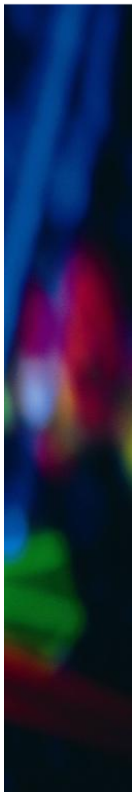
ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

ARTS ET MÉTIERS ParisTech



2016-ENAM-0026

École doctorale n° 432 : Science des Métiers de l'ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École Nationale Supérieure d'Arts et Métiers

Spécialité " Génie Industriel "

présentée et soutenue publiquement par

Baptiste HERVÉ

le 30 juin 2016

Méthodologie de conception d'une plateforme de services intelligents

Directeur de thèse : **Améziane AOUSSAT**
Co-encadrement de la thèse : **Jean-François OMHOVER**

Jury

Mme. Julie STAL-LE CARDINAL, Professeur, LGI, Ecole Centrale Paris	Président
M. Daniel BRISSAUD, Professeur, G-SCOP, Université de Grenoble	Rapporteur
M. Bertrand ROSE, Maître de conférences HDR, UDR, Université de Strasbourg	Rapporteur
M. Améziane AOUSSAT, Professeur, LCPI, Arts & Métiers ParisTech	Examinateur
M. Jean-François OMHOVER, Maître de conférences HDR, LCPI, Arts & Métiers ParisTech	Examinateur
M. Marc TRELA, Docteur, e.l.m. leblanc (groupe Bosch)	Examinateur
M. François VUILLAUME, e.l.m. leblanc (groupe Bosch)	Invité

T H È S E

Arts et Métiers ParisTech - Campus de Xxxx
Nom de l'Unité de recherche

Remerciement

Je souhaite tout d'abords remercier les membres du Jury pour leur temps passé pour évaluer mes travaux de recherche et pour leurs retours constructifs.

D'autre part, cette thèse ne serait rien sans l'encadrement que j'ai reçu. C'est pourquoi je souhaite remercier chaleureusement Améziane Aoussat et Jean-François Omhover pour leur implication tout au long de cette recherche, qui ont sus discuter et structurer l'ensemble de mes idées depuis les débuts et jusqu'à ma soutenance.

Je souhaite également remercier l'ensemble de mes collègues et encadrants industrielles de l'entreprise e.l.m. leblanc pour la confiance qu'ils ont pu me faire pour un sujet aussi innovant et stratégique dans le développement de l'entreprise. C'est l'engagement, le dynamisme et la motivation de l'ensemble des équipes de l'entreprise qui ont rendu ces travaux pertinents et consistants.

Ensuite, je souhaite remercier les cinq master de recherche que j'ai pu encadrer pour leurs projets de fin d'étude et qui ont participé activement à l'élaboration de cette thèse. Je souhaite également remercier mes collègues doctorant en entreprise, Paul et Carine, pour la diversité des échanges que nous avons pu avoir sur des sujets transverses.

Merci à tous les membres du LCPI pour leur accueil dans le laboratoire, pour le partage et pour la convivialité des moments passés dans cette équipe.

Enfin, merci à mes amis et ma famille pour leur soutien dans cette période particulière qu'est la thèse de doctorat.

Sommaire

Introduction générale **9**
Objectif et problématique de recherche **9**
Structure du document **11**

1. Positionnement et enjeux de la recherche **13**
1.1. Enjeux de la recherche **13**
1.1. Positionnement scientifique **14**
1.1.1. Le Laboratoire de Conception de Produits et d'Innovation (LCPI) 14
1.1.2. Rattachement aux axes de recherche du laboratoire 14
1.2. Contexte industriel **16**
1.2.1. e.l.m. leblanc 16
1.2.2. Projet Thermibox 17
1.2.3. Problématique industrielle 19
1.2.4. Projet Hybride 2.0 20
1.3. Conclusion du positionnement **22**

2. État de l'art **25**
2.1. Les objets connectés **27**
2.1.1. Introduction à l'internet des objets 27
2.1.2. Les technologies de l'internet des objets 28
2.1.3. Les architectures « plateforme » 31
2.1.4. Synthèse sur l'internet des objets 32
2.2. Les sciences de la conception **34**
2.2.1. Méthodologies de conception 35
2.2.2. L'ingénierie concourante et le *Design for X* 37
2.2.3. La conception des logiciels 37
2.2.4. Synthèse sur la science de la conception 40
2.3. Le service **41**
2.3.1. Définition du service 41
2.3.2. Méthodologies et bonnes pratiques de la conception de services 44
2.3.3. Les outils et techniques de la conception de service 47
2.3.4. Synthèse sur le service 51
2.4. *Machine Learning* **52**
2.4.1. Positionnement du *Machine Learning* 52
2.4.2. Méthodologies et bonnes pratiques du *Machine Learning* 53
2.4.3. Les outils et techniques du *Machine Learning* 58
2.4.4. Synthèse sur la découverte de connaissances 62
2.5. Synthèse bibliographique et constat **64**

3. Problématique de recherche **67**
3.1. Enoncé de la problématique **67**
3.2. Formulation des hypothèses **69**
H1 : L'approche service, moteur de la conception 70

H2 : Un développement des fonctions intelligentes intégré 71

4. Expérimentations 73

4.1. Expérimentation 1 : Conception amont par le service 75

4.1.1. Objectifs 75

4.1.2. Protocole 76

4.1.3. Application 78

4.1.4. Analyse des résultats et conclusion 87

4.2. Expérimentation 2 : Conception intégrée des fonctions intelligentes 93

4.2.1. Objectifs 94

4.2.2. Démarche 94

4.2.3. Cadre d'application et contexte 95

4.2.4. Application 101

4.2.5. Analyse des résultats et conclusion 116

4.3. Expérimentation 3 : Conception générique de fonctions intelligentes par le service 121

4.3.1. Objectifs 121

4.3.2. Démarche 122

4.3.3. Cadre d'application et contexte 123

4.3.4. Application 127

4.3.5. Analyse des résultats et conclusion 139

4.4. Conclusion des expérimentations 143

5. Apports et discussion 147

5.1. Modèle de conception par le service 148

5.2. Dynamique de conception détaillée des fonctions intelligentes 151

5.3. Les apports industriels 156

Conclusion de la thèse 160

Références 164

Annexe A : Projet Hybride 2.0 172

Annexe B : Projets de Master Recherche 175

Annexe C : Projet TOP79 177

Annexe D : Projet « *Hackfest* Microsoft » 179

Annexe E : Exemple de personas générés en phase de clarification de l'expérimentation 1 181

Annexe F : Exemple de fiches concept issues de la phase de génération de concept de l'expérimentation 1 182

Annexe G : *Service blueprint* pour les fonctionnalités du client final 183

Annexe H : Service blueprint pour le service après-vente 185

Annexe I : Exploration des données réalisées dans l'étape 2 de l'expérimentation 1 187

Annexe J : Résultats de l'atelier créatif effectué lors de l'étape 3 de l'expérimentation 2 189

Introduction générale

Ce mémoire de thèse de doctorat est le fruit de trois années de recherche dans le cadre d'un format CIFRE, à cheval entre le laboratoire LCPI (Laboratoire de Conception de Produit et Innovation) des Arts et Métiers ParisTech et l'entreprise e.l.m. leblanc, filiale française du groupe Bosch pour la fabrication, la vente et la maintenance d'appareils de chauffage mural à gaz.

Nous y abordons le thème d'un transfert technologique entre les sciences informatiques et l'industrie, où les technologies de l'information et de la communication se révèlent comme un vecteur porteur de l'innovation en entreprise. Effectivement, le paysage industriel se montre favorable à l'introduction de ces technologies à travers l'internet des objets. L'idée portée par la mise sur le marché de produits communicants, tel un réseau, laisse imaginer des perspectives fonctionnelles nouvelles pour ces produits.

Cependant, la transition n'est pas triviale. D'un point de vue technologique, nous pouvons observer une certaine maturité sans pour autant faire émerger de standards. Autrement dit, la technologie est prête pour créer de telles solutions, mais sa structure et son dimensionnement nécessitent de lourds investissements. Du point de vue de l'entreprise, l'introduction de telles solutions marque également un tournant dans son activité et la proposition de nouvelles fonctionnalités l'amène à reconsidérer son organisation.

Nos travaux de recherche portent ici sur l'étude de cette transition sous la facette méthodologique pendant la conception de ces nouveaux produits. C'est-à-dire que nous souhaitons proposer une démarche spécifique à la conception de produits connectés et suffisamment générique pour qu'elle puisse s'appliquer dans de nombreux domaines de l'industrie. Cette démarche doit alors favoriser l'innovation pour les entreprises souhaitant digitaliser leur offre de produits.

Une fois cette démarche explicitement caractérisée, nous observerons sa mise en application dans un cadre industriel. Les résultats ainsi obtenus seront analysés afin d'en tirer des enseignements.

Objectif et problématique de recherche

Nombre de médias font dorénavant référence au « *big data* », au « *cloud computing* » ou encore à « l'internet des objets » comme des éléments de rupture dans le domaine de l'informatique, sonnant le glas de l'internet tel que nous le connaissons aujourd'hui. Ces termes font effectivement référence à des améliorations technologiques simplifiant la mise en place d'infrastructures vouées à la gestion de flux d'information au travers d'internet tout en augmentant drastiquement les capacités de traitement des systèmes impliqués. Ce phénomène est mis en lumière par un accroissement toujours plus important du nombre de terminaux connectés à internet. L'année 2008 marque un tournant dans cette dynamique avec un nombre d'objets connectés à internet supérieur à la population mondiale. Ce phénomène est en croissance perpétuelle et on estime que 100 milliards d'objets seront connectés d'ici à 2020.

A l'instar des géants de l'internet tels que Google, Amazon, Facebook et Apple, l'ensemble des domaines de l'industrie souhaitent profiter de cette dynamique afin de se faire leur place sur le marché

et innover par le biais de ces technologies. Cependant, ces perspectives technologiques, fortement portées par l'informatique, sont en marge des paradigmes communs aux industries lourdes.

C'est au regard de ces constatations que s'initie notre problématique industrielle. Nous pouvons l'exprimer par la nécessité des entreprises à comprendre les leviers d'innovation qui leur sont offerts au travers des technologies de l'information et de la communication. Celles-ci, en plein essor, montrent de nombreuses synergies possibles pour les entreprises dont l'activité principale est organisée autour de la vente de produits. Cependant, les difficultés de ces entreprises apparaissent très rapidement lorsqu'elles tendent à embrasser une dynamique purement technologique pour laquelle elles ne possèdent pas les compétences nécessaires. L'objectif de nos recherches consiste donc à guider et structurer les démarches à mettre en place afin de favoriser l'innovation dans la mise sur le marché d'une offre de produits connectés.

Structure du document

Ce mémoire s'articule autour de cinq grandes parties comme le montre la figure suivante

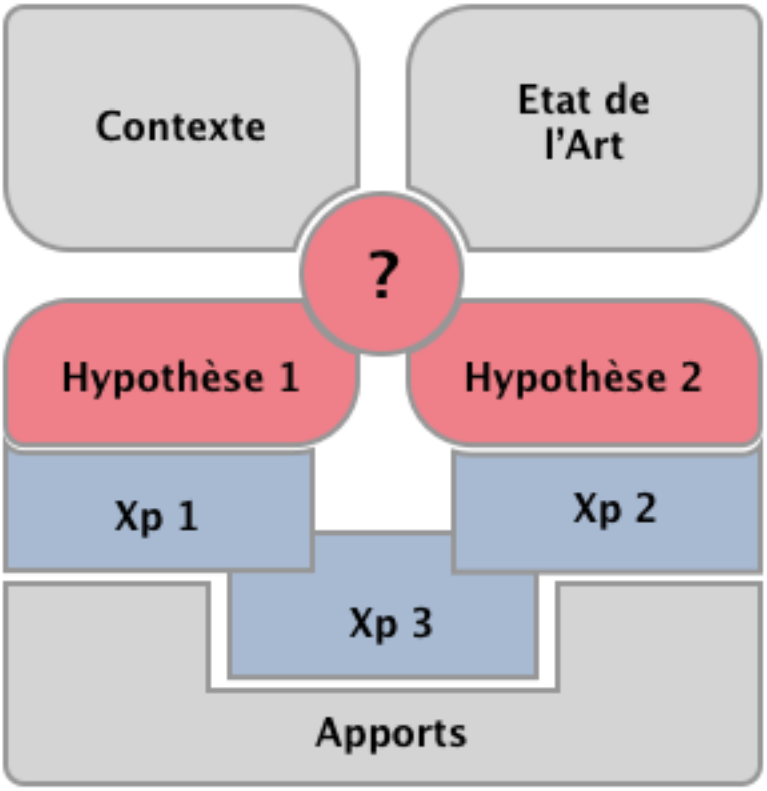


Figure 1 : Plan du mémoire

Nous commencerons par positionner nos travaux dans l'environnement. Cela comprend une facette académique avec les axes de recherche du laboratoire qui nous accueille, ainsi qu'une facette industrielle où nous présenterons l'entreprise qui supporte nos travaux. Ensuite, nous présenterons le cadre de nos travaux au travers d'un état de l'art. Nous nous reposerons sur celui-ci afin d'exprimer notre problématique de recherche pour laquelle nous proposerons deux hypothèses de résolution. Ces hypothèses seront alors testées dans le cadre de trois expérimentations que nous présenterons. Avant de conclure, nous prendrons un peu de hauteur sur les travaux réalisés afin de dégager les enseignements de nos recherches.

1. Positionnement et enjeux de la recherche

Cette thèse effectuée sous le format CIFRE s'inscrit dans un environnement double, avec d'une part une approche scientifique auprès du Laboratoire de Conception de Produit et d'Innovation (LCPI) et d'autre part une approche industrielle au sein du groupe Bosch, dans l'entreprise e.l.m. leblanc. Afin de cerner les intérêts pluriels de cette recherche, nous présenterons ici ces deux parties qui ont accueilli nos travaux, centrées sur des enjeux communs.

1.1. Enjeux de la recherche

En effet, la dualité des parties encadrantes de cette recherche reflète la dualité des enjeux que nous portons. Dans ce mémoire nous focalisons nos efforts sur l'étude des objets connectés et sur leur mise en œuvre dans l'industrie. Ce sujet d'étude est d'un intérêt certain car il est dépeint dans les médias, aussi bien spécialisés que grand public, comme étant fortement porteur pour l'industrie de demain. Cet élan technologique pourrait alors marquer un tournant dans l'industrie telle que nous la connaissons aujourd'hui. Cependant, la maturité des technologies, des sciences et des entreprises dans leurs savoirs faire ne semble pas suffisante pour mener à bien, de manière triviale, de tels projets. L'objet de cette étude est alors porté par ce besoin de clarifier les nouveaux champs des possibles issus de ces nouvelles technologies afin de favoriser l'innovation dans les entreprises.

C'est pourquoi, d'un point de vue scientifique, nous nous intéressons plus particulièrement à l'étude des méthodologies liées à la conception de tels systèmes. Effectivement, le sujet de notre attention présentant encore des zones d'ombre se doit d'être clarifié et défini afin de faire émerger les caractéristiques et les propriétés qui lui sont propres. C'est alors sur cette base que nous pouvons analyser les mécanismes à mettre en œuvre pour pouvoir s'approprier les technologies afin de servir un but préalablement convenu.

D'autre part, et au regard de ces guides méthodologiques apportés par la science de la conception, nous envisagerons le problème avec l'œil d'un industriel afin de comprendre comment ces connaissances peuvent être transférées dans son activité. Les technologies qui composent l'internet des objets sont issues des domaines proches de l'informatique, de la communication et de l'information. Des domaines sans conteste loin des compétences développées dans les industries lourdes, mais néanmoins nécessaires pour concevoir ces systèmes en phase avec le domaine d'activité de l'entreprise. C'est pourquoi afin de favoriser l'innovation autour de ces technologies, il nous semble primordial de considérer ce second axe afin d'identifier et de localiser les différents apports des domaines de compétence impliqués au cours de la conception.

1.1. Positionnement scientifique

Suivant cette perspective, nous sommes encadrés dans nos recherches par le Laboratoire de Conception de Produit et d'Innovation. Nos recherches s'inscrivent donc dans un environnement propice à l'analyse des principes de l'innovation ainsi que dans une continuité de travaux relatifs à l'intégration de nouvelles technologies dans la conception. Dans cette partie nous présenterons notre positionnement.

1.1.1. Le Laboratoire de Conception de Produits et d'Innovation (LCPI)

Depuis sa création en 1973, le LCPI des Arts et Métiers ParisTech se positionne en génie industriel dans l'analyse des méthodes industrielles pour conduire l'innovation. Les recherches qu'il supporte sont alors reliées à « l'optimisation du processus de conception et d'innovation » en gardant une vision transverse à l'ensemble de ses composantes. Cela se traduit par trois axes principaux tel que le montre la figure 2 ci-dessous. Un axe métiers vise à extraire les spécificités, telles que des connaissances, des pratiques ou des outils pouvant contribuer à l'évolution du processus de conception. Le deuxième axe concerne les processus et vise à formaliser et maîtriser leurs phases de manière transverse. Enfin, le dernier axe concerne les technologies, considérées comme un support et une voie d'enrichissement de la conception.

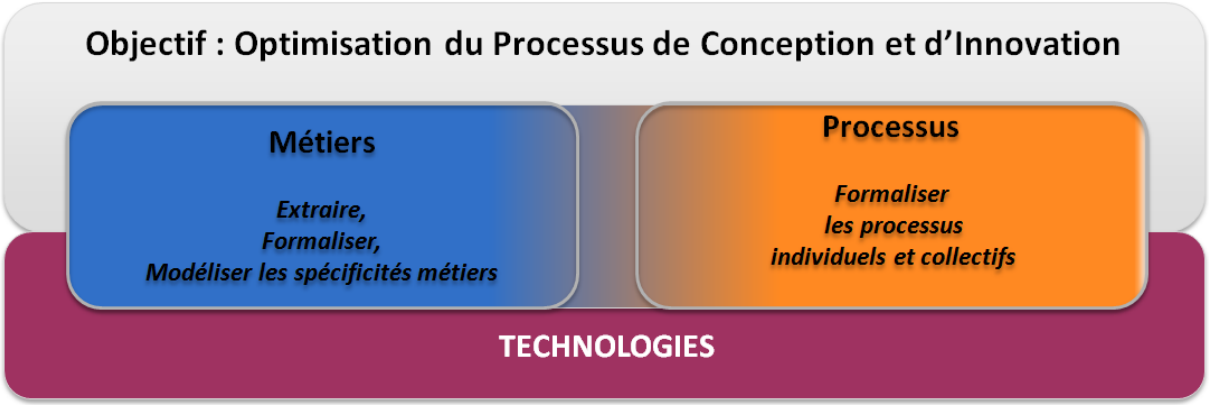


Figure 2 : Positionnement de la recherche au Labratoire de Conception de Produits et Innovation

Nous nous positionnons ici au croisement de ces trois axes dans la maîtrise d'une technologie venant ouvrir des potentiels d'innovation pour l'industrie. Nous nous intéresserons à la fois à l'impact de cette technologie sur la conduite de processus de conception ainsi qu'à l'implication nécessaire des métiers dans sa mise en œuvre.

1.1.2. Rattachement aux axes de recherche du laboratoire

Notre étude prend place dans une continuité de recherches portées sur la réalisation de systèmes dans le domaine du numérique et valorisant la mise en place de solutions intelligentes. Cette dynamique de travaux a été initiée par (Lim 2003) avec ses recherches portées sur le design et l'ergonomie pour intégrer l'utilisateur dans la conception d'une interface de téléviseur interactif. Suite à ses travaux se sont succédé un ensemble de recherches sur la mise au point d'outils numériques pour supporter les activités des concepteurs. Nous pouvons alors citer les systèmes tels que TRENDS (Mougenot 2008), GENIUS (Kim 2011) ou SKIPPI (Bongard 2013). Ces trois systèmes ont la particularité de disposer d'une composante que nous pouvons qualifier d'intelligente, où la structuration et le traitement des données suivent les préceptes de l'intelligence artificielle afin de supporter le concepteur dans l'une des tâches qu'il doit accomplir.



Figure 3 : GENIUS (à gauche) et TRENDS (à droite)

Au regard des autres projets réalisés au sein du LCPI, cette thèse trouve son originalité dans la cible à laquelle s'adressent les solutions développées. Autrement dit, les précédents projets utilisent les technologies de l'information principalement comme support du processus de conception. Leurs utilisateurs sont donc restreints à une poignée d'initiés dans le cadre d'une activité de conception spécifique. Ces projets ont le mérite de capter et de supporter des mécaniques cognitives complexes engagées dans cette activité, cependant, ils ne s'adressent pas au marché de masse.

Notre recherche vise à construire des systèmes s'adressant de manière directe aux utilisateurs finaux des produits de l'entreprise, dans toute leur diversité. L'objectif de cette recherche n'est donc pas uniquement orienté vers la maîtrise de cette technologie en elle-même pour assister l'utilisateur dans une activité spécifique. En effet, nous portons également notre intérêt vers la proposition de nouvelles fonctionnalités de nos produits sur cette base technologique. Nous nous orientons alors également dans la recherche du besoin des utilisateurs et des usages attendus du système.

1.2. Contexte industriel

Nos travaux sont également soutenus dans le monde industriel, encadrés au sein de l'entreprise e.l.m. leblanc. Cette dernière dispose d'un environnement favorable au déroulement d'un projet innovant suite à la thèse de (Trela 2014) qui a fait naître une réelle culture de l'innovation au sein de cette entreprise.

1.2.1. e.l.m. leblanc

Créée en 1932 par Marcel Leblanc, l'entreprise e.l.m. leblanc s'est spécialisée dans la production d'appareils pour le chauffage individuel et la production d'eau chaude sanitaire au gaz. Installée dès son origine à Drancy, près de Paris, elle est passée par étapes du stade artisanal au stade industriel. En 1996, elle rejoint la division Thermotechnique du groupe Bosch, leader européen en systèmes de chauffage et de production d'eau chaude.



Figure 4 : Secteur d'activité de Bosch Thermotechnique

L'entreprise e.l.m. leblanc, avec 700 collaborateurs, représente une entreprise modeste au sein d'un groupe qui en pèse plus de 200 000. Néanmoins, elle s'inscrit pleinement dans les démarches et la culture d'innovation portée par le groupe. En effet, avec plus de 3 800 demandes de dépôt de brevet en 2009, le groupe Bosch est un leader technologique mondial dans ses secteurs d'activité.

Cette thèse prendra place dans le cadre d'un projet d'innovation, la « Thermibox », détaillé dans la partie suivante. Ce projet se déroulera sur le site de Drancy, dans le service en charge de réaliser les études de pré-développement pour l'ensemble des nouveaux produits de la marque. Ce service dispose à la fois d'une expertise sur les méthodes et les processus liés à l'innovation, mais également d'une forte expertise technique sur ses produits et sur les contraintes industrielles qui leurs sont liées.

Afin d'assurer la pertinence des développements, ce service d'innovation entretient un contact particulier avec l'ensemble des autres services de l'entreprise. Notamment, le marketing et la qualité se comportent tels des clients d'une étude en spécifiant les attentes d'un produit vis-à-vis des contraintes de leurs domaines. D'autre part, le service après-vente, la logistique et les ventes interviennent tout au long

de l'étude afin de s'assurer de la bonne intégration des nouveaux produits dans le processus de fabrication, de distribution et de maintien existant.

D'autre part, ce service dispose également de contacts particuliers avec des entités extérieures à l'entreprise. Elle entretient notamment des contacts avec d'autres entreprises du groupe Bosch afin d'échanger sur des domaines de compétences particuliers ou de faire part de retours d'expérience vis-à-vis de certaines technologies. Nous pouvons par exemple y retrouver ici des entités maîtresses des technologies liées à l'électronique de contrôle des appareils ou bien des entités de recherche centrales autour des technologies et techniques de valorisation de la donnée. Le service de pré-développement entretient également des contacts avec des entreprises extérieures au groupe afin de co-développer des applications d'une technologie ou bien de partager des compétences dans le cadre d'une application telle une relation client-fournisseur classique.

Ceci constitue l'environnement industriel dans lequel notre étude prendra place. Afin de définir le contexte direct des applications auxquelles nous nous intéressons, nous abordons maintenant l'historique que possède l'entreprise dans les produits de l'internet des objets.

1.2.2. Projet Thermibox

La Thermibox est un projet précurseur en termes de connectivité. Effectivement, développée en 2004 et mise sur le marché en 2006, la Thermibox est un système de communication en charge de faire le lien entre une chaudière à gaz et le service après-vente de l'utilisateur dans le but d'améliorer l'efficacité de ce dernier. Ce système est composé d'un système d'information, permettant le traitement des informations, et d'un boîtier communiquant chargé de la collecte des données à l'interface de l'automate de la chaudière. Vendu sous forme de contrat de maintenance amélioré, la Thermibox offre la possibilité de suivre le comportement de la chaudière en temps réel afin d'en améliorer la maintenance. Les fonctionnalités du système reposent sur deux éléments : la remontée des informations en cas de panne et la prévention des pannes.

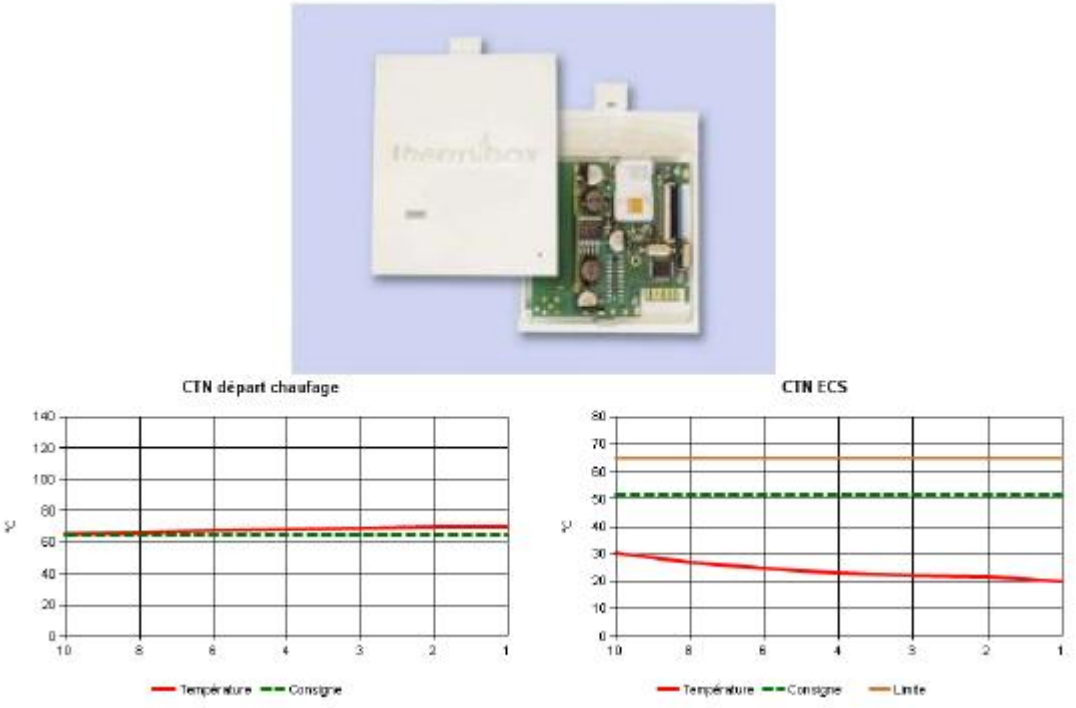


Figure 5 : Système Thermibox

Grace à ce système, lorsqu'une chaudière se met en défaut et cesse de fonctionner, l'opérateur du service après-vente est immédiatement informé. Il dispose alors d'informations complémentaires caractérisant la panne, dont le code erreur de la chaudière et une boîte noire décrivant les dernières secondes de fonctionnement de l'appareil. L'opérateur peut alors juger de la gravité de la panne et prendre les décisions qui s'imposent. Cela pouvant aller du simple contact avec le client pour s'assurer que son système fonctionne correctement au déclenchement d'une intervention d'un technicien afin de remettre en état l'appareil.

D'autre part, le système collecte périodiquement les données de comportement du produit en fonctionnement. Les données de fonctionnement d'un appareil sont alors mises en relation avec l'ensemble des données du parc afin de les comparer. Le système apprend alors ce qu'est un « fonctionnement normal » pour un produit et met en place un système de limites de fonctionnement afin d'identifier des dérives et valeurs anormales. Lorsque de tels modes de fonctionnement sont détectés, le système notifie l'opérateur du service après-vente d'un potentiel problème sur l'appareil.

Un retour d'expérience de la part des membres de l'entreprise qui ont vu grandir cette solution nous a permis de mettre en lumière un certain nombre de points positifs du système ainsi que des lacunes. Dans les points positifs, nous pouvons retenir que :

- Cette première version constitue une validation technologique pour ces nouvelles fonctionnalités. Précurseur dans le domaine des objets connectés, ce système permet de créer de nouveaux services basés sur des données comportementales des produits installés.

- La réaction du client face à ces types de fonctionnalités est majoritairement positive. En effet, lorsque le service après-vente prend les devants en appelant le client pour faire face à une panne, cela semble le rassurer et il montre une certaine satisfaction envers le service proposé.

- Le nombre d'interventions sur les chaudières équipées de ce système a significativement baissé. Autrement dit, ce type de système permet d'améliorer la rentabilité du service après-vente.

Cependant, le service proposé n'a pas rencontré le succès escompté. Nous pouvons alors pointer du doigt certaines lacunes, notamment dans la manière dont ce service est proposé au client. Nous pouvons alors dresser le bilan suivant :

- L'offre de service s'adresse à la mauvaise cible. Nous entendons par là que ce service, facturé au client dans le contrat de maintenance, ne lui propose pas de fonctionnalités clairement perceptibles. De plus, l'introduction de la notion de risque de défaillance est mal perçue relativement à l'image de la marque qui vise à avoir un haut niveau de qualité.

- Le système est déjà vieillissant et peu évolutif. Nous avons déjà évoqué le fait que les technologies numériques évoluent très vite, contrairement aux chaudières qui disposent d'une durée de vie de quinze ans. Cette faiblesse est d'autant plus importante que la plateforme, de par sa conception, ne permet pas d'évolution pour intégrer les derniers produits de la marque.

- Le système peut s'avérer être difficilement utilisable par l'opérateur du service après-vente. Que ce soit par les connaissances à engager pour établir un diagnostic, ou par certaines notions d'usage, l'interface de la solution est difficilement utilisable par des personnes qui ne sont pas expertes.

Au regard de cette première expérience dans les objets connectés, et dans le contexte des transformations qui s'opèrent dans le paysage industriel actuel, e.l.m. leblanc a jugé stratégique la reprise

de ce projet afin de l'améliorer et de renforcer sa position dans ce domaine. Ce nouveau développement représente le cœur de notre projet de thèse. Ce dernier se structure au sein d'un projet collaboratif, l'Hybride 2.0, que nous présenterons par la suite.

1.2.3. Problématique industrielle

En premier lieu, et avant de présenter le contexte du projet dans lequel nous inscrivons nos travaux, nous pouvons clairement expliciter la nécessité de nos recherche au travers de notre problématique industrielle. Effectivement, fort de la première expérience en termes d'internet des objets de l'entreprise, le produit « Thermibox », que nous avons présenté précédemment, montre le potentiel que peuvent apporter ces technologies pour l'entreprise, et cela même si son introduction sur le marché n'a pas reçu le succès escompté. Afin de renforcer cette approche et de pérenniser la création de nouvelles activités pour l'entreprise par ce biais, nous devons alors répondre à la question générale suivante :

Comment ouvrir le domaine d'activité de l'entreprise de manière durable par la fourniture de nouveaux services en passant par la réalisation d'un système d'objets connectés ?

Cette question très générale se structure alors autour de différentes composantes qui décomposent le problème et qui constitueront nos objectifs de recherche.

Tout d'abord, nous pouvons nous interroger sur **le but de ce système**. Effectivement, au regard de l'activité de l'entreprise, qui tourne autour de la production et de la vente de chaudières à gaz, nous cherchons à garder une certaine forme de consistance entre ces nouveaux services et l'activité de l'entreprise. La question étant ici de savoir comment nous pouvons identifier les bons services, en adéquation avec l'activité de l'entreprise, pour lesquels notre système pourra apporter de nouveaux éléments.

Ensuite se pose la question de **la constitution du système**, et cela à différents niveaux. Tout d'abord, et suivant les premières expressions du besoin de l'entreprise, nous souhaitons proposer des accessoires embarquant la connectivité. Ces accessoires doivent alors être rétro-compatibles avec le plus de produits du catalogue de l'entreprise possible et ils ne doivent pas complexifier l'utilisation ou l'installation des produits. Autrement dit, notre fourniture de connectivité doit passer par un module indépendant du produit chez l'utilisateur et pour lequel notre seule interface disponible est un bus de données propriétaire, nourri par le produit lui-même. Par la suite, nous pouvons également nous interroger sur la fourniture de ces données collectées au travers de nos nouveaux services et de la plus-value que peut apporter la connaissance de l'entreprise à ce niveau. Nous pouvons imaginer ici la mise en place de systèmes experts gérant des règles de traitement des données brutes reçues de la part des produits, tout comme des systèmes d'apprentissage automatisés. Quoi qu'il en soit, cette facette de notre problématique industrielle nous renvoie à la notion de choix technologiques à faire sur différents niveaux du système. Ces choix seront alors fortement conditionnés par la nature des services que nous souhaitons mettre en place pour l'entreprise.

Enfin, nous cherchons ici à **proposer une activité durable** pour l'entreprise. Nous entendons par là que la nature même du système, portée sur la gestion de flux d'informations issues d'une multitude de produits doit permettre une certaine flexibilité dans la mise en place de nouveaux services. Nous sommes ici amenés à réfléchir à la constitution d'un système pour un périmètre fonctionnel que nous déterminerons. De ce fait, et afin de pérenniser l'activité de l'entreprise sur le long terme, les choix technologiques que nous sommes amenés à faire doivent alors permettre une certaine flexibilité. Dans

cette optique cela nous offre la possibilité de garder un périmètre fonctionnel de la solution extensible à de nouveaux services qui n'avaient pas été envisagés au départ.

1.2.4. Projet Hybride 2.0

Le projet Hybride 2.0 est un projet de développement industriel d'un appareil hybride communiquant. Conscient du manque de solutions multi-énergies sur le marché des chaudières dans le résidentiel, ce projet propose de concevoir et développer un système de chauffage hybride monobloc (combinant une pompe à chaleur électrique et une chaudière gaz à condensation), et pour ce faire de lever les verrous technologiques relatifs aux questions d'encombrement restreint, d'atteinte d'une haute efficience énergétique et de minimisation des nuisances sonores pour un équipement couplant une pompe à chaleur à alimentation électrique et une chaudière gaz à condensation avec son système de pilotage intégré selon les caractéristiques respectives des énergies disponibles. Le réseau d'expertise constitué dans le cadre de l'appel à projet permettra de valider et d'enrichir le contenu technique du projet. Ce projet vise également à relier le monde de l'énergie auquel e.l.m. leblanc appartient à celui des Technologies de l'Information et de la Communication.

L'objectif du projet Hybride 2.0 est de mener à son terme une étude de la faisabilité et de développer un appareil de chauffage hybride compact de seconde génération ayant les propriétés suivantes :

- Performances énergétiques supérieures (COP 7/35 de 4) permettant en particulier d'atteindre le niveau de 5kWh/m²/an de production d'énergie à partir de sources renouvelables
- Equipé d'une régulation intelligente permettant d'anticiper les besoins énergétiques de l'habitat et d'optimiser le couplage entre les parties pompe à chaleur et chaudière gaz à condensation
- Connecté à une plateforme de services
- Niveau acoustique très inférieur à la version actuelle (de l'ordre de 9 dBA de moins – soit 8 fois moins bruyant) et haute qualité sonore, y compris en régime transitoire

Pour mener à bien le projet, le travail s'organise autour de quatre axes de recherche devant nous amener à terme à la fourniture de prototypes fonctionnels testés tel que le montre la figure 6. Ces axes sont segmentés en deux avec en premier lieu l'optimisation du produit, dans son acoustique et dans ses cycles thermodynamiques. Les deux axes suivants visent à étudier la fourniture de nouveaux services autour des produits au travers de la valorisation des données qu'il génère. Ces deux derniers axes sont au cœur de nos travaux. Nous partageons ceux-ci entre la fourniture d'un produit physique réalisant la communication, aussi appelé plateforme *Machine to Machine* (M2M), et l'analyse des moyens à mettre en œuvre pour réaliser les fonctions de service.

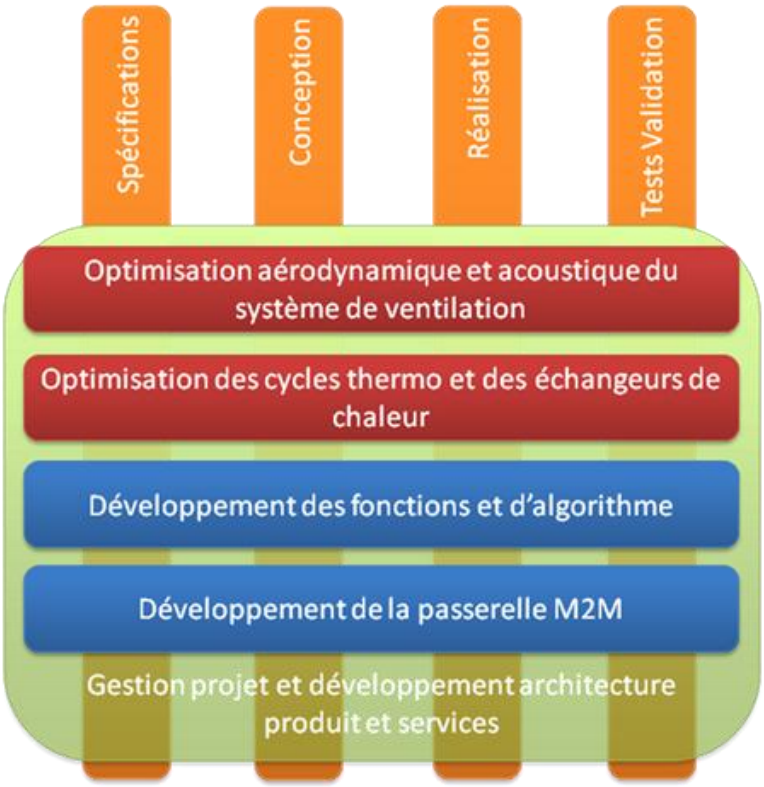


Figure 6 : Axes de recherche du projet Hybride 2.0

Pour ce faire, le projet s'organise autour de nombreux acteurs issus de monde académique comme industriel. Nous y retrouvons :

- Trois thèses de doctorat encadrées chaque fois par des laboratoires spécialisés pour les axes Acoustique, Thermodynamique et Informatique.
- Un universitaire expert pour l'encadrement de la démarche de résolution de problème (utilisation de TRIZ)
- Une PME pour la conception et la réalisation de groupes froids
- Une PME pour la conception et la réalisation de cartes électroniques
- Deux grands groupes pour la gestion du projet de développement et le suivi des démonstrateurs terrain

Plus de détails sur le projet et son organisation sont mis à disposition dans l'annexe A.

1.3. Conclusion du positionnement

Pour conclure sur ce contexte et ainsi introduire l'état de l'art, nous pouvons remarquer que la conception d'objets connectés nous amène à reconsidérer notre vision traditionnelle du produit. En effet, les technologies de l'information et de la communication permettent d'étendre le périmètre fonctionnel du produit car, comme nous l'avons évoqué, les utilisateurs d'internet et des technologies numériques sont de plus en plus nombreux et seront sensibles à une interaction avec leur système. En outre, l'expérience de la première Thermibox nous montre que, même si la technologie a fait ses preuves, la mise en place d'une offre de service cohérente et acceptable par les clients pose certains problèmes et doit être réfléchie. C'est pourquoi, nous nous attacherons ici à répondre à la question suivante :

Comment concevoir un réseau d'objets connectés permettant d'ouvrir l'activité de l'entreprise de manière durable à la fourniture de nouveaux services ?

Au regard de cette problématique industrielle, nous considérerons alors notre système comme étant à la fois :

- Un produit physique, avec l'objet en lui-même défini par des fonctionnalités et un comportement et accompagné de produits supports des fonctions de connectivité. Nous retrouvons ici les systèmes d'information et les capteurs ou systèmes de collecte de données.
- Un service, qui englobe et propose les nouvelles fonctionnalités liées à la connectivité à un client. Ce service, basé sur les informations du produit, constitue une offre sur le marché devant répondre aux besoins clients et être en adéquation avec les valeurs de la marque.
- Un système de données, chargé de produire les informations qui seront valorisées par le service. Ce système doit être en mesure de répondre au besoin client sur la base des mesures réalisées sur le produit physique.

C'est alors sous cette forme que nous considérerons le point de vue du concepteur dans l'ensemble de notre étude. Dans la suite de ce document, nous nous intéresserons donc à ces types de systèmes, complexes par leurs multiples facettes, en nous focalisant sur les besoins de structure méthodologique afin d'en réaliser au mieux la conception. Cette approche part donc de l'analyse de la technologie, en mettant ses paradigmes et ses contradictions en avant, afin d'apporter les optimisations nécessaires au processus de conception.

2. État de l'art

Suite à notre introduction et à la clarification de l'environnement dans lequel est conduite cette recherche, notamment au travers de la volonté d'une entreprise à proposer une nouvelle offre de service liée à l'introduction d'un nouveau produit connecté sur le marché, nous parcourrons l'ensemble des sujets de la littérature afin de mettre en évidence les informations pertinentes pour répondre à notre problématique industrielle. Cela nous amènera à nous poser de nombreuses questions. Comment un tel système se structure-t-il ? Qui devraient-être les utilisateurs du service ? Comment se justifierait une offre de produits connectés sur le marché et qu'apporterait-elle de plus pour ses utilisateurs ?

Ces questions sont légitimes dans le cadre du développement d'un produit nouveau, et c'est pourquoi nous commencerons notre étude par une revue de la littérature. Cette revue a pour objectif de cadrer l'environnement dans lequel prend place notre recherche et d'extraire des éléments de réponse communément admis par nos pairs. Par la suite, nous nous baserons sur cet état de l'art afin d'identifier les interrogations qui restent en suspens, ce qui dirigera, justifiera et alimentera nos travaux de recherche.

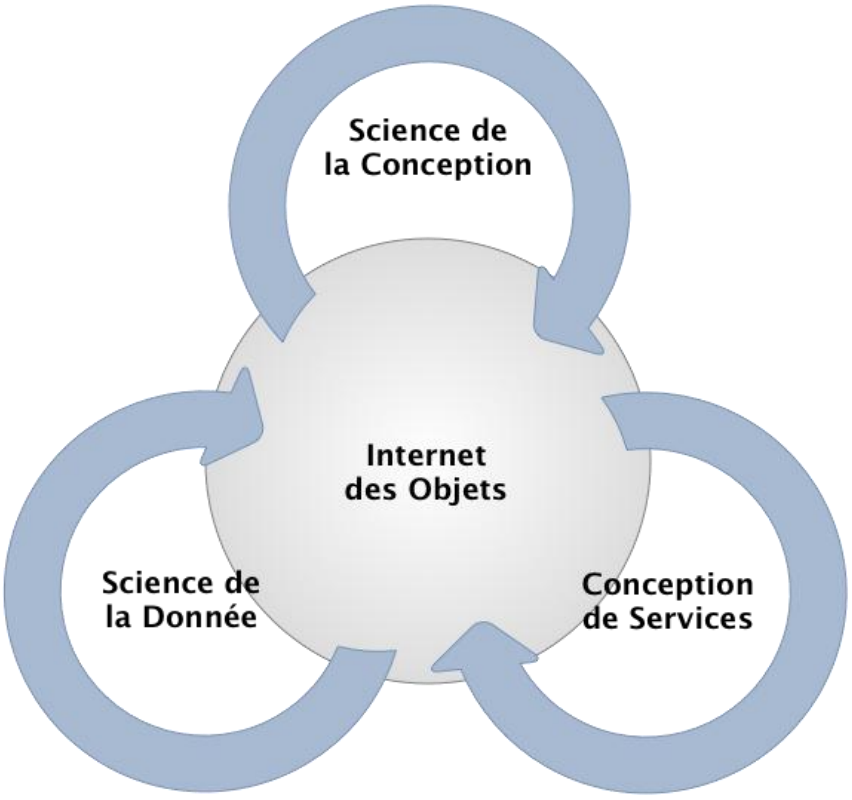


Figure 7 : Plan de l'état de l'art

Cette partie se structure autour de quatre grandes parties tel que le montre la figure 7. Nous nous intéresserons tout d'abord à la nature de notre objet d'étude en investiguant ce que l'internet des objets représente dans la littérature. Cela structurera alors notre problème et nous permettra d'orienter la suite de nos recherches au travers des trois axes que nous avons identifiés. La deuxième partie s'intéressera aux sciences de la conception, qui offrent des moyens pour faire face à des problèmes complexes de conception en passant par différents niveaux d'abstraction du système technique. Cela nous permettra alors d'extraire une trame et de nombreuses bonnes pratiques nous permettant de guider les acteurs de la conception dans leur activité. Ensuite, nous nous intéresserons plus spécifiquement aux voies de génération et de proposition de valeurs. La troisième partie de cet état de l'art traitera de la conception de service, qui s'intéresse à dégager les particularités dont dispose par nature notre système. Cela nous permettra alors d'envisager la prise en compte des besoins de nos utilisateurs. Puis, dans la quatrième et dernière partie, nous identifierons les sciences de la donnée comme prometteuses pour permettre la réalisation de fonctionnalités à forte valeur ajoutée au sein de notre système. Ces sciences mettent à disposition de nombreuses techniques et offrent un cadre méthodologique afin de manipuler et de travailler des flux d'information pour les valoriser au mieux.

2.1. Les objets connectés

> *Mardi, 16h20. Comme tous les mardis, Héloïse est pressée car son emploi du temps est serré. [...] Au moment de quitter son bureau, Héloïse reçoit sur son Smartphone un message dont la catégorie indique clairement la provenance, son appartement. Curieuse par nature, mais aussi inquiète – les messages en provenance de son appartement peuvent être provoqués par des intrusions –, Héloïse l'ouvre dans l'ascenseur. Elle se rassure vite... il s'agit d'une suggestion diététique qui émane des produits alimentaires dont elle dispose chez elle [...]. En l'occurrence, une boîte de petits pois a pris l'initiative, en concertation avec les autres vivres également présentes dans le réfrigérateur, le congélateur et les placards, de lui suggérer une idée de repas pour ce soir.*

Cette petite fiction proposée par Philipe Gautier dans l'introduction de son ouvrage (Gautier & Gonzalez 2011) projette le lecteur dans un univers où l'internet des objets est une réalité. Il fait ici transparaître l'évolution du quotidien que cela pourrait impliquer pour les futurs utilisateurs au travers d'applications concrètes, autonomes et intelligentes.

En 2016 cependant, de tels scénarii semblent encore être de la pure science-fiction. Néanmoins, c'est un monde vers lequel nous semblons tendre et c'est pourquoi nous commencerons notre état de l'art par positionner notre étude par rapport aux découvertes et aux tendances que nous avons pu dégager dans la recherche. Relativement à notre problématique industrielle, nous chercherons ici à faire émerger les caractéristiques principales d'un système d'internet des objets, principalement en termes de constitution. Pour cela, nous tâcherons de définir l'internet des objets dans un premier temps afin de mieux percevoir à quoi cette notion renvoie et de comprendre ce qu'elle implique. Puis, nous dégagerons de nos lectures une prise de parti technologique, mettant en avant une infrastructure commune à ces systèmes. Nous pourrons finalement conclure par l'extraction de propriétés inhérentes à de tels systèmes par analogies avec d'autres domaines de l'industrie.

2.1.1. Introduction à l'internet des objets

Ancré dans une société mettant l'accent sur la promotion de la connectivité dans le monde, (Zheng 2011) décrit l'internet des objets comme une tendance technologique qui vise à montrer que tout objet ou personne peut prétendre à être connecté l'un à l'autre.

La définition précise de ce secteur d'activité reste floue de par la diversité des métiers et des technologies qu'elle englobe. Nous tâcherons d'expliciter cette complexité par la suite, néanmoins, (Bassi & Horn 2008) s'intéressent aux deux termes dénommant ce secteur qui nous permettent d'ores et déjà d'exposer l'idée principale. « L'internet » renvoie à une notion de réseau commun et standardisé alors que le mot « objets » nous renvoie à l'ensemble hétérogène des artefacts de notre vie quotidienne.

D'un point de vue technologique, (Fan 2010) localise le défi dans notre capacité à créer le lien physique, logique et harmonisé entre le monde réel des objets et le monde virtuel des applications. Cela nous amène à un paradigme triple dont (Atzori 2010) révèle la raison par la nature pluridisciplinaire de cette approche. Tout d'abord un paradigme orienté vers les objets eux-mêmes, visant à embarquer la connectivité dans l'objet. Cela implique des contraintes telles que la compacité des puces, leurs ressources en énergie ou encore l'identification unique des objets dans le réseau. Ensuite, nous distinguons un paradigme orienté vers l'internet, visant à intégrer ces objets dans l'infrastructure existante

du web. Nous y retrouvons ici des problématiques protocolaires d'adressage des objets ainsi que des problématiques liées à la sécurité des flux de données. Et enfin, nous avons un paradigme orienté vers le savoir, visant à clarifier l'hétérogénéité des objets présents sur un même réseau. Nous retrouvons ici des problématiques liées à la classification de familles d'objets et à la standardisation de leurs caractéristiques et fonctionnalités afin d'en diversifier les usages.

Le phénomène de croissance rapide observé sur l'internet des objets est poussé principalement par deux facteurs environnementaux. D'un côté, les usages d'internet se démocratisent. Même si le rapport de (IW Stats 2009) montre que de fortes inégalités territoriales persistent dans l'accès aux technologies d'internet, de nombreux efforts sont déployés depuis une décennie pour les démocratiser (Gombault 2011). D'un autre côté, l'industrie et les chercheurs s'emploient à développer les technologies de connectivité embarquée pour les objets. Ce secteur est communément appelé *Machine to Machine* (M2M) (Huang 2010).

De ce fait, cela ouvre de nombreuses perspectives pour repenser les produits et services tels que nous les connaissons actuellement. (Atzori 2010) met ici en avant un ensemble de domaines d'application dans lesquels l'internet des objets pourrait entraîner une révolution. Parmi eux, nous retrouvons :

- Les logements intelligents (ou *Smart home*) pour lesquels réside un enjeu pour la gestion des ressources afin de minimiser leur consommation tout en maximisant la satisfaction de l'utilisateur.
- Les villes intelligentes (ou *Smart Cities*) où de nombreuses infrastructures sont déployées notamment pour le trafic automobile comme la voirie ou les parkings. Ces derniers sont des objets porteurs de potentielles améliorations de la qualité de vie des habitants.
- Le transport, comme l'aviation ou l'automobile pour qui l'aspect sécurité reste un enjeu majeur et perfectible par l'ajout d'intelligence embarquée dans nos systèmes.
- Le contrôle environnemental où les objets connectés peuvent être utilisés afin d'améliorer nos connaissances sur des processus naturels difficiles d'accès tel que le contrôle de l'activité sismique ou des fonds marins. Cela pouvant par la suite être mis en œuvre dans un dispositif de sécurisation des populations.
- La santé où les objets connectés peuvent améliorer le suivi d'un patient dans son environnement de vie quotidienne tout en gardant un contact avec un médecin.
- La sécurité où les objets connectés peuvent mener à des scénarii alternatifs et moins intrusifs que le déploiement massif de caméras de surveillance.
- La production, où l'apport d'informations sur l'ensemble du cycle de vie des produits ouvre de nombreuses perspectives d'amélioration de la qualité ou d'optimisation des coûts.
- L'industrie énergétique (avec le *Smart Grid*) ou les chaînes logistiques en général, pour lesquelles l'activité dépend fortement de la demande. Ces dernières voient un potentiel d'amélioration important dans le fait de maîtriser leurs flux internes et les différents maillons de la chaîne.

Cette liste n'est pas exhaustive. Néanmoins, à la lumière de ces potentielles applications, l'internet des objets et considéré par (Thierer & Castillo 2015) comme un marché colossal, principalement basé sur l'optimisation des systèmes et structures existants et disposant d'un potentiel de croissance exponentiel. Néanmoins, le principal frein au développement industriel de ce secteur réside dans l'intégration de ce lien entre monde physique et monde virtuel.

2.1.2. Les technologies de l'internet des objets

Comme nous l'avons précédemment évoqué, l'enjeu principal de l'internet des objets est de créer une passerelle entre le monde réel des objets et le monde virtuel des applications. Les solutions proposées dans la littérature, et particulièrement par (Kellmereit & Obodovski 2013), nous amènent à considérer

notre système suivant ses capacités à collecter, transmettre et analyser des données et cela avec la perspective de créer des réseaux, tels des maillages, d'ensembles hétérogènes d'objets (daCosta 2013).

Selon (Spiess 2009) et d'un point de vue pratique, l'architecture disposant des meilleures capacités pour structurer notre problème est dite orientée service : *Service Oriented Architecture* (SOA). Cette structure décompose le système en une suite d'interactions entre différents éléments et utilisateurs du système. D'un point de vue fonctionnel, cela revient à décrire une fonction par une succession d'interactions avec les transformations d'état du système qui leurs sont associées. Elle a pour avantage de faciliter son intégration dans les systèmes d'information déjà existants en s'appuyant sur de des standards d'échange et des processus propres aux métiers dans la couche applicative (De Deugd 2006 ; Pasley 2005).

Pour ce faire, l'ensemble des modèles proposés reposent sur une segmentation hiérarchique de couches entre les objets et les applications (Tan 2010 ; Wu 2010) généralement composée de cinq couches illustrées dans la figure 8 : perceptive, réseau, intermédiaire, applicative et métier. L'ensemble de ces couches doivent coexister pour faire fonctionner un système d'internet des objets. Nous passerons ci-dessous en revue l'ensemble de ces couches afin d'en extraire la vocation, le périmètre et d'identifier quelles sont les grandes tendances de la recherche dans chacune d'elles.

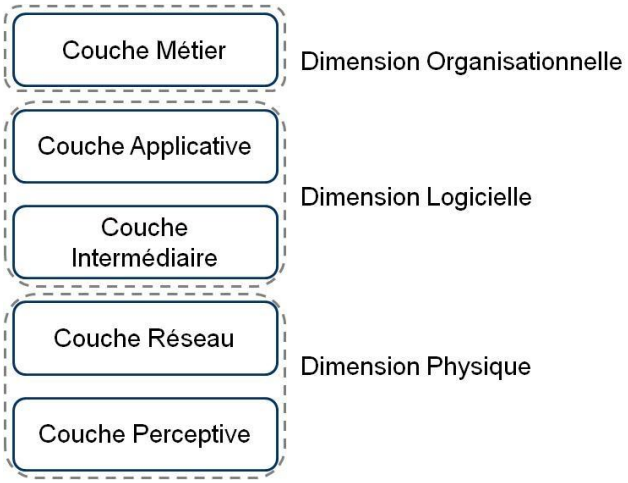


Figure 8 : Architecture générique d'un système d'internet des objets

Par cette vision, nous pouvons d'ores et déjà affirmer que l'implémentation technique d'un tel système ne constitue pas un problème en soi, et notre intérêt se porte principalement sur la couche métier et l'adaptation de ces technologies à un usage précis. Pour décrire plus en détails cette structure, nous pouvons considérer trois dimensions qui la composent : une dimension physique permettant la mise en réseau des objets, une dimension logicielle hébergeant les fonctions de service du système et une dimension organisationnelle représentant l'exploitation des objets communicants ancrée dans une activité métier.

La dimension physique

Cette dimension physique de l'internet des objets assure deux rôles principaux avec sa couche perceptive et sa couche réseau. La couche perceptive correspond à l'ensemble des capteurs et d'actionneurs ainsi que les fonctionnalités constituant l'objet en lui-même. En parallèle, la couche réseau représente les moyens matériels destinés à rendre les objets communicants. Cela inclue les éléments physiques de l'objet ainsi que toute l'infrastructure permettant l'acheminement de l'information.

À titre d'exemple, nous retrouvons dans cette partie le cas des puces RFID, des identifiants par radiofréquence. Ces puces sont une évolution du code barre permettant d'identifier un produit sans la nécessité d'actions manuelles pour scanner l'objet. Cette technologie s'intègre dans notre vision de l'internet des objets dans certains cas d'usage nécessitant le contrôle de flux ou une identification de masse à un instant donné.

À ce niveau, la recherche porte plus particulièrement sur la compacité et la portabilité des systèmes de communication. L'enjeu ici est de minimiser les coûts en énergie des communications et d'augmenter la compacité des puces. Ce travail va de pair avec les efforts à fournir pour encapsuler des données dans un protocole de communication. Le fonctionnement des objets sur un réseau tend à se caractériser par de fréquents messages à faible quantité de contenu. (daCosta 2013) pointe le fait que cela va à l'encontre des évolutions du protocole standard de l'internet tel que nous le connaissons (TCP/IP) qui nous amène à faire transiter de plus en plus gros contenus. Pour une même quantité de données fixée, le protocole nous fait transmettre de nombreuses informations contextuelles qui coûtent à la couche perceptive. Néanmoins, une communauté de chercheurs (6LoWPAN, pour *IPv6 over Low-Power Wireless Personal Area Networks)* propose des solutions pour optimiser cette mise en réseau des objets. Ces chercheurs travaillent également activement pour intégrer ce nouveau paradigme des objets connectés dans les évolutions de l'internet.

La dimension logicielle

La dimension logicielle de l'internet des objets représente les moyens d'intégration des fonctionnalités de service du système au travers de deux couches. La couche intermédiaire permet la représentation du parc d'objets au sein d'une infrastructure informatique. Elle concerne la mise à disposition pour les fonctionnalités de l'ensemble des caractéristiques des différents objets du système en gérant les contraintes induites par les réseaux de communication. Cela consiste à abstraire notre réseau d'objets sous forme de modèles virtuels dans le but de gérer l'ensemble des interactions permises par nos objets. La couche applicative quant à elle correspond à la transformation des données générées par l'ensemble des autres strates pour répondre à un besoin et à un usage. Autrement dit, elle héberge l'ensemble des transformations générant de la valeur relativement à nos fonctions de service.

Nous retrouvons ici la problématique du facteur d'échelle nécessaire pour gérer la croissance de notre réseau d'objets. Pour y faire face, l'industrie du web, avec toute sa maturité, propose la location de machines (IaaS pour *Infrastructure as a Service*) ou bien de briques logicielles (PaaS ou SaaS pour *Platform* et *Software as a Service*) en fonction de l'usage. Ce modèle d'exploitation des serveurs expliqué entre autre par (Wu 2010) est appelé vulgairement « le *Cloud* » et permet au concepteur du système de dépenser uniquement la puissance informatique qu'il consomme dans son application. Cela est donc une voie de réponse pour amortir le facteur d'échelle qu'il doit maîtriser avec la croissance de son système.

La seconde problématique de cette couche vient du fait de devoir traiter un grand nombre de données simultanément. Problématique commune à l'exploitation de données dans l'industrie en général, nous pouvons l'identifier par le terme *Big Data*. D'un point de vue général, cette notion implique selon (Jacobs 2009) que les technologies usuelles de stockage, d'accès et de traitement des données atteignent leurs limites face à la quantité d'informations et ne permettent pas leur valorisation au sein d'un système. Dans ce sens également la recherche avance et propose la mise en parallèle des traitements sur plusieurs serveurs. Elle peut être mise en lumière par (Dittrich & Quaiané-Ruiz 2014) avec la technologie *Hadoop*.

La dimension organisationnelle

La dimension organisationnelle de l'internet des objets, représentée par la couche métier, correspond à la strate décisionnelle du développement du réseau et à l'exploitation des données générées. C'est depuis cette couche que ces technologies s'inscrivent dans une stratégie d'entreprise afin de permettre d'alimenter les écosystèmes dans lesquels les objets sont installés.

Le principal axe de recherche dans cette couche réside dans la valorisation des services que peuvent accomplir ce type de réseaux d'objets. Elle se fait principalement par la construction d'un plan de rentabilité, ou *Business Model*, qui confronte le coût du système à la valeur ajoutée qu'il dégage. Les coûts peuvent être ajustés suivant des paramètres de conception de l'ensemble des couches précédentes, alors que les revenus, eux, sont propres au domaine d'application.

L'ensemble des couches préalablement décrites constitue une architecture typique, dite de « plateforme », où la génération de valeur se fait par le passage de flux dans une multitude de briques structurelles de manière séquentielle. Ce type d'architecture comporte des caractéristiques particulières que nous tâcherons d'expliciter.

2.1.3. Les architectures « plateforme »

La notion de plateforme est largement utilisée dans le langage courant et parfois de manière abusive. Afin de mieux comprendre les caractéristiques typiques de ces systèmes, nous nous devons de délimiter la notion de plateforme telle que nous l'entendons ici. Pour ce faire, nous devons déjà dissocier l'objet plateforme d'une architecture de système organisée autour de plateformes. Issue du domaine de l'électronique (Balarin 1997 ; Chang & Martin 1999), où il existe de fortes interactions entre les composants physiques des produits et leur exploitation par les logiciels, la notion de plateforme intervient dans la volonté de minimiser le temps et les efforts à fournir afin de proposer une solution répondant à un besoin particulier sur le marché. Cela est permis et se concrétise par abstraction du système suivant deux axes (Chang & Martin 1999). D'une part le besoin auquel répondent le système et les moyens qui y sont attribués, et d'autre part les moyens de communication et d'intégration du système. Cela est appelé l'orthogonalisation des intérêts par (Balarin 1997 ; Keutzer 2000). Maintenant, si nous voulons définir de manière générique la notion de plateforme, nous pouvons nous référer à (Sangiovanni 2001) qui l'explicite de la manière suivante :

- Une architecture de type plateforme est une abstraction d'un système dont la structure peut être soustraite à une multitude de sous-systèmes indépendants, les plateformes, minimisant la complexité locale du système.
- Où, une plateforme constitue un système ou un sous-système pouvant rapidement être adapté ou paramétré par son utilisateur afin répondre à de nouveaux usages et lui permettre un déploiement rapide de sa solution.

Par ces efforts de représentation, un système s'inscrivant dans ce cadre dispose de propriétés particulières affectant principalement les coûts, les temps de développement et les performances. (Ferrari 1999) insiste sur le fait que cela provient de la réutilisabilité des composants développés pour permettre un maximum de flexibilité sur les applications que nous pouvons leur trouver. La figure 9 montre qu'une instance de plateforme permet de répondre à un large spectre d'applications. Il se trouve alors que l'intégration par un concepteur d'un nouveau cadre applicatif pour un système plateforme se localise, une fois le système identifié comme étant pertinent pour nos applications, dans son paramétrage et son interfaçage avec le reste des éléments du système.

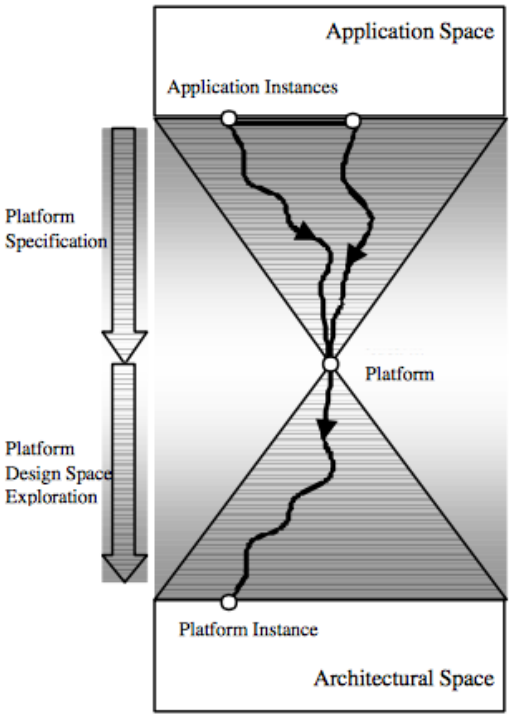


Figure 9 : Concept de plateforme extrait de (Ferrari 1999)

Les produits électroniques et les ordinateurs sont reconnus comme disposant d'un fort potentiel de modularité, notamment dû à la prédominance de l'aspect logiciel qui les compose. Néanmoins, les avantages de ce concept de plateforme sont reconnus dans l'ensemble des secteurs de l'industrie et ses principes y ont été transposés au niveau de la stratégie d'entreprise (Cusumano 2010). (Simpson 2006 ; Jiao 2007) l'expose sous l'angle du développement de familles de produits, par lequel une gestion stratégique de la conception des produits et du portefeuille produit d'une entreprise lié à une approche modulaire permet une diversification et une personnalisation des offres d'une entreprise fortement amortie par un grand facteur d'échelle sur les composants clés d'une architecture produit. C'est ainsi que nous retrouvons ces mêmes concepts transposés dans l'industrie lourde telle que l'automobile (Muffatto 1999 ; Bremmer 1999).

2.1.4. Synthèse sur l'internet des objets

Dans cette partie, nous avons pu dans un premier temps donner une définition large de l'internet des objets et présenter l'ensemble des notions sous-jacentes. Nous avons ensuite dégagé une infrastructure représentant au mieux ces systèmes. Et enfin, nous avons fait émerger les caractéristiques qu'une telle architecture nous apporte.

Nous pouvons en retenir que l'internet des objets est amené à toucher de nombreux domaines de l'industrie en proposant des optimisations des systèmes des activités des industriels. De plus, la croissance de œ domaine semble propice, avec des technologies de plus en plus matures avec le développement du M2M et avec une démocratisation massive des usages d'internet.

D'autre part, de nombreux exemples d'applications relatives à l'internet des objets convergent vers une architecture de type plateforme pour laquelle les différentes briques technologiques sont clairement identifiées. Cette infrastructure et l'ensemble des éléments qui la composent ne sont pas encore optimisés pour toutes les applications possibles, néanmoins, cela démystifie les visions utopistes de l'internet des objets et localise les problèmes que nous sommes amenés à rencontrer, sans pour autant nous bloquer.

Le déploiement d'un tel système demande de faire appel à de nombreuses connaissances et de nombreux domaines technologiques différents. Néanmoins, le système de plateforme nous offre la possibilité d'intégrer facilement des connaissances de sources externes à notre système. Dès lors, la concentration de nos efforts peut alors être focalisée sur les interfaces entre les différentes plateformes afin de créer la cohérence nécessaire au déploiement d'un tel système sur le marché.

Enfin, sur l'ensemble des strates composant l'architecture de notre système et connaissant les opportunités et les freins technologiques de notre environnement, l'une des parties les plus fragiles semble être sa valorisation. Autrement dit, les efforts permettant de lui créer une place sur le marché. Cela passe donc par une analyse fine de la valeur ajoutée du système et des moyens mis en œuvre pour la fournir aux utilisateurs.

2.2. Les sciences de la conception

Nous venons de voir précédemment la dynamique et la structure du domaine industriel dans lequel nous inscrivons nos recherches avec son apparente complexité. Nous nous attardons maintenant dans cette partie sur l'activité de conception et les moyens que la littérature nous offre afin de guider le concepteur au travers de cette activité. Pour cela, la recherche a largement contribué à formaliser et rationnaliser cette activité de conception afin de simplifier la manipulation de systèmes complexes et la gestion des connaissances (Simon 1969). Effectivement, (Dorst & Cross 2001) pointent le fait que tout au long du cycle de conception le problème et la solution co-évoluent nécessairement à cause du caractère habituellement mal défini du problème initial. De ce fait, le concepteur doit adopter une stratégie afin de se représenter et de répondre au problème auquel il fait face. Ces stratégies entraînent le concepteur dans un va et vient constant entre l'espace des solutions et l'espace des problèmes, et ce, à différents niveaux de granularité du problème initial.

Communément appelées DTM, pour *Design Theories and Methodologies*, la littérature regorge de solutions pour aider le concepteur. Afin d'y voir plus clair, nous pouvons nous reposer sur la représentation de (Lahonde 2010), basée sur (Tomiyama 1997) et (Cross 2008) distinguant quatre niveaux d'abstraction permettant la classification des DTM. Parmi ces quatre niveaux, nous retrouvons :

- **Les Théories** qui désignent une idée ou un champ de connaissances dans lequel s'inscrit la réflexion autour de la conception. Principalement basées sur l'observation et l'expérience, elles proposent des modèles de représentation éloignés des applications concrètes.
- **Les Méthodologies** qui représentent des guides pour les concepteurs. Elles proposent un cheminement d'actions permettant de procéder au mieux à la conception d'un artefact dans le cadre de la théorie préalablement définie.
- **Les méthodes et outils** qui sont les briques élémentaires directement applicables par le concepteur afin de lui permettre la réalisation des objectifs exigés par sa démarche.

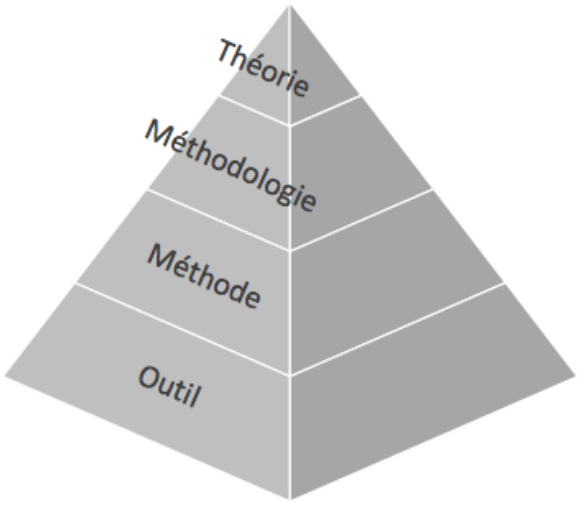


Figure 10 : Classification des théories et méthodologies de la conception

Dans le cadre de notre étude, nous souhaitons intégrer les paradigmes des objets connectés que nous venons de voir dans la partie précédente au contexte théorique afin de dégager une méthodologie

pertinente et de permettre la création de systèmes innovants. C'est pourquoi, nous exposerons dans la suite de cette partie une revue des concepts méthodologiques présents dans la littérature.

2.2.1. Méthodologies de conception

Afin de focaliser nos investigations autour des principales méthodologies de la conception intéressantes pour étayer notre propos, nous pouvons nous reposer sur les classifications faites des DTM. Une organisation autour de six axes est proposée par (Finger & Dixon 1989) dans lesquels nous retrouvons :

- Les modèles descriptifs du processus de conception,
- Les modèles prescriptifs de conception,
- Les modèles pour la conception assistée par ordinateur,
- L'environnement, la représentation et le langage dans la conception,
- L'analyse et le support de décision dans la conception et
- La conception basée sur le cycle de vie.

Ici, les modèles prescriptifs correspondent à notre besoin de support méthodologique pour le concepteur. Parmi ces derniers, l'auteur en extrait les caractéristiques représentatives de ce type de modèles :

- Le processus de conception est une progression itérative à travers six principales phases.
    - L'identification du besoin,
    - l'expression de spécifications,
    - la formulation de concepts,
    - la sélection de concepts,
    - la conception détaillée et
    - la production, vente et maintenance.
- Le concepteur fonctionne suivant trois modes.
    - La divergence où le concepteur doit étendre le périmètre de son problème et des solutions techniques potentielles
    - La transformation où le concepteur décompose son problème ainsi que ses objectifs et décide des choix d'orientation
    - La convergence où le concepteur réduit les incertitudes afin de faire émerger un unique concept

Sur cette base, (Tomiyama 2009) propose un large panel de méthodologies qu'il présente et discute. Parmi elles, deux groupes se distinguent particulièrement. Les méthodologies communément « acquises » telles que la conception systématique de (Pahl and Beitz 2007), la conception axiomatique de (Suh 1990) ou l'approche de (Ullman 1992) et celle de (Ulrich & Eppinger 2007). Ces méthodologies se distinguent des méthodologies dites « acquises et utilisées » comme le *Design for X*, la conception totale (Pugh 1991) ou l'ingénierie concourante. Cette segmentation est faite suivant les considérations suivantes :

- Les méthodologies acquises sont largement utilisées dans la littérature et dans l'éducation. Elles se focalisent plus sur la conception amont et s'approchent plus des concepts plus abstraits de la conception. Cela leur donne un caractère très générique.
- Au contraire, les méthodologies acquises et utilisées sont plus employées dans l'industrie. Cela s'explique principalement par leur vocation à se focaliser sur des objectifs plus concrets tels que l'optimisation des coûts ou de la qualité.

L'approche de conception systématique proposée par (Pahl and Beitz 2007), représentée par la figure 11, s'illustre comme étant la méthodologie générique la plus utilisée, à la fois dans l'éducation et dans l'industrie. Cette dernière est représentée par un modèle organique où l'activité de conception se positionne au centre du cycle de vie du produit. Le processus de conception ainsi introduit rapporte et positionne les principales tâches ainsi que les prises de décisions autour de quatre grandes phases.

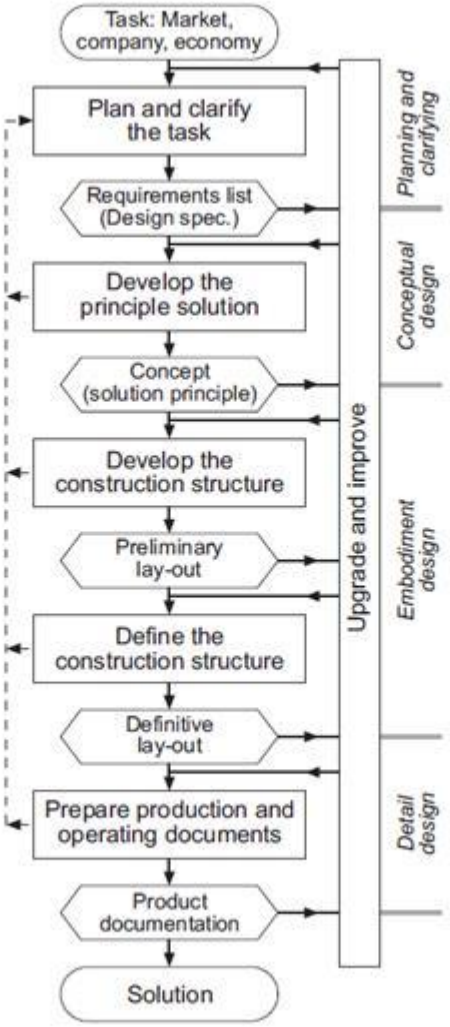


Figure 11 : Modèle de la conception systématique (Pahl & Beitz 2007)

- Phase 1 : Planification et clarification

Cette première phase permet la capitalisation de l'ensemble des informations nécessaires au démarrage d'un projet de conception. Cela comporte des informations fonctionnelles concernant le besoin auquel le produit doit répondre ainsi que des informations opérationnelles avec la clarification des tâches et du temps nécessaire pour mener à bien cette activité.

- Phase 2 : Génération de concepts

Cette seconde phase vise à déterminer le premier principe de solution apporté pour répondre au besoin précédemment défini. Cela réside principalement dans une abstraction du problème de base permettant de le structurer autour de représentations intermédiaires.

- Phase 3 : Conception générale

Dans cette phase, des propositions d'architecture générale de notre solution sont proposées et évaluées suivant la réponse apportée au besoin exprimé, de la compatibilité des différents éléments de notre système ainsi que de la viabilité financière.

- Phase 4 : Conception détaillée

Enfin, la dernière phase fait émerger l'ensemble des spécifications et des documents permettant de lancer la production de notre système, et ce en se focalisant sur l'ensemble des éléments constituant notre système afin d'en définir tous les paramètres de conception comme les formes ou les dimensions.

2.2.2. L'ingénierie concourante et le *Design for X*

Notre cadre d'étude montre, par sa nature, la nécessité d'intégrer la dominante pluridisciplinaire dans notre approche méthodologique. Effectivement, la construction d'une plateforme d'exploitation d'objets connectés fait intervenir de nombreux métiers et domaines de l'ingénierie. Au regard de cela, deux DTM se détachent des autres. L'ingénierie concourante (ou CE pour *Concourant engineering*) et le *Design for X* (ou DfX). Ces deux approches ont pour vocation l'intégration de connaissances dans le cycle de développement de produits.

D'une part, l'ingénierie concourante fait apparaître la nécessité d'une conception transversale à de nombreux métiers et pôles de compétence aussi divers que l'ergonomie, le design, l'ingénierie ou le marketing par exemple (Aoussat 1990 ; Lecoq 1992). Il est démontré qu'une telle approche favorise l'innovation, améliore la qualité et l'acceptation des artefacts produits en plus de fournir un cadre idéal à la conduction d'un projet de conception en équipe pluridisciplinaire (Aoussat 2000).

D'autre part, le *Design for X* est une approche générique regroupant des démarches qui sont pour chacune d'elles orientées vers l'amélioration d'un objectif précis, représenté par « X » (Huang 1996). Cet objectif cible peut être très variable (Tichem 1997), ciblant une étape du cycle de vie du produit tel que le *Design for Manufacturing and Assembly* pour la production, ou bien une propriété du produit ou du processus de développement tel que le *Design for Cost* ou le *Design for Flexibility*. Pour ce faire, les différentes DfX proposent au concepteur une stratégie afin de créer un environnement de développement favorable à l'atteinte des objectifs.

Les DfX et le CE sont deux approches différentes, cependant elles convergent sur les moyens mis en œuvre pour améliorer l'environnement de développement. En effet, elles tendent toutes deux vers une amélioration notable de la conception amont par l'apport d'informations métier complémentaires émanant de l'ensemble des phases du cycle de vie du produit (Bralla 1996). Cela permet aux concepteurs d'être plus pertinents dans la génération et la sélection des concepts, ce qui a pour effet direct la réduction des efforts tout au long du cycle de vie du produit.

2.2.3. La conception des logiciels

Pour terminer cette veille sur les outils de la conception et au regard de la nature multiple des systèmes d'internet des objets, nous nous intéressons ici plus particulièrement à leurs applications dans le domaine de la conception logicielle. Effectivement, d'un point de vue théorique, les artefacts logiciels disposent de quelques caractéristiques qui les distinguent des produits physiques. Ceux-ci ont un caractère modulaire extrêmement fort ainsi qu'un potentiel pour le prototypage simple et rapide. C'est dans ce cadre que s'opposent dans la pratique deux démarches : l'approche en V et l'approche agile.

Le modèle de cycle en V, hérité du modèle en cascade proposé par (Royce 1970), est un processus séquentiel de conception largement répandu dans l'industrie pour permettre de limiter le nombre de retours dans les étapes de développement. Ce formalisme met en parallèle les efforts liés à la création

d'un nouveau logiciel avec les phases de déploiement et d'intégration. De ce fait, si nous prenons l'exemple de la phase de conception détaillée, les livrables attendus de cette phase sont les spécifications d'une fonction ainsi que la description des tests qui devront être passés une fois la fonction réalisée. Cela constitue alors une approche dont le principal avantage est la gestion de la qualité. Néanmoins, sa mise en œuvre est souvent jugée trop rigide pour conduire un projet de développement et c'est pourquoi nous l'opposons aux approches dites agiles.

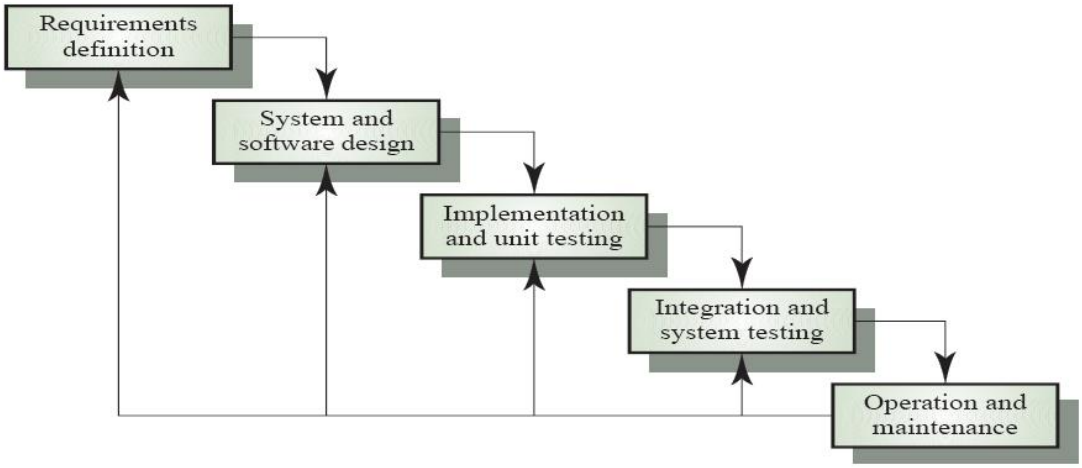


Figure 12 : Modèle de conception en cascade (Royce 1970)

Les approches agiles ont été formalisées pour la première fois par (Boehm 1988) avec son modèle en spirale. Par définition, l'agilité offre de la souplesse et de la résilience au changement. Ces modèles se veulent itératifs et incrémentaux. Autrement dit, ils reprennent les mêmes phases que le processus en V, mais les organisent autour de livraisons successives de versions. Ce modèle se propose de fournir alors plusieurs versions de prototypes pouvant être soumises au regard critique du client tout au long du développement. Cette succession de cycles de développement et d'évaluation rend alors l'environnement de conception plus résilient aux demandes de modifications et à l'évolution.

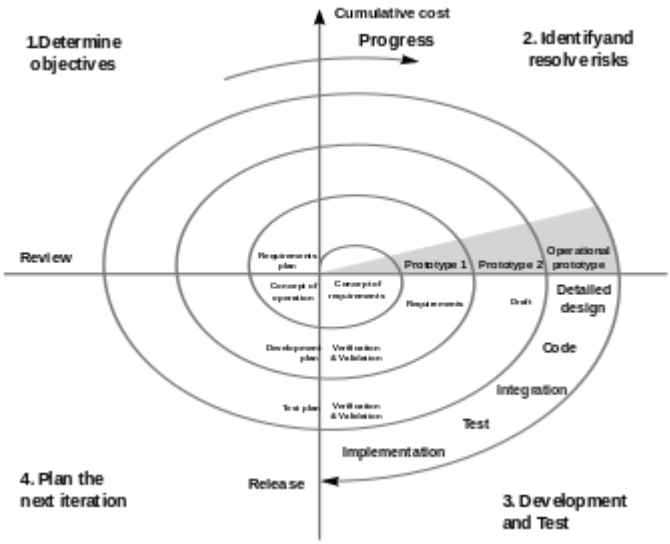


Figure 13 : Modèle de conception en spirale (Boehm 1988)

Parmi les approches agiles, nous pouvons retrouver le *Extreme Programming*, aussi dit XP (Beck 2000), ou encore SCRUM (Schwaber & Sutherland 2011). La communauté scientifique s'accorde à définir par

le biais du manifeste pour l'agilité de (Fowlerand & Highsmith 2001) douze principes vers lesquels ces approches convergent.

1. La priorité la plus importante est la satisfaction du client par la fourniture amont et continue de prototypes à haute valeur ajoutée.
2. Le changement est accepté dans le développement, et ce, même s'il survient tardivement.
3. Les livraisons de prototypes s'opèrent en cycle court, entre deux semaines et deux mois, et sous la forme d'applications concrètes.
4. Les personnes du métier ainsi que les développeurs travaillent ensemble tout au long du projet au travers des points journaliers.
5. Le projet doit se construire autour de personnes motivées pour lesquelles il est nécessaire de créer un environnement propice avec support et confiance.
6. Le moyen le plus efficace pour transmettre de l'information reste le face à face.
7. La progression du projet s'observe par la mesure d'un logiciel fonctionnel.
8. Le rythme est soutenable pour l'équipe de développement.
9. Une attention continue doit être portée à la technique et à la qualité de la conception.
10. La simplicité est un principe essentiel.
11. Les meilleurs concepts émergent d'une équipe indépendante dans son organisation.
12. Les potentiels d'amélioration continue de l'équipe projet doivent être discutés régulièrement.

Concernant SCRUM plus particulièrement, (Schwaber & Sutherland 2011) explicitent les profils de l'équipe de développement, les événements ainsi que les artefacts qui composent la méthode. Parmi les profils, nous retrouvons l'équipe de développement, disposant de l'ensemble des compétences transverses au projet afin de le réaliser. D'autre part, le *product owner* est le responsable des spécifications. Il est en charge de traduire et prioriser les besoins du client. Enfin, le *scrum master* est, tel un manager, la personne en charge de fournir au projet un environnement propice au développement.

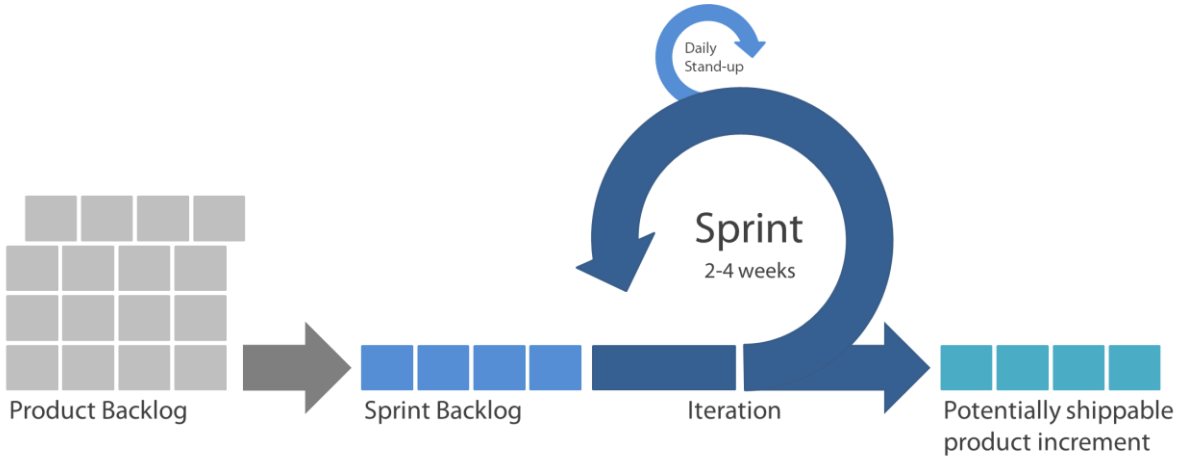


Figure 14 : Représentation de la méthodologie SCRUM

En ce qui concerne les événements, l'intégralité de la méthode s'organise autour des itérations appelées *Sprint*. L'ensemble des acteurs du projet se réunissent en amont d'un *sprint* pour un *sprint planning* afin de définir le périmètre fonctionnel à embarquer dans l'itération et qui constituera l'objectif de l'équipe. Ensuite, chaque jour un point est fait lors du *daily scrum* sur les avancées de chacun dans le but de répartir les tâches et d'identifier les problèmes potentiels. Enfin, lorsque le temps du *sprint* est écoulé, deux nouveaux événements ont lieu. Le *sprint review* qui est une démonstration de la nouvelle version livrée faite en présence de tous les membres du projet, ainsi que le client ou le futur utilisateur du

système. Et enfin, la rétrospective est un point entre l'ensemble des membres du projet afin qu'ils auto-évaluent le travail réalisé et qu'ils proposent des leviers d'amélioration.

Pour finir, les seuls artefacts de SCRUM sont les spécifications. Dès le commencement du projet, une description du livrable final doit être exprimée. Ce cahier des charges est appelé le *product backlog*. Ce dernier a pour caractéristique, d'une part d'être le seul document de référence pour regrouper l'ensemble des besoins, et d'autre part d'être évolutif sur toute la durée du projet. Les éléments de spécification qui le composent doivent alors disposer d'une granularité de détail différente en fonction de leur priorité. Un extrait de ce document est prélevé pour définir le livrable du *sprint*, appelé le *sprint backlog*, et doit être composé des éléments les plus prioritaires pour l'utilisateur avec une description suffisante à leur création. Au contraire, les éléments les moins prioritaires restent à ce stade grossièrement esquissés et seront détaillés au cours du projet.

2.2.4. Synthèse sur la science de la conception

Dans cette partie, nous nous sommes intéressés aux moyens exposés dans la littérature permettant de guider un concepteur dans son activité. Nous sommes dans un premier temps passés par une revue des méthodologies de conception. Puis nous nous sommes arrêtés sur les principes de l'ingénierie concourante et le *Design for X* avant de terminer sur les principaux schémas démocratisés dans la conception de logiciels.

Nous pouvons alors conclure que notre besoin d'un support méthodologique pour le concepteur doit se traduire par la fourniture d'un modèle de conception prescriptif. Parmi ces derniers, la littérature met en avant des modèles suffisamment génériques pour nous inspirer dans notre cas particulier. De plus, ces modèles sont suffisamment utilisés dans l'enseignement et dans l'industrie afin de favoriser le transfert et l'appropriation.

D'autre part, afin de répondre aux problématiques issues de la gestion de différentes connaissances et l'intégration des contraintes issues de différents domaines, le *Design for X* ainsi que l'ingénierie concourante semblent proposer des pistes prometteuses si nous pouvons y intégrer les domaines de compétences liés à l'internet des objets.

Enfin, et dans l'espoir de s'intégrer au mieux aux bonnes pratiques de la conception logicielle, nous retenons deux approches. La première est l'approche en V qui assure un développement de haute qualité, et la seconde est l'approche dite agile, dont nous avons exposé la souplesse. Ces deux approches ont été retenues dans le but principal d'accroître l'efficience du développement en augmentant la résilience d'un projet face aux modifications et en s'assurant que la solution proposée soit la plus pertinente possible pour ses utilisateurs.

2.3. Le service

Lors de l'établissement de notre cadre de travail sur l'internet des objets, nous avons pu constater que les motivations associées au développement de tels systèmes appartenaient à la couche métier et étaient, dans l'état actuel de la recherche, difficilement perceptibles. Nous l'avons ainsi représentée comme la couche supérieure de notre architecture système (cf. figure 8) en affirmant qu'elle était la clé de la pérennité du système d'un point de vue fonctionnel et économique. Suivant notre problématique industrielle, nous poussons dans cette partie les investigations pour clarifier les moyens de percevoir le but d'un système d'internet des objets. De plus, en raison de la nature particulière de ces systèmes, où la génération de valeur réside principalement dans un flux d'information, et non pas uniquement dans la fourniture d'un artefact, nous prenons le parti d'investiguer la notion de service.

Nous nous attacherons ici à solidifier notre approche de conception, actuellement plus orientée vers les produits, par la fourniture d'un cadre théorique en définissant et caractérisant le paradigme des services. Puis nous investiguerons les propositions et constats de la littérature sur les bonnes pratiques et les outils mis en œuvre pour concevoir des services.

2.3.1. Définition du service

Tout d'abord, la place des services dans l'économie grandit de plus en plus dans les pays développés et représente selon un rapport de l' (INSEE 2012) 75% des emplois en 2009. Cette augmentation est principalement due au transfert de certaines entreprises, depuis les années 80, d'une stratégie principalement orientée produit, vers une logique de service (Baines 2006). Dans la littérature, ce phénomène est appelé « la servuction des produits » et est rendu possible par une augmentation graduelle des offres de services des entreprises autour de leurs produits (Oliva 2003). Effectivement, (Potts 1988) met en évidence la disponibilité de nombreuses ressources dans le cycle de vie d'un produit ayant un fort potentiel d'amélioration de l'offre de service d'une entreprise. Cette transition, d'une pensée orientée produit vers une pensée orientée service, implique une reconsidération complète de la stratégie de l'entreprise et, de ce fait, implique d'importants efforts afin de changer ses processus internes ainsi que son organisation.

Afin de pouvoir comprendre le paradigme du service et de manipuler des éléments de leur conception, nous tâcherons ici d'en comprendre mieux la nature et d'en dégager les caractéristiques principales. Or, il règne encore un état de doute et d'incertitude dans la définition d'un service (Goldstrein 2002 ; Jones 2005) et notamment dans les principes d'innovation (Gadrey 1991). Nous pouvons tout de même retenir la définition énoncée par (Kotler 1977) :

> *Un service peut être considéré comme tout acte ou action qu'un parti peut offrir à un autre, tout en étant fondamentalement intangible et ne résultant en rien qui puisse être possédé. Cette production pouvant ou non être supportée par un produit physique.*

Cette définition reflète bien certains aspects des services propres à leur nature, tout en gardant un point de vue générique. Néanmoins, dans les faits, il existe autant de définitions du service qu'il existe de concepteurs. C'est donc au regard de cette affirmation que nous envisageons de construire, basée sur une revue de la littérature, une caractérisation du service suffisante pour appréhender leur conception.

Caractérisation du service

Historiquement, Le premier secteur d'activité à avoir perçu les différences entre produits et services dans l'industrie est le marketing. Le premier indicateur de ce phénomène dans la littérature réside dans l'évolution des « 4Ps » issus du concept du *marketing mix* (Produit, Prix, Promotion et Place). Utilisé pour définir une offre sur le marché, ce concept s'est vu ajouter trois P additionnels avec : Personnes, Processus et manifestation Physique (Booms 1981). Cette évolution tend à tenir compte des interactions entre les clients et les fournisseurs du service dans leur définition d'offres présentes sur le marché. De plus, le marketing a identifié des caractéristiques, plus dépendantes du service en lui-même que de sa représentation comme offre sur le marché. (Palmer 2003) nous propose les propriétés de service suivantes :

- *Intangible* : Autrement dit, le service n'est pas un artefact, palpable par le client. Un service est immatériel et difficilement perceptible.

- *Variable* : Un même service peut prendre plusieurs formes en fonction des clients. On peut dire ici qu'un service est *customisable* ou adaptatif aux besoins des clients.

- *Périssable* : Un service s'inscrit dans le temps, avec une date de début et une date de fin.

- *Impossible à procéder* : Un service n'engage la procession de rien.

- *Avec une forte relation client / fournisseur* : Contrairement à la vente de produit, qui est ponctuelle, le service vise à créer une relation constante entre client et fournisseur.

De plus, en considérant le haut niveau d'interaction entre le client et fournisseur lors de la réalisation du service, selon (Vargo 2006, 2008) les échanges entre les entreprises et les clients peuvent être utilisés comme retour d'information afin d'améliorer le service. Par ce fait, avec des échanges chargés d'une grande quantité d'informations, nous pouvons estimer que les services sont plus à même de collecter et manipuler des données. Par conséquent, nous considérerons également la valeur dans l'échange comme une caractéristique complémentaire des services.

D'autre part, si nous souhaitons extraire les attributs du service afin d'en dégager un modèle, (Tomiyama 2001) prend en considération deux acteurs principaux : Le fournisseur du service et le récepteur (client) du service. Alors, nous pouvons modéliser la valeur ajoutée du service par un flux d'informations ou de produits physiques entre le fournisseur du service et le client visant à changer l'état de satisfaction du client. Néanmoins, au regard de notre précédente définition du service ce modèle semble limité par le fait que les échanges soient modélisés par un flux simple entre le fournisseur et le client. De manière générale, nous considèrerons la transformation de valeur au sein du service comme une suite d'interactions entre le client et le fournisseur (Moreli 2002).

De plus, et en fonction du service proposé par le fournisseur, la transformation de la valeur ajoutée se fait par le biais d'un certain nombre d'autres acteurs, considérés comme des parties prenantes du service. Nous considérerons ces partenariats et l'ensemble de ces échanges entre deux types d'acteurs comme frontière délimitant l'environnement interne du service.

Ainsi l'ensemble des services est fourni dans l'environnement extérieur, ce qui comprend selon (Rae 2009) un milieu des affaires, d'autres acteurs et un contexte technologique. De plus, l'ensemble de ces éléments, dont le client, définissent le milieu extérieur du service, et disposent donc d'un point de vue partiel sur la réalisation du service. Cette limite de visibilité est principalement due au nombre limité d'interactions proposées par le fournisseur lors de la fourniture du service, mais cela est également dû à

la communication limitée que le fournisseur propose sur la transformation de la valeur ajoutée, et des processus engagés par lui-même et par les autres parties prenantes.

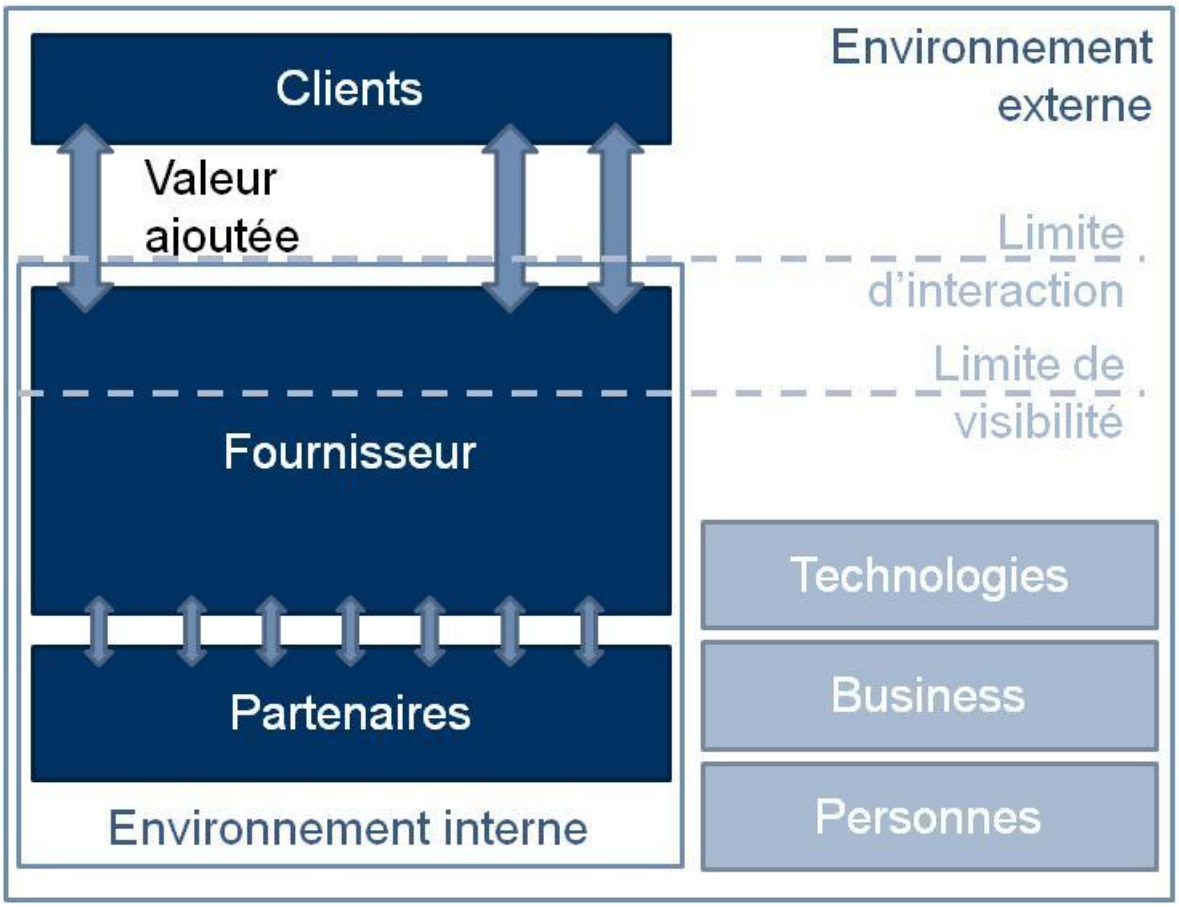


Figure 15 : Modèle descriptif du service

Pour résumer, notre modèle de service est construit d'un point de vue systémique au regard des propriétés propres du service que nous avons identifiées précédemment. Ce modèle se base sur trois principaux composants :

- Les acteurs, ou parties prenantes, du service peuvent être le fournisseur, le client ou un intervenant extérieur issu d'un partenariat.
- Les interactions, communément appelées « points de contacts » dans le domaine du *design* d'expérience, représentent la séquence du service.
- Le service pend place dans un certain environnement. Un environnement extérieur qui définit le cadre dans lequel le service est distribué, ainsi qu'un environnement intérieur dans lequel le service est produit.

Les motivations liées à la création d'une nouvelle offre de service

Au regard de cette définition, nous pouvons illustrer les motivations que peut avoir une entreprise à introduire une nouvelle offre de service. Dans un premier temps, cela implique la création d'un lien fort et constant entre l'entreprise et son client. Mis en lumière par (Demirkan 2008) ces liens offrent la possibilité d'avoir un contact plus fréquent avec son client et nécessitent un effort de la part de l'entreprise pour mieux gérer sa relation avec ce dernier. Néanmoins, alors que cela représente un effort considérable de la part de l'entreprise pour implanter un tel lien avec son client, (Gwinner 1998) et

(Kinard 2006) soulignent le fait que cela constitue une réelle opportunité pour l'entreprise d'améliorer la loyauté d'un client déjà existant pour la marque. En effet, des contacts réguliers avec son client offrent de nombreuses possibilités de collecter plus de données au sujet de sa satisfaction et de son comportement. Cela constitue un moyen efficace d'ajuster son offre de service et sa stratégie en fonction d'un retour direct de la part du client.

Ensuite, dans le cadre des domaines avec un fort niveau de compétitivité, les services peuvent être de bons vecteurs de distinction. Effectivement, l'innovation dans le milieu des services peut proposer de nouvelles valeurs ajoutées. Idéalement cela permet d'ouvrir de nouveaux horizons pour les entreprises et offre un moyen similaire à la stratégie « océan bleu » pour gérer sa stratégie d'entreprise tel que (Tan 2010) nous l'illustre. De plus, comme nous l'avons vu lors de l'établissement de notre modèle de service, la plus grosse partie du service se cache derrière ce que nous avons appelé sa limite de visibilité. (Heskett 1997) souligne ce point en ajoutant que cela rend alors les services difficiles à imiter.

Enfin, (Anderson 1997) montre que, comparés aux activités à logique orientée vers la vente de produits, les services offrent des moyens de génération de revenus beaucoup plus confortables pour l'entreprise. D'un côté, les flux de revenus générés par les services sont continus, car le client et le fournisseur du service sont liés par un contrat, et donc, les transactions financières ne se réduisent plus à des achats ponctuels de produits de consommation. D'un autre côté, les marges générées par les offres de services sont majoritairement considérées comme plus élevées que celles liées à la vente de produit.

2.3.2. Méthodologies et bonnes pratiques de la conception de services

Dans un second temps, et sur la base de la définition du service que nous venons de proposer, nous investiguons ici plus en détail les méthodologies et les bonne pratiques liées à la conception de services. Cependant, la littérature n'est pas aussi riche de propositions en termes de méthodologies prescriptives dédiées à la conception de service. La notion de service pur exprimé par (Chase 1981), peut en partie expliquer ce phénomène en catégorisant les problématiques de conception comme étant purement organisationnelles et seulement portées par la stratégie d'entreprise. Néanmoins, nous pouvons obtenir de la littérature de nombreuses bonnes pratiques liées à la conception de service ainsi que des outils principalement basés sur des représentations formelles du service.

D'une part, la littérature met en avant certaines règles inhérentes à la conception de services ainsi que des bonnes pratiques. (Stickdorn 2011) notamment, met en lumière le fait que l'activité conception de service gagne en pertinence si l'on respecte certains principes. L'activité de conception de service se doit d'être :

- *Centrée sur l'utilisateur* : C'est-à-dire que le service doit être conçu à travers les yeux de l'utilisateur.
- *Co-créative* : L'ensemble des parties prenantes du service doit être inclus dans le processus de conception du service.
- *Séquencée* : Le service doit être vu comme étant un ensemble d'actions en relation les unes avec les autres et se structurant dans le temps.
- *Matérialisée* : Le caractère intangible du service doit être perceptible par le client par le biais d'artefacts.
- *Holistique* : la totalité de l'environnement du service doit être considérée lors des phases de conception.

D'autre part, nous pouvons mettre en relation trois grands concepts méthodologiques majoritairement présents dans la littérature. Le premier, proposé par (Tomiyama 2001) est basé sur la réalisation d'un modèle de flux. Le second proposé par (Osterwalder 2004), basé sur le *business model*. Le dernier introduit par (Shostack 1977) est basé sur le service *blueprint*.

Le modèle de flux, issu de la méthodologie proposée par (Tomiyama 2001) est principalement dérivé des concepts que nous retrouvons dans la conception axiomatique (Suh 1990). Le modèle tend à décrire les acteurs du service ainsi que les flux de valeurs entre ces différents acteurs par le biais de fonctions et de paramètres de conception. Et ce afin de pouvoir soulever les leviers de conception sur l'ensemble du système de service lui permettant de maximiser la satisfaction du client. Ce modèle de service constitue une bonne solution dans l'optique de transférer au service les outils de la conception de produit. Parmi ces derniers, nous pouvons notamment utiliser plus facilement l'analyse fonctionnelle interne. Les limites de cette méthodologie apparaissent lorsque les interactions entre les acteurs deviennent trop complexes. En effet, ce modèle ne permet pas d'intégrer le service comme une suite d'interactions client-fournisseur séquencée dans le temps et avec de nombreux points de contacts.

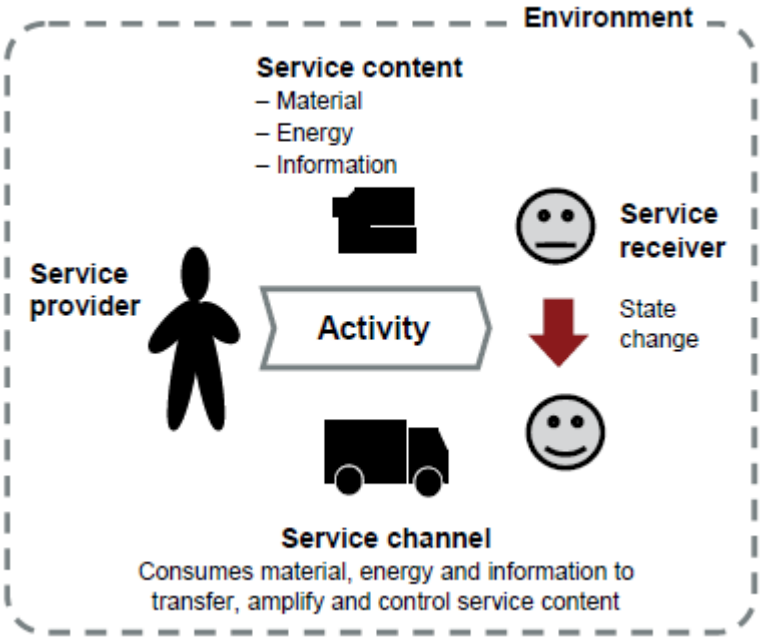


Figure 16 : Modèle de flux pour la conception de produit-service (Tomiyama 2001)

Le *business model canvas*, issu de la méthodologie proposée par (Osterwalder 2011) se base sur la définition de l'offre de service. Ce dernier positionne la valeur ajoutée en élément central avec la transformation de valeurs du côté fournisseur et la distribution du côté client. Cela rend possible la distinction des leviers envisageables pour optimiser les coûts et les revenus brassés par le service. De ce fait, cette méthodologie permet, et nous pousse, à innover dans chaque partie de la chaîne de valeur. De plus, cette manière de procéder à l'aide de ce modèle semble très efficace dans les stades amont du développement dans le cas où nous voulons porter l'innovation dans les moyens mis en place pour livrer le service au client. Cela dit, effectuer la transition entre le stade du concept de service vers l'implémentation semble difficile. Effectivement, même si les coûts semblent être de bonnes métriques pour juger du potentiel d'efficacité d'un service, cette méthodologie ne permet en aucun cas de souligner les problématiques techniques ou les leviers d'amélioration pour la réalisation ou l'implémentation d'un service ainsi décrit.

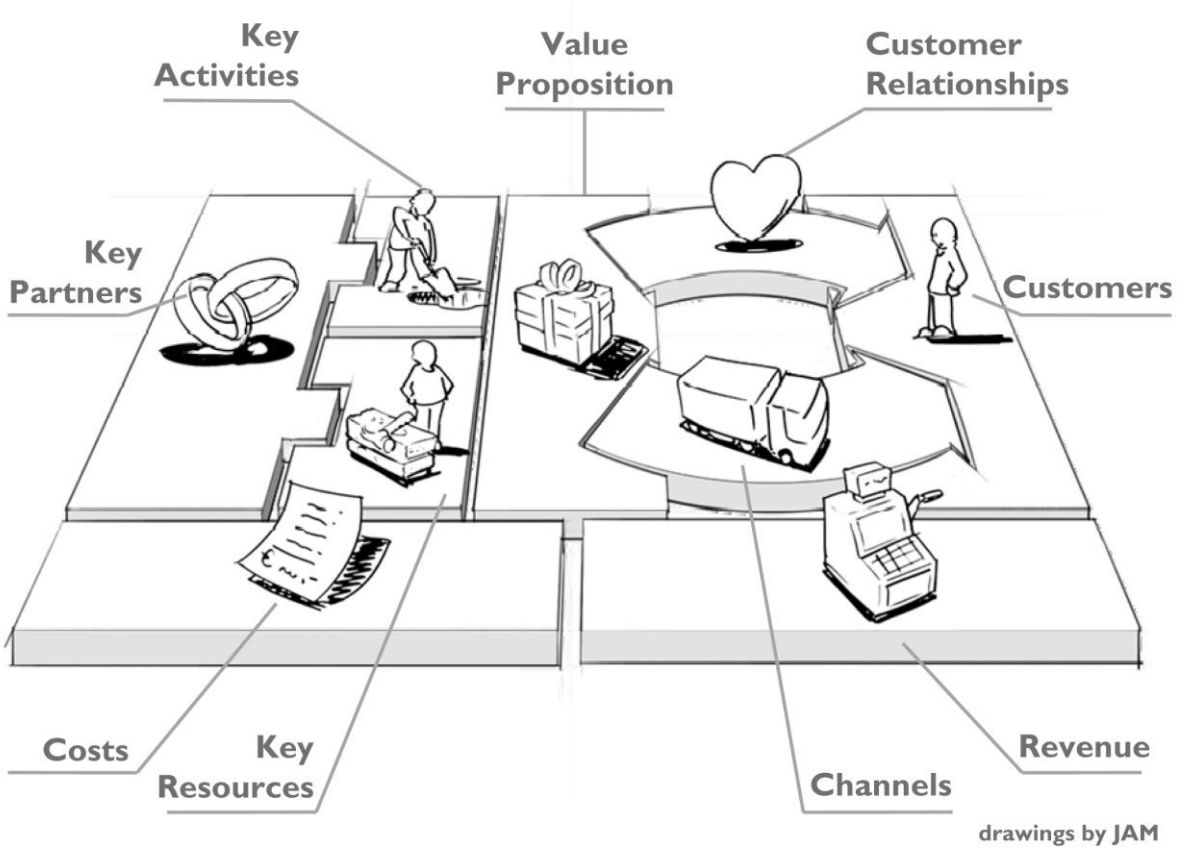


Figure 17 : Exemple de *Business model canvas*

Le *service blueprint* , issu de la méthodologie de (Shostack 1977) est basé sur la description du service sous forme de processus suivant plusieurs dimensions : les acteurs qui ont un rôle à jouer dans le service ainsi que l'ensemble des interactions qu'ils peuvent avoir les uns avec les autres, la séquence de tâches que chacun doit réaliser pour le bon déroulement du service, les matérialisations que prend le service pour le client tout au long de son déroulement, puis les processus et le matériel support à la réalisation du service. L'ensemble de ces éléments est organisé en respectant le niveau de visibilité du service avec un milieu intérieur, privé, et un milieu extérieur, public. Contrairement aux deux autres concepts que nous avons décrits ci-dessus, ce modèle semble épouser la description du service que nous avons préalablement proposée. De plus, ce modèle semble disposer d'un haut niveau de flexibilité et d'évolutivité. Autrement dit, le niveau de détail avec lequel nous pouvons renseigner le modèle semble pouvoir s'améliorer progressivement en fonction de l'évolution de la vision que nous portons sur notre système au cours du développement.

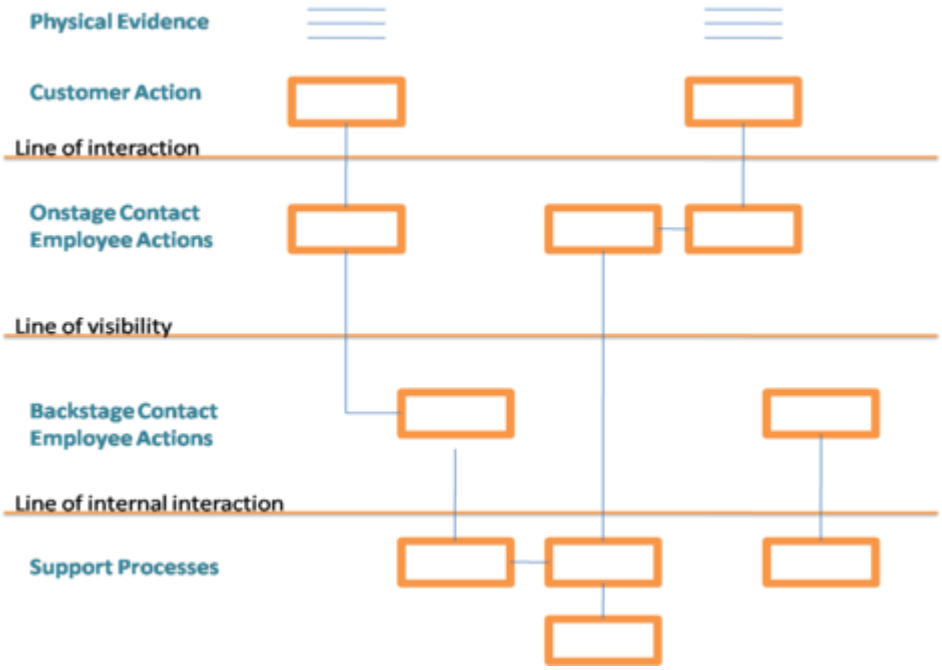


Figure 18 : Exemple de *service blueprint*

Enfin, nous pouvons trouver certains auteurs tels que (Saco 2008) et (Moritz 2005) ayant mis en place une méthodologie de conception prescriptive propre à la conception de service. Ces derniers proposent un processus de conception itératif reprenant les principales phases standard auxquelles ils ont associé un panel d'outils spécifiques à la conception de service. Il nous semble alors pertinent de considérer une méthodologie standard de la conception produit telle que la conception systématique de (Pahl & Beitz 2007) afin de l'adapter au paradigme du service et dont la clé de voûte réside dans la séquence d'outils et techniques à mettre en œuvre.

2.3.3. Les outils et techniques de la conception de service

Dans l'optique d'assister le concepteur dans sa tâche de création de service, la littérature recense de nombreux outils de la conception. Ces outils peuvent alors être utilisés tout au long du cycle de conception dans le but de faire évoluer le point de vue du concepteur sur le service qu'il veut créer. Et ce, en tenant compte des multiples aspects du service, tels que l'environnement dans lequel il sera livré, les attentes et besoins utilisateur ou les ressources nécessaires à son déploiement.

Dans ce cadre, (Diana 2009) nous propose une classification de ces outils permettant de représenter les services suivant 5 grandes familles :

- *Les textes* comme les matrices de motivation ou le *story telling*, ces méthodes permettent de représenter le service sous forme d'un rapport écrit.

- *Les graphes* : comme les grilles d'analyse de tâches ou la *mind map*, ces outils permettent de positionner des éléments du système de service les uns par rapport aux autres et permettent souvent de visualiser les connexions entre les différents éléments du système.

- *Les narrations* : comme les personas ou les posters, ces outils visent à représenter le service sous forme d'une collection d'événements qui, placés dans un certain ordre, racontent une histoire.

- *Les jeux* : comme le « *lego serious play* » ou les jeux de rôle, ils représentent le service sous forme d'activité où le concepteur est un acteur.

- *Les modèles* : comme le magicien d'Oz ou le prototype d'expérience, ils visent à représenter le service sous forme de prototypes fonctionnels.

Une revue exhaustive de ces outils et méthodes a été faite par (Tassi 2009). Parmi ces derniers, nous retrouvons le *blueprint* ainsi que le *business model* que nous avons déjà cités précédemment. Afin de répondre à notre besoin méthodologique, nous sélectionnons ici un certain nombre d'outils que nous détaillerons. Cette sélection est faite au regard de la pertinence des outils eux-mêmes, mais, nous considérerons également leurs synergies les uns avec les autres ainsi que leur potentiel à pouvoir encapsuler l'information nécessaire au concepteur pour l'assister dans son activité.

La carte des parties prenantes

Le principe de cette méthode, introduite par (Morelli 2006) est de faire l'inventaire des acteurs pouvant tenir un rôle dans le service. Il est alors possible de décrire la place de chaque acteur par rapport aux autres et pour ce faire, nous séparons notre espace des parties prenantes en deux parties. D'une part avec les acteurs internes, qui peuvent être certains départements de l'entreprise fournisseur du service ou bien d'autres entreprises du groupe. Cela permet principalement de cartographier l'ensemble des ressources et des forces d'action qui sont en présence dans l'entreprise. D'autre part, nous pouvons dissocier les acteurs extérieurs pouvant être impliqués dans la livraison de service. Il est alors possible de considérer plusieurs sous-groupes. Principalement, nous devons positionner ici le client cible du service, qui éventuellement peut être complété par l'ensemble des autres acteurs pouvant intervenir dans le cycle de vie du produit. De plus, nous pouvons positionner ici les concurrents, les principales parties prenantes du service, les fournisseurs ainsi que l'ensemble des acteurs évoluant dans un domaine analogue au service que nous souhaitons concevoir. Pour finir, cette carte nous permet de cartographier l'ensemble des connaissances qui peuvent être amenées à être utilisées en les représentant par des potentielles parties prenantes de la conception du service. Cela permet également de modéliser les relations qu'entretiennent les acteurs, et donc d'identifier des risques liés au projet. Enfin, cette carte peut être le support de suivi des actions en cours avec certains acteurs identifiés, telles que des études de marché, des enquêtes de satisfaction ou des co-développements.

Les personas

La méthode suivante, les personas, est introduite par (Cooper 1999), elle est utilisée et reconnue dans le cadre de la conception centrée sur l'utilisateur (Goodwin 2011). Cette dernière consiste à représenter les clients types auxquels s'adresse notre offre de service. Cela est fait dans le but d'être utilisé en tant que

référence dans les phases suivantes. Nous avons précédemment décrit l'activité de conception de service comme devant être centrée sur l'utilisateur et comme devant prendre en compte son avis dès la phase de conception. Or, cela peut parfois s'avérer difficile de divulguer de l'information sur le concept de service que nous voulons réaliser à des personnes autres que les membres actifs du projet afin de préserver sa confidentialité. Les personas sont un moyen de recueillir les informations comportementales des utilisateurs afin que les concepteurs puissent en tenir compte dans l'ensemble des phases du développement en jouant sur l'empathie.
Dans les faits, ce phénomène d'empathie est maximisé par une grande personnification du modèle de client créé par les personas. En effet, les personas sont des clients fictifs caractérisés par des comportements et des usages types de notre population de clients cibles. Ils sont alors identifiés par leur nom, leur âge, leur classe sociale et leur position vis-à-vis de l'environnement dans lequel le service sera amené à prendre place. De plus, afin de donner plus de profondeur aux personnages, leurs activités journalières, leurs comportements et leurs sentiments envers l'environnement du service et les technologies utilisées sont décrites.
La construction de tels outils peut être facilement réalisée avec un travail préparatoire permettant de capter l'ensemble des caractéristiques des clients cibles. Ce travail préliminaire peut prendre la forme d'études de marché ou d'interviews avec les consommateurs. Les fiches de personas sont alors créées afin de pouvoir incorporer l'opinion et les besoins des utilisateurs dans l'activité de conception sans pour autant devoir les faire intervenir dans le processus et ainsi compromettre la confidentialité du projet.

La carte système

Le principe de cette méthode, introduite par (Morelli 2007) est de diviser la chaîne de valeur du système de service que nous souhaitons développer en termes de flux d'énergie, d'information, de matériel ou d'argent entre les acteurs et les parties techniques du système. Par conséquent, cela nous propose une vue d'ensemble des systèmes techniques à mettre en place pour la réalisation du service.
Comme les principaux acteurs du service ont déjà dû être identifiés, du côté des clients aussi bien que du côté des partenaires et des fournisseurs, cet outil nous permet de représenter la chaîne de valeurs proposée par le service. En dégageant la valeur du service d'un point de vue besoin client, puis d'un point de vue fournisseur, le concepteur peut alors identifier les activités clés du service ainsi que les solutions techniques à développer dans le but de supporter cette activité. Par exemple, si une entreprise veut démarrer une activité de vente en ligne, cette méthode permettra probablement de dégager le transport et la logistique comme activité clé, avec comme support technique un site internet référençant l'ensemble de ses produits ainsi que leur disponibilité dans son stock.
Une fois ce travail réalisé, cet outil permet au concepteur de mesurer l'effort nécessaire à la réalisation du service, en termes de connaissances à capitaliser, de temps et de ressources à déployer. D'autre part, pour les prochaines étapes du service, cette représentation du système de service permet au concepteur de faire évoluer sa vision sur son concept de service et de diviser les efforts à réaliser en groupes élémentaires.

Les user stories

Les *user stories* littéralement traduit « les histoires utilisateurs », sont utilisées dans le but de formaliser des idées. Ces dernières visent à décrire une idée avec plus de détails concernant le concept sous-jacent. Effectivement, le principe de cette méthode est d'exprimer une idée comme étant une volonté de la part d'un client avec les bénéfices que ce dernier en tire.
Le principe de cette méthode est donc d'agréger une idée dans un contexte d'usage et de bénéfice. Le fait de faire intervenir un client dans la génération de l'idée fait jouer l'empathie du concepteur envers ses clients. C'est pourquoi cet outil semble particulièrement bien adapté avec l'usage des personas lors

des phases de génération d'idées. De plus, le fait de forcer le concepteur à décrire le bénéfice attendu de son idée permet à l'équipe de conception de juger très rapidement de la pertinence de celle-ci.

Les cartes concepts

Le principe de cet outil est de fédérer l'ensemble des idées générées lors de phases créatives sous le couvert de concepts communs. Et ce de la manière la plus lisible possible. Effectivement, cet outil a pour finalité de résumer les idées générées afin de préparer leurs évaluations. Evaluations qui sont des phases essentielles nous permettant de centrer nos efforts de développement sur les éléments du service utiles et valorisables pour le client ainsi que rentables pour l'entreprise.
Cette étape de formalisation et de préparation peut s'avérer longue et difficile, mais elle est néanmoins indispensable afin de rendre l'évaluation ludique et agréable, assurant ainsi le confort de l'équipe d'évaluation. Elle maximise en outre la pertinence des résultats qui s'en dégagent.
Ces cartes se composent d'un nom associé à une image, qui les rendent reconnaissables facilement par l'équipe chargée de l'évaluation. Puis la carte décrit le concept qu'elle met en lumière suivant trois axes. Dans un premier temps en proposant une brève description avec une liste des principales caractéristiques du concept. Puis, elle expose la valeur ajoutée que propose ce concept d'un point de vue aussi bien qualitatif que financier. Enfin, elle liste les spécifications techniques qui ont été imaginées afin de réaliser la fonction. Selon le cas, ces cartes peuvent également être attachées à un dient particulier.
Dans les faits, le rassemblement des idées autour de concepts fédérateurs peut s'avérer complexe par la présence d'oppositions dans les idées générées, ou simplement par un trop grand nombre d'idées. Il est alors indispensable, afin d'encapsuler le maximum d'informations possible, de catégoriser les idées en fonction de leur granularité, ou niveau de détail. Les idées les plus détaillées seront énoncées dans la fiche concept comme étant une caractéristique qui cadre les principes que ces idées représentent.

Le parcours utilisateur

Le principe de cette méthode est de représenter une journée type du consommateur où l'influence du service est alors modélisée en termes d'impact et d'usage.
La majeure partie des efforts pour la réalisation de cet outil a normalement déjà été faite lors de l'identification des activités clés du service et avec la réalisation des cartes concepts. La réalisation du parcours utilisateur n'est plus qu'une traduction de ces deux éléments formalisant le concept du service sous la forme d'une succession de tâches utilisateur. Les personas peuvent également être utiles à la réalisation de cette représentation en forçant le concepteur à se projeter dans l'utilisation d'une solution en gardant la chronologie des étapes par lesquelles il doit passer pour accéder au service.
Une fois le parcours utilisateur modélisé, les points de contact avec l'utilisateur ainsi que les liens entre les concepts apparaissent d'eux-mêmes. Ce qui nous donne la première matérialisation de l'offre de service dans son ensemble. Cet outil permet alors d'avoir la première vue d'ensemble sur le service, ce qui nous offre alors la possibilité d'échanger plus simplement avec d'autres personnes extérieures à l'équipe de conception. Et ce dans le but de pouvoir dégager des contraintes ou des spécifications qui ne paraissaient pas évidentes au premier abord.

Le service blueprint

Cette méthode, dont le principe a déjà été discuté, reprend le parcours utilisateur dans son principe, et intègre également les notions de matérialisation physique du service pour le client, ainsi que les activités clés et les supports techniques à mettre en place pour réaliser la transformation de la valeur ajoutée.
Le *blueprint* une fois formalisé permet à la fois d'identifier les efforts à déployer afin de développer l'ensemble des briques fonctionnelles nécessaires à la livraison du service, mais il constitue également une représentation du service permettant d'entamer la communication avec les fournisseurs et les partenaires sur les spécifications des systèmes techniques.

Le business model

Cette méthode, également déjà évoquée dans la partie précédente, permet de rapprocher les coûts du service avec ses revenus attendus. Nous avons vu que le *business model* permet d'innover sur la manière de mettre en place une offre de service. Néanmoins, il offre également un support de suivi pour le développement d'un service. En effet, tout au long du cheminement vers le *blueprint*, et même au-delà, la vision du concepteur est amenée à changer, ces changements pouvant fortement impacter la rentabilité du projet. C'est pourquoi, afin de supporter la prise de décision concernant l'évolution du concept du projet, le *business model* offre un support de réflexion indispensable.

2.3.4. Synthèse sur le service

Dans cette partie, nous avons tenté de nous approprier la notion de service afin de percevoir ce qui le différencie du produit physique et d'établir des passerelles pour servir notre propos. Dans cette optique, nous avons commencé par définir ce que nous entendions par service. Puis nous avons fait état des guides et des bonnes pratiques attachées à leur conception avant de présenter une sélection d'outils pouvant nous être utiles.

Il nous apparaît alors ici que la principale caractérisation d'un service réside dans son intangibilité, ce qui semble convenir partiellement à notre système, principalement basé sur la valorisation de flux d'information. Néanmoins, cette volatilité dans la perception des services ne la rend pas incompatible avec l'offre d'un produit classique. Bien au contraire, celui-ci s'avère d'ailleurs être complémentaire et permet à l'entreprise d'opérer une transition vers des nouvelles perspectives de revenus.

En ce qui concerne les approches de conception, la littérature propose moins de modèles de processus permettant de guider le concepteur dans la création de nouveaux services, que de modèles applicables à la conception de produit. Nous pouvons observer que la plupart sont axés vers des outils et méthodes permettant de représenter et de formaliser l'offre de service. Notre attention est ici tout particulièrement portée vers le *service blueprint* dont le potentiel dans notre cas semble important. Effectivement, ce dernier propose de définir le service conformément à notre vision du système, à savoir une séquence d'interactions avec l'utilisateur. De plus, ce dernier semble suffisamment flexible et évolutif pour supporter une approche itérative et progressive dans sa réalisation.

Cela étant dit, les approches purement orientées vers le service semblent être extrêmement performantes afin de capter les besoins des clients et utilisateurs. Néanmoins, nous pouvons noter de gros manques en ce qui concerne l'implémentation et le déploiement d'une offre de service. C'est pourquoi il nous semble pertinent d'utiliser cette approche en phase amont de la conception, le tout au sein d'un modèle de processus prescriptif général.

2.4. *Machine Learning*

Enfin, dans cette dernière partie de l'état de l'art, nous partons du principe que la conception d'un système basé sur l'exploitation des flux d'information au sein d'une infrastructure complexe nous amène naturellement à étudier les outils de valorisation de données. Effectivement, lors de la génération de concepts pour la réalisation d'un système d'internet des objets, les concepteurs et experts vont naturellement s'orienter vers une architecture fonctionnelle pour des services permettant d'une part de valoriser les compétences de leurs métiers au sein de l'offre produit qu'ils proposent. D'autre part, la connexion entre l'ensemble des produits vise à créer un système de contrôle des produits basé sur une sorte d'intelligence collective où l'information collectée peut être valorisée sous forme de connaissance afin de permettre une meilleure adaptation des produits à leur environnement.

C'est pourquoi nous souhaitons dans cette partie tenir compte des systèmes de découverte automatique de connaissance. Relativement à notre problématique industrielle, nous cherchons par là à clarifier la constitution du système. Nous chercherons dans cette partie à positionner un système d'internet des objets par rapport aux technologies de la science des données afin d'en dégager l'impact sur les fonctionnalités, l'architecture, mais également sur la manière de conduire la conception.

De ce fait, nous chercherons ici à explorer les opportunités qui nous sont offertes au travers du domaine de l'intelligence computationnelle. Dans un premier temps, nous tâcherons de nous positionner et de comprendre les tendances qui régissent ce domaine. Ensuite, nous investiguerons les modèles méthodologiques qui sont recommandés pour mettre au point de tels systèmes. En dernier lieu, nous essaierons de comprendre les outils et techniques à disposition pour le concepteur afin de prototyper et réaliser de tels projets.

2.4.1. Positionnement du *Machine Learning*

Lors de nos recherches préliminaires, de nombreux termes sont apparus pour décrire les techniques d'extraction de connaissances dans les données. Parmi eux, nous pouvons citer la fouille de données, l'extraction de connaissances, la découverte d'information, l'archéologie des données ou encore le traitement de schéma de données. C'est à la lumière de l'ensemble de ces termes que nous sentons la nécessité de nous positionner vis-à-vis de ce domaine. Dans cette optique, nous ne garderons que trois termes que nous expliciterons : le *Machine Learning* (ML), le *Data Mining* (DM) et le *Knowledge Discovery in Databases* (KDD). Ces trois termes renvoient à des compétences, des techniques et des processus transverses aux mathématiques appliquées, aux statistiques et aux sciences computationnelles.

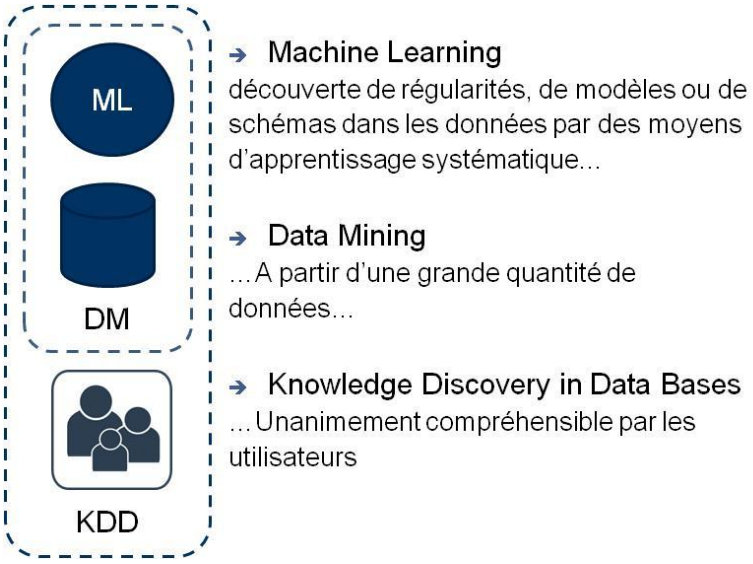


Figure 19 : Positionnement du *Machine Learning*

Historiquement, le *Machine Learning* introduit par (Samuel 1959) est considéré comme une part entière des sciences computationnelles. Cette science à présent mature a pour principal objet la découverte de régularités, de modèles ou de schémas dans les données par des moyens d'apprentissage systématique (Michell 1977). Cela afin de permettre à une machine la réalisation de tâches complexes difficilement réalisables par l'application d'algorithmes classiques.

Par la suite, et en parallèle du développement des systèmes de bases de données dans les années 1980, le *Data Mining* est un terme fortement employé par les statisticiens et désigne selon (Klosgen & Zytkow 1996) la mise en application d'algorithmes ayant pour vocation l'analyse ou la découverte de schémas à partir d'une grande quantité de données. Cela de manière spécifique pour la fourniture d'un jeu de données et la réponse à un cas d'usage particulier. Elle met en avant les méthodes développées dans le cadre du ML, mais s'attache également au processus de mise en œuvre avec les traitements amont à faire sur les données, l'évaluation des modèles ainsi que la prise en compte des performances computationnelles liées à la génération de modèles.

Dans le même temps, le *Knowledge Discovery in Databases* apparaît en 1989 avec (Pieatetsky-Shapiro 1991). Par la suite, (Fayyad 1996) définit la découverte de connaissances dans les bases de données comme un processus non trivial d'identification de schémas nouveaux, valides, potentiellement utiles et unanimement compréhensibles dans les données. Contrairement au DM, et même si les algorithmes appliqués sont similaires, le KDD se dégage par l'intégration de l'utilisateur et la réduction de la nécessité d'une connaissance poussée dans les sciences de la données pour pouvoir utiliser ces technologies. Cela repose sur deux axes principaux : la sélection de la bonne méthode d'analyse des données et la restitution de la connaissance de manière explicite (Frawley 1992 ; Fayyad 1996b).

2.4.2. Méthodologies et bonnes pratiques du *Machine Learning*

Les principes de résolution de problèmes dans le cadre du KDD et du DM ont été mis en exergue par la formalisation de nombreux processus méthodologiques. Ces derniers se présentent comme des modèles à phases itératives similaires aux méthodologies agiles du logiciel. Cependant, ces modèles intègrent les contraintes et caractérisent les efforts du concepteur suivant des tâches qui sont propres aux problématiques du domaine de la science des données. Nous exposons dans cette partie les processus de conception de la science des données afin d'en clarifier leur impact dans le cadre de leur intégration dans un projet de conception d'un système d'internet des objets.

Parmi ces nombreuses méthodologies disponibles dans la littérature, (Krugan & Musilek 2006) nous proposent un état de l'art des plus couramment utilisées telles que celles développées par (Fayyad 1996c), (Cabena & al. 1998) CRISP-DM (Shearer 2000 ; Wirth & Hipp 2000) ou (Cios & al. 2000 ; Cios & Kurgan 2005). De œ fait, et malgré certaines divergences dans les démarches comme le positionnement du choix des techniques, (Krugan & Musilek 2006) dégagent de cet état de l'art un modèle générique en six phases.

1) La compréhension du domaine d'application,
2) La compréhension des données,
3) La préparation des données et l'identification des techniques de fouille,
4) L'application des techniques sélectionnées et la création du modèle,
5) L'évaluation du modèle,
6) La consolidation des connaissances et le déploiement.

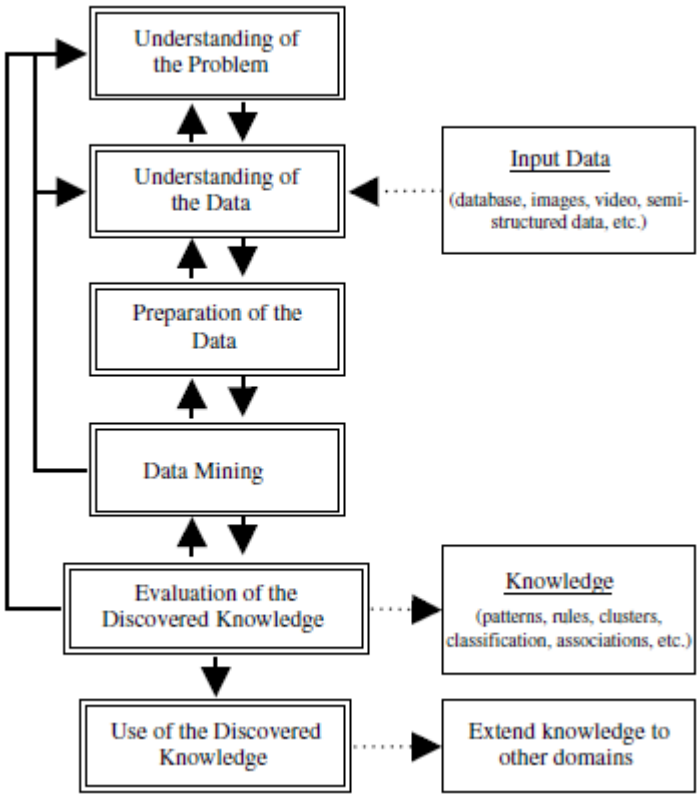


Figure 20 : Méthodologie de conception de la science des données (Krugan & Musilek 2006)

Avant d'entrer dans les détails de chacune de ces phases, il est bon de noter que (Krugan & Musilek 2006) proposent une estimation de la dispersion des efforts à fournir, autrement dit du temps passé, au travers de ces six phases. La principale caractéristique qui ressort de cette estimation est qu'en moyenne 50% des efforts sont localisés dans la phase de préparation des données. Cela s'explique principalement par deux raisons. La première tient à la faible qualité des données généralement disponibles dans les entreprises. Cela impose des efforts considérables afin d'évaluer et de rectifier les données. D'autre part, cela s'explique par la conception même des techniques de fouille. Ces dernières, afin de fonctionner, nous demandent de positionner nos observations dans un format particulier comme les données transactionnelles ou temporelles. Ceci demande nécessairement un effort de mise en forme. Cela étant dit, nous pouvons désormais expliciter chacune des phases en décrivant œ à quoi elles nous préparent.

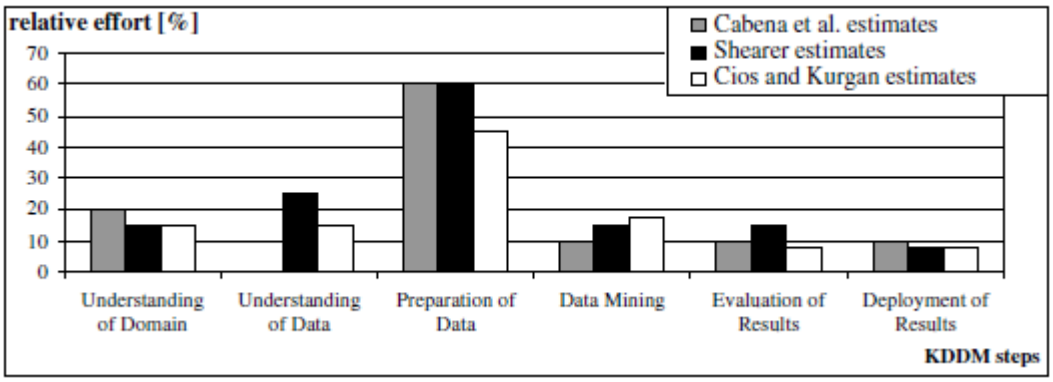


Figure 21 : Effort relatif aux étapes du processus de découverte de connaissances extrait de (Maimon & al. 2005)

Phase 1 : Compréhension du domaine d'application

Dans cette partie, il convient d'établir les spécifications de l'étude à réaliser. Pour ce faire, la personne en charge de l'étude, l'analyste doit s'imprégner des problématiques spécifiques au domaine d'application. L'analyste doit comprendre ici de quels leviers il dispose afin d'améliorer les affaires du demandeur d'étude. Il est d'usage d'investiguer à la fois l'univers des problèmes ainsi que l'univers des solutions.

En ce qui concerne la partie problème, nous devons ici définir une bonne question à nous poser afin de réaliser l'étude. Cette question doit faire référence à une problématique concrète pour le client, mais doit également faire sens pour l'analyste afin de déployer les bons outils lors du traitement statistique des données.

D'autre part, investiguer l'univers des solutions permet à l'analyste de se projeter dans le déroulement de l'étude avec une idée claire sur l'usage attendu de la solution par le client. Cette notion d'usage joue un rôle important sur les solutions qu'il devra mettre en œuvre afin de répondre à la question du client. Nous pouvons noter par exemple le besoin ou non de dynamisme ainsi que la mise en forme finale des résultats fournis par le système.

Enfin, lors de cette phase, l'analyste doit pouvoir dégager des moyens d'évaluation de la solution qu'il proposera. Cela peut prendre la forme d'indicateurs de performance permettant de mesurer la capacité de la solution proposée à répondre à la question du client de l'étude.

Phase 2 : Compréhension des données

L'étape suivante vise à prendre en main ce qui sera la substance de l'étude de fouille, à savoir, les données. Lors de cette phase, l'analyste se doit de comprendre ce qu'il devra manipuler afin de répondre à la problématique. Ces données sont des mesures d'un objet ou d'un système du monde réel. L'analyste doit alors savoir quels sont les types de données disponibles, quelle est leur qualité et quelle est la représentation physique de chaque variable.

Les données peuvent être de type quantitatives, avec des ratios, des intervalles, des ordinaux, des valeurs numériques ou qualitatives tels que des booléens, des valeurs nominales ou des catégories. Ces différents types peuvent alors influencer les traitements à effectuer. Leurs représentations physiques permettent d'évaluer le potentiel de chaque variable à pouvoir répondre à la problématique étudiée. Des échanges avec le client de l'étude sont alors nécessaires afin d'apporter des informations sur son activité et sur les moyens déployés pour collecter les données.

De plus, en parcourant les données, l'analyste peut commencer à évaluer leur représentativité et à repérer la présence de données invalides dans le jeu qui lui est fourni. Cela peut constituer une base afin d'estimer l'effort à mettre en œuvre pour la préparation des données pour le traitement.

La représentativité, fortement corrélée au nombre de données disponibles, définit si chacune des variables présentes dans le jeu de données client offre une vision réelle du comportement du système. Cela comprend la qualité de la distribution d'une variable, mais également, la distribution de l'ensemble des variables avec leurs dépendances. Si des données viennent à manquer, l'image du comportement du système que nous pouvons retrouver dans les données peut être fortement impactée et dans ce cas une phase de collecte de données complémentaires peut être demandée au client. Le cas échéant, un manque de représentativité des données peut également entraîner une restriction du périmètre fonctionnel étudié et impacter fortement la validité des résultats d'une étude.

D'autre part, à cause des imperfections du monde réel, et notamment de celles des moyens de mesure et de collecte, les données invalides dans un jeu de données peuvent prendre plusieurs formes :

- *Les données manquantes ou inexistantes* : Qui sont des trous ou des décalages dans les données. Nous nous retrouvons donc avec des champs non renseignés.

- *Les données aberrantes* : Qui sont des données très éloignées de leurs homologues.

- *Les inconsistances* : Qui sont des données représentant le même objet portant des noms différents. Ce phénomène est commun lorsque nous traitons des données venant de plusieurs sources.

- *Les biais* : Qui sont des erreurs récurrentes pour une même variable pouvant être dues à un mauvais étalonnage d'un appareil de mesure.

- *Les données polluantes* : Qui sont présentes sous forme de valeurs difficilement interprétables pour une variable. Cela peut être dû à un problème de copie, d'usage détourné du système ou encore de format tel que les dates.

Par ailleurs, il est également bon de vérifier, pour chaque variable, la granularité – ou le niveau de détail, le domaine de validité d'une variable, les relations entre les variables et les combinaisons permises.

Lors de cette phase, l'analyste doit enrichir au maximum sont point de vue sur les données afin d'en rendre compte au client et de savoir si les données qui semblent être invalides sont bien des erreurs ou si ces phénomènes sont naturels.

Phase 3 : Préparation des données et identification des techniques de fouille

Une fois les données comprises et le problème bien cadré, l'étape suivante revient à préparer les données pour le traitement. Ces préparations doivent être faites dans l'optique d'un traitement ciblé. Le choix des techniques d'analyse doit être fait au regard du rendu attendu par l'utilisateur ainsi que son futur usage. Une mise en forme de l'échantillon de données est alors nécessaire car les différents algorithmes de traitement possèdent leur propre sensibilité : aux types des variables d'entrée, aux valeurs manquantes ou aux valeurs aberrantes. Le cas échéant l'application d'un algorithme de traitement pourra être biaisée, voire impossible.

Dans la majorité des cas, l'application d'un algorithme nécessitera la création d'un tableau de données unique avec des types de variables spécifiques. Dans ce cadre, nous retrouvons quelques standards :

- *Le pivot* : Cette méthode consiste à changer la disposition du tableau de données brutes par le changement de notre variable de référence. Si par exemple nous disposons de données sous forme d'une série temporelle, avec une liste de transactions définie par un client et un mode de paiement, notre intérêt peut se porter sur les dernières transactions de nos clients pour chaque mode de paiement. Nous passerons alors d'une série temporelle à un tableau ordonné par client.

- *L'extraction de caractéristiques* : Certaines variables disponibles dans l'ensemble de nos données brutes ne peuvent disposer que d'un faible intérêt pour notre étude dans leur état actuel. Il est parfois plus pertinent d'effectuer des ratios, des différences ou des opérations propres au métier du client afin d'obtenir de nouvelles variables plus représentatives du comportement du système étudié.

- *Le remplacement ou la suppression des valeurs invalides* : Si nous avons statué avec le client sur l'invalidité de certaines valeurs, il est d'usage de les transformer afin de limiter les erreurs d'interprétation des résultats ou simplement afin de pouvoir appliquer l'algorithme. Nous pouvons alors simplement retirer les individus de nos données d'entrée ou alors tenter de les remplacer par des données valides. Dans ce dernier cas, nous devons alors estimer une valeur cohérente, soit par rapport aux autres valeurs de cette même variable en prenant la moyenne ou la médiane, soit en nous reposant sur les autres variables en tâchant de prédire quelle pourrait être la valeur la plus probable pour notre individu manquant.

- *La discrétisation* : Cette opération vise à changer le type de la variable que nous voulons traiter. Les principales transformations nous permettent de jongler entre des valeurs numériques, nominales ou binaires.

- *La normalisation* : Cette opération a pour but de réduire l'effet de pondération que nous pouvons observer entre certaines variables et qui est dû à leurs domaines de validité qui peuvent être plus ou moins grands. Cette phase de normalisation est particulièrement importante lorsque nous souhaitons appliquer un traitement qui se base sur une mesure de distance ou de similarité entre les individus présents dans nos données.

Une fois la préparation réalisée afin de s'assurer que notre algorithme est applicable, nous devons séparer notre jeu de données en deux. D'une part avec un échantillon qui servira à créer notre modèle, avec d'autre part un échantillon qui nous servira à tester et à valider notre modèle.

Phase 4 : Application des techniques et création du modèle

Une fois nos données prêtes à être traitées, nous pouvons maintenant extraire les connaissances qu'elles renferment par l'application de l'algorithme pour lequel nos données ont été préparées. Nous ne détaillerons pas ici l'ensemble des méthodes applicables pour créer un modèle de connaissance, ce qui sera fait dans la prochaine partie de cet exposé.

Phase 5 : Évaluation du modèle

Une fois notre modèle de connaissance extrait, nous nous devons de le valider, et ce de deux manières. D'une part en validant le modèle en termes de performances, puis, d'autre part, en analysant les résultats vis-à-vis de la problématique initiale.

La performance d'un modèle se mesure par le biais de nombreuses métriques (Japkowicz 2006), principalement basées sur une différence entre une prédiction et la valeur réelle pour un panel d'individus connus. À titre d'exemple, nous y retrouvons la précision, autrement dit le taux de bonnes

prédictions du modèle ainsi que la stabilité, à savoir la vitesse à laquelle le modèle converge. Ces différentes métriques de performance peuvent être optimisées en jouant à la fois sur le paramétrage de l'algorithme, mais aussi sur la qualité des données d'entrée.

En ce qui concerne la faculté du modèle à répondre à la problématique, une analyse des résultats doit se faire avec le client, qui pourra alors juger si le modèle fournit des résultats satisfaisants et si ses performances sont suffisantes pour être appliquées dans son activité.

Phase 6 : Consolidation des connaissances et déploiement

Une fois le modèle validé par le client, il est alors possible de le déployer dans une solution industrialisable. Cette phase vise alors à intégrer le processus de traitement sélectionné au flux d'informations géré par le client. Il est alors bon d'adapter les résultats que procure le modèle à l'usage que l'on souhaite faire de l'information recueillie.

2.4.3. Les outils et techniques du *Machine Learning*

Maintenant que nous avons étudié la mise en œuvre d'une démarche d'extraction de connaissance à partir de données, nous pouvons nous intéresser plus en détails aux méthodes de traitement constituant la boîte à outils du concepteur de tels systèmes. Nous ne pouvons néanmoins pas être exhaustif dans la description des différentes techniques. D'une part en raison de leur nombre, et d'autre part du fait des nombreuses implémentations différentes de concepts similaires. À titre d'exemple, nous pouvons citer ID3, C4.5 ou CART qui sont trois algorithmes permettant la génération d'arbres de décision, mais reposant eux-mêmes sur des principes mathématiques différents. De ce fait, nous avons décidé de baser notre veille autour d'une classification en quatre grandes catégories fonctionnelles issues de différentes lectures (Fayyad 1996c ; Han & Kamber 2006 ; Maimon & al. 2005 ; Pyle 1999 ; Fry 2004) ; les techniques prédictives, les techniques descriptives, les outils de visualisation et les techniques de transformation des données. C'est dans ce contexte que nous pouvons explorer chacune de ces catégories afin d'expliciter quels sont les outils les plus utilisés, dans quelle phase du processus de conception ils sont utilisés et dans quel but. Cela nous permettra alors de clarifier l'impact fonctionnel que peut avoir la science des données sur un système d'internet des objets.

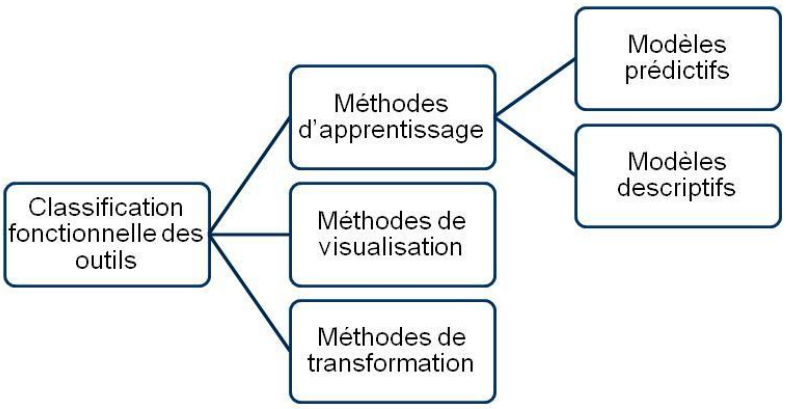


Figure 22 : Classification fonctionnelle des outils du *Machine Learning*

Techniques prédictives

Les techniques prédictives font partie des méthodes de découverte de connaissance, elles visent à extraire des motifs ou des règles implicites automatiquement à partir de données. Selon (Fayyad 1996c), elles se distinguent des méthodes descriptives par le fait de chercher à prédire le comportement futur d'un système observé. D'autre part, nous pouvons raffiner cette description en accord avec (Han & Kamber 2006 ; Maimon & al. 2005) en introduisant la notion d'algorithme supervisé. Autrement dit, ces techniques se basent sur un ensemble connu de comportements du système observés afin d'en inférer un modèle nous permettant de prédire son comportement futur.

Ces types d'algorithmes peuvent remplir deux tâches distinctes :

- *La classification* : Les méthodes de classification sont parmi les plus nombreuses. Nous pouvons citer k-NN (*k-Nearest Neighbors*), SVM (*Support Vector Machine*), *Naives Bayes* ou les arbres de décision. Le but de ces méthodes est d'apprendre, puis de prédire la classe d'un individu (variable nominale) en fonction des autres variables.

- *La régression* : Les algorithmes de régression sont utilisés afin de prédire une valeur numérique en fonction des autres valeurs de notre jeu de données. Principalement, cette méthode tâchera de construire son modèle en faisant correspondre les valeurs de la variable que nous souhaitons prédire avec les variations des autres variables décrivant chaque individu.

À la frontière entre ces deux familles de méthodes, nous retrouvons les *réseaux de neurones.* Ces méthodes sont un cas d'application particulier des méthodes de régression. Le format du schéma extrait par ces algorithmes est un réseau composé de nombreux neurones, où chacun correspond à l'application d'une fonction élémentaire affectant une portion de l'espace des solutions. Pendant la phase d'apprentissage, l'algorithme tâchera de minimiser les écarts entre la réalité et la prédiction faite, tout comme les algorithmes de régression. Cependant, ils sont également communément utilisés afin de réaliser des tâches de classification.

Techniques descriptives

Les techniques descriptives, faisant également partie des méthodes de découverte de connaissance, ont pour vocation de faire émerger des modèles représentatifs du comportement du système observé et de les représenter d'une manière compréhensible par l'être humain.

Nous pouvons noter que certaines techniques prescriptives possèdent une fonction descriptive du modèle estimé, mais la tâche de description revient généralement à des techniques dites d'apprentissage non supervisées. Autrement dit, ces techniques permettent de générer des modèles de connaissance en créant des ensembles homogènes d'individus. On qualifie ces techniques de « non supervisées » car elles sont autonomes dans la génération du modèle et elles ne dépendent pas de la fourniture d'une variable à prédire.

Nous retrouvons parmi ces techniques :

- *Le clustering* : Il permet d'identifier des groupes d'individus similaires dans notre jeu de données.

- *Les règles d'association* : Cette méthode permet de trouver, pour des variables nominales, quelles sont leurs valeurs associées les plus fréquentes. Cela permet d'obtenir un modèle de prédiction d'une variable nominale sous la forme « SI <condition> ALORS <valeur> ».

- *Modèle de séquences généralisées* : Cet outil permet de manipuler des séquences. Autrement dit des séries temporelles d'états de notre objet d'étude. Grâce à cette méthode, nous pouvons identifier quelles sont les suites d'états que prend notre objet le plus fréquemment.

Nous pourrions ajouter ici les techniques de visualisation. Néanmoins, étant utilisées de manière systématique de la part des concepteurs, quelle que soit l'application métier envisagée, nous leur consacrons une catégorie à part entière.

Techniques de transformation

Les techniques de transformation représentent un panel d'outils standards fréquemment utilisés par les analystes afin de mesurer la pertinence du jeu de données ou d'effectuer des manipulations de préparation des données en vue d'un traitement spécifique. Néanmoins, nous leurs réservons une famille, notamment au regard du temps nécessaire à la préparation des données. Certaines de ces techniques peuvent être utiles afin de faire de l'extraction de caractéristiques sur nos observations ou bien afin d'améliorer la qualité générale de notre jeu de données. En outre, certaines techniques présentées ici constituent en elles-mêmes la réponse à certains besoins.

Nous pouvons retrouver ici les traitements suivants :

- *Les matrices de corrélation et covariance* : Ces deux méthodes nous permettent de trouver les dépendances entre nos variables. En d'autres termes, ces matrices nous montrent quelles variables évoluent simultanément et si elles le font dans le même sens, ou en opposition.

- *La détection de valeurs aberrantes* : Souvent basée sur des notions de distance ou de densité, la détection de valeurs aberrantes permet de trouver dans nos données les individus qui semblent être très éloignés de l'ensemble de la population.

- *Décomposition en tendance, bruit et saisonnalité* : Cette méthode vise à décrire une série temporelle en fonction de trois facteurs. Le premier étant la tendance, obtenue par l'application d'une régression qui sera détaillée dans les outils prédictifs. Puis nous pouvons isoler le bruit, qui est souvent une perturbation à haute fréquence, pouvant être obtenu en lissant les données. Puis la saisonnalité représente le comportement périodique propre de notre variable.

- *L'analyse en composantes principales* : Cet outil permet à l'origine de réduire le nombre de dimensions (variables) que nous avons dans notre espace de données. Pour ce faire, cette analyse trouve les axes dans lesquels la dispersion de nos données est maximale. Nous pouvons nous en servir ici afin de déterminer quels sont les axes dans notre espace de données qui représentent au mieux l'information qu'elles contiennent.

- *Les tests de vérification* : Ces tests, issus du domaine des statistiques, sont en charge de vérifier la correspondance des données vis-à-vis d'un postulat de base. Par exemple, le test du χ^2 (chi-deux) permet de vérifier l'appartenance d'une variable à une loi de distribution donnée. L'ANOVA (ou Analyse de la variance) permet quant à elle de vérifier la distribution de deux échantillons de données afin de s'assurer qu'ils sont issus d'une même population.

Techniques de visualisation

Les techniques de visualisation de données représentent selon (Ware 2000) un intérêt certain dans le cadre de l'extraction de connaissances.

> *Le système visuel humain peut percevoir des modèles avec énormément de ressources et de subtilités. L'œil et le cortex optique du cerveau se*

comportent comme un lien à très haut débit d'information vers le centre cognitif humain, à l'image d'une multitude de machines en parallèle. À un plus haut niveau d'abstraction, perception et cognition sont intimement liées, c'est pourquoi les mots « comprendre » et « voir » sont synonymes.

Et effectivement, pour notre propos, ces techniques sont utilisées par les concepteurs en charge de l'extraction de connaissances. Les représentations visuelles peuvent alors être considérées comme de potentiels prototypes guidant les concepteurs dans la réalisation de son traitement. D'autre part, et c'est la raison pour laquelle nous aurions pu associer la visualisation aux techniques descriptives, ces représentations peuvent constituer un moyen d'interaction avec l'utilisateur afin de supporter une prise de décision.

Nous pouvons distinguer 5 grandes catégories de représentations (Börner 2013):

- *Les tableaux* : Ils visent à disposer des objets de représentation, tels que des données ou des graphiques, sous forme de matrices ou de listes. Ceux-là ont l'avantage de pouvoir ordonner nos objets de représentation en fonction d'index, et donc d'améliorer la lisibilité lorsque beaucoup d'objets sont disposés les uns à côté des autres.

- *Les graphiques* : ll s'agit de la famille la plus fournie en standards. Nous y retrouvons l'ensemble des outils couramment utilisés dans le processus d'extraction de connaissance afin de représenter des données. Parmi eux, nous avons l'histogramme, le *box plot* ou le *scatter plot.* Ces graphiques ont la particularité de représenter des données suivant des axes et des échelles bien définis.

- *Les diagrammes* : Ce sont des représentations graphiques sans repères d'échelle, contrairement aux graphiques. Parmi eux, nous retrouvons notamment des diagrammes de représentation de flux largement utilisés pour modéliser un processus. Mais nous pouvons également retrouver de nombreux autres types de représentations, tel qu'un plan de ligne de bus ou un diagramme à secteur.

- *Les réseaux* : Ce sont des représentations qui visent à montrer les relations entre les différents individus d'un jeu de données. Les données sont donc représentées par des éléments graphiques qui sont mis en relation les uns avec les autres. Parmi eux, nous retrouvons les représentations en forme d'arbre ainsi que les diagrammes nœuds-liens.

- *Les cartes* : Ce sont des représentations qui utilisent des données géographiques (longitude et latitude) afin de positionner les données. Ces graphiques sont souvent associés à un fond de carte représentant les frontières entre les différents lieux (régions ou pays) ou certains éléments particuliers de la zone représentée (fleuves, reliefs, routes, ...)

Dans l'ensemble de ces représentations, nos données doivent être transcrites dans un langage graphique. L'encodage des données doit se faire en fonction du type de variable que nous avons ainsi que des caractéristiques graphiques à disposition. Lors de la représentation dans un espace en deux dimensions, (Börner 2013) fait état de paramètres tels que la forme, la couleur ou la position qui sont plus ou moins bien adaptés à la nature plus ou moins qualitative ou quantitative des données.

Enfin, (Tufle 1983 ; Tufle 1990) nous propose de suivre un certain nombre de bonnes pratiques dans le cadre de la représentation graphique. Ces bonnes pratiques limitent les risques de mauvaises perceptions de l'information par l'utilisateur. Effectivement, la majorité des graphiques doivent nous

permettre d'effectuer des comparaisons ou des analyses. Certains éléments tels que les échelles, les labels, les formes ou les textures utilisées peuvent nuire à la lisibilité du graphique proposé. De plus, et dans certaines mesures, ces éléments polluants peuvent même nuire à la confiance que porte l'utilisateur envers l'information que l'on souhaite lui transmettre. Ces bonnes pratiques peuvent être exprimées de la manière suivante :

- Choisir un format de représentation adapté aux données et aux questions qu'elles doivent soulever.
- Les échelles proposées doivent avoir un sens et un équilibre en vue de l'information transmise par le graphique.
- Utiliser des nombres et des mots dans de justes proportions pour accompagner le graphique.
- Montrer un niveau accessible de détails.
- Toujours s'assurer d'avoir une bonne qualité narrative, une histoire à raconter au travers des données.
- Construire des graphiques avec suffisamment de détails techniques sur son domaine d'application, mais les utiliser avec précaution.
- Eviter les effets de décoration qui peuvent polluer la lecture du graphique.

2.4.4. Synthèse sur la découverte de connaissances

Dans cette partie nous avons étudié les technologies de découverte de connaissance et leur mise en œuvre. À la lumière du paysage technologique dans lequel nous sommes baignés, les concepteurs s'y réfèreront très certainement lors de la génération de concepts. Nous souhaitions alors par cette revue, clarifier l'impact de ces technologies sur la constitution technique et fonctionnelle d'une plateforme d'internet des objets. Pour ce faire, nous nous sommes tout d'abord attachés à nous positionner vis-à-vis de la science des données. Puis nous avons fait émerger une méthodologie de conception de la mise en œuvre de systèmes de découverte de connaissances acceptée par la littérature. Enfin, nous avons proposé une vue d'ensemble des outils et techniques liés à la découverte de connaissances basée sur leurs fonctions.

En premier lieu, nous pouvons constater que la science des données dispose d'une large palette d'outils permettant la manipulation de différentes typologies de données et permettant de répondre à différents buts. Ces outils reposent sur des principes mathématiques complexes, mais sont disponibles par le biais de modules logiciels du marché. Nous retrouvons ici la notion de plateforme de traitement de la donnée, avec des environnements tels que R ou Python pour lesquels il existe de nombreuses librairies spécifiques au traitement de données.

D'autre part, la mise en application des outils afin de créer des modèles de prédictions ainsi que l'évaluation de la qualité de ces modèles sont des procédures relativement simples et standardisées. La complexité du développement d'un modèle provient principalement du prétraitement nécessaire de notre jeu de données afin de le rendre compatible avec un algorithme préalablement sélectionné.

Les méthodologies d'application d'algorithmes de découverte de connaissances semblent matures et unanimement adoptées par l'académie et l'industrie. Cependant, les activités de conception gravitent autour des jeux de données, et non autour des utilisateurs et de leurs besoins. Ces démarches restent focalisées sur la fourniture d'un traitement de qualité alors que son usage et ses raisons d'exister sont un

peu oubliés. Prise en compte dans les phases amont du développement, la compréhension des besoins des utilisateurs reste floue et peu cadrée.

Enfin, issue d'un besoin de créer des systèmes autonomes, la découverte de connaissances (KDD) ouvre la voie vers un repositionnement de l'utilisateur au centre du développement. En effet, l'une des facettes de cette démarche réside dans la représentation des modèles prédictifs afin de rendre explicites les motifs découverts dans les données. À la lumière de ces efforts fournis pour opérer un transfert plus simple de ces technologies vers leurs utilisateurs, nous pouvons être confiants dans les potentiels d'intégration de ces champs technologiques dans un projet plus vaste d'exploitation de données tels que les systèmes d'internet des objets.

2.5. Synthèse bibliographique et constat

Dans cet état de l'art, nous cherchions à apporter des éléments de réponse à la problématique industrielle que nous avions préalablement énoncée. Celle-ci nous questionne sur les moyens que l'entreprise doit mettre en œuvre afin d'ouvrir son domaine d'activité à de nouveaux services basés sur la réalisation d'un système d'objets connectés. Cela nous pousse alors à nous interroger sur le but d'un tel système, sur sa composition et sur la pérennité techno-économique de la nouvelle activité que cela crée pour l'entreprise.

C'est alors dans cette optique que nous nous sommes attachés dans un premier temps à définir l'environnement dans lequel s'inscrit le sujet de notre étude, à savoir, l'internet des objets. Puis, dans un second temps, nous nous sommes intéressés aux méthodes présentes dans la littérature et ayant pour vocation de structurer l'activité de conception. Ensuite, nous avons investigué les caractéristiques faisant les particularités de notre système. D'une part, par la nature même des éléments qui le composent, où le caractère intangible des flux d'information nous a conduits vers le concept de service. D'autre part, par les technologies disposant d'un fort potentiel de génération de valeur ajoutée face aux attentes que pourraient avoir de futurs utilisateurs de tels systèmes, ce qui nous a amené aux systèmes de découverte de connaissance dans les bases de données.

Cet état de l'art nous amène à la conclusion que le concept d'internet des objets fait appel à la mise en œuvre de systèmes complexes, où de multiples métiers et connaissances doivent être engagés. A lui seul, cet argument justifie la nécessité de guide et de support méthodologique pour les concepteurs. La littérature expose alors de nombreuses possibilités principalement issues de la conception de produit. Parmi elles, nous pouvons retenir qu'il n'existe pas d'approche dédiée et unanimement utilisée pour la conception d'offres produits-services dans le numérique. Cependant, des modèles suffisamment génériques sur lesquels nous pouvons nous reposer existent. Ceux-là formalisent et cadrent les besoins de la conception quelle que soit la nature de l'artefact à concevoir.

D'autre part, la nature de notre système favorise l'intégration de domaines de connaissances variés avec une architecture de type plateforme, où chacune des strates peut être conçue indépendamment pour répondre à un panel d'usages puis intégrée pour de nombreux cas particuliers. Le corolaire de cette propriété nous amène à voir notre propre système comme une plateforme, où une application spécifique peut être conçue pour répondre à un besoin client de manière suffisamment générique pour l'ouvrir par la suite à d'autres clients disposant de besoins similaires.

Concernant les différentes plateformes mises en lumière par la littérature, l'ensemble des domaines technologiques impliqués propose des réponses afin de satisfaire les besoins de tels systèmes. Même si de forts potentiels d'amélioration persistent, l'environnement technologique actuel est suffisamment mature pour voir émerger des écosystèmes connectés. Néanmoins, nous avons identifié les faiblesses de ces plateformes dans les couches hautes, plus proches des fonctionnalités et des usages que des technologies.

D'une part nous avons pu percevoir un manque dans le rôle que tient le domaine d'application dans ces systèmes. Ce qui nous a conduits à nous positionner vis-à-vis des services, montrant que ces derniers correspondent au mode de valorisation que nous attendons dans l'écosystème du domaine d'application. Autrement dit, une valorisation via un flux d'information intangible et par le biais de contrats périssables et impossibles à posséder. Les démarches proposées dans la littérature semblent très performantes pour capter et formaliser le besoin des clients. Néanmoins, ces approches sont

insuffisantes pour développer notre offre de bout en bout. C'est pourquoi il nous semble pertinent d'avoir une approche service dans les phases amont de la conception, le tout intégré dans une démarche plus générale.

Enfin, les technologies de valorisation de flux d'information et les systèmes de découverte de connaissance dans les bases de données semblent cloisonnés à la science des données. Nous avons tout de même pu constater que les outils et techniques sont accessibles par le plus grand nombre par le biais de plateformes de traitement de la donnée. Néanmoins, il est difficile, pour un non spécialiste, de percevoir les tenants et les aboutissants d'un cycle de conception d'un système de valorisation de la donnée. De plus, la donnée étant au cœur d'un système d'internet des objets, mais également au cœur du cycle de développement d'algorithmes de découverte de connaissance, nous pouvons nous attendre à avoir de fortes interactions entre ces deux systèmes. Cela met en avant notre besoin d'intégration de ces sciences au sein du développement d'un écosystème d'objets connectés. Afin de préparer cette intégration, nous avons pu, d'une part, faire un parallèle avec la conception logicielle et avec des démarches itératives et incrémentales. D'autre part, nous avons identifié les bonnes pratiques de la conception favorisant l'intégration de la connaissance en amont de la conception.

3. Problématique de recherche

Dans cette partie, c'est fort de cet état de l'art et au regard des enjeux industriels que nous exposerons les éléments de notre problématique de recherche. Nous chercherons notamment à découvrir dans cette étude les leviers qui peuvent pousser une entreprise principalement orientée sur la vente de produits à rendre ses produits connectés. Autrement dit, immergé dans un paysage technologique plein d'opportunités, sous quelles conditions l'investissement d'une entreprise pour rendre ses produits communicants lui apporterait un nouveau vecteur d'innovation.

Pour ce faire, nous mettons alors en perspective cet objectif avec les constats qui se dégagent de notre état de l'art afin d'énoncer une problématique de recherche et d'inférer des hypothèses de résolution.

3.1. Enoncé de la problématique

Pour élaborer cette problématique, nous pouvons commencer par rappeler ce que nous avons préalablement démontré.

- Un réseau d'objets constitue un système complexe dont la conception nécessite un support méthodologique adapté afin de cadrer, structurer et formaliser les efforts du concepteur au sein d'un environnement favorable à son activité.
- Ce réseau est structurellement constitué d'un enchevêtrement de briques technologiques où les efforts du concepteur se localisent particulièrement dans les couches supérieures relatives à l'usage de ces systèmes.
- Le service révèle un fort potentiel pour considérer l'utilisateur dans les phases amont de la conception, mais se révèle insuffisant à l'approche de la réalisation.
- Les technologies de découverte automatique de connaissance et de fouille de données semblent indispensables pour répondre au besoin des utilisateurs et doivent être abordées avec une approche intégrée dans le cycle de conception.

En considérant notre système comme une plateforme, nous ne nous limitons pas à la fourniture d'un produit autonome sur le marché, mais nous nous ouvrons à la fourniture d'un ou plusieurs services de manière extensible dans le temps et dans les usages. Cela nous ouvre alors les portes d'un véritable support de promotion de l'innovation pour le moyen et le long terme.

En outre, et afin de répondre aux besoins présents et d'anticiper les besoins futurs, il nous semble difficile de ne pas établir de parallèle entre notre système de flux d'information et les technologies de découverte de connaissance. Cela nous ouvre les portes de nombreux services pouvant être qualifiés d'intelligents, sous réserve de prendre en compte leur mise en œuvre en amont, pendant la conception.

C'est au travers de cette perspective que nous pouvons exprimer notre problématique de recherche de la manière suivante :

Comment concevoir une plateforme innovante de l'internet des objets par une approche intégrant simultanément les dimensions produit, service et données ?

Afin de répondre à cette problématique, nous nous reposons sur les constats préalablement explicités afin de formuler des hypothèses de résolution.

3.2. Formulation des hypothèses

De ce fait, la question ainsi formulée nous entraîne vers la fourniture d'un support méthodologique prescriptif pour le concepteur. Basée sur un modèle unanimement considéré comme générique, nous pouvons proposer une approche intégrée où le service et la découverte de connaissance dans les bases de données sont mis à l'honneur. Nous formulerons alors ici nos hypothèses de résolution suivant ces deux axes.

H1 Le service permet de définir et de structurer notre problème dans les phases amont de la conception.

H2 Les systèmes de découverte de connaissance dans les données doivent être pris en compte parallèlement au développement du reste du système afin d'assurer la faisabilité des fonctions intelligentes.

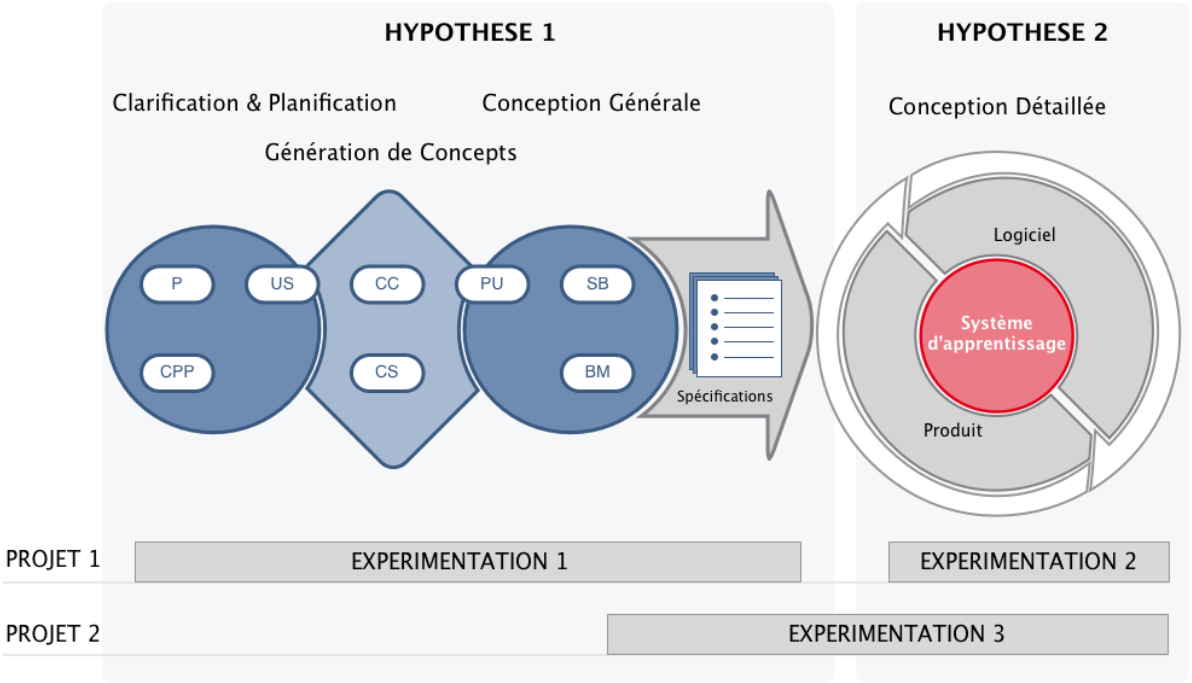


Figure 23 : Positionnement des hypothèses de recherche

Nous pouvons dès à présent représenter nos hypothèses au sein d'un processus de conception tel que le montre la figure 23. La première hypothèse H1 s'intéresse aux phases amont de la conception, alors que la seconde hypothèse H2 tient uniquement compte de la conception détaillée. C'est alors que nous pouvons noter, de par leurs positionnements, que ces deux hypothèses sont complémentaires au regard de notre problématique. Autrement dit seule la validation conjointe de nos deux hypothèses nous permettra d'apporter une réponse satisfaisante.

Dans le but de mieux percevoir quels en sont les tenants et les aboutissants, nous tâcherons par la suite de formaliser et d'expliciter chacune de ces hypothèses dans leur contexte.

H1 : L'approche service, moteur de la conception

Cette première hypothèse tend à promouvoir une approche service dans les phases amont de la conception. Effectivement, et comme nous avons pu le voir dans l'état de l'art, les approches liées à la conception de service disposent d'un potentiel incontestable pour encapsuler le besoin client et pour fournir au concepteur des voies de représentation afin de rendre sensible les leviers de conception d'un système par nature intangible.

C'est pourquoi nous développons l'hypothèse suivante :

Une approche service, dans le cadre du développement d'une plateforme d'exploitation d'un réseau d'objets connectés permet de structurer les phases amont. Elle permet d'une part de favoriser l'acceptation du marché, et donc l'innovation, en caractérisant le système en phase avec les besoins clients. Mais elle permet également d'effectuer la bascule entre la conception amont, abstraite, et la conception détaillée, beaucoup plus concrète.

C'est sous ce regard que nous souhaitons montrer que le service est moteur de la conception. C'est lui qui insuffle le premier mouvement vers une offre produit-service optimale. C'est également lui qui dresse les premières esquisses du service qui devront constituer notre base de spécification, tout en donnant au concepteur les moyens de faire des choix technologiques et d'établir des priorités fonctionnelles.

Toujours dans l'optique de fournir au concepteur un outil méthodologique prescriptif et opérationnel, nous nous basons sur notre état de l'art afin de fournir une démarche complète de qualification du service pour le concepteur. Pour cela, nous nous basons sur les outils et méthodes que nous avons identifiés comme pertinents pour notre cas. Nous pouvons de ce fait considérer le service *blueprint* comme une finalité, ce dernier permettant de matérialiser le service en tenant compte des acteurs avec leurs interactions, des tâches et procédures internes ainsi que des supports techniques nécessaires. Associé à un *business model*, il doit nous offrir un outil comparatif pour le choix des technologies à utiliser. Puis, associé à une évaluation du besoin et des concepts, il doit nous permettre d'établir aisément la première liste de spécifications avec leurs priorités pour les différentes couches du système, qu'elles soient physiques ou logicielles.

H2 : Un développement des fonctions intelligentes intégrées

Pour cette seconde hypothèse, nous nous positionnons dans la conception détaillée du système, à la suite de la formalisation du périmètre fonctionnel. Lors de cette phase, nous adoptons une démarche agile, représentée par un développement itératif et incrémental, où doivent être considérés les produits physiques dans la couche perceptive et réseau aussi bien que les logiciels dans la couche intermédiaire et l'application. L'agilité permet alors de se focaliser sur les priorités des clients de manière démonstrative, avec la fourniture de nombreux prototypes, tout en nous laissant suffisamment de flexibilité afin de gérer les interactions entre les différentes couches.

Ce mode de fonctionnement n'est toutefois pas problématique en soi, dans la mesure où l'intégralité des fonctions et leurs priorités ont pu préalablement être dégrossies. Bien au contraire, ce fonctionnement constitue même un environnement favorable à la démonstration que nous souhaitons mettre en place ici. En effet, nous nous intéressons à la mise en place des fonctions intelligentes du système. Or nous pouvons d'ores et déjà observer une contradiction dans le fait de vouloir développer un outil de découverte de connaissance basé sur un flux d'information qui n'existe pas encore, alors qu'un large historique serait d'ordinaire nécessaire avant de prétendre pouvoir commencer. Cependant, la qualité de nos modèles est influencée de manière directe par la constitution de ce flux d'information. A titre d'exemple, nous pouvons considérer le niveau de résolution des données transmises par les objets. Plus la résolution est haute et plus nous avons de chances de pouvoir extraire des schémas dans nos données et d'être précis dans nos modèles. D'autre part, et afin d'assurer cette résolution, l'ensemble de la chaîne d'acheminement de l'information doit se voir dimensionner en conséquent. Autrement dit, pour une même fonction, la prise en compte des technologies de découverte de connaissance lors du développement peut ici assurer la faisabilité de la fonction et sa rentabilité.

C'est dans cette optique que nous pouvons développer notre seconde hypothèse de la manière suivante :

> La pérennité d'un service reposant sur des fonctionnalités intelligentes, au sein d'un écosystème d'objets connectés, est assurée par une mise en œuvre agile des solutions technologiques. C'est donc par l'intégration d'experts dans le domaine de la science des données que nous serons à même de spécifier le système de manière optimale.

De ce fait, nous souhaitons démontrer ici qu'il est possible d'opérer des transferts de connaissance entre les différents représentants de l'ensemble des couches de notre système. C'est alors par ces transferts, couplés à la fourniture de prototypes, que le système doit pouvoir converger vers une solution faisable et optimale.

Nous souhaitons ici démontrer qu'il est possible de proposer un cadre méthodologique agile dans lequel les échanges entre les différentes parties prenantes et les efforts sont structurés afin de favoriser la convergence de notre système vers son optimum. En effet, l'agilité nous offre la possibilité de garder de la souplesse vis à vis des modifications, cependant, chaque modification coûte et particulièrement si la partie physique de notre système est impliquée. L'agilité révèle ici ses limites et c'est pourquoi notre démarche doit s'avérer suffisamment efficace pour freiner le moins possible le développement.

4. Expérimentations

Dans cette partie, et suite à l'expression de la problématique, nous verrons, par une suite de trois expérimentations, comment nous confrontons nos hypothèses de recherche à une réalité industrielle afin de les valider ou de les infirmer. Pour ce faire, nous nous positionnerons dans un premier temps pour les deux premières expérimentations dans le cadre du projet de re-conception d'une solution de chaudière connectée telle que nous l'avons discuté dans l'introduction de ce document. Puis, dans un second temps, nous sortirons de ce projet particulier afin de confronter nos découvertes à un problème distinct dans le but d'éprouver la résilience de nos propositions au contexte initial.

À la lumière de notre problématique de recherche, nous pouvons positionner nos expérimentations de la manière illustrée par la figure 24 ci-dessous.

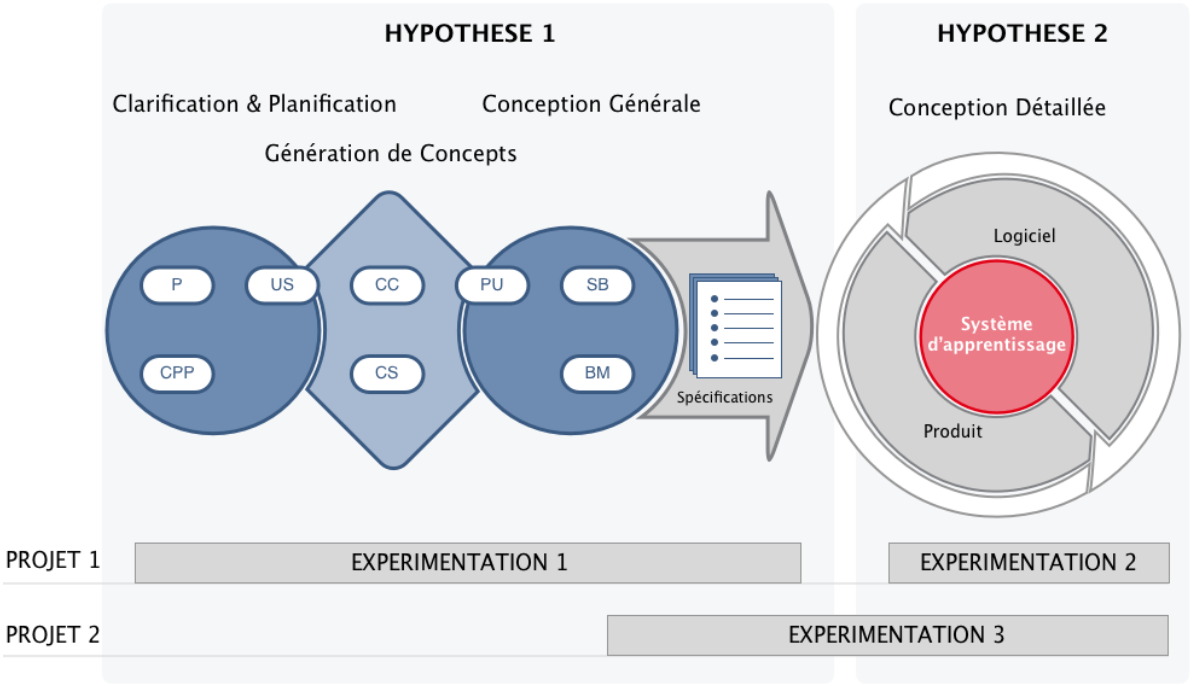


Figure 24 : Positionnement des expérimentations

Plus concrètement, la première expérimentation s'intéresse à la validité de la première hypothèse qui porte sur l'intérêt de passer par une approche service au démarrage de la conception afin de mieux percevoir notre problème. Pour ce faire, nous nous baserons sur la mise en application outils du service que nous avons fait émerger dans l'état de l'art en phase amont du processus de conception. Nous présenterons alors un séquencement de ces outils permettant d'aboutir à une conception générale satisfaisante et optimale dans une optique de réalisation.

La deuxième expérimentation vise à démontrer la véracité de la seconde hypothèse. Celle-ci pointe la nécessité d'aborder une approche intégrée dans la conception détaillée et le développement du système,

où les aspects produit physique et logiciel évoluent conjointement sous l'influence des systèmes de découverte de connaissance. Nous nous baserons alors sur la suite du projet de chaudière connectée afin de montrer les interdépendances entre les différents systèmes en explicitant les problèmes que cela soulève et les risques pour la pérennité du système. Nous proposerons alors une approche intégrée qui vise à prendre en compte les contraintes de conception incluses par les systèmes de découverte de connaissance en amont, et cela dans le but de favoriser le succès du déploiement des fonctions intelligentes du système.

Enfin, lors d'une troisième et dernière expérimentation, nous sortirons du cadre du projet précédent afin de vérifier que les solutions que nous proposons dans le cadre des deux premières expérimentations ne sont pas dépendantes du contexte. Nous nous positionnerons ici dans un projet de connectivité à l'échelle de l'entreprise où nous chercherons à développer un système intelligent en charge d'améliorer la production de l'usine. Par cette démarche, nous renforcerons alors les apports liés à la seconde hypothèse dans un premier temps, en montrant différentes facettes des dépendances entre les systèmes de découverte de connaissance et le système global. D'autre part, nous nous appuierons sur cette expérimentation pour montrer la complémentarité entre les deux hypothèses et la nécessité d'une approche service, en se référant aux outils précédemment utilisés.

4.1. Expérimentation 1 : Conception amont par le service

Dans cette partie nous exposerons puis discuterons de la première expérimentation relative à l'hypothèse 1 sur la nécessité d'intégrer en amont une approche service dans la conception d'un système d'objets connectés. Pour ce faire, nous détaillerons tous d'abord les objectifs de cette expérimentation. Puis nous présenterons la démarche mise en œuvre pour tenter d'atteindre ces objectifs. Ensuite, nous présenterons l'application que nous avons faite de cette démarche. Et enfin, nous analyserons les résultats afin de conclure sur la validité de notre hypothèse.

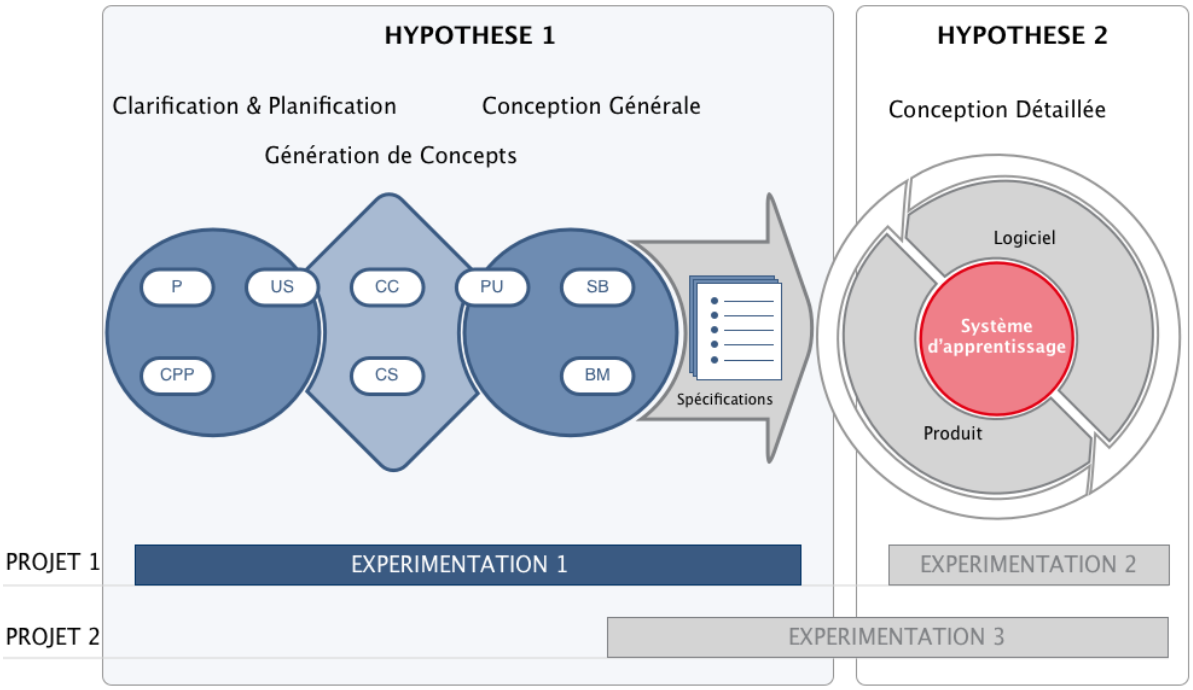


Figure 25 : Positionnement de l'expérimentation 1

4.1.1. Objectifs

Dans cette expérimentation, nous nous situons en phase amont de la conception de notre système. Nous lançons le projet avec la seule idée de rendre nos produits connectés, en l'occurrence dans notre contexte industriel des chaudières. L'objectif de cette partie est de déterminer un schéma logique pour le concepteur afin de définir au mieux son système et son problème de conception. En effet, lors de notre état de l'art, nous avons mis en lumière le fait qu'il paraît judicieux de réaliser le développement de notre système sur la base d'un fonctionnement agile. Or, nous avons vu, avec l'exemple de la méthodologie SCRUM, que le lancement des itérations incrémentales de développement nécessite une préparation amont afin de définir le périmètre de notre projet au travers de spécifications générales de notre système, appelées *product backlog*.

Pour notre sujet, au regard de la nature du projet, poussé par la technologie, il nous est donc très important de considérer cette phase de la conception afin de mieux comprendre quels seraient les usages attendus de notre système et les fonctionnalités clés que doit fournir notre système. Cela doit donc passer par une investigation de nos futurs utilisateurs et une analyse de leur besoin.

Pour ce faire, nous avons pris parti et démontré dans l'état de l'art que la conception de service se prêtait bien à l'exercice de par la nature intangible de notre système et de par sa faculté à analyser les besoins clients. L'objectif de cette expérimentation est donc, à l'aide des outils mis en lumière dans l'état de l'art, de proposer une démarche permettant d'encapsuler au mieux le besoin dient et de préparer la phase de développement. Nous mettrons alors notre démarche à l'épreuve dans notre cadre industriel de développement d'une chaudière connectée afin d'évaluer son impact sur le projet de conception et de discuter de sa pertinence vis-à-vis de notre première hypothèse de recherche.

4.1.2. Protocole

Afin de découvrir le périmètre fonctionnel de notre système ainsi que de comprendre et formaliser le besoin de nos utilisateurs, nous mettons en place une démarche de conception. Cette démarche prend part dans les trois premières phases de la conception, à savoir la clarification et la planification, la génération de concepts et la conception générale du système. Cette démarche se base sur un séquencement d'outils que nous avons présentés dans l'état de l'art.

Cette expérimentation c'est déroulée sur une période de 10 mois en sollicitant un total de 11 participants aux profils variés. Notons également que ces travaux ont étés au cœur de projets de fin d'études pour trois étudiants en master de recherche décrits en Annexe B.

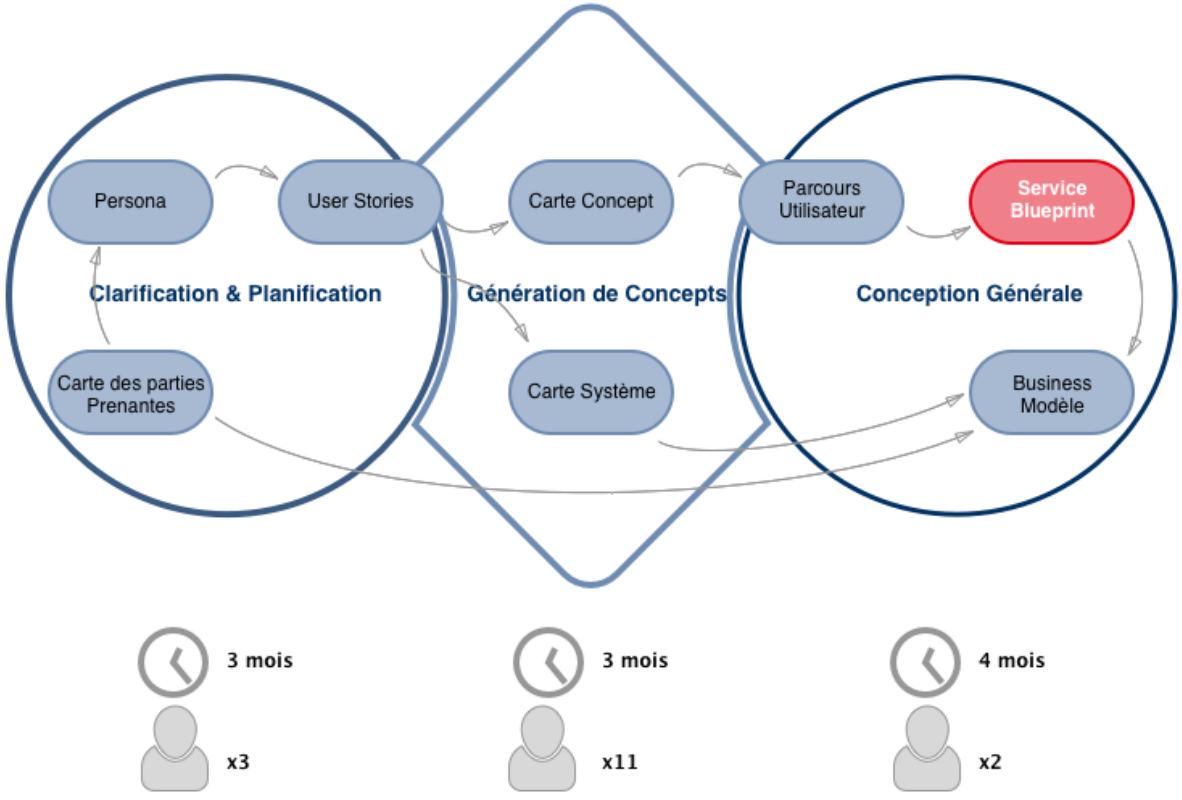


Figure 26 : Protocole de l'expérimentation 1

Tel que l'illustre la figure 26, ce cheminement que nous proposons commence lors de la clarification et de la planification par la formalisation du réseau des potentiels futurs acteurs du service au sein d'une carte des parties prenantes. Cette phase a été réalisée sur une durée de 3 mois avec trois intervenants, tous membres de l'équipe innovation de l'entreprise. L'investigation auprès des acteurs que nous pouvons identifier comme futurs clients peut être résumée et représentée sous forme de personas.

Ces derniers peuvent alors être utilisés comme sode pour la génération de concept. D'une part pour la phase créative de génération d'idées sous forme de *user stories* et d'autre part pour leur formalisation et leur évaluation sous forme de carte concept. Cela a été réalisé sur 3 mois avec 11 participants d'horizons divers. Nous y retrouvons notamment :

- 3 membres du service innovation
- 1 membre de l'équipe marketing
- 1 expert informatique
- 1 expert en télécommunication
- 2 experts en maintenance
- 1 fournisseur d'énergie
- 2 utilisateurs

Ensuite, avec la liste des concepts à embarquer dans la solution, nous pouvons commencer à définir d'un côté la structure du système technique au travers d'une carte système, puis de l'autre côté, détailler le service au travers du parcours utilisateur. Ces deux derniers éléments peuvent alors être détaillés dans une vue d'ensemble à l'aide de *service blueprint*. D'autre part, et pour clôturer la conception générale de notre système, nous pouvons, à l'aide de l'ensemble des éléments précédemment générés, positionner notre projet au sein d'une approche financière avec un *business model* afin d'en évaluer la rentabilité. Cette phase a été réalisée sur une période de 4 mois avec deux participants membres du service innovation de l'entreprise.

4.1.3. Application

Afin d'expliciter l'application de notre démarche précédemment proposée, nous présenterons, au travers des trois phases de la conception que nous détaillerons ici, la mise en application des outils que nous avons sélectionnés. Nous nous attacherons particulièrement à démontrer l'utilité de l'outil vis-à-vis de l'objectif général que nous nous sommes fixé pour la formalisation du besoin tout en explicitant les efforts nécessaires à sa mise en œuvre.

Phase 1 : Clarification et planification

La première étape de la conception est la clarification et la planification, où il nous faut découvrir et cadrer le contexte du projet. Nous partons donc ici de la première idée : réaliser un système permettant de connecter les chaudières afin de dégager une nouvelle proposition de valeur par le biais des flux de données que nous pouvons générer. Nous voulons alors ici nous positionner dans un écosystème où ces flux de données font sens et peuvent être valorisés. C'est dans cette optique que nous optons pour une représentation des acteurs constituant cet écosystème avec la carte des parties prenantes.

La carte des parties prenantes regroupe l'ensemble des acteurs, internes ou externes, qui peuvent jouer un rôle dans le service. Les contacts internes représentent l'ensemble des parties avec lesquelles nous disposons d'un contact direct à partir du positionnement du projet dans l'entreprise, autrement dit le service de développement. Elles sont principalement constituées des différents services de l'entreprise ou du groupe pouvant avoir un intérêt direct pour le système, comme le service après-vente constructeur ou les autres usines de production en Europe. L'environnement externe représente quant à lui l'ensemble des entreprises constituant le paysage industriel actuel pouvant jouer un rôle ou tirer parti de notre système. Nous y retrouvons les concurrents de notre marque, les utilisateurs finaux des produits, les entreprises de maintenance et de plomberie ou bien encore les différents fournisseurs de services tels que les entreprises de télécom ou de fourniture d'énergie.

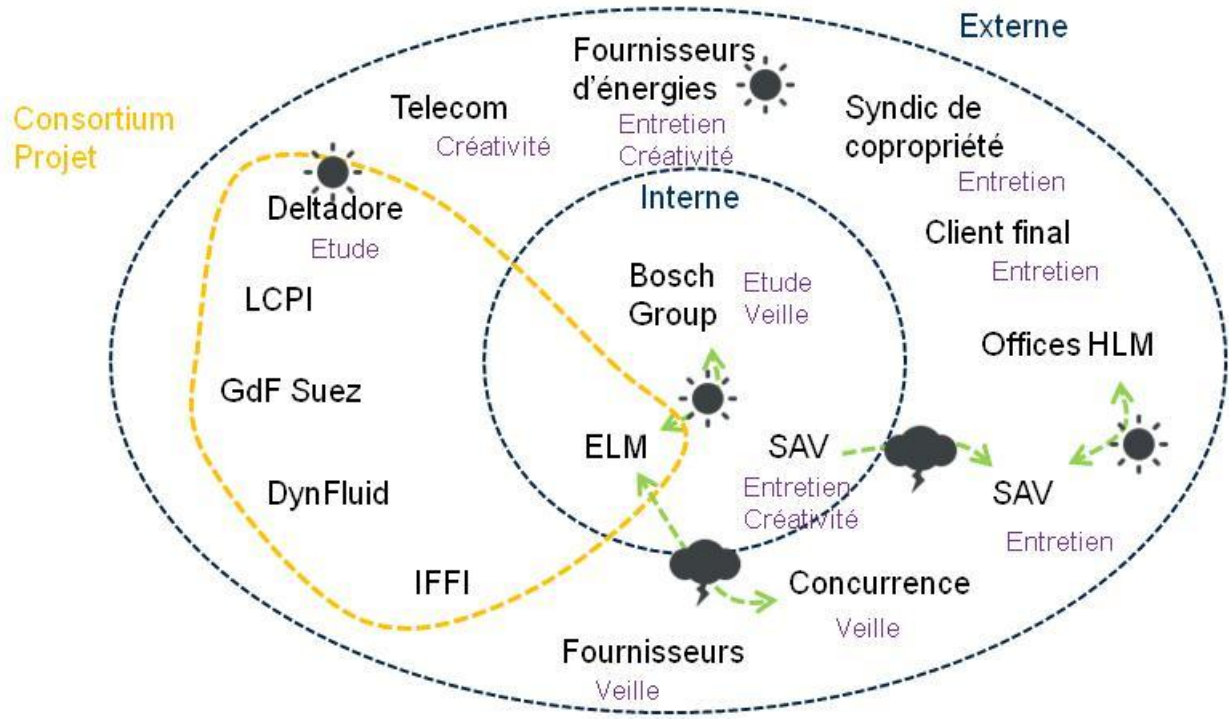


Figure 27 : Carte des parties prenantes réalisée lors de l'expérimentation 1

Le fait de dresser une première représentation de l'écosystème dans lequel notre système peut évoluer nous permet alors une première projection de celui-ci afin de servir une activité ou un métier. Nous

pouvons déjà à ce stade voir apparaître un périmètre fonctionnel se dessiner autour des énergies et de la maintenance. De plus, et plus axé vers un point de vue projet, cet outil nous permet de cartographier l'ensemble des investigations à mener pour évaluer les besoins des utilisateurs. Nous avons étayé notre représentation sur la figure 27 avec les actions réalisées auprès de chacun des acteurs. Nous y retrouvons de nombreuses opérations de veille technologique grâce auxquelles nous avons pu voir émerger des tendances fonctionnelles telles que le pilotage à distance ou la maîtrise et l'optimisation des consommations. Cette veille a également permis d'identifier de nombreuses entreprises pouvant intervenir au sein du projet en tant que partenaire pour développer ou intégrer certaines plateformes du système. Nous retrouvons également dans les actions d'investigation auprès des acteurs du service de nombreux entretiens, conduits principalement auprès des potentiels clients du système. Ces entretiens nous ont permis de sonder les différents acteurs sur leurs usages de la connectivité, sur leur perception de ces technologies et sur leurs espérances vis-à-vis des fonctionnalités.

D'autre part, cela nous a permis de dégager au sein de ce paysage un ensemble de risques et d'opportunités, représentées sur la figure 27 par des soleils ou des nuages. Nous retrouvons par exemple de fortes opportunités avec les fournisseurs d'énergie qui investissent énormément dans le développement de nouveaux services basés sur des compteurs intelligents, pouvant à terme se manifester par un projet de co-développement. Des risques ont également été mis en évidence, comme par exemple avec les relations entre notre service après-vente constructeur et les entreprises de maintenance indépendantes qui, de par leur concurrence, peuvent se voir réticentes au partage de certaines informations propres à leur métier. Ces risques et opportunités mis à jour se sont avérés être des éléments décisionnels clés dans la sélection des concepts et la mise en place d'une architecture du système.

Enfin, et dans le but de capitaliser les efforts d'investigation auprès des potentiels clients du système, nous avons résumé les données collectées au travers de personas. Cinq personas représentant les clients finaux, tel que l'expose la figure 28, ont été réalisés sur la base de profils sociologiques, comportementaux et technologiques qui ont émergé pendant les phases d'entretiens. De la même manière, nous avons également réalisé des personas des mainteneurs afin d'illustrer leur segmentation avec, de la même manière, une représentation des attributs qui les représentent. Ces personas nous ont permis de condenser et de diffuser les principales informations que nous avons pu collecter dans nos phases d'investigation auprès de l'équipe projet. Ce qui nous a permis d'améliorer la phase de génération de concept que nous sommes désormais prêts à aborder.



Figure 28 : Ensemble des personas générées suite à la phase 1. (Un exemple de personas est présenté en Annexe E)

Phase 2 : Génération de concepts

Une fois le périmètre clarifié, nous devons définir et expliciter les fonctionnalités attendues de notre système lors de la phase de génération de concepts. Cette phase nous amène donc, par une approche créative, à définir et caractériser les axes fonctionnels principaux de notre système. Nous nous reposons alors ici sur une succession de séances de créativité avec l'ensemble de l'équipe projet, accompagnés parfois de certains représentants des utilisateurs potentiels lorsque cela était possible. Ces séances ont alors été nourries par nos investigations précédentes qui ont permis d'axer nos réflexions autours des sujets fédérateurs.

La phase de génération d'idées s'est organisée autour de 5 jours d'ateliers créatifs, pendant lesquels nous avons focalisé notre attention sur les sept sujets suivants :

- La maintenance préventive et curative
- La gestion de parc
- Le conseil
- L'installation
- Le pilotage
- La capitalisation de connaissances
- La régulation intelligente

Ces séances ont étés menées au sein d'une équipe pluridisciplinaire comprenant :

- Un animateur externe disposant d'une expertise orientée vers les systèmes d'information
- Le responsable innovation d'e.l.m. leblanc
- Le responsable informatique d'e.l.m. leblanc
- Le responsable marketing produit d'e.l.m. leblanc
- Deux experts du service en cours de développement

D'autre part, et en fonction des sujets abordés, nous avons également intégré des profils utilisateurs afin de les faire réagir dans ces ateliers créatifs. Nous y retrouvons particulièrement :

- Deux mainteneurs pour le sujet de la maintenance
- Un énergéticien pour le sujet du conseil
- Deux utilisateurs pour traiter du sujet du pilotage

Les idées générées pendant ces cinq jours d'ateliers créatifs ont étés mises sous la formes d'*user stories*. Ainsi, une idée générée est formulée de la manière suivante :

> *En tant que (client cible), je veux (fonction), dans le but de (objectif).*

À titre d'exemple, nous pouvons citer la situation suivante :

> *En tant que service après-vente, je veux pouvoir justifier à mon client les opérations de maintenance préventive à réaliser, dans le but de le rassurer.*

Cette formulation présente de nombreux avantages. Le premier étant de situer de manière systématique l'utilisateur ou l'acteur concerné par l'idée dans le premier terme. D'une part, cela nous permet d'organiser les idées autour des utilisateurs qu'elle cible. Mais, d'autre part, cela permet aussi à l'auteur

de l'idée de faire preuve d'empathie lors de la génération. Cela montre également les synergies de cette formulation avec les personas définies dans la phase précédente pouvant être la cible de la fonction. Ensuite, alors que le second terme de la *story* exprime le contenu de la fonction, le dernier la justifie en décrivant l'intérêt qu'elle propose au client cible. Cette justification est importante pour l'auteur qui, au moment de rédiger son idée, fait une nouvelle fois preuve d'empathie pour évaluer l'impact de cette idée dans l'environnement de l'utilisateur. Sur le moment, cela permet d'évaluer très rapidement la pertinence de l'idée. Autrement dit, si l'auteur n'est pas en mesure de trouver une raison à cette idée, c'est qu'elle n'a peut-être pas lieu d'être. D'autre part, cette justification permet, dans la suite du projet et en le confrontant avec les investigations terrain que nous avons menées, de mesurer la priorité des idées les unes par rapport aux autres.

A la fin de cette phase créative, nous obtenons un total de 550 idées organisées en 61 sous-parties. Ces dernières mettent en scène un total de 15 acteurs de services différents. Nous pouvons alors noter que ces idées sont nombreuses, mais définissent le besoin de manières très différentes, avec des niveaux de granularité variables dans la description des fonctionnalités. D'autre part, certaines idées sont orientées vers l'usage et l'aspect purement fonctionnel du système, alors que d'autres pointent davantage les caractéristiques techniques auxquelles il doit répondre. Afin d'illustrer notre propos, nous pouvons donner à titre d'exemple la paire de *User Stories* suivante concernant le service après-vente dans le thème de la maintenance curative, où la première reste fonctionnelle et générique et où la seconde est beaucoup plus précise et pointe vers des solutions techniques.

> ***En tant que*** *service après-vente,*
> ***je veux*** *pouvoir avoir une remontée des pannes sur le terrain en temps réel,*
> ***dans le but de*** *mieux anticiper mes interventions de maintenance.*

> ***En tant que*** *service après-vente,*
> ***je veux*** *que les problèmes détectés par les chaudières soient hiérarchisés et que seuls les plus importants et les plus graves me soient communiqués directement par SMS ou par courriel. Les autres doivent être visibles sur mon tableau de bord,*
> ***dans le but de*** *garantir l'efficacité de mon service et de limiter le nombre de sollicitations directes.*

Après cette phase de divergence sur le besoin des utilisateurs, nous nous sommes chargés d'extraire les principaux concepts du service. Nous avons alors convergé suivant deux axes. Un axe purement fonctionnel dans lequel nous nous sommes attachés à définir notre service tel qu'il devrait être d'un point de vue client. Puis, un second axe technologique où nous nous sommes chargés de construire la première esquisse architecturale du système technique.

Afin de nous représenter le besoin, nous avons formalisé les concepts principaux de notre service au travers des cartes concepts. Ces cartes reprennent les idées principales du service grâce, d'abord, à un titre explicite. Puis nous y ajoutons les détails caractérisant ses fonctions, de par les fonctionnalités attendues et par les perceptions qualitatives et financières attendues des clients. Nous y ajoutons également des potentiels risques qui ont pu être mis à jour par la phase de clarification ou au cours des discussions pendant la phase de génération d'idées. La figure 29 donne à voir un exemple de l'une de ces fiches concepts représentant le concept de restitution et de visualisation des données du système pour le client. Un extrait de ces fiches concepts est présenté plus en détail dans l'Annexe F.



Figure 29 : Exemple de fiche concept générée dans la phase 2 de l'expérimentation 1

Sur la base des 550 idées générées, nous avons pendant cette phase isolé 300 idées uniques, illustrées autour de 33 fiches concepts réparties principalement autour des usages des services après-vente et des utilisateurs finaux des produits. Une fois formalisés, ces concepts ont pu être évalués par un panel d'utilisateurs mixte, comprenant des experts du domaine ainsi que des novices. Lors de cette évaluation, nous avons demandé à notre panel de s'identifier à l'un des personas que nous avions générés dans la phase de clarification puis, selon son profil, de positionner les cartes concept les unes par rapport aux autres dans un tableau, suivant une échelle de pertinence. D'autre part, nous avons également réalisé le même travail avec l'équipe projet, et ce, sans l'utilisation des personas, afin de juger la pertinence des concepts avec une vision avérée du domaine d'application et du marché. Ainsi, nous avons pu construire des profils de notes, comme l'illustre l'exemple de la figure 30, nous permettant de sélectionner et de prioriser les concepts les plus pertinents pour la suite de nos travaux.

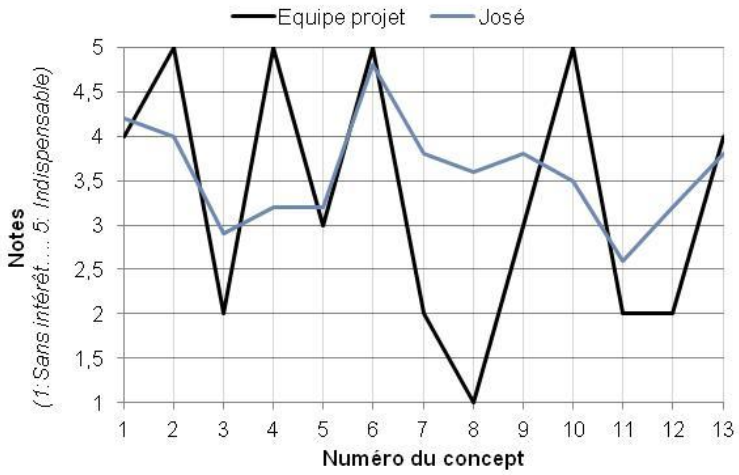


Figure 30 : Profil de notes issu de l'évaluation des concepts en phase 2 de l'expérimentation 1

Parmi les concepts principaux retenus suite à l'évaluation, nous retenons les suivants :

Pour le client final,

- La modification et la création de programmes horaire
- La régulation en fonction de conditions extérieures

- L'optimisation du paramétrage du système
- Le contrôle à distance de l'appareil
- La visualisation des données du système

Pour le service après-vente,

- Le pré-diagnostic des interventions à distance
- Le suivi des paramètres d'état de l'appareil
- La possibilité d'interventions préventives
- La communication client et l'implication du client dans la maintenance de son appareil
- La gestion des pièces détachées
- Les actions à distance du service après-vente

Et pour e.l.m. leblanc,

- Le suivi des essais terrain
- La création d'indicateurs qualité
- La capitalisation en données marché

En définitive, et pour clore cette phase de génération de concepts, nous avons pu mener nos premières réflexions sur la construction d'une architecture du système technique sur la base des idées à caractère technologique et sur les informations. Le fruit de cette réflexion a alors pu être formalisé par le biais d'une carte système. Cet exercice, dont le premier résultat est exposé avec la figure 31, met en lumière plusieurs éléments. Tout d'abord, nous retrouvons l'architecture en forme de plateforme où le flux d'information transite du produit vers l'utilisateur en passant par un ensemble de blocs fonctionnels. Puis, ces blocs sont divisés en deux parties. La partie locale, représentant le domaine physique, puis la partie distante avec le traitement et la restitution de l'information. Enfin, nous pouvons grâce à ce modèle nous représenter l'ensemble des connaissances à capitaliser afin de déployer notre service. Nous pouvons notamment souligner la mise en évidence d'un besoin de créer un contrôleur externe au produit et la nécessité de clarifier et définir une interface de collecte de données entre la couche distante de notre système et notre produit.

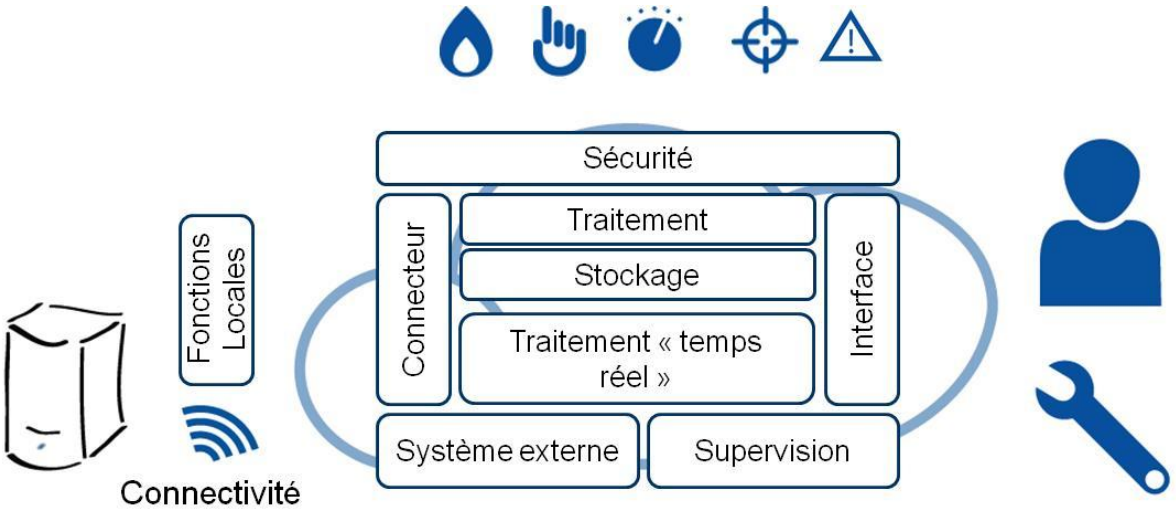


Figure 31 : Carte système réalisée en phase 2 de l'expérimentation 1

Phase 3 : Conception générale

Finalement, lorsque le périmètre fonctionnel du système est enfin défini et que l'architecture technique est esquissée, il nous faut faire émerger les caractéristiques particulières de notre système sous la forme de spécifications lors de la phase de conception générale. Le but de cette phase est de matérialiser le service dans son ensemble afin de pouvoir discuter l'impact de ce nouveau service dans l'activité quotidienne des clients, d'extraire les caractéristiques nécessaires de notre système technique afin de le rendre compatible avec la fourniture du service que nous avons imaginé, et également de pouvoir représenter le service d'une manière suffisamment précise pour pouvoir traduire l'image que nous nous en faisons en spécifications techniques pour débuter son développement. Nous faisons alors, une fois de plus, face à un besoin de formalisation et de structure de nos concepts afin de les rendre plus explicites.

La première étape de formalisation passe par la mise en perspective de nos concepts vis-à-vis de l'activité de l'utilisateur, en les détaillant sous la forme d'activités basées sur les fonctionnalités identifiées pour chaque concept et en les séquençant les uns par rapport aux autres. Cet effort peut être réalisé à l'aide du parcours utilisateur client. À titre d'exemple, nous pouvons représenter le parcours utilisateur de la première fonction concernant l'utilisateur final du produit. La figure 32 met en scène un de nos personas à travers le concept, dont nous avons présenté la fiche précédemment, par une séquence d'actions menant l'utilisateur à l'accès aux fonctions. Cette méthode projette également ce concept vers une suite logique d'actions représentée par d'autres concepts. Nous pouvons noter ici le fait que les fonctions de ces concepts se basent sur un usage préalable du produit, que l'utilisateur doit s'identifier sur une plateforme. Les fonctions se matérialisent par la visualisation des consommations afin de mieux les comprendre. Par la suite, ces fonctions nous amènent vers les concepts d'optimisation et de paramétrage de l'installation.

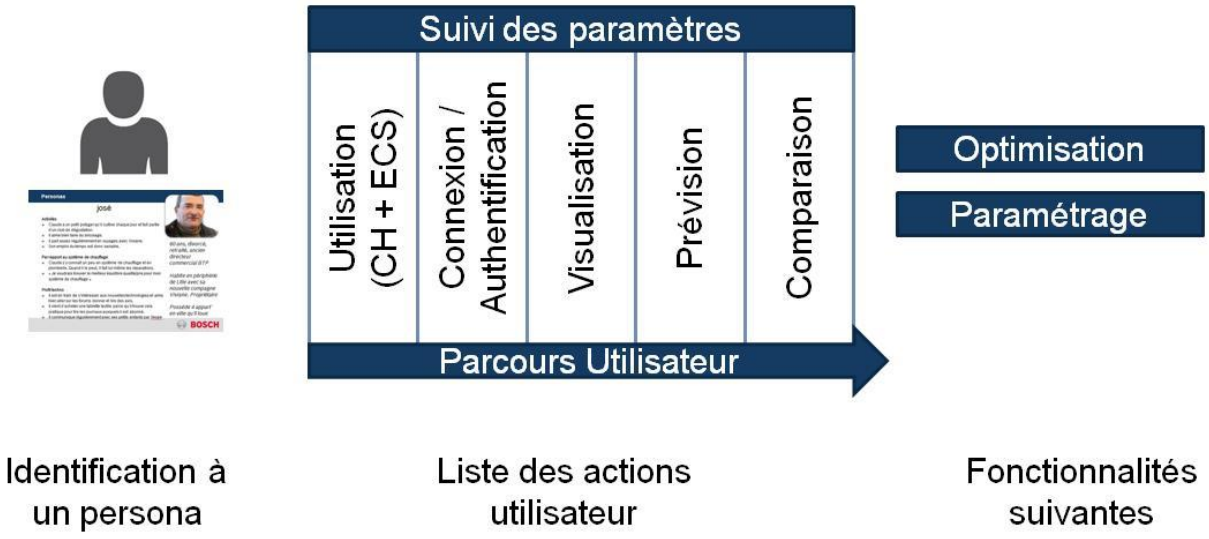


Figure 32 : Parcours utilisateur réalisé en phase 3 de l'expérimentation 1

Afin de finaliser la matérialisation de notre service et à l'aide de l'ensemble des informations que nous avons préalablement collectées et structurées, nous pouvons nous lancer dans la création de notre *service blueprint*. L'étape précédente de structuration des concepts au sein des parcours utilisateurs nous a permis d'étirer nos concepts dans une dimension temporelle. Nous nous attachons alors avec cet outil à continuer ce travail avec une granularité de détail plus importante. Mais le principe de cet outil vient de sa vocation à étirer les concepts sur un axe des acteurs du service où le principal intérêt réside dans la représentation des interactions qui les lient. Nous pouvons alors dissocier les clients ou

utilisateurs du service du fournisseur afin de mettre en avant les points de contact entre ceux-ci. Cela constituant le principal vecteur de qualité perçu par les clients, la force de cet outil réside principalement dans la faculté à manipuler ces points de contact dans le cadre de la conception de ce service. D'autre part, cet outil fait également apparaître l'ensemble des mécaniques engagées dans la fourniture de ce service. Cela passe par l'implication des acteurs extérieurs à la fourniture même du service, mais participant activement à la production de valeur restituée à l'utilisateur. Cela permet notamment de planifier et de mesurer l'effort humain et la connaissance à capitaliser afin de mettre en œuvre le service. Enfin, le dernier niveau de détail de cet axe apparaît avec les technologies support à mettre en œuvre pour réaliser chacune des phases du service. Ces technologies se retrouvent à deux niveaux. Au niveau client avec la matérialisation du service, autrement dit avec la définition des artefacts qui seront vecteurs de la transmission de la valeur issue de la fonction du service. Puis, au niveau du fournisseur avec la définition des éléments techniques support aux agents du service pour réaliser leurs tâches. Cela pouvant être une base de données, un logiciel, une procédure ou encore de l'outillage.

Nous représentons dans la figure 33 un extrait du *blueprint* pour le cas que nous avons abordé précédemment avec la visualisation des paramètres client. L'ensemble des *blueprints* sont développés plus en détail dans l'Annexe G pour le client final et dans l'Annexe H pour le service après vente. Notre service s'intègre ici dans une relation existante entre les deux clients principaux, les utilisateurs et leur service après-vente, qui représentent nos deux niveaux d'utilisateurs. Pour chacun d'entre eux, nous détaillons les voix de matérialisation du service. En dessous, nous détaillons le rôle du fournisseur de service, séparé en deux par la ligne de visibilité client, le *front office* et le *back office*. Enfin, nous détaillons les supports techniques nécessaires à la fourniture des fonctions, où nous retrouvons l'ensemble des capteurs et données nécessaires, les systèmes d'information avec leurs processus ainsi que les algorithmes de traitement devant être sollicités.

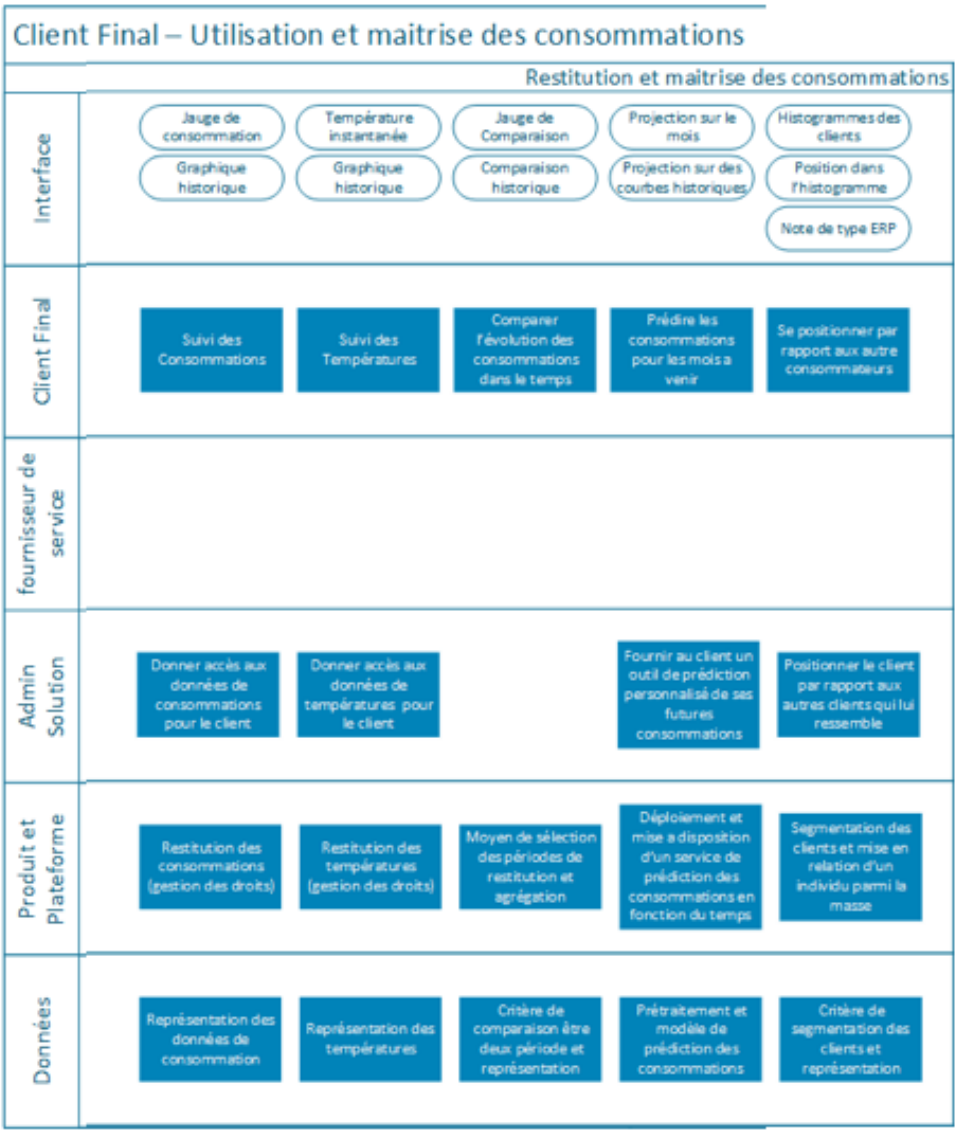


Figure 33 : Extrait du service blueprint pour la fonction de suivi des consommations pour le client final

La formalisation de ce service sous forme de *blueprint* nous permet d'alimenter de nombreuses discussions et de défendre nos choix technologiques auprès de fournisseurs et d'experts. Effectivement, nous pouvons identifier et positionner les caractéristiques techniques de notre système relativement aux tâches et aux fonctions du service. Nous avons alors pu dissocier et spécifier ce qui avait trait à la réalisation d'un système de capteurs et actionneurs pour la collecte de l'information, ce qui avait trait à la réalisation d'un système d'information pour la manipulation et la restitution de celle-ci, et ce qui avait trait à la réalisation d'algorithmes de traitement pour exploiter la donnée afin d'en extraire l'information nécessaire à l'utilisateur.

De ce fait, les discussions portées par ce mode de représentation auprès des futurs utilisateurs et des fournisseurs ont permis de faire émerger les spécifications générales de notre système, les points de forte valeur ajoutée ainsi que les priorités à porter au développement. Dans les priorités, nous avons alors pu identifier la réalisation du socle matériel comme étant le plus urgent afin de disposer de notre réseau de capteurs et d'actionneurs pour supporter la suite des développements. La rédaction de cahiers des charges pour chacun des domaines techniques a donc pu être réalisée dans cette optique. Le modèle de

service nous a également facilement conduits à des prototypes d'interfaces de la plateforme, où la navigation et la hiérarchisation de l'information ont été basées sur les détails suffisants fournis ici. L'exemple de notre concept de visualisation des paramètres pour le client y est alors matérialisé dans la figure suivante.

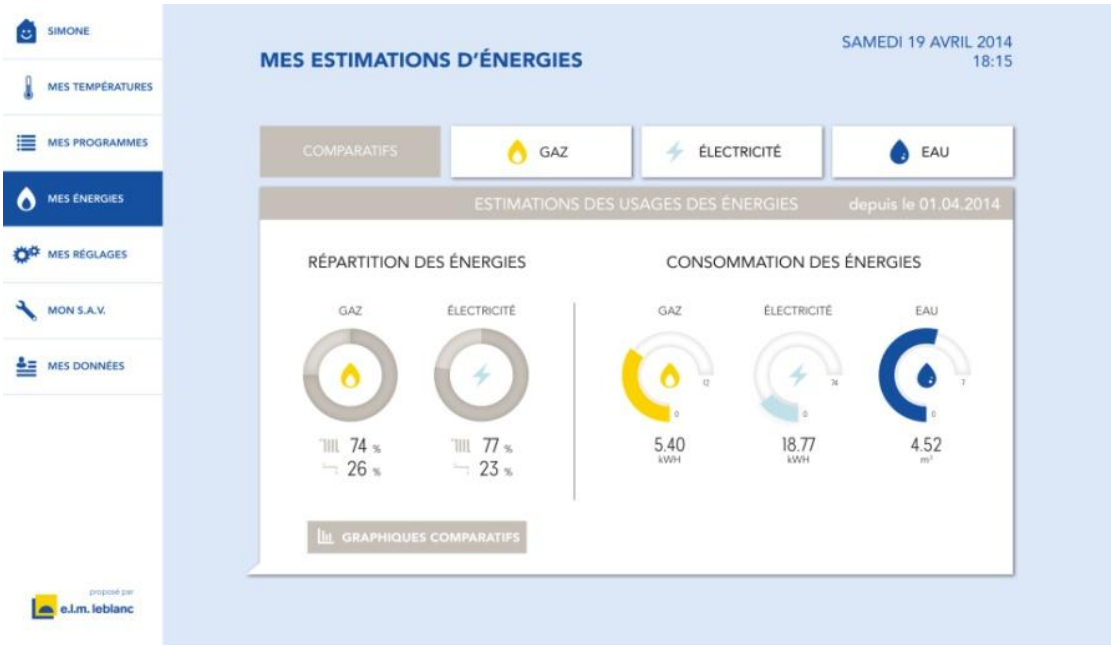


Figure 34 : Prototype d'interface pour la fonction de suivi des paramètres client

La constitution de prototypes d'interfaces s'est révélée primordiale dans la perception de l'acceptation client et dans la réalisation d'un cahier des charges. Conjointement avec l'ensemble des fournisseurs, nous avons alors pu bâtir un modèle de rentabilité. La mise en perspective des gains estimés du service avec les coûts de développement et d'exploitation du service nous a permis d'une part d'ajuster les choix technologiques, mais également d'argumenter la pertinence du projet dans la stratégie générale de l'entreprise afin de d'obtenir l'autorisation de lancer les développements.

4.1.4. Analyse des résultats et conclusion

Cette première expérimentation avait pour vocation de démontrer la pertinence d'une approche orientée service dans l'attaque des phases amont de la conception, et particulièrement dans notre cadre, pour la conception d'un système d'exploitation d'objets connectés. Effectivement, notre volonté d'innover, poussée par la technologie, nous force à concentrer nos efforts dans ces premières phases de conception sur la découverte du client et de ses besoins ainsi que sur la maîtrise des nombreuses connaissances technologiques que notre système nous amène à manipuler. Cela constituant notre première hypothèse de résolution, nous présenterons ici les résultats que nous avons obtenus suite à l'application de la démarche que nous proposons, nous discuterons des intérêts et des limites de cette démarche avant de conclure sur la validation de notre première hypothèse.

Résultats

Suite à la conduite de cette expérimentation, nous disposions d'une modélisation du service au travers du *blueprint* réalisé, ainsi que d'une analyse de la rentabilité du projet. Ces deux éléments ont été le support de discussion avec les partenaires industriels pour le choix des solutions technologiques pour chaque couche de notre système. Cela nous a amené à améliorer la vision que nous pouvions avoir de notre système, tel que l'expose la figure 35. De ce fait, nous avons également pu raffiner les hypothèses

de notre modèle de rentabilité, ce qui nous a alors permis d'argumenter pour officialiser le développement de ce projet et de débloquer les investissements nécessaires.

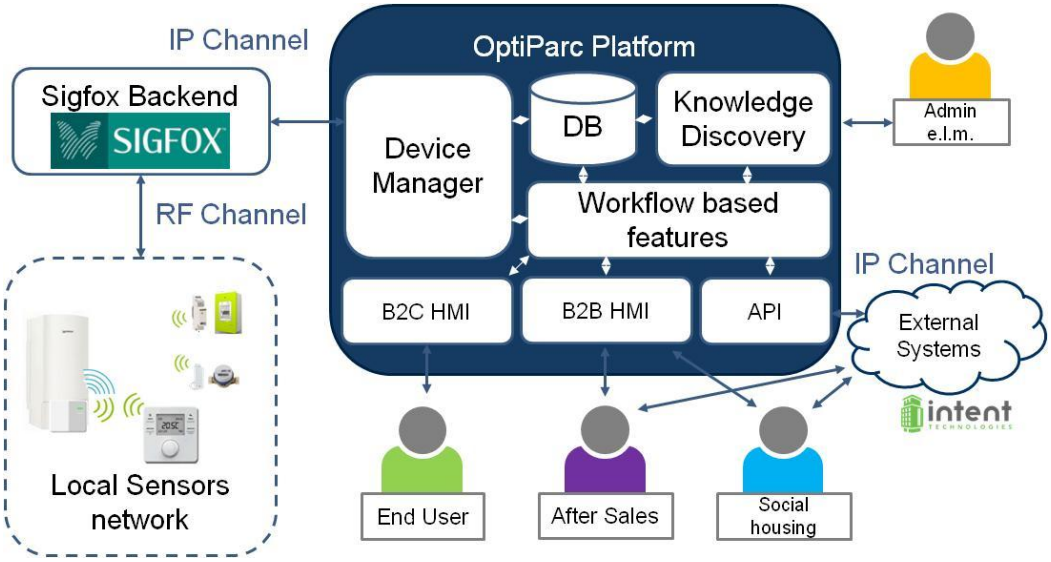


Figure 35 : Carte système de la solution Optibox

Nous avons alors pu initier le développement de notre système par les éléments identifiés comme prioritaires avec la partie physique de la plateforme. Nous avons alors développé le thermostat d'ambiance illustré par la figure 36 dont les caractéristiques sont en accord avec nos besoins pour la réalisation des fonctions connectées. Nous y retrouvons notamment :

- La mesure des consommations et sa remontée au dient pour répondre au concept de suivi des paramètres du système
- Le contrôle de la température ambiante et de l'eau chaude au travers des programmes horaire pour les besoins de notre service en termes de pilotage
- Un algorithme de régulation tenant compte des besoins en chauffage de l'utilisateur et des conditions climatiques afin de permettre l'optimisation des consommations



Figure 36 : Thermostat d'ambiance développé comme support physique de la connectivité dans l'offre Optibox

D'autre part, en parallèle du développement de ce thermostat d'ambiance et sur la base de notre modèle de service, nous avons entamé la conception de la partie logicielle en travaillant sur les interfaces dédiées à nos utilisateurs, les clients finaux des produits, ainsi que les mainteneurs. Exposées dans la figure 37, ces interfaces constituent notre premier prototype réel de solution avec lequel nous avons pu approcher les clients afin de leurs présenter nos propositions de service et capter leurs motivations. Ce travail a permis d'avancer une fois de plus dans la matérialisation du service et dans la description détaillée de nos attentes pour les développeurs, dans le but de réaliser la plateforme logicielle de notre système.

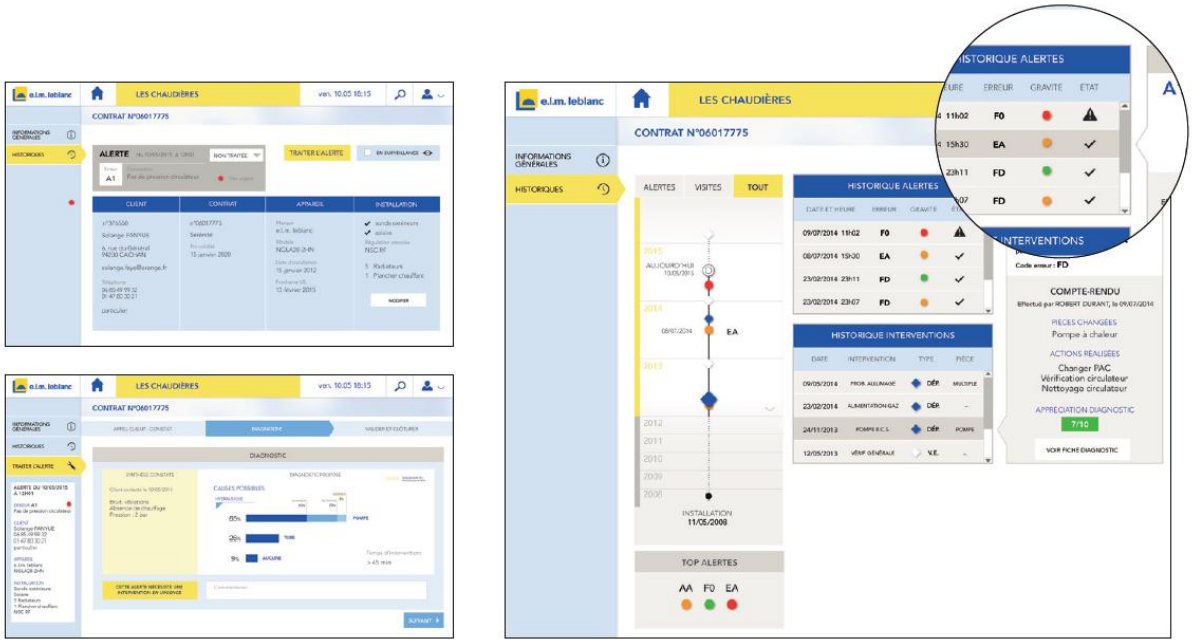


Figure 37 : Prototype d'interface pour les fonctionnalités des services après vente

Afin de conclure sur ces résultats, nous pouvons observer que l'approche que nous avons mise en place ici afin de formaliser le périmètre fonctionnel du système a simplifié la suite du déroulement du projet. Cela nous a permis de construire un cahier des charges formel de notre système sur la base des besoins de nos clients. Cependant, nous nous attendions à retrouver de nombreuses dépendances entre la conception du produit physique et celle de la plateforme informatique. Or, le niveau de détail de notre modèle de service était alors suffisamment précis à ce stade pour pouvoir mener de front le développement des deux parties du système.

Intérêts et limites

Par ces résultats nous pouvons voir que grâce à cette approche nous avons pu avancer dans le développement de notre système de service. Cependant, nous pouvons également discuter ici des intérêts et des limites que présente la démarche que nous avons proposée et suivie.

Tout d'abord concernant les avantages, nous avons pu observer que le déroulement de cette démarche s'est fait de manière très douce avec une avancée progressive vers un modèle de service consistant. Effectivement, cette démarche est basée sur une sélection d'outils que nous avions réalisée à la suite de notre état de l'art, puisque les outils ont été sélectionnés pour leurs synergies dans une séquence logique de conception. Cela se concrétise par une capitalisation de connaissances maximale dans le passage d'un outil à un autre et une minimisation des efforts entre chaque étape. Nous pouvons tout de même noter que les efforts de synthèse et de formalisation sont importants. Nous les retrouvons principalement dans

la réalisation des cartes concepts ou du *blueprint* où la quantité d'informations à traiter est importante et où nous pouvons retrouver des idées contradictoires à équilibrer.

Ensuite, nous avons pu noter la force du *blueprint* dans la séparation entre les acteurs et le système technique. Cela a constitué un atout primordial dans la communication que nous avons pu faire sur nos intentions, et particulièrement dans l'implication des différents partenaires. Cela à permis de localiser l'impact d'un choix technologique sur un besoin utilisateur clairement exposé et scénarisé.

Enfin, cette mise en perspective de ces technologies avec de nombreux usages et de nombreux utilisateurs potentiels très en amont a permis, même si ces utilisateurs n'ont pas été considérés comme prioritaires, d'estimer leurs besoins dans une solution globale et d'établir les parallèles avec des fonctionnalités prioritaires. Cela nous a offert de nombreuses opportunités de partenariats par la suite, lorsque les premiers prototypes fonctionnels furent disponibles, avec un discours éprouvé et des voies d'intégration déjà imaginées. Nous pouvons ici prendre l'exemple d'échanges réalisés avec plusieurs énergéticiens, que nous avions écartés de notre périmètre fonctionnel car les retombées de leurs cas d'usages n'étaient pas assez concrètes pour l'entreprise. Or, ces échanges se tenant alors que les développements de la plateforme étaient bien avancés, nous avons su nous positionner très rapidement dans leur vision de fourniture de service au travers des systèmes d'information et de leurs activités relatives aux compteurs connectés. Cela nous a très rapidement menés à la réalisation de tests dans leurs laboratoires et cela pourrait, dans un futur proche, se concrétiser par le lancement d'un projet collaboratif européen.

Conclusion

Suite à cela, et au regard de l'ensemble de ces résultats, nous pouvons conclure à la validation de notre première hypothèse de recherche. Le déroulé de cette expérimentation montre comment une approche service peut être moteur dans la conduite d'un projet porté par les flux d'information entre les produits et leurs utilisateurs. La démarche que nous avons mise en place a montré son efficacité pour avancer dans le projet et guider le concepteur vers un modèle de service pertinent dans l'optique d'en assurer sa réalisation. Puis, l'ensemble des efforts déployés en amont pour définir ce service ont pu être valorisés, à la fois pendant cette conception amont, mais également par la suite, lors du développement, de sa mise en place et de la recherche de nouveaux partenaires et utilisateurs pour étendre le périmètre fonctionnel du système.

Nous avons ici fait en sorte que cette approche nous mène à la conduite d'un développement agile, recommandé afin de négocier au mieux les différentes dépendances que nous pouvons noter entre la réalisation du produit physique et celle de la plateforme logicielle associée. Cependant, nous avons fait face à de nombreuses incertitudes relatives à la réalisation des fonctions intelligentes, telles que l'amélioration de la maintenance ou l'optimisation des consommations. Effectivement, la majorité des fonctions du système font appel à un flux d'information entre produit et utilisateur ou conduisent à la structuration d'échanges entre utilisateurs, contrairement aux fonctions intelligentes qui reposent principalement sur un historique sur lequel on souhaite capitaliser. C'est au travers de cette observation que nous avons construit notre seconde hypothèse, et donc conduit les prochaines expérimentations que nous présenterons par la suite.

4.2. Expérimentation 2 : Conception intégrée des fonctions intelligentes

Suite à la présentation de la première expérimentation, qui nous a conduits à la validation de notre première hypothèse, nous présenterons ici la conduite de la seconde expérimentation. Celle-ci a pour vocation de tester la seconde hypothèse de recherche concernant la nécessité d'aborder une approche de conception intégrée dans la conception détaillée du service afin de prendre en compte au mieux le caractère particulier des fonctionnalités intelligentes dans le système.

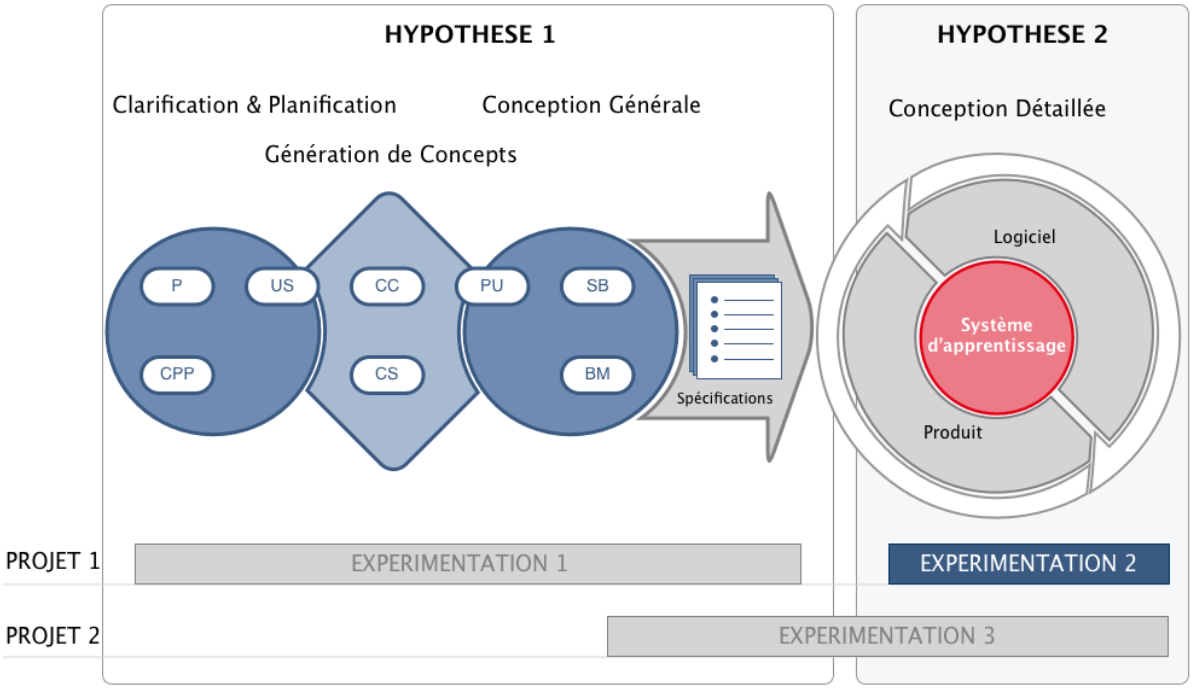


Figure 38 : Positionnement de l'expérimentation 2

Pour cette seconde expérimentation, nous nous positionnons à la suite de la première. À savoir que nous intervenons ici sur le même projet de développement d'une plateforme de chaudières connectées pour lequel nous disposons d'un cahier des charges explicite des fonctions du service que nous souhaitons réaliser ainsi que des premiers prototypes physiques et logiciels. Dans le cadre de cette conception détaillée itérative, nous portons ici notre intérêt sur l'introduction des fonctionnalités relatives à l'anticipation de la maintenance. Ces fonctionnalités disposent d'un caractère intelligent dans le sens où nous devons extraire des données brutes de nos chaudières les attributs nécessaires à l'identification et à la prédiction d'une défaillance.

Afin de présenter au mieux les tenants et aboutissants de cette nouvelle phase expérimentale, nous détaillerons dans un premier temps les objectifs. Nous définirons ensuite notre démarche, dont nous exposerons le contexte ainsi que le déroulement avant de conclure sur les résultats de cette expérimentation au regard de notre problématique de recherche initiale.

4.2.1. Objectifs

Dans cette expérimentation, nous nous situons dans les phases de conception détaillée et de réalisation de notre système. Lors de l'expression de notre problématique de recherche, nous sommes partis du postulat que le développement des fonctions intelligentes, basé sur la science des données, était dépendant de la conception détaillée de la plateforme physique et logicielle. Ceci constitue le fondement de notre hypothèse 2. Afin de la vérifier, nous proposons alors d'aborder une approche agile de la conception détaillée focalisée sur les systèmes de découverte de connaissance au sein de laquelle nous intégrons l'ensemble des métiers impliqués.

L'objectif de cette expérimentation est donc de mettre en pratique cette proposition dans un cadre industriel afin de vérifier la véracité de notre seconde hypothèse de recherche. Nous nous attarderons particulièrement sur les choix de conception concernant la réalisation de notre système physique que nous amènent à faire cette approche, nous permettant par la suite de valider la présence de dépendance entre les systèmes de découverte de connaissance et le reste de notre système.

D'autre part, et en raison du cadre particulier de la mise en application des méthodologies issues de la science des données intervenant ici très en amont de la conception, nous observerons l'intérêt d'opérer une démarche intégrant l'ensemble des métiers relatifs aux fonctionnalités du système que nous souhaitons mettre en place. Nous souhaitons ici modéliser la dynamique insufflée par l'ensemble des parties au cours de la conception des fonctionnalités intelligentes. Nous souhaitons identifier quels peuvent être les apports des différents domaines de connaissance dans le cadre d'un projet de conception de système intelligent. Qu'il s'agisse de dynamiques d'ouverture du problème et de convergence vers des solutions techniques ou de dynamiques de blocage. Et cela, en localisant ces phénomènes dans la séquence d'étapes du processus méthodologique que nous proposons.

4.2.2. Démarche

Afin d'atteindre ces objectifs, nous mettons en place une démarche de recherche et d'analyse. D'une part, nous mettrons ici en application l'approche méthodologique que nous avons identifiée pour la réalisation de système de découverte de connaissance (cf. paragraphe 2.4.2) au sein duquel nous solliciterons les corps de métier essentiels à la bonne conduite du projet. Ce processus itératif et incrémental prendra place dans le cadre industriel du développement de produits communicants que nous détaillerons par la suite. Nous baserons alors nos observations sur les interactions entre le développement des fonctions intelligentes et du produit physique afin de conclure sur les dépendances entre les technologies issues de la science des données et la conception du reste des composants du système.

D'autre part, et afin de percevoir la dynamique collaborative entre les différents métiers impliqués dans la conception, nous mettrons ici en application ce processus que nous avons identifié en tenant compte des apports que peuvent fournir les domaines d'expertise, qu'il s'agisse d'experts issus de la science des données, du service ou du domaine d'application, ce qui est mis en évidence par la figure 39.

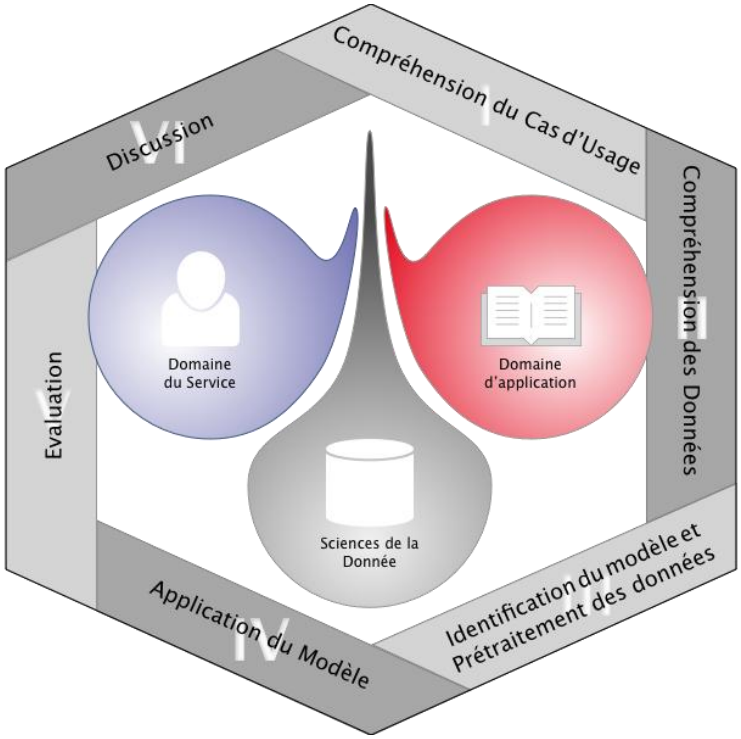


Figure 39 : Démarche de conception pour l'expérimentation 2

Dans le cadre de cette phase de conception et afin de formaliser cette collaboration entre les domaines d'expertise, nous suivons une grille de lecture bidimensionnelle où nous projetons nos observations des dynamiques collaboratives du projet. Ces dernières s'organisent suivant un axe temporel, représenté par la séquence des étapes de la méthodologie ; puis suivant un axe des connaissances, en identifiant le domaine moteur de l'évènement. Nous identifions alors ces observations par leur valeur vis-à-vis du projet : soit par l'apport d'une dynamique positive, en nous amenant vers une solution ou vers de nouvelles applications ; soit par une dynamique négative, apportant de nouveaux problèmes. Puis, afin de compléter cette vision, nous ferons référence aux supports ou aux méthodes mis en œuvre pour instaurer cette dynamique. En dernier lieu, et de manière plus globale, nous porterons un regard critique sur la concrétisation de cette conception dans le projet.

4.2.3. Cadre d'application et contexte

Avant de présenter la mise en application de cette démarche, nous souhaitons contextualiser le cadre dans lequel cette expérimentation prend place. Nous nous situons ici dans la continuité de notre première expérimentation. Néanmoins, et entre temps, nous avons pu commencer la réalisation du produit physique, constituant notre réseau de capteurs et d'actionneurs prenant la forme dans notre cas d'un thermostat d'ambiance. Notre première expérimentation a pointé du doigt l'ensemble des fonctionnalités devant être réalisées par ce produit, avec notamment le pilotage de la chaudière sur la température ambiante et la température extérieure, la programmation horaire, la mesure de la consommation, mais également par la collecte de données d'état de la chaudière pour améliorer la maintenance et l'intégration d'un module communicant pour échanger des données avec une plateforme de traitement.

D'autre part, nous disposons à ce stade d'une liste exhaustive de spécifications pour mener à bien le développement de la plateforme : le *product backlog*. Parmi l'ensemble de ces objets de spécifications, nous retrouvons les modes d'exploitation des différentes données devant être collectées par le produit comme les moyens de restitution des consommations pour le client final. Néanmoins, nous retrouvons

également des fonctions beaucoup plus abstraites à ce stade, souvent liées au caractère intelligent du système, comme le diagnostic et l'identification des pannes pour l'aide à la maintenance ou l'optimisation des réglages. Ces fonctions, dont la réalisation reste encore à définir ne nous permettent pas de mesurer leur impact sur l'ensemble du système. Le plus gros risque du projet réside dans le fait de ne pas pouvoir définir l'ensemble des processus de collecte de données permettant d'assurer la faisabilité de ces fonctions et qui seront implémentés dans le produit. De plus, et grâce à l'aide du modèle de rentabilité de notre système, nous avons identifié que le principal centre de coûts de notre système résidait dans la quantité d'informations échangées dans le réseau d'objet.

C'est alors que nous pouvons nous interroger sur le meilleur moyen de remonter l'ensemble des données permettant la réalisation des fonctions intelligentes que nous avons préalablement identifiées. En effet, si nous nous placions du point de vue de la science des données, nous aurions tendance à remonter le plus d'informations possible afin de pouvoir maximiser nos chances de réaliser les fonctions intelligentes avec la meilleure précision possible. En revanche, d'un point de vue industriel la solution nous amenant à limiter au maximum le nombre de transactions de données sur le réseau est préférable. C'est donc au regard de ce constat que la seconde hypothèse de notre recherche prend tout son sens. En effet, nous souhaitons ici spécifier notre protocole de remontées de données en assurant la faisabilité de notre fonction du service tout en minimisant son coût.

C'est par le biais de ces premiers constats que nous avons initié la conception de ces fonctions. Nous présenterons par la suite plus précisément nos fonctions liées à la maintenance. Puis, et au regard du caractère itératif de notre méthodologie, nous présenterons l'ensemble des trois itérations que nous avons réalisées et sur lesquelles repose notre raisonnement afin de pouvoir focaliser par la suite notre analyse sur le séquencement des étapes du modèle, en considérant parallèlement l'ensemble de nos itérations.

Contexte : Fonctions intelligentes d'identification et de prédiction des pannes

Afin d'éprouver notre approche et d'en faire ressortir les tenants et les aboutissants, nous l'appliquerons dans le cas du développement d'une fonction intelligente de notre système pour la reconnaissance de panne et la proposition de diagnostic pour le service après-vente. L'analyse du service représentée par la figure 40 nous montre que cette fonction repose principalement sur deux sources de données : les données de service du service après-vente et sur un flux de données émanant de la chaudière et transcrivant son état de fonctionnement en vue de la maintenance. La fonction doit également pouvoir être restituée sous forme d'une liste de pièces détachées accompagnée d'un protocole à suivre pour le mainteneur.

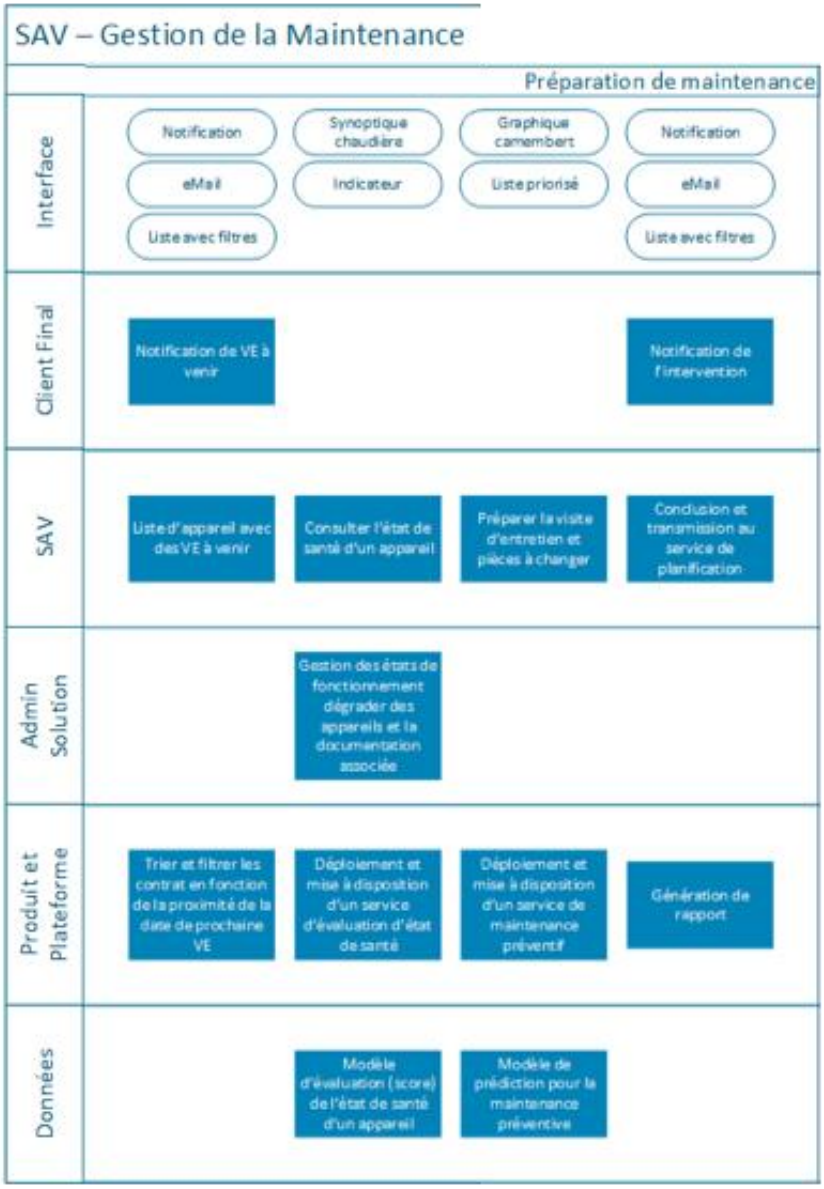


Figure 40 : Extrait du service *blueprint* des fonctionnalités liées à la maintenance pour les services après-vente

Notre principale implication ici est de définir l'ensemble des données nécessaires et suffisantes afin de réaliser cette fonction. Or, l'automate de notre produit générant et mettant à disposition un grand nombre d'informations, nous devons alors sélectionner convenablement celles que nous souhaitons transmettre à notre plateforme de service au regard de leur traitement et notamment afin d'assurer la faisabilité de nos fonctions.

Nous avons appliqué notre approche méthodologique, de nature itérative et incrémentale, dans une optique de recherche-actions (REF), dans le sens où nous sommes entrés dans le problème au sein d'une première itération afin de le découvrir et de mieux le percevoir. Chaque itération permet alors de reconsidérer notre problématique de conception initialement mal définie, pour nous permettre de rebondir sur de nouvelles pistes de résolution.

Cette mise en application, principalement orientée sur la donnée, n'étant pas communément appliquée aussi tôt dans la conception de systèmes d'information, nous intégrons ici l'ensemble des métiers de l'entreprise pouvant être concernés. Nous y retrouvons :

- Le domaine de la science des données, avec des compétences principalement basées sur la fouille de données et la mise en application de ses techniques.
- Le domaine du service, possesseur de la connaissance globale du système, avec notamment sa composante organisationnelle et l'ensemble des leviers d'exploitation des données.
- Le domaine d'application de notre périmètre fonctionnel, pour lequel nous retrouvons la maintenance, avec ses techniciens sur le terrain et ses services supports de gestion et planification. Mais également les métiers liés au développement produit, et le service de gestion de la qualité.

Protocole :

Dans le cadre de notre étude, nous présenterons l'application de notre démarche suivant l'axe du modèle méthodologique, étape par étape, en focalisant notre attention sur l'intervention des métiers dans chacune des phases. Afin de clarifier le contexte d'application, nous présenterons ici les trois itérations par lesquelles nous sommes passés et leur séquencement afin d'offrir une vision panoramique sur le déroulé de l'expérimentation. La figure 41 montre le séquencement de ces itérations, sur lequel nous reviendrons plus en détail par la suite.

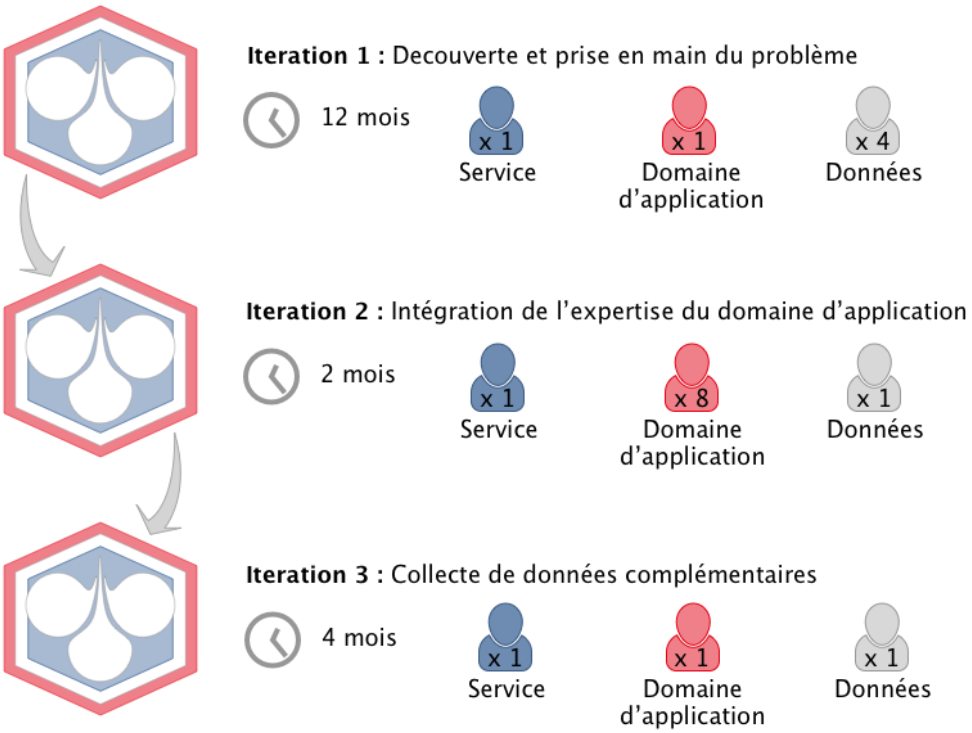


Figure 41 : Protocole de l'expérimentation 2

Itération 1 : Découverte du problème et investigation des pistes de solution

Cette première itération s'est initiée par la fourniture d'un jeu de données composé de un an de données de fonctionnement de plus de cinq cents produits sur le terrain. L'analyse de ces données s'est faite dans le cadre d'un partenariat avec d'autres entreprises et unités de recherche du groupe Bosch. Cette itération s'est étalée sur une période 12 mois avec une équipe de 6 personnes : un expert du service, un expert du domaine d'application et quatre analystes. L'objectif principal de cette itération était de découvrir le problème, de prendre connaissance des cas d'usage et des données afin d'évaluer les perspectives de réalisation d'un système de découverte de connaissance. Cela a été réalisé autour des compétences de la science des données au sein d'un projet de recherche interne au groupe Bosch présenté plus en détail dans l'Annexe C.

Cela nous a amené à évaluer la représentativité des données que nous avions à notre disposition, l'analyse des approches par lesquelles nous souhaitions avancer au regard des cas d'usage. Puis à réfléchir à la mise en forme de ces données dans le cadre de l'exploitation que nous avions préalablement définie.

Cette première itération nous a conduits à une approche satisfaisante d'un point de vue service sur la traduction des données en vue d'une transmission dans un système d'exploitation. Néanmoins, l'analyse de ces dernières et la réalisation des algorithmes n'ont pas pu être menées à bien en raison d'un manque de représentativité des données vis-à-vis du cas d'usage. Effectivement, n'étant pas en mesure de lier le comportement d'un produit avec un mode de défaillance clairement identifié, nous avons pris la décision de nous tourner vers les métiers du domaine d'application pour nous aider dans le cadre d'une seconde itération.

Itération 2 : Reconnaissance d'une panne par une approche métier

Dans le cadre de cette seconde approche, nous avons pris le parti de faire intervenir l'ensemble des métiers du domaine d'application afin d'intégrer leurs connaissances au sein de notre système. Cette recherche s'est tenue sur une période de 2 mois pendant laquelle nous avons fait intervenir 8 participants, tous issus du domaine d'application. Chacun des participants disposait d'une expertise spécifique en relation avec notre sujet d'étude tel que le développement de produit, la maintenance ou la qualité. Cela s'est réalisé principalement autour d'un atelier créatif que nous avons préparé, en nous focalisant sur les cas les plus pertinents de détection et d'identification.

Cette démarche nous a conduits à une représentation plus claire des modes de défaillances potentielles de nos appareils associés à des heuristiques de détection. Ces derniers sont représentés sous la forme de mesures comportementales des produits dans leur fonctionnement. Ces heuristiques ont par la suite été testés sur la base de nos données disponibles.

Néanmoins, face à la limitation des données disponibles fournies par nos appareils, nos experts ont alors émis des doutes sur la faisabilité de certains cas sans passer par l'ajout de capteurs complémentaires. Cela dit, et ayant la confirmation d'une altération du comportement du système pour l'ensemble de ces cas problématiques, nous nous sommes alors réorientés vers une approche donnée dans une troisième itération, basée sur la création de notre propre jeu de données spécifique à la détection de ces cas particuliers.

Itération 3 : Reconnaissance de panne par une approche données

Pour cette troisième et dernière itération, nous sommes partis du postulat que nos indicateurs de mesure du comportement du produit étaient suffisants pour assurer la détection des altérations du comportement des produits dans le cas des modes de défaillance jugés problématiques par nos experts. Cela nous a conduit à réaliser un plan d'essai pour assurer la collecte de données produit dans un

environnement de fonctionnement contrôlé, apportant la représentativité manquante à notre première approche pour la mise en application d'algorithmes de classification. Cette étude a été réalisée sur 4 mois avec l'aide de 3 participants : un expert sur le service, un expert dans les données et un expert du domaine d'application. Nous avons fait intervenir ici le responsable des essais en laboratoire afin de nous aider dans la collecte des données et pour représenter le domaine d'application. L'expertise en données est ici représentée par un étudiant en master de recherche effectuant son stage de fin d'études dont le contenu est décrit dans l'Annexe B.

Cette approche a montré la pertinence de nos indicateurs par la mise en évidence de cette altération du comportement produit au travers de nos données. Et même si nous n'avons pas pu produire un modèle d'identification et de prédiction de pannes au travers de ces trois itérations, nous avons pu évaluer la faisabilité de notre fonction et l'impact que peux avoir le développement de notre produit dans ce sens.

4.2.4. Application

Maintenant, afin de présenter la mise en application de notre démarche méthodologique, nous détaillons ici l'ensemble des actions menées au sein de ce développement étape par étape, suivant la séquence définie par notre modèle. Pour chacune de ces étapes, nous analyserons alors quels ont été les efforts déployés, quelles connaissances ont été sollicitées dans sa réalisation et quels moyens ont étés mis en œuvre afin de formaliser les résultats. Nous faisons alors abstraction ici de la dimension temporelle réelle du projet en traitant l'ensemble des trois itérations simultanément afin de nous focaliser sur les critères d'analyse de la dynamique collaborative au sein du projet que nous avons mis en avant dans la présentation de notre démarche.

Étape 1 : Compréhension du cas d'usage

La première étape fut la présentation du cas d'usage. Cette étape consiste à expliciter le cadre d'application afin d'en dégager les bénéfices que nous pouvons tirer d'une analyse des données disponibles. L'idéal étant ici d'extraire le besoin et les critères d'acceptation associés au domaine d'application et à la fonction du service que nous souhaitons mettre en place.

Dans cette démarche et particulièrement dans cette expérimentation, l'ensemble des activités que nous avions préalablement menées pour définir notre service nous a été des plus bénéfique, car nous sommes à présent en mesure de définir notre besoin et de détailler les gains que nos clients peuvent tirer d'un tel système de gestion de connaissances. L'expertise du service, en accord avec le domaine d'application est alors le principal moteur de cette phase de définition du cas d'usage.

La fonction que nous ciblons ici est double. Nous y retrouvons le concept d'identification des pannes et le concept de prédiction des pannes. Ces deux concepts touchent notre service dans deux perspectives différentes pour l'amélioration de la maintenance, le premier pour la maintenance curative et le second pour la maintenance préventive. Afin de transmettre nos besoins, nous avons rédigé et présenté un brief lors du premier atelier organisé au sein du projet dont la figure 42 représente le principal propos.

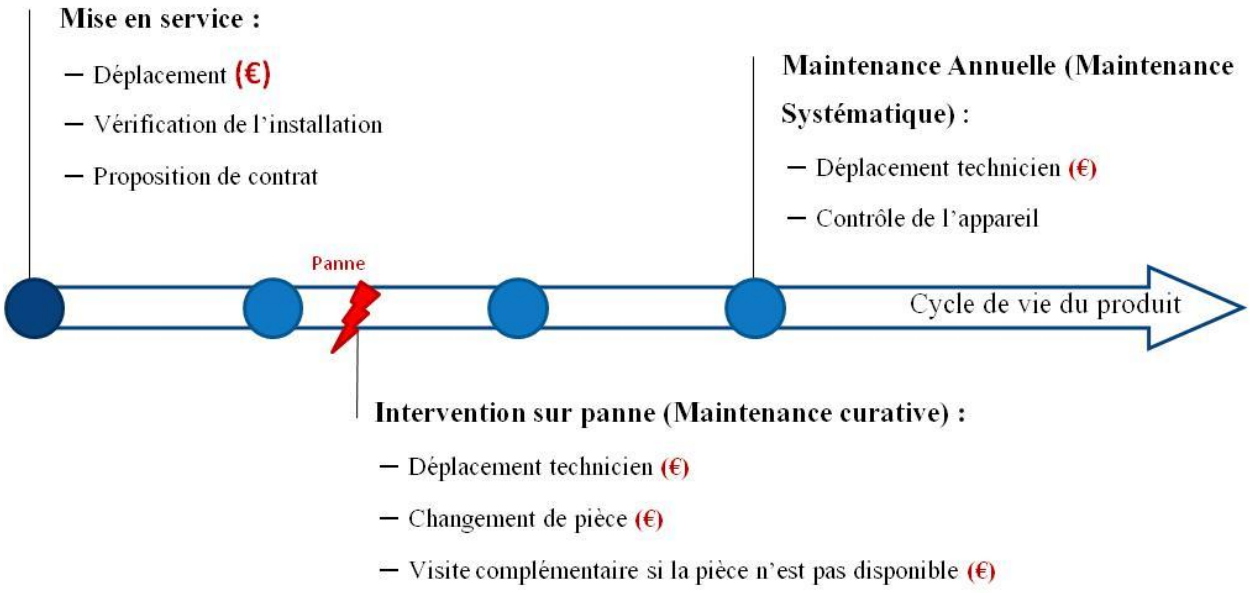


Figure 42 : Schéma explicatif du cas d'usage de l'expérimentation 2

La maintenance des chaudières s'organise de deux manières. D'une part, nous organisons une visite annuelle réglementée pendant laquelle le mainteneur doit contrôler le bon fonctionnement de l'appareil et effectuer un certain nombre d'opérations de nettoyage. D'autre part, le mainteneur peut être amené à se déplacer chez un client en dehors du cadre de ces visites annuelles lorsque l'appareil est en panne. Cela demande au mainteneur d'effectuer un certain nombre d'opérations de maintenance afin de remettre la chaudière en service et, potentiellement, cela peut entraîner le changement d'une ou de plusieurs pièces détachées. Les efforts déployés par le mainteneur pour l'ensemble de ces visites et opérations de maintenance, qu'elles soient curatives ou préventives, sont plus ou moins prises en charge, et donc plus ou moins rentables pour l'entreprise de service après-vente, en fonction du contrat souscrit par le client.

Dans ce modèle d'activité métier, nous pouvons alors situer le gain de chacune de nos fonctions :

- L'identification des pannes : Lorsque le mainteneur est amené à se déplacer chez un client afin de remettre en service son appareil de chauffage, le système doit être en mesure de lui indiquer l'ensemble des pièces détachées pouvant être changées. Avec cette information, le mainteneur s'assure de pouvoir régler le problème en une fois et de minimiser le nombre déplacements qu'il aura à faire chez son client en anticipant les efforts de maintenance à réaliser et les stocks de pièces détachées à prévoir lors de sa visite.
- La prédiction des pannes : Lors de la visite annuelle du mainteneur chez son client, il doit pouvoir avoir une vision de l'état de santé de l'appareil qu'il entretient avec une projection sur ses potentiels risques de panne dans l'année à venir. Cela lui permettrait de pouvoir anticiper les opérations de maintenance à réaliser au cours de cette visite annuelle obligatoire, et donc, par ce fait, de limiter les coûts de maintenance de ce contrat en minimisant le nombre de déplacements entre deux visites annuelles.

Au-delà de la description purement fonctionnelle de notre périmètre, l'analyse préalable que nous avons faite de notre service et de l'exploitation de ces données de maintenance fait apparaître des contraintes en termes de réalisation. D'une part, nous observons des contraintes fonctionnelles dues à l'expression du gain de ces fonctions pour l'utilisateur, avec notamment le fait de devoir identifier les pièces détachées potentiellement mises en cause, et non uniquement les modes de défaillance, afin de pouvoir être valorisé au sein du service. D'autre part, nous observons des contraintes techniques de réalisation, mises en évidence par la réalisation d'un plan de rentabilité du système. Nous pouvons ici donner pour exemple la fréquence de transmission que nous devons impérativement limiter à un maximum d'une par heure afin de limiter les coûts liés à la communication.

D'autre part, nous avons ici réalisé un état de l'art sur la conception de systèmes de maintenance afin de nous positionner et de justifier notre approche. Nous pouvons voir par la figure 43 que le domaine de la maintenance propose des approches similaires à notre volonté de superviser des produits afin d'en améliorer la maintenabilité, par la maintenance conditionnelle. Ce type de maintenance est basé sur la supervision d'indicateurs de fonctionnement du produit pour lesquels nous identifions des limites de validité. Un dépassement de ces limites entraîne la génération d'une alarme préventive indiquant une dérive ou un dysfonctionnement de l'appareil suivi. Cette approche métier ressemble beaucoup à notre cas d'application, cependant sa mise en application se base sur une identification des modes de défaillance, afin de prévoir un ensemble de capteurs nécessaires à la supervision du comportement du produit. Cela est alors jugé incompatible avec notre objectif de n'utiliser que les données mises à disposition par l'automate de notre appareil.

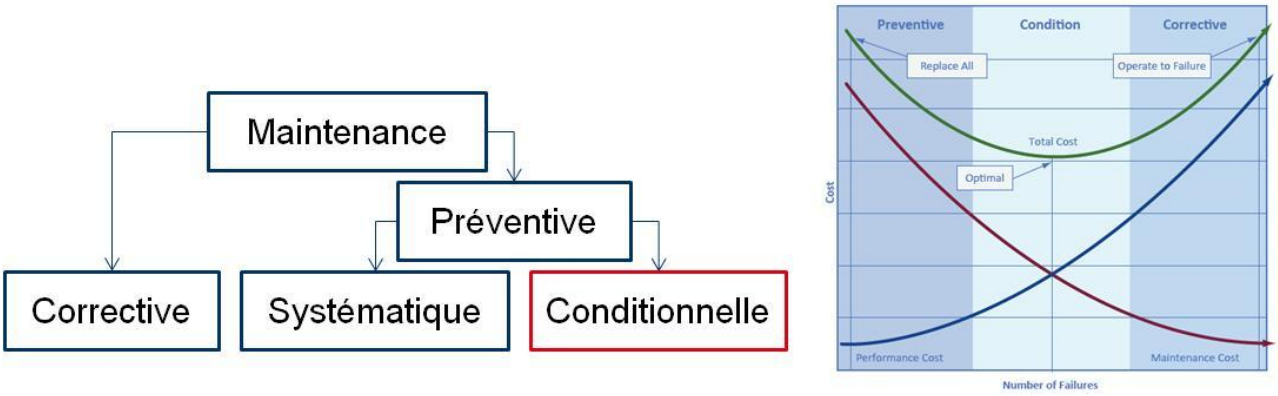


Figure 43 : Positionnement du cas d'usage de l'expérimentation 2 dans le domaine de la maintenance

Pour conclure sur cette première étape, nous pouvons observer dans le tableau 1 l'importance des études réalisées autour du service. Ce dernier permet de définir très clairement le périmètre de l'étude, à l'aide de l'analyse préliminaire que nous avons réalisée dans l'expérimentation précédente. Le domaine d'application apporte des pistes de réponse en identifiant notre problème de conception à une typologie de problèmes déjà connus et pour lesquels des solutions existent. Enfin, le domaine d'application et la science des données peuvent pointer du doigt des problématiques déjà rencontrées dans des projets similaires afin d'extraire des priorités du projet.

Service	Domaine d'application	Science des données
Service *blueprint*	**Heuristique**	**Retour d'expérience**
• Définition du cas d'usage	• Définition du cas d'usage	• Priorités
• Fonctions	**Retour d'expérience**	
• Usage	• Priorités	
Modèle Economique		
• Priorités		

Table 1 : Résumé de l'étape 1

Étape 2 : Compréhension des données

Après nous être positionnés et avoir défini le périmètre de notre étude, la deuxième étape de notre processus méthodologique est la compréhension des données. Ces dernières constituant la substance principale du système en cours de conception, il nous faut ici construire une image de ce qui nous est rendu disponible et de leur représentativité au sein du produit. C'est pourquoi les intervenants du domaine d'application sont les principaux moteurs de cette étape.

En se basant sur les spécifications de l'interface de communication de nos produits, le travail a été ici principalement d'identifier les champs intéressants afin de servir notre propos. Poussée par le métier, cette étape a permis de restreindre notre univers de plus de 180 champs à une vingtaine. D'autre part, cela nous a permis de donner du sens à ces données vis-à-vis du système technique duquel elles proviennent. Nous pouvons alors noter que les données principales retenues sont :

- La puissance du brûleur
- La consigne en puissance du brûleur
- La température d'eau en sortie du brûleur
- La consigne de température d'eau en sortie du brûleur
- La température d'eau en sortie de la chaudière pour le sanitaire
- La consigne de température d'eau en sortie de la chaudière pour le sanitaire
- Le débit d'eau sanitaire

- Le statut de fonctionnement chauffage
- Le statut de fonctionnement sanitaire
- La position de la vanne de distribution d'eau pour le chauffage ou le sanitaire
- La position de la vanne gaz
- Le statut de fonctionnement du ventilateur du mélange air-gaz
- Le statut de démarrage et de création d'un train d'étincelles
- Le statut de présence de flamme
- Le courant d'ionisation de la flamme
- Le statut de fonctionnement de la pompe
- La valeur de modulation en puissance de la pompe
- Les statuts d'erreur de l'appareil
- Le statut de définition des erreurs appareil

En parallèle, nous retrouvons un ensemble d'autres données nous donnant plus de précisions sur le type d'appareil suivi, sachant que l'ensemble de ces valeurs ne sont pas présentes pour tous les appareils comme le débit en sanitaire par exemple.

Suite à l'extraction de cette liste de données pertinentes pour notre étude, nous avons également proposé de les représenter sur un schéma synoptique du produit comme le montre la figure 44 dans le but d'étayer nos réflexions sur les traitements envisageables.

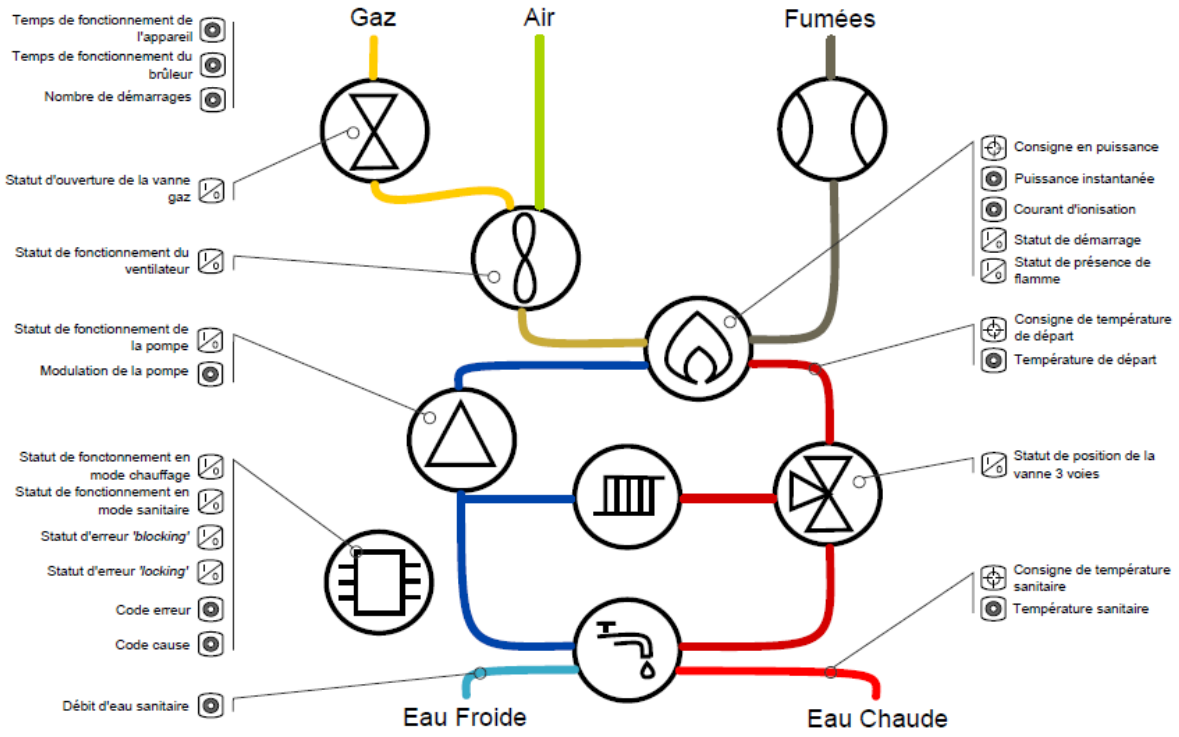


Figure 44 : Cartographie des données du produit disponibles pour la réalisation de l'expérimentation 2

D'autre part, nous souhaitons étayer également les informations disponibles concernant les erreurs des chaudières. Ces dernières sont effectivement l'unique représentation que nous avons d'une panne au travers des données des systèmes. Nous avons pu voir dans la définition du cas d'usage (cf. Etape 1) que ce qui nous intéresse ici est de nous concentrer sur les pièces potentiellement incriminées dans le cas d'une défaillance. Or, suite à la capitalisation des connaissances que nous avons pu extraire de l'ensemble de la documentation de ces codes erreur et suite à de nombreux échanges avec nos experts,

nous pouvons voir, tel que le montre la figure 45, que ces codes ne sont pas systématiquement des images de défaillances relatives à une seule pièce détachée. Effectivement, seulement peu d'entre eux sont en mesure de nous cibler une pièce précise. La majorité décrit un simple comportement anormal ou à risque, pouvant avoir de nombreuses causes.

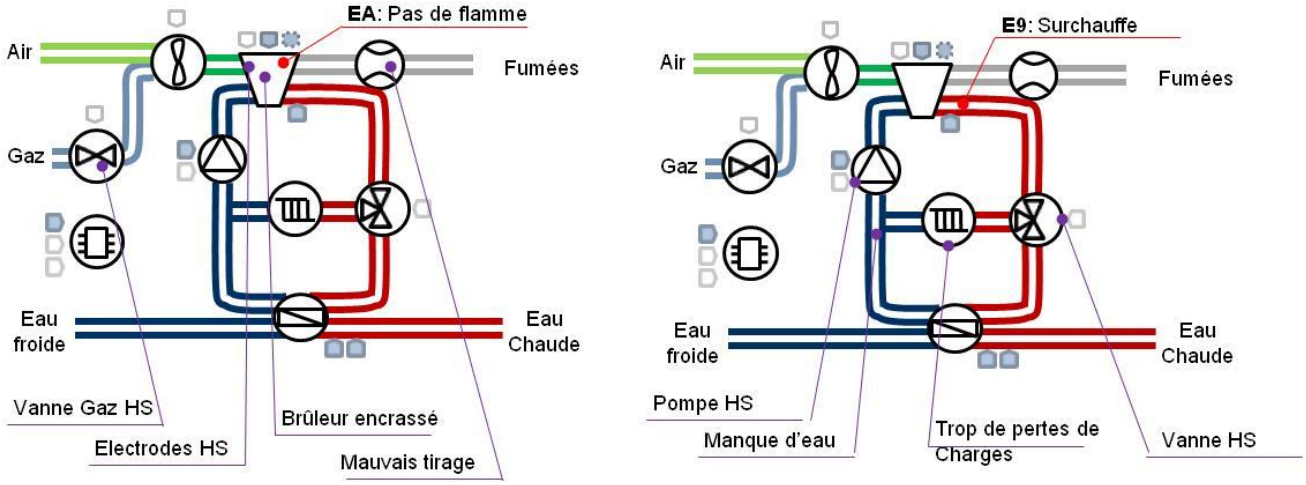


Figure 45 : Représentation des erreurs du produit et de leurs causes

Suite à la livraison des données provenant du projet dans lequel nous nous situons au sein du groupe Bosch, nous avons, avec l'aide des experts des données, commencé à appréhender ce que ces attributs représentent et comment ils nous permettent de définir notre système. Nous avons alors été en mesure, par différentes approches, de nous représenter les données. Nous pouvons ici mettre en évidence :

- Les différents profils de sollicitation des produits en fonction de leur lieu d'installation et la dépendance des différents utilisateurs au fonctionnement observé du produit dans les données (cf. Annexe I).
- Les procédures internes de démarrage et d'arrêt de l'appareil avec une analyse de transition des statuts (cf. Annexe I)
- Ou encore la représentation du fonctionnement de l'appareil par un découpage par cycles (figure 46)

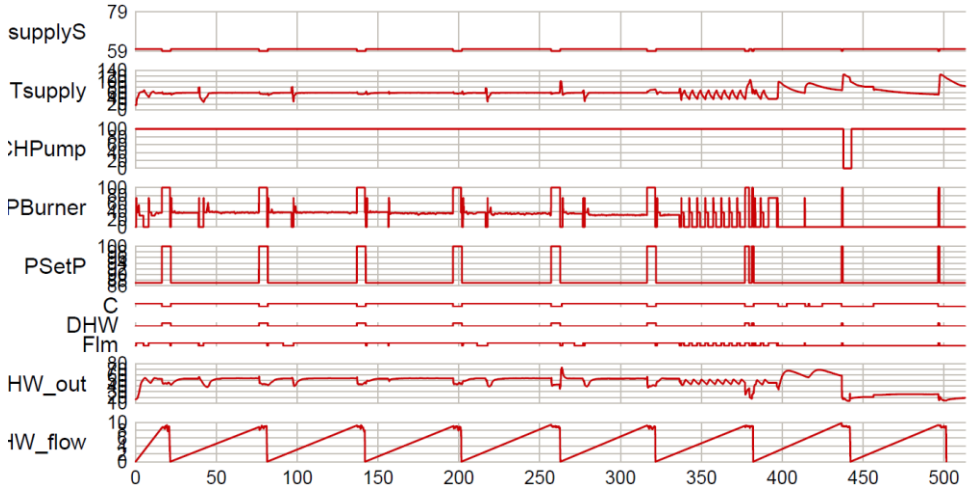


Figure 46 : Représentation du fonctionnement d'un appareil dans les données

Ces travaux de prise en main des données nous ont notamment permis de mettre en évidence les problématiques de perte significative d'information liées à la nécessité de transmettre un minimum d'une donnée par heure comme le montre la figure 47. Dans le cas d'une agrégation simple des données, nous perdons alors l'ensemble des schémas de comportement des produits, nous permettant par la suite de détecter des dysfonctionnements. Cela nous pousse alors à réfléchir par la suite à la création d'indicateurs pertinents pour assurer une représentation suffisante du comportement du système.

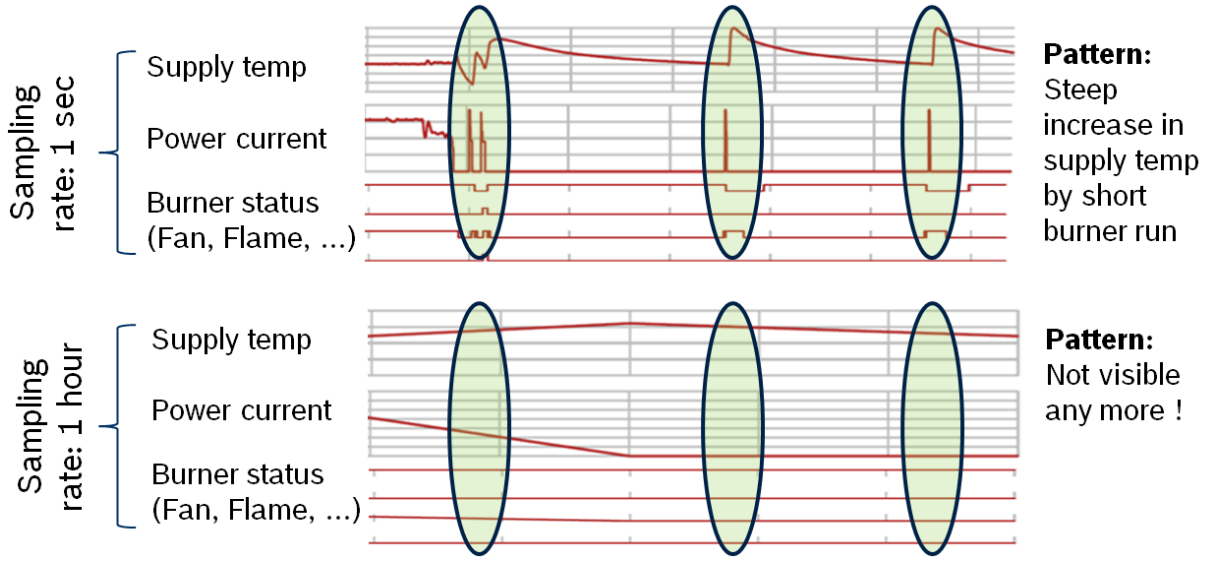


Figure 47 : Illustration de la problématique de collecte des données des appareils

Nous pouvons conclure sur cette phase en appuyant sur l'importance de ces différentes représentations faites du système. Le domaine d'application peut lister et expliciter l'ensemble des données disponibles alors que la science des données peut en évaluer la représentativité. Ces représentations résumées dans le tableau 2 sont très efficaces afin d'assurer les échanges entre les trois différents domaines d'expertise, en plus de nous avoir permis par la suite de focaliser les débats sur les réalités concrètes de notre problème.

Service	Domaine d'application	Science des Données
	Carte Système • Source les données • Explicite l'ensemble des données • Filtre les données les plus pertinentes	**Statistique Générale** • Représentativité des données disponibles

Table 2 : Résumé de l'étape 2

Étape 3 : Identification du modèle et prétraitement des données

Ensuite, nous entrons dans la phase d'identification du modèle et de prétraitement des données. Elle consiste à sélectionner les voies de traitement par lesquelles nous souhaitons passer afin de répondre à notre besoin. Le choix du modèle doit être fait relativement à la forme que doit prendre le résultat pour permettre son exploitation au sein du service. Mais cela doit également être fait relativement au format des données disponibles. Le cas échéant, cette phase consiste à choisir un modèle adapté et à préparer les données afin de les rendre adaptables à la mise en application de l'algorithme.

Comme nous l'avons vu dans l'état de l'art, c'est l'étape la plus chronophage de ce cycle de développement. Elle est d'autant plus longue ici que nous y décrivons les trois approches différentes par lesquelles nous sommes passés au cours des trois itérations que nous avons réalisées. Nous dissocions ici nos efforts en deux parties. Dans un premier temps nous évoquerons l'approche métier que nous avons adoptée dans le deuxième cycle afin de dégager les heuristiques propres au domaine d'application sur la base de nos formalisations du problème. Ensuite, nous aborderons les efforts que nous avons mis en œuvre dans le cadre de l'analyse des données afin de les préparer à l'application d'un algorithme d'apprentissage dans les cycles 1 et 3.

Avec l'ensemble des éléments que nous avons formalisés lors des deux étapes précédentes, nous sommes maintenant prêts à mener un atelier avec un panel d'experts. Cet atelier prend place dans l'étape d'identification du modèle et la réalisation du prétraitement, car nous pousserons ici nos experts à produire des résultats similaires à ceux que nous pouvons attendre d'un algorithme issu de la science des données, en identifiant des schémas de fonctionnement caractéristiques à un état de notre système. Cependant, et contrairement aux algorithmes, nous partons de ces schémas de fonctionnement pour redescendre à un moyen de les mesurer et de les quantifier, dans l'optique de pouvoir l'automatiser par la suite au sein de notre système.

Pour cela, nous avons constitué un panel d'experts disposant de nombreuses connaissances sur le fonctionnement des produits, sur les fonctions de régulation des automates qui les contrôlent et sur les modes de défaillance les plus observés sur le terrain. Nous retrouvons dans notre panel trois types de profils :

- Deux ingénieurs du service d'innovation. Ces participants connaissent l'ensemble des solutions technologiques du marché et les principes physiques qui les supportent. D'autre part, ils disposent aussi d'une faculté particulière à se projeter dans des problèmes desquels ils ne sont pas familiers sans inertie psychologique.
- Un ingénieur du développement, connaissant d'une part parfaitement l'ensemble des produits du catalogue avec leurs variantes technologiques et d'autre part l'ensemble des fonctionnalités du logiciel embarqué dans les automates.
- Cinq ingénieurs et techniciens de la maintenance venant de plusieurs horizons, tel que le responsable des retours qualité, des techniciens de maintenance, un opérateur de la ligne téléphonique d'assistance pour les professionnels et un contrôleur de gestion pour le service après-vente. Ces profils connaissent parfaitement les défauts de nos appareils, leurs occurrences et les moyens à mettre en œuvre afin de les détecter et de les traiter.

Afin d'alimenter les discussions dans ce sens, nous avons sélectionné deux codes représentatifs de toute la complexité de l'identification des pannes. Effectivement, la majorité des codes erreur pointent vers le dysfonctionnement d'une pièce, comme une sonde ou une vanne. Les codes sélectionnés sont, d'une part les plus présents sur le terrain, dont nous avons des exemples dans notre jeu de données, d'autre part les plus complexes par la multiplicité des causes possibles. Ces codes représentent les défauts suivants :

- Le code EA apparaît lors d'un défaut d'allumage de l'appareil
- Le code E9 apparaît lors de la présence d'une surchauffe en sortie du brûleur

Dans le but de faire converger rapidement les discussions autour des centres d'intérêts que nous avons ici, à savoir la modélisation des comportements d'un produit sur la base des données disponibles avant une panne spécifique incriminant un mode de défaillance en particulier, nous avons fourni à l'ensemble des participants les représentations du système et des codes erreurs avec leurs causes.

Cet atelier s'est tenu pendant une demi-journée que nous avons séparée en deux pour traiter indépendamment chaque code erreur. Suite à une brève présentation des participants et des documents que nous avons préparés, chaque partie s'est déroulée en deux étapes. La première étape consistait à prendre connaissance des modes de défaillance que nous avions mis en évidence vis-à-vis d'un code erreur, de s'assurer que les informations fournies étaient correctes et complètes, puis de les classer suivant leur potentiel d'observation et d'identification. Effectivement, dans l'ensemble des modes de défaillance, certains pouvaient ne pas être directement liés à l'appareil en lui-même, comme une coupure de réseau de gaz, ou bien, pouvaient entraîner un comportement trop soudain pour être observé, comme l'arrivée d'un court-circuit. La seconde étape a été de faire le lien entre les modes de défaillance potentiels et les données de l'appareil afin d'aboutir à des moyens de mesure quantitatives du comportement ciblé.

A la suite de cet atelier, nous avons pu faire apparaître deux faits. Le premier sur l'émergence de solutions proposées afin de percevoir les comportements liés aux défauts d'allumages. Le second sur l'impossibilité décrite par les experts de pouvoir espérer identifier la cause d'une surchauffe. Dans le cadre de cette itération, nous nous concentrerons sur la mise en évidence des heuristiques métiers dans les données, puis nous traiterons dans notre troisième itération les cas jugés comme impossibles. Nous pouvons d'ailleurs exposer les résultats de notre atelier dans le tableau suivant :

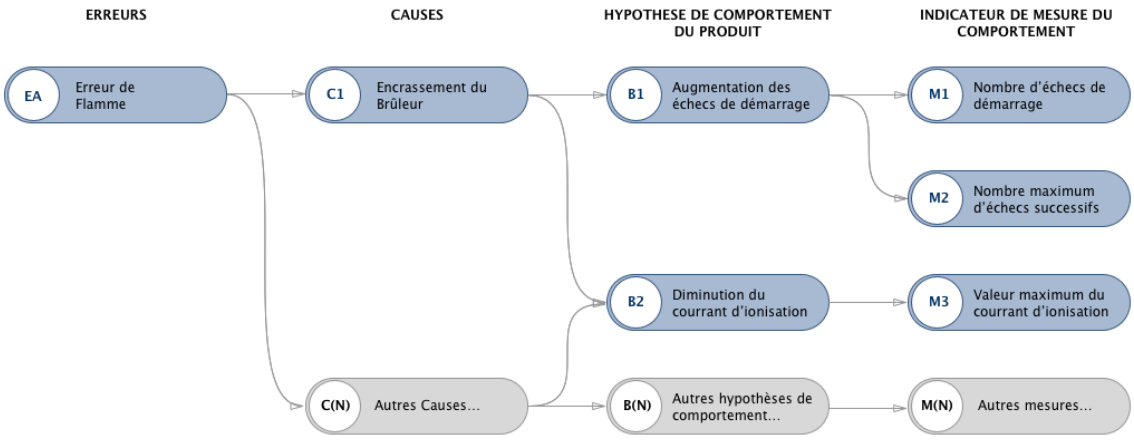


Figure 48 : Extrait des résultats de l'atelier créatif avec le domaine d'application

Ensuite, concernant la première et la troisième itération, où nous travaillerons sur l'applicabilité d'un algorithme d'apprentissage sur nos données produit, la sélection d'un algorithme de classification nous a semblé le plus pertinent afin de permettre l'identification d'un problème d'un appareil en état d'erreur, puis de juger la prédictibilité de cet état. Ensuite, le choix précis de l'algorithme importe peu, mais la mise en forme de nos jeux de données pour son application est cruciale. La base de données étant composée de données de type transactionnelles – c'est-à-dire que nous disposons d'une ligne pour

chaque changement de valeur d'un champ – nous devons transformer ces données en un tableau traduisant les conditions de fonctionnement de l'appareil.

Pour ce faire, nous avons abordé plusieurs approches permettant d'extraire des caractéristiques comportementales des appareils en fonctionnement. La première approche a consisté à reprendre l'analyse préliminaire de données présentées dans la partie précédente.

La deuxième approche vise à nous représenter notre système au travers d'un modèle physique, en appliquant les règles de la thermodynamique. Dans notre cas, la formule proposée fait le lien entre les températures mesurées et les capacités thermiques de notre fluide, ici notre eau de chauffage. Cela nous semble très intéressant car nous pourrions ici nous abstraire des caractéristiques thermiques particulières de chaque logement et nous focaliser sur certains problèmes de maintenance liés à cette eau de chauffage. Néanmoins nous ne pouvons pas conclure sur l'élaboration de ce modèle car, en l'absence de certaines données telles que la température ambiante du logement ou la température de l'eau dans les circuits de chauffage, nous sommes dans l'obligation de faire trop d'hypothèses que nous ne pouvons vérifier d'aucune manière.

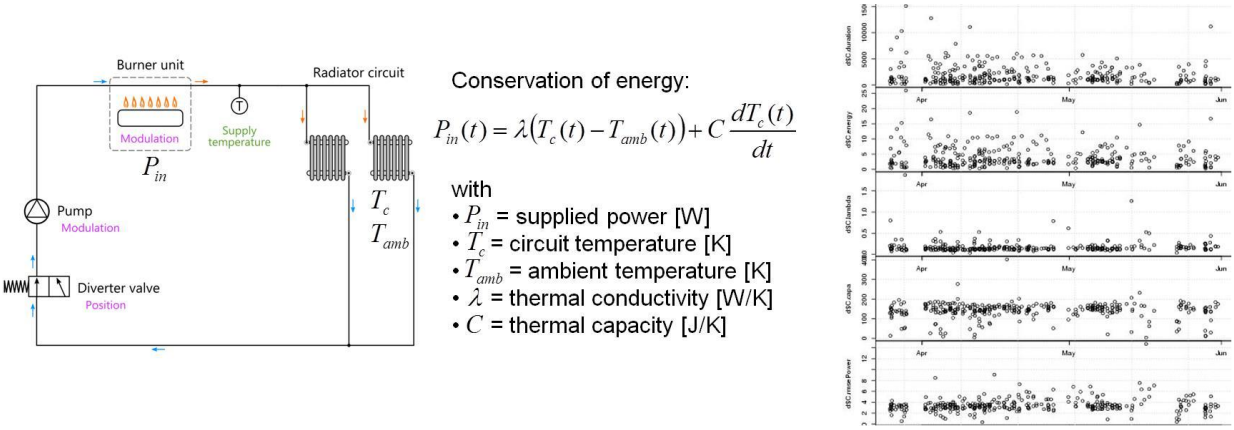


Figure 49 : Représentation de l'analyse par un modèle physique du système de chauffage

La troisième et dernière approche que nous détaillerons ici vise à découper le fonctionnement de l'appareil au travers ces nombreux cycles de fonctionnement. Nous avons effectivement ici, et tel que le montre la figure 46 un découpage du fonctionnement des appareils jonché d'arrêts et de redémarrages suivant les sollicitations qu'il subit en chauffage et en sanitaire. Nous pouvons alors par ce biais caractériser l'ensemble de ses cycles tel que le montre la figure 50, avec des pentes, des durées, des énergies, des dépassements de consigne etc. Cette représentation par cycles permet alors d'avoir une image du fonctionnement qualitatif de nos appareils dans le temps. D'autre part, c'est par ce type de raisonnement que nous avons pu traduire les résultats de notre atelier avec les experts afin d'y ajouter leurs représentations des comportements spécifiques à l'identification des cas complexes.

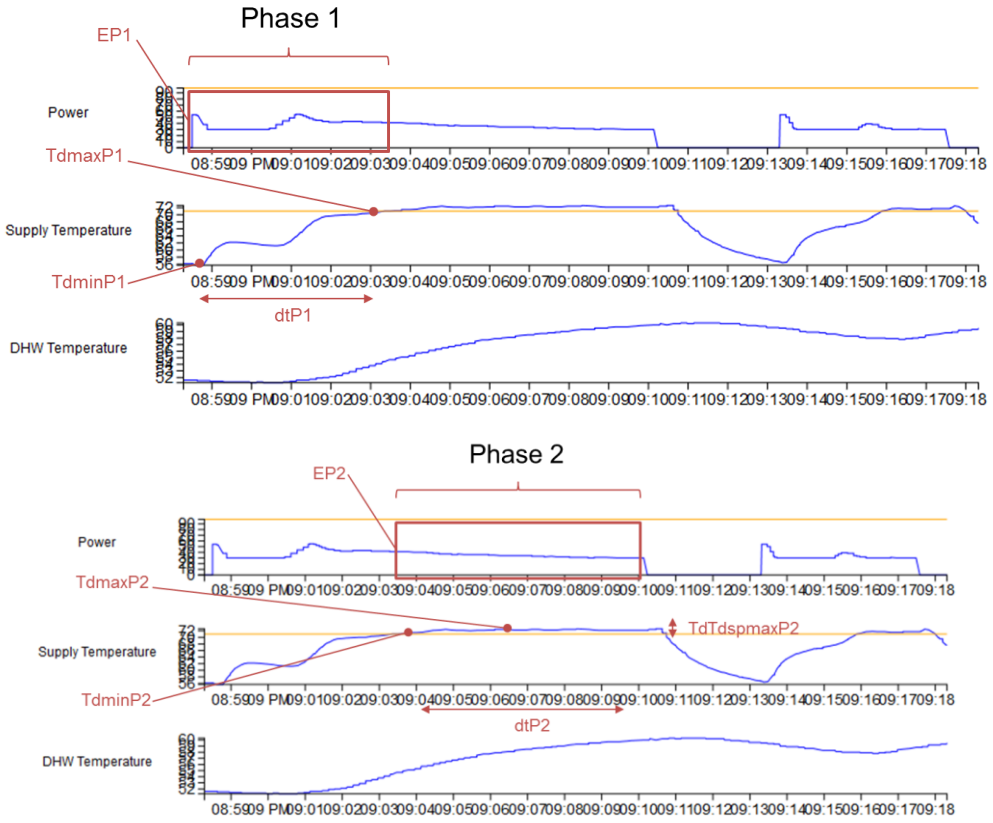


Figure 50 : Exemple de création d'indicateurs de fonctionnement pour chaque cycle des appareils

C'est alors sous cet angle que nous avons pu construire un tableau unique afin de permettre l'application de modèles de classification. Nous avons alors cumulé ces trois approches afin de créer un tableau décrivant le fonctionnement d'une chaudière au travers d'une ligne de caractéristique par jour. En revanche, ne disposant pas de l'information des erreurs effectives et l'incrimination de pièces dans le jeu de données fournie par le groupe, nous avons fait le choix de l'alimenter par l'ajout de nouvelles données issues de tests en environnement contrôlé.

Pour ce faire, nous avons créé un plan d'expérience que nous avons suivi afin de collecter l'ensemble des données dans des cas de fonctionnement normal du produit ainsi que dans des cas de fonctionnement dégradés. Nous nous sommes d'ailleurs basés sur les remarques de nos experts lors de notre atelier afin de nous focaliser sur la mise en évidence de comportements jugés difficiles à prédire. Nous prenons le cas ici d'un manque d'eau dans le système, pouvant entraîner une surchauffe. Pour ce faire, nous avons alors installé un appareil instrumenté dans le laboratoire d'endurance de l'entreprise, tel que le montre la figure 51. Nous l'avons fait fonctionner sur des périodes de trois heures en faisant varier :

- La pression d'eau dans son circuit, afin de créer des modes de fonctionnement dégradés
- La consigne en température
- Le temps et le nombre de demandes de fonctionnement sanitaire
- Le temps et le nombre de demandes de fonctionnement en chauffage
- Le taux de refroidissement de l'eau de chauffage
- Le débit de puisage en sanitaire



Figure 51 : Installation d'un appareil instrumenté pour réaliser de la collecte de données complémentaires

La mise en application du prétraitement établie dans les deux premières itérations nous a alors permis de constituer un échantillon de données dont nous connaissons la défaillance et dont nous pourrons par la suite tester la prédictibilité. C'est ce que nous montrerons dans les étapes suivantes.

Néanmoins, nous pouvons conclure sur cette étape avec les apports des différents métiers résumés dans le tableau 3. Les choix des modèles se sont faits d'une manière conjointe entre le domaine d'application, à l'aide de l'heuristique préalablement identifiée, et de la science des données ; et l'ensemble des techniques d'apprentissage connues et compatibles avec le cas d'usage. Par la suite, les efforts sont principalement concentrés autour de la fourniture d'indicateurs de comportement du produit, également appelé « attributs », nous observons une forte collaboration entre le domaine d'application et la science des données. Le domaine d'application tend à dégager des caractéristiques pertinentes au regard de leurs connaissances autour du cas d'application alors que la science des données dispose de nombreuses techniques fréquemment utilisées dans ce but. D'autre part, dans le cas d'un manque de représentativité préalablement identifié, le domaine d'application joue un rôle déterminant dans l'apport de consistance dans nos données par la construction de plan d'essai.

Service	Domaine d'application	Science des Données
	Heuristique • Définition d'une approche **Indicateurs** • Extraction de caractéristiques pertinentes • Formatent les données en adéquation avec le cas d'usage **Plan d'expérience** • Apporte de la consistance pour faire face au manque de représentativité des données disponibles	**Technique d'apprentissage** • Définit une approche **Techniques de prétraitement** • Extraction de caractéristiques du produit • Formatent les données en adéquation avec le cas d'usage

Table 3 : Résumé de l'étape 3

Étapes 4 et 5 : Création du modèle et évaluation

L'étape suivante de notre cycle de développement réside dans la mise en application des techniques d'apprentissage que nous avons identifiées et sélectionnées dans la phase précédente. Cela consiste à appliquer l'algorithme d'apprentissage aux données disponibles afin d'en déduire un modèle. Dans notre cas d'application et afin de simplifier l'exposition des résultats, nous illustrons ici l'étape suivante également, consistant à évaluer les résultats obtenus. Cette étape permet d'évaluer les performances de l'algorithme par rapport aux données d'entrée que nous avons sélectionnées. Nous pouvons néanmoins noter que cette évaluation est purement associée à la science des données. Nous n'évaluons pas ici les performances du modèle généré vis-à-vis du service que nous avons explicité dans l'étape 1.

Nous dissocions une nouvelle fois ici l'approche orientée vers le domaine d'application de celle orientée vers les données.

Concernant notre approche issue du domaine d'application, une fois nos heuristiques découvertes, nous passons maintenant à leur application. Nous nous attarderons dans cette étape à leur réalisation et leur mise en évidence. Pour ce faire, nous exposerons le premier principe identifié par notre panel d'experts, à savoir, la détection d'un encrassement du brûleur, pour lequel nous nous chargerons de traduire les mesures identifiées par les experts en procédures de prétraitement par des moyens de calcul réalisables en temps réel avec, si cela est possible, une relation au sein d'un cycle de fonctionnement du produit. De plus, nous souhaitons matérialiser ces indicateurs dans l'optique de pouvoir vérifier les hypothèses énoncées par notre panel d'experts. Pour ce faire, nous nous orientons vers une validation basée sur la visualisation, en sélectionnant un cas avéré de notre problème dans les données du groupe, et en exposant de manière visuelle le comportement décrit dans l'hypothèse énoncée par nos experts.

Dans notre cas, l'hypothèse de nos experts était :

> *En cas d'un encrassement d'un brûleur entraînant une erreur de flamme, l'appareil doit voir son nombre d'échecs de démarrage augmenter progressivement et son courant d'ionisation diminuer progressivement avant l'arrivée de la panne.*

La première mesure concerne le nombre de tentatives de démarrage avant l'obtention d'une flamme. Cette mesure s'appuie sur le statut de démarrage dans les données, un booléen correspondant à l'activation du train d'étincelles pour allumer la chaudière. Ce comportement avait déjà été perçu dans la première itération par l'analyse des transitions de statut. A présent, nous pouvons établir un critère précis de mesure afin de la positionner sous un format adéquat pour l'application d'un modèle de classification. Et nous pouvons, à l'aide des indications fournies par notre panel d'experts, établir une procédure de calcul. Typiquement :

> *Si le statut de démarrage disparaît et que le statut de flamme n'est pas présent pendant un cycle, alors le nombre d'échecs de démarrage augmente de 1.*

La seconde mesure concerne le courant d'ionisation. Ce dernier est une valeur mesurée par une électrode au cœur du brûleur donnant une image de la qualité de la combustion et dépendante de la puissance de fonctionnement. Nous pouvons noter une relation forte entre ce courant d'ionisation et la puissance de fonctionnement du brûleur. Afin de faciliter notre mise en œuvre, nos experts nous ont conseillé de traiter ce cas du courant d'ionisation uniquement dans le fonctionnement de la chaudière en sanitaire, car les puissances étaient plus fortes et plus constantes. Alors les valeurs mesurées pour les

courants seraient plus représentatives et plus facilement comparables entre elles. Nous avons alors généralisé cette approche en proposant la procédure de calcul suivante :

> *Si la puissance de la chaudière arrive à un régime stabilisé et que la valeur actuelle de la puissance est supérieure à la valeur déjà en mémoire, il faut alors conserver la valeur de puissance et du courant comme référence.*

Nous mettons donc en application ces procédures de calcul de nos indicateurs, puis nous sélectionnons un cas dans nos données présentant un cas d'erreur de flamme afin de représenter l'évolution de ces facteurs dans la période précédant la panne. Nous pouvons alors, pour une même installation, représenter le nombre de d'échecs de démarrage ainsi que l'évolution du courant d'ionisation par la figure 52.

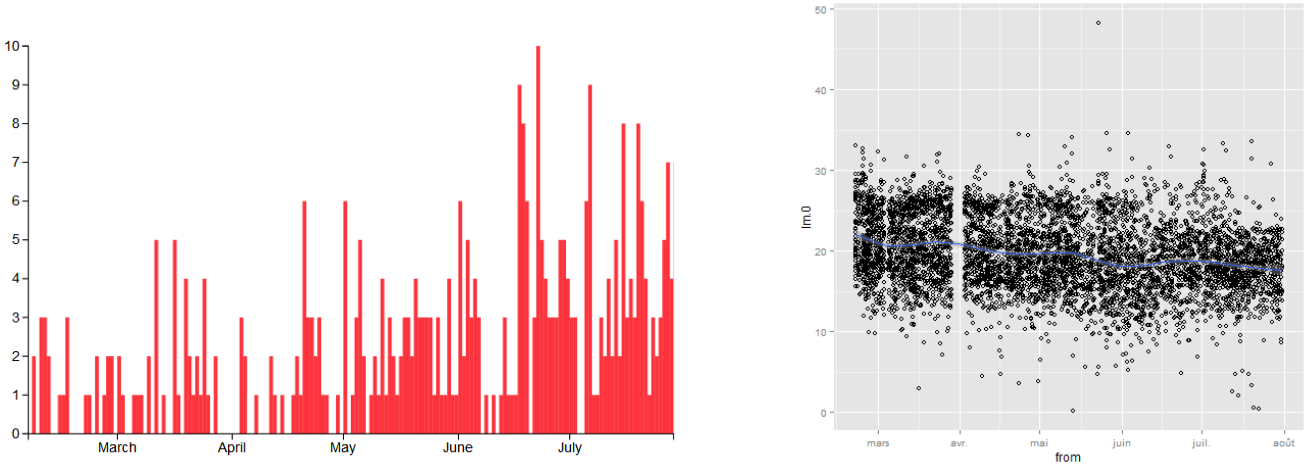


Figure 52 : Illustration visuelle des comportements produits avant l'arrivée d'une erreur EA avec respectivement de gauche à droite, le nombre d'échecs de démarrage et le courant d'ionisation

Ces deux figures nous montrent effectivement que le nombre d'échecs de démarrage a tendance à augmenter alors que le courant d'ionisation a tendance à diminuer avant l'apparition de l'erreur. Ceci nous permet alors d'une part de valider l'hypothèse émise par notre panel d'experts, et d'autre part de valider notre procédure de prétraitement des données afin de fournir des indicateurs pertinents.

En ce qui concerne notre approche orientée vers l'exploitation de la donnée, nous n'avons pas pu conclure sur la première itération sachant que nous ne disposions pas de l'information nécessaire à la création d'un modèle répondant à notre cas d'usage. Cependant, notre troisième itération a permis de nous fournir cette information dans un cas d'application déterminé. Nous avons alors pu, sur cette base, expérimenter la reconnaissance de panne par la création d'un arbre de décision, puis nous avons souhaité évaluer l'importance de chaque indicateur par une analyse en composante principale.

La création d'un arbre de décision basé sur notre nouveau jeu de données, illustré par la figure 53, nous a permis de montrer qu'il était possible de créer un modèle capable de classer l'ensemble de nos échantillons suivant leur niveau de pression en fonction des indicateurs que nous avions préalablement définis. Et ce sans passer par une étape de validation du modèle, car nous ne nous intéressons pas ici à produire un modèle directement exploitable dans notre système mais plutôt à montrer la faisabilité de la prédiction, nous pouvons observer que suivant 4 critères, l'algorithme arrive à identifier le niveau de pression dans lequel le système fonctionne avec seulement 6% d'erreurs. Cela montre donc la pertinence de ces critères dans le cadre de cette fonction, et particulièrement dans le cas particulier lié à l'identification d'une perte de pression dans le système.

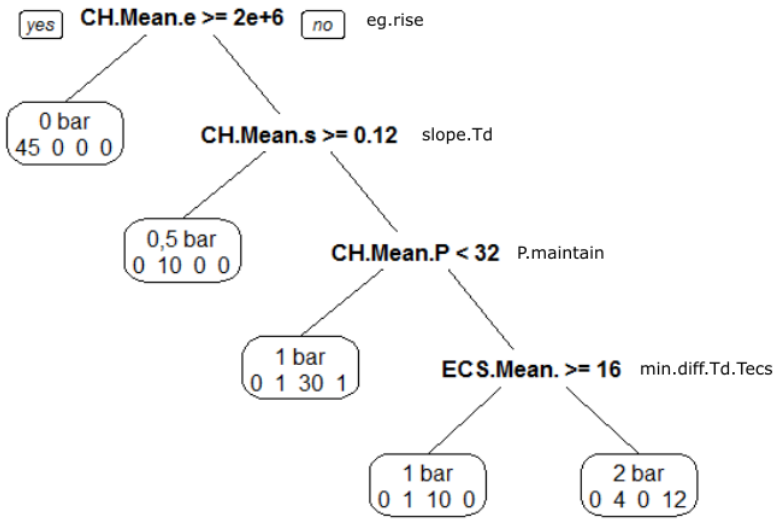


Figure 53 : Arbre de décision pour identifier un manque de pression généré sur la base des données complémentaires collectées

D'autre part, nous avons cherché à approfondir notre analyse pour comprendre l'implication de nos indicateurs dans la discrimination des situations de fonctionnement. L'analyse en composantes principales, représentée dans la figure 54 nous permet de découvrir les axes représentant au mieux la variance de notre jeu de données. La collecte ayant été faite dans un environnement contrôlé, où la variation principale est le fonctionnement du système suivant le mode de défaillance ciblé, nous y observons, tel que nous l'espérions, une segmentation claire de nos échantillons suivant la première composante. Par cette observation, nous pouvons remonter à l'implication de nos indicateurs et leur représentativité relative à ce cas particulier par l'analyse de cette composante. Cette analyse montre alors la prédominance de certains indicateurs dans la création de cette composante principale.

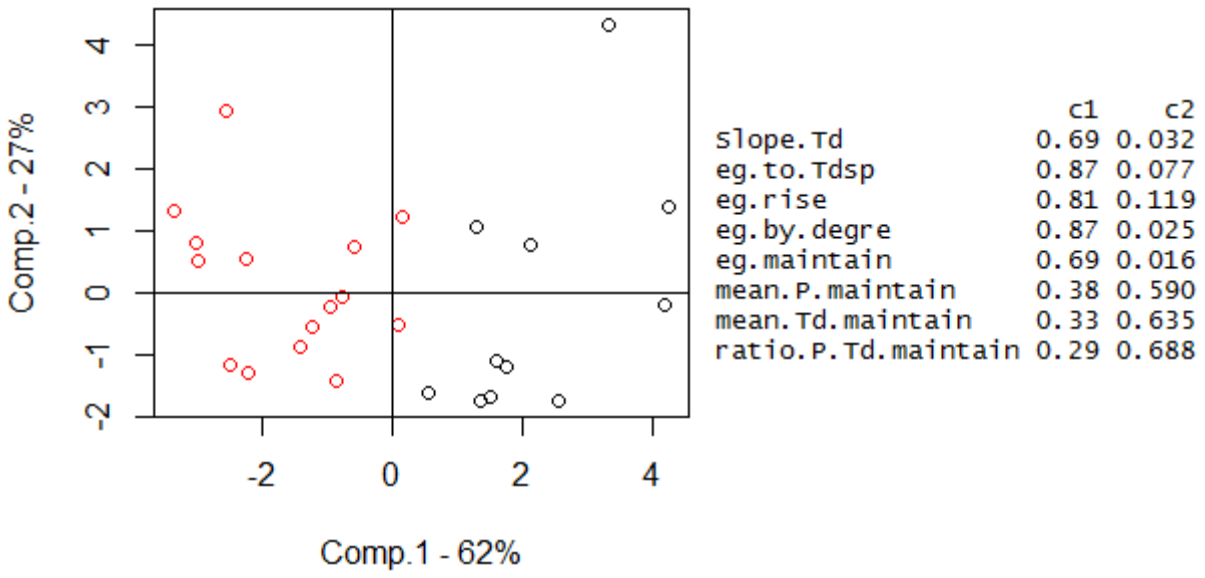


Figure 54 : Analyse en composantes principales avec nos observations (Rouge pour le manque de pression, Noir pour le fonctionnement nominal) et la composition des axes

Par cette approche, nous pouvons alors valider la pertinence de nos indicateurs vis-à-vis du cas d'usage que nous avons sélectionné.

D'autre part, nous pouvons ici conclure avec le tableau 4 sur les dynamiques collaboratives autour de cette étape. Nous pouvons noter que seule la science des données est impliquée dans la réalisation des modèles, chargée de mettre en application l'ensemble des hypothèses que nous avions émises dans les étapes précédentes. L'évaluation se fait par la suite par une validation visuelle des hypothèses ou par une mesure formelle de la performance d'un modèle.

Service	Domaine d'application	Science des Données
	Représentation • Validation des hypothèses	**Calcul des écarts** • Mesure des performances **Représentation** • Explicite les modèles

Table 4 : Résumé des étapes 4 et 5

Étape 6 : Discussion

La dernière étape de cette approche méthodologique consiste à prendre du recul sur les résultats générés pendant une itération afin de conclure sur les possibilités d'implémentation des solutions partielles développées ou sur les efforts à mettre en œuvre pour pouvoir poursuivre le projet. Cela consiste notamment à reboucler les résultats obtenus dans les deux étapes précédentes avec la définition du cas d'usage. Dans le cadre de notre étude, c'est lors de cette étape que nous avons pu construire la séquence d'itérations par laquelle nous sommes passés.

La première itération a montré ses limites dans le manque de qualité au sein des données vis-à-vis du cas d'usage ciblé. En effet, il nous a semblé difficile de pouvoir prédire des évènements avec une si faible représentativité des cas possibles de panne par l'approche que nous avons abordée ici. Le jeu de données ne fournissait pas les ressources nécessaires à l'élaboration d'une solution répondant à notre besoin initial dont les critères de réussite ont été clairement identifiés par le service. D'autre part, et même s'il ne nous était pas possible de conclure favorablement sur la caractérisation que nous avons faite d'une chaudière dans les données à ce stade, ce mode de représentation expose les besoins de notre système à manipuler les données afin de pouvoir être compatible avec une exploitation correspondant au cas d'usage, ici la classification. Ces conclusions nous ont alors poussés vers une approche orientée métier dans la seconde itération afin de répondre à notre besoin de connaissances plus importantes sur les pannes de nos appareils.

Lors de la deuxième itération, nous avons fait émerger des heuristiques du métier d'application de notre fonction afin de pallier aux manques de vérité terrain et de mieux percevoir la faisabilité technique de notre système. Cette étude a montré des axes d'amélioration pour la prévention de certains modes de défaillances ciblés de nos appareils. Ces améliorations se concrétisent dans la procédure de prétraitement de nos jeux de données et par l'ajout de nouvelles mesures reflétant un comportement à risque de nos systèmes. Cependant, nos experts ont montré une forte réticence dans le cas de l'identification des causes liées à certaines pannes. Les comportements associés étant établis, les mesures que nous pouvons en faire ne nous sont pas disponibles dans notre jeu de données car les capteurs associés ne sont pas présents dans les appareils, tel qu'un capteur de pression ou de débit. Néanmoins, nous avons pu constater que le comportement général de nos appareils devrait être affecté. C'est pourquoi, dans une troisième itération, nous avons focalisé notre attention sur la faisabilité de l'identification d'une surchauffe par une approche plus orientée vers les systèmes de données.

Enfin, lors de la dernière itération, nous nous sommes redirigés vers une approche orientée données en sélectionnant un cas d'usage problématique mis en évidence par nos experts et où nous étions à l'initiative de la création de notre propre jeu de données, représentatif de l'exploitation que nous souhaitions en faire. Cela nous a démontré la faisabilité d'une fonction complexe pour les experts à l'aide des techniques de la science des données. Nous avons alors pu déterminer que les indicateurs définis dans la première et la seconde itération étaient pertinents.

Nous pouvons alors conclure sur cette dernière étape en résumant l'ensemble des éléments que nous venons de voir dans le tableau suivant. Nous pouvons notamment noter un retour du service, chargé de faire une évaluation des résultats obtenus vis-à-vis du cas d'usage afin de statuer sur la possibilité d'intégration des éléments développés ici dans le système d'internet des objets global.

Service	Domaine d'application	Science des données
Evaluation • Faisabilité de la fonction **Système technique** • Cotation de l'effort nécessaire pour l'implémentation de la fonction	**Indisponibilité des données** • Limitation des approches possibles	**Manque de vérité terrain** • Réduction du périmètre

Table 5 : Résumé de l'étape 6

4.2.5. Analyse des résultats et conclusion

Avec cette seconde expérimentation, nous avions pour objectif de démontrer la pertinence d'une approche intégrée tenant compte de la science des données dans le développement agile de notre système d'objets connectés. Effectivement, la réalisation de fonctions intelligentes fait appel à des concepts technologiques et méthodologiques non conventionnels dans le cadre d'une conduite de projet standard. C'est pourquoi nous souhaitions ici mettre en évidence les dépendances entre les différentes couches technologiques de notre système ainsi que l'ensemble des connaissances métiers concernées par ces fonctions. Cela constituant notre seconde hypothèse de résolution, nous présenterons ici les résultats que nous avons obtenus suite à l'application de la démarche que nous proposons, nous discuterons des intérêts et des limites de notre approche avant de conclure sur la validation de notre seconde hypothèse.

Résultats

Tout d'abord, suite à l'application de notre démarche de conception au travers de ces quelques itérations, nous avons pu nous construire un ensemble d'attributs représentatifs du comportement de notre produit vis-à-vis des attentes que nous pouvions en avoir, plus particulièrement dans le cas de l'amélioration de la maintenance. La construction de ces attributs a su, non seulement nous aider à définir le modèle que nous devions obtenir de nos produits au sein du système, mais cela a également permis de démontrer la faisabilité de nos fonctions, ce qui était attendu dans cette phase du projet.

Nous pouvons noter que ces itérations ont eu des impacts positifs sur les deux parties de notre système. Concernant la partie physique du produit, ces attributs une fois définis ont étés caractérisés et spécifiés afin de pouvoir être calculés localement, permettant ainsi de limiter la quantité d'informations à transmettre dans nos canaux de transmission de la donnée. Puis, du point de vue de la plateforme logicielle, cette démarche de conception nous a permis de mieux percevoir l'utilisation que nous souhaitions faire d'une plateforme de découverte de connaissance dans notre système en définissant les modèles répondant à notre besoin tout en intégrant leur création et leur interrogation dans les scénarii d'usage.

Intérêts et limites

Au regard des résultats obtenus, démontrant les dépendances entre notre système de traitement de la donnée et les couches produits et logicielles, nous pouvons remettre en perspective la proposition d'approche que nous avons faite afin d'en distinguer les intérêts et les limites. Effectivement, nous avons repris la démarche de conception issue de la science des données sur laquelle nous avons introduit l'ensemble des acteurs et des connaissances relatives à la fonction à l'étude.

Cela nous permet alors de mettre en lumière les dynamiques que l'introduction de ces disciplines induit dans le cycle de conception. Pour cela, nous pouvons positionner sur le tableau 6 l'ensemble des apports liés aux trois disciplines que nous avions préalablement sélectionnées.

Etapes	Service	Domaine d'application	Science des données
1	**Service *blueprint*** • Définition du cas d'usage • Fonctions • Usage **Modèle économique** • Priorités	**Heuristique** • Définition du cas d'usage **Retour d'expérience** • Priorités	**Retour d'expérience** • Priorités
2		**Carte système** • Source les données • Explicite • Filtre	**Statistique générale** • Représentativité
3		**Indicateurs** • Extraction de caractéristiques • Formatent **Plan d'expérience** • Comble des manques de représentativité	**Techniques de prétraitement** • Extraction de caractéristiques • Formatent
4		**Heuristique** • Définition d'une approche	**Technique d'apprentissage** • Définit une approche
5		**Représentation** • Validation	**Calcul des écarts** • Mesure des performances **Représentation** • Explicite les modèles
6	**Evaluation** • Faisabilité **Système technique** • Implémentation	**Indisponibilité des données** • Limitation des approches possibles	**Manque de vérité terrain** • Réduction du périmètre

Table 6 : Résumé des dynamiques collaboratives entre les métiers dans l'expérimentation 2

Tous d'abord, la discipline de la science des données représente les connaissances des différentes techniques de mise en œuvre d'algorithmes d'apprentissage. Son principal apport dans notre cas d'application réside dans le choix d'un traitement répondant à notre besoin et les indications qui en découlent afin de traduire et orienter notre problème de conception. Cela nous a permis de sélectionner un axe d'investigation, puis, par la suite de sélectionner un cadre de représentation de notre problème orientant et structurant nos réflexions. Enfin, nous pouvons noter que les principaux freins rencontrés par les experts de ce domaine étaient relatifs à la qualité et à la représentativité des données. La conduite de notre troisième itération démontre la nécessité d'une connaissance complémentaire sur l'état du système sans laquelle nous ne pouvons pas construire notre raisonnement. Cette connaissance, que nous

appellerons la « vérité terrain », nous permet de constituer une base de faits sur laquelle peut s'appuyer notre système pour apprendre à reconnaître les différents états du système qui nous intéresse.

Ensuite, nous avons fait intervenir les métiers du domaine d'application représentant l'ensemble des connaissances pouvant être en lien avec la fonction développée. Ses apports, que nous avons pu identifier dans le cycle de développement, sont nombreux. Tout d'abord dans les phases de définition et de description, où nous devons comprendre le cadre d'application et le système auxquels les données sont attachées. Ce transfert, facilité par la création d'un modèle descriptif de processus ou de système, permet à la fois de faire comprendre les leviers dont nous disposons pour maximiser la génération de valeur et également ce sur quoi nous pouvons nous baser pour la réaliser. Ensuite, nous avons pu voir que la réalisation de notre fonction peut projeter ces experts du métier dans un univers de problèmes connus d'où nous pouvons extraire de la connaissance afin de la transcrire et de la valoriser directement dans notre système. Cependant, des limites apparaissent dès lors que notre problème ne repose pas sur des faits maîtrisés, ici sur des phénomènes physiques ou mécaniques.

Enfin, nous avons la discipline du service qui représente la part organisationnelle des connaissances apportées au cycle de développement avec une vision globale de la fonction sur laquelle nous nous focalisons. Ces apports ici sont plus minces et sont principalement liés aux efforts de définition du cas d'usage où il peut être ici une sorte de catalyseur, permettant de centrer nos efforts sur l'essentiel en dégageant des leviers d'action principaux et des priorités. L'aspect service est également moteur pendant les phases de discussion en fin de cycle afin de capitaliser sur les découvertes mises en avant pendant l'itération et de réagir sur la planification de l'itération prochaine.

Ensuite, concernant les limites de cette approche, nous pouvons remarquer que, malgré le fait de se situer dans la phase de conception détaillé de ces fonctions, nous n'avons pas été en mesure de produire un algorithme de reconnaissance de panne. Les données de sortie de notre approche concernent la mise en conformité de l'ensemble de notre système à la réalisation future de la fonction intelligente, qui pourra être faite par la suite, basée sur un historique dont nous disposerons une fois notre produit sur le marché.

Nous pouvons clore cette analyse en remarquant que nous avons opté pour une intégration des métiers à cause du fait que nous n'avions pas de données disponibles afin mener un cycle de conception classique de fouille de données. Cependant, nous pouvons noter qu'à deux reprises nous avons été débloqués par l'apport de jeux de données lors de ces itérations. Nous pouvons alors considérer que ces réflexions sont poussées par le domaine d'application et structurées par le service, mais elles restent néanmoins soutenues par l'apport de jeux de données permettant d'avancer au travers de nombreux prototypes de solutions intermédiaires.

Conclusion

Au regard de cette analyse, nous pouvons conclure sur cette expérimentation par la validation de notre seconde hypothèse. Effectivement, nous avons démontré ici les dépendances et les impacts que peuvent avoir les technologies de traitement de données sur l'ensemble de notre système. D'autre part, nous avons également démontré l'intérêt d'aborder une approche pluridisciplinaire afin de faire face à ce problème de conception en exposant les dynamiques que chaque métier peut apporter.

Cependant, nous pouvons nous interroger à présent sur la généricité de cette approche. En effet, pour ces deux premières expérimentations nous avons traité un problème de conception spécifique pour lequel nous avons mis au point notre propre approche. Il est alors légitime de se questionner sur la dépendance que peut avoir notre support méthodologique au contexte dans lequel nous l'appliquons. C'est donc pourquoi nous testerons alors cette approche dans une troisième et dernière expérimentation dans un cadre nouveau.

4.3. Expérimentation 3 : Conception générique de fonctions intelligentes par le service

Suite à la seconde expérimentation, nous avons pu valider notre seconde hypothèse, à savoir que la science des données doit être prise en compte parallèlement au développement du reste du système afin d'assurer la faisabilité des fonctionnalités intelligentes. Cela nous permet d'ores et déjà de répondre à notre problématique de recherche en validant l'approche méthodologique que nous proposons. Néanmoins et dans le cadre de cette dernière expérimentation, nous souhaitons conforter nos résultats en testant la résilience de notre modèle de développement de fonctions intelligentes au contexte dans lequel il prend place. Nous présenterons donc ici la mise en application de notre modèle de conception dans le cadre d'un projet différent, tel que le montre la figure 55.

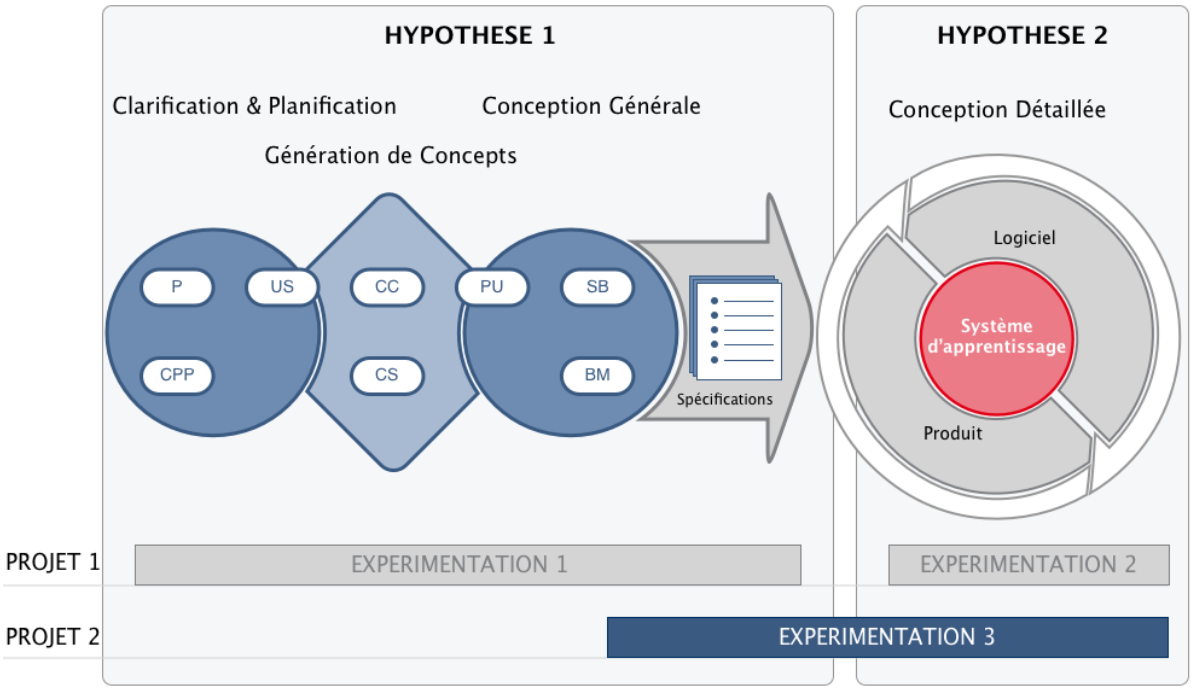


Figure 55 : Positionnement de l'expérimentation 3

Nous présenterons ici plus en détail les objectifs de cette expérimentation. Nous discuterons ensuite la démarche adoptée, nous présenterons le contexte d'application avant de détailler son application puis de conclure par une discussion sur les résultats obtenus.

4.3.1. Objectifs

Cette dernière expérimentation a pour vocation d'éprouver notre modèle de conception de fonctions intelligentes dans un cadre de projet différent, et cela afin de valider sa généricité. Cela nous permettra d'une part de conforter la validation de notre seconde hypothèse, puis d'autre part, cela nous permettra également de montrer d'autres facettes liées à l'intérêt d'aborder une approche intégrée dans le développement de fonctions intelligentes.

L'objectif est donc ici fortement similaire à l'objectif de la deuxième expérimentation. Soit, la mise en application de la démarche de conception issue de la science des données dans un cadre déterminé en faisant intervenir différents domaines d'expertise. Le but étant d'observer les apports des différents

domaines de compétence face au problème de conception ainsi que les freins auxquels ils sont confrontés, puis d'analyser l'intérêt d'impliquer toutes ces connaissances pendant le développement, à quel moment et dans quel but. Cela doit donc nous permettre de renforcer la validation de notre seconde hypothèse de résolution, portant sur l'intérêt d'intégrer l'ensemble des métiers dans la conception détaillée des fonctionnalités intelligentes et de montrer la généricité de notre raisonnement. Cette application nous permet également de révéler de nouvelles facettes des apports de chacun des métiers au cours du développement.

D'autre part, la mise en application dans ce nouveau projet nous permettra de montrer la complémentarité de nos hypothèses. Effectivement, nous commençons ici le projet de conception à partir de la première idée sans formalisation claire du besoin. Nous pourrons alors, par ce fait, analyser l'impact du domaine du service sur la mise en place d'une fonctionnalité intelligente avec plus de recul que lors de la deuxième expérimentation.

4.3.2. Démarche

Afin de servir ce but, nous appliquerons donc notre modèle méthodologique incrémental et itératif relatif à la conception de système de découverte de connaissance (figure 56) dans le cadre d'un nouveau projet industriel. La démarche que nous suivons dans cette expérimentation est strictement identique à celle que nous avons présentée lors de l'expérimentation 2 (cf. 4.2.2). Nous impliquerons dans le développement l'ensemble des experts des métiers impliqués dans la conception. Nous distinguerons les métiers du domaine du service, de la donnée et du domaine d'application. Nous observerons alors de la même manière les influences des différents métiers pour chacune des phases de notre modèle de conception.

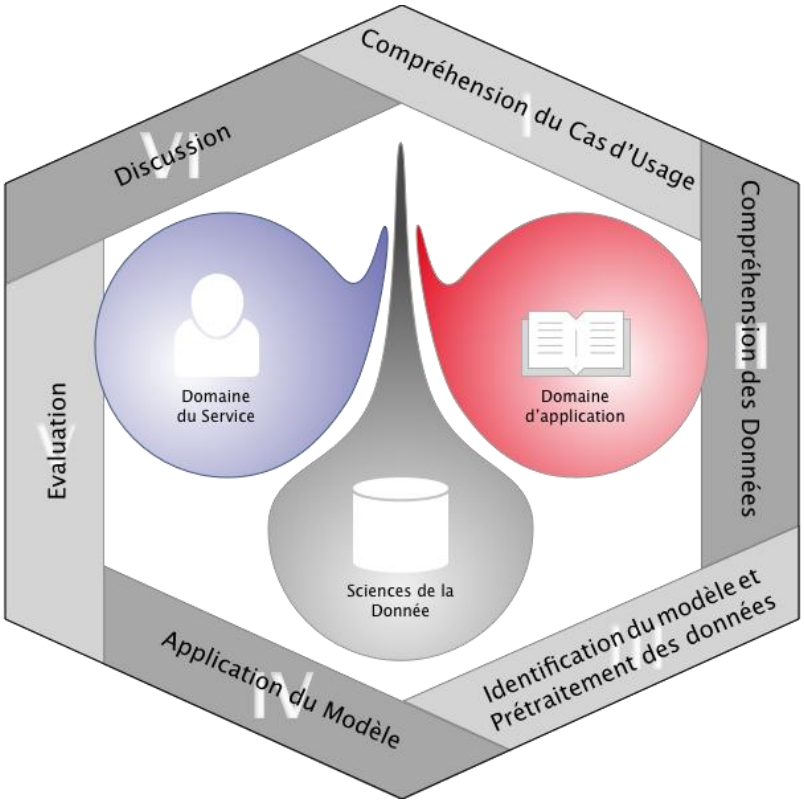


Figure 56 : Démarche de conception des expérimentations 2 et 3

4.3.3. Cadre d'application et contexte

Avant de présenter l'application de notre modèle, nous souhaitons tout d'abord présenter ce nouveau projet et son cadre d'application. Puis, au regard de la nature itérative de notre approche, nous présenterons également la succession des différentes itérations que nous avons pu faire dans ce cadre afin de pouvoir nous en abstraire par la suite, et de présenter l'application de notre modèle au travers des axes de lecture que nous proposons.

Contexte : Industrie 4.0

Nous avons basé l'ensemble de notre démonstration précédente dans le cadre d'un projet visant à créer du service autour de produits connectés. Cette première démarche de valorisation de l'internet des objets au sein de l'entreprise a eu de nombreuses retombées, notamment la mise en perspective de ces technologies dans un cadre plus large avec l'industrie 4.0. En référence avec les révolutions industrielles avec l'apparition des machines-outils en 1784, la production de masse des années 1870 et l'usage des technologies de l'information et des premiers contrôleurs dans les années 1970, le 21ème siècle est porté par la mise en réseau et la démocratisation de l'internet. Suivant cette perspective, le projet d'industrie 4.0 vise à repenser le fonctionnement de l'entreprise au travers de l'interopérabilité de ses services et de ses systèmes d'information. Ce projet constitue donc un vaste chantier, dont notre chaudière connectée représente l'une des briques.

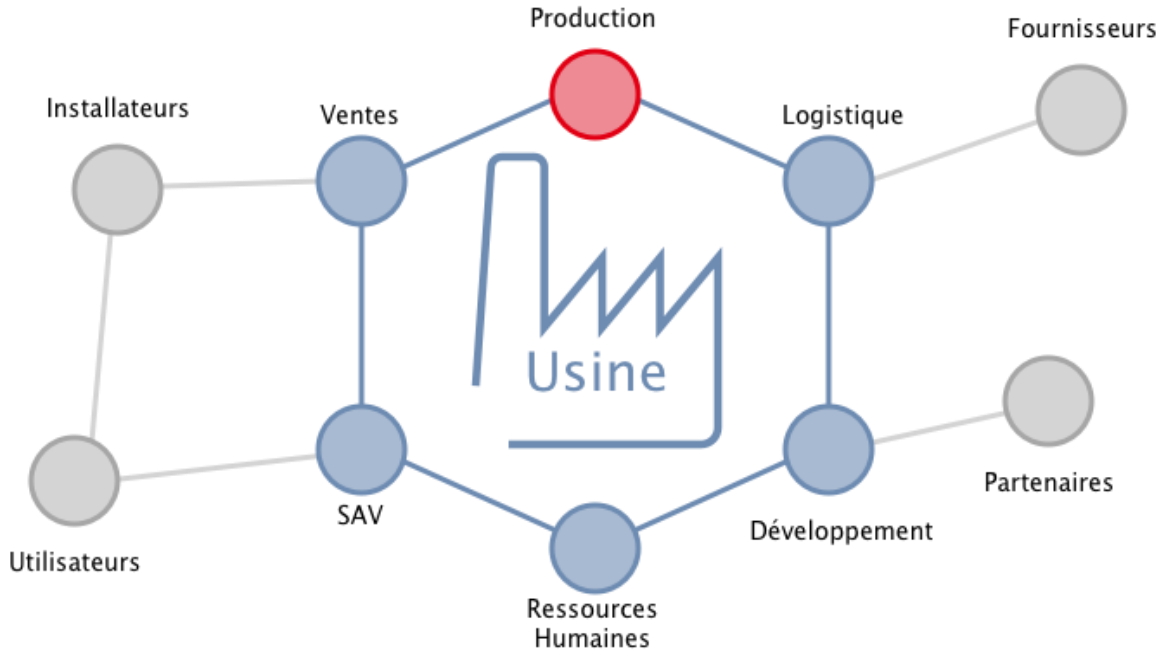


Figure 57 : Illustration du cas d'usage de l'expérimentation 3

Loin de nous la prétention de conduire une révolution par cette expérimentation, nous nous positionnons dans ce cadre avec l'idée suivante : utiliser les données générées lors de la production, pour améliorer les processus de production. Par cette idée simple nous souhaitons entrer dans le problème afin de généraliser les découvertes que nous avons pu faire dans le projet précédent et dans l'idée de commencer à construire une infrastructure technique permettant de supporter nos ambitions. La restriction du périmètre de notre étude à un service nous permet alors de prendre en considération les principales problématiques qui lui sont propres et d'identifier les ressources qu'il génère. Cela afin de les valoriser pour son propre usage, avant d'envisager les opportunités présentes dans les interactions avec d'autres services.

Des premiers efforts ont déjà été déployés concernant l'infrastructure qui pourrait accueillir nos futurs services par la digitalisation des supports des processus de l'entreprise. Nous retrouvons sur un ensemble d'écrans, de totems ainsi qu'une suite logicielle en cours de développement visant entre autre à collecter et partager les informations issues des plans d'action pour l'amélioration continue de l'ensemble des services. Ce qui constitue, en plus de l'ensemble du système de production de l'usine, notre voie principale de distribution de nos services.

Au regard de la nature du projet, nous considérons ici les trois mêmes pôles d'expertise :

- L'expertise en science des données, portée par nous-mêmes ainsi que par deux ingénieurs en science des données mandatés pour nous aider dans le cadre de la deuxième itération.
- Les experts du domaine d'application, ou les experts métier qui sont cette fois issus du service de production, avec les méthodes, la qualité et la maintenance des lignes.
- L'expertise en service est portée par le service informatique de l'entreprise qui gère l'ensemble de ce projet et qui maîtrise l'infrastructure et son usage.

Protocole :

Cette expérimentation s'est alors déroulée autour de trois itérations comme le montre la figure 58, dont nous pouvons préciser le contenu et le séquencement de la manière suivante.

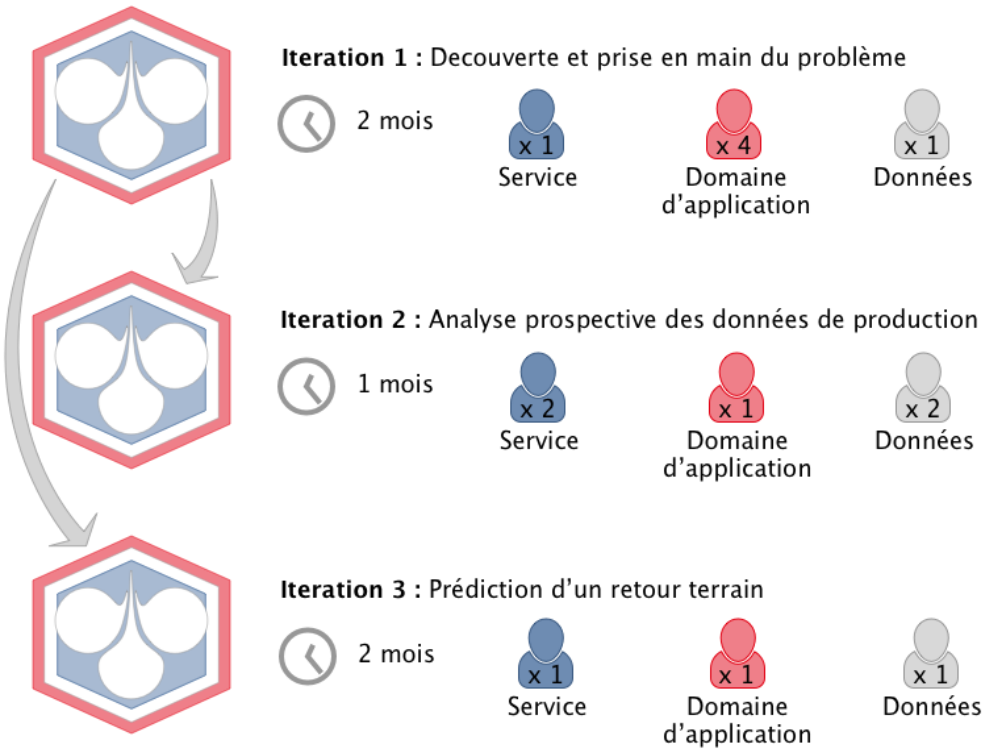


Figure 58 : Protocole de l'expérimentation 3

Itération 1 : Découverte et prise en main du problème

Le projet a été initié au cours de cette première itération à la lumière de deux faits. La mise à disposition d'une copie de la base données de production contenant l'ensemble des mesures faites en ligne lors de l'assemblage de nos appareils, et la volonté du domaine d'application d'améliorer le contrôle de la qualité en ligne, concrétisé par la fourniture d'un heuristique de mesure et d'évaluation de la qualité propre au groupe Bosch. Ces travaux ont été réalisés sur une période de 2 mois avec le concours de six

intervenants : un responsable du service, un responsable des données et quatre représentants du domaine d'application d'horizons divers. Parmi ces derniers, nous retrouvons le responsable qualité, le responsable de la production de l'usine, le responsable de la maintenance des lignes de production ainsi que le responsable des méthodes de production.

Nous avons alors ici mis en œuvre cette procédure de contrôle des processus afin de fournir au domaine d'application un tableau de bord exposant l'évolution de la stabilité et de la capabilité des différents postes constituant les lignes d'assemblage. Nous nous sommes alors attachés à mettre en évidence les problèmes que nous pouvions identifier sur chacun des postes. Ce livrable a alors été présenté aux représentants du domaine d'application qui ont conclu à une impossibilité d'exploiter les informations que nous leur apportons.

Bien que cela nous amène à remettre en cause notre proposition, la présentation d'une première approche de valorisation des données de production a été source d'inspiration pour ces experts du domaine d'application qui ont alors pu nous rediriger vers un besoin plus clairement exprimé. D'une part, l'information ainsi proposée, trop complexe et trop peu lisible, ne correspond pas aux usages que nous pouvions attendre d'un tel système. Ce qui nous a poussés à nous refocaliser sur l'usage dans le cadre de la seconde itération. D'autre part, l'ouverture de la discussion sur les voies d'exploitation de données conduisant à une amélioration notable et concrète du fonctionnement de l'usine nous a conduit à l'identification de fonctions potentielles du système portant le plus de valeurs, nos priorités. Nous avons ici identifié une voie d'amélioration possible en croisant deux sources de données, entre les données issues de la production et les retours terrains dans l'optique de pouvoir prédire ces derniers en sortie de ligne, ce qui constitue l'orientation de notre troisième itération.

Itération 2 : Analyse prospective des données de production

Comme nous venons de l'exprimer, la première itération nous oriente vers une reconsidération des usages que nous pouvons faire de ce type de système au sein de notre service, constituant l'objectif principal de notre deuxième cycle de conception. L'itération s'est déroulée sur une période d'un mois avec une équipe de cinq participants : deux experts du service, un expert du domaine d'application et deux experts de la science des données. Nous avons ici fait appel à une entreprise externe spécialisée dans la fouille de données dans le cadre d'un projet décrit plus en détail dans l'Annexe D. L'objectif était de clarifier les usages, notamment par les voies de restitution de l'information, afin de les utiliser comme points de départ pour conduire les choix de traitement. L'expertise en science des données a donc été ici mise à l'épreuve en tant que moteur de proposition de solutions de traitement compatible avec l'usage attendu du système.

Cette itération s'est ici soldée par la proposition d'une voie nouvelle de traitement satisfaisant les besoins du cas d'usage, suite à de nombreux échanges avec les experts du service et les experts du domaine d'application. De plus, l'ensemble des efforts réalisés pour mettre en place ce traitement, notamment dans le cadre de la préparation du jeu de données, a inspiré l'ouverture du sujet vers de nouvelles fonctions pour le système.

Itération 3 : Prédiction des retours terrain

Pour finir, cette troisième et dernière itération a été initiée suite à la confrontation de nos premiers résultats au regard de l'expertise du domaine d'application, qui nous a ouvert à de nouveaux cas d'usage basés sur le croisement des données que nous avons pu manipuler avec les données de la qualité, comportant les retours terrain de nos appareils. Cette itération s'est déroulée sur une période de quatre mois avec trois participants : un expert du service, un expert du domaine d'application et un expert de la science des données. L'idée principale derrière ce croisement est de pouvoir identifier un appareil qui serait susceptible de comporter un problème qualité et de conduire à un retour terrain avant sa sortie de

l'usine. Ce cas d'usage est hors du périmètre que nous nous sommes préalablement fixé, en étendant nos données à un autre service. Néanmoins, celui-ci a été identifié comme étant rapidement applicable, facilement démontrable et dont les retombées pouvaient être fortement bénéfiques pour l'entreprise.

Le but de cette itération a donc été de retravailler notre jeu de données en y intégrant celles de la qualité, puis d'appliquer des algorithmes d'apprentissage, par la classification, afin d'extraire un modèle de prédiction des retours terrain. L'idée étant de pouvoir interroger ce modèle pour chaque produit sortant de la production afin de détecter ceux pour lesquels peut persister un problème qualité. Suite à cette détection, nous devons mettre en place des mesures de prise en charge de cette alerte afin d'intégrer ce nouvel outil dans le système en place d'amélioration continue de l'usine.

4.3.4. Application

Nous pouvons désormais présenter la mise en application de notre démarche méthodologique suivant la séquence d'étapes définies par notre modèle. Nous analyserons alors quels ont été les efforts déployés, les connaissances sollicitées et les moyens de formalisation mis en œuvre dans chaque étape. Nous faisons alors ici abstraction de la dimension temporelle réelle du projet en traitant l'ensemble des trois itérations simultanément afin de nous focaliser sur les critères d'analyse que nous avons mis en avant dans la présentation de notre démarche.

Étape 1 : Compréhension du cas d'usage

La première étape de notre approche méthodologique consiste à prendre connaissance du cas d'application dans lequel nous nous situons. Cela permet alors de définir le périmètre fonctionnel et les critères de réussite de notre projet.

Or, partant ici d'une idée très générale et ne disposant pas pour cette expérimentation d'une évaluation complète du service et des voies d'exploitation potentielles de l'information, cette phase a été négligée lors de notre entrée dans le problème. Cette négligence représente par ailleurs la principale limitation de notre première itération, nous amenant à reconsidérer sa formalisation par la suite.

L'idée principale de cette expérimentation est d'améliorer la qualité de la production et d'apporter des informations complémentaires afin de supporter l'amélioration continue de l'usine. Cette expression du besoin est extrêmement vague, c'est pourquoi, nous pouvons la raffiner au travers des moyens mis en œuvre pour mesurer cette performance en production. Cet indicateur de qualité de production est appelé « efficience », mettant en perspective les capacités de la production à atteindre un objectif relativement aux moyens déployés pour y parvenir. Cet indicateur constitue donc un équilibre entre efficacité et rentabilité. C'est au travers de cette perspective d'amélioration de l'efficience de l'usine que notre recherche de solutions doit prendre place.

D'autre part et sans le formaliser autant que lors de la première expérimentation, nous nous sommes intéressés à l'analyse de l'activité des principaux utilisateurs du système. À savoir les opérateurs en ligne, les superviseurs des lignes et les services support tels que la qualité. Nous pouvons déduire de cette analyse que :

- Les opérateurs en ligne se réunissent chaque matin avec leurs superviseurs afin de faire un point sur la journée passée et évoquer les actions en cours pour l'amélioration continue de l'usine. Ce point journalier constitue notre principal point d'entrée afin de leur transmettre de l'information complémentaire pour favoriser l'amélioration continue.
- Les superviseurs sont continuellement à l'écoute des problèmes associés à la production. Leur principale fonction est de supporter et réguler l'ensemble des incidents pouvant survenir sur une ligne.
- Enfin, les services supports n'ont pas de temps prédéfini et de tâches récurrentes associés à la surveillance et au maintien du bon fonctionnement de la production. Néanmoins, ces derniers peuvent être amenés à prendre du temps sur requête au sein d'un projet d'amélioration continue.

Enfin, pour compléter cette analyse d'activités, nous avons pu identifier différents outils pouvant nous servir de point de contact. Notamment dans le cadre des points journaliers d'amélioration continue, où les équipes se réunissent autour d'écrans représentés dans la figure 59 sur lesquels est mis à disposition un logiciel spécifique permettant d'effectuer le suivi des indicateurs de performance de l'usine et d'assurer le suivi des actions en cours. Cet outil logiciel étant à l'initiative du service informatique de l'usine, nous avons pour objectif de créer un système d'apprentissage pouvant par la suite y être intégré.



Figure 59 : Système de suivi des indicateurs de performance de l'usine

Par ailleurs, et cette fois plus en relation avec le cas d'application de la troisième itération, sur la prédiction d'un retour qualité, des outils sont également présents sur les lignes afin de suivre la production en temps réel. Ce logiciel de suivi peut alors être envisagé pour restituer de l'information sur la ligne, pendant l'assemblage du produit.

Nous pouvons alors conclure sur cette première étape avec le tableau 7. Principalement initié par le domaine d'application, la définition du cas d'usage s'est structurée autour de la mise en avant d'un heuristique de contrôle des processus de production. Par la suite, nous avons affiné la description de fonction en passant par le service et en définissant plus précisément l'usage que nous souhaitions d'un tel système d'exploitation des données de production.

Service	Domaine d'application	Science des données
Modèle de processus • Définition du cas d'usage • Usage attendu de la fonctionnalité	**Heuristique** • Définition du cas d'usage **Retour d'expérience** • Priorités	

Table 7. Résumé de l'étape 1

Étape 2 : Compréhension des données

La deuxième étape du processus vise à mieux comprendre les sources de données disponibles. Cette étape a été réalisée lors des deux premières itérations afin de mieux comprendre la substance sur laquelle nous pouvions travailler. Ce travail s'est avéré particulièrement important dans cette expérimentation sachant que notre périmètre fonctionnel est moins bien défini comparé à notre cadre expérimental précédent, une bonne compréhension des données est alors nécessaire afin de permettre aux acteurs de la conception issus des sciences de la conception d'être force de proposition dans les traitements et la valorisation de ces données.



Figure 60 : Exemples de manipulations réalisées en production et dont une trace est collectée dans les bases de données

Travaillant sur les bases de données utilisées et générées en production, nous nous sommes tout d'abord intéressés à l'exploration des champs et des différentes tables présentes dans la base. Puis, nous nous sommes chargés de mettre en lumière cette interprétation par la formalisation d'un modèle représentant les principales données au sein du processus d'assemblage de nos produits. Après discussion et validation de notre modèle de représentation par les experts du domaine d'application, nous avons alors illustré la constitution de notre système par le schéma de la figure 61.

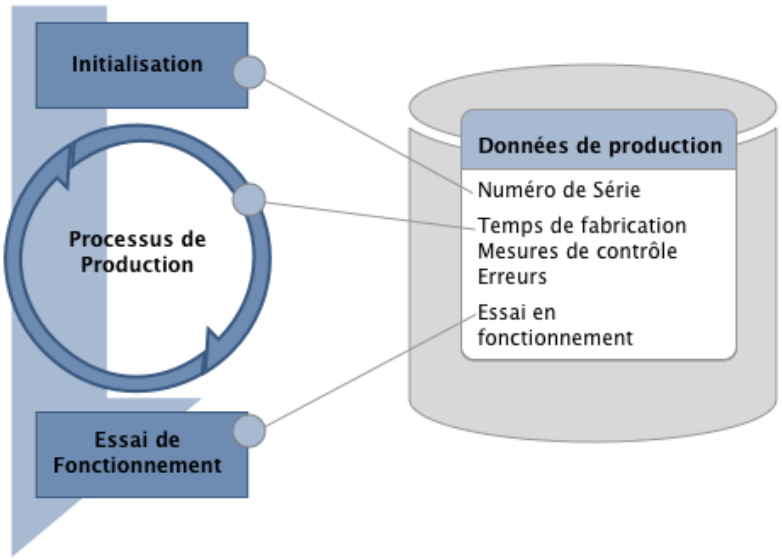


Figure 61 : Principe de collecte des données sur le processus de fabrication

Nous pouvons expliquer l'ensemble des données de la manière suivante, en suivant le cheminement de production d'un produit.

- La création d'un numéro de série constitue la première étape de la production. Ce numéro de série correspond à notre identifiant unique pour chaque assemblage en cours de production. De plus cette identifiant comprend de nombreuses informations, telles que la date de création, la ligne de production ou la référence produit.
- Ensuite, l'assemblage passe dans une suite de postes. Ces postes représentent des ateliers où un opérateur effectue une suite de transformations. Nous pouvons alors ici identifier la date de sortie d'un poste pour chaque assemblage ainsi que la séquence par laquelle les produits sont passés.
- Au sein de ces postes, la transformation est réalisée au travers d'un ensemble de tâches. Les difficultés rencontrées par un opérateur sur ces tâches peuvent transparaître au travers d'erreurs

qui sont remontées dans le système. D'autre part, certaines tâches sont constituées de mesures réalisées par l'opérateur afin de contrôler certains éléments techniques du système. Ces mesures sont alors qualifiées par une valeur et une validité, conditionnées par des bornes supérieures et inférieures.

- Enfin, le dernier poste par lequel passe notre assemblage est un banc d'essai. Ce banc permet de tester le bon fonctionnement du produit fini. Nous y enregistrons alors ici certaines constantes de fonctionnement comme les émissions de gaz, les temps de mise en sécurité, etc.

D'autre part, nous disposons également pour notre dernier cycle de conception des données de retour terrain de nos appareils. Ces données nous apportent de nombreuses informations sur les conditions d'installation du produit, l'installateur, la localisation ou encore la date et la cause de la panne.

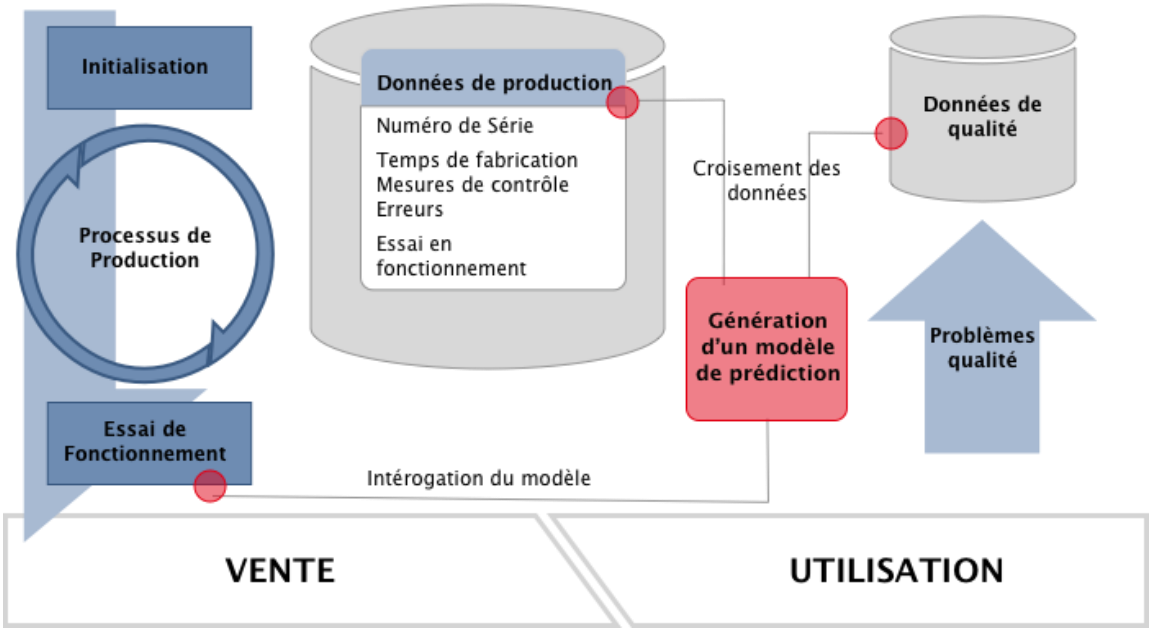


Figure 62 : Principe d'exploitation des retours terrain

L'idée étant pour cette itération de prédire l'arrivée d'un retour terrain lors de la fin de la sortie d'un produit en production (figure 62), nous croisons les bases de données à l'aide du numéro de série de chaque produit. La présence ou non d'un produit dans les données de qualité nous indique la présence ou non d'un défaut qualité observé pour ce produit. Cette dernière information représente alors l'état que nous souhaitons prédire du produit.

Pour conclure sur cette seconde étape, nous pouvons la résumer par le tableau 8. Nous remarquons que seul le domaine d'application n'a pas été impliqué ici.

Service	Domaine d'application	Science des données
	Carte système • Lister l'ensemble des données disponibles • Expliciter les données • Filtrer les données utiles	

Table 8 : Résumé de l'étape 2

Étape 3 : Identification du modèle et prétraitement des données

L'étape suivante consiste à identifier un modèle de traitement de la donnée, et de constituer la mise en forme de notre jeu de données dans le format nécessaire à l'application du modèle sélectionné. Une nouvelle fois, comme nous sommes passés ici par trois cycles itératifs abordant chacun une approche différente, nous dissocions nos trois itérations pour en présenter le contenu.

Notre première itération a été initiée par la découverte d'un heuristique commun au groupe Bosch décrivant des méthodes de contrôle de la qualité des processus. L'objectif de notre itération était alors d'adapter ces méthodes afin de les appliquer dans nos données de production. Le principe de cette méthode, dérivée de la méthodologie « six sigma », est d'évaluer la stabilité et la capabilité des processus sur la base d'un échantillon de mesures réalisées sur un produit fini. L'idée principale est donc ici d'adapter ces méthodes de calcul à l'ensemble de nos postes de mesure sachant que nous disposons ici de l'intégralité des mesures effectuées sur nos produits. Afin d'être plus clairs sur le principe de cette méthode, nous pouvons exprimer les deux critères de la manière suivante :

- La stabilité consiste à comparer, sur une période de mesure donnée, la valeur d'une moyenne mobile avec la dispersion générale de l'échantillon. Autrement dit, nous définissons des limites, basées sur la moyenne générale et l'écart type de notre échantillon, devant cadrer notre moyenne mobile dans le cas d'un système stable. Nous pouvons alors définir ces limites puis générer des alertes de stabilité en cas de dépassement.
- La capabilité consiste, elle, à observer le positionnement de notre échantillon considéré par sa moyenne et sa variance vis-à-vis des limites admissibles de notre processus. L'observation des valeurs calculées en capabilité nous montre les marges dont nous disposons dans notre processus de mesure au regard des tolérances de notre processus. De la même manière que pour la stabilité, nous pouvons ici établir des limites admissibles de mesure afin de générer des alertes de capabilité.

Disposant maintenant de l'ensemble des informations nécessaires à la mise au point du modèle, nous appliquons ici la mise en forme décrite par les formules représentatives de chacune de ces mesures, avec les limites associées. Nous sélectionnons alors une période afin de l'appliquer et d'exposer les résultats par la suite aux représentants du domaine d'application.

Concernant notre deuxième itération, la manière d'entrer dans le problème n'était pas définie en amont, ce qui s'est précisé lors d'un atelier d'une journée organisé avec des représentants des trois domaines de compétence, à savoir trois représentants du service, deux représentants du domaine d'application ainsi que l'équipe d'ingénieurs de la donnée. L'idée était ici de partir de la formulation du besoin, avec notamment l'expression de l'efficience comme ligne directrice, ainsi que du support de représentation des données disponibles afin de converger vers une approche de valorisation de la donnée nous permettant de tendre vers une application concrète aidant à l'amélioration continue de l'usine.

Cet atelier nous a permis de conclure sur l'importance des erreurs au sein du processus de fabrication et sur l'intérêt de pouvoir les prévoir et les anticiper. C'est pourquoi nous avons convergé vers un modèle de régression basé sur le nombre d'erreurs survenues par jour. Nous pouvons alors nous projeter dans un problème purement orienté données, avec la nécessité de réaliser un modèle de prédiction, reposant sur des séries temporelles représentant l'évolution du nombre d'erreurs par jour au cours du temps.

Cette approche a alors nécessité la mise en forme des données en agrégeant par jour le nombre d'erreurs survenues en segmentant notre échantillon par lignes de production et par postes. Les itérations rapides entre cette phase de prétraitement de l'information et les phases d'application et d'évaluation du modèle nous ont également conduits à une nécessité d'identifier et de filtrer certaines

valeurs aberrantes présentent dans notre jeu de données. L'équipe experte de la donnée a alors été force de proposition afin de trouver un moyen d'identifier ces valeurs aberrantes spécifiques à notre typologie de données, à savoir, une série temporelle. La procédure de filtrage, représentée par la figure 63, nous permet alors d'identifier des pics apparaissant comme des potentielles erreurs du jeu de données. Ces derniers peuvent représenter des phases de tests en interne lors de procédure de déploiement de nouveaux équipements en ligne. Ce filtrage a permis d'éliminer un biais important dans les prédictions du modèle comme nous le verrons dans les prochaines étapes.

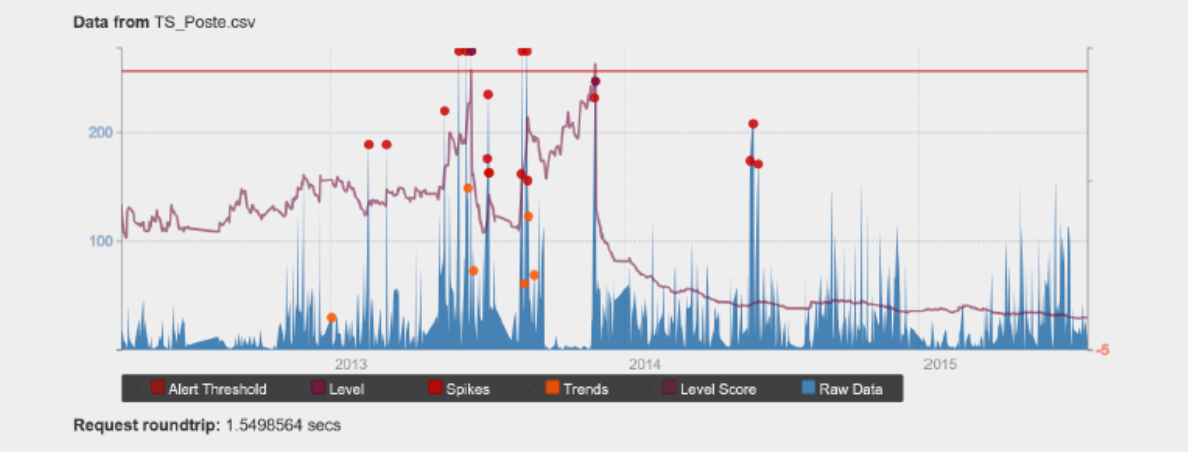


Figure 63 : Visualisation du nombre d'erreurs en ligne sur quatre ans et de leur prétraitement

Enfin, pour notre dernière itération, notre problème relève de la classification. Nous voulons ici prédire la classe de chaque chaudière en sortie d'usine vis-à-vis du retour sous garantie. Nous créons alors ici un tableau avec un individu par numéro de série sortant de l'usine et auquel nous lions les attributs le caractérisant. Parmi ces attributs, nous retrouvons d'une part, notre prédicteur, soit un booléen indiquant si ce numéro de série est présent dans notre tableau des retours terrain ou non, puis l'ensemble des caractéristiques le reliant à la production. Nous sélectionnons ici l'ensemble des informations présentes aux bancs d'essai afin de caractériser la qualité de production du produit.

Pour conclure sur cette étape, nous pouvons dégager les apports de chaque métier résumés dans le tableau 9. Nous pouvons remarquer que l'extraction de caractéristiques est uniquement faite par le domaine d'application, sur la base des méthodes d'analyse qui lui sont propres. La science des données a joué un rôle important dans la préparation des données avant la mise en application des algorithmes d'apprentissage, notamment avec l'application d'une technique de détection des valeurs aberrantes.

Service	Domaine d'application	Science des données
	Heuristique • Définition d'une approche **Indicateurs** • Extraction de caractéristiques pertinentes • Formatent les données en adéquation avec le cas d'usage	**Technique d'apprentissage** • Définit une approche **Techniques de prétraitement** • Filtrage des valeurs aberrantes • Formate les données en adéquation avec les techniques d'apprentissage et le cas d'usage

Table 9 : Résumé de l'étape 3

Étapes 4 et 5 : Création du modèle et évaluation

Après avoir mis en forme notre jeu de données, nous appliquons les traitements que nous avons sélectionnés puis nous évaluons les résultats obtenus au regard de l'application que nous souhaitons en faire. Une fois de plus nous dissocions l'ensemble de nos trois itérations afin d'en présenter les résultats séparément.

Pour notre première itération, où nous souhaitions adapter une heuristique de contrôle de la qualité à notre processus de production et nos données, nous avons réalisé un tableau de bord illustré par la figure 64 représentant de manière visuelle nos indicateurs. Le but de cette représentation était de pouvoir fournir un support nous permettant d'exposer le principe de calcul et de soulever les potentielles d'application de cette approche de traitement de la donnée auprès de la production pour leur fournir un outil d'analyse complémentaire. Ce support se compose donc d'une sélection des mesures prises au cours de la production, segmentées par lignes et par postes, où nous affichons les graphiques issus des calculs de stabilité et de capabilité de la mesure. Afin de faciliter la compréhension de l'impact de ces calculs nous avons mis en place une notion d'alerte suivant les différents critères décrits dans l'approche.

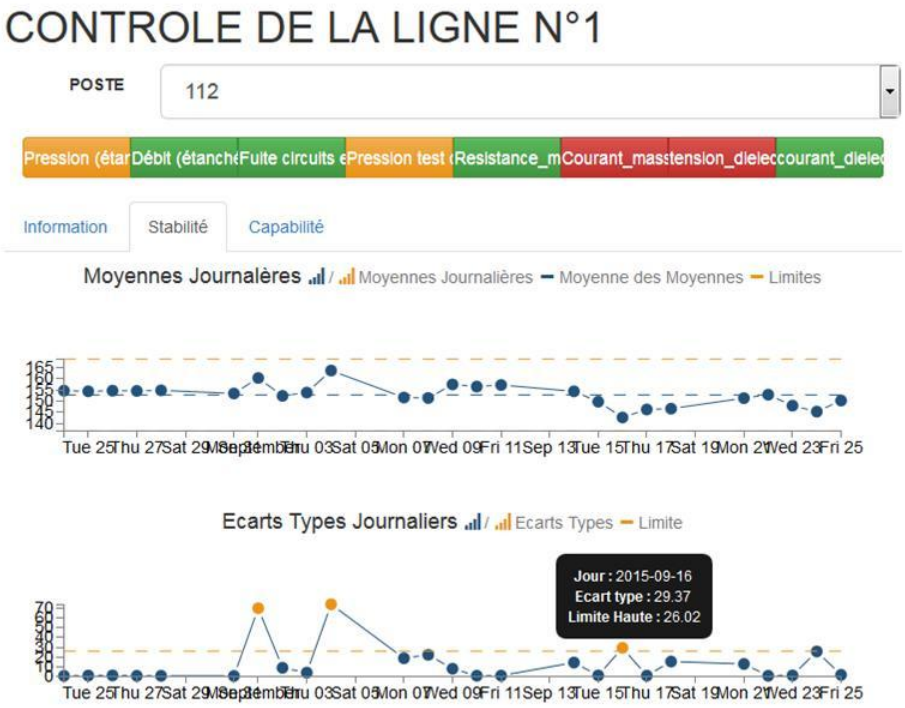


Figure 64 : Ecran du tableau de bord proposé dans le l'expérimentation 3

Suite à la présentation faite aux représentants du domaine d'application, ici les responsables de la production, de la qualité et de la maintenance des lignes, nous avons pu évaluer cet outil potentiel et conclure à sa non applicabilité pour les opérateurs. Nous pouvons expliquer cela par plusieurs remarques ayant étés faites. D'une part, cet outil a été jugé trop complexe, avec trop d'informations présentes pour les opérateurs de ligne, nécessitant trop de temps pour les déchiffrer et analyser les causes. Néanmoins, les termes utilisés ici, avec les mesures de stabilité et de capabilité étaient familières au domaine d'application en s'adressant à des problématiques concrètes et bien intégrées. Par la suite, notre approche a essuyé des remarques sur le potentiel d'analyse et l'interprétation que nous pouvions faire des résultats. Le processus de fabrication repose dans son intégralité sur des manipulations manuelles où les opérateurs en difficulté se soumettent souvent à des boucles d'essai-erreur dans les

prises de mesures. Ce phénomène pollue notre jeu de données avec des mesures à vide ou des mesures d'étalon n'étant pas clairement identifiées dans le jeu de données. Ceci nous amène à conclure que la représentativité des mesures se retrouve biaisée dans le cadre d'une analyse de stabilité et nécessite la mise en place d'un filtrage sophistiqué afin de pouvoir être interprété. Enfin, nos experts ont également pointé le manque de proactivité de l'approche, ne permettant pas aux opérateurs de se projeter dans l'amélioration à court terme de leur jour ou semaine de travail à venir.

La seconde itération nous a également menés à la fourniture d'un tableau de bord présentant la démarche mise au point et les résultats obtenus dans la prédiction des erreurs survenues par poste de fabrication. Cette représentation illustrée par la figure 65 montre la précision avec laquelle nous pouvons prévoir le nombre d'erreurs survenues en ligne sur un poste de fabrication sélectionné. L'évaluation se fait ici dans le cadre de l'application du modèle pour une période connue de production et où la mesure de l'écart entre la prédiction et la valeur réelle constitue notre indicateur de précision. Au-delà de cela, une analyse fine du modèle montre également sa faculté à restituer les pics d'erreur à venir dans la production. Ces pics impactant le plus la production, leur prédiction constitue alors une forte valeur ajoutée de notre système.



Figure 65 : Visualisation des résultats liés à l'application du modèle de prédiction du nombre d'erreurs en ligne

Nous pouvons d'autre part noter sur la réalisation de ce modèle que de nombreuses itérations ont pris place entre l'étape de prétraitement des données et cette étape d'application et de validation du modèle. Les premières applications du modèle nous on fait noter la présence d'un biais important entre les valeurs prédites et les valeurs réelles. C'est alors par la mise en place d'un prétraitement de filtrage des pics, apparaissant comme des anomalies et présentés dans l'étape précédente, que la précision du modèle a pu être améliorée progressivement jusqu'au résultat présenté.

Enfin, concernant notre dernière itération, nous avons ici appliqué un ensemble d'algorithmes d'apprentissage en classification afin de pouvoir prédire le retour terrain d'un appareil au moment de sa sortie de ligne de production. Un peu de la même manière que pour le cas précédent, nous appliquons

les mécanismes d'apprentissage sur une partie des données afin de créer le modèle de classification, que nous appliquons par la suite sur la seconde partie des données que nous avons laissées de côté. À partir de la comparaison entre les prédictions faites par notre modèle avec les classes réelles de nos individus, nous pouvons alors évaluer les performances de celui-ci. Cette évaluation peut alors être synthétisée dans une matrice telle que celle que de la figure 66, la matrice de confusion. Ce cas représente l'évaluation d'un modèle de SVM appliqué à un jeu de données dans lequel nous retrouvons à proportion égale des cas présentant un retour terrain ou non. Nous avons entraîné notre modèle sur trente pourcents des données, et testé sur les soixante-dix pourcent restants. Nous y observons une précision alors suffisamment importante pour que ce modèle soit présenté au domaine d'application et envisager sa mise en production.

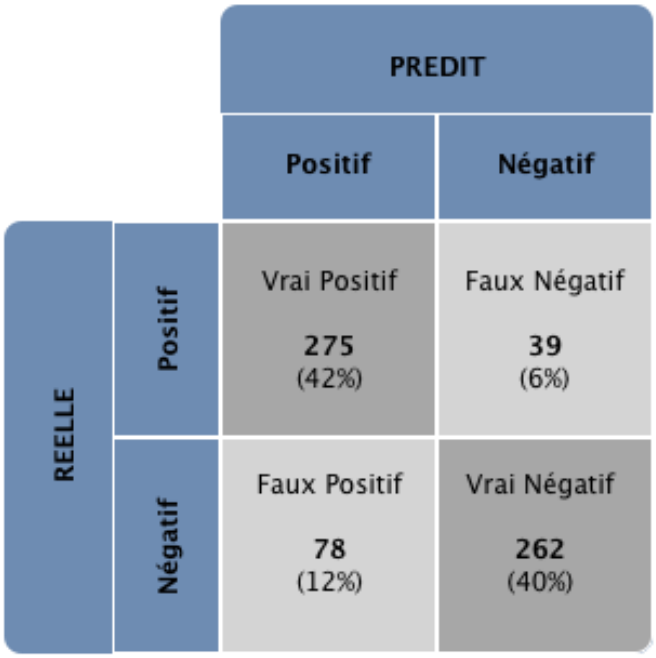


Figure 66 : Matrice de confusion pour l'évaluation du modèle de prédiction des retours terrain

Pour conclure sur cette étape, nous pouvons exprimer les dynamiques collaboratives comme les présente le tableau 10. Une fois de plus, la science des données joue un rôle prédominant dans cette étape avec la mise en application de l'ensemble des algorithmes que nous avons sélectionnés. Elle a aussi pour responsabilité l'évaluation formelle des performances des modèles. Le domaine d'application pour sa part a eu la possibilité de faire une évaluation visuelle du comportement des indicateurs, ou attributs, que nous avons sélectionnés. D'autre part, la génération d'un premier prototype dans notre première itération a permis l'évaluation des usages également.

Service	Domaine d'application	Science des données
	Représentation • Validation des hypothèses **Prototype** • Validation des usages	**Calcul des écarts** • Mesure des performances **Matrice de confusion** • Mesure des performances **Représentation** • Explicite les modèles • Confirme les performances

Table 10 : Résumé des étapes 4 et 5

Étape 6 : Discussion

Suite à cette présentation des résultats, la dernière étape consiste à confronter nos résultats à la réalité industrielle afin de décider des suites à donner au projet. C'est donc par cette phase, une nouvelle fois, que s'est construite la suite d'itérations que nous avons suivie. Nous pouvons alors décrire l'ensemble des remarques et décisions faites suite à chaque cycle de conception.

Lors de la première itération et suite à la confrontation de nos premiers résultats avec l'expertise du domaine d'application, nous avons conclu à l'impossibilité de mettre en application notre tableau de bord dans un cadre industriel. Ces conclusions sont principalement portées par l'inadéquation de notre proposition avec les usages que nous pouvons attendre d'un tel système d'exploitation de la donnée. Cela nous a donc conduits à l'approche de notre deuxième itération, en prenant plus le temps de définir en amont les utilisateurs, leurs besoins et les points de contact par lesquels nous pouvons les satisfaire. Cela nous amène donc à une démarche visant à créer un système intégré dans l'activité de notre domaine d'application, étant les activités relatives à la production.

Néanmoins, nous pouvons également noter que les experts du domaine ont été très réceptifs aux opportunités que pouvaient nous ouvrir ces types d'exploitation. Un discours alimenté par une représentation concrète des ressources disponibles assisté par un référent sur les connaissances de la science des données, pouvant réagir sur les principes d'exploitation et la faisabilité technique au cours des échanges, nous a permis d'ouvrir notre première proposition à de nouveaux axes de recherche. Notamment, nos réflexions ont pu converger rapidement sur le cas d'usage que nous avons exposé dans le cadre de notre troisième itération, présentant une disponibilité des données, une faible complexité et un potentiel de gain important.

Le deuxième cycle d'itération a montré des résultats très satisfaisants, avec une bonne qualité du modèle de prédiction généré et un axe original et pertinent de valorisation de la donnée. Cette phase a réellement montré l'importance des phases amont de la conception par la description du cas d'usage et des données disponibles par les acteurs du service et les experts du domaine d'application. Les résultats apportés dans ce cadre sont une nouvelle fois trop spécifiques et ne sont pas directement utilisables par les opérateurs de la production. Cependant, la prédiction réalisée par le modèle ici développé présente de réels gains pour l'organisation de la production. Il a alors été acté lors de cette étape de discussion que ce modèle serait implanté dans l'outil support de l'amélioration continue de l'entreprise afin de restituer l'information relative aux erreurs à venir sur les postes, mais sous une autre forme.

D'autre part, la latitude laissée aux experts de la science des données, permise par une contextualisation précise et une expression claire des critères de réussite du projet, leur a permis d'être force de proposition dans les traitements possibles. Nous pouvons noter ici le choix d'un prétraitement des anomalies qui, une fois présenté aux potentiels utilisateurs lors de la restitution finale, a permis d'imaginer de nouvelles fonctionnalités possibles du système, basées sur ce mode d'identification. A l'origine sélectionnée pour correspondre à la nature des données, ici des séries temporelles, une brève explication du fonctionnement de cet algorithme de détection d'anomalies a suffi aux représentants du domaine d'application pour se projeter dans d'autres cas d'application. Nous pouvons citer ici l'application de ce même algorithme en temps réel, non pas pour créer un filtre, mais pour générer des alertes afin d'informer un responsable de ligne d'une anomalie dans le fonctionnement d'un poste.

Enfin, la dernière itération concerne la création de notre modèle de prédiction des retours qualité en sortie de production. La solution proposée a séduit le domaine d'application par son originalité, sa simplicité d'utilisation et les perspectives que cela peut apporter. Il a été acté dans cette étape d'intégrer l'interrogation de ce modèle à la fin du processus de fabrication. Néanmoins, deux questions demeurent. La première porte sur les impacts réels de cette prédiction sur l'efficience d'une ligne de

production. Autrement dit, l'évaluation que nous avons faite de notre modèle nous donne des notions formelles de ces performances, mais nous donne peu de lisibilité sur l'impact de son usage en production. Suite à cela, nous avons alors pris la décision de traduire les résultats que nous offre la matrice de confusion en indicateur de valeur monétaire par les experts du service. Cela se traduit par la réalisation de scénarii, comme le montre la figure 67, représentant ce que signifie et ce qu'implique une bonne ou une mauvaise prédiction dans une perspective organisationnelle. Cette démarche permet alors de faire le lien entre l'évaluation de notre modèle avec un plan de rentabilité afin d'apporter un critère unanime de validation des performances.

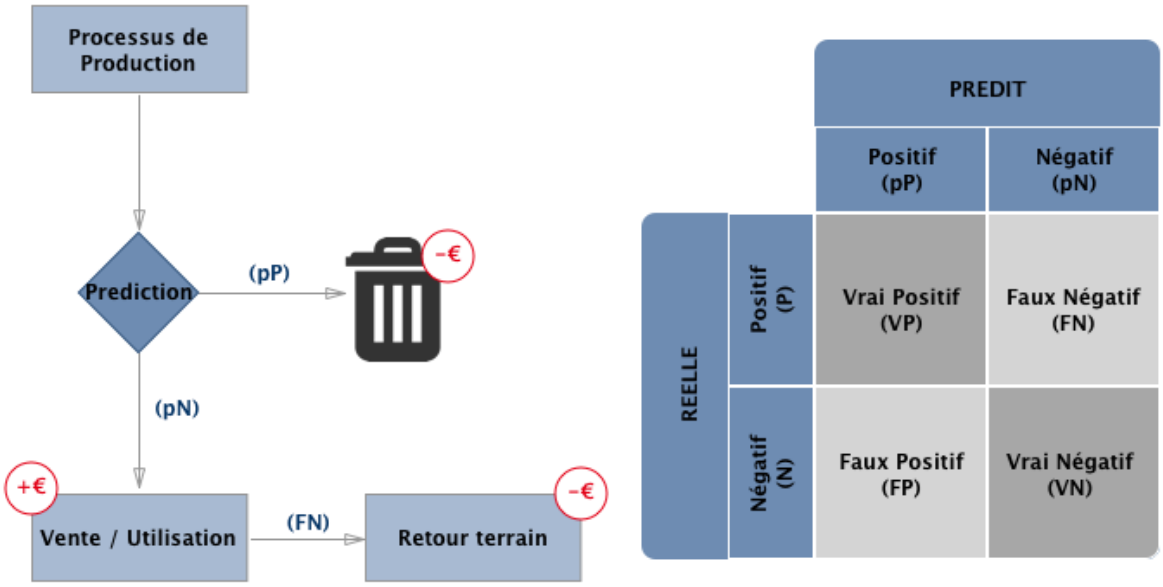


Figure 67 : Relation entre la matrice de confusion et le modèle économique

Nous pouvons alors, à l'aide de ce modèle économique, nous représenter la viabilité de l'algorithme en comparant les scénarii d'utilisation :

Sans prédiction :

$$Gain_{init} = (N_{\%} + P_{\%}) . €^{+}_{vente} - P_{\%} . €^{-}_{retour}$$

Avec prédiction :

$$Gain_{prediction} = pN_{\%} . €^{+}_{vente} - FN_{\%} . €^{-}_{retour} - pP_{\%} . €^{-}_{recyclage}$$

pour lesquels :

$Gain_{init}$: $Modèle\ de\ gain\ économique\ sans\ algorithme\ de\ détection\ des\ retours\ terrain$
$Gain_{prediction}$: $Modèle\ de\ gain\ économique\ estimé\ avec\ un\ algorithme\ de\ détection\ des\ retours\ terrain$
$€^{+}_{vente}$: $Marge\ moyenne\ d'une\ vente\ de\ produit$
$€^{-}_{retour}$: $Coût\ moyen\ d'un\ retour\ terrain$
$€^{-}_{recyclage}$: $Coût\ moyen\ d'un\ recyclage\ produit$
$N_{\%}$: $taux\ de\ négatifs$
$P_{\%}$: $taux\ de\ positifs$

$pN_{\%}$: $taux\ de\ prédictions\ négatives$
$pP_{\%}$: $taux\ de\ prédictions\ positives$
$FN_{\%}$: $taux\ de\ faux\ négatifs$

D'autre part, notre seconde interrogation porte sur les changements nécessaires de l'infrastructure du système d'information déjà présent afin de pouvoir entraîner et interroger un modèle de prédiction. Ces mécaniques engagées sont assez bien standardisées et ne doivent pas poser de problèmes majeurs. Cependant, notre système d'information dédié à la production étant fermé, nous devons mettre en place un plan d'action afin d'y ajouter une machine dédiée à l'entraînement et à la mise a disposition de modèles. Nous souhaitons alors profiter de cette dynamique afin de penser aux standards définissant l'architecture que nous pouvons mettre en place afin de gérer les activités futures de l'entreprise autour de la science des données pour l'ensemble de ses différents services.

Nous pouvons ici conclure sur la dernière étape en résumant les dynamiques collaboratives entre les métiers tel que le montre le tableau 11. Nous pouvons noter ici la prédominance du service, avec la particularité dans notre cas d'une mise en relation formelle entre les métriques d'évaluation des performances des modèles et les indicateurs du plan de rentabilité de la fonction. D'autre part, et contrairement à nos observations sur le projet précédent, les retours du domaine d'application et de la science des données sont des ouvertures du périmètre fonctionnel de notre cas d'usage.

Service	Domaine d'application	Science des données
Evaluation • Faisabilité de la fonction • Evaluation économique de la fonction **Système technique** • Cotation de l'effort nécessaire pour l'implémentation de la fonction	**Prototype** • Ouverture à de nouveaux cas d'usage • Ouverture à de nouvelles sources de données	**Techniques de traitement** • Ouverture à de nouveaux cas d'usage

Table 11 : Résumé de l'étape 6

4.3.5. Analyse des résultats et conclusion

Cette dernière expérimentation avait pour objectif de démontrer la généralisation de notre approche intégrée tenant compte de la science des données dans le développement agile de notre système d'objets connectés. Ou tout du moins, nous souhaitions éprouver notre approche de développement dans un cadre de projet différent afin de valider nos précédentes conclusions dans un contexte distinct. C'est pourquoi nous nous sommes positionnés dans un cadre nouveau de valorisation de données afin de conduire un projet suivant l'approche pluridisciplinaire que nous proposions dans notre seconde hypothèse de résolution. Dans le but de valider notre démarche, nous présenterons ici les résultats que nous avons obtenus suite à son application, nous discuterons des intérêts et des limites de notre approche avant de conclure sur la validité de nos hypothèses.

Résultats

Suite à l'application de notre approche dans le cadre de cette expérimentation, nous avons fait émerger des axes concrets de valorisation de la donnée sur la base d'une idée très peu définie à l'origine. Cela a eu différentes retombées au sein du projet et au sein de l'entreprise.

Tout d'abord, et contrairement à l'expérimentation précédente, nous avons fait émerger au cours de ces trois itérations, deux modèles directement exploitables en l'état, sous réserve de quelques ajustements. L'usage de ces modèles doit être intégré dans les évolutions de nos outils logiciels internes de la production. Cela passe notamment par l'évolution de l'architecture des systèmes d'information de l'entreprise pour permettre l'intégration progressive des plateformes de fouille de données.

D'autre part, les membres du domaine d'application de cette expérimentation ont été fortement sensibilisés aux possibilités offertes par ces méthodologies de traitement de l'information. Nous avons pu observer une ouverture d'esprit plus large sur la prise en compte de ces outils lorsqu'il s'agit de dimensionner un système de collecte de données. Cela s'est observé lors de la refonte du poste des bancs d'essai, où des liens ont été établis entre ces bancs et les indicateurs que nous suivons sur le terrain avec notre thermostat connecté pour la maintenance.

Enfin, suite aux résultats prometteurs que ce projet nous a apportés, nous avons poursuivi nos efforts afin d'étendre cette démarche à d'autres services de l'entreprise. Nous retrouvons par exemple les ventes, la logistique, les ressources humaines ou encore le marketing pour lesquels nous pouvons entrevoir des perspectives de valorisation.

Intérêts et limites

En parallèle de ces résultats, nous pouvons analyser le déroulé de notre application afin d'extraire les intérêts et les limites que cette approche a constitué dans notre projet. Et pour ce faire, nous devons tout d'abord pointer les différences que ce projet présente par rapport au cadre de notre deuxième expérimentation. Celles-ci se distinguent par deux caractéristiques principales.

Dans un premier temps nous pouvons remarquer que nous ne disposions pas, pour ce cas d'application, d'étude fine du service dans lequel s'inscrit notre développement. C'est alors en commençant le développement sans se soucier du contexte d'utilisation des résultats que ce manque d'information sur le service s'est fait le plus sentir. L'analyse préalable du service permet non seulement de dégager des fonctionnalités pertinentes pour le domaine d'application, mais permet également de contextualiser ces fonctions dans des scénarii d'utilisation desquels découlent les usages et presque implicitement les critères d'acceptation des utilisateurs.

Dans un second temps, nous pouvons remarquer que ce cadre d'application est plus conventionnel. Nous nous rapprochons ici d'une méthodologie de fouille de données classique où nous partions d'un

jeu de données appartenant à un service de l'entreprise et pour lequel nous souhaitions valoriser le contenu. De ce fait, ce projet a été poussé dans une dynamique tout à fait différente dont nous avons matérialisé les apports sur le tableau 12.

Etape	Service	Domaine d'application	Science des données
1	**Modèle de processus** • Définition du cas d'usage • Usage	**Heuristique** • Définition du cas d'usage **Retour d'expérience** • Priorités	
2		**Carte système** • Liste les données disponibles • Explicite • Filtre	
3		**Indicateurs** • Extraction de caractéristiques • Formatent	**Techniques de prétraitement** • Filtrage des valeurs aberrantes • Formatent
4		**Heuristique** • Définition d'une approche	**Technique d'apprentissage** • Définit une approche
5		**Représentation** • Validation des hypothèses	**Calcul des écarts** • Mesure des performances **Matrice de confusion** • Mesure des performances **Représentation** • Explicite les modèles • Confirme les performances
6	**Evaluation** • Faisabilité de la fonction • Evaluation économique de la fonction **Système technique** • Implémentation	**Prototype** • Validation des usages • Ouverture à de nouveaux cas d'usage • Ouverture à de nouvelles sources de données	**Techniques de traitement** • Ouverture à de nouveaux cas d'usage

Table 12 : Résumé des dynamiques collaboratives entre les métiers de l'expérimentation 3

Nous pouvons remarquer, au-delà des apports certains du service que nous avons déjà abordé, que les libertés laissées aux experts de la science des données dans le choix des approches nous ont amenés à de nombreuses ouvertures sur les typologies des solutions que nous pouvions apporter. Ce phénomène s'est particulièrement exprimé dans notre deuxième itération, avec l'expression de nouvelles fonctions imaginées par le domaine d'application suite à la présentation de la méthode de prétraitement utilisée pour préparer les données à la création du modèle. Nous pouvons également noter que ce phénomène nécessite l'implication intense des métiers sollicités dans le cadre d'application et donc de mener une démarche intégrée telle que nous la proposons.

Conclusion

En conclusion, au regard de cette analyse, nous pouvons déduire deux constats de cette expérimentation. D'une part, celle-ci nous permet de démontrer l'aspect générique de notre approche, pouvant être appliquée à de nouveaux contextes. De plus, nous pouvons remarquer que même si le cas que nous proposons ici paraît suffisant pour être traité par une approche classique de la fouille de données, nous mettons en lumière de nouvelles dynamiques du projet où une approche intégrée prend

tout son sens. D'autre part, cette expérimentation expose la complémentarité de nos deux hypothèses. À savoir que la conception amont orientée vers le service semble nécessaire afin de préparer les cycles de développement itératifs. Et même si la démarche que nous avons proposée dans la première partie de cette thèse semble particulièrement lourde à mettre en place pour notre cas relativement simple, nous pouvons cibler les éléments primordiaux par la définition de la fonction, par son périmètre et l'usage que l'on souhaite en faire, et par la mise en place d'un modèle de rentabilité pour en assurer l'évaluation.

4.4. Conclusion des expérimentations

Dans cette partie, nous avons pu parcourir les trois expérimentations que nous avons mises en place afin de confronter nos hypothèses de résolution avec notre problématique de recherche. Nous pouvons alors rappeler que cette problématique visait à déterminer comment nous pouvions concevoir une plateforme d'exploitation d'objets connectés par des services intelligents. Nous avons basé notre réponse à cette problématique autour de deux axes, distinguant les phases amont de la conception détaillée du système comme le rappelle la figure suivante.

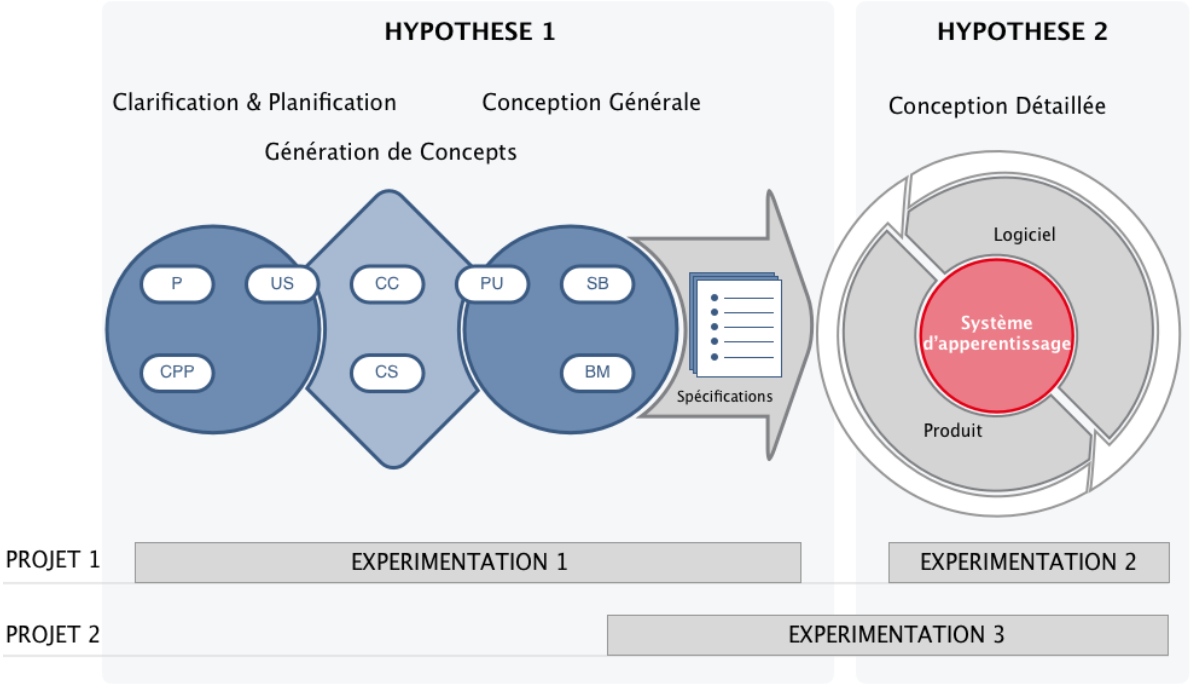


Figure 68 : Positionnement des expérimentations relativement aux hypothèses

La première hypothèse affirme qu'une approche orientée vers une analyse du service en amont de la conception était favorable au projet. Afin de confirmer cette affirmation, nous avons mis en place, dans le contexte de la première expérimentation, une démarche de conception permettant de guider le concepteur dans la structuration de son projet. Nous avons démontré ici son efficacité pour faire émerger les acteurs du service, découvrir leurs besoins et établir les premiers cahiers des charges avec leurs priorités. C'est par cette première application que nous avons pu valider notre première hypothèse avec la nécessité de travailler notre service dans les phases amont.

Puis, notre seconde hypothèse vise à introduire la science des données, au même titre que les domaines de connaissance du logiciel et du produit physique dans les cycles de développement pendant la phase de conception détaillée. Nous avons alors démontré les interactions et les intérêts que pouvait comporter l'intégration de cette nouvelle compétence dans le cadre de notre deuxième expérimentation. Cette dernière a d'une part montré les problèmes liés à la conception de fonctions intelligentes, se basant sur un modèle de données de nos produits dont la pertinence se doit d'être étudiée afin de garantir la pérennité de la solution vis-à-vis de ses fonctions sur le long terme. D'autre part, nous avons mis en avant l'impact que cette approche avait sur notre système, avec ici l'apport d'importants éléments de spécification concernant la couche physique et la couche réseau de notre système par la constitution

d'indicateurs de comportement du produit relatifs à la maintenance. Ces constatations nous ont alors permis de valider notre seconde hypothèse, affirmant la nécessité d'aborder une approche intégrée de la science des données pour la conception détaillée des fonctionnalités intelligentes du système.

Enfin, nous avons mis au point une troisième et dernière expérimentation, appliquant notre démarche de conception détaillée de fonctions intelligentes dans un cadre nouveau afin d'éprouver la généricité de notre approche. Cela nous a confirmé l'implication de chacun des métiers dans le cycle de développement et a mis en évidence la complémentarité de nos hypothèses. Effectivement, cette approche nouvelle a montré des dynamiques différentes, tout en confirmant les apports de chaque métier. De plus, nous avons constaté dans ce cadre la nécessité de concevoir une fonction en considérant l'exploitation que l'on souhaitait en faire dans son contexte. Cela se traduit par un retour vers une analyse préliminaire du service, celle-ci n'étant pas présente initialement. Cette expérimentation nous amène donc à confirmer l'aspect générique de notre approche générale, avec une étude amont du service, complétée par une conception détaillée intégrée avec la science des données.

Cela conclue notre phase expérimentale, et nous amène à poursuivre notre étude en exposant nos découvertes vis-à-vis de notre positionnement dans la recherche et dans l'industrie.

5. Apports et discussion

Afin de capitaliser et de synthétiser le contenu de cette recherche, nous présentons dans cette partie les différents apports auxquels cette recherche nous amène d'un point de vue scientifique et industriel. Si nous reprenons notre positionnement dans l'internet des objets (figure 69), par lequel nous faisons émerger la complexité de notre système par la multiplicité de ses niveaux et des domaines de compétences impliqués, alors nous pouvons discuter des apports de nos phases expérimentales.

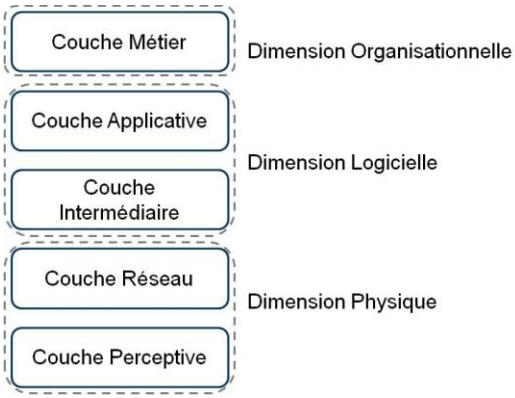


Figure 69 : Rappel de notre prositionnement dans l'internet des objets

Cette complexité entraîne la nécessité de suivre une conception structurée pour laquelle nous proposons un modèle prescriptif de conception comme le montre la figure 70. Nous considérons dans ce modèle une approche service pour les phases amont de la conception, constituant notre premier apport. Puis, pour réaliser la conception détaillée, nous proposons une démarche itérative dans laquelle nous intégrons l'ensemble des métiers du service. Le cas particulier de la conception des fonctionnalités intelligentes est le sujet de notre second apport dans lequel nous identifions les synergies et les dynamiques collaboratives au travers des différents métiers. Enfin, nous détaillons les retombées sur l'entreprise nous ayant fourni l'environnement adéquat pour la conduite de nos recherches.

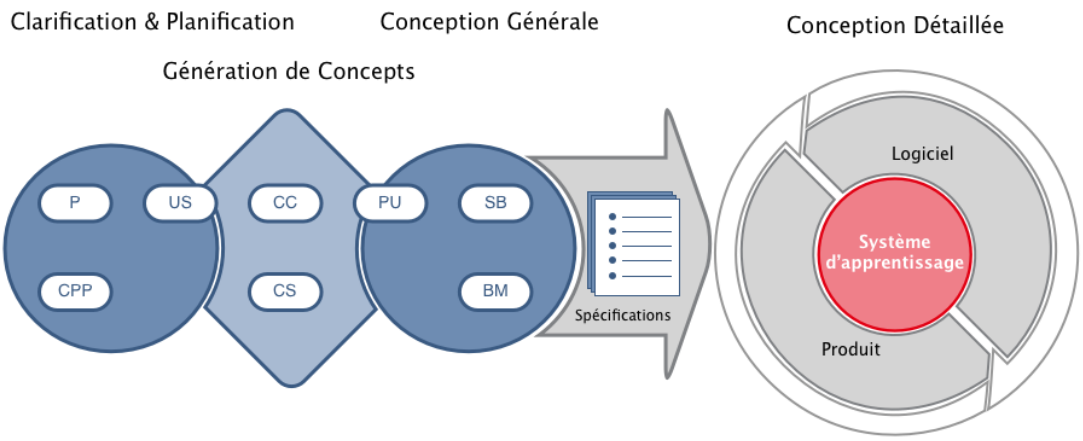


Figure 70 : Modèl générale de conception d'une plateforme de service intelligents

5.1. Modèle de conception par le service

Notre premier apport est constitué d'un modèle prescriptif de conception amont d'un système d'internet des objets. Celui-ci peut se représenter par la figure 71 que nous avions préalablement mise en avant dans la partie d'exposition de notre problématique et de nos hypothèses de notre recherche. Ce modèle de processus reprend les premières phases de la conception suivant (Pahl & Beitz 2007), à savoir, la clarification du périmètre et la planification, la génération de concepts et la conception générale. Nous avons positionné sur ces phases les méthodes et techniques les plus pertinentes pour faire face à la nature de notre problème, en sachant que le service *blueprint* a été identifié comme l'outil disposant du meilleur potentiel de représentation de notre système.

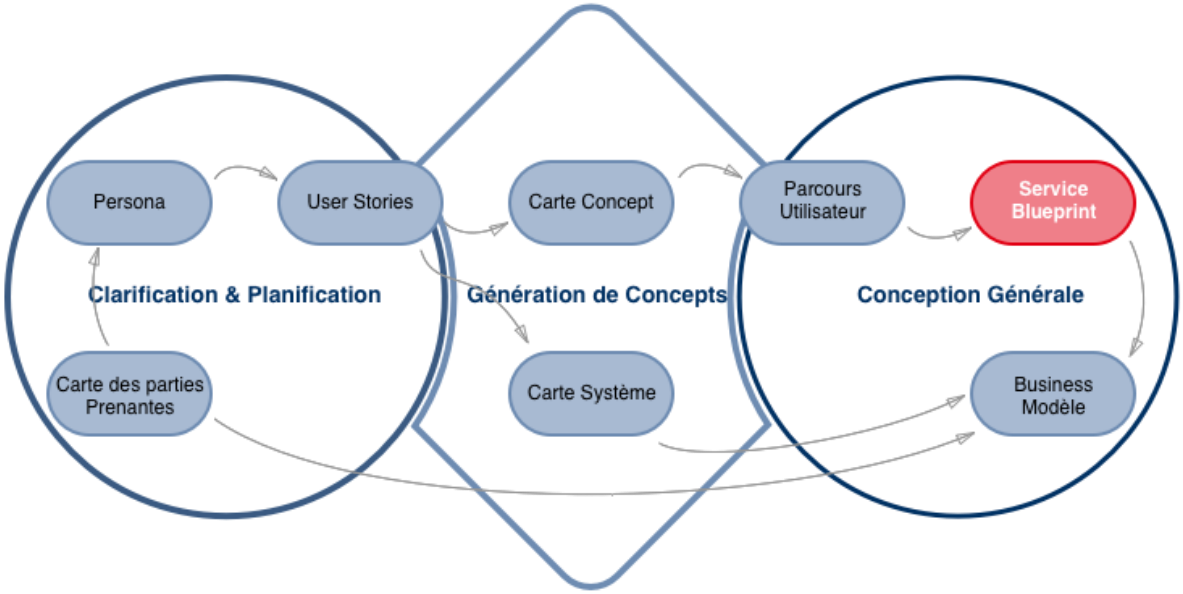


Figure 71 : Modèle méthodologique de conception amont par le service

La force de cet apport réside dans sa faculté à structurer le problème de conception en accord avec les caractéristiques particulières d'un projet d'internet des objets, comparativement aux autres modèles de conception de produits ou de services. Nous pouvons noter qu'un projet d'internet des objets se distingue fréquemment par une démarche centrée sur la technologie, pouvant être mise en opposition avec une démarche de conception issue d'une demande du marché ou d'un besoin utilisateur. Nous partons ici d'une base technologique disposant d'un potentiel d'amélioration fonctionnelle du produit ou d'amélioration de nature organisationnelle de l'activité de l'entreprise. Nous constatons également la nature immatérielle de notre système. Autrement dit, la génération et la transmission de valeur vers le client ne résident pas dans la fourniture unique d'un artefact matériel. Enfin, nous avons identifié que les bonnes pratiques de la conception orientent le concepteur vers une démarche de type agile, avec une réalisation progressive de notre système suivant de nombreuses itérations. Cependant, cela ne concerne que la réalisation de notre système, sans tenir compte de la préparation nécessaire au lancement des phases itératives.

Conformément à ces observations, nous pouvons alors identifier les apports de notre modèle méthodologique au travers des points suivants :

La séparation des phases amont avec la conception détaillée et la réalisation

Effectivement, notre proposition de démarche s'intéresse non seulement au développement de notre système, mais prend également en compte le fait que nous sommes positionnés dans une démarche centrée autour des technologies, et qu'il nous est nécessaire de mieux comprendre le potentiel réel de notre technologie au sein de notre domaine d'application et vis-à-vis de nos clients cibles avant de pouvoir entreprendre sa réalisation. Ce positionnement répond donc à un manque dans le cadre des méthodologies agiles, dont on a constaté la force pour maintenir un certain degré de flou au sein d'un projet afin de nous permettre de nous focaliser sur l'essentiel pour notre client. Cependant, l'ensemble de ces démarches ne détaille pas les efforts nécessaires en amont de la conception afin de préparer le terrain pour la mise en application des itérations. Et cela est d'autant plus vrai pour nous, qui ne connaissons a priori pas nos clients ni leurs besoins.

La prise en compte de la nature intangible du système par une approche service

Notre système, dont la valeur réside principalement dans les flux d'information, ne peut pas être considéré comme un produit pour lequel la fourniture se fait de manière unique dans le temps et dans l'espace. C'est la raison pour laquelle nous nous sommes positionnés au travers d'une démarche orientée service, ne considérant pas ici la fourniture d'un artefact à un instant précis, mais la mise en place d'offres de service omniprésentes entre clients et fournisseurs. Pour ce faire, la suite d'outils que nous avons proposés permet de faire face aux différentes facettes du système en tenant compte des acteurs, du système technique, des interactions, de l'environnement d'affaire, du besoin client et des fonctionnalités associées. Cette démarche est construite par une succession d'outils fournissant à l'équipe projet de nombreux modèles descriptifs, représentant l'ensemble de ses dimensions dans le temps. Cela s'est avéré particulièrement efficace pour comprendre notre système et son potentiel puis pour cadrer l'environnement de développement.

La conception détaillée au croisement de plusieurs disciplines

Enfin, notre système dans son ensemble fait appel à de nombreuses compétences. Nous tenons compte du besoin au travers des représentants d'un domaine d'application, faisant émerger les leviers sur lesquels ces technologies peuvent influer afin d'améliorer l'activité de ce domaine. L'aspect organisationnel est pris en compte par le service. Il intervient autour de deux axes principaux, la structure du service et l'évaluation de la rentabilité du système. Puis, l'ensemble des métiers en relation avec la dimension technologique du système interviennent avec l'ensemble des acteurs pour les différentes connaissances du logiciel, pour les connaissances produit ou encore pour les connaissances de la partie réseau. La génération de valeur étant transverse à l'ensemble des couches du système, la qualité de service proposée est donc en relation avec l'implication de ces connaissances. La proposition de méthode permet alors de modéliser notre système et de représenter notre service de manière à faciliter les échanges entre ces différents représentants des compétences afin qu'ils puissent se positionner suivant notre proposition d'offre de service. Cela permet alors d'intégrer au mieux les représentants de ces compétences tout au long de la spécification du système.

Les limites du Service Blueprint

D'autre part, en marge de ces points positifs, nous pouvons également noter certaines limites de notre modèle de service avec le *Service Blueprint*. Effectivement, même si dans notre cas d'application, sont usage s'avers être des plus juste, nous pouvons lui reprocher certains manques pouvant desservir sa généricité. Le Service Blueprint est un modèle permettant de représenter un nouveau service de manière simple, exhaustive et flexible. Cependant, celui-ci ne peut être considéré comme une fin en soi. Autrement dit, ce modèle ne permet ni la mise en place d'une nouvelle organisation au sein de l'entreprise, ni la réalisation d'un produit. Ces deux parties doivent faire l'objet d'études

complémentaires pour être réalisées. Le *Service Blueprint* est considéré ici comme un plan d'architecte, nous amenant l'ensemble des informations nécessaires sur la structure du service pour la conception détaillée du système physique et organisationnel.

Enfin, pour conclure, nous pouvons pointer du doigt les limites de cet outil pour modéliser des systèmes organisationnels trop complexes. Nous entendons ici le fait que le Service Blueprint est efficace dans la modélisation d'interactions entre un client et un fournisseur. Seulement, si le service doit disposer d'une structure organique avec grand nombre de clients cibles et de fournisseurs différents, pour desquels la fourniture de la valeur se traduit en de multiples interactions, alors le modèle perd tous ces atouts de simplicité et de clarté. Cela peut être contourné par la séparation du modèle suivant différentes représentations de fonctionnalités indépendantes. Cependant, nous perdons alors dans ces conditions la vue globale de notre service.

5.2. Dynamique de conception détaillée des fonctions intelligentes

Par la suite, au sein du périmètre fonctionnel de notre système, sont apparues des fonctionnalités nous amenant à considérer une certaine forme d'intelligence système. Autrement dit, nous nous sommes heurtés, lors de la conception détaillée de ce système, à la nécessité de mettre en place des processus liés à l'apprentissage autonome de la machine et à la prise de décision. Cela faisant intervenir une nouvelle fois une discipline complémentaire, avec des experts en science des données. Cela nous a permis de mettre en lumière une forte dépendance, non explicite, entre les données transitant dans le système et les moyens à mettre en œuvre pour réaliser nos fonctions. En se positionnant dans une démarche classique de conception de systèmes de découverte de connaissance, nous faisons alors face à un manque de données issues de notre système, constituant la principale amorce de cet apport. Nous proposons ici une démarche intégrée dans laquelle nous avons fait intervenir principalement trois domaines d'expertise afin de développer ces fonctionnalités.

Nous pouvons établir un constat sur l'identification de ces fonctions disposant d'un caractère intelligent. Effectivement, ces dernières apparaissent spontanément dans la phase de génération de concepts et se distinguent des autres par des verbes et adjectifs particuliers faisant appel aux technologies d'apprentissage autonome. Nous pouvons alors mettre en lumière ici certains cas auxquels nous avons été confrontés :

- La prédiction, avec les mots clé : Anticiper, prévoir, appréhender ou prévenir
- La reconnaissance de cas, avec : Identifier, préconiser ou préparer
- La mise en avant de schémas, avec : Comprendre, simplifier, extraire, améliorer ou optimiser

Lors de la conception générale de ces fonctions, nous avons tendance à nous orienter vers un processus de conception issue de la littérature en science des données tels que ceux mis en lumière par (Krugan & Musilek 2006). Ces modèles sont optimisés pour les experts de la science des données, mais montrent de réelles faiblesses dans leur application pour la conception d'un nouveau système de données (incluant la collecte et le traitement).

Nous proposons avec le protocole expérimental que nous avons utilisé dans les expérimentations 2 et 3 une approche pluridisciplinaire, où l'ensemble des métiers en lien avec la fonction sont impliqués. Cette approche permet de minimiser l'impact des points bloquants au cours de la conception en apportant de la connaissance complémentaire sur le projet ou le système étudié. De la même manière que pour l'ingénierie concourante ou les DfX, cette approche permet de minimiser les efforts à mettre en place afin d'obtenir une solution optimale (Bralla 1996). Cela se traduit dans notre cas par une diminution du nombre d'itérations nécessaires pour obtenir une solution pouvant être utilisée dans le service.

Afin de systématiser cette approche, nous avons extrait les dynamiques induites par l'ensemble des participants et des métiers au sein du processus de conception. Nous pouvons alors représenter cet apport sous la forme du graph présenté par la figure 72 avec, pour chaque phase du processus, les contributions apportées par chaque acteur et le support sur lesquelles elles se matérialisent.

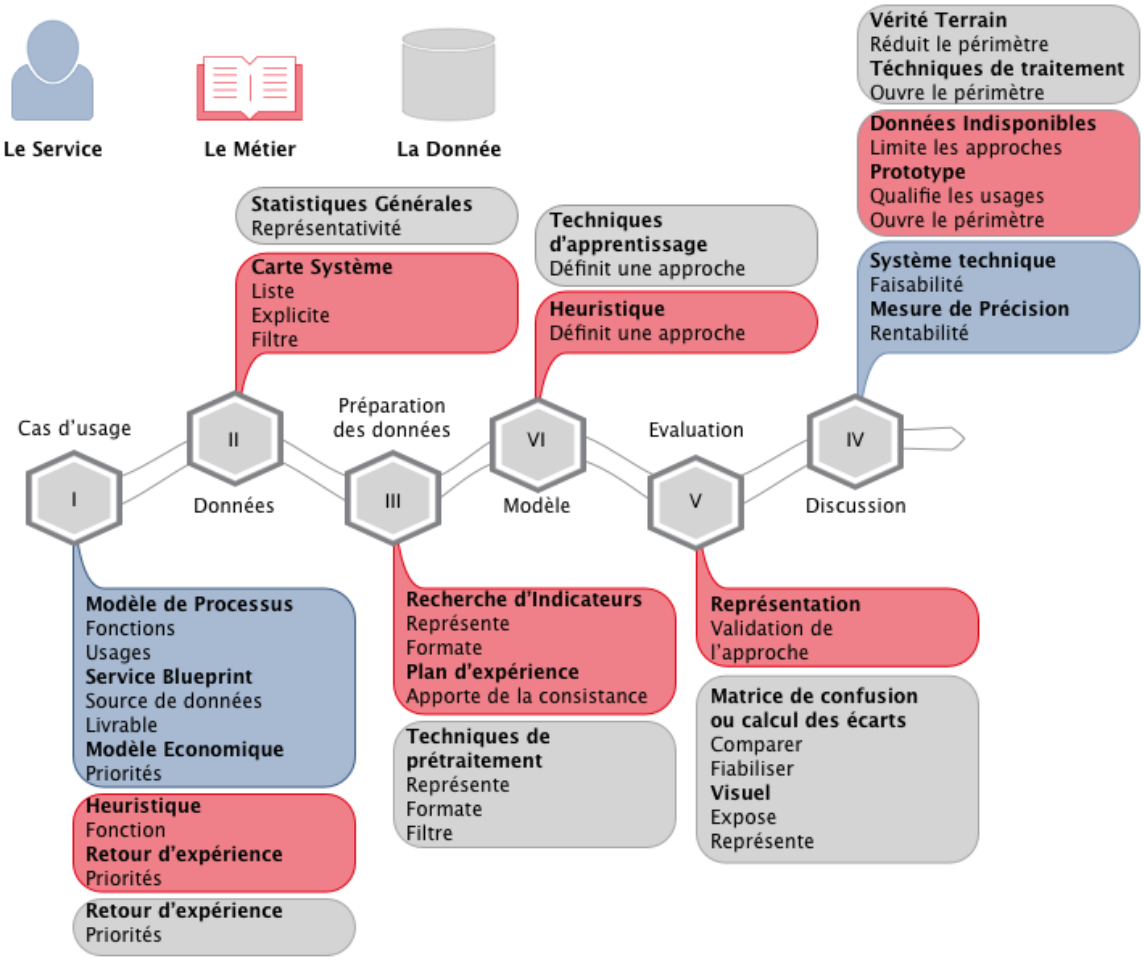


Figure 72 : Illustration des dynamiques collaboratives entre les métiers du service, de la donnée et du domaine d'application dans le cycle de conception d'algorithmes d'apprentissage

Pour plus de détails, nous revenons maintenant phase par phase sur les dynamiques que nous avons ici identifiées.

Phase 1 : Compréhension du cas d'usage

Cette première phase s'appuie particulièrement sur les connaissances et les travaux pouvant être réalisés par le service. Autrement dit, la contextualisation et la définition de cette étude passe par une analyse du cadre organisationnel de l'application que nous souhaitons ici mettre en place. Nous avons identifié que la mise en œuvre du *service blueprint* nous permet de clarifier les acteurs engagés et les sources de données pouvant être utilisées. De plus, un modèle plus détaillé du processus de mise en œuvre de la fonction au sein de l'activité de son utilisateur nous permet de caractériser avec plus de détails les usages que nous pouvons en attendre. Enfin, le service peut proposer un modèle économique lié à cette fonction. Cela permet de bien comprendre son intégration dans l'activité de l'utilisateur en dégageant les leviers potentiels de cette fonction sur l'écosystème dans lequel elle prend place. Ces leviers identifiés, nous pouvons dresser la liste des cas d'usage qui présentent le plus gros potentiel de rentabilité pour l'entreprise et qui constituent nos priorités.

Nos deux autres domaines d'expertise peuvent ici nous faire part de leur retour d'expérience autour de cas d'usage ciblés afin d'en évaluer la faisabilité et de raffiner notre analyse préliminaire ainsi que nos priorités. Le domaine d'application peut déjà avoir une expérience d'amélioration de son service similaire et identifier les problématiques les plus intéressantes à traiter. Cela pouvant être des problématiques techniques, tout comme des problématiques d'intégration dans la vie du service. D'autre

part, la science des données dispose aussi d'expériences pouvant être valorisées par analogie et à moindre frais dans le cadre d'un cas d'usage que nous pouvons ici sélectionner. Dans les deux cas, nous pouvons raffiner notre perception du domaine d'application afin de nous focaliser sur des cas apportant des gains rapides répondant à des problématiques réelles du domaine d'application.

Phase 2 : Compréhension des données

Lors de cette deuxième phase, les connaissances métier sont alors motrices pour comprendre les données disponibles. Cela de deux manières, afin de pouvoir les expliciter, leur donner sens ; puis afin de les filtrer et d'éliminer les attributs pouvant être considérés comme parasites face au problème posé. Nous avons mis en évidence différents modèles de représentation pouvant être utiles ici suivant la nature du système sur lequel nous portons notre intérêt : des modèles de systèmes, décrivant l'ensemble de leurs sous-parties et de leur agencement les uns par rapport aux autres dans le cas de l'étude d'un système physique, puis des modèles de processus ou de flux décrivant les acteurs, leurs rôles et leurs interactions dans le cas d'une étude d'un système organisationnel. Ces modèles permettent alors de nous représenter concrètement le contenu des bases de données dans l'environnement du système observé avec les formalités de collecte et de stockage. Cela permet par la suite d'évaluer plus facilement le potentiel que ces bases de données peuvent renfermer pour répondre à notre cas d'application.

Ensuite, la science des données peut ici commencer à manipuler les données, évaluer leur contenu et juger des efforts de mise en forme que nous pourrons avoir à concéder. Parmi cela nous pouvons évaluer la représentativité de chaque attribut avec son format, sa résolution, sa distribution, sa variabilité ou encore ses valeurs manquantes. Cela peut venir compléter les efforts fournis par les connaissances métier pour évaluer la pertinence de certains champs et contribuer à évaluer les efforts de mise en forme nécessaires pour la phase suivante.

Phase 3 : Identification d'un modèle et préparation des données

Cette troisième phase permet de préparer nos données en accord avec une approche sélectionnée. Nous discuterons plus en détail l'implication de nos trois domaines de compétence dans le choix de l'approche dans la partie suivante. Nous voyons ici l'implication de ces derniers pour faciliter la préparation des données où la science des données et le domaine métier jouent deux rôles distincts.

Si nous nous intéressons au domaine de la science des donnés, ce domaine dispose d'une vision assez claire du formalisme nécessaire des données pour pouvoir appliquer une approche d'apprentissage. De plus, ce domaine offre un ensemble de techniques et de savoir-faire permettant la préhension et la manipulation des données afin de les accommoder à un formalisme défini et nécessaire.

Ensuite, lors de cette phase de prétraitement, les connaissances du métier concernant le système d'étude s'avèrent être primordiales dans le prétraitement des données. Nous avons notamment montré leur investissement dans la création d'indicateurs. Autrement dit dans la création de croisements entre les différentes données brutes afin de faire ressortir un caractère comportemental du système pertinent pour le cas d'usage ciblé. Cela nous permet d'isoler potentiellement de la matière utile difficilement perceptible de l'ensemble des données, mais également de faire fi de bruits que nous pouvons retrouver dans certains attributs et que le métier peut contribuer à réduire en nous guidant vers un échantillonnage pertinent.

D'autre part, nous avons vu que le métier pouvait être sollicité afin d'améliorer la qualité des données sur une petite échelle en préparant des plans d'expérience pour la collecte de nouveaux jeux. Nous avons vu que cela ne permet pas de réaliser notre fonction intelligente, néanmoins, cela nous offre la possibilité de désamorcer une situation pour laquelle nous ne disposions pas d'enregistrements de la vérité que nous souhaitions étudier. Un plan d'essai, dans un environnement contrôlé et pour une

échelle réduite, organisé autour du phénomène que nous souhaitons observer nous conforte dans la faisabilité de notre fonction.

Phase 4 : Application du modèle

Concernant cette quatrième phase, l'application du modèle étant très mécanique, son analyse présente un intérêt réduit. Ainsi, ce qui nous importe ici est son choix, pour lequel nous sommes de nouveau opposés à deux approches entre le métier et les données.

Pour la science des données, la description du cas d'usage va pouvoir nous orienter vers une typologie de techniques que nous avons décrite dans l'état de l'art. A savoir, des modèles descriptifs ou prescriptifs par classification ou par régression. Une fois cette dépendance au cas d'usage résolu, la technique en elle-même peut être choisie en fonction de la typologie de nos données. Et si, suite à cette segmentation, le nombre de solutions possibles est encore élevé, l'ensemble des modèles peuvent alors être testés au travers d'itérations rapides entre cette phase et la suivante afin de sélectionner le plus performant.

D'un point de vue métier, le choix d'une approche peut également se faire en analysant le cas d'usage. Effectivement si le problème que nous posons initialement peut être identifié à une typologie de problèmes connue par le métier, alors nous pouvons mettre en œuvre des solutions connues pour leur efficacité dans ce contexte. Néanmoins, si nous ne sommes pas directement passés par ce biais pour résoudre notre problème, c'est qu'il existe potentiellement de nombreuses limites. Nous avons pu mettre en avant ici des limites d'interprétation des données ou encore des limites d'autonomie dans la mise en place d'une solution. Cependant, ces approches forment une base solide pour l'interprétation du problème et cela permet de nous ouvrir de nouveaux axes de recherche pour le résoudre.

Phase 5 : Évaluation du modèle

Ensuite, la phase d'évaluation vient de paire avec l'application du modèle. Pour celle-ci, nous observons deux possibilités de mise en œuvre.

D'une part, nous avons la possibilité de visualiser les résultats d'un traitement. Cela peut prendre la forme d'un schéma extrait par l'application du modèle lorsque cela est possible, comme un arbre de décision. Cela peut également prendre la forme d'une représentation de notre jeu de données par la mise en évidence du traitement que nous venons de lui faire subir. Nous pouvons nous baser sur l'exemple de la représentation de nos indicateurs, cela permet alors, de manière visuelle, d'évaluer si notre traitement vérifie une hypothèse que nous avions émise avant sa mise en application. Grâce à cette approche visuelle nous pouvons donc évaluer la pertinence d'une solution, et cette approche présente l'avantage de montrer de manière explicite ce qu'apporte l'application de notre modèle, et de manière compréhensible par tous.

D'autre part, la science des données pousse à une évaluation formelle des performances des modèles par une mesure des écarts entre les valeurs prédites et les valeurs réelles. Cette quantification des écarts permet notamment la mise en relation entre les performances du modèle et notre analyse économique. Effectivement, si nous prenons le cas de l'application d'un modèle de classification, nous pouvons alors quantifier l'impact d'une bonne ou d'une mauvaise prédiction sur le domaine d'application. Cela nous offre donc la possibilité de mettre en place, à l'aide du service, un plan de rentabilité pour justifier le bénéfice économique du déploiement du modèle.

Phase 6 : Discussion

Enfin, la phase de discussion permet de faire le point sur les avancées apportées par l'itération afin d'identifier et de planifier les actions à mettre en œuvre par la suite pour faire avancer notre projet. Suite

au déroulement d'une itération, nous avons pu noter que chacune des parties se focalisait sur certains sujets que nous pouvons présenter de la manière suivante.

Suite à l'application de nombreuses approches n'aboutissant pas à une solution convenable pour répondre au besoin du cas d'usage, la science des données nous amène à reconsidérer notre jeu de données. Cela pouvant être dû à un problème qualitatif, où la collecte et le stockage ne s'effectue pas de manière a permettre un traitement suffisamment précis par la suite ; ou à un problème quantitatif, avec un manque de représentativité de l'étendue des états que peut prendre notre système. D'autre part, et un peu en marge de ces problèmes, nous pouvons également identifier ici des manques d'information sur la donnée que nous souhaitons prédire. Ces discussions peuvent entraîner des croisements avec de nouvelles sources de données ou des actions long terme avec le service sur les moyens mis en œuvre pour collecter les informations.

Du point de vue du domaine d'application, le manque de données clés dans son interprétation du comportement des systèmes à l'étude peut le pousser à restreindre le périmètre de la fonction que nous souhaitons développer. Cela peut entraîner la restriction du périmètre de la fonction dans l'attente d'actions long terme également pour améliorer la collecte de la donnée, ou bien, de recentrer le problème suivant un nouvel axe de recherche épaulé par les autres domaines de compétence. D'autre part, le domaine d'application ouvre généralement le périmètre de la fonction à la mise en œuvre et à l'usage du système, considérant ainsi les ressources nécessaires à l'interprétation des résultats du modèle et l'analyse des retombées que peut avoir ce système au long terme pour son activité. Nous pouvons ici noter que la mise en œuvre de modèles non explicites peut être source de frustration pour le domaine d'application, ne pouvant pas exploiter les lois utilisées dans le modèle afin de comprendre le phénomène étudié et de déceler des leviers d'amélioration long terme de son activité.

Enfin, du point de vue du service, les notions de mise en œuvre interviennent dans le cadre d'un déploiement dans une architecture système générale. Se posent alors les questions d'organisation des flux d'information, de leur restitution à l'utilisateur et de l'ordonnancement des phases de création de modèle et de génération de prédiction. Cette systématique d'apprentissage doit être réalisée en assurant le suivi sur la qualité de la prédiction du système tout au long son exploitation. Cette qualité de prédiction peut alors être mise en relation une nouvelle fois avec un modèle économique afin d'évaluer le gain de ce système non pas dans des termes mathématiques, mais d'une manière explicite par les responsables du domaine d'activité. D'autre part, fort de cette expérience et par la vision panoramique qu'entretient le service sur l'ensemble du domaine d'activité, ce dernier peut faire de nouvelles propositions de cas d'usage auxquelles les approches retenues dans une itération peuvent répondre.

5.3. Les apports industriels

Enfin, le dernier type d'apport est relatif à l'industrie dans laquelle nous inscrivons nos travaux, en d'autres termes pour l'entreprise d'e.l.m. leblanc et le groupe Bosch. Nous pouvons alors dissocier ces apports d'une part relativement à la proposition de valeur de l'entreprise, puis, d'autre part relativement à la connaissance maîtrisée par l'entreprise et la culture qu'elle développe autour d'elle.

Concernant la proposition de valeur de l'entreprise, notre étude est arrivée dans un tournant réglementaire dans le monde du chauffage favorisant l'installation de chaudières à condensation au détriment des chaudières conventionnelles. Cela constituait un risque majeur de perte de parts de marché pour l'entreprise. Cependant, ces recherches sur les objets connectés nous ont permis de faire valoir la haute technicité de nos produits et notre savoir-faire sur l'ensemble du cycle de vie du produit incluant son utilisation et sa maintenance. D'autre part, la structure fonctionnelle du service, avec des offres orientées vers les professionnels du chauffage, nous a permis d'attaquer de nouveaux marchés en nous adressant aux acteurs force de prescription pour notre marque avec de nouvelles propositions de valeurs nous différenciant de la concurrence.

En outre, la création de nouveaux services se basant sur l'utilisation de nos produits positionne l'entreprise dans un nouveau paradigme de vente. Effectivement, les revenus d'e.l.m. leblanc, avant nos travaux, reposaient uniquement sur des transactions uniques lors de la vente d'appareils. La proposition de service crée un lien continu entre l'entreprise et ses clients avec une source de revenus continus tout au long de l'utilisation des services. Cela offre, en plus de la proximité créée avec nos clients, un modèle de revenus plus confortable, permettant d'atténuer les fluctuations saisonnières des volumes de ventes que nous pouvons subir.

D'autre part, le positionnement proche des parties prenantes du système, comprenant les futurs utilisateurs, que nous avons tenu tout au long du développement de ces nouveaux services, entretient l'image innovante de l'entreprise et favorise la mise en place de nombreux partenariats. À titre d'exemple, l'intégration des connaissances en électronique de régulation a ouvert deux nouveaux projets répondant à des demandes marché spécifiques pour lesquelles e.l.m. leblanc serait en mesure de se positionner rapidement. Concernant l'intégration des utilisateurs, cela a conduit au lancement d'un partenariat avec un service après-vente externe dans lequel nous avons remporté un appel d'offre pour la fourniture et le maintien d'un parc de chaudières communicantes. Ce cas est le seul ayant à l'heure actuelle abouti à la concrétisation d'une vente, cependant trois autres preuves de concept sont en cours avec d'autres partenaires, dont un pour l'ouverture d'un projet collaboratif européen.

Enfin, concernant les apports de nouvelles connaissances dans l'entreprise, nous avons su démocratiser les usages de la fouille de données au sein de l'entreprise. Cela, grâce à l'investissement fait sur la plateforme de traitement pour nos objets connectés, nous permet d'imaginer étendre la création de fonctionnalités intelligentes pour l'ensemble du cycle de vie du produit. Nous avons commencé à établir cette dynamique au travers de notre dernière expérimentation sur l'amélioration de la production. Nous avons alors poussé cette réflexion pour l'ensemble des services de l'entreprise tels que la logistique, les ventes ou les ressources humaines. Ces efforts ont montré, avec la production, que l'introduction de la fouille de données a convaincu et ses perspectives sont prises en compte dans l'établissement des stratégies d'amélioration.

Au regard de ces efforts, l'entreprise e.l.m. leblanc cherche également à les faire valoir au sein du groupe Bosch et est devenue l'usine de référence dans ce domaine. Elle prend donc la responsabilité du

développement des systèmes d'amélioration des usines du groupe par la création de services intelligents. Cela comporte donc la responsabilité des infrastructures, des connaissances et du développement des cas d'applications en assurant leur déploiement pour l'ensemble des usines du groupe.

Conclusion de la thèse

L'objectif premier de ces travaux de thèse était d'accompagner les entreprises dans leur appropriation des dynamiques industrielles portées par les technologies de l'information et de la communication. Cela avait pour but d'ouvrir de nouveaux leviers d'innovation pour des entreprises dont les activités étaient a priori loin du monde du numérique. Ces travaux résultent en la fourniture d'un guide méthodologique, pragmatique et inspiré des standards de la science de la conception, permettant d'organiser les efforts des concepteurs afin d'avancer dans cette direction.

Si nous revenons tout d'abord sur notre cas d'application industriel, l'opération menée au cours de ces trois ans semble couronnée de succès. Même si il est un peu tôt pour conclure sur le fait que nos travaux se traduisent par de réelles innovations, nous observons l'ouverture de l'entreprise à de nouveaux marchés dans lesquels nos solutions connectées sont très bien accueillies. De plus, et suite à l'introduction de ces nouvelles technologies, l'entreprise semble vouloir étendre la mise en application de ces procédés, plus seulement pour les produits finis, mais pour l'ensemble des activités jalonnant le cycle de vie des produits.

Suite à ces constatations, nous pouvons conclure sur l'objet de notre étude, à savoir l'apport méthodologique nécessaire au concepteur afin de parvenir à innover au travers de ces technologies. Revenons d'abord sur le parti pris d'aborder une approche initiée par le service. Celle-ci a joué un rôle déterminant dans l'acceptation des utilisateurs et cela s'explique par une concentration des efforts dans les phases amont afin de mieux les comprendre. Comprendre quel est leur besoin et comment un système, quel qu'il soit, pourrait y répondre au mieux tout en s'intégrant dans les activités qui leurs sont propres. Ces efforts du concepteur ont alors permis de créer un cahier des charges représentatif d'un idéal de solutions pour les clients auxquelles elles s'adressent.

Par la suite, un découpage strict des différentes plateformes qui composent le système permet de faciliter la gestion de la complexité du système. Ce découpage permet de considérer indépendamment chacun des sous-systèmes, en se concentrant sur ce qui le rend utile pour le client. Enfin, une gestion agile du projet permet d'appréhender au mieux l'ensemble des dépendances entre ces plateformes, notamment entre la partie physique et la partie logicielle. Cela s'avère d'autant plus vrai si l'une de ces deux parties est déjà existante et nécessite d'être modifiée pour répondre à notre besoin.

Enfin, il semble indispensable de considérer l'intégration de la science des données au sein du développement de la plateforme logicielle pour la réalisation de certaines fonctions présentant un caractère intelligent. Ces typologies de fonctions sont de plus en plus omniprésentes dans le monde qui nous entoure et c'est spontanément que les utilisateurs nous y poussent lors de l'analyse du besoin. L'intégration de ces sciences n'est pas triviale lorsque nous travaillons sur un projet complètement nouveau. Cependant, nous avons montré qu'une collaboration étroite entre les experts du domaine d'application, de la science des données et du service pouvait conduire à favoriser leur mise en application en les considérant au plus tôt dans la conception du système. De plus, et afin d'optimiser ces échanges, nous avons cartographié les dynamiques apportées par chacun de ces domaines. Cela devrait par la suite nous aider à minimiser le nombre d'itérations nécessaires à la conception de ces fonctions.

Ce dernier point offre les plus importantes perspectives de recherche pour la suite de nos travaux. En effet, les deux hypothèses sur lesquelles nous nous sommes penchés pendant ces travaux sont complémentaires, et il est impératif de se représenter un service avant d'en aborder le contenu. Néanmoins, c'est dans la réalisation des fonctionnalités intelligentes que nous trouvons le potentiel le plus important d'amélioration. La science des données s'ouvre de plus en plus au monde industriel en proposant un nombre grandissant de plateformes de traitement de la donnée simples à utiliser et à intégrer dans des systèmes complexes. Cependant, la réalisation d'un cas d'usage à l'aide de ces technologies nécessite de comprendre les théories sur lesquelles ces technologies sont basées, se référant souvent à des concepts abstraits et produisant des résultats peu démonstratifs. Nous pouvons alors imaginer travailler sur des moyens simplifiant la collaboration entre les métiers, principalement basés sur des relations entre des typologies de fonctions et des typologies de solutions. De telles analyses peuvent conduire à de nombreuses améliorations du processus de conception. Nous pouvons imaginer que cela pourrait être un support de réflexion lors de différentes phases créatives, aussi bien pour formaliser le besoin que pour rechercher des solutions. Cela permettrait également de considérer en amont la réalisation des systèmes de collecte et de stockage de la donnée au regard de ce que nous souhaiterions en faire à l'avenir.

Références

Adian Palmer. (2003). Marketing of services. In *Marketing book* (pp. 585–609).

Anderson, E. W., Fornell, C., & Rust, R. T. (1997). Customer Satisfaction, Productivity, and Profitability: Differences between Goods and Services. *Marketing Science*, *16*(2), 129–145. doi:10.1287/mksc.16.2.129

Aoussat, A. (1990). *La pertinence en innovation: nécessité d'une approche plurielle*. Arts & Métiers ParisTech.

Aoussat, A., Christofol, H., & Le Coq, M. (2000). The New Product Design: A Transverse Approach. *Journal of Engineering Design*, *11*, 399–417.

Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A Survey. *Computer Networks*, *54(15)*, 2787–2805.

Baines, T. S., Lightfoot, H. W., & Kay, J. M. (2006). The Servitization of Manufacturing: a Review of Literature, 1–11.

Balarin, F., Watanabe, Y., Hsieh, H., Lavagno, L., Passerone, C., & Sagiovanni-Vincentlli, A. (2003). Metropolis: An Integrated Electronic System Design Environment. *Computer*, *36*, 45–52.

Bassi, A., & Horn, G. (2008). Internet of Things in 2020 : A Roadmap fot the Future. *European Commission : Information Scociety and Media.*

Beck, K. (2000). *Extreme Programming Explained*. (A. Professional, Ed.).

Boehm, B. W. (1988). A Spiral Model of Software Development and Enhancement. *Computer*, *21*, 61–72.

Bongard, K. (2013). *Bringing the User Experience to Early Product Design : from Idea Generation to Idea Evaluation*. Arts & Métiers ParisTech.

Booms, B. H., & Bitner, M. J. (1981). Marketing Strategies and Organization Structures for Service Firms. *Marketing of Services, American Marketing Association, Chicago, IL,*, 51– 67.

Börner, K. (2013). *Visual Insights*.

Bralla, J. G. (1996). Design for Excellence. *McGraw-Hill Professional Publishing*.

Bremmer, R. (1999). Cutting-Edge Platforms. *Financial Times Automotive World*, *9*, 30–38.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0 Step-by-Step Data Mining Guide.

Chase, R. B. (1981). The Customer Contact Approach to Services: Theoretical Bases and Practical Extensions. *Operations Research*, *29*, 698–706.

Cios, K. J., & Kurgan, L. A. (2005). Trends in Data Mining and Knowledge Discovery. *Advanced Techniques in Knowledge Discovery and Data Mining*, 1–26.

Cios, K. J., Teresinska, A., Konieczna, S., Potocka, J., & Sharma, S. (2000). Diagnosing Myocardial Perfusion from PECT Bull's-Eye Maps-A Knowledge Discovery Approach. *IEEE Engineering in Medicine and Biology Magazine*, *19*, 17–25.

Cooper, R. G. (1999). *Product Development for the Service Sector: Lessons from Market Leaders*. (B. Books, Ed.).

Cross, N. (2008). *Engineering Design Methods: Strategies for Product Design*. *Design* (Vol. 1, p. 230).

Cusumano, M. (2010). Technology Strategy and Management The Evolution of Platform Thinking. *Communications of the ACM*, *53*, 32–34.

DaCosta. (2013). *Rethinking the Internet of Things*.

De Deugd, S., Carroll, R., Kelly, K., Millett, B., & Ricker, J. (2016). SODA: Service Oriented Device Architecture. *IEEE Pervasive Computing*, *3*, 94–96.

Demirkan, H., Kauffman, R. J., Vayghan, J. a., Fill, H.-G., Karagiannis, D., & Maglio, P. P. (2008). Service-Oriented Technology and Management: Perspectives on Research and Practice for the Coming Decade. *Electronic Commerce Research and Applications*, *7*(4), 356–376.

Diana, C., Pacenti, E., & Tassi, R. (2009). Visualtiles : Communication Tools for (Service) Design. In *De-Thinking Service Re-Thinking Design, Oslo, Norway, 24-26 November 2009* (p. 75).

Dittrich, J., & Quiané-Ruiz, J. A. (2012). Efficient Big Data Processing in Haddop MapReduce. *Roceedings of the VLDB Endowment*, *5*, 2014–2015.

Dorst, K., & Cross, N. (2001). Creativity in the Design Process: Co-Evolution of Problem–Solution. *Design Studies*, *22*, 425–437.

Fan, T., & Chen, Y. (2010). A Scheme of Data Management in the Internet of Things. In *Network Infrastructure and Digital Content, 2010 2nd IEEE International Conference* (pp. 110–114). IEEE.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996c). Knowledge Discovery and Data Mining: Towards a Unifying Framework. *KDD*, *96*, 82–88.

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996a). *Advances in Knowledge Discovery and Data Mining*.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996b). From Data Dining to Knowledge Discovery in Databases. *AI Magazine*, 37.

Ferrari, A., & Sangiovanni-Vincentelli, A. (1999). System Design: Traditional Concepts and New Paradigms. *Iccd*, 2.

Finger, S., & Dixon, J. R. (1989). A Review of Research in Mechanical Engineering Design. Part I: Descriptive, Prescriptive, and Computer-Based Models of Design Processes. *Research in Engineering Design*, *1*, 51–67.

Fowler, M., & Highsmith, J. (2001). The Agile Manifesto. *Software Development*, *9*, 28–35.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine*, *13*, 57.

Fry, B. J. (2004). Computational Information Design. *SciencesNew York*, *Ph. D*(May 1997), 170.

Gadrey, J. (1991). Le Service n'est pas un produit : quelques implications pour l'analyse économique et pour la gestion. *Politiques et Management Public*, *9*(1), 1–24.

Gautier, P., & Gonzaler, L. (2011). *L 'Internet des objets.*

Goldstein, S. M., Johnston, R., Duffy, J., & Rao, J. (2002). The service Concept: the Missing Link in Service Design Research ? *Journal of Operations Management*, *20*(2), 121–134.

Gombault, V. (2011). *Deux ménages sur trois disposent d'internet chez eux*. INSEE.

Goodwin, K. (2011). *Designing for the Digital Age: How to Create Human-centered Products and Services*.

Gwinner, K., Gremler, D., & Bitner, M. (1998). Relational Benefits in Services Industries: the Customer's Perspective. *Journal of the Academy of Marketing*.

Han, J., & Kamber, M. (2006). Data Mining: Concepts and Techniques. *University of Illinois at Urbana-Champaign*.

Heskett, J. L., Sasser, W. E., & Schlesinger, L. A. (1997). *The Service Profit Chain*. *Handbook of Service Science* (Vol. 38, p. 320). Free Press.

Huang, G. (1996). Design for X-Concurrent Engineering Imperatives. *Springer, Berlin*.

Huang, Y., & Li, G. (2010). Descriptive Models for Internet of Things. In *Intelligent Control and Information Processing (ICICIP)* (pp. 483–486).

IWStats. (2009). INTERNET USAGE STATISTICS The Internet Big Picture. *World Internet Users and 2015 Population Stats*.

Jacobs, A. (2009). The Pathologies of Big Data. *Communications of the ACM*.

Japkowicz, N. (2006). Why Question Machine Learning Evaluation Methods. *AAAI Workshop on Evaluation Methods for Machine Learning*, 6–11.

Jiao, J. R., Simpson, T. W., & Siddique, Z. (2007). Product Family Design and Platform-Based Product Development: A State-of-the-Art Review. *Journal of Intelligent Manufacturing*, *18*, 5–29.

Jones, S. (2005). Toward an Acceptable Definition of Service. *Software, IEEE*, (June), 87–93.

Kellmereit, D., & Obodovski, D. (2013). *The Silent Intelligence - the Internet of Things*.

Keutzer, K., Rabaey, J. M., & Sangiovanni-Vincentelli, A. (2000). System-Level Design: Orthogonalization of Concerns and Platform-Based Design. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions*, *19*, 1523–1543.

Kim, J. (2011). *Modeling Cognitive and Affective Processes of Designers in the Early Stages of Design : Mental Categorization of Information Processing*. Arts & Métiers ParisTech.

Kinard, B. R., & Capella, M. L. (2006). Relationship Marketing: the Influence of Consumer Involvement on Perceived Service Benefits. *Journal of Services Marketing*, *20*(6), 359–368.

Klosgen, W., & Zytkow, J. (1996). Knowledge Discovery in Databases Terminology. *Advances in Knowledge Discovery and Data Mining*, 573–592.

Kotler, P., & Connor, R. a. (1977). Marketing Professional Services. *Journal of Marketing*.

Kurgan, L. A., & Musilek, P. (2006). A Survey of Knowledge Discovery and Data Mining Process Models. *Cambridge University Press*, *21*, 1–24.

Lahonde, N. (2010). *Optimisation du Processus de Conception: proposition d'un modèle de selection des méthodes pour l'aide à la décision*. Arts & Metiers ParisTech.

Le Coq, M. (1992). *Approche intégrative en conception de produits*. Arts & Metiers ParisTech.

Lim, D. (2003). *Modélisation du processus de conception centrée utilisateur, basée sur l'intégration des méthodes et outils de l'ergonomie cognitive : Application à la conception d'IHM pour la télévision interactive*. Arts & Métiers ParisTech.

Maimon, O., & Rokach, L. (2005). Data Mining and Knowledge Discovery Handbook. *New York: Springer*.

Martin, G., & Chang, H. (2012). *Winning the SoC Revolution: Experiences in Real Design*. (Springer Science & Business Media, Ed.).

Mitchell, T. M. (1977). Version Spaces: A Candidate Elimination Approach to Rule Learning. *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, *1*, 305–310.

Morelli, N. (2002). Designing Product/Service Systems: A Methodological Exploration1. *Design Issues*, *18*(3), 3–17.

Morelli, N. (2006). Developing New Product Service Systems (PSS): Methodologies and Operational Tools. *Journal of Cleaner Production*, *14*, 1495–1501.

Moritz, S. (2005). *Service Design: Practical Access to an Evolving Field*.

Mougenot, C. (2008). Modélisation de la phase d'exploration du processus de conception de produits, pour une créativité augmentée. *Sciences New York*.

Muffatto, M. (1999). Introducing a Platform Strategy in Product Development. *Journal of Production Economics*, *60*, 145–153.

Oliva, R., & Kallenberg, R. (2003). Managing the Transition from Products to Services. *International Journal of Service Industry Management*, *14*(2), 160–172.

Osterwalder, A. (2004). The Business Model Ontology: A Proposition in a Design Science Approach. *Institut d'Informatique et Organisation. Lausanne.*

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K.-H. (2007). *Engineering Design: A Systematic Approach*. (K. Wallace, Ed.)*Springer* (p. 617). Springer.

Pasley, J. (2005). How BPEL and SOA are Changing Web Services Development. *IEEE Internet Computing*, *9*, 60.

Piateski, G., & Frawley, W. (1991). Knowledge Discovery in Databases. *MIT Press*.

Potts, G. W. (1988). Exploit your Product's Service Life Cycle. *Harvard Business Review*, *66*(5), 32–36.

Pugh, S. (1991). Total Design: Integrated Methods for Successful Product Engineering. *New York, Addison - Wesley*.

Pyle, D. (1999). *Data Preparation for Data Mining*.

Rae, J. (31Volts). (2009). *Innovation is Served*. Minister of Economic Affairs, Netherlands.

Royce, W. W. (1970). Managing the Development of large Software Systems. *Proceedings of IEEE WESCON*, *26*, 1–9.

Saco, R. M., & Goncalves, A. P. (2008). Service Design: An Appraisal. *Design Management Review*, *19*, 10.

Samuel, A. L. (19959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*, 210–229.

Sangiovanni-Vincentelli, A., & Martin, G. (2001). Platform-Based Design. *IEEE Design & Test of Computers*, *6*, 23–33.

Schwaber, K., & Sutherland, J. (2011). The Scrum Guide. *Scrum Alliance*.

Shostack, G. L. (1977). Breaking Free from Product Marketing. *The Journal of Marketing*, 73–80.

Simon, H. A. (1969). *The Sciences of the Artificial*.

Simpson, T. W., Siddique, Z., & Jiao, R. J. (2006). Product Platform and Product Family Design: Methods and Applications. *Springer Science & Business Media*.

Spiess, P., Karnouskos, S., Guinard, D., Savio, D., Baecker, O., Souza, L. M. S. D., & Trifa, V. (2009). SOA-based Integration of the Internet of Things in Enterprise Services. *Web Services, 2009. ICWS 2009. IEEE International Conference*, 968–975.

Stickdorn, M., & Schneider, J. (2010). *This is Service Design Thinking*.

Suh, N. P. (1990). The Principle of Design. *New York: Oxford University Press*.

Tan, A. (2010). Service-Oriented Product Development Strategies, (June).

Tan, L. (2010). Future Internet: The Internet of Things. *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference*, *5*, 376–380.

Thierer, A., & Castillo, A. (2015). Projecting the Growth and Economic Impact of the Internet of Things. *SSRN 2618794*.

Tichem, M. (1997). *A Sesign Coordination Approach to Design for X*. Delft University of Technology.

Tomiyama, T. (1997). A Note on Research Directions of Design Studies. *ICED'97, 11th International Conference on Engineering Design*.

Tomiyama, T. (2001). Service Engineering to Intensify Service Contents in Product Life Cycles. *Proceedings Second International Symposium on Environmentally Conscious Design and Inverse Manufacturing*, 613–618.

Tomiyama, T., Gu, P., Jin, Y., Lutters, D., Kind, C., & Kimura, F. (2009). Design Methodologies: Industrial and Educational Applications. *CIRP Annals-Manufacturing Technology*, *58*, 543–565.

Trela, M. (2014). *Optimisation des performances d'innovation : Une approche combinant inventivité technique et recherche du succès commercial.* Arts & Métiers ParisTech.

Tufle, E. (1983). The Visual Display of Quantitative Information. *CT Graphics, Cheshire*.

Tufle, E. (1990). *Envisioning Information*.

Ullman, D. G. (1992). *The Mechanical Design Process*. (N. Y. McGraw-Hill, Ed.).

Ulrich, K. T., & Eppinger, S. D. (2007). Product Design and Development. *New York, McGraw - Hill*.

Vargo, S. L., & Lusch, R. F. (2008). Service-Dominant Logic: Continuing the Evolution. *Journal of the Academy of Marketing Science*, *36*(1), 1–10.

Vargo, S. L., Lusch, R. F., & Morgan, F. W. (2006). Historical Perspectives on Service-Dominant Logic. In R. F. Lusch (Ed.), *The servicedominant logic of marketing Dialog debate and directions* (Vol. 468, pp. 29–42). M.E. Sharpe.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a Standard Process Model for Data Mining. *The 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, 29–39.

Wu, M., Lu, T. L., Ling, F. Y., Sun, L., & Du, H. Y. (2010). Research on the Architecture of Internet of Things. In *Advanced Computer Theory and Engineering (ICACTE)* (pp. 484–487). IEEE.

Zheng, J., Simplot-Ryl, D., Bisdiklan, C., & Mouftah, H. (2011). The Internet of Things. *IEEE Communications Magazine. Novembler 2011.*, 30–31.

Annexe A : Projet Hybride 2.0

Il s'agit d'un projet de développement industriel d'un appareil hybride communiquant. Le projet est co-labellisé par les pôles Advancity et Systématic. Les problématiques du projet sont liées à :

- L'optimisation des performances d'un générateur de chaleur innovant (20% d'équivalent CO2 et de consommation d'Energie Primaire en moins face à une chaudière à condensation)
- L'augmentation de la part d'énergie renouvelable utilisée (objectif RT)
- La réduction du bruit (puissance sonore et qualité acoustique)
- La mise en œuvre de la théorie d'innovation TRIZ
- L'anticipation des consommations d'énergie à venir
- La gestion intelligente du chauffage pour l'optimisation du confort et la maîtrise de l'énergie à l'échelle de la maison et du bâtiment

Le projet Hybride 2.0, conscient du manque de solutions multi-énergies sur le marché des chaudières dans le résidentiel, propose de concevoir et développer un système de chauffage hybride monobloc (combinant une pompe à chaleur électrique et une chaudière gaz à condensation), et pour ce faire de lever les verrous technologiques relatifs aux questions d'encombrement restreint, d'atteinte d'une haute efficience énergétique et de minimiser les nuisances sonores pour un équipement couplant une pompe à chaleur à alimentation électrique et une chaudière à gaz à condensation avec son système de pilotage intégré selon les caractéristiques respectives des énergies disponibles. Le réseau d'expertises constitué dans le cadre de l'appel à projet permettra de valider et d'enrichir le contenu technique du projet. Ce projet vise également à relier le monde de l'énergie auquel e.l.m. leblanc appartient à celui des Technologies de l'Informatique et des Communications.

La mise en adéquation avec l'appel à projet doit permettre d'accélérer le développement en s'inscrivant dans des dynamiques plus générales nationales et internationales. Ce projet est une diversification et une création d'activité pour les PME concernées. Le financement public constitue un élément déterminant dans le développement de ce projet. La non obtention de ces fonds rendrait sa réalisation problématique, eu égard aux sommes que représente le projet Hybride 2.0 pour les 3 PME concernées.

L'objectif du projet Hybride 2.0 est de mener à son terme une étude de la faisabilité et le développement d'un appareil de chauffage hybride compact de seconde génération ayant les propriétés suivantes :

- Performances énergétiques supérieures (COP 7/35 de 4) permettant en particulier d'atteindre le niveau de 5kWh/m²/an de production d'énergie à partir de sources renouvelables
- Equipé d'une régulation intelligente permettant d'anticiper les besoins énergétiques de l'habitat et d'optimiser le couplage entre les parties pompe à chaleur et chaudière gaz à condensation
- Connecté à une plateforme de services
- Niveau acoustique très inférieur à la version actuelle (de l'ordre de 9 dBA de moins – soit huit fois moins bruyant) et haute qualité sonore, y compris en régime transitoire

Pour mener à bien le projet, l'approche suivante est retenue :

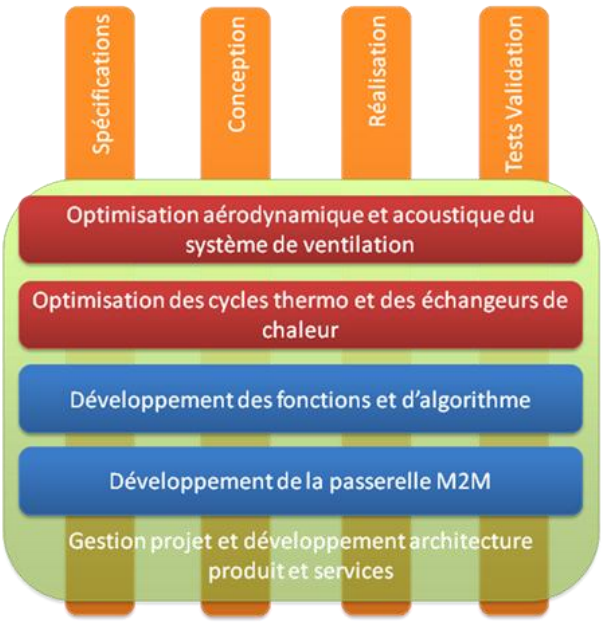
- Décomposition du projet en 4 axes majeurs : Aéro-Acoustique / Thermodynamique / Régulation Développement de fonctions et d'algorithmes / Développement d'une passerelle M2M
- Trois thèses de doctorat encadrées par des laboratoires spécialisés pour les axes Acoustique, Thermodynamique et Algorithmique.
- Encadrement de la démarche de résolution de problème par un universitaire expert (utilisation de TRIZ)
- Une PME pour la conception et la réalisation de groupes froids
- Une PME pour la conception et la réalisation de cartes électroniques
- Deux grands groupes pour la gestion du projet de développement et le suivi des démonstrateurs terrain

La figure ci-dessous représente le découpage du projet selon quatre thématiques :

- Optimisation acoustique
- Optimisation thermodynamique
- Développement algorithmique
- Développement de passerelle M2M

et quatre grandes phases de déroulement du projet :

- Spécifications
- Conception
- Réalisation
- Tests et validation



Pour être plus précis, la réalisation du projet a été segmentée en 31 tâches réparties sur l'ensemble des acteurs du projet et sur les quatre lots d'avancement comme le montre le tableau ci-dessous. Les travaux présentés dans le cadre de cette thèse sont représentés par les axes de développement caractérisés en

bleu, à savoir le développement des fonctions et d'algorithmes et le développement d'une passerelle M2M.

Partenaire	Lot 1 Spécifications	Lot 2 Conception	Lot 3 Fabrication	Lot 4 Tests
DynFluid Laboratoire de Dynamique des Fluides	Tâche 1 : CDCF produit (PAC et thermodynamique) **Tâche 3** : Etat de l'art turbomachines et acoustique **Tâche 4** : CDCF banc d'essai	**Tâche 9** : Solutions aéroacoustiques optimisées **Tâche 10** : Caractérisation numérique **Tâche 11** : Conception banc	**Tâche 19** : Fabrication banc **Tâche 20**: Réalisation de maquettes (aéraulique) et prototypes de machines	**Tâche 26** : Tests et validation de solutions (aéraulique et psychoacoustique) et caractérisation expérimentale
IFI	Tâche 1 : CDCF produit (PAC et thermodynamique) **Tâche 5** : Etat de l'art cycles thermos et échangeurs **Tâche 6** : CDCF banc d'essai	**Tâche 12** : Solutions échangeurs et cycles **Tâche 13** : Conception banc	**Tâche 21** : Fabrication banc **Tâche 22** : Réalisation de maquettes (thermodynamique) et prototypes de machines	**Tâche 27** : Tests et validation de solutions (cycle et échangeur)
Labo CPI	Tâche 2 : CDCF services **Tâche 7** : Etat de l'art algorithmique	**Tâche 14** : Conception services et algorithmes Tâche 15: Conception plateforme de service	**Tâche 23** : Prototypage des services et algorithmes	**Tâche 28** : Tests et validation des nouveaux services
DELTA DORE	Tâche 1 : CDCF produit Tâche 2 : CDCF services **Tâche 8** : CDCF M2M	**Tâche 16**: Conception M2M Tâche 15: Conception plateforme de service	**Tâche 24**: Intégration de la passerelle M2M	**Tâche 29** : Tests et validation passerelle M2M
GEL-KIT	Tâche 1 : CDCF produit (PAC et thermodynamique)	**Tâche 17** : Conception banc de test PAC Tâche 18 : Conception architecture	Tâche 25 : Réalisation de maquettes fonctionnelles **Tâche 21** : Formation syst. thermodynamiques	
GDF SUEZ	Tâche 1 : CDCF produit Tâche 2 : CDCF services			**Tâche 30** : Essais démonstrateurs terrains et enceinte climatique Tâche 31 : Tests et validation de solutions
e.l.m. leblanc	**Tâche 0** : Gestion de projet **Tâche 1** : CDCF produit **Tâche 2** : CDCF services	Tâche 9 & 12 : Solutions alternatives optimisées Tâche 14 : Concept. services **Tâche 15**: Conception plateforme de service **Tâche 18** : Conception architecture	**Tâche 25** : Réalisation de maquettes fonctionnelles Tâche 21 : Formation syst. thermodynamiques	**Tâche 31** : Tests et validation de solutions (appareil complet et services)

Annexe B : Projets de Master Recherche

Au cours du projet Hybride 2.0 et avec l'implication du LCPI, 5 projets de fin d'études ont été conduits avec des étudiants de master de recherche. Ces projets ont participé à la réalisation des expérimentations 1 et 2 de cette thèse. Nous présentons dans cette annexe les étudiants ayant participé à ce projet, leurs sujets de recherche et leur implication dans les travaux présentés dans nos expérimentations.

Chloé Romat

Profil :	Ingénieur
Sujet :	Conception de services innovants pour la maîtrise des énergies
Livrables :	- Processus d'obtention du *service blueprint* - Première version du *service blueprint*
Lien avec la thèse :	Cette recherche a constitué la base de travail pour la réalisation de l'expérimentation 1. Nous avons travaillé ici sur le choix des outils, la mise en application des méthodes, la collecte des idées, la formalisation des concepts et la réalisation de la première version du *service blueprint*.

Guillaume Gorse

Profil :	Designer
Sujet :	Méthode agile et expérience utilisateur pour la conception d'une interface d'un service innovant
Livrables :	- Processus de conception d'interface agile - Prototype d'interface pour les fonctions du client final
Lien avec la thèse :	Cette recherche basée sur le *service blueprint* a été mise en avant à la fin de l'expérimentation 1. Cela constitue une base de validation pour la pertinence des spécifications rendues par l'analyse du service lorsque nous passions dans l'étape de conception détaillée.

Rodrigo Coronado

Profil :	Ingénieur
Sujet :	Conception d'un système pour l'analyse des habitudes de consommation
Livrables :	- Processus de conception agile d'un algorithme d'apprentissage - Analyse des profils d'utilisation dans les données - Mise en lumière de certaines inconsistances dans les données disponibles
Lien avec la thèse :	Cette recherche a été menée en amont de l'expérimentation 2. Une partie des travaux réalisés dans le cadre de ce projet a servi d'amorce pour la première itération de notre expérimentation. La mise en lumière de certaines inconsistances dans les données nous a permis de trouver des solutions palliatives et d'améliorer par la suite la qualité de nos livrables.

Eloïse Rulquin

Profil :	Designer
Sujet :	Conception agile d'une interface orientée métiers
Livrables :	- Processus de conception agile de conception d'une interface - Prototype d'interface pour les fonctions du service après-vente
Lien avec la thèse :	De la même manière que le projet de Guillaume Gorse, cette recherche basée sur le *service blueprint* a été mise en avant à la fin de l'expérimentation 1. Cela constitue une base de validation pour la pertinence des spécifications rendues par l'analyse du service lorsque nous passions dans l'étape de conception détaillée.

Maxime Hemery

Profile :	Ingénieur
Sujet :	Fouille de données appliquée à un projet de maintenance prévisionnelle
Livrables :	- Modèle économique d'une fonction de maintenance intelligente - Processus de conception d'un modèle de données basé sur une collecte complémentaire dans un environnement contrôlé - Évaluation des performances des attributs calculés
Lien avec la thèse :	Cette recherche représente la troisième itération réalisée dans l'expérimentation 2. Cela nous a permis de mettre en place la validation de nos indicateurs avec la collecte de nouvelles données. Nous avons également pu considérer l'interprétation des indicateurs de performance d'un modèle de données dans un plan de rentabilité.

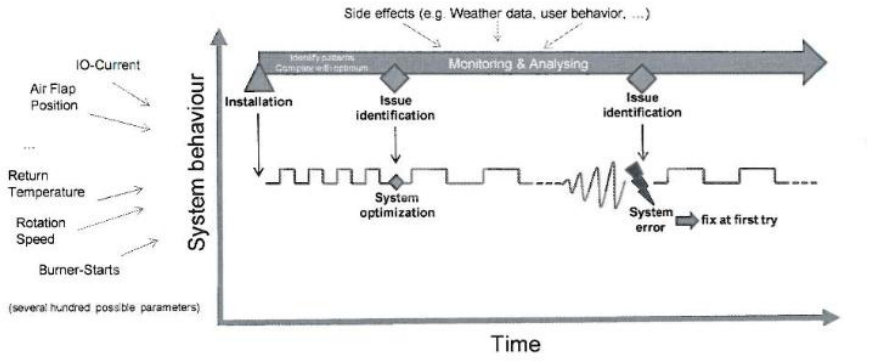
Annexe C : Projet TOP79

Ce projet s'inscrit dans la vision de l'internet des objets et le positionnement de Bosch Thermotechnologie pour fournir de nouveaux services à ses clients, installateurs comme utilisateurs finaux, par la mise en réseau de ses appareils de production de chauffage et d'eau chaude.

Dans cette dynamique, le nombre de produits connectés présents dans les catalogues du groupe augmente tous les ans, ainsi que la quantité de données que ces produits génèrent. Les technologies du *Big Data* et du *Data Mining* constituent la base pour la fourniture de ces nouveaux services. L'objectif du projet TOP79 est de développer deux fonctions principales de ces services sur la base des données déjà présentes dans le groupe. Ces deux fonctions sont l'identification et la prédiction des pannes.

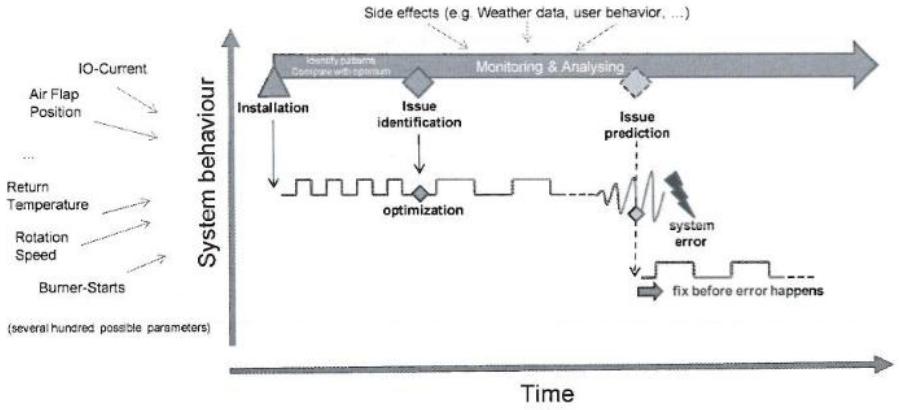
Le but de la fonction d'identification des pannes consiste à :

- Identifier les appareils dont les performances sont faibles
- Identifier les causes probables d'une panne basée sur les données de fonctionnement collectées
- Classer les possibles causes d'une panne afin de constituer une liste de recommandations pour l'utilisateur



Le but de la fonction de prédiction de panne consiste à :

- Identifier les potentiels problèmes pouvant survenir sur un appareil avant leur apparition.



Ces deux fonctions ne peuvent pas être considérées séparément. Cependant la logique voudrait réaliser l'identification des pannes avec une grande précision en premier avant de commencer à s'intéresser à la prédiction de ces pannes.

Les livrables de ce projet s'organisent de la manière suivante :

Livrable	Tâche / Description
Disponibilité des données	- Collecte de la vérité terrain - Labellisation des données
Importation et transformation des données	- Visualisation des données - Création de jeux de données pour l'entraînement, le test et la validation des modèles - Développement d'un outil d'importation des données brutes
Première revue des résultats avec Bosch Thermotechnologie	- Préparation de présentation et de démonstration montrant le statut du projet et les prochaines étapes
Prototypage d'un environnement de traitement	- Préparation d'une plateforme de traitement pour accueillir les fonctionnalités - Test de l'acceptation client
Développement de modèles I	- Suivi et optimisation des algorithmes - Analyse des causes des pannes - Évaluation de l'acceptation de l'algorithme
Deuxième revue des résultats avec Bosch Thermotechnologie	- Préparation d'une plateforme de traitement pour accueillir les fonctionnalités - Test de l'acceptation client
Développement de modèles II	- Prédiction des causes - Évaluation de l'acceptation de l'algorithme
Troisième revue des résultats avec Bosch Thermotechnologie	- Préparation d'une plateforme de traitement pour accueillir les fonctionnalités - Test de l'acceptation client
Intégration des modèles	- Intégration des résultats dans la plateforme de Bosch Thermotechnologie - Recommandation pour l'intégration et la mise à jour des modèles - Formation et documentation
Conclusion finale	- Résumé des résultats du projet - Retour d'expérience

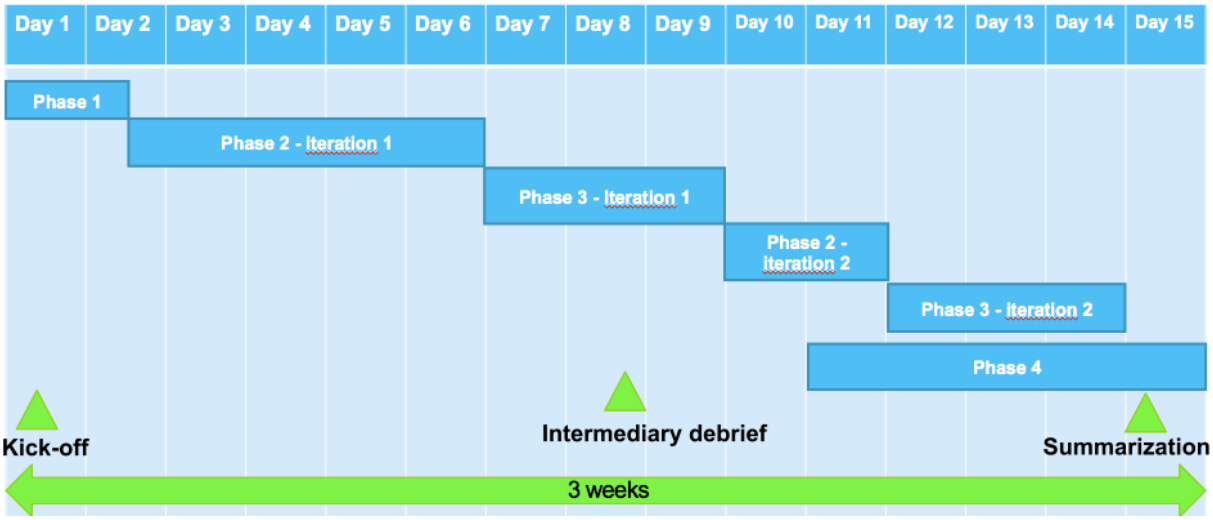
Annexe D : Projet « *Hackfest* Microsoft »

Ce projet s'inscrit dans une logique exploratoire de données. Ici, l'ensemble des données de production exposées dans l'expérimentation 2. L'idée est de faire appel à des connaissances extérieures en science des données pour mener à bien une preuve de concept d'utilisation des technologies du *Machine Learning* pour un cas d'application ciblé.

Le projet s'articule autour des objectifs et livrables suivants :

Nom	Description
Rédaction d'un cahier des charges client	Réalisation d'un document détaillant : - Un résumé des cas d'usage envisagés avec leurs buts et objectifs - Les pistes envisagées pour réaliser chaque cas d'usage Cela représente les idées et décisions prises lors d'un premier atelier.
Développement de modèles	Apprentissage d'un modèle pour le cas d'usage retenu
Visualisation des résultats	Fourniture d'un tableau de bord représentant les résultats de l'application du modèle
Présentation des résultats	Présentation de l'ensemble des résultats du projet ainsi que d'un retour d'expérience sur le développement du modèle.

La réalisation de ce projet s'inscrit dans une période de temps prédéfinie, ici trois semaines. Sur cette base de temps sont réparties deux itérations de développement et trois revues de projets.



Les itérations sont structurées suivant quatre phases de développement, dont les activités sont présentées ci-après.

Phase	Activité de Microsoft	Activité du client
Phase 1	- Définition des critères de succès en accord avec le client - Sélection d'un cas d'usage - Sélection des sources de données pertinentes - Exploration des données	- Définition de la valeur métier relative aux cas d'usage - Fourniture d'un jeu de données - Transfert de compétence sur la compréhension des données et leur exploration
Phase 2	- Analyse fine de la documentation relative aux données - Importation des données dans un environnement d'analyse - Préparation des données - Statistique descriptive des données	- Participation à l'exploration des données
Phase 3	- Création d'un modèle - Entraînement et test du modèle sur les données sélectionnées - Ajustement des paramètres de l'algorithme d'apprentissage - Présentation des résultats intermédiaires	- Participation à la validation intermédiaire du modèle
Phase 4	- Création d'un tableau de bord incluant les résultats du modèle - Présentation des résultats de l'itération	- Participation à l'intégration du modèle dans les tableaux de bord

Annexe E : Exemple de personas générés en phase de clarification de l'expérimentation 1

Arnaud



28 ans, marié et bientôt papa, prof d'anglais au lycée

Habite à Paris dans un appartement du 13ème ***(locataire)*** *avec Virginie sa compagne*

Attentes principales : Impact environnemental, Qualité, Coût

Activités

- Arnaud aime la lecture et les jeux vidéos.
- Il aborde avec ses élèves des sujets d'actualités comme le développement durable.
- Il sort régulièrement avec ses collègues de travail et avec ses amis dans Paris.

Par rapport au système de chauffage

- « J'aimerai mieux contrôler ma consommation, niveau coûts et impact écologique mais de manière simple. »
- Il est locataire et aimerait que son propriétaire fasse des travaux d'isolation.
- Il ne connait pas particulièrement le fonctionnement d'une chaudière.

Profil techno

- Il possède un Smartphone et est présent sur les réseaux sociaux (Twitter, AgregIn) et visite régulièrement les sites d'actualités et forums sur lesquels il laisse régulièrement des commentaires

BOSCH

Tuyen (alias Tutu)



23 ans, étudiante en master d'histoire

Habite en colocation à Grenoble

Attentes principales : Coût, Impact environnemental

Activités

- Tuyen aime la randonnée et les sports d'aventure entre amis
- Elle fait de la danse tous les mardis soir
- Elle donne des cours particuliers le samedi matin pour gagner un peu d'argent
- Elle milite au sein d'une association pour l'amélioration de la qualité de vie

Par rapport au système de chauffage

- Tuyen n'y connaît rien et ses coloc' non plus.
- Elle aimerait que ça lui coûte le moins possible.
- Elle est sensible à son impact écologique et cherche à limiter son « bilan carbone perso »

Profil techno

- Elle possède un ordinateur portable pour prendre ses cours et un Smartphone. Si un jour, elle a un peu plus d'argent, « j'achèterai une tablette sans hésiter ! »
- Elle est très présente sur les réseaux sociaux : Facebook, Twitter, Instagram… Elle tient aussi un blog alimenté de temps en temps

BOSCH

Annexe F : Exemple de fiches concept issues de la phase de génération de concept de l'expérimentation 1

Client **2. Choix, modification et création de programmes horaires**

- Fonctionnalités :
  Avoir des suggestions de paramétrage selon : le type d'habitation, le profil du client, des règles métiers…
  Avoir plusieurs configurations proposées
  Pouvoir créer des programmes selon mon niveau (débutant, avancé, expert) avec une aide interactive
  Pouvoir partager ces programmes sur les réseaux sociaux
- Valeur qualitative :
  Personnalisation, Appropriation du système par la création de programmes, Favorisation du dialogue inter-clients
- Valeur financière :
  Eventuellement limite du gaspillage d'énergie donc économie d'argent
- Notion de temps :
  Réalisable entièrement maintenant
  Marché mûr



5 BOSCH

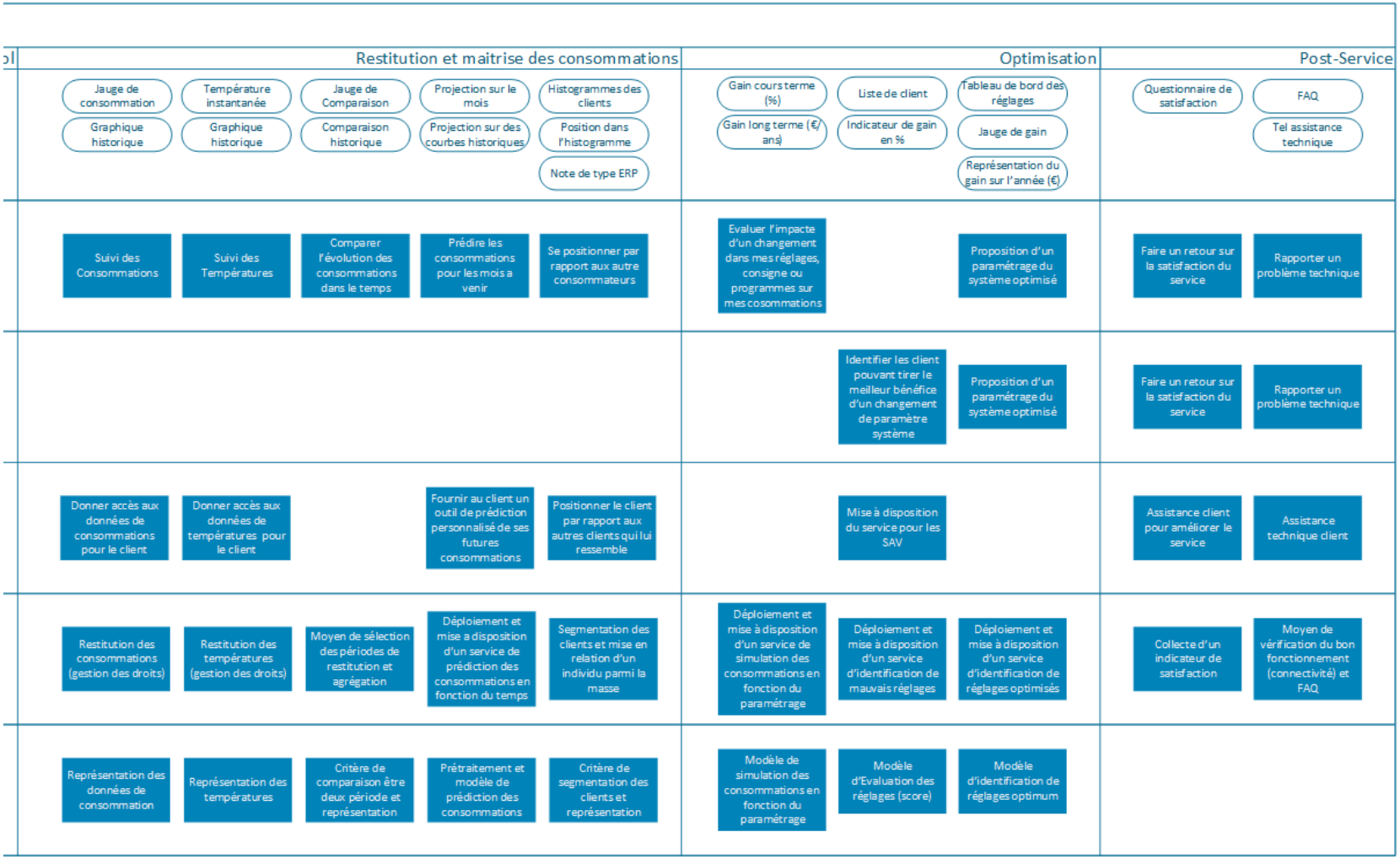
SAV **4. Pré-diagnostic des interventions (à distance)**

- Fonctionnalités:
  - Qualifier à distance les pannes et les défauts
  - Anticiper les actions a faire sur une intervention (pièces a changer, temps d'intervention…)
- Valeur qualitative:
  - Permettre au technicien de mieux préparer son matériel avant une intervention
  - Valorisation des compétences du SAV
- Valeur financière:
  - Limiter les visites qui aboutissent uniquement à des devis (et donc limiter les revisites)
  - Offrir la possibilité de mieux gérer les stocks mobiles pour les techniciens
- Notion de temps:
  - Réalisable une fois les règles de diagnostic connues.

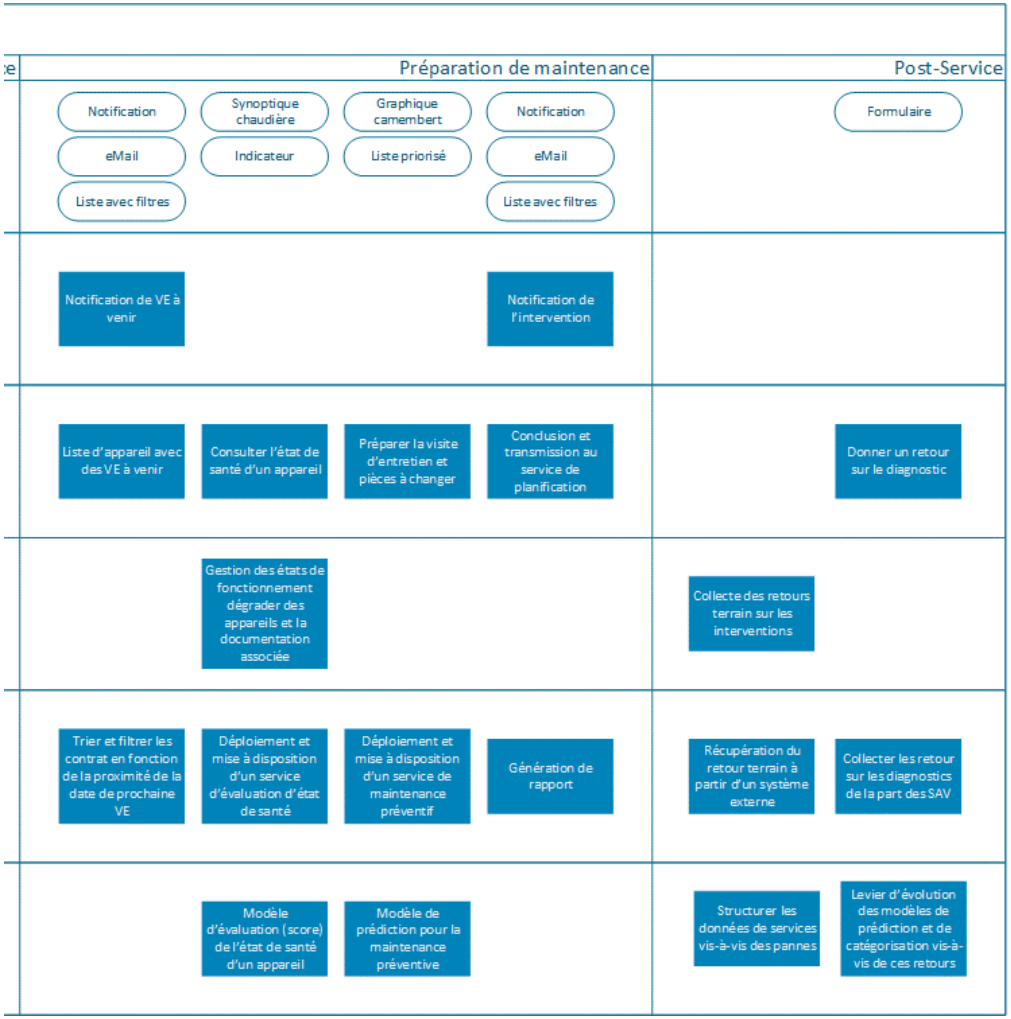


BOSCH

Annexe G : *Service blueprint* pour les fonctionnalités du client final


ol
Restitution et maitrise des consommations
Optimisation
Post-Service
Jauge de consommation
Température instantanée
Jauge de Comparaison
Projection sur le mois
Histogrammes des clients
Graphique historique
Graphique historique
Comparaison historique
Projection sur des courbes historiques
Position dans l'histogramme
Note de type ERP
Gain cours terme (%)
Liste de client
Tableau de bord des réglages
Gain long terme (€/ans)
Indicateur de gain en %
Jauge de gain
Représentation du gain sur l'année (€)
Questionnaire de satisfaction
FAQ
Tel assistance technique
Suivi des Consommations
Suivi des Températures
Comparer l'évolution des consommations dans le temps
Prédire les consommations pour les mois a venir
Se positionner par rapport aux autre consommateurs
Evaluer l'impacte d'un changement dans mes réglages, consigne ou programmes sur mes cosommations
Proposition d'un paramétrage du système optimisé
Faire un retour sur la satisfaction du service
Rapporter un problème technique
Identifier les client pouvant tirer le meilleur bénéfice d'un changement de paramètre système
Proposition d'un paramétrage du système optimisé
Faire un retour sur la satisfaction du service
Rapporter un problème technique
Donner accès aux données de consommations pour le client
Donner accès aux données de températures pour le client
Fournir au client un outil de prédiction personnalisé de ses futures consommations
Positionner le client par rapport aux autres clients qui lui ressemble
Mise à disposition du service pour les SAV
Assistance client pour améliorer le service
Assistance technique client
Restitution des consommations (gestion des droits)
Restitution des températures (gestion des droits)
Moyen de sélection des périodes de restitution et agrégation
Déploiement et mise a disposition d'un service de prédiction des consommations en fonction du temps
Segmentation des clients et mise en relation d'un individu parmi la masse
Déploiement et mise à disposition d'un service de simulation des consommations en fonction du paramétrage
Déploiement et mise à disposition d'un service d'identification de mauvais réglages
Déploiement et mise à disposition d'un service d'identification de réglages optimisés
Collecte d'un indicateur de satisfaction
Moyen de vérification du bon fonctionnement (connectivité) et FAQ
Représentation des données de consommation
Représentation des températures
Critère de comparaison être deux période et représentation
Prétraitement et modèle de prédiction des consommations
Critère de segmentation des clients et représentation
Modèle de simulation des consommations en fonction du paramétrage
Modèle d'Evaluation des réglages (score)
Modèle d'identification de réglages optimum

Annexe H : Service blueprint pour le service après-vente

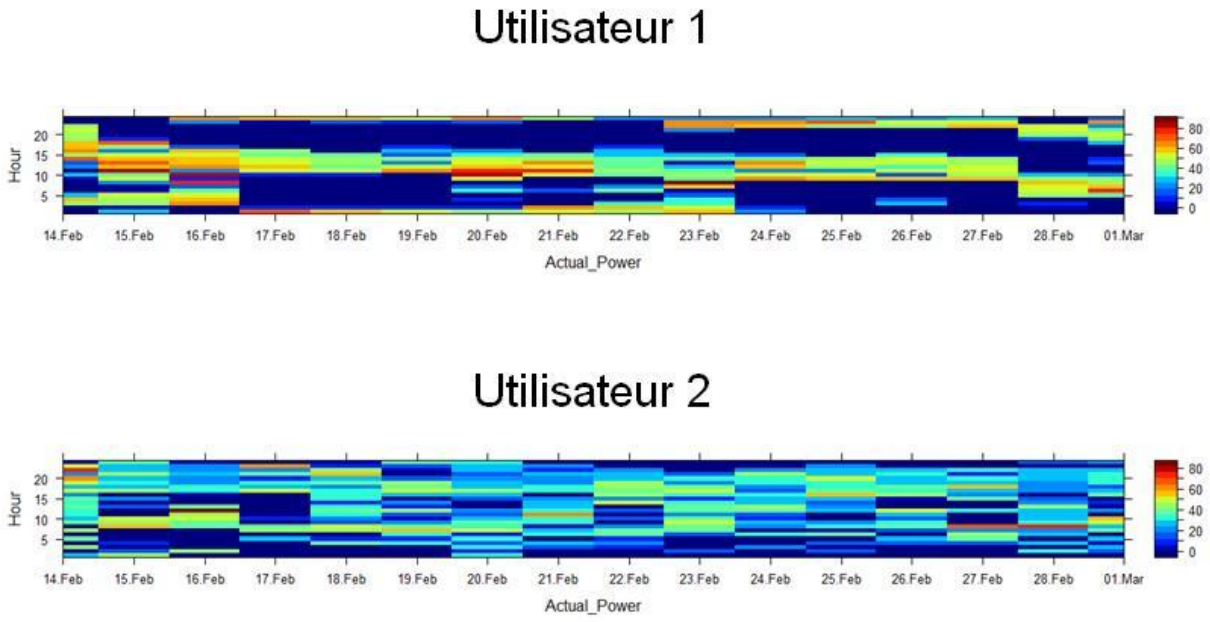

Préparation de maintenance
Post-Service
Notification
eMail
Liste avec filtres
Synoptique chaudière
Indicateur
Graphique camembert
Liste priorisé
Notification
eMail
Liste avec filtres
Formulaire
Notification de VE à venir
Notification de l'intervention
Liste d'appareil avec des VE à venir
Consulter l'état de santé d'un appareil
Préparer la visite d'entretien et pièces à changer
Conclusion et transmission au service de planification
Donner un retour sur le diagnostic
Gestion des états de fonctionnement dégrader des appareils et la documentation associée
Collecte des retours terrain sur les interventions
Trier et filtrer les contrat en fonction de la proximité de la date de prochaine VE
Déploiement et mise à disposition d'un service d'évaluation d'état de santé
Déploiement et mise à disposition d'un service de maintenance préventif
Génération de rapport
Récupération du retour terrain à partir d'un système externe
Collecter les retour sur les diagnostics de la part des SAV
Modèle d'évaluation (score) de l'état de santé d'un appareil
Modèle de prédiction pour la maintenance préventive
Structurer les données de services vis-à-vis des pannes
Levier d'évolution des modèles de prédiction et de catégorisation vis-à-vis de ces retours

Annexe I : Exploration des données réalisée dans l'étape 2 de l'expérimentation 1

L'étape 2 de l'expérimentation 1 visait à comprendre les données qui nous ont été fournies dans le cadre de l'itération 2. Dans ce but, nous avons, avec les experts en science des données, mené plusieurs approches afin de décrire le contenu du jeu de données. Parmi ces approches, nous décrivons dans ces annexes les travaux réalisés sur la génération de profils utilisateurs et l'analyse de transition de statut.

Génération de profils utilisateur :

Nous avons cherché à caractériser par cette approche les différences que nous pouvions observer dans le fonctionnement des appareils en fonction de la demande de l'utilisateur. Pour cela, nous disposons dans les données d'un identifiant unique pour chaque produit afin de les distinguer les uns des autres, un attribut représentant le taux de modulation du brûleur (la puissance instantanée fournie par l'appareil) ainsi qu'un index temporel de référence. Nous pouvons alors extraire de cela des cartes, ou des profils d'utilisation comme le montre la figure ci-dessous.



Ces profils montrent des différences nettes entre les fonctionnements des appareils installés chez l'utilisateur 1 et 2. Nous pouvons noter que l'utilisateur 1 chauffe peu de temps dans la journée, mais à des puissances élevées. Alors que l'utilisateur 2 chauffe très souvent dans la journée, mais avec des puissances faibles. Après confrontation avec les experts du domaine d'application, nous pouvons conclure que cela est principalement dû à l'installation et à l'usage. L'utilisateur 2 possède probablement un régulateur d'ambiance plus perfectionné que l'utilisateur 1, des radiateurs probablement différents et

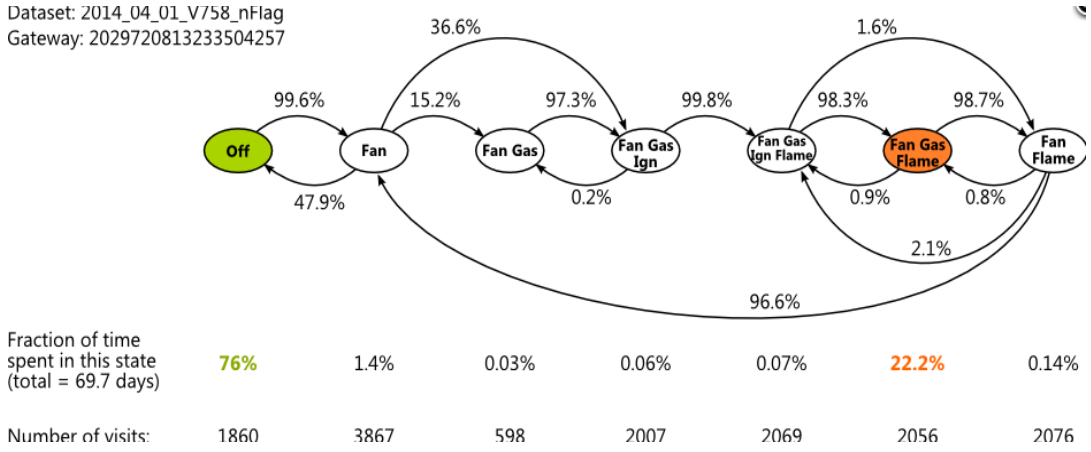
il semble avoir un besoin plus important en terme de chauffage dans le temps. Cependant, dans le cadre de la maintenance, nous pouvons noter que le comportement de l'appareil chez l'utilisateur 2 semble plus sain pour éviter la fatigue de l'appareil.

Analyse de transition de statut :

D'autre part, nous avons cherché à modéliser les procédures d'allumage et d'arrêt des appareils. Pour cela, nous avons à disposition dans les données un ensemble de booléens représentant l'état du système à chaque instant. Parmi ces états, nous avons :

- Fan : Fonctionnement du ventilateur
- Gas : Vanne gaz ouverte
- Ign : Présence d'un train d'étincelles dans le brûleur
- Flame : Détection de flamme dans le brûleur

À travers ces attributs, nous pouvons caractériser quel est l'état de l'appareil à tout instant et définir les transitions que nous pouvons observer dans ces états au cours du temps. Cette analyse nous conduit à la modélisation représentée dans la figure ci-dessous.



Par cette analyse, nous pouvons observer les procédures standard du produit, tel qu'un arrêt avec la suite :

$$FanGasFlame \rightarrow Fan\ Flame \rightarrow Fan \rightarrow Off$$

Cependant, nous pouvons noter certaines transitions atypiques telles que :

$$FanGasFlame \rightarrow FanFlame \rightarrow FanGasFlame$$

Selon les experts du domaine d'application, la cotation des ces singularités et leur fréquence d'apparition peuvent être le signe d'un disfonctionnement d'un appareil. Cette représentation des états de l'appareil nous a alors conduits par la suite à la caractérisation des transitions conduisant à l'apparition d'un défaut.

Annexe J : Résultats de l'atelier créatif effectué lors de l'étape 3 de l'expérimentation 2

Lors de l'étape 3 de l'expérimentation 2, nous avons focalisé notre attention avec les experts du domaine d'application sur la caractérisation de deux cas de génération de codes d'erreur sur les appareils : le défaut d'allumage et la surchauffe. Afin d'expliquer comment nous pouvions mesurer et prévenir ces erreurs, nous avons cherché à décrire ces erreurs de la manière suivante :

- Description de la cause pouvant engendrer le défaut ciblé
- Caractérisation du comportement du produit induit par l'apparition progressive de la cause identifiée
- Mesure pouvant être prise pour percevoir ce comportement dans les données

Nous présentons ainsi les résultats obtenus dans cet atelier dans les deux tableaux suivants.

CAS : Défaut d'allumage – Erreur EA

Cause	Comportements	Mesures
Encrassement du brûleur	- Multiplication du nombre de tentatives d'allumage avant apparition de la flamme - Moins bonne combustion du gaz	- Nombre d'apparitions du statut de démarrage ne conduisant pas à une flamme - Nombre d'apparitions du statut d'allumage successif avant l'apparition de la flamme - Diminution progressive de la valeur du courant d'ionisation
Mauvaise évacuation des gaz brûlés	- Perte de la flamme en fonctionnement - Moins bonne combustion du gaz	- Perte de la flamme alors qu'une demande de chauffage est toujours présente et que la consigne n'est pas atteinte - Diminution progressive de la valeur du courant d'ionisation
Mauvaise évacuation des condensats	- Mauvaise évacuation des gaz brûlés - Diminution des échanges thermiques vers l'eau de chauffage	- Perte de la flamme alors qu'une demande de chauffage est toujours présente et que la consigne n'est pas atteinte - Diminution progressive de la valeur du courant d'ionisation - Une énergie plus importante nécessaire pour élever la température de l'eau de chauffage
Court-circuit des électrodes d'allumage	- Pas de flamme créée pendant l'allumage	- Pas de mesures possibles

CAS : Surchauffe – Erreur E9

Cause	Comportements	Mesures
Perte de pression dans le circuit de chauffage	- Montée en température plus rapide - Dépassement plus fréquent et plus important de la température de consigne - Cycles de fonctionnement plus courts et plus nombreux	- Vitesse de montée en température après un démarrage du brûleur - Dépassement de la consigne pour la température d'eau en sortie du brûleur - Temps de fonctionnement d'un cycle - Nombre de cycles de fonctionnement
Perte de pression ou membrane percée dans le vase d'expansion	- Fuite dans la soupape de sécurité - Perte de pression dans le circuit de chauffage	- Vitesse de montée en température après un démarrage du brûleur - Dépassement de la consigne pour la température d'eau en sortie du brûleur - Temps de fonctionnement d'un cycle - Nombre de cycles de fonctionnement
Augmentation des pertes de charge dans le circuit de chauffage	- Diminution du débit - Montée plus rapide en température en sortie du brûleur - Diminution des températures en sortie du brûleur plus lente après la disparition de la flamme	- Vitesse de montée en température après un démarrage du brûleur - Vitesse de chute de la température après un arrêt du brûleur
Mauvais fonctionnement de la pompe	- Diminution du débit - Montée plus rapide en température en sortie du brûleur - Diminution des températures en sortie du brûleur plus lente après la disparition de la flamme	- Vitesse de montée en température après un démarrage du brûleur - Vitesse de chute de la température après un arrêt du brûleur

MÉTHODOLOGIE DE CONEPTION D'UNE PLATEFORME DE SERVICES INTELLIGENTS

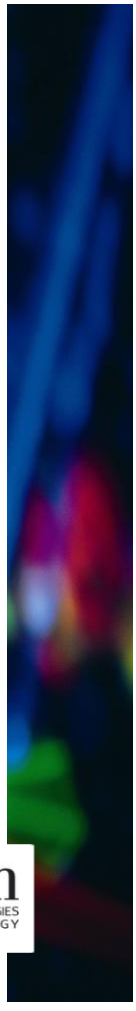
RESUME : Dans un paysage industriel de plus en plus tourné vers le numérique, les opportunités des entreprises ne manquent pas pour innover et répondre à une demande jusqu'alors inaccessible. C'est dans ce cadre que l'internet des objets apparait comme un élan technologique a fort potentiel. Ce levier d'innovation, basé sur la valorisation de flux de données, sont par nature intangible et c'est pourquoi nous les considérons ici comme des services. Cependant, les concepteurs doivent faire face ici à un univers complexe où de nombreux domaines d'expertise et de connaissance sont engagés. C'est pourquoi nous proposons dans cette thèse un modèle méthodologique de conception mettant en scène le service, l'expertise métier et les technologies de découverte de connaissance de manière optimisé pour concevoir à l'internet des objets. Ce modèle de conception a été éprouvé chez e.l.m. leblanc, entreprise du groupe Bosch, dans le développement d'un appareil de chauffage connecté et de ses services.

Mots clés : Conception, méthodologie de conception, innovation, internet des objets, science des données, data mining, conception de service, orienté données, gestion des connaissances

DESIGN METHODOLOGY FOR SMART SERVICES PLATEFORMS

ABSTRACT : In a more and more numeric oriented industrial landscape, the business opportunities for companies to innovate and answer needs inaccessible yep are increasing. In this framework, the internet of things appears as a high potential technology. This innovation lever, where the value-creation is principally based on the data, is not tangible by nature and this is the reason why we conceder it as a service in this thesis. However, the designer has to face a complex universe where a high number expertise and knowledge are engaged. This is the reason why we propose in this thesis a design methodology model organizing the service, the domain knowledge and the data discovery technologies in an optimized process to design the internet of things. This model has been experienced at e.l.m. leblanc, company of the Bosch group, in the development of a connected boiler and its services.

Keywords : design, design methodology, innovation, internet of things, data science, data mining, service design, data driven, knowledge management



ARTS ET MÉTIERS ParisTech

ParisTech INSTITUT DES SCIENCES ET TECHNOLOGIES PARIS INSTITUTE OF TECHNOLOGY