multidimensionnelle en recherche d'information ».
- Murray P.C. 1996. « Who owns knowledge management », *Information, Knowledge and Document Management Technology*, KM Metazine.
- Munir, Kamran, et M. Sheraz Anjum. 2018. « The Use of Ontologies for Effective Knowledge Modelling and Information Retrieval ». *Applied Computing and Informatics* 14(2): 116-26.
- Muthu, S., Mendonca Devadasan SR, P.S., et G. Sundararaj. 2001. « Pre-Auditing through a Knowledge Base System for Successful Implementation of a QS 9000 Based Maintenance Quality System ». *J. Qual. Maint. Engng* 7(2): 90–103.
- Napoli, A., Lieber J, et R. Curien. 1996. *Classification-Based Problem-Solving in Case Based Reasoning European Workshop on Advances in Case-Based Reasoning Springer*. Berlin Heidelberg.
- N.A.S.A. 1999. *Failure Modes, Effects and Criticality Analysis (FMECA). Practice No. PD-AP-1307*.
- Nazakat A. et J. Hoing. (2018). Failure Detection and Prevention for Cyber-Physical Systems Using Ontology-Based Knowledge Base. *Computers*. 7. 68. 10.3390/computers7040068.
- NIST Big Data Public Working Group Definitions and Taxonomies Subgroup. 2015. « NIST Big Data Interoperability Framework: Big Data Taxonomies ». *National Institute of Standards and Technology* 2: 31,.
- NIST Data Public Working Group, Definitions and Taxonomies Subgroup. 2015. « Nist Big Data Interoperability Framework: Definitions ». *National Institute of Standards and Technology* 1: 32,.
- Nonaka, I. et Takeuchi, et H. 1995. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford university press.
- Norvig, S.J. Russell, et P. 1995. *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence.
- Novins, P. et Armstrong, et R. 1998. « Choosing Your Spots for Knowledge Management ». *Perspectives on Business Innovation* 1: 45–54.
- Oberle, Daniel et al. 2007. « DOLCE Ergo SUMO: On Foundational and Domain Models in the SmartWeb Integrated Ontology (SWIntO) ». *Journal of Web Semantics* 5(ue 3): 156–174,.
- Oberle, M., Issah A., et Dias J. 2018. « Ontology-Based Reference Data Model. Work Package 6 Simulation and Decision Support Systems ». *Projet « ScalABLE* 4(0).
- Octaviani, Dewi, Andri Pranolo, et Mohd Othman. 2015. *RDB2Onto: An Approach for Creating Semantic Metadata from Relational Educational Data*.

- Olson, David L. 2018. « View of IJPR Contributions to Knowledge Management in Supply Chains ». *International Journal of Production Research* 56(1-2): 733–742,.
- Olson, D.L., et D. Wu. 2017. *Predictive Data Mining Models*. Springer.
- Ozbayrak, Mustafa, et Robert Bell. 2003. « A knowledge-based decision support system for the management of parts and tools in FMS ». *Decision Support Systems* 35: 487-515.
- Palmer, Claire et al. 2018. « Interoperable Manufacturing Knowledge Systems ». *International Journal of Production Research* 56(8): 2733–2752,.
- Panetto, H., M. Dassisti, et A. Tursi. « Onto-Pdm: Product-Driven Ontology for Product Data Management Interoperability within Manufacturing Process Environment ». *Advanced Engineering Informatics* 26;2012: 334–348.
- Paquette, G. 2002. *Modélisation des connaissances et des compétences*. Saite-Foy, Quebec: Presses de l’université du Quebec.
- Paquette, Gilbert. 2014. *cours INF 6500 Système à base de connaissances dans les organisations*.
https://inf6500.teluq.ca/teluqDownload.php?file=2014/07/SBC_Texte3.pdf.
- Parasuraman, R., et V. Riley. 1997. « Humans and Automation: Use, Misuse, Disuse, Abuse ». *Human Factors* 39(2): 230–253.
- Patwardhan, Amit, Ajit Verma, et Uday Kumar. 2016. « A Survey on Predictive Maintenance Through Big Data ». In *Lecture Notes in Mechanical Engineering*, , 437-45.
- Peres, Ricardo, Andre Rocha, Paulo Leitão, et J. Barata. 2018. « IDARTS – Towards intelligent data analysis and real-time supervision for industry 4.0 ». *Computers in Industry* 101: 138-46.
- Philippot, A., M. Sayed-Mouchaweh, V. Carre-Menetrier, et B. Riera. 2006. « Decentralized Approach to Diagnose Manufacturing Systems ». In *The Proceedings of the Multiconference on « Computational Engineering in Systems Applications »*, , 912-18.
- Poslad, S. 2007. « Specifying Protocols for Multi-Agent System Interaction ». *ACM Trans. Autonom. Adapt. Syst* 2, 4, Article 15: 24.
- Prax, J.-Y. 2000. « Le Guide Du Knowledge Management ». *Dunond*.
- Psyché, V. 2007. *Rôle des ontologies en ingénierie des ELAH : cas d’un système d’assistance au design pédagogique*. Thèse de doctorat, Université du Québec à Montreal.
- Qi, Q., et F. Tao. 2018. « Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison ». *IEEE Access* 6: 3585–3593,.
- Qiao, Lihong et al. 2017. « An ontology-based modelling and reasoning framework for assembly sequence planning ». *International Journal of Advanced Manufacturing Technology* 94(9-12): 4187-97.
- Rabah, N Ben et al. 2016. « Intelligent Case Based Decision Support System for Online

- Diagnosis of Automated Production System ». In *European Workshop on Advanced Control and Diagnosis (ACD)*, Journal of Physics: Conference Series, Lille, France, 012009. <https://hal.archives-ouvertes.fr/hal-01899933> (16 novembre 2020).
- Rachuri, S. et al. 2003. « Object-Oriented Representation of Electro- Mechanical Assemblies Using UML ». *NIST*.
- Rajkovic, V., J. Efstathiou, et M. Bohanec. 1987. « A Concept of Rule-Based Decision Support Systems ». In *Optimization Models Using Fuzzy Sets and Possibility Theory*, Theory and Decision Library, éd. J. Kacprzyk et S. A. Orlovski. Dordrecht: Springer Netherlands, 292-99. https://doi.org/10.1007/978-94-009-3869-4_20 (16 novembre 2020).
- Ramis, Borja et al. 2016. « Product, process and resource model coupling for knowledge-driven assembly automation / Kopplung von Produkt-, Prozess- und Ressourcenmodell für die wissensgetriebene Montageautomation ». *at - Automatisierungstechnik* 64.
- Rasmussen, J. 1993. « Diagnostic Reasoning in Action ». *IEEE Transactions on Systems, Man, and Cybernetics* 23(4): 981–992
- Reichow, D. 1991. « Fault Management in a Modern Airliner ». *IFAC Proceedings Volumes* 24(6): 1-8.
- Rezaie, K. et al. 2007. « Using Extended Monte Carlo Simulation Method for the Improvement of Risk Management ». *Consideration of relationships between uncertainties. Mathematics and Computation* 190(2): 1492–1501.
- Rezaie, K., A. Gereie, B. Ostadi, et M. Shakhsheniaee. 2008. « Safety Interval Analysis: A Risk-Based Approach to Specify Low-Risk Quantities of Uncertainty for Contractor's Bid Proposals. Computer & ». *Industrial Engineering*, 56(1): 152–156.
- Ritou, M. et al., éd. 2007. « Instrumentation de la broche par des capteurs inductifs : application à la mesure d'effort et à l'étude du comportement des broches UGV ». In *J Hascoet 10ème Colloque National AIP PRIMECA*,.
- . 2018. « Influence of Spindle Condition on the Dynamic Behavior ». *CIRP Ann Manuf Technol* 67: 419–422.
- . 2019. « Knowledge-Based Multi-Level Aggregation for Decision Aid in the Machining Industry ». *CIRP Annals - Manufacturing Technology, Elsevier* 68(1): 475–478.
- Rowley, J. 2007. « The Wisdom Hierarchy: Representations of the DIKW Hierarchy ». *Journal of Information and Communication Science* 33(2): 163–80.
- Rumbaugh, J. et al. « OMT: Modélisation et Conception Orientées Objet », Masson, Prentice Hall, 1995.
- Saaksvuori, Antti, et Anselmi Immonen. 2008. *Product Lifecycle Management*. 3^eed. Berlin Heidelberg: Springer-Verlag. <https://www.springer.com/gp/book/9783540781738> (16 novembre 2020).

- Sacco, M. et al. 2011. « Virtual Factory Manager. Vol 6774 ». In *Virtual and Mixed Reality - Systems and Applications, Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, 397–406.
- Saddem, Ramla, et Alexandre Philippot. 2014. « Causal Temporal Signature from diagnoser model for online diagnosis of Discrete Event Systems ». *Proceedings - 2014 International Conference on Control, Decision and Information Technologies, CoDIT 2014*: 551-56.
- Sanfilippo, E.M., et S. Borgo. 2015. « Feature-Based Modelling and Information Systems for Engineering ». In *AI*IA 2015: AI*IA 2015 Advances in Artificial Intelligence*, Cham: Springer, 151–163.
- Sanya, I.O., et E.M. Shehab. 2015. « A Framework for Developing Engineering Design Ontologies within the Aerospace Industry ». *International Journal of Production Research* 53(8): 2383–2409.
- Schlenoff, C. et al. 2000. *ISO-18629 The Process Specification Language (PSL): Overview and Version 1.0 Specification, NISTIR 6459*. Gaithersburg, MD: National Institute of Standards and Technology.
- Schoenmackers, Stefan, Oren Etzioni, Daniel S. Weld, et Jesse Davis. 2010. « Learning First-Order Horn Clauses from Web Text ». In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 1088–1098.
- Schreiber, Guus et al. 2001. *Knowledge Engineering and Management – The CommonKADS Methodology*.
- Schwartz, Stéphane, Irène Abi-Zeid, et Nicole Tourigny. 2007. « L'ingénierie des connaissances pour la modélisation du raisonnement dans une tâche de diagnostic: Application à la recherche et sauvetage ». *Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration* 24: i-xvi.
- Shim, J., M. Warkentin, et J. Courtney. 2002. *Past, present, and future of decision support technology (Decision Support System 33)*.
- Silberschatz, Abraham, et Alexander Tuzhilin. 1995. *On Subjective Measures of Interestingness in Knowledge Discovery*.
- Simon, H.A. 1965. « How to Make a Business Decision ». *Archer, E.R.*: 54–61.
- Sirin, Evren et al. 2007. « Pellet: a practical OWL-DL reasoner ». *Web Semantics: Science, Services and Agents on the World Wide Web* 5: 51-53.
- Song, F., G. Zacharewicz, D. Chen. 2013. « An ontology-driven framework towards building enterprise semantic information layer ». *Adv. Eng. Inform.*, 27 (1), pp. 38-50.
- Soon, G.K., C. Kim On, P. Anthony, et A.R. Hamdan. 2019. « A Review on Agent Communication Language ». In *Computational Science and Technology, Lecture Notes in Electrical Engineering*, éd. Rayner Alfred, Yuto Lim, Ag Asri Ag Ibrahim, et Patricia Anthony. Singapore: Springer, 481-91.

- Stankovic, J.A. 2014. « Research Directions for the Internet of Things ». *IEEE Internet of Things Journal* 1(1): 1–7.
- Steels, L. 1990. « Components of Expertise ». *AI Magazine*: 27–49.
- Stentoft, Jan, Kent Wickstrøm Jensen, Kristian Philipsen, et Anders Haug. « Drivers and Barriers for Industry 4.0 Readiness and Practice: A SME Perspective with Empirical Evidence ». : 10.
- Struss, P., A. Malik, et M. Sachenbacher. 1996. « Qualitative Modeling Is the Key to Automated Diagnosis ». *IFAC Proceedings* 29(ue 1): 6365–6370,.
- Sureephong, Pradorn, Nopasit Chakpitak, Yacine Ouzrout, et Abdelaziz Bouras. 2008. « An Ontology-Based Knowledge Management System for Industry Clusters ». In *International Conference on Advanced Design and Manufacture (ICADAM 2008)*, Sanya, China.
- Tari, Luis. 2013. « Knowledge Inference ». In *Encyclopedia of Systems Biology*, éd. Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, et Hiroki Yokota. New York, NY: Springer, 1074-78. https://doi.org/10.1007/978-1-4419-9863-7_166 (15 novembre 2020).
- Terkaj, W., G. Pedrielli, et M. Sacco. 2012. « Virtual Factory Data Model ». In *Proceedings of the Workshop on Ontology and Semantic Web for Manufacturing*, Graz, Austria.
- Teti, R., K. Jemielniak, G.E. O'Donnell, et D. Dornfeld. 2010. *Advanced Monitoring of Machining Operations. CIRP Annals- Manufacturing Technology*.
- Toro, C., I. Barandiaran, et J. Posada. 2015. « A Perspective on Knowledge Based and Intelligent Systems Implementation in Industrie 4.0 ». *Procedia Computer Science* 60: 362–70.
- Uddin, M.K., A. Dvoryanchikova, A. Lobov, et J.L.M. Lastra. 2011. « An Ontology-Based Semantic Foundation for Flexible Manufacturing Systems ». In *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, Melbourne, VIC, 340–345,.
- Uhlmann, Eckart, Abdelhakim Laghmouchi, Claudio Geisert, et Eckhard Hohwieler. 2017. « Decentralized Data Analytics for Maintenance in Industrie 4.0 ». *Procedia Manufacturing* 11: 1120-26.
- Vernadat, F. 2002. « UEML: Towards a Unified Enterprise Modelling Language ». *International Journal of Production Research* 40(17): 59.
- Villemin, F.Y. 2008. *Techniques de Représentation des Connaissances*.
- Vrba, Pavel, Miloslav Radakovič, Marek Obitko, et Vladimír Mařík. 2011. « Semantic Technologies: Latest Advances in Agent-Based Manufacturing Control Systems ». *International Journal of Production Research* 49(5): 1483–1496,.
- Wakulicz-Deja, Alicja, Agnieszka Nowak-Brzezińska, et Tomasz Jach. 2011. « Inference Processes in Decision Support Systems with Incomplete Knowledge ». In *Rough Sets*

- and Knowledge Technology*, Lecture Notes in Computer Science, éd. JingTao Yao, Sheela Ramanna, Guoyin Wang, et Zbigniew Suraj. Berlin, Heidelberg: Springer, 616-25.
- Wang, W.M., et C.F. Cheung. 2011. « A Narrative-Based Reasoning with Applications in Decision Support for Social Service Organizations ». *Expert Systems with Applications* 38: 3336–3345.
- Wang, Zhiqiang, Mathieu Ritou, Catherine Cunha, et Benoit Furet. 2019. « Classification contextuelle pour système d'aide à la décision pour machines-outils ». *Colloque National S-mart/AIP-PRIMECA, Apr.*
- Weggeman, Mathieu. 1997. *Kennismanagement: inrichting en besturing van kennisintensieve organisaties*. Scriptum Management.
- Weiß, Gerhard. 2001. « Cognition, Sociability, and Constraints ». In *Balancing Reactivity and Social Deliberation in Multi-Agent Systems*, Lecture Notes in Computer Science, éd. Markus Hannebauer, Jan Wendler, et Enrico Pagello. Berlin, Heidelberg: Springer, 217-35.
- Westkmper, E. 2013. Towards the Re-Industrialization of Europe: A Concept for Manufacturing for 2030 *Towards the re-industrialization of Europe: A concept for manufacturing for 2030*
- Wu, X., Goepp, V. & Siadat, A. 2020. « Concept and engineering development of cyber physical production systems: a systematic literature review ». *Int J Adv Manuf Technol* 111, 243–261. <https://doi.org/10.1007/s00170-020-06110-2>
- Xiong, N., Funk P, et T. Olsson. 2012. Case-Based Reasoning Supports Fault Diagnosis Using Sensor Information *2nd International Workshop and Congress on EMaintenance*.
- Yang, H., W. Li, K. Liu, et J. Zhang. 2011. « Knowledge-Based Clinical Pathway for Medical Quality Improvement ». In *Information Systems Front.* DOI 10.1007/S10796-011-9307-z.
- Young, R.I.M. 2007. « Manufacturing Knowledge Sharing in PLM: A Progression towards the Use of Heavy Weight Ontologies' ». *International Journal of Production Research* 45(7): 1505–1519 10 1080 00207540600942268.
- Zadeh, L. A. 1965. « Fuzzy Sets ». *Information and Control* 8(3): 338-53.
- Zemirline, A., L. Lecornu, B. Solaiman, et A. Ech-cherif. 2008. « An Efficient Association Rule Mining Algorithm for Classification ». In *Artificial Intelligence and Soft Computing – ICAISC 2008*, Lecture Notes in Computer Science, éd. Leszek Rutkowski, Ryszard Tadeusiewicz, Lotfi A. Zadeh, et Jacek M. Zurada. Berlin, Heidelberg: Springer, 717-28.
- Zhang, B., et L. 1992. *Theory and Applications of Problem Solving 9, North-Holland Amsterdam*.
- Zhang, R., J. Lu, et G. Zhang. 2011. « A Knowledge-Based Multi-Role Decision Support System for Ore Blending Cost Optimisation of Blast Furnaces ». *European Journal of*

Operational Research 215: 195–203.

Zighed, D., et R. Rakotomalala. 2000. « Graphes d'induction : apprentissage automatique et data mining ». *Hermes*: 82–88,.

Zins, C. 2007. « Conceptual Approaches for Defining Data, Information, and Knowledge ». *Journal of the American Society for Information Science and Technology* 58(4): 479–493.

Zsidisin, G.A., S.A. Melnyk, et G.L. Ragatz. 2005. « An Institutional Theory Perspective of Business Continuity Planning for Purchasing and Supply Management ». *International Journal of Production Research* 43(16): 3401–3420,.

Site web

[Site Web frenchfab] www.lafrenchfab.fr

[Site Web Techopedia] www.techopedia.com/definition/29951/big-data-platform

[Site Web GDT, 1999] www.gdt.oqlf.gouv.qc.ca

[Site web phpMyAdmin] www.phpmyadmin.net

[Site Web CIGREF, 2000] www.cigref.fr

Liste de annexes

Annexe 1: Le consortium du projet

Annexe 2: Diagrammes SADT - analyse des processus industrielles

Annexe 3: Les langages de modélisation

Annexe 4: Analyse des normes et standards

Annexe 1 : Le consortium du projet

Le Laboratoire des Sciences du Numérique de Nantes

Le LS2N est une nouvelle Unité Mixte de Recherche (UMR 6004) créée en janvier 2017 qui résulte de la fusion des UMR IRCCyN (UMR 6597 : Institut de Recherche en Communications et Cybernétique de Nantes), et LINA (UMR 6241 : Laboratoire d'Informatique de Nantes Atlantique). Réunissant 450 personnes au cœur des sciences du numérique, ce laboratoire a pour ambition de faire progresser significativement la visibilité de la recherche en Cybernétique et Informatique à Nantes. Les talents scientifiques du LS2N participent à la révolution numérique de notre société sur les sujets scientifiques et techniques qu'elle met en œuvre.

Deux équipes du LS2N sont impliquées dans ce projet de recherche :

IS3P (Ingénierie des Systèmes : Produits, Processus, Performances) est une équipes du LS2N qui a pour but de développer des méthodes et outils pour la modélisation, l'analyse, la conception et le pilotage des systèmes sociotechniques. La gestion des connaissances, l'amélioration des performances et des systèmes d'information métier sont au cœur des problématiques de recherche de l'équipe.

RoMas (Robots and Machines for Manufacturing, Society and Services) est une équipe du LS2N qui s'intéresse notamment à l'optimisation des procédés de fabrication (dont l'UGV) et au développement d'outils nécessaires à la conception, la mise au point, l'exploitation de robots, de machines, de systèmes ou de moyens pour des applications industrielles ou de la vie courante.

Le groupe Europe Technologies

une entreprise de taille moyenne employant 350 personnes pour 53 M€ de chiffre d'affaires en 2015. L'activité est répartie sur de nombreux secteurs comme : Aéronautique (41%), Naval et Défense (20 %), l'énergie (20 %), l'automobile (10 %) et le reste dans l'alimentaire, le médical et l'industrie lourde. Le groupe est composé de 8 filiales organisées en 3 business units : BU Technologique, BU mécanique et BU composite. Ses filiales sont impliquées dans de nombreuses activités de recherche par le développement de procédés, la qualification des procédés, l'outillage et les machines de la fabrication jusqu'au service après-vente.

Dans le projet SmartEmma Europe technologies a en charge la mise en place, l'intégration et le contrôle de la nouvelle plateforme SmartEmma connectée au système de surveillance Emmatools.

Mecachrome

L'industrie Mecachrome France, un leader de l'industrie mécanique, avec un chiffre d'affaires de 335M€ et 2400 employés. Elle est présente dans l'automobile (16 %), l'aéronautique (81 %), la défense et l'énergie (3 %). Mecachrome est fournisseur de grandes entreprises aéronautiques telles qu'Airbus, Safran, Dassault, Rolls-Royce, ... Mecachrome produit de grandes pièces pour les structures aéronautiques (structure d'ailerons, fuselage...) structure, structure et nacelle du pylône moteur, baies de trains d'atterrissage, ...) et pièces pour l'aéronautique.

Airbus

Airbus est une référence internationale dans le secteur aérospatial. Ils conçoivent, fabriquent et livrent des avions commerciaux, des hélicoptères, des moyens de transport militaires, des satellites et des lanceurs de premier plan, tout en fournissant des services de données, de navigation, de communications sécurisées, de mobilité urbaine et d'autres solutions à leurs clients à l'échelle mondiale.

Grâce à une stratégie prospective fondée sur les technologies de pointe, le numérique et l'excellence scientifique, Ils visent un monde mieux connecté, plus sûr et plus prospère.

Le siège social de Airbus se trouve à Blagnac (Toulouse, France). Cette entreprise fut fondée en tant que consortium par plusieurs fabricants européens à la fin des années 1960. Filiale d'EADS, elle a été renommée Airbus Group en 2014 puis Airbus en 2017. Avec 130 000 salariés et un chiffre d'affaires de 66,8 milliards d'euros en 2017, le groupe est l'un des champions mondiaux du secteur aéronautique et spatial, toutes catégories confondues.

Annexe 2 : Diagramme SADT - analyse des processus industrielles

Ci-dessous les différents diagrammes SADT qui ont été développés pour le projet afin d'analyser la chaîne numérique industrielle chez Mécachrome. Parmi les étapes les plus importantes de la fabrication d'une pièce aéronautique c'est l'industrialisation puisqu'elle comporte toute la partie conversion du modèle CAO en modèle FAO, la génération des trajectoires, de la gamme de fabrication...etc. Mais surtout la génération du code G nécessaire pour l'usinage de la pièce. L'enchaînement et la relation entre ces étapes sont représentés par le modèle SADT suivant :

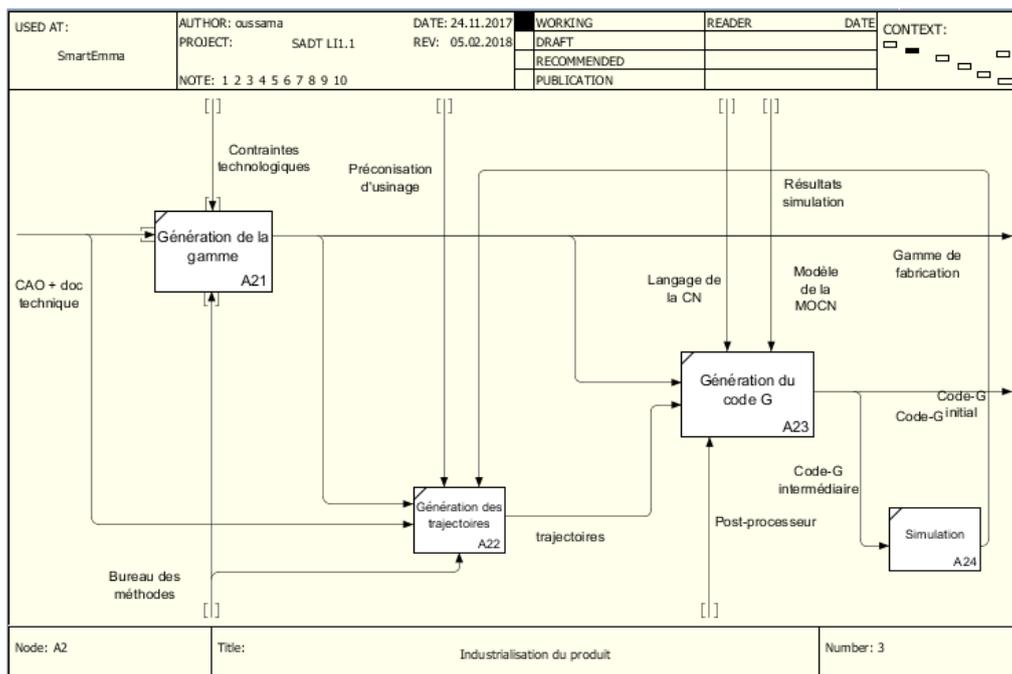


Figure 80. Industrialisation du produit

Une étape intermédiaire très importante aussi qui se situe entre l'industrialisation et la fabrication est la planification et la préparation de la production puisqu'elle permet de bien piloter les opérations d'usinage à travers l'ordre de fabrication mais aussi, elle permet de lancer le processus d'approvisionnement de la matière première.

La planification et la préparation de la production peut être représentée par le modèle SADT suivant :

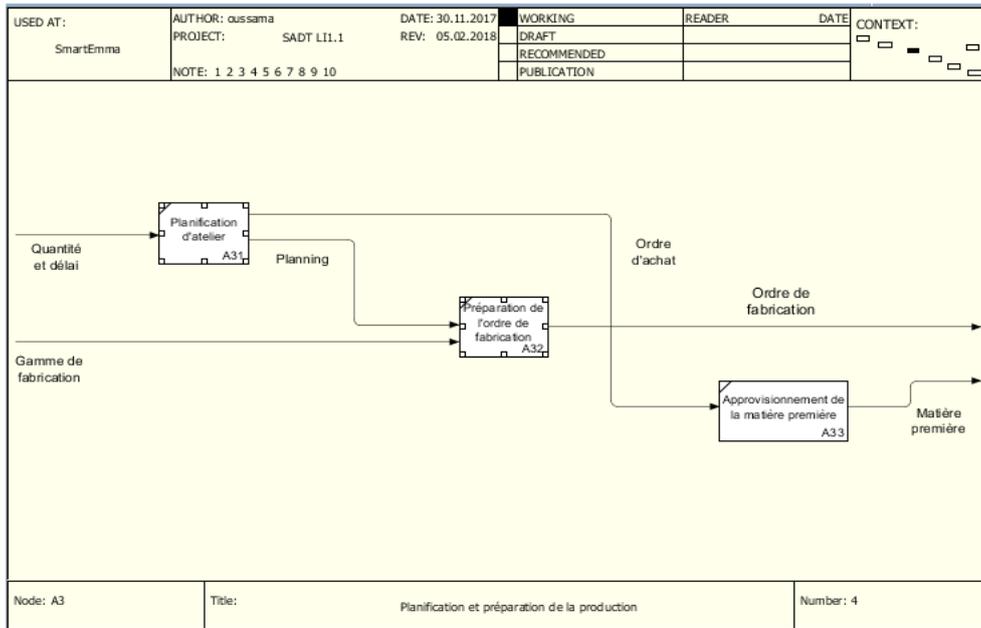


Figure 81. Planification et préparation de la production

L'étape de base dans tout ce processus de réalisation c'est la fabrication du produit qui en se basant sur une gamme de fabrication, un code G et de la matière première permet de faire dans un premier temps l'homologation du code ISO, c'est-à-dire de fabriquer la première pièce et de bien vérifier le bon fonctionnement du code développé à travers l'évaluation de la fiabilité et la conformité de la première pièce usinée, et par la suite de lancer l'usinage des autres pièces. Ou de faire les modifications nécessaires au niveau du code G en se basant sur le rapport qualité de la première pièce. Le Diagramme SADT de la partie homologation du code G est le suivant

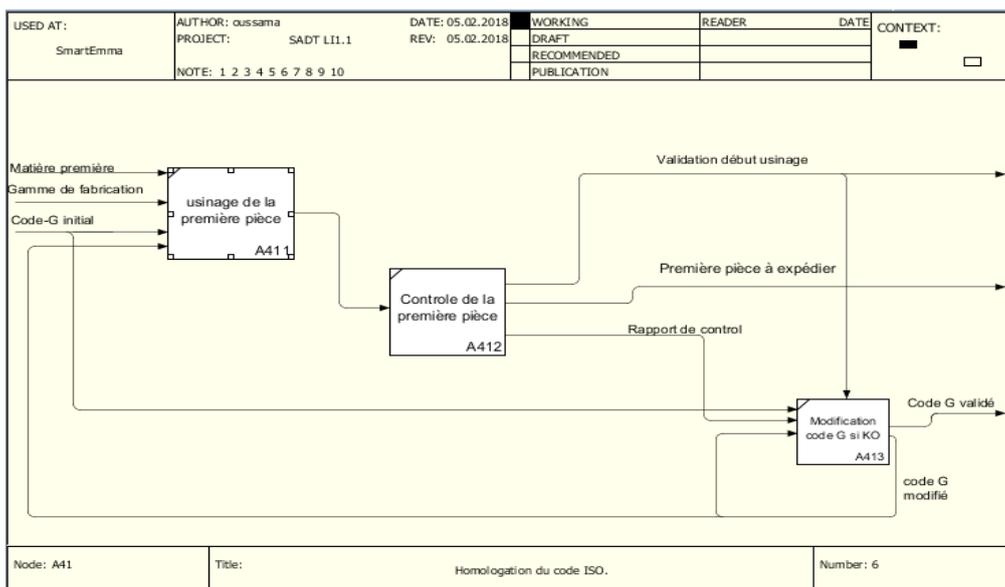


Figure 82. Homologation du code ISO

Le diagramme SADT suivant montre une vue globale de l'activité Fabrication du produit :

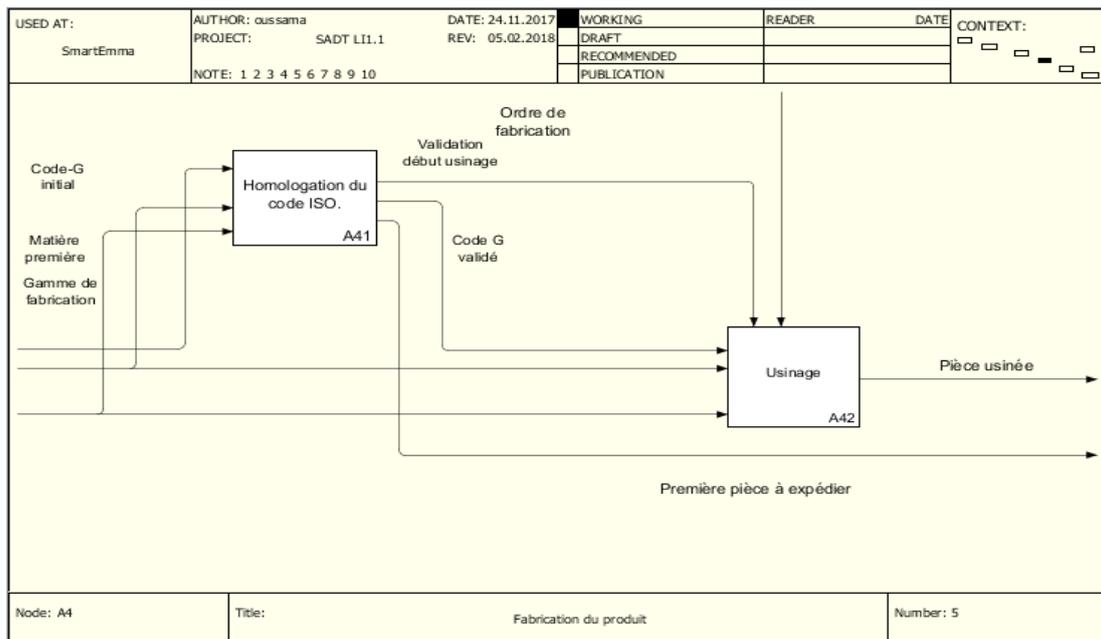


Figure 83. Fabrication du produit

Le processus d'usinage se compose de plusieurs étapes en amont et en aval en plus de la réalisation des opérations d'usinage par la machine à outil sur la matière première. Ces étapes sont représentées dans le modèle SADT suivant :

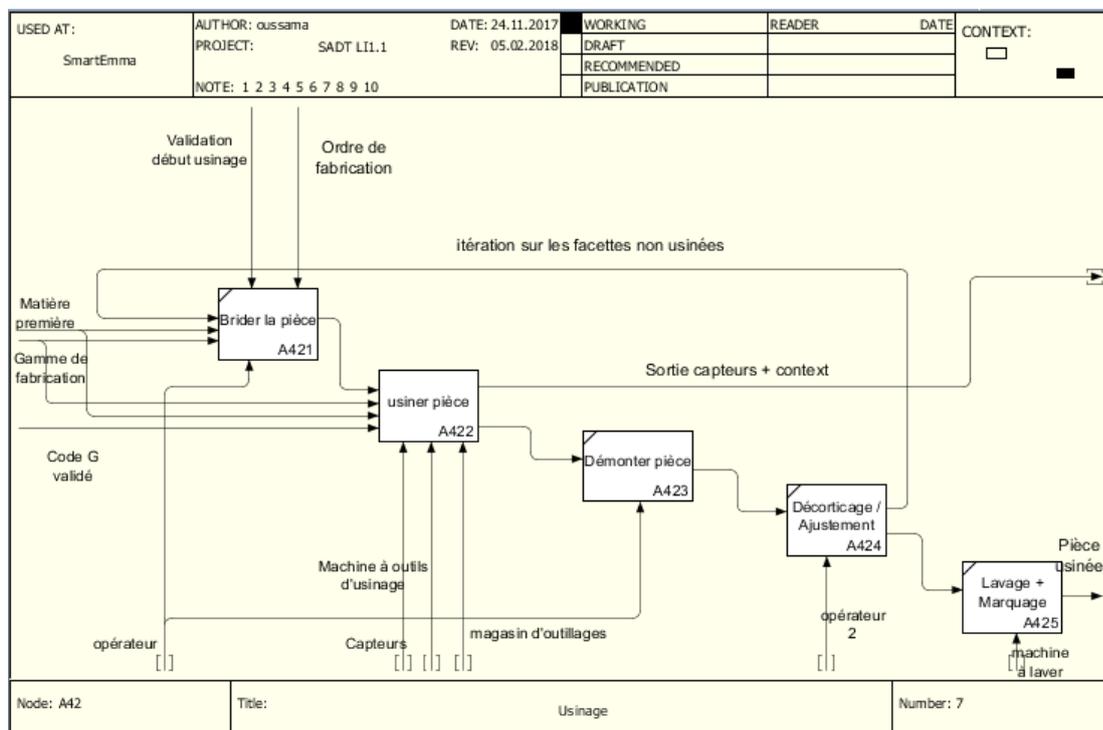


Figure 84. Usinage

La dernière étape avant l'expédition du produit c'est le contrôle de la qualité du produit et éventuellement le parachèvement dans les cas de non-conformité des pièces usinées toute en respectant l'ordre de fabrication, et afin de fournir un produit final qualifié avec son rapport de contrôle qualité. Chez Mécachrome les étapes de contrôle sont représentées par le diagramme SADT suivant

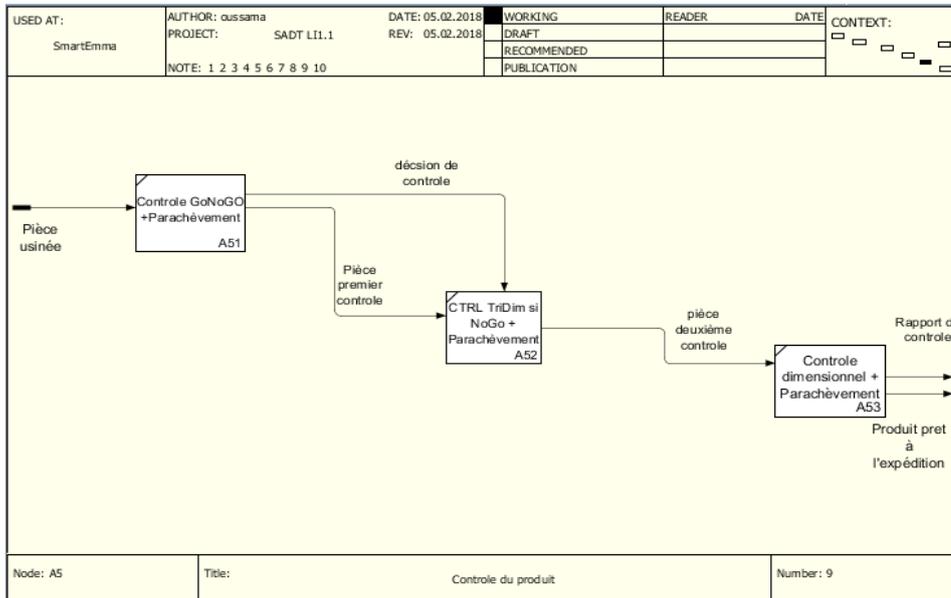


Figure 85. Contrôle du produit

L'implantation de l'EmmaTools dans Mécachrome est montrée dans le SADT représenté dans la figure suivante :

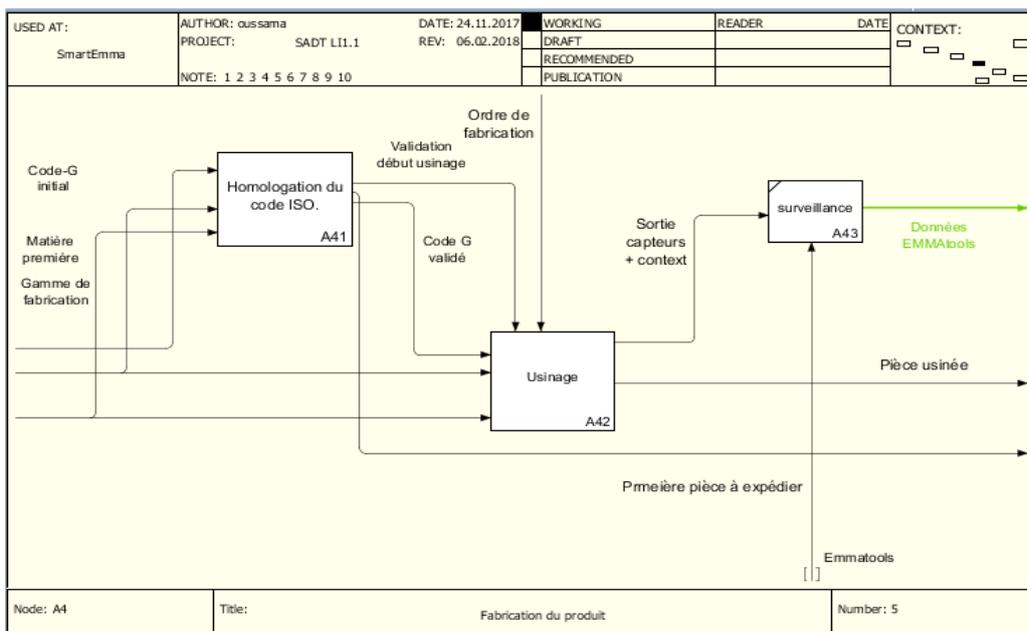


Figure 86. As-Is avec l'EmmaTools

Annexe 3 : Les langages de modélisation

Un langage de modélisation est généralement graphique. Pour développer ces graphes, les normes et standards sont mises en place pour définir les différents types de diagrammes incluant : des symboles pour représenter les concepts et les objets de base, des liens pour représenter les relations entre ces objets et un ensemble d'annotations pour les cardinalités, les contraintes, etc. Il peut aussi être textuel contrôlé par des standards qui fixent des mots clés et des symboles afin de construire des expressions et des relations interprétables par les ordinateurs. Ils existent des langages de modélisations disponibles dans les deux versions graphique et textuelle, par exemple EXPRESS-G et EXPRESS (ISO 10303-11).

L'OMG, le consortium américain de standardisation, a développé plusieurs langages de modélisation, dont nous pouvons citer à titre d'exemple : *UML*, *BPMN*, *SysML*, *DMN*, *BMM*, *OCL*, *MDA*, *QVT*, *XMI*, etc. D'autres langage et méthodes de modélisations ont été développé tels que : les méthodes IDEF, les VSM pour cartographier les chaînes de valeur, le MOT, les logigrammes, etc.

Les recherches dans le domaine de la modélisation des connaissances sont souvent liées aux web sémantiques [Berners-Lee et al., 2001], d'où le recours à l'utilisation de ses langages spécialement conçus pour les données, généralement par le NIST (National Institute of Standards and Technology) par exemple : RDF, XML et OWL [McLean et al., 2005 ; McGuinness et Harmelen, 2004]. Parmi les méthodes les plus efficaces pour assurer la compréhension sémantique, il y a les ontologies [Vrba et al., 2011].

Afin de comparer entre les différents langages de modélisation et choisir les plus adéquats pour chaque projet et ses contraintes, il faut se baser sur plusieurs critères. Ci-dessous, une méthode de comparaison est présentée. Elle est basée sur les typologies des modèles (fonctionnel, structurel, comportemental) pour comparer entre les langages de modélisation les plus utilisées.

- Un modèle fonctionnel permet de décrire le comportement de fonctionnement des composants individuels pour la réalisation de la fonction principale par le concepteur au système.
- Un modèle structurel, consiste à la caractérisation de la structure physique d'un système physique, en décrivant ses différentes parties et leur interconnexions.
- Un modèle comportemental décrit l'interaction entre les composants pour prévoir et contrôler le comportement d'un système physique.

Le tableau 1 compare l'adaptabilité de quelques langages de modélisation aux typologies des modèles décrites ci-dessus :

Tableau 7. Synthèse des langages de modélisation [Le Duigou, 2010]

	SADT	BPMN	XML	Express	UML	SysML
Fonctionnel	Non	Non	Non	Non	Partiel	Oui
Structurel	Non	Non	Oui	Oui	Oui	Oui
Comportemental	Oui	Oui	Non	Non	Oui	Oui

Dans le cadre de ces travaux de recherche, nous avons eu recours à l'utilisation de plusieurs langages de modélisation de plus en plus formels au fur et à mesure de l'avancement des développements. Pour la première phase d'analyse, afin d'étudier la chaîne numérique industrielle et comprendre les flux d'information circulants en entreprise, les diagrammes IDEFO (Icam DEFinition for Function Modeling) ont été utilisés. Ces diagrammes sont aussi appelés SADT, et ont permis de décrire la situation AS-IS chez les partenaires industriels du projet.

Par la suite, après identification et analyse des différentes sources d'informations disponibles, une première structuration formelle des données a été réalisée en utilisant le diagramme de classes du langage « UML ». Sur la base des modèles de données des nouveaux modèles, dit de connaissances, possédant un niveau d'abstraction plus poussé, seront développés.

La modélisation ne s'arrête pas à ce niveau, l'augmentation du besoin de granularité et les développements sémantiques ont impliqué le recours aux ontologies. Elles permettront le raisonnement et l'aide à la décision en intégrant des règles métier, qui génèrent les nouvelles taxonomies de concepts et les nouvelles formes de connaissances. Enfin, pour garantir la genericité des modèles développés et de la solution proposée, il est apparu primordial d'utiliser les normes et standards.

Ces différents langages de modélisation ainsi que normes intégrées seront détaillés dans la suite de ce chapitre.

La méthode IDEFO :

Cette technique connue aussi avec le nom SADT a été développée en 1977 aux Etats Unis, au profit de Softech par Doug Ross. En 1982, elle a été introduite en Europe par Michel Galinier (Waldner, 1993). IDEF0 est un standard permettant la description graphique d'un système complexe ou d'un processus opératoire à travers l'analyse fonctionnelle descendante. Autrement, l'analyse commence du général qui décrit une vision globale de système « niveau A-0 » vers le plus détaillé « niveaux A i j k ».

La méthode IDEF0 est utilisée dans plusieurs domaines : La télécommunication, l'avionique, l'armée, etc. C'est une méthode d'analyse par niveaux successifs d'approche descriptive d'un ensemble quel qu'il soit. Elle permet la spécification fonctionnelle de systèmes complexes. On peut appliquer IDEF0 à la gestion d'une entreprise tout comme à un système automatisé. Et en plus, c'est une méthode très optimale pour la communication, puisqu'elle a pour objectif de permettre les échanges aisés avec l'utilisateur ainsi de favoriser le travail en équipe.

Un actigramme SADT donne une vue fonctionnelle de l'activité (qu'est ce qu'elle fait ?). L'activité est représentée dans un actigramme par une boîte rectangulaire à laquelle est associée quatre types de données.

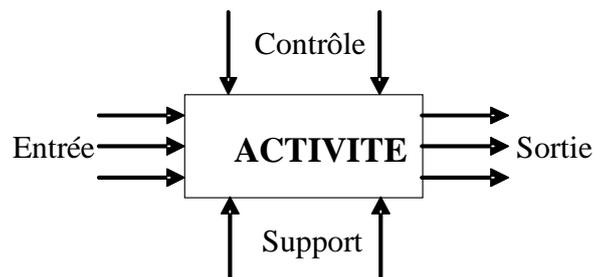


Figure 87. Le concept d'activité dans SADT

Chaque type de donnée est représenté par une flèche décrivant le sens (entrée/sortie) des données :

- Les entrées : C'est les données consommées, ou le flux de matière ou d'information intrant, qui sera transformé par l'activité.
- Les sorties : C'est les données résultantes, où le flux de matière ou d'information résultat de la transformation, ça sera soit des flux utiles ou des rebuts.
- Les contrôles : Ce sont en général, les données qui régissent l'activité. Ils décrivent les objectifs et les informations qui déclenchent l'activité.
- Le support : Ces données caractérisent les moyens et les ressources impliqués dans la réalisation de l'activité.

Le deuxième aspect est l'aspect hiérarchique par niveau de détail. Pour aborder progressivement la complexité, la méthode adopte une démarche descendante où le diagramme supérieur décrit une vision globale de système qui sera "éclaté" dans le niveau inférieur sur plusieurs boîtes, formant un nouveau sous actigramme SADT.

UML (Unified Modeling Language, langage de modélisation unifiée)

UML (Unified Modeling Language)

UML est un langage graphique de modélisation orienté objet issu de la fusion des méthodes objet OMT (Object Modeling Technique) et OOSE (Object Oriented Software Engineering), [Booch, 1996]. C'est un langage formel défini par l'organisme qui vise à standardiser et promouvoir le modèle objet : OMG (Object Management Group).

Le formalisme UML a été initialement éprouvé pour représenter la connaissance dans un projet européen dénommé ESPRIT. [Dhouieb,2016] redistingue les apports du formalisme UML par rapport à la représentation de la connaissance cités dans [Cranefield, 2006] :

- Une modélisation statique des objets pertinents du domaine
- Une modélisation dynamique des comportements
- Une compréhension facile par l'homme et le système grâce à la représentation graphique des diagrammes
- Une extraction de connaissance possible à partir des modèles grâce au raisonnement sur leur contenu.

Pour plus de sémantique, on parle de UML/OCL (Object Constraint Language). L'utilisation d'OCL consiste à l'attribution des contraintes et des règles sur un diagramme afin de compléter les spécifications d'une modélisation. Particulièrement, les contraintes OCL sont utilisées dans le diagramme de classe UML [OMG, 2003].

UML est un langage graphique de modélisation orienté objet issu de la fusion des méthodes objet OMT (Object Modelling Technique) et OOSE (Object Oriented Software Engineering), [Booch, 1996]. C'est un langage formel défini par l'organisme qui vise à standardiser et promouvoir le modèle objet: OMG (Object Management Group).

Le formalisme UML a été initialement éprouvé pour représenter la connaissance dans un projet européen dénommé ESPRIT. [Dhouieb,2016] redistingue les apports du formalisme UML par rapport à la représentation de la connaissance cités dans [Cranefield, 2006] :

- Une modélisation statique des objets pertinents du domaine
- Une modélisation dynamique des comportements
- Une compréhension facile par l'homme et le système grâce à la représentation graphique des diagrammes
- Une extraction de connaissance possible à partir des modèles grâce au raisonnement sur leur contenu.

UML présente treize types de diagrammes afin de répondre aux besoins des utilisateurs et pour décrire les différentes vues et comportements du système. Il est possible de les regrouper selon les aspects fonctionnels et architecturaux [Muller, 1997], [Roques, 2003]. Parmi elles le diagramme de classe est d'un grand intérêt pour ces travaux de recherche. Il permet de faire une représentation d'un domaine en utilisant les concepts d'objet et de classe. Un objet est une entité qui a des propriétés et réalise des méthodes. Une classe est un regroupement d'objets de même nature sous le même nom générique à un niveau conceptuel. Le concept de classe permet d'implémenter les relations d'héritage pour classifier les différents types d'objets dans un domaine (exemple différents types d'outils d'usinage dans notre cas).

Pourquoi UML ?

Plusieurs sont les avantages qui justifient le choix du langage UML: La simplicité de l'utilisation, la modularité des différents types de modèles et surtout l'adaptabilité et la possibilité de l'utilisation dans plusieurs types de domaine. L'approche orientée objet permet une représentation unifiée des différents éléments d'un système, ce qui correspond à notre besoin pour la modélisation des différentes entités du système. Dans nos modèles, nous utilisons les diagrammes de classes. Ce type de représentation nous permet de définir la structure interne de notre système ainsi que les relations entre les différents composants de point de vue statique.

Le choix du langage UML avec son approche orientée objet, et surtout le choix des diagrammes de classes, pour la représentation des différents objets de notre système avec la liste enrichissante d'attributs qui les caractérisent, est très adéquat pour les développements du framework global attendu par ce projet, vu la possibilité d'exploitation de ce type de modèle dans des applications informatiques [Belkadi, 2006].

Les généralisations ou les propriétés d'héritage dans les diagrammes de classes offrent deux avantages, le premier est le regroupement de plusieurs éléments selon leur nature principale. Le deuxième avantage c'est que de point de vue modèle de connaissances les héritages permettent

de développer le maximum de taxonomies d'un concept bien précis et on parle ainsi d'ontologies.

Pour plus de sémantique, on parle de UML/OCL (Object Constraint Language). L'utilisation d'OCL consiste à l'attribution des contraintes et des règles sur un diagramme afin de compléter les spécifications d'une modélisation. Particulièrement, les contraintes OCL sont utilisées dans le diagramme de classe UML [OMG, 2003].

Ontologie

Une ontologie est une technique de représentation et d'organisation des connaissances d'un domaine d'application précis en utilisant des mécanismes de raisonnement et d'inférence. [Octaviani, 2015 ; Munir, 2017]. Dans la littérature plusieurs définitions des ontologies ont été proposées, on peut citer celles proposées par [Gruber, 1993] :

« L'ontologie est une spécification explicite d'une conceptualisation » et « une ontologie apporte les moyens pour décrire explicitement la conceptualisation sous-jacente aux connaissances représentées dans une base de connaissances » proposée par [Gómez-Pérez et Benjamins, 1999]. Suivant [Psyché, 2007], « les ontologies ont pour rôle de fournir un système de concepts fondamentaux du domaine afin de construire une base solide sur laquelle sont construites des bases de connaissances partageables et utilisables ». L'ensemble de ces concepts fondamentaux est décrit dans [Gruber, 1993 ; Gomez-Perez, 1999].

1. Les concepts ou les classes : ce sont des notions permettant la description d'une tâche, d'une fonction, d'une action ou d'un processus de raisonnement. Ils peuvent être abstraits ou concrets, élémentaires ou composés, réels ou fictifs. Habituellement, les concepts sont organisés en taxonomie. Une taxonomie est une hiérarchie de concepts reliés entre eux en fonction de critères sémantiques particuliers.
2. Les relations : représentent un type d'interaction entre deux ou plusieurs concepts. Des exemples de relation sont « subclass-of » ou encore « connected-to ».
3. Des fonctions : sont un cas particulier de relation à n éléments où le dernier élément de la relation est unique pour les n-1 premiers éléments.
4. Les axiomes : permettent de définir la sémantique des termes (classes, relations), leurs propriétés et toutes contraintes quant à leur interprétation. Ils sont définis à l'aide de formules bien formalisées de la logique du premier ordre en utilisant les prédicats de l'ontologie.

5. Les instances : utilisées pour représenter les éléments.

Elles sont utilisées dans plusieurs domaines pour assurer la communication, la structuration et la réutilisation des connaissances [Gruninger et Lee 2002 ; Sureephong et al. 2008], et pour assurer l'interopérabilité [Young et al. 2007] entre les différentes parties qui composent un système. Elles sont ainsi considérées comme des « repositories de connaissances » nécessaire pour le développement des systèmes à base de connaissance [Sanya al., 2015].

Une première analyse de la littérature montre que la modélisation des connaissances métier consiste principalement à modéliser les objets de l'entreprise qui participent, produisent, interagissent, transmettent la connaissance. [Labrousse, 2004] définit l'objet de l'entreprise comme « une entité constitutive de l'entreprise et/ou manipulée par elle et qui joue un rôle dans son fonctionnement ». Pour modéliser l'ensemble des connaissances industrielles et en se basant sur cette notion d'objet, plusieurs approches ont été développées dans la littérature. Chaque approche regroupe l'ensemble d'objets et de concepts de base jugés nécessaires pour modéliser le système globale et autour desquelles s'intègrent les différentes connaissances disponibles. Plusieurs ont été développés dans la littérature, nous nous intéressons ici aux approches développées dans le domaine de la production.

Annexe 4 : Analyse des normes et standards

La première catégorie la plus intéressante pour nos travaux c'est « 25.080 - Machines-outils ». Nous retrouvons ici plusieurs normes qui traitent l'environnement des machines à outils en général qui font partie du « 25.080.01 » :

- La norme « ISO 229:1973 » : s'intéresse aux vitesses et aux avances (valeurs nominale, tolérances (sur l'avance par minutes et par tour))
- Les normes « ISO/TR 17243-X:2014 » : s'intéressent aux broches pour machines-outils
- Les normes « ISO 230-X » : présentent des codes d'essai des machines-outils (des effets thermiques, de l'émission sonore, des exactitudes géométriques, des positionnements, des vibrations, etc.)
- Les normes « ISO 16090-X:2017 » : traitent la sécurité des machines-outils
- Les normes « ISO 14955-X:2018 Machines-outils » : s'intéressent à l'évaluation environnementale
- Finalement nous retrouvons plusieurs autres définitions dans la catégorie « 25.080.01 » qui concernent: Les symbolisations des indications figurant sur les machines, l'évaluation environnementale, la sécurité, les conditions de fonctionnement des machines, etc.

D'autres sous-catégories de la « 25.080 », regroupent un ensemble de conditions d'essai des machines, pour assurer le contrôle de la précision et le contrôle de l'exactitude de plusieurs machines à caractéristiques très différentes à savoir, qui sont :

- 25.080.10 - Tours
- 25.080.20 - Machines à aléser et à fraiser
- 25.080.30 - Machines à brocher
- 25.080.40 - Machines à percer
- 25.080.50 - Machines à rectifier et à polir
- 25.080.99 - Autres machines-outils

** La deuxième sous catégories que nous jugeons très intéressante pour nos travaux c'est « 25.100 - Outils de coupe »

Nous retrouvons ici plusieurs normes qui traitent l'environnement des outils de coupe en général qui font partie du « 25.100.01 » :

- Les normes : « ISO 3002-X » : s'intéressent aux grandeurs de base pour la coupe et la rectification (système de référence, angles de l'outil et angles en travail, brise-copeaux, grandeurs géométriques et cinématiques en usinage, forces, énergie et puissance, etc.)
- Les normes : « ISO 5610-X :2014 » : s'intéressent aux porte-plaquette à queue rectangulaire pour plaquettes amovibles
- Les normes : « ISO 5611-X :2015 » : s'intéressent aux cartouches du type A pour plaquettes amovibles
- Les normes : « ISO 13399-X :2006 : Représentation et échange des données relatives aux outils coupants (relatives aux éléments coupants, aux outils, aux attachements, aux éléments d'assemblage, aux systèmes de connexion, Informations graphiques, des modèles 3D, etc.)
- Finalement nous retrouvons plusieurs autres définitions dans la catégorie « 25.100.01 » qui concernent : outils tournants à queue cylindrique, classification et application des matériaux durs de coupe, Beaucoup de définitions sur tout ce qui concerne les plaquettes amovibles

D'autres sous-catégories de la « 25.100 », regroupent un ensemble de normes qui traitent les différentes caractéristiques des outils utilisés dans les opérations d'usinage de base : tournage, fraisage, taraudage, etc.

Ces sous-catégories sont principalement :

- 25.100.10 - Outils de tour
- 25.100.20 - Outils à fraiser
- 25.100.25 - Outils pour machines à raboter et machines à brocher
- 25.100.30 - Forets, outils à chanfreiner et alésoirs
- 25.100.40 – Scies
- 25.100.50 - Tarauds et filières
- 25.100.60 – Limes
- 25.100.70 – Abrasifs

La troisième sous catégories traitée c'est : « 25.040 - Systèmes d'automatisation industrielle »

Nous retrouvons ici plusieurs normes qui traitent l'environnement des systèmes d'automatisation industrielle en général qui font partie du « 25.040.01 » :

- La norme « ISO/TR 11065:1992 » présente un glossaire des termes d'automatisation industrielle

Un ensemble de normes traite le sujet de l'interopérabilité :

- Les normes « ISO 11354-X » : traitent les exigences relatives à l'établissement d'un processus d'interopérabilité pour les entreprises de fabrication

- Les normes « ISO 16100-X » : étudient le profil d'aptitude du logiciel de fabrication pour interopérabilité

- Les normes « ISO 16300-X » : s'intéressent à l'interopérabilité des unités d'aptitude pour la fabrication de solutions d'application

Un ensemble de normes traite le sujet de la relation environnementale :

- Les normes « ISO 20140-X » : assurent l'évaluation de l'efficacité énergétique et autres facteurs de fabrication des systèmes qui influencent l'environnement

Un ensemble de normes traite les indicateurs de la performance clé (KPI) :

- Les normes « ISO 22400-X » : s'intéressent aux indicateurs de la performance clé pour le management des opérations de fabrication (concepts et terminologie, Définitions et descriptions, Description de l'acquisition des données relatives aux séquences opérationnelles, Management de l'énergie)

Un ensemble de normes présente des dictionnaires techniques :

- La norme « ISO 22745-1:2010 » présente des dictionnaires techniques ouverts et leurs applications aux données permanentes (principes fondamentaux, Vocabulaire, Représentation du dictionnaire, Interface de recherche de dictionnaire, Modes opératoires de maintenance d'un dictionnaire technique ouvert, Recherche de données, Représentation des données)

Un ensemble de normes traite les installations distribuées:

- Les normes « ISO 23570-X : Installation distribuée dans les applications industrielles (Capteurs et organes de commande, Bus de communication hybride, Bus de distribution d'énergie)

Finalement nous retrouvons plusieurs autres définitions dans la catégorie « 25.100.01 » qui concernent: Sécurité des machines, Les concepts et règles pour modèles d'entreprise, les

architectures des communications en temps réel, Classification des objets et codes pour les classes, etc.

Nous retrouvons aussi plusieurs normes qui s'intéressent aux centres d'usinage qui font partie de la «25.040.10 »

- Les normes « ISO 10791-X » : présentent des conditions d'essai pour les centres d'usinage (Essais géométriques, Précision et répétabilité de positionnement des palettes porte-pièces, Précision et répétabilité de positionnement des axes linéaires et rotatifs, Précision des vitesses et interpolations, Exactitude des pièces d'essai usinées, Évaluation des performances, Évaluation des temps opératoires de changement d'outils et de changement de palettes, Évaluation des déformations thermiques, etc.)

D'autres normes s'intéressent à l'environnement des machines à commande numérique qui font partie de la «25.040.20 »

- La norme : « ISO 3592:2000 » fournit des informations de sortie des processeurs CN -
- Structure de fichier et format de langage

- La norme : « ISO 4342:1985 » traite les données d'entrée des processeurs CN --
Langage de référence de base pour programme de pièce

- La norme : « ISO 4343:2000 » traite les informations de sortie des processeurs CN --
Instructions post-processeur

- La norme : «ISO 13041-1:2004 » présente des conditions d'essai des tours à commande numérique et des centres de tournage (Essais géométriques, Précision et répétabilité, exactitudes des vitesses, des pièces usinées, évaluation des performances, etc.)

- Les normes « ISO 14649-X définissent un modèle de données pour les contrôleurs numériques informatisés (Données de la machine-outil pour les procédés de coupe, Outils pour le tournage, Outils pour fraiseuses, Données des procédés généraux, etc.)

Une dernière catégorie de « 25.040 » qui nous intéresse particulièrement regroupe des normes qui traitent l'environnement des mesures et du contrôle des processus industriels qui font partie de la «25.040.40 »

- Les normes « ISO 10303-X » : traitent la représentation et l'échange de données de produits (Principes fondamentaux, Méthodes de description « Express », Méthodes de mise en application, Liaison du langage C++, C et Java TM à l'interface d'accès aux données normalisées, Liaison EXPRESS à XMI, Représentations XML de schémas et de données

EXPRESS, Evaluation de la conformité, Ressources génériques intégrées, Ressources d'application intégrées, Conception de l'emballage, Aménagement des navires, Données fonctionnelles et représentation schématique pour les installations industrielles, Configuration spatiale d'usine, Systèmes d'ingénierie, Cycle de vie du produit, Gestion des modèles 3D d'ingénierie, Identification du produit, Tolérances dimensionnelles et géométriques, Conditions de surface, Données fonctionnelles, Gestion des données de produit, Topologie du réseau, etc.)

- Les normes « ISO 15531-X » : s'intéressent aux données de gestion de la fabrication industrielle (Modèle d'information des ressources, Modèle du temps, Modèle de suivi des flux et échange des données de la fabrication, Modélisation de l'information pour l'acquisition des données d'atelier, etc.)

- Les normes « ISO 15926-X » : traitent l'intégration de données de cycle de vie pour les industries de "process" (Données de référence pour la géométrie et la topologie, Méthodologie de modèle, mise en œuvre pour l'intégration de systèmes distribués, Mise en œuvre du langage d'ontologie du Web (OWL), Ontologie d'intégration de cycle de vie représentée dans le langage d'ontologie du Web (OWL), etc.)

- Les normes « 18828-X » : présentent les procédures normalisées pour l'ingénierie des systèmes de production (Flux d'informations dans les processus de planification de la production, Indicateurs de performance clé dans les processus de planification de la production)

- Les normes « ISO 3511-X » : s'intéressent aux fonctions et à l'instrumentation pour la mesure et la régulation des processus industriels (Principes de base, les diagrammes d'interconnexion d'instruments, les fonctions calculatrices, etc.)

- Les normes « ISO 8000-X » : traitent la qualité des données (Vocabulaire, Gestion de la qualité des données, Évaluation de la maturité organisationnelle des processus, Gestion de la qualité des données, Échange des identificateurs qualité, Provenance, Exactitude, Complétude, Cadre de management de la qualité, etc.)

- ...

La dernière sous catégories c'est : « 25.060 - Systèmes de machines-outils ». Nous retrouvons ici, principalement, plusieurs normes qui traitent les éléments standards et autres dispositifs des systèmes de machines-outils, qui font partie du « 25.060.10 » :

- Les normes « ISO 8526-1 » et « ISO 8526-2 » présentent des éléments standards pour machines-outils, et principalement les palettes.

- D'autres normes traitent les éléments standards pour la construction des machines-outils (tables de déplacement, têtes multibroches, Plateaux tournants porte-pièces, Montage au sol, consoles, unités de broche de perçage, unités de broche de fraisage, unités de broche de tournage, etc.)

- ...

Titre : titre Développement d'un outil à base de connaissances pour l'aide à la décision dans le contexte de l'Industrie 4.0 : Application au diagnostic des machines d'usinage à grande vitesse

Mots clés : aide à la décision, modélisation des connaissances, ontologie, industrie 4.0, diagnostic.

Résumé : Dans le cadre de sa quatrième révolution, le monde industriel subit une forte digitalisation dans tous les secteurs d'activité. Les travaux de recherche de cette thèse s'intègrent dans un contexte de transition vers l'industrie du futur, et plus spécifiquement dans les industries d'usinage mécanique. Ces travaux de recherche répondent ainsi à la problématique d'intégration données et connaissances industrielles, comme support aux systèmes d'aide à la décision. L'approche proposée est appliquée au diagnostic de défaillance des machines d'usinage connectées. Cette thèse propose, dans un premier temps, un cadre conceptuel pour la structuration de bases de données et de connaissances hétérogènes, nécessaires pour la mise en place du SAD.

Grace à une première fonction de traçabilité, le système capitalise la description des caractéristiques de tous les événements particuliers et les phénomènes malveillants pouvant apparaître au moment de l'usinage. La fonction de diagnostic permet de comprendre les causes de ces défaillances et de proposer des solutions d'amélioration, à travers la réutilisation des connaissances stockées dans l'ontologie du domaine et un raisonnement à base de règles métiers. Le système à base de connaissances proposé est implémenté dans un Framework global d'aide à la décision, développé dans le cadre du projet ANR collaboratif appelé Smart Emma. Une application pratique a été faite sur deux bases de données réelles provenant de deux industriels différents.

Title: A Knowledge-based tool development for decision support in the context of Industry 4.0: Application to the diagnosis of high-speed machining machines

Keywords: knowledge management, ontology, DSS, industry 4.0, diagnostic.

Abstract: Within the context of its fourth revolution, the industrial world is undergoing a strong digitalization in all sectors. This research work is integrated in a context of transition towards the industry of the future, specifically in the mechanical machining industries. These studies answer the problematic of industrial data and knowledge integration, to sustain the functioning of decision-support systems. The proposed approach is used to diagnose the failure of connected machining machines. This thesis proposes, in a first step, a conceptual framework for the structuring of heterogeneous knowledge and data bases, necessary to implement the DSS.

Through a first traceability function, the system capitalizes the description of the characteristics of all particular events and malicious phenomena that may appear during machining. The diagnostic function allows to understand the causes of these failures and to propose improvement solutions, through the reuse of knowledge stored in the ontology and a rule-based reasoning. The proposed knowledge-based system is implemented in a global Decision Support Framework, developed as part of the ANR collaborative project called Smart Emma. A practical application has been made on two real databases from two different industrials.