



Une démarche de gestion stratégique et opérationnelle du changement dans le contexte de l'exploitation avancée de données massives internes aux organisations

Mémoire

Ndeye Gueye

Maîtrise en informatique
Maître ès sciences (M. Sc.)

Québec, Canada

© Ndeye Gueye, 2017

Une démarche de gestion stratégique et opérationnelle du changement dans le contexte de l'exploitation avancée de données massives internes aux organisations

Mémoire

Ndeye GUEYE

Sous la direction de :

Bernard Moulin, directeur de recherche

Résumé

L'objectif de ce mémoire est d'élaborer une démarche de gestion du changement destinée aux organisations qui souhaitent mettre en place les transformations nécessaires pour tirer parti des techniques d'exploitation avancée de données massives. Le domaine de transformation choisi dans le cadre de ce mémoire est l'exploitation des données internes à une organisation pour développer une maturité technologique et organisationnelle, notamment en ce qui concerne l'orientation-client de son modèle d'affaires. Cette évolution de maturité a pour but de suivre la transformation de la société actuelle, fortement influencée par le mouvement Big Data. La démarche proposée s'adresse à des organisations dont la maturité technologique et organisationnelle est éloignée de celle des compagnies leaders de cette transformation « digitale ». Nous proposons un cadre méthodologique stratégique et opérationnel permettant aux organisations d'effectuer les transformations de culture, d'organisation, de gestion et d'adoption de techniques pour se lancer dans l'exploitation avancée de leurs données internes qui sont en forts volumes et mal structurées. L'application pratique de notre cadre méthodologique est illustrée dans le cas d'une grande compagnie d'assurance canadienne sur une période de deux ans. Nous avons également réalisé un projet pilote de 9 mois visant à tester la phase de mise en œuvre de la gestion stratégique du changement en exploitant les données-clients internes à cette compagnie d'assurance. Nous avons élaboré, raffiné et testé ces démarches en plusieurs itérations afin de proposer des méthodes détaillées, illustrées par des actions qui conduisent à des résultats concrets et qui mettent en évidence certains enjeux ou difficultés liés à l'adoption de ce type d'approches. À l'issue de notre projet, la démarche proposée a été adoptée par la compagnie à des fins d'opérationnalisation.

Table des matières

Résumé	i
Liste des tableaux	vii
Liste des figures	viii
Remerciements	x
1 Chapitre 1 : Mise en contexte du mémoire	1
1.1 Introduction	1
1.2 Présentation du contexte historique – une nouvelle révolution industrielle	2
1.2.1 L'apparition de technologies « disruptives »	2
1.2.2 Une transformation de la société en cours	5
1.2.3 Les organisations, au cœur de la transformation digitale actuelle.	8
1.3 Présentation du contexte théorique – l'exploitation des données massives	11
1.3.1 Les techniques de stockage	12
1.3.2 Les techniques de chargement, de traitement et de modélisation	15
1.3.3 Les techniques de visualisation	17
1.3.4 Les techniques d'analyse de données	20
1.4 Revue synthétique des nouvelles démarches d'exploitation de données destinées aux organisations	22
1.4.1 Approches méthodologiques	22
1.4.2 Modèles de maturité	25
1.5 La problématique de recherche	28
1.6 Présentation du contexte pratique – le projet de recherche	30
1.6.1 Les origines de notre projet de recherche appliqué BMP en partenariat avec iA	30
1.6.2 Présentation du secteur des assurances - historique et enjeux	31
1.6.3 Présentation de l'Industrielle Alliance (iA)	32
1.6.4 Démarche appliquée	37
1.6.5 Planification initiale du projet	38
1.7 Conclusion	39

2	Chapitre 2 : Démarche proposée aux organisations	40
2.1	Introduction	40
2.2	Présentation globale de la démarche proposée	40
2.2.1	La démarche de gestion du changement en quelques mots	40
2.2.2	Détails sur le déroulement du cycle de gestion stratégique du changement	43
2.2.3	Détails sur le déroulement du cycle d'opérationnalisation du changement	46
2.3	Raffinement de la préparation à la mise en œuvre de la gestion stratégique du changement	48
2.3.1	Présentation globale de cette phase	48
2.3.2	L'évaluation de la maturité	52
2.4	Raffinement de la mise en œuvre et de l'évaluation intermédiaire de la gestion stratégique du changement	54
2.4.1	Présentation globale de cette phase	54
2.4.2	L'élaboration de la vision stratégique du changement	56
2.4.3	L'élaboration du plan d'action	58
2.4.4	Mise en place de projets pilotes	60
2.5	Raffinement de l'évaluation finale du cycle de gestion stratégique du changement	63
2.5.1	Présentation globale de cette phase	63
2.6	Conclusion du chapitre	66
3	Chapitre 3 : Exemple de mise en œuvre de projets pilotes - cas du croisement de données clients	68
3.1	Introduction	68
3.2	Présentation de la démarche expérimentale généralisée de croisement de données clients	68
3.2.1	Présentation globale de la démarche	68
3.2.2	Présentation schématique détaillées des étapes clés de chaque phase	69
3.3	Phase 1 : Collecter	71
3.3.1	La méthode proposée	71
3.3.2	Les actions menées à l'iA	71
3.3.3	Les résultats obtenus	72

3.3.4	Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA	74
3.4	Phase 2 : Comprendre	75
3.4.1	La méthode proposée.....	75
3.4.2	Les actions menées à l'iA.....	75
3.4.3	Les résultats obtenus.....	76
3.4.4	Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA	77
3.5	Phase 3 : Cibler	78
3.5.1	La méthode proposée.....	78
3.5.2	Les actions menées à l'iA.....	79
3.5.3	Les résultats obtenus.....	79
3.5.4	Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA	85
3.6	Phase 4 : Préparer	86
3.6.1	La méthode	86
3.6.2	Les actions menées à l'iA.....	87
3.6.3	Les résultats obtenus.....	87
3.6.4	Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA	90
3.7	Phase 5 : Analyser	91
3.7.1	La méthode	91
3.7.2	Les actions menées à l'iA.....	93
3.7.3	Les résultats obtenus.....	94
3.7.4	Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA	97
3.8	Phase 6 : Interpréter.....	98
3.8.1	La méthode	98
3.8.2	Les actions menées à l'iA.....	98
3.8.3	Résultats obtenus	98
3.8.4	Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA	100
3.9	Les itérations à l'origine de l'élaboration de la démarche expérimentale proposée	101

3.9.1	Itération et leçons tirées de la phase de collecte	101
3.9.2	Itération et leçons tirées de la phase de compréhension des données	102
3.9.3	Itération et leçons tirées de la phase de définition de la cible	102
3.9.4	Itération et leçons tirées de la phase de préparation.....	103
3.9.5	Itération et leçons tirées de la phase d'analyse	103
3.9.6	Itération et leçons tirées de la phase d'analyse d'interprétation	104
3.10	Conclusion du chapitre.....	106
4	Chapitre 4 : Opérationnalisation du projet pilote et leçons apprises	109
4.1	Introduction	109
4.2	Retour sur la fin du cycle de gestion stratégique du changement à l'iA : l'application de l'évaluation finale du cycle.....	110
4.2.1	Évaluation finale du plan d'action et de la vision du changement compte tenu des résultats du projet pilote.....	110
4.2.2	Recommandations pour le lancement du cycle opérationnel	112
4.3	Contexte de l'opérationnalisation du projet pilote : présentation de la mission réalisée avec une équipe marketing d'iA	113
4.3.1	Contexte de la mission	113
4.3.2	Déroulement de la mission.....	113
4.4	Opérationnalisation de la phase « Cibler » dans la mission	114
4.4.1	La démarche opérationnelle et les opérations menées.....	114
4.4.2	Résultats obtenus	115
4.4.3	Les enjeux opérationnels liés à cette phase	116
4.5	Opérationnalisation de la phase « Préparer » dans la mission	116
4.5.1	La démarche opérationnelle et les opérations menées.....	116
4.5.2	Résultats obtenus	118
4.5.3	Les enjeux opérationnels liés à cette phase	119
4.6	Opérationnalisation de la phase « Analyser et interpréter » dans la mission.....	120
4.6.1	La démarche opérationnelle et les opérations menées.....	120

4.6.2	Résultats obtenus	121
4.6.3	Les enjeux opérationnels liés à cette phase	123
4.7	Bilan de la mission réalisée	123
4.7.1	Retour sur les attentes initiales	123
4.7.2	Retour sur les KPIs	124
4.8	Conclusion du chapitre.....	124
5	Chapitre 5 : Conclusion du mémoire et perspectives.....	127
5.1	Bilan sur le travail de recherche réalisé	127
5.2	Perspectives sur la démarche et la problématique du mémoire	131
	Bibliographie	133

Liste des tableaux

Tableau 1.1:Modèle Data Flux.....	26
Tableau 1.2: Modèle Delta	26
Tableau 1.3: Modèle EY (Hollander, Hertz, et Wassink 2013)	27
Tableau 1.4 : Planning initial du projet de recherche	38
Tableau 2.1:Planning de l'application de la démarche proposée appliquée au cas de l'iA.....	44
Tableau 3.1: Le concept de "client" selon chaque secteur d'affaires d'iA	77
Tableau 3.2: Modèle analytique exploité pour l'expérimentation du projet croisement de l'information	89
Tableau 3.3 : Résultat du processus d'analyse 1	97
Tableau 3.4 : Résultats et interprétations des activités d'analyse et de fouilles de données	99

Liste des figures

Figure 1.1: Les 12 technologies les plus disruptives de notre société actuelle (Manyika et al. 2016)	3
Figure 1.2 : Panorama des entreprises de l'économie collaborative - version 1.0 (Owyang 2014)	7
Figure 1.3 : Portrait des techniques et approches d'exploitation de données traditionnelles et émergentes	12
Figure 1.4: Représentation schématique du théorème de CAP (Brewer 2001)	14
Figure 1.5: Architecture HDFS. Source : http://hortonworks.com/apache/hdfs	16
Figure 1.6: Représentation schématique de l'architecture MapReduce tirée de (Hurwitz et al. 2013)	17
Figure 1.7: Cartographie du réseau souterrain du réseau des infrastructures d'exploitation du gaz naturel de la région du Tennessee aux États-Unis	19
Figure 1.8 : Processus général de constitution de cubes OLAP (Stonebraker 1999)	20
Figure 1.9 : Architecture des MKIS	23
Figure 1.10 : Exemple d'approche dite "Big Data" (Fogelman-Soulié et Lu 2016)	24
Figure 1.11: Autre exemple d'approche dite « Big Data »	25
Figure 1.12 : Organisation interne des principaux secteurs d'activité de l'iA (maison mère)	35
Figure 1.13: Structure d'accompagnement de l'architecture d'entreprise d'iA	36
Figure 2.1 : Diagramme de la démarche globale proposée	41
Figure 2.2: Diagramme de raffinement du cycle 1 - Préparation à la mise en œuvre	49
Figure 2.3 : Diagramme de raffinement du cycle 1 - Mise en œuvre et évaluation	54
Figure 2.4: Diagramme de l'étape de mise en place de projets pilotes appuyant le plan d'action stratégique ..	60
Figure 2.5 : Diagramme de la mise en place de projets pilotes - cas de l'iA	62
Figure 2.6 : Diagramme de raffinement du cycle 1 - Évaluation finale du cycle	63
Figure 3.1 : Présentation synthétique de l'approche expérimentale proposée pour l'analyse de des données ..	69
Figure 3.2 : Schéma détaillé de l'approche expérimentale proposée	70
Figure 3.3: Périmètres de collecte des données d'iA choisies pour le projet expérimental	72
Figure 3.4 Modèle conceptuel simplifié de croisement de l'information client	76
Figure 3.5 : Synthèse des préoccupations soulevées par les gestionnaires de l'iA	80
Figure 3.6: Schématisation du plan d'action en lien avec le croisement de données clients	81
Figure 3.7: Modèle d'exploitation de données clients proposé par Thomas Davenport dans son ouvrage (T. H. Davenport, Harris, et Morison 2010)	82
Figure 3.8 : Interface SSIS	84
Figure 3.9: Interface Power Bi Desktop	84
Figure 3.10: Architecture d'exploitation de données choisie pour le projet expérimental	88
Figure 3.11: Interface de Rapidminer	90

Figure 3.12 : Démarche d'analyse de données	91
Figure 3.13: Représentation du découpage d'un objectif d'analyse	93
Figure 3.14: Représentation conceptuelle de l'ensemble de données choisi	95
Figure 3.15: Visuels graphiques et métriques profilant la clientèle d'IAAH	95
Figure 3.16 : Approche itérative de la démarche du projet pilote de croisement de l'information client	101
Figure 3.17 : Histogramme représentant la répartition des clients (tous secteurs confondus) par catégorie d'âge	104
Figure 3.18 : Représentation d'une interprétation possible de la figure 3.16	105
Figure 3.19: Représentation d'une autre interprétation possible de la figure 3.16	106
Figure 4.1: Démarche de gestion du changement dans l'exemple spécifique d'iA	109
Figure 4.3: L'étape de mise en place de projet pilote à l'iA	111
Figure 4.4: Architecture opérationnelle d'exploitation de données	117
Figure 4.5: Modèle en voûte de l'entrepôt de données opérationnel mis en place pour la mission	118
Figure 4.6 : Portrait des clients possédant un produit REEE et au moins un produit d'IAAH	121
Figure 4.7 : Répartition de la profession des clients en fonction de l'âge d'achat d'un REEE	122

Remerciements

J'adresse mes remerciements aux personnes qui ont contribué à la réalisation de ce mémoire.

En premier lieu, mes remerciements vont aux Fonds de recherche du Québec - Nature et technologies (FRQNT) et au Conseil de recherche en sciences naturelles et en génie (CRSNG) qui ont offert le programme de bourses de recherche d'innovation en milieu de pratique (BMP Innovation) dont j'ai bénéficié pour la réalisation de ce mémoire.

En second lieu, je tiens à remercier grandement mon directeur de recherche le Professeur Bernard Moulin, Professeur titulaire au département informatique et de génie logiciel de la faculté de Sciences et Génie de l'Université Laval. Bernard Moulin a été pour moi plus qu'un directeur de recherche. C'est un mentor tant professionnellement, qu'humainement. Il a été d'une générosité, d'un soutien dans les moments difficiles et d'une bienveillance qui m'ont permis de grandir humainement et de me dépasser dans mon travail.

J'aimerais également adresser mes remerciements à l'entreprise partenaire et sponsor de ce projet l'Industrielle Alliance. Je remercie tout particulièrement à Guy Daneau, Jean-François Cloutier et Marc Morin pour m'avoir donné la chance de me développer au sein de leur entreprise et de m'avoir apporté, tout au long du projet, leur confiance, support et écoute afin de mener ce projet dans les meilleures conditions qui soient.

Enfin, je tiens à remercier toutes les personnes de la DAC, de la VPSI et des différents secteurs d'affaires de l'Industrielle Alliance, qui m'ont consacré aussi généreusement du temps pour répondre à mes questions et m'apporter leur aide pour faire avancer ce projet de manière favorable.

Un merci spécial également aux membres du personnel universitaire de l'Université Laval de m'avoir accompagné et conseillé tout au long de ma formation de maîtrise.

1 Chapitre 1 : Mise en contexte du mémoire

1.1 Introduction

L'émergence des nouvelles technologies de l'information est un sujet au cœur du débat actuel sur l'avenir des organisations. En effet, 81% des dirigeants estiment que les avancées technologiques transformeront leur entreprise dans les cinq prochaines années (Rocard 2014). Deux facteurs clés en sont à l'origine : l'apparition du phénomène des données massives, encore appelé le mouvement « Big Data » et la transformation de notre société vers l'utilisation très large des technologies de l'information et de communication (TIC), souvent appelé « le digital ».

Lorsqu'on parle de données massives on considère des volumes de données non seulement énormes, mais aussi des données en constante augmentation. Cet « amas » considérable de données devient de plus en plus difficilement stockable par les méthodes traditionnelles, comme les bases de données relationnelles, car une compagnie a en moyenne l'équivalent de 427 fois la quantité de données jamais répertorié par la Bibliothèque du Congrès des États-Unis, alors que cette bibliothèque possède la plus grande collection de documents au monde (T. Davenport 2014).

On parle également de données de moins en moins structurées et d'une grande variété de formats (tweet, message texte, vidéo, audio, contenu web). Par exemple, Facebook stocke plus de données photographiques que l'ensemble des pixels que la société Kodak a pu traiter dans toute son existence (T. Davenport 2014).

Enfin, le Big Data est souvent associé à la génération de données à très grande vitesse. Par exemple, on capture aujourd'hui, quotidiennement, une quantité de vidéos plus importante que celle capturée pendant les 50 premières années de l'histoire de la télévision (T. Davenport 2014).

Plus récemment 2 nouvelles caractéristiques du Big Data se sont ajoutées : la véracité, désignant la qualité et l'intelligibilité des données, et la valeur, c'est-à-dire la capacité du Big Data à répondre aux besoins d'affaires.

Un autre facteur d'évolution important est la digitalisation : les objets digitaux facilitent la lecture de vastes quantités d'informations en des périodes de temps très courtes et au bénéfice de personnes localisées dans différentes parties du monde (Tapscott 1996). De plus, la technologie a rendu possible le phénomène de virtualisation : il est possible aujourd'hui de convertir des objets physiques et tangibles en objets virtuels (Tapscott 1996).

La notion de valeur des données, qui sera un des points importants considérés dans ce mémoire, gravite autour des notions d'information et de connaissance. Le concept d'information n'est pas nouveau, mais le processus selon lequel celle-ci est collectée, stockée, traitée, analysée et interprétée l'est. Tout au long de ce mémoire, ce sera principalement à ce nouveau processus que l'on s'intéressera, au-delà du mouvement Big Data qui s'y rattache.

A contrario, le concept de connaissance dans le contexte de la « digitalisation » est relativement nouveau et devient progressivement l'essence d'un changement sociétal auquel nous sommes confrontés aujourd'hui. Pourquoi ? Car la connaissance permet de gérer et d'exploiter l'information (Tapscott 1996).

Le phénomène Big Data et l'Économie dite « du Digital » sont traités dans de nombreux articles et très souvent présentés de façon très élogieuse et comme étant les promesses d'un monde meilleur compte tenu des récents changements technologiques de notre société. Ces changements sont portés par des compagnies technologiquement très avancées qui façonnent progressivement le monde industriel.

Cependant, la plupart des organisations, dites plus « traditionnelles », n'ont pas pu anticiper ce tournant et ne sont pas aujourd'hui en mesure de réagir efficacement à l'arrivée de ce qui est fortement pressenti comme étant l'aube d'une nouvelle révolution industrielle.

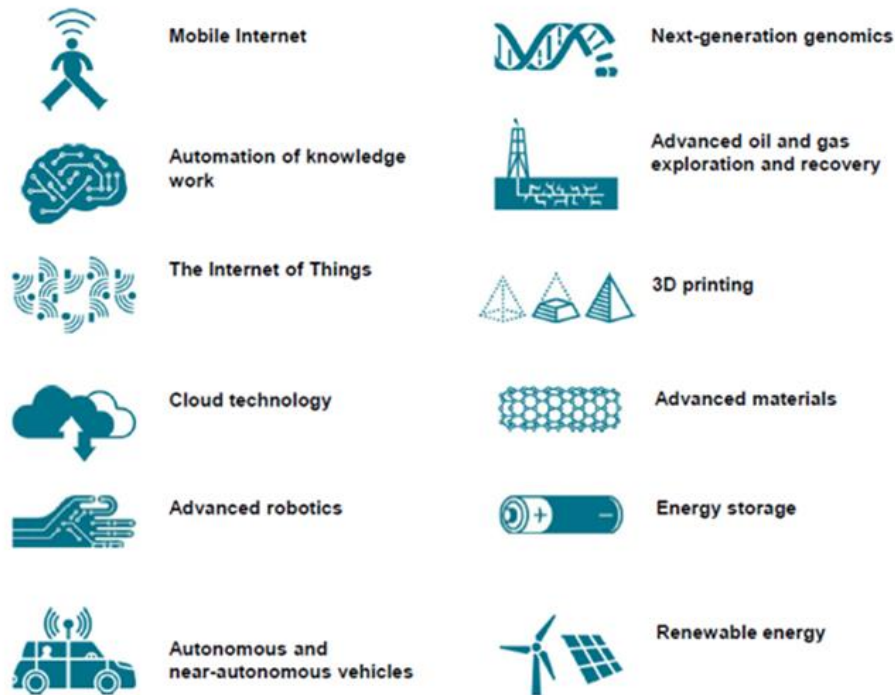
Dans ce chapitre d'introduction de notre travail de recherche, nous présenterons d'abord le contexte historique de cette « nouvelle révolution industrielle ». Puis nous présenterons le contexte théorique en lien avec l'exploitation des données massives qui sont un des défis très actuels de cette révolution et qui sont pertinents à notre recherche. Comme ce travail s'est fait en milieu pratique, nous présenterons le secteur des assurances et services financiers dans lequel œuvre la compagnie Industrielle Alliance qui a parrainé ce projet et offert le milieu d'accueil et les problèmes pratiques à considérer. Cette longue mise en contexte, nous permettra de présenter la problématique, les objectifs visés et la démarche qui ont orienté notre recherche.

1.2 Présentation du contexte historique – une nouvelle révolution industrielle

1.2.1 L'apparition de technologies « disruptives »

On appelle technologie disruptive, une nouvelle technologie ou une combinaison de technologies existantes qui peuvent être à l'origine de changements majeurs de paradigmes affectant des produits technologiques existants ou à l'origine de la création de produits entièrement nouveaux (Manyika et al. 2016). De nombreuses compagnies de consultation spécialisées en technologies de l'information s'intéressent à ces technologies disruptives. La firme de consultation McKinsey dresse un tableau des technologies majeures des dernières années dans son rapport de 2013 qui est illustré dans la Figure 1.1 (Manyika et al. 2016).

12 Disruptive Technologies



Source: McKinsey Global Institute

Figure 1.1: Les 12 technologies les plus disruptives de notre société actuelle (Manyika et al. 2016)

Comme l'illustre Figure 1.1 on retrouve 12 technologies disruptives majeures : l'Internet Mobile (IM), l'Automatisation de la Collecte de Connaissances (ACC), l'Internet des objets (Internet of Things ou IoT), l'Informatique Nuagique (Cloud), la robotique avancée, les véhicules autonomes ou semi-autonomes, la génomique nouvelle génération, l'exploration et la régénération des ressources naturelles, l'impression 3D, les matériaux de pointe, le stockage énergétique et les énergies renouvelables. Toutes ces technologies ne sont pas à détailler nécessairement pour comprendre dans quel contexte historique s'applique ce mémoire, mais il est néanmoins nécessaire de présenter certaines d'entre elles.

Un appareil internet mobile est un appareil de taille intermédiaire entre le smartphone et l'ordinateur portable, offrant un accès sans fil à Internet. Ces appareils sont conçus pour fournir du divertissement de l'information et des services basés sur la localisation de l'utilisateur pour une utilisation personnelle ou professionnelle. Ils permettent la communication à deux voies et le partage d'information en temps réel (Wikipédia 2016). Parmi les appareils mobiles les plus utilisés aujourd'hui on retrouve par exemple les smartphones ou encore les tablettes électroniques (« Tablet PC »).

L'Automatisation de la Collecte de Connaissances (ACC) découle des avancées en Intelligence Artificielle, en Apprentissage Automatique ainsi qu'en Interface en Langage Naturel et consiste en l'automatisation de certains processus manuels et ou intellectuels considérés jusqu'à présent comme pratiquement impossibles à réaliser par une machine (Manyika et al. 2016). Un exemple d'application d l'ACC est la reconnaissance vocale.

En ce concerne l'Internet des objets (IoT), un article paru en 2014 dans le *IEEE Transactions on Industrial Informatics Journal* propose une définition qui servira de référence dans le cadre de ce mémoire : « *In the application systems based on IoT, physical entities are connected to the information systems through sensors or tags, which become the representations of physical entities in information systems. The information systems cannot transfer physical entities directly; instead, the representations of the physical entities are transferred to realize the interactions of the physical entities.* » (Xu et al. 2014). On trouve aujourd'hui beaucoup d'applications IoT dans le secteur de la santé, notamment pour faciliter le suivi des patients dans un hôpital. En effet, plusieurs objets contenant des capteurs sensoriels sont utilisés pour transmettre des informations sur l'état des patients. Ils peuvent être contrôlés à distance par les médecins. On assiste progressivement à la prolifération et à la commercialisation de masse d'objets dits IoT, comme par exemple les montres connectées à distance à un appareil mobile.

Enfin, l'informatique nuagique permet de virtualiser l'information générée localement et de distribuer son stockage de manière à offrir d'avoir des performances qui s'adaptent dynamiquement, voire en temps réel au besoin. Il est important de préciser que l'informatique nuagique est une extension du « Grid Computing », de l'informatique distribuée et parallèle (Zhang et al. 2010).

Aux fins de ce mémoire, il n'est pas nécessaire d'avoir une compréhension complète de ces concepts, mais on gardera à l'esprit que ces technologies, toutes cumulées, génèrent beaucoup plus de données que par le passé : un rapport de 2015 de la firme de télécommunication Cisco fait état du fait que 66% du trafic d'adresse IP provient des appareils mobiles et des appareils sans fil, et que ce trafic atteindra les 18 Go par capitale nationale d'ici à 2019 (Trestian et al. 2017).

Tous ces constats créent un fort contraste, voire une rupture, avec le passé car, il y a encore quelques décennies de cela, les systèmes informatiques avaient des capacités limitées en termes de puissance, de stockage, de performance, de vitesse de transfert entre les systèmes; l'information n'était pas dématérialisée et était diffusée entièrement via les médias et les compagnies de télécommunication (Lake et Drake 2014).

Cette nouvelle situation crée également davantage des opportunités de développement de notre société. En effet, l'internet mobile offre la capacité et la possibilité aux utilisateurs « mobiles » d'accéder à de nombreux services en ligne, n'importe où et n'importe quand. L'Automatisation de la Collecte de Connaissances pourrait permettre de faciliter des tâches dans certaines professions qui sont encore actuellement très manuelles et peu stimulantes; et ceci grâce à l'internet des objets, à de nouveaux moyens de capture d'information qui

s'offrent à nous et sont à l'origine de l'amélioration des services publics comme la santé. On assiste aussi à la création de nouveaux services répondant à des besoins jusqu'à présent non suffisamment adressés (par exemple avoir un suivi médical hautement personnalisé, qui soit le plus préventif possible).

De plus, on note l'apparition de nouvelles offres de services dans le secteur privé comme par exemple Amazon qui offre un service d'informatique et de stockage nuagique aux entreprises : le paiement du service calculée au mois selon l'utilisation en termes de stockage serveur, bande-passante et mémoire CPU (Fang et al. 2017).

Tout ceci nous ramène à l'idée qu'une nouvelle révolution industrielle se profile et prend racines par le biais des développements récents de ces technologies et ceci dans divers secteurs d'activité, ce qui est à l'origine de nouveaux comportements de consommation, de produits et services. Nous reviendrons là-dessus, dans la section 1.2.2.

Cependant, cet essor de technologies disruptives soulève des questions importantes : Comment exploiter efficacement les informations générées par ces technologies ? Quel impact économique aura cette transformation sociétale pour les organisations ? De quels moyens celles-ci disposent-elles pour suivre cette transformation et y prendre part ?

De plus, on observe que les avancées technologiques liés à l'accroissement des capacités de stockage supérieures évoluent plus rapidement que celles liées à l'accroissement de plus grandes puissances de calcul. Cela signifie que l'augmentation des capacités de stockage des appareils ne s'apparie pas nécessairement avec une augmentation de la puissance de calcul et de traitement des données stockées. Les grands volumes de données deviennent donc de plus en plus difficiles à traiter et à analyser, ce qui renforce l'idée que, pour être en mesure de générer de la connaissance à partir de grands volumes de données, il est nécessaire de repenser la façon dont on exploite ces amas de données (Lake et Drake 2014).

1.2.2 Une transformation de la société en cours

Notre société est en cours de transformation et cette transformation s'explique souvent par l'essor de technologies disruptives dont nous avons parlé à la section 1.2.1. Cette transformation souvent est appelée dans la littérature « la transformation digitale » ou l'« ère du digital ». Ce changement touche plusieurs sphères de notre société.

Tout d'abord, c'est le domaine des sciences qui bénéficie directement de ces technologies pour faire de nouvelles avancées dans le domaine de la recherche ou optimiser certains aspects du processus de recherche (collecte de données, constitution de jeu de données test, outils de simulations).

Ensuite, comme cela a été introduit dans la section 1.2.1, de nombreuses initiatives visent à bonifier le système de santé où l'accès aux soins, le suivi des patients et le traitement des maladies, par l'usage de

technologies avancées de support à la décision, ainsi que par de nouveaux outils de collecte de données sur les patients.

On observe également, un développement de la connaissance grâce aux perspectives, pratiques et usages qu'apportent ces nouvelles technologies qui permettent de partager de l'information et de faciliter l'accès à l'apprentissage; par exemple l'explosion de l'usage des plateformes de e-learning.

On voit aussi de plus en plus de systèmes de « gestion intelligente » des villes et des municipalités où les nouvelles technologies sont exploitées pour optimiser les services rendus, tels que, la gestion du réseau de transport et des stationnements, la préservation du patrimoine culturel, la collecte des ordures, pour nommer quelques exemples.

L'objectif de ce mémoire n'est pas de rentrer dans le détail de cette transformation mais d'identifier les impacts organisationnels qui en résultent. Néanmoins, nous recommandons la lecture du livre « Digital Transformation and Global Society » qui offre un recueil d'articles très pertinents faisant état de cette transformation (Chugunov et al. 2016).

D'un point de vue économique, de nouveaux paradigmes de consommation font aussi leur apparition. L'un d'entre eux est appelé l'économie collaborative. En effet, l'économie collaborative est un mode de consommation de biens et services où les particuliers se coordonnent pour acquérir ou distribuer une ressource en échange de redevances ou d'autres formes de compensation, comme par exemple par le troc qui n'implique pas de compensation monétaire (Belk 2014). L'économie collaborative est basée sur le principe qu'au lieu d'acheter et de posséder des choses, les consommateurs choisissent d'accéder à des biens et préfèrent payer pour l'expérience d'y avoir accès temporairement (Belk 2014). Cette économie connaît une croissance importante car elle exploite justement des technologies disruptives comme par exemple l'internet mobile, la géolocalisation, les réseaux sociaux.

Les biens et services que l'on retrouve dans ce nouvel « écosystème » sont nombreux. La **Erreur ! Source du renvoi introuvable.** offre une vision globale du panorama des services que l'on retrouve aujourd'hui dans le cadre de cette nouvelle économie de consommation. Y figure en particulier des services comme le prêt entre particuliers, le financement participatif, le Wi-Fi partagé, l'agriculture communautaire, l'échange de services de proximité, et le covoiturage courte distance. Des compagnies comme Airbnb, Uber, Zipcar, Wikipedia, YouTube, Flickr, Facebook, Freecycle ou encore Twitter ont été les précurseurs de ce mouvement.



Figure 1.2 : Panorama des entreprises de l'économie collaborative - version 1.0 (Owyang 2014)

La force de cette économie est qu'elle répond directement aux besoins du consommateur car proviennent du consommateur lui-même tandis qu'auparavant les services habituels étaient assurés, en grande partie, par des organisations traditionnelles. Ces compagnies avaient la main sur le marché, la compétition était moins rude et les moyens à la disposition des clients pour évaluer la qualité des services étaient moins nombreux qu'aujourd'hui.

La transformation de notre société dans tous ces domaines offre de nouvelles possibilités pour :

- Améliorer l'efficacité des traitements médicaux, voire de sauver plus de vies
- Améliorer la surveillance des activités industrielles
- Développer de nouveaux produits pour faciliter la vie quotidienne (exemple avec le Google Car)
- Prendre de meilleures décisions

On constate également que les règles de l'offre et de la demande se redéfinissent et changent le statu quo d'avant la crise économique et financière de 2008 (Belk 2014). Ces nouvelles initiatives de partage à succès remettent de plus en plus en question le fonctionnement des industries établies et mettent en avant l'idée que la consommation collaborative permet d'acheter moins et de se diriger vers la propriété partagée plutôt que vers la propriété individuelle, bien plus coûteuse.

On peut ainsi se demander comment les différents acteurs de notre société (individus, gouvernements, organisations) s'adapteront à ces changements qui se profilent rapidement et plus particulièrement comment les organisations traditionnelles maintiendront leur place sur leurs marchés face à ce genre d'entreprises innovatrices qui intègrent de nouvelles technologies aussi rapidement.

Un autre fait saillant en lien avec ces nouvelles évolutions est la tendance qui semble s'établir chez les consommateurs (surtout les plus jeunes) de considérer comme normaux (et presque des normes de services attendus) les services personnalisés et à la pointe de la technologie. Notre société devient, par conséquent, de plus en plus dépendante de cette qualité de service et a des attentes de plus en plus élevées au fil du temps. Cette tendance observée devient un véritable enjeu pour les industries et les organisations qui ne sont pas encore en mesure de répondre à ces attentes (Lake et Drake 2014).

Finalement notons que ces changements technologiques et sociétaux soulèvent un nouvel enjeu de plus en plus critique : le respect de la vie privée et des informations personnelles des individus. Il est important de préciser que cet enjeu ne sera pas traité dans le cadre de ce mémoire mais il existe de nombreux ouvrages et articles qui traitent du sujet, notamment l'ouvrage de Daniel J Solove « The digital person » (Solove 2004).

1.2.3 Les organisations, au cœur de la transformation digitale actuelle.

Dans les deux sections précédentes nous avons fait référence à l'importance de la nouvelle révolution industrielle qui est en cours et il est intéressant d'identifier quels en sont les principaux acteurs. En effet, toute révolution est menée par des leaders qui forcent la marche et qui « guident la masse populaire » (Jensen 1993). Si on fait l'analogie avec la révolution industrielle du 19ème siècle, les innovations en matière de production industrielle introduites par les entreprises leaders du marché à cette époque ont conduit à la transformation de la société (par exemple Ford avec l'invention de la ligne d'assemblage mécanique en 1913, basée sur la découverte scientifique de la dynamo).

Cette analogie amène à penser qu'une découverte de portée universelle, appliquée d'abord par des leaders du marché avant qu'elle soit largement adoptée à l'échelle des industries, serait la clé pour libérer un fort potentiel de transformation (Carlsson 2004). En est-il de même aujourd'hui ?

Pour tenter de répondre à cette question on peut évoquer un certain nombre de constats actuels. En premier lieu on observe que de nouveaux modèles organisationnels font leur apparition. En effet, de nombreuses organisations nées à l'ère du digital sont très actives dans la recherche et le développement, en exploitant de façon innovatrice les technologies au point de devenir des leaders de la recherche faite dans leur domaine d'activité. Des compagnies comme Google, Amazon, Facebook, Yahoo!, Apple, Amazon ou encore Netflix sont des leaders de l'industrie en matière d'innovations, essentiellement grâce à une exploitation de plus en plus sophistiquée des données (Lake et Drake 2014).

Un autre type de changement organisationnel qu'il est pertinent de présenter ici concerne une culture d'entreprise axée sur la réponse aux besoins du clients, qu'ils soient explicites ou implicites. Le terme habituellement utilisé pour qualifier cette nouvelle culture organisationnelle est le « modèle d'orientation client » ou de « Customer Centricity ». Nous allons mettre l'accent sur ce type de modèle dans ce mémoire.

Trois caractéristiques majeures permettent de définir le concept d'orientation-client :

- Le fait d'utiliser une grande quantité disponible de données variées pour entretenir des relations individualisées avec la clientèle malgré des segments de clients disparates, des canaux de distribution et des médias multiples; et ceci en rassemblant des données pertinentes sur les clients et prospects : leur caractéristiques, préférences, attentes, demandes, motivations d'achat et leur comportement (Brashear, Gebauer, et Kowalkowski 2012).
- Le fait d'impliquer activement le client dans le processus marketing et d'innovation afin d'aboutir à la co-crédation de produits et services qui bénéficieront à la clientèle (Lamberti 2013).
- Le fait de déplacer les efforts organisationnels de la simple offre de produits et services vers une offre « d'expérience client personnalisée ».

Dans ce contexte, la gestion de l'information et l'utilisation de techniques avancées d'analyse et de visualisation de données en support à la prise de décision stratégique et opérationnelle deviennent de plus en plus une priorité pour les organisations.

Ces nouveaux modèles organisationnels sont en contraste avec les modèles qui sont en place dans la génération précédente d'organisation et qui étaient davantage orientés-produit. Ainsi leur structure organisationnelle était fondée, pour la plupart, sur la production de groupes de produits offerts aux clients (Brashear, Gebauer, et Kowalkowski 2012). L'organisation est alors souvent découpée en parties (« directions », « filiales ») qui visent chacune à produire une gamme de produits ou services de façon autonome et de manière compétitive. C'est ce qu'on appelle une « organisation en silos », où peu d'échanges et de collaboration prennent place entre les parties de l'organisation qui sont en compétition entre elles. Ce manque d'intégration conduit habituellement à un décalage entre les attentes du client et la vision de l'offre de produits et service de la compagnie (Brashear, Gebauer, et Kowalkowski 2012).

Notons aussi que la recherche et les progrès scientifiques en analyse avancée de données étaient portés, jusqu'à récemment, en grande partie par des institutions scientifiques et le milieu de la recherche universitaire (Lake et Drake 2014). Maintenant les avancées sont faites essentiellement par les leaders de l'industrie tels que Google, Amazon et Uber.

Dans ces organisations les données-client servaient principalement aux besoins opérationnels et celles-ci étaient stockées dans de grosses bases de données et gérées par des systèmes administratifs structurés en silos jusqu'à ce que des compagnies comme Google, Apple, Yahoo!, Amazon s'en servent pour développer de nouvelles perspectives d'offre de produits et services.

Grâce à ces nouvelles perspectives de nouvelles entreprises souvent appelées des « Digital Native Organizations », s'inspirant de ces modèles à succès, émergent et entrent petit à petit en concurrence avec des piliers de l'industrie de la génération précédente. L'ouvrage « Business Model Innovation: Concepts, Analysis, and Cases » de Allan Afuah présente le modèle d'affaires de plusieurs de ces nouvelles organisations (Afuah 2014).

Il est important d'ajouter que ces nouvelles entreprises, qui entrent en force sur le marché, mettent l'accent sur l'orientation client, sur le besoin de développer une intimité avec chaque client en tentant de créer un climat de dialogue collaboratif pour co-crée une expérience client qui a de la valeur pour le client et pour la compagnie (Lamberti 2013). Ces compagnies misent sur le fait que l'information est devenue un important vecteur de différenciation. Elles sont conscientes qu'intégrer rapidement les informations tirées des données collectées au sein de leurs processus d'affaires permettra de développer un avantage concurrentiel non négligeable et de se positionner au plus près de la demande du marché.

A contrario, les structures organisationnelles plus traditionnelles ne supportent pas activement la mise en place de processus relationnels avec le client en exploitant leur données internes. Or ces processus d'analyse facilitent l'arrimage de produits et services avec le client.

On observe néanmoins que de plus en plus de compagnies, traditionnelles, tendent à évoluer pour créer cette intimité avec le client. De ce fait, ce sont les entreprises qui n'ont pas encore fait cette transition qui courent le plus de risques de perdre des parts de marchés au profit des compagnies ayant créé une relation plus personnalisée avec leur clientèle (Brashear, Gebauer, et Kowalkowski 2012).

Le dernier enjeu important en lien avec ces évolutions récentes concerne la qualité des données. Celle-ci pose un gros enjeu actuellement, principalement pour les organisations traditionnelles qui souhaitent exploiter davantage leurs données internes pour rejoindre la demande du marché. En effet, dans ces organisations gérées en silos les systèmes d'information qui supportent la collecte des données détenues par ces organisations n'ont pas été conçus pour être exploités de manière intégrée comme le font les compagnies digitales leaders du marché. Habituellement, leurs bases de données ne sont pas intégrées, ayant souvent été créées à des époques différentes, avec des objectifs différents, et sans souci de la qualité des structures et formats de données pouvant faciliter l'interopérabilité. Aussi, un gros effort de transformation des données internes au sein des organisations traditionnelles est donc nécessaire pour augmenter la qualité des données et permettre la production d'analyses fiables sur des données intégrées.

Ce mémoire ne traitera pas en détails cet enjeu, bien qu'il soit mentionné à plusieurs reprises. Cependant, de nombreux ouvrages et articles traitent de ce sujet, notamment l'article : « *The moderating effect of normative commitment on the service quality-customer retention relationship* » de Fullerton (Fullerton 2014).

1.3 Présentation du contexte théorique – l'exploitation des données massives

Comme on l'a vu dans la partie précédente, notre société est en pleine transformation du fait du développement de nouvelles technologies disruptives. Ces technologies disruptives génèrent des données structurées et non structurées (vidéos, images, messages textes, fichiers audio, etc.) en grands volumes, à haute fréquence et rarement avec un haut niveau de qualité. Ce phénomène récent est associé dans la littérature au terme « données massives » ou « Big Data ». À ce titre, nous recommandons la lecture de l'ouvrage "*Big data: A revolution that will transform how we live, work, and think.*" de Mayer-Schönberger et Cukier qui présente en détails tous les aspects essentiels permettant de bien comprendre ce qui caractérise ce phénomène mais également les enjeux qu'il soulève (Mayer-Schönberger et Cukier 2013). Face à cette prolifération de données et à l'expansion du phénomène Big Data qui l'accompagne, la nécessité de comprendre comment exploiter différemment ces données est devenue une question primordiale pour les organisations. Celles-ci doivent non seulement accroître leur capacité à exploiter plus efficacement les données qu'elles possèdent, mais aussi à collecter de nouvelles données pertinentes afin de développer une offre de produits et services permettant d'enrichir leur relation d'affaires avec le client. Un grand nombre d'ouvrages, articles et travaux de recherche s'intéressent aux techniques et approches destinés aux organisations qui souhaitent tirer parti des données massives.

Tout d'abord, précisons que ce qu'englobe le terme « exploitation de données » car il servira de référence dans tout le reste du mémoire. Pour cela, rappelons ce qu'est le phénomène Big Data en nous appuyant sur une définition proposée dans un article du journal scientifique *IEEE Transactions on Knowledge and Data Engineering* (Wu et al. 2014).

Pour être en mesure d'explorer les relations complexes et évolutives entre les données et d'en exploiter le potentiel, il est important de tenir compte du stockage, du chargement et traitement, de la modélisation, de la visualisation (ou représentation) et enfin de l'analyse de données. Dans la suite de ce mémoire nous utiliserons l'expression « exploitation de données » pour faire référence à cet ensemble de concepts/techniques. De ce fait, il est important de décliner ces notions selon les techniques et approches proposées qui feront l'objet des sous-sections suivantes de ce chapitre.

De plus, les techniques d'exploitation de données utilisées aujourd'hui ne sont pas toutes apparues récemment. Les changements que l'on observe actuellement ne sont pas uniquement le fait de nouvelles technologies. En effet, certains domaines scientifiques ont été bonifiés au fil du temps et les avancées technologiques et techniques réalisées dans ce contexte de nouvelle révolution industrielle ont permis leur développement.

L'objectif de ce chapitre n'est pas de décrire de manière précise et exhaustive toutes ces techniques mais de présenter synthétiquement celles qui sont d'intérêt pour notre mémoire. Dans la Figure 1.3 nous présentons

un portrait des principales techniques d'exploitation de données suivant les quatre grands domaines que sont le stockage, le chargement – traitement – modélisation, la visualisation et l'analyse. De plus, la figure est séparée en deux parties, la partie supérieure listant les techniques traditionnelles, la partie inférieure listant les techniques dites « émergentes », le tout placé dans un contexte historique.

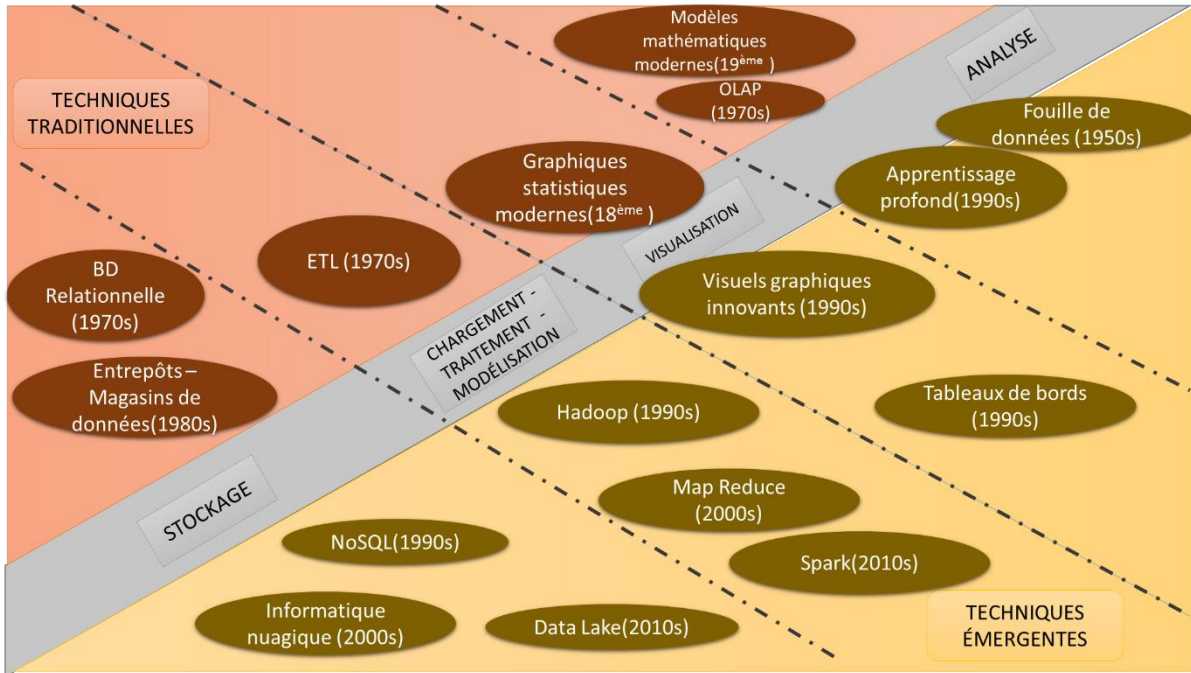


Figure 1.3 : Portrait des techniques et approches d'exploitation de données traditionnelles et émergentes

Certaines de ces techniques sont présentées dans les sous-sections suivantes.

1.3.1 Les techniques de stockage

Les techniques de structuration des données permettent de définir une architecture ou un squelette de la source de données pour permettre de manipuler le plus efficacement possible.

Dans les techniques traditionnelles on retrouve :

- Les bases de données relationnelles
- Les entrepôts et magasins de données

Les bases de données permettent de structurer et exploiter les données organisées suivant un modèle relationnel.

L'information y est stockée dans un ensemble de tables nommées distinctement, chaque table a une collection de colonnes spécifiques (appelées aussi attributs) et à l'intérieur d'une colonne on trouve des entrées contenant des données de même type (Dictionary of Computing - Oxford Reference 2016). L'objectif principal

de ces bases de données est de fournir des données structurées exploitables simultanément par divers systèmes informatisés. L'inconvénient de ces bases de données est que le haut degré de structuration des données rend le système peu flexible aux changements. En effet, des changements de modèle rendent la base de données non opérationnelle et nécessitent qu'elle soit réorganisée et rechargée (Dictionary of Computing - Oxford Reference 2016). Cet inconvénient apparaît quand on veut exploiter des données peu ou pas structurées, de formats hétérogènes, typiques des données massives.

Un entrepôt de données, quant à lui, est "un système de base de données conçu pour supporter des données archivées, les analyses et rapports subséquents aux processus de l'organisation. Il contient à la fois des données actuelles et historiques mais ne supporte pas le traitement de données transactionnelles » (Butterfield et Ngondi Ekembe 2016). Les magasins de données, contrairement aux entrepôts, ont pour vocation de stocker des données relatives à un ou plusieurs sujets d'analyses spécifiques. Ils répondent en général à un besoin précis d'analyse de données et sont alimentés à partir des entrepôts. Les entrepôts et les magasins de données sont structurés selon des modèles dits dimensionnels et sont conçus pour livrer des performances élevées d'agrégation et de traitement de requêtes d'exploration des données. On les exploite avec les techniques dites « OLAP » (« OnLine Analytical Processing »)

L'inconvénient des magasins de données est lié à leur multiplication au sein des départements d'une organisation, ce qui amène rapidement à des problèmes de synchronisation de l'information entre les différents magasins et entrepôts (Wiederschain 2011).

Pour comprendre davantage ces notions, nous recommandons la lecture de l'ouvrage "The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling" de Ralph Kimball l'un des initiateurs de l'exploitation d'entrepôts de données visant à répondre à des besoins organisationnels.

Face à ces limitations des BD relationnelles et des entrepôts de données de nouvelles techniques ont vu le jour permettant notamment d'offrir plus de flexibilité dans le stockage. Parmi ces techniques on retrouve :

- Le Not Only SQL (NoSQL)
- L'informatique nuagique (Cloud)
- Le Data Lake

Le Not Only SQL (NoSQL) correspond à un ensemble de technologies conçues pour améliorer les bases de données en termes de facilité de conception, de scalabilité et de disponibilité. Elles deviennent de plus en plus les technologies de choix pour traiter de grands ensembles de données et pour supporter des applications web (Butterfield et Ngondi Ekembe 2016).

La plupart des systèmes NoSQL n'adhèrent pas aux principes ACID appliqué aux BD relationnelles (Atomicité, Consistance, Isolation, Durabilité). Les bases NoSQL sont des systèmes qui peuvent travailler dans une configuration de traitement de données CP (Cohérente et Résistante au partitionnement), CA (Cohérente et

Disponible) ou AP (Disponible et Résistante au partitionnement) dépendamment du besoin d'exploitation que l'on a (voir Figure 1.4) (Brewer 2001).

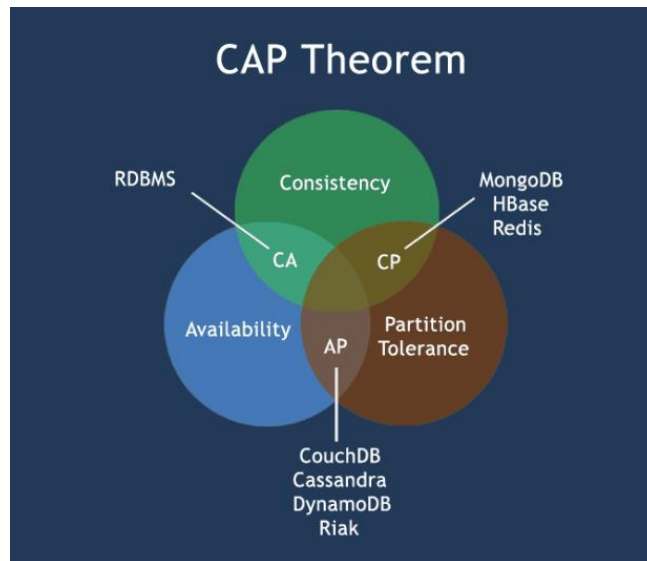


Figure 1.4: Représentation schématique du théorème de CAP (Brewer 2001)

MongoDB, HBase, Redis, CouchDB, Cassandra, DynamoDB et Riak sont des produits commerciaux connus sur le marché de Base de Données NoSQL. À noter que dans la Figure 1.4 le terme RDBMS désigne les bases de données relationnelles.

Les bases de données NoSQL ne revêtent pas une seule forme. Ceci s'explique assez bien par le fait que, comme d'après le théorème de CAP, les caractéristiques de performances de ces BD varient en fonction du besoin. On retrouve 4 principaux types de base de données NoSQL : modèle Clé/Valeur, modèle Orienté Colonne, Orienté Document et le modèle Orienté Graphe. Pour plus de détails concernant ces différents types de modèles, nous suggérons l'ouvrage « Les bases de données NoSQL : Comprendre et mettre en œuvre » de Rudi Bruchez (Bruchez 2013).

Pour ce qui est de l'informatique nuagique (ou « cloud computing »), cette technique a été présentée brièvement dans la section 1.2.1. Pour rappel, c'est une approche informatique dans laquelle l'utilisateur final se connecte à un réseau distant d'ordinateurs appelé « le cloud » afin d'exécuter des programmes, stocker des données, etc. Ceci permet aux utilisateurs d'accéder à de grandes capacités de stockage et de calcul provenant de plusieurs endroits sur la planète sans avoir à posséder ou maintenir ces ressources localement par eux-mêmes (Butterfield et Ngondi Ekembe 2016).

Le cloud offre ainsi de nouvelles possibilités en termes de capacité d'analyse de grands volumes de données et permet de réduire considérablement les temps de calcul. Cette technologie permet également de réduire nettement les frais d'infrastructure TI. En effet, aujourd'hui, des compagnies comme Amazon, Cisco,

Salesforce, Yahoo, Facebook, Microsoft and Google, Cisco proposent des offres de services cloud « à la demande » : l'entreprise cliente paie uniquement pour ce qu'elle consomme. L'article «A Comprehensive Survey on Cloud Computing » de Alam, Pandey, et Rautaray dresse un portrait actuel des applications possibles des technologies, reliées aucloud (Alam, Pandey, et Rautaray 2015).

1.3.2 Les techniques de chargement, de traitement et de modélisation

Ces techniques ont pour but de transformer les données issues de sources de données (bases de données, systèmes de fichiers) afin de les rendre plus facilement manipulables au moment de la visualisation et de l'analyse.

Pour ce qui est des techniques plus traditionnelles on présentera principalement les ETL (Extract, Transform, Load) qui sont « des processus en charge des opérations qui ont lieu en amont d'une architecture d'entrepôt de données. » (Butterfield et Ngondi Ekembe 2016). L'extraction consiste à collecter les données issues de différentes sources qui sont ensuite déposées dans un espace intermédiaire appelé le « transit » ou « staging area » où les données seront transformées. Ainsi, elles feront l'objet de plusieurs opérations de normalisation, nettoyage de manière à respecter l'intégrité et les contraintes d'utilisation de celles-ci. Enfin, les données vont être chargées dans l'entrepôt où elles seront exploitées (on dit aussi « consommées ») par des vues, magasins de données ou encore par des systèmes applicatifs (Liu et Özsu 2009).

Pour connaître le détail des opérations impliquées dans le processus d'ETL , il est possible de se référer à l'ouvrage de Wiederschain aux pages 206 à 209 (Wiederschain 2011).

Le processus d'ETL est un processus très consommateur de ressources, plus particulièrement pour les phases d'extraction et de transformation, si on veut constituer des sources de données exploitables à des fins d'analyses avancées. En effet, la gestion des processus d'ETL constitue 60 à 80% des projets menés dans le domaine de l'Intelligence d'Affaires (Misra, Saha, et Mazumdar 2013). Afin de réduire les efforts de performance et de déploiement de telles solutions de nouvelles technologies ont fait leur apparition progressivement :

- Hadoop
- Map Reduce
- Spark

Sur le site web d'Apache (licence sous laquelle a été développé le système Hadoop) on définit Hadoop de la manière suivante : « Hadoop Distributed File System (HDFS) est une approche versatile, résiliente et agglomérée pour gérer des fichiers dans un environnement volumineux. HDFS est un service de données qui permet de traiter des données fort volumineuses, très véloces.

Ceci est permis grâce au fait que les données sont écrites une fois et ensuite lues plusieurs fois, contrairement aux répétitions multiples d'écritures puis de lectures. »

La figure ci-dessous présente de manière synthétique les fondements de l'architecture HDFS.

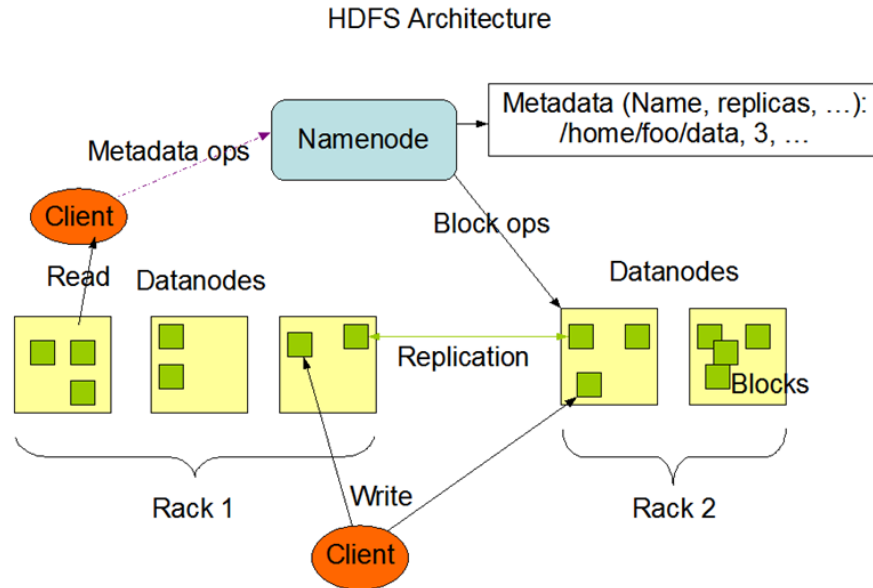


Figure 1.5: Architecture HDFS. Source : <http://hortonworks.com/apache/hdfs>

L'architecture est composée de « Name Node » et de « Data Nodes » qui s'exécutent sur un regroupement de disques durs virtuels appelés des « racks » distribués sur plusieurs serveurs (les « clients ») (Hurwitz et al. 2013).

Quant à MapReduce, il a été conçu pour être un modèle de programmation générique. Ce modèle combine deux capacités provenant de langages de programmation existants : la fonction « map » et la fonction « reduce ». La fonction map prend chaque élément d'une liste de clés -valeurs et produit de nouvelles listes de clés-valeurs grâce à une fonction d'association. La fonction reduce regroupe l'ensemble des opérations à réaliser dans une liste et retourne une valeur basée sur les opérations réalisées pour former la liste en sortie (Hurwitz et al. 2013). La figure ci-dessous schématise le fonctionnement du MapReduce et l'application de ces deux fonctions.

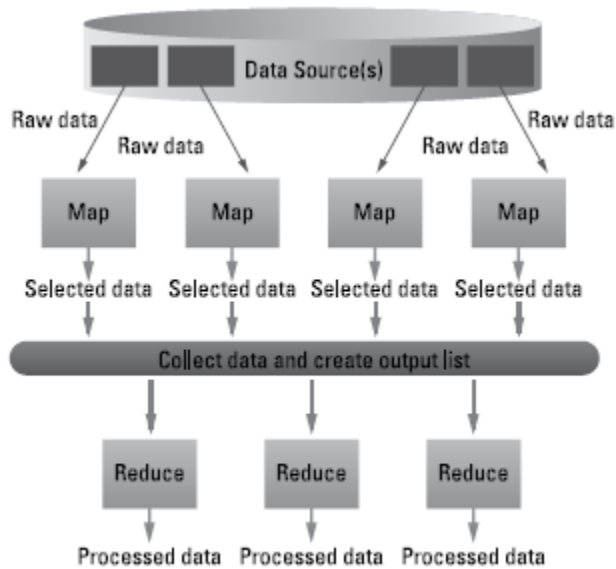


Figure 1.6: Représentation schématique de l'architecture MapReduce tirée de (Hurwitz et al. 2013)

L'ouvrage "Big Data for Dummies" est recommandé car il explique de manière simple et concise le fonctionnement plus détaillé de HDFS et Map Reduce (Hurwitz et al. 2013).

Bien que MapReduce et Hadoop, qui sont des modèles distribués spécialisés, permettent une meilleure gestion des traitements de données par lots, ils ne permettent pas de faire des traitements de données en temps réel. La nécessité de créer d'autres modèles distribués spécifiques à d'autres besoins a, de ce fait, rapidement émergé. Cette expansion de modèles spécialisés conduit à un nouveau problème, car mis ensembles ils peuvent devenir complexes à interconnecter et donc devenir inefficaces. Le moteur Spark offre une récente réponse à ce problème. En effet, Spark offre un moteur unifié pour le traitement distribué. Son modèle est similaire à celui de MapReduce mais possède une couche de partage de données appelée « Resilient Distributed Datasets » (RDDs). Grâce à cette surcouche Spark peut capturer une large portion de traitements qui nécessitait auparavant de faire appels à des moteurs séparés s'appuyant sur des technologies telles que SQL, le temps réel, l'apprentissage automatique et le traitement de graphes (Zaharia et al. 2016).

L'article "Apache Spark: A Unified Engine for Big Data Processing" dont est tiré cette définition approfondit les principes de base du moteur Spark (Zaharia et al. 2016).

1.3.3 Les techniques de visualisation

Pour ce qui en est des techniques de visualisation et d'analyse, elles remontent aux techniques de représentation statistiques proposées dès les années 1800. Mais elles ont fortement évolué, et de nouvelles approches de visualisation des données ont émergé récemment.

En effet, le défi qui se pose aujourd'hui est que, compte tenu des volumes de données que l'on est capable de stocker il devient de plus en plus difficile de visualiser ces grands volumes de données d'une façon qui soit facilement analysable et interprétable par l'être humain. On voudrait en effet être capable d'automatiser certains processus d'analyse lorsque le volume de données devient trop important notamment :

- Mettre en évidence des points représentant des valeurs aberrantes
- Faire apparaître l'existence de points dont les positions suivent un modèle dont les propriétés sont connues
- Comparer les résultats courants avec des résultats antérieurs
- Identifier la forme visuelle la plus adaptée aux types d'analyse que l'on veut réaliser
- Réaliser une série de tâches d'ajustement du rendu visuel des données (par exemple, le choix de l'échelle de représentation des données en fonction de l'amplitude des valeurs des données que l'on veut visualiser)

De nombreux ouvrages et articles s'intéressent au développement de nouvelles techniques de visualisation. On peut considérer notamment deux types de techniques pertinentes à utiliser dans ce contexte : les techniques représentation visuelle et les techniques de visualisation interactive. La représentation visuelle permet à l'utilisateur de distinguer les principaux modèles ou schémas représentatifs d'un vaste ensemble de données, de manière statique afin de déterminer les zones d'exploration à approfondir pour prendre des décisions. La visualisation interactive offre quant à elle, la possibilité de manipuler et d'explorer les données que l'utilisateur visualise pour lui permettre de renforcer sa compréhension des données observées (Zudilova-Seinstra, Adriaansen, et van Liere 2014).

Pour en savoir davantage sur ces nouvelles techniques visualisation de l'information, la lecture des ouvrages suivants est fortement recommandée :

- "Data Visualization: Principles and Practice, Second Edition" (Telea 2014) qui aide à bien faire la distinction entre les différents types de visualisation que l'on peut exploiter pour représenter des données et ce qui les caractérisent
- "Trends in interactive Visualization" (Zudilova-Seinstra, Adriaansen, et van Liere 2014) est une revue de littérature des nouveaux concepts, avancées scientifiques et applications lié à la visualisation interactive. C'est un ouvrage qui contient de nombreuses figures illustratives permettant de faciliter la compréhension du lecteur.

Un autre type de technique de visualisation consiste à exploiter des tableaux de bords. Le terme « tableau de bord » vient d'une analogie avec l'organisation des postes de pilotage des automobiles ou des avions dont l'objectif est de visualiser rapidement les principales mesures liées au bon fonctionnement du véhicule.

Un principe similaire est utilisé pour les tableaux de bords organisationnels : il s'agit de présenter les informations consolidées des valeurs actuelles et passées de la performance de l'entreprise afin d'aider des gestionnaires à prendre des décisions pour atteindre des objectifs fixés (Pappas et Whitman 2011). Les outils de représentation visuelle et de visualisation interactive de l'information varient selon divers éléments fondamentaux : les couleurs (variations de teintes), les formes et la position des éléments (Keim 2002). Des exemples, illustrant ce que sont des rapports ou tableaux de bord, sont présentés dans le chapitre 3.

Parmi les techniques émergentes que l'on retrouve aujourd'hui en matière de visualisation on observe de plus en plus l'utilisation de systèmes de représentation des données géographiques mettant en œuvre des systèmes d'information géographique (SIG). Grâce aux SIG, une carte géographique physique peut être traduite en un document de cartographie et être décomposée en des objets distincts. L'information sur les formes et localisations géographiques peuvent être ainsi combinées en une variété d'attributs incluant l'utilisation des terrains, la densité de population et le type d'infrastructures (routes, gaz, réseau de télécommunication, etc.) (Kwakkel et al. 2014). Grâce à cette décomposition géographique en couches superposables dans les SIG, il est possible de créer des combinaisons de couches d'information de manière dynamique et de s'adapter ainsi aux besoins d'analyse.

La Figure 1.7 présente deux représentations différentes de la cartographie du réseau souterrain du réseau des infrastructures d'exploitation du gaz naturel de la région du Tennessee aux États-Unis réalisée grâce à un SIG (Yana 2016).

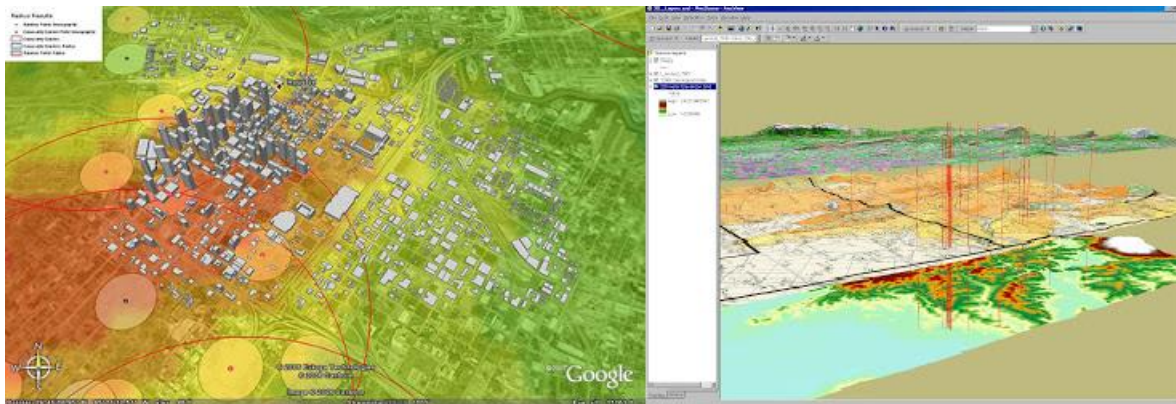


Figure 1.7: Cartographie du réseau souterrain du réseau des infrastructures d'exploitation du gaz naturel de la région du Tennessee aux États-Unis

1.3.4 Les techniques d'analyse de données

L'analyse de données est un processus qui fait appel à différentes techniques permettant de tirer de l'information interprétable d'un ensemble de données grâce à la mise en relation de ces données. Le domaine de l'analyse de données est en forte croissance depuis plusieurs années et évolue de plus en plus rapidement depuis l'émergence du mouvement Big Data. Certaines des techniques liées à l'analyse de données seront présentées ou en allant des plus traditionnelles vers les plus émergentes.

Pour ce qui est des techniques plus traditionnelles, il est pertinent de présenter le concept de cubes OLAP (OnLine Analytical Processing). Les cubes OLAP correspondent à une structure de données facilitant l'analyse de données agrégées grâce à des modèles multidimensionnels souvent représentés par un cube (Butterfield et Ngondi Ekembe 2016) : le cube de données est formé d'arêtes représentant les axes d'observations et sur chaque arête, une graduation est choisie afin d'observer les données à un niveau adéquat de granularité (Teste 2009). La Figure 1.8 illustre le processus de constitution de ces cubes.

On voit bien comment des bases de données extraites des systèmes de production sont rassemblées dans des entrepôts de données (voir aussi section 1.3.1). De ces entrepôts de données, on extrait des données qui sont organisées dans une base de données multidimensionnelles (le cube OLAP) pour être explorées avec des techniques de requêtes et visualisations afin de faciliter la prise de décision.

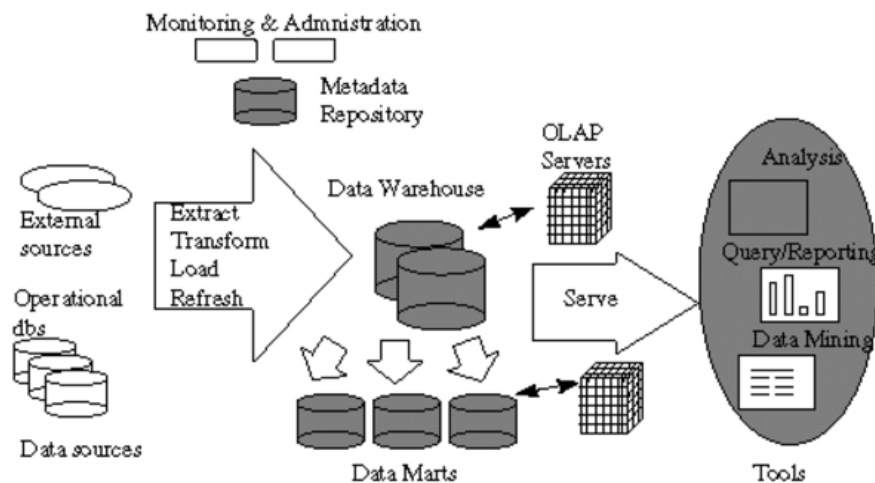


Figure 1.8 : Processus général de constitution de cubes OLAP (Stonebraker 1999)

Les analyses faites grâce à l'exploitation de cubes OLAP consistent à suivre des indicateurs qui sont représentés dans le cube par des cellules dans un espace défini par différentes arêtes du cube (les axes d'analyse). A partir de cette modélisation, il est possible d'effectuer différents types d'opérations correspondant à différents angles d'analyses.

Parmi ces opérations on compte principalement :

- Le Pivotement qui consiste à inter-changer 2 dimensions du cube
- Le Forage vers le bas qui permet de descendre dans la hiérarchie de la dimension (et donc chercher plus de détails)
- Le Forage vers le haut pour remonter dans la hiérarchie de la dimension (et donc d'abstraire des résultats par agrégation)
- Le Forage latéral servant à passer d'une dimension à une autre

L'OLAP permet aux analystes, gestionnaires et au personnel exécutif d'accéder rapidement, de manière consistante et interactive à une variété de vues possibles de l'information. Des géants de l'édition logicielle proposent aujourd'hui des solutions basées sur la technologie OLAP. On peut notamment citer Oracle et SAP. Cette technologie se décline de plus en plus sous différentes formes: multidimensionnelle, relationnelle, hybride, spatiale ou encore « Desktop » (« Online Analytical Processing » 2016).

Dans ce qu'on trouve de plus émergent, en termes d'analyse de données dans les entreprises, on voit apparaître de plus en plus l'utilisation de techniques dites de « Data Mining » (fouille de données). Le processus de Data Mining consiste en l'exploitation d'un ensemble de techniques statistiques et de programmation dans le but de découvrir de manière automatique des nouveaux modèles de données permettant de faire des analyses avancées à grande échelle. Une des approches qui fait l'objet de nombreuses recherches actuellement est l'apprentissage automatique dont le but est de construire des programmes qui s'améliorent au fur et à mesure de l'expérience tirée du traitement d'ensembles de données. Cet apprentissage se fait notamment à partir d'une base de données d'apprentissage qui est composée de cas déjà analysés, traités et validés (Begoli et Horey 2012).

L'encyclopédie libre Wikipédia offre une explication claire de l'origine du développement de cette technique :

La difficulté réside dans le fait que l'ensemble de tous les comportements possibles compte tenu de toutes les entrées possibles devient rapidement trop complexe à décrire (on parle d'explosion combinatoire) dans les langages de programmation disponibles. On confie donc à des programmes le soin d'ajuster un modèle permettant de simplifier cette complexité et de l'utiliser de manière opérationnelle. De plus, ce modèle est adaptatif, de façon à prendre en compte l'évolution de la base des informations pour lesquelles les comportements en réponse ont été validés, ce que l'on appelle apprendre ; ceci permet d'auto-améliorer le système d'analyse ou de réponse (commande adaptative...), ce qui est une des formes que peut prendre l'intelligence artificielle (« Apprentissage automatique » 2016).

Pour avoir une compréhension plus approfondie du processus du Data Mining et des différentes techniques et algorithmes associés, l'ouvrage « Data Mining for Business Analytics: Concepts, Techniques, and Applications in XLMiner » est à la fois complet, facile de compréhension et comporte un aspect très applicatif que nous recommandons au lecteur de ce mémoire (Shmueli, Patel, et Bruce 2016).

Depuis la fin des années 90 (Schmidhuber 2013), le domaine de l'apprentissage automatique a fait de grands progrès notamment avec le développement du Deep Learning (apprentissage profond) qui fonctionne avec un apprentissage à plusieurs niveaux de détails (ou de représentations des données) permettant d'améliorer fortement les résultats d'analyse, notamment en traitement d'images ou encore en analyse de signaux sonores.

1.4 Revue synthétique des nouvelles démarches d'exploitation de données destinées aux organisations

1.4.1 Approches méthodologiques

On a vu dans les sections précédentes que plusieurs défis attendent les organisations plus traditionnelles qui veulent suivre les nouvelles tendances d'exploitation de données pour participer au changement sociétal et industriel qu'on observe actuellement. Dans les dernières années, des approches méthodologiques ont été proposées pour guider les organisations dans leur démarche d'exploration de données. Les approches qui seront mentionnées dans cette section sont :

- Le Master Data Management (MDM)
- Le Knowledge Discovery (KD)
- Les Systèmes d'Information Marketing (MKIS)

« Le Master Data Management (MDM) a pour objectif de lier tous les fichiers de données en un seul fichier maître fournissant ainsi une version unique de l'information. Cela garantit l'intégrité des données à travers des architectures de systèmes variés et des applications déployées dans toute l'organisation. Cela permet également de partager des données entre plusieurs entités de l'organisation. » (Fogelman-Soulié et Lu 2016). Le concept MDM est une force pour l'organisation principalement pour unifier les données de systèmes déjà existants, chacun ayant une architecture définie au moment de son implémentation.

Le concept de Knowledge Discovery désigne un ensemble d'activités visant à extraire de l'information « interprétable » à partir d'ensembles de données complexes.

Cette approche se situe au croisement de plusieurs domaines scientifiques : l'intelligence artificielle, les statistiques et la visualisation de données (Begoli et Horey 2012). À noter que le concept de Knowledge Discovery est sous-jacent au processus de fouille de données (voir section 1.3.4)

Voici les principes clés du KD :

- La mise en place de pratique de collecte, stockage et de structuration de données complexes
- L'application de méthodes d'analyse qui permettent de comprendre des informations (implicites ou « cachées ») que recèlent les données collectées
- Identifier la signification des informations tirées de l'analyse des données

Un des enjeux liés au KD est l'automatisation et la formalisation de ces activités à un niveau opérationnel : on ne veut pas rester indéfiniment dans l'exploration expérimentale. Un autre enjeu lié à la gestion de la sécurité des flux de données qui transitent entre ces différentes activités.

Enfin, les systèmes d'information marketing (MKIS) permettent de fournir à une entreprise un flot d'information organisé qui aide à soutenir ses activités marketing : ils répondent aux besoins de collaboration, d'analyse et d'efficacité opérationnelle (Harmon 2003). Ce flot d'information facilite la prise de décision en tenant compte des équipements, des outils et bases de données, mais aussi des procédures, méthodes et ressources nécessaires au système pour adresser les besoins de l'organisation. Les besoins collaboratifs, analytiques et opérationnels sont ainsi desservis.

L'article de Harmon (Harmon 2003) développe de manière complète les atouts et limites de ces systèmes de gestion de l'information. La Figure 1.9, tirée de cet article, présente les bases de l'architecture de ces systèmes.

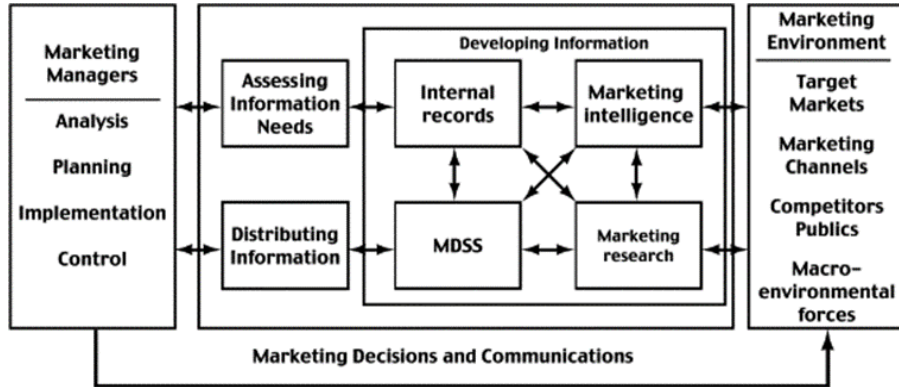


Figure 1.9 : Architecture des MKIS

La limite de cette approche est qu'elle touche un volet très spécifique de l'organisation (le marketing), sa généralisation à d'autres secteurs d'activité étant moins évidente.

Considérant toutes ces approches on peut se demander comment une organisation peut tirer parti de tous ces concepts et quelle démarche adoptée pour les intégrer ? On trouve dans la littérature plusieurs types d'approches qui tentent de répondre à cette question.

Parmi ces approches on peut citer :

- La gouvernance de données
- Les approches dites « Big Data »
- Les modèles de maturité

La gouvernance de données consiste en la spécification d'un canevas (ou ensemble de directives) définissant les droits de décision et les responsabilités des personnes/fonctions en lien avec l'information gérée au sein d'une organisation ou d'un département dans le but d'inciter à la mise en place de comportements adaptés

pour l'évaluation, la création, le stockage, l'utilisation, l'archivage et la suppression de l'information. La gouvernance de données permet ainsi de définir les rôles, mesures et processus de gestion de données efficaces et utiles à l'atteinte des objectifs de la compagnie (Morabito 2015).

Cette démarche est néanmoins destinée à des organisations dont la réflexion en lien avec l'information qu'elle possède est bien entamée et qui ont une idée claire de la façon dont la gestion de l'information va servir aux besoins de l'entreprise.

Les approches dites « Big Data », quant à elles, visent spécifiquement à adresser le volet d'exploitation de données. On trouve à la Figure 1.10 un exemple assez classique de démarche proposée :

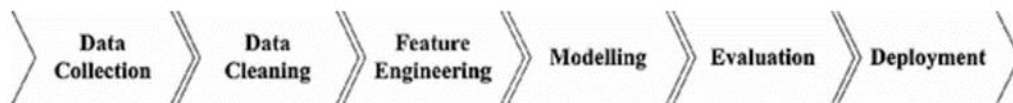


Figure 1.10 : Exemple d'approche dite "Big Data" (Fogelman-Soulié et Lu 2016)

L'étape de « Data Collection » sert à collecter des données disponibles pour répondre aux besoins d'exploitation. L'étape de « Data Cleaning » vise à évaluer la qualité des données et le besoin de rehaussement de celle-ci. L'étape de « Feature Engineering » transforme les données pour les rendre plus facilement manipulables par des techniques de modélisation. Par exemple dans le cas de l'utilisation d'algorithmes d'apprentissage automatique, il est important d'éliminer les valeurs nulles. La phase de « Modelling » permet d'appliquer le modèle de structuration de l'information. Suite à cela, on évalue les résultats de la modélisation des données (étape « Evaluation ») et déploie les techniques et modèles qui se sont avérés efficaces à une plus grande échelle (étape «Deployment »). L'inconvénient de ce type d'approche est qu'elle n'offre pas une vision intégrée de l'exploitation de données qui tienne compte à la fois des impacts organisationnels, des capacités et de la maturité de l'entreprise ainsi que des ressources disponibles. Il s'agit d'approches plutôt théoriques.

Un autre type d'approche proposée pour le Big Data est illustré à la Figure 1.11. L'emphase est davantage mise sur les étapes de réflexion préalables à l'intégration des techniques d'exploitation de données massives. On vise ainsi à déterminer si le Big Data répond réellement aux besoins de l'entreprise, à modéliser une structure d'entreprise adaptée à un tel changement, à évaluer les investissements et les risques que cela implique, à piloter des projets - test pour valider ou non la décision d'adopter cette démarche. Il s'agit d'une méthodologie plus riche que celle présentée par la Figure 1.10 et qui tient compte des impacts potentiels sur l'organisation. Cependant, elle fournit encore des recommandations d'ordre général. Dans le cadre de ce mémoire nous voulons proposer une démarche aussi complète et qui soit adaptée à la réalité du quotidien des entreprises.

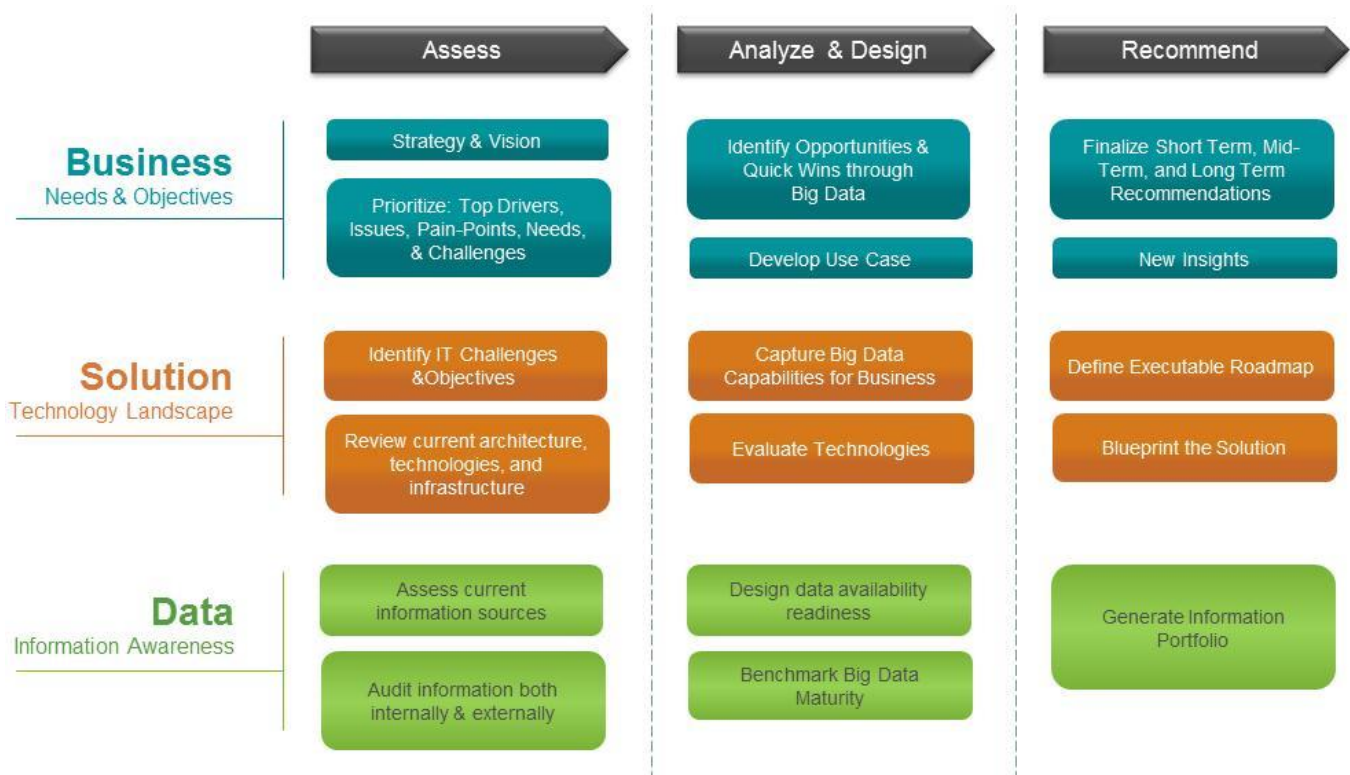


Figure 1.11: Autre exemple d'approche dite « Big Data »

1.4.2 Modèles de maturité

En ce qui concerne les modèles de maturité, il en existe une multitude proposée dans la littérature, il est donc important de bien choisir ceux que l'on juge les plus appropriés aux transformations que l'on vise à accomplir en matière d'exploitation de données. Les modèles choisis sont les suivants :

- Le modèle Data Flux (Morabito 2015)
- Le modèle Delta (T. H. Davenport, Harris, et Morison 2010)
- Le modèle orienté client EY (Hollander, Hertz, et Wassink 2013)

Le modèle Data Flux vise à présenter les principaux stades d'évolution d'une organisation pour augmenter la valeur et la qualité des données.

Le Tableau 1.1 reprend les éléments clés qui caractérisent chaque phase (« stage ») proposée dans ce modèle :

Tableau 1.1:Modèle Data Flux

Level of maturity	Characteristics
Undisciplined (Stage 1)	There is redundant and duplicate data, divergent sources and records at this stage. The main problem is that bad data and information will lead to bad decisions, and lost opportunities
Reactive (Stage 2)	This is the beginning of data governance. Improvements are experienced at the department level
Proactive (Stage 3)	The enterprise recognizes the value of a unified view of information and therefore begins thinking about Master Data Management (MDM). IT and business groups collaborate at this stage. At this phase the culture is also preparing to change
Governed (Stage 4)	Information is unified across the enterprise. Business needs drive IT projects. The enterprise has a sophisticated data strategy and framework. Employees have realized that information is a key enterprise asset

Ce modèle de maturité ne suffit pas à lui tout seul car la problématique d'exploitation des données est plus large que la seule notion de qualité de l'information.

Le modèle DELTA présenté dans le Tableau 1.2 mesure 5 compétences analytiques indispensables (« success factors ») : la nature des données de l'entreprise (« Data »), l'approche d'exploitation de l'information (« Enterprise »), la proactivité en matière d'analyse (« Leadership »), la cible analytique choisie (« Targets ») et les ressources humaines d'analyse à disposition (« Analysts »). Il attribue un niveau progressif selon ces compétences. Ces niveaux sont : Beginner, Localized, Aspiring, Capable and Leader

Tableau 1.2: Modèle Delta

Success Factor	Stage 1 Analytically Impaired	Moving to:			
		Stage 2 Localized Analytics	Stage 3 Analytical Aspirations	Stage 4 Analytical Companies	Stage 5 Analytical Competitors
Data	Inconsistent, poor quality, poorly organized	Data useable, but in functional or process silos	Organization beginning to create centralized data repository	Integrated, accurate, common data in central warehouse	Relentless search for new data and metrics
Enterprise	n/a	Islands of data, technology, and expertise	Early stages of an enterprise-wide approach	Key data, technology and analysts are central-ized or networked	All key analytical resources centrally managed
Leadership	No awareness or interest	Only at the function or process level	Leaders beginning to recognize importance of analytics	Leadership support for analytical competence	Strong leadership passion for analytical competition
Targets	n/a	Multiple disconnected targets that may not be strategically important	Analytical efforts coalescing behind a small set of targets	Analytical activity centered on a few key domains	Analytics support the firm's distinctive capability and strategy
Analysts	Few skills, and these attached to specific functions	Isolated pockets of analysts with no communication	Influx of analysts in key target areas	Highly capable analysts in central or networked organization	World-class professional analysts and attention to analytical amateurs

L'inconvénient de ce modèle est qu'il reste encore très général et ne rend pas suffisamment compte des problèmes ou difficultés que peut rencontrer pour une organisation dans l'application d'une telle démarche.

Enfin, le modèle d'orientation client EY est axé sur les 5 stades majeurs de compréhension d'une entreprise vis-à-vis de sa clientèle dans les domaines suivants :

- Créer des processus qui adressent les besoins de segments de clients bien définis
- Livrer une expérience client positive à tous les moments du cycle de vie du client
- Maintenir un dialogue riche et constructif avec la clientèle
- Entretenir une culture d'entreprise qui place le client au cœur de ses préoccupations

Un exemple des stades 1 et 5 de ce modèle est présenté dans le Tableau 1.3.

Tableau 1.3: Modèle EY (Hollander, Hertz, et Wassink 2013)

	Stage 1: Undeveloped		Stage 5: Leading
Process	Product push approach; little or no customer segmentation; products developed ad hoc; "One size fits all" service footprint	Stage 4: Advancing Stage 3: Developing Stage 2: Emerging	Tailored product offering based on discrete customer segments; real-time customer insight supports personalized interactions; predictive analytics employed to adjust offers and service actions accordingly; multi-channel approach based on behaviors and buying preferences
People	No ownership of the customer within the organization; compensation drives inconsistent behaviors; customer-facing employees not supported by back office functions; lack of authority to resolve customer issues at point of contact		Customer-centric culture pervasive at all levels of the organization; customer insight is a key component of the company's strategic agenda; employees take clear ownership for customer problems and manage issues through to resolution
Technology	Technology strategy not aligned with customer-centric culture; legacy systems limit ability to meet customer requirements; customer data trapped in organizational silos; lack of enterprise-wide data strategy hampers integration		Unified vision for data management across both enterprise and business partners; master data is managed as a corporate asset; proactive measures are in place or ongoing data quality improvements; the organization embraces customer-facing technologies as integral to channel effectiveness
Brand and market positioning	Low brand equity among target customers – known but not valued		High brand recognition and strong positive associations among target customers, based on real evidence and reputation for customer focus
Customer segmentation	Broad, undifferentiated offering, not specific to any distinct customer segments		Clearly defined target segments based on robust understanding of current and potential customer lifetime value
Distributor management	Distributor relationships driven by volume, little or no provider involvement in sale or post-sale customer activity		Highly selective approach to distributors driven by customer and distributor profitability; partnership approach with selected distributors with joined up customers engagement model
Propositions	Broad range of traditional life products: complex, inflexible, opaque, capital-intensive with provider-financed commission		Focused product range based on insight into needs of profitable customer segments; flexible design with transparent pricing, supporting fee-based advise and/or customer self-service
Service delivery	Transactional and admin-driven, reactive to customer and distributor requests; low-touch approach with no proactive contact or value added interaction		Relevant and value-adding customer contact through variety of media including phone and online self-service tools and information
Customer lifetime value management	No attempt to manage customer lifetime value beyond standard service model		Explicit understanding of customer lifetime value with value-adding interventions at key stages including lapse, life-stage changes, maturity, retirement etc., tailored to needs and value of specific customer segments
Culture	New business volume and market share are key metrics; enforce those customers not actively managed or measured		Customer lifetime value management is key metric: clear focus on meeting needs of profitable customer segments across their whole life cycle

Tout comme le modèle Delta il s'agit d'une approche théorique et pas suffisamment applicative. Toutes ces approches sont complémentaires mais incomplètes prises séparément. De plus, elles n'adressent pas toujours les besoins des organisations traditionnelles qui ont peu d'aisance dans l'exploitation de données. En effet, la

plupart des approches proposées actuellement sur le marché proposent d'exploiter les données internes des entreprises en offrant des modèles calqués sur des entreprises leaders qui sont des exemples de réussite. Qu'en est-il des organisations qui n'ont pas les prérequis nécessaires pour suivre ces exemples ? Que se passe-t-il quand on a affaire à une compagnie aux activités très diversifiée où la définition d'objectifs d'exploration communs va à l'encontre du modèle d'affaires actuel de l'entreprise ? Dans un tel contexte ces démarches recommandées ont peu de valeur pour ces entreprises ou offrent peu de possibilités d'application concrète.

Ces organisations sont pourtant celles qui sont les plus susceptibles d'être dépassées par toutes les transformations actuelles si elles ne sont pas capables de les suivre.

1.5 La problématique de recherche

Résumons ici les constats et enjeux actuels que nous avons identifiés pour l'exploitation avancée de données et quelles approches sont proposées pour les traiter :

- Une nouvelle révolution « dite du digital » est en marche : elle est le fruit de l'effervescence de nouvelles technologies qui transforment nos modes de vie et produisent des données en grande quantité
- Les changements liés à cette transformation touchent plus particulièrement les organisations qui historiquement sont les vecteurs des bouleversements qui affectent notre société
- À l'heure actuelle les compagnies leaders de cette nouvelle ère ont une maturité technologique qui leur confère un fort avantage concurrentiel
- Plusieurs approches et démarches ont été proposées pour permettre aux entreprises plus traditionnelles d'effectuer la transformation nécessaire afin de s'adapter aux nouveaux paradigmes de consommation résultant de la révolution du digital
- La plupart des démarches d'exploitation de données proposées aujourd'hui ne sont pas adaptées à des entreprises qui sont technologiquement moins avancées en matière d'exploitation de leurs données et qui en sont au stade de la réflexion sur la façon de s'adapter
- Les méthodes les plus abouties en matière d'exploitation avancée des données ne sont généralement pas directement applicables au quotidien des organisations les moins avancées technologiquement : elles ne tiennent pas compte de l'incontournable transition que doivent négocier ces organisations à divers niveaux (culture d'entreprise, gouvernance, vision d'affaires, relations à la clientèle, intégration des systèmes et données internes, nouveaux processus d'affaire intégrant une perspective d'intelligence d'affaires)
- De plus, ces approches sont essentiellement focalisées sur la collecte et l'exploitation de données externes (disponibles sur internet, les médias sociaux) et peu sur l'enrichissement de la valeur des données internes de l'organisation

Compte tenu de ces constats, nous visons dans ce mémoire à résoudre le problème suivant : Quelle stratégie et démarche de gestion du changement proposer aux organisations qui souhaitent exploiter le potentiel de leur données internes, mais qui n'ont ni la maturité technologique et organisationnelle ni les prérequis pour appliquer directement les techniques et approches émergentes qui sont offertes actuellement sur le marché ?

Dans ce mémoire, nous proposons une démarche de gestion stratégique et opérationnelle du changement qui tient compte de l'état de maturité technologique et organisationnel de l'entreprise en matière d'exploitation de données internes. Cette démarche a également pour but d'adresser les enjeux des entreprises moins matures par rapport aux leaders de la transformation « digitale » actuelle.

L'iA est un cas idéal d'illustration de cette démarche car :

- L'organisation possède un grand volume de données internes (lié à son activité d'affaires entièrement basée sur des éléments immatériels)
- Les données qu'elle collecte sont de natures variées (rapports, enregistrements audio, données transactionnelles, historiques de ventes, données de références)
- Les services qu'elle offre à sa clientèle (placements financiers, protection des risques des individus, couverture santé, investissements immobiliers, protection de groupe, etc.) sont diversifiés et gérés par des secteurs distincts et aux fonctionnements indépendants, ce qui offre une richesse de cas d'application pour mener des études en termes de maturité technologique, de processus d'affaires (PA), ou encore de gestion de l'information

Ainsi, les objectifs de ce mémoire sont de :

- 1) Apporter aux organisations dont la maturité technologique et organisationnelle est éloignée de celle des compagnies leaders de la transformation « digitale » actuelle de notre société, un certain nombre de connaissances sur ce qu'il est nécessaire de faire pour rendre effective cette transformation;
- 2) Préciser les changements à effectuer pour permettre à l'organisation d'exploiter ses données internes (dont plusieurs caractéristiques relèvent du phénomène « Big Data » - masses considérables de données, non intégrées, de formats différents, de qualités disparates);
- 3) Proposer une démarche systématique pour gérer ce changement en tenant compte de l'évolution du niveau de maturité de la compagnie;
- 4) Appliquer cette démarche à un cas concret pour une entreprise (développer l'orientation-client de son modèle d'affaires);
- 5) Montrer comment cette démarche peut s'appliquer concrètement au cas d'une entreprise.

1.6 Présentation du contexte pratique – le projet de recherche

1.6.1 Les origines de notre projet de recherche appliqué BMP en partenariat avec iA

Le projet de recherche à l'origine de ce mémoire est issu du programme de bourses de recherche d'innovation en milieu de pratique (BMP Innovation) et est offert conjointement par le Fonds de recherche du Québec - Nature et technologies (FRQNT) et le Conseil de recherche en sciences naturelles et en génie (CRSNG).

Ce programme est décrit de la manière suivante : « par ce programme de bourse en milieu de pratique, le FRQNT désire soutenir le développement de culture d'innovation par la création et le renforcement de liens de collaboration entre le milieu universitaire et les organisations des secteurs public (excepté les ministères), parapublic et privé; par la réalisation de recherches portant sur des thèmes prioritaires pour ces organisations et par l'insertion d'étudiants-chercheurs dans ces milieux de pratique » (Site internet du Fonds de recherche du Québec – Nature et technologies (FRQNT)).

Ainsi, les principaux objectifs de ce programme sont de :

- Favoriser l'accroissement des compétences en innovations des entreprises du Québec
- Créer des réseaux de collaboration universités-entreprises
- Aider les étudiants des 2^{ème} et 3^{ème} cycles à acquérir de nouvelles expériences et compétences personnelles et professionnelles

Pour le programme spécifiquement lié à ce mémoire les parties prenantes impliquées directement dans ce projet sont le département informatique de l'Université Laval, la compagnie d'assurance canadienne Industrielle Alliance (iA) et une étudiante du 2^{ème} cycle inscrite à la maîtrise en informatique avec mémoire. Les bases de ce partenariat entre l'Université Laval et l'Industrielle Alliance, ainsi que le sujet du programme de recherche ont été définis dans les grandes lignes préalablement au lancement du projet.

En effet, le mandat en milieu pratique a été formulé initialement de la façon suivante :

« Le rôle du candidat sera :

- D'explorer les concepts permettant d'exploiter les informations disponibles à l'Industrielle Alliance, Assurance et Services Financiers (IAASF). Que ce soit des informations internes ou provenant de sources externes
- D'identifier les informations disponibles à l'interne aussi bien que le potentiel d'utilisation de données externes de différentes natures. Le tout dans le but d'enrichir la connaissance actuelle, pour dégager des tendances, pour prévoir des changements, etc.
- D'identifier des opportunités concrètes d'exploitation de ces informations pour un ou plusieurs secteurs d'affaires d'IAASF
- La mise en place d'infrastructure et d'outils pour démontrer les opportunités est aussi envisagée

- Coté encadrement, un superviseur se chargera de mettre en contact le candidat avec les ressources internes d'IAASF selon les besoins. Le candidat aura accès à plusieurs personnes de la Vice-Présidence des Services Informatiques. Il aura accès à des personnes clés du côté affaire selon les besoins.

- Le candidat aura à sa disposition un espace de travail dans les bureaux d'IAASF »

Le groupe de recherche d'informatique cognitive de la faculté de sciences et génie de l'Université Laval a été retenu pour le partenariat avec l'Industrielle Alliance (iA). Ce choix s'explique par le fait que l'informatique cognitive s'intéresse aux systèmes complexes de traitement de l'information, capables d'acquérir, mettre en œuvre et transmettre des connaissances. Ce groupe de recherche est dirigé par le professeur Bernard Moulin : son rôle est de veiller à l'avancement et au bon déroulement des travaux de recherche. Il assure un suivi régulier des interactions avec le milieu pratique et partage ses recommandations et son expérience quant aux choix d'orientation du projet.

Les principales raisons pour lesquelles l'Industrielle Alliance (iA) a démontré un intérêt pour ce projet sont les suivantes :

- L'iA avait la volonté de travailler en partenariat avec une université pour avoir une nouvelle perspective sur certains aspects du fonctionnement de son organisation
- Un constat des hauts gestionnaires d'un retard technologique dans le domaine de l'exploitation de données par rapport à d'autres concurrents du secteur a été formulé
- La compagnie pressentait la nécessité de mieux comprendre l'impact potentiel du mouvement Big Data sur le secteur et la compagnie
- L'iA avait un grand intérêt pour la mise en place d'une démarche exploratoire, démarche fortement compatible avec l'environnement de recherche universitaire

1.6.2 Présentation du secteur des assurances - historique et enjeux

Le processus d'assurance permet de protéger un individu, une association ou une entreprise contre les conséquences financières et économiques liées à la survenance d'un risque ou événement aléatoire particulier. Le contrat (ou police) d'assurance a pour but le transfert de risques: l'assuré cède à la compagnie d'assurances un risque par définition aléatoire.

Lorsque le contrat prévoit une franchise, l'assuré conserve une partie du risque à sa charge. La compagnie d'assurance accepte le risque en échange de la prime (ou cotisation) de l'assuré (« Assurance », Wikipedia, 2016)

L'assurance se décompose en deux branches : la vie et la non-vie. La branche vie consiste en la protection de la mortalité et des rentes; la non-vie englobe la protection contre les accidents et certaines maladies, ainsi que la protection des biens et dommages.

De plus, le secteur des assurances est un secteur soumis à d'importants contrôles par des organismes de réglementation agréés. Au Canada, il s'agit de l'Autorité des Marchés Financiers (AMF).

En ce qui concerne le secteur de manière plus générale, il faut s'avoir que le Canada représente 7.5 % de la valeur du marché des assurances en Amériques (États-Unis, Canada, Mexique) (Marketline 2016). Les 3 compagnies leaders du marché sont Assurant basée à Kingston, Great-West Lifeco basée à Winnipeg et Intact basée à Toronto.

Les principaux enjeux auxquels ce secteur est confronté sont les suivants :

- La réglementation évolue constamment et force les compagnies à constamment revoir leur offre de produits et services pour qu'elles restent conformes à la réglementation
- Il s'agit d'un secteur qui contrairement au secteur de la vente au détail, par exemple, a une approche de vente plus centrée autour de la vente de produits que sur la satisfaction des besoins du client
- C'est une industrie dont les parts de marché sont visées par des start-ups de la nouvelle économie digitale. Un exemple de cette nouvelle tendance est la compagnie *The climate corporation* qui offre aux agriculteurs un nouveau type d'assurance basé sur l'analyse avancée des données météorologiques et qui leur propose des moyens d'optimiser leurs récoltes. La start-up vaut actuellement 930 millions de dollars (US) (Lake et Drake 2014).
- Le secteur de l'assurance subit actuellement une crise de confiance de la part de ses clients. La confiance des clients en leur compagnie d'assurance semble être en déclin (plus particulièrement depuis la crise mondiale financière de 2008), à une période où la fidélisation est un enjeu fort pour les organisations (Lake et Drake 2014). Elles doivent donc reconsidérer sérieusement leur rôle auprès de leur clientèle comme l'ont fait d'autres secteurs avant eux (comme, par exemple le secteur de la vente au détail).

1.6.3 Présentation de l'Industrielle Alliance (iA)

L'Industrielle Alliance, fondée en 1892, est la 4e plus importante société d'assurance de personnes au Canada et gère un actif d'une valeur de 106,9 milliards de dollars. Elle est inscrite à la Bourse de Toronto sous le symbole IAG et le cours de son action s'élève aujourd'hui à 42,05 \$CAD. Elle compte plus de 4900 employés au Canada et possède un portefeuille client de plus de 4 millions de Canadiens et de Canadiennes.

Enfin, son réseau de distribution est composé de plus de 25 000 représentants. La compagnie regroupe la maison mère (iA) et ses filiales. C'est principalement la maison mère qui sera présenté dans ce mémoire.

Son offre de services s'adresse à deux cibles : les particuliers et les entreprises. Pour les particuliers, la mission d'IA est d'offrir à ses clients et à leurs proches une protection financière en cas de décès, d'invalidité et de maladie, et de les aider à atteindre une autonomie financière en vue de la retraite ou de la réalisation de projets qui leur tiennent à cœur.

Ainsi, elle propose 8 types de services aux particuliers :

- L'assurance vie et maladie : la vocation des assurances vie est de garantir le versement d'une certaine somme d'argent (capital ou rente) lorsque survient un événement lié à l'assuré : son décès ou sa survie. Il existe une distinction entre l'assurance en cas de décès dite « assurance décès » qui verse le capital ou la rente en cas de décès et l'assurance en cas de vie (aussi appelée assurance sur la vie), qui verse un capital ou une rente en cas de vie à échéance du contrat (si décès avant l'échéance rien n'est dû à la succession) (Wikipedia)
- L'épargne et retraite : l'assurance propose de gérer les opérations relatives aux produits d'épargne émis et de générer des rendements sur ces placements en vue de la retraite ou pour réaliser des projets futurs (achat d'un appartement, achat d'une voiture, héritage, etc.)
- Le placement financier : le placement financier consiste à bloquer pendant une certaine durée un certain montant d'épargne dans une opération financière pouvant apporter un gain. (Wikipedia)
- Les fonds mutuels : les fonds mutuels sont des fonds de placement réunissant l'argent de plusieurs investisseurs. Les administrateurs du fonds doivent administrer l'argent selon un document officiel (charte) décrivant comment l'argent sera placé. Par exemple, un fonds d'obligations canadiennes devra placer l'argent uniquement dans ce genre d'obligations. La description du fonds spécifie généralement des pourcentages qui doivent être maintenus selon les types de titres. Par exemple, un fonds pourrait avoir 20 % d'obligations du gouvernement canadien, 40 % d'obligations de compagnies canadiennes et 40 % d'actions de compagnies canadiennes.
- Le prêt hypothécaire : une assurance prêt hypothécaire est nécessaire lorsqu'une personne contracte un prêt correspondant à plus de 80 % du prix d'achat ou de la valeur marchande d'une propriété (le moindre des deux). La mise de fonds minimale à être fournie par l'acheteur est alors de 5 % plutôt que 20 % pour un prêt non assuré (prêt conventionnel). Le montant de la prime d'assurance varie en fonction du montant de la mise de fonds.
- L'assurance automobile et habitation qui vise à protéger les intérêts financiers des clients en cas de dégradation du véhicule ou de l'habitation
- Le service d'investissement aux immigrants : il permet aux travailleurs et leurs familles d'immigrer au Canada en faisant un investissement passif de cinq ans auprès du Gouvernement. Les fonds investis sont garantis par le Gouvernement et la résidence permanente qui est accordée aux candidats est inconditionnelle. Il s'agit d'un service intermédiaire financier et facilitateur dans le cadre du programme québécois pour immigrants investisseurs et du programme canadien pour immigrants investisseurs. (Industrielle Alliance)

Pour les entreprises ou organismes collectifs, iA offre des régimes d'assurance et de retraite en fonction des besoins spécifiques aux collectifs. Elle propose ainsi 3 types de services :

- L'assurance collective : l'assurance collective de personne couvre, en vertu d'un contrat-cadre, les personnes adhérant à un groupe déterminé et, dans certains cas, leur famille et/ou les personnes à leur charge. Grâce à ce contrat l'assuré se fait promettre moyennant le paiement d'une prime, une prestation de l'assureur qui, prenant en charge un ensemble de risques, les compense en cas de réalisation d'un événement tel qu'un décès, accident, invalidité, maladie, etc. (Industrielle Alliance)
- L'épargne et retraite collective : ce service regroupe l'ensemble des mécanismes d'épargne constitué au sein d'une entreprise et présentant des avantages fiscaux et sociaux. Elle intègre donc : l'épargne salariale (plan d'épargne d'entreprise, épargne retraite, intéressement, participation, ...) et la retraite supplémentaire
- Une assurance accident pour les groupes

L'Industrielle Alliance (iA) est organisée selon des secteurs d'activités. Parmi ces secteurs d'activités on distingue : les secteurs des services « corporatifs » qui desservent l'ensemble de l'organisation, les « lignes d'affaires » qui desservent le client final avec l'offre de divers produits détaillés plus haut, ainsi que la Vice-Présidence des Service Informatique (VPSI). La figure ci - dessous illustre cette organisation des secteurs de l'iA.

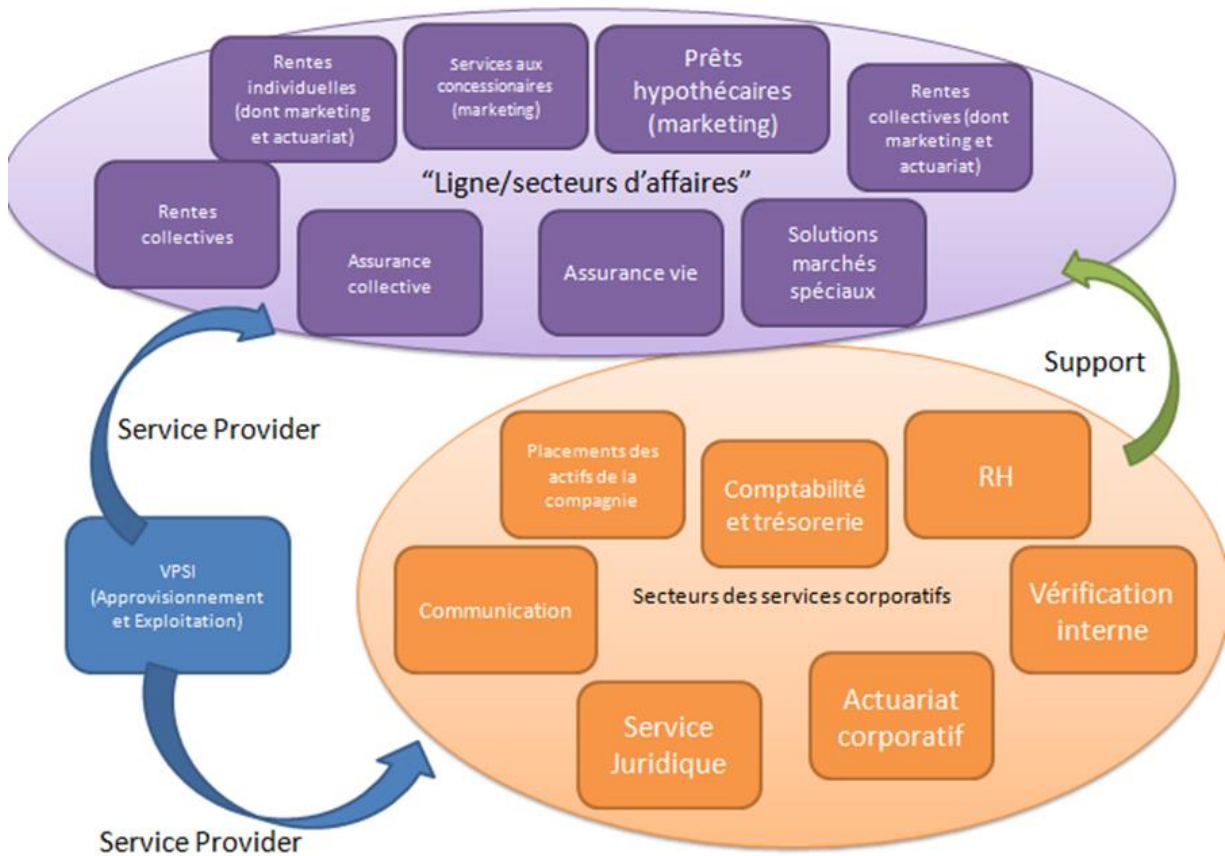


Figure 1.12 : Organisation interne des principaux secteurs d'activité de l'IA (maison mère)

La VPSI joue un rôle de fournisseur de services informatiques et de gestion de l'information aux autres secteurs d'activité. La structure transversale à toute la VPSI est appelée l'architecture d'entreprise. L'équipe d'architecture d'entreprise a un rôle de conseil et d'accompagnement dans la coordination des opérations « Informatiques » de l'IA. Comme l'illustre la Figure 1.13, l'architecture d'entreprise est à la fois transversale à toute l'activité de l'entreprise et à la fois plus spécialisée, par secteur d'activité.

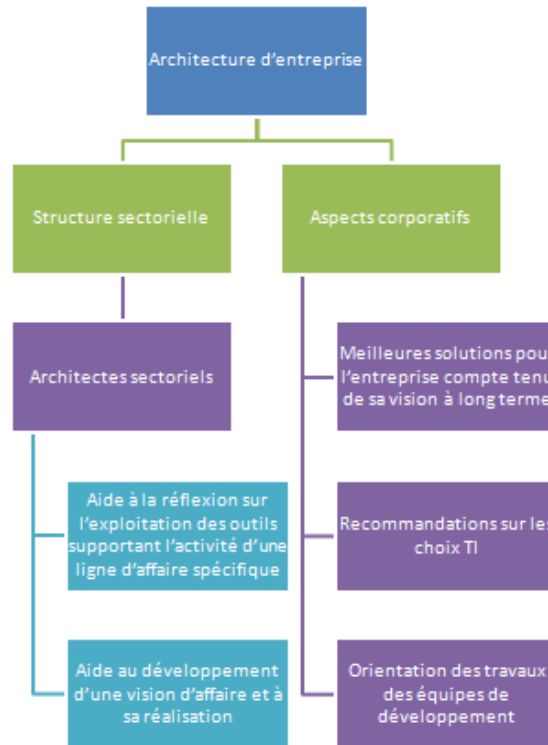


Figure 1.13: Structure d'accompagnement de l'architecture d'entreprise d'iA

L'architecture d'entreprise, répond aux demandes des lignes d'affaires sans effectuer de « contrôle » sur la cohérence de ces demandes vis-à-vis des objectifs stratégiques globaux de l'entreprise : elle joue plutôt un rôle de conseil auquel les lignes d'affaires peuvent faire appel.

Comme cela a été mentionné précédemment, le modèle organisationnel de la compagnie est basé sur une grande autonomie des différentes lignes d'affaires. On pourrait qualifier cette organisation de « modèle en silos » avec un fonctionnement des lignes d'affaires complètement dissocié à plusieurs niveaux : les processus d'affaires, les architectures systèmes, les modèles de bases de données et les valeurs dominantes de gestion de la documentation, etc. Ce fonctionnement très « libre » des secteurs d'activité couplé à l'absence de gouvernance d'entreprise est un réel enjeu pour la réussite de notre projet d'implantation de techniques avancées de gestion de données massives.

De plus, comme la plupart des compagnies du secteur, l'iA a un modèle d'affaires très orienté produit qui focalise ses efforts stratégiques sur les produits qu'elle propose sur le marché plutôt que sur les clients qui sont susceptibles d'acheter ses produits. Dans cette configuration stratégique la mesure du succès se fait via le portefeuille de produits, les ventes qui en découlent et les parts de marchés détenues grâce à ces produits.

Le danger de ce modèle d'affaire est le risque important pour l'iA d'avoir une vision organisationnelle « aveuglée » par cette approche produit au détriment d'une évolution qui permettrait de coordonner ses activités pour servir en priorité les besoins du client.

Un autre élément essentiel à présenter pour comprendre les enjeux auxquels font face l'entreprise est son réseau de distribution. En effet l'iA possède 25 000 représentants répartis dans plus de 50 agences à travers tout le Canada. Ces représentants, aussi appelés des agents, sont en charge de la vente et de la promotion des produits de la compagnie : ce sont eux qui ont une relation directe avec la clientèle. Ces représentants peuvent soit vendre exclusivement des produits de l'iA, on parle dans ce cas d'agents dits « carrière », soit vendre des produits de plusieurs compagnies dont l'iA et dans ce cas on parle d'agents « courtage ».

Enfin, la culture en silos de l'iA a créé progressivement, d'une part, un fort manque d'intégration des filiales au Système d'Information de la maison mère, et d'autre part, de grandes disparités de maturité technologique entre les différentes lignes d'affaire de l'IA. De ce fait, certaines lignes d'affaires disposent d'outils d'automatisation des tâches et d'analyse de données plus modernes, tandis que d'autres fonctionnent encore majoritairement par saisie manuelle.

1.6.4 Démarche appliquée

La particularité de ce mémoire de recherche est qu'il a bénéficié de la complémentarité des milieux académiques (le milieu de recherche universitaire) et pratique (le travail avec des intervenants de l'iA). En effet, ces deux approches menées en parallèle ont permis, à partir de la réalisation de plusieurs revues de littératures réalisées, de la participation à plusieurs colloques scientifiques et à l'observation approfondie du fonctionnement des différents secteurs d'affaires d'iA, d'identifier précisément les étapes manquantes dans les démarches proposées dans la littérature, et ce afin qu'elles puissent s'appliquer au cas de l'iA. Grâce à ce travail de préparation à la fois théorique et pratique, il nous a été plus facile de dégager les critères, techniques et recommandations pour élaborer une nouvelle démarche plus complète et qui prenne en considération à la fois l'hétérogénéité de la maturité de l'organisation et les bonnes pratiques suggérées dans la littérature.

De plus, nous avons fait une étude de faisabilité pour déterminer si l'organisation était bien en mesure de mener le projet à son terme : nous avons cherché à déterminer si l'organisation disposait des compétences, des ressources et des fonds nécessaires pour notre projet. Pour cela il a été important d'analyser, d'une part les risques liés au lancement du projet, mais également les risques de renoncer à le mettre en œuvre. Parallèlement à cela, l'accompagnement et l'encadrement de notre directeur de recherche à l'université ainsi que le choix de cours universitaires directement en lien avec le projet ont également contribué à faciliter ce travail.

Pour prendre connaissance du contexte de l'entreprise vis-à-vis de l'exploitation des données internes de l'entreprise, nous avons conduit une série d'entrevues. Elles ont permis de recueillir des faits et des opinions sur la situation actuelle de l'entreprise et de mieux identifier les besoins des utilisateurs vis-à-vis de notre projet comparativement aux besoins exprimés par les gestionnaires. L'objectif premier de ces rencontres fût

de clarifier les objectifs du projet. L'objectif sous-jacent était de rendre compte régulièrement de l'état d'avancement du projet et aussi de notre compréhension du contexte du projet. Cet objectif était primordial car il s'agissait aussi d'un moyen de s'assurer du soutien de nos responsables vis-à-vis de notre démarche, voire de leur intervention sur certains aspects de notre mission.

C'est à la suite de ces étapes, que nous avons soumis une première version de la démarche qui serait immédiatement testée à l'iA après approbation par les gestionnaires responsables du projet. Cette première version a ensuite été raffinée et testée au cours de plusieurs itérations conçues et testées à nouveau et ce jusqu'à l'obtention d'une version que nous avons jugée satisfaisante et suffisamment généralisée pour s'appliquer au cas d'autres entreprises. C'est cette version de la démarche qui a est présentée aux chapitres 2 et 3.

1.6.5 Planification initiale du projet

La planification du projet a été menée sur une période de deux ans comme présenté dans le Tableau 1.4 qui présente les sept grandes phases du travail réalisé (et leurs principales étapes). Les périodes temporelles sont présentées en colonnes avec le sigle des sessions couvrant chacune une période de 4 mois (ainsi A14 correspond à la session d'automne 2014 du mois de septembre au mois de décembre, H15 correspond à la session hiver 2015 de janvier à avril, E15 correspond à la session d'été 2015 de mai à août).

Tableau 1.4 : *Planning initial du projet de recherche*

	A14	H15	E15	A15	H16	E16	A16
Analyse préliminaire au projet (étude du milieu pratique)							
Faire une revue de littérature sur l'exploitation de données massives en entreprise et des démarches existantes							
Comprendre le fonctionnement du milieu pratique							
Définir les objectifs et enjeux du projet							
Conduire des entrevues avec les gestionnaires d'affaires parties prenantes au projet							
Présenter les résultats de l'analyse préliminaire							
Sensibiliser les acteurs du projet aux enjeux							
Réviser les objectifs du projet							
Planification du projet							
Présenter la revue de littérature réalisée							
Prioriser des objectifs du projet							
Mener des ateliers d'idéation avec les gestionnaires d'affaires							
Identifier un projet pilote à réaliser							
Analyse préliminaire projet expérimental							
Mettre à jour la revue de littérature sur l'exploitation de données massives en entreprise et des démarches existantes							
Mener des ateliers avec les experts TI pour identifier les enjeux du projet							
Établir un budget et identifier les ressources du milieu pratique à faire intervenir							
Réalisation du projet expérimental							
Collecter les données							
Comprendre les données collectées							
Raffiner la cible							
Préparer les données à analyser							
Analyser et interpréter les résultats							
Validation du projet expérimental							
Présenter les résultats du projet expérimental aux parties prenantes							
Présenter la démarche et les recommandations en vue d'une opérationnalisation du projet expérimental							
Opérationnalisation du projet expérimental							
Rédaction mémoire							

Le projet s'est décomposé en 7 phases principales (présentées en différentes couleurs dans le Tableau 1.4

- Analyse préliminaire au projet (étude du milieu pratique) ou étude de faisabilité du projet
- Planification du projet de recherche BMP
- Analyse préliminaire projet expérimental suite aux résultats de l'étude de faisabilité réalisée
- Réalisation du projet expérimental en milieu pratique avec une équipe dédiée sur place
- Validation du projet expérimental par les mandataires et parties prenantes au projet
- Opérationnalisation du projet expérimental en milieu pratique (à noter que cette phase s'étend au-delà de la période d'écriture de ce mémoire mais a été amorcée dans l'entreprise suite à la validation du projet expérimental mené dans le cadre de ce projet)

Les étapes qui ont composées chaque phase sont détaillées dans le tableau, leur découpage dans le temps exact n'a pas été représenté car jugé moins pertinent pour la compréhension globale du lecteur.

1.7 Conclusion

Nous avons présenté dans ce chapitre les contextes historiques, théoriques et pratiques nécessaires à la compréhension de la problématique adressée dans ce mémoire. En effet, ce mémoire a pour objectif, non pas d'apporter une nouveauté par rapport aux techniques et technologies en lien avec l'exploitation de données mais de proposer une stratégie efficace et applicables aux organisations. On souhaite également proposer une approche qui soit la plus complète possible en adressant les préoccupations concrètes des dirigeants d'organisations plus traditionnelles désirant faire un premier pas vers l'exploitation avancée de leurs données internes et commencer à développer la maturité nécessaire pour être progressivement en mesure de s'adapter à l'évolution actuelle du marché. L'application de cette démarche au cas de l'Industrielle Alliance, qui est un cas réel d'entreprise sur lequel est basé l'écriture de ce mémoire, a permis de mettre au point et raffiner la démarche globale proposée au chapitre 2 et de la mettre en pratique dans le cadre d'un projet expérimental en milieu pratique qui sera présenté au chapitre 3. Grâce aux résultats obtenus, des premières retombées positives ont pu être observées à l'iA, notamment la volonté des gestionnaires de la compagnie de démarrer l'opérationnalisation de la stratégie de changement que nous avons proposée. Le chapitre 4 présentera une des premières phases de cette opérationnalisation. Enfin, la conclusion (chapitre 5) nous permettra de tirer un bilan de l'expérience réalisée et de proposer des voies d'amélioration.

2 Chapitre 2 : Démarche proposée aux organisations

2.1 Introduction

On rappelle que l'objectif de ce mémoire est d'élaborer une démarche qui se veut pratique et qui s'adresse à des gestionnaires, chef de projets ou dirigeants qui souhaitent suivre et intégrer les transformations actuelles du marché et de la société. Le domaine de transformation choisi dans le cadre de ce mémoire est l'exploitation des données internes pour développer une approche d'affaires orientée vers le client en tenant compte de l'environnement externe fortement influencé par le mouvement *Big Data*. La démarche proposée a aussi pour objectif de s'adresser à des organisations qui sont moins matures par rapport aux compagnies leaders de leur industrie en regard de l'analyse des données.

Dans le chapitre 1, les principales techniques, démarches existantes dans la littérature des dernières années ont été présentées. Le constat a été fait qu'il existe actuellement peu de démarches proposées qui soient complètes, appliquées à des cas d'entreprises réels et à la portée des entreprises qui débutent dans le domaine de l'exploitation de données. Dans ce chapitre, nous proposons une nouvelle approche pour le développement d'un cycle de gestion du changement stratégique et opérationnel liée à la gestion des données internes en tenant compte de la complexité qui peut en découler. La gestion stratégique du changement, qui est notre contribution principale dans ce mémoire, sera raffinée selon des phases et étapes clés et sera illustrée par l'exemple de ce qui a été réalisé à l'iA.

2.2 Présentation globale de la démarche proposée

2.2.1 La démarche de gestion du changement en quelques mots

La démarche de gestion de changement que nous proposons est composée de deux cycles que nous proposons à toute organisation qui souhaite adopter un changement organisationnel important dans un contexte d'exploitation des données avancées. Le premier cycle est plus stratégique et précède un second cycle plus opérationnel. Pour chacun de ces cycles nous avons établi un nombre de phases P1.i et P2.i qui les découpent de manière logique et séquentielle. Chacune des phases représente un morceau clé de la démarche que nous avons mise au point.

De plus, cette démarche a été construite et raffinée en se basant sur les actions que nous avons menées à l'IA et sur l'expérience que nous en avons tirée. Elle a ensuite été généralisée pour qu'elle puisse servir de guide à d'autres entreprises. Le nom de chaque phase est inspiré des étapes clés que l'on retrouve communément dans la littérature lorsqu'on parle de gestion de projet (Rota 2009). La Figure 2.1 illustre, de manière synthétique, notre démarche sous forme de diagramme.

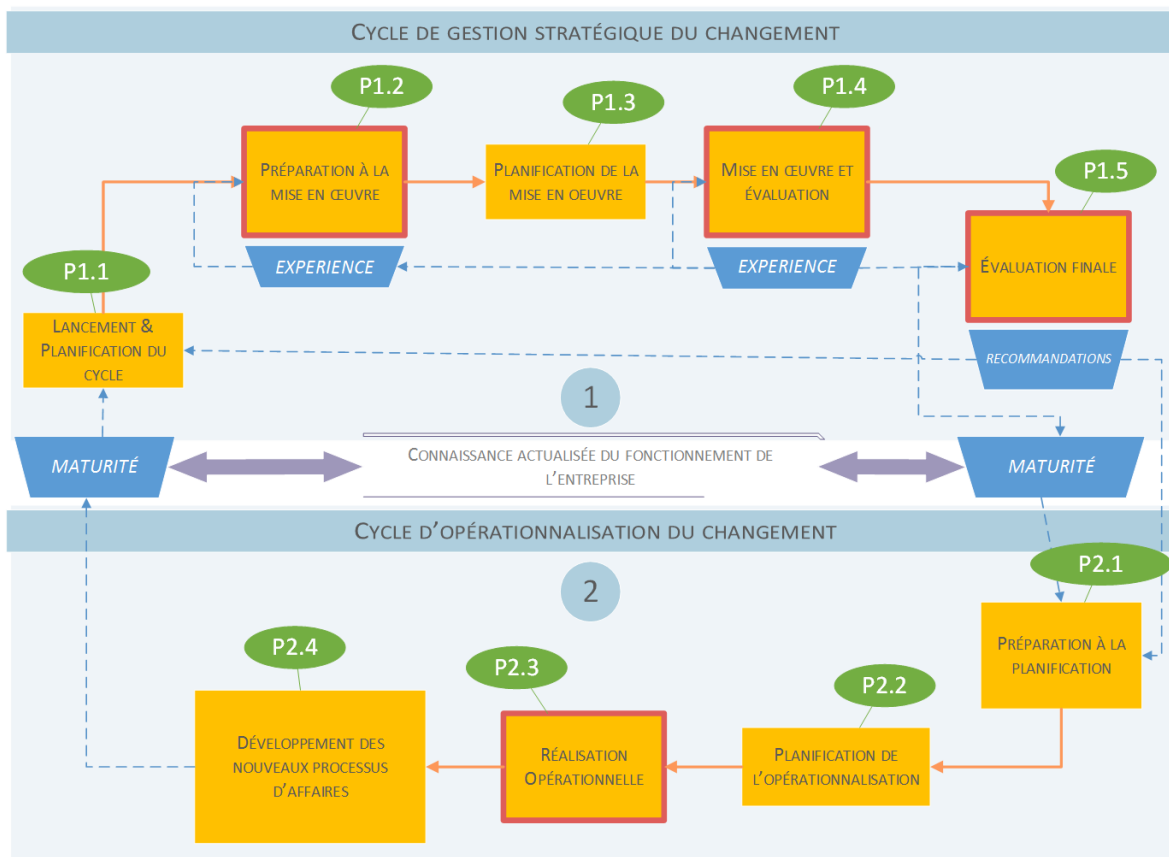


Figure 2.1 : Diagramme de la démarche globale proposée

Les rectangles qui composent le diagramme mettent en évidence les différentes phases du cycle. Les phases du cycle encadrés en orange feront l'objet d'un raffinement dans le cadre de ce chapitre ou d'un chapitre ultérieur :

- Les phases *préparation à la mise en œuvre* et *évaluation finale* seront raffinées dans ce chapitre
- La phase *mise en œuvre et l'évaluation* fera l'objet d'un raffinement et sera illustrée au chapitre 3
- Un exemple de la phase *réalisation opérationnelle* sera développé dans le chapitre 4

Les trapèzes représentent des accumulations de connaissances qui sont réutilisés par d'autres phases : les flèches bleues en pointillés matérialisent le fait que ces accumulations alimentent ces phases de manière progressive ou itérative. Les doubles flèches violettes illustrent l'idée que la maturité développée au travers de chacun des cycles accroît la connaissance de l'entreprise au cours de son fonctionnement et apporte de nouvelles perspectives d'évolution.

Un scénario possible d'illustration de l'application de la démarche pourrait être le suivant :

- Une compagnie décide d'exploiter davantage les données qu'elle possède
- Cette entreprise est consciente qu'au vu de son niveau de maturité estimé, il s'agit d'un changement important dans le fonctionnement de son organisation
- Elle commence donc par établir une stratégie de changement (cycle 1 sur la Figure 2.1)
- La mise en œuvre de la stratégie de changement amène l'organisation à développer la maturité stratégique nécessaire pour opérationnaliser le changement attendu mais aussi pour développer une meilleure connaissance de son fonctionnement
- Une fois ce premier cycle achevé, les recommandations issues de l'évaluation finale du cycle 1 peuvent d'une part alimenter de nouvelles réflexions stratégiques, qui déclenchent à nouveau le cycle, et d'autre part servir d'intrants pour le cycle d'opérationnalisation du changement (cycle 2)
- De la même façon que la mise en œuvre de la stratégie de changement fait gagner en maturité la compagnie, la réalisation opérationnelle du changement et son déploiement dans les processus d'affaires de la compagnie permettent de franchir une nouvelle étape dans l'évolution de la maturité et permet d'accroître davantage la compréhension du fonctionnement de l'entreprise.

Il est important de préciser que chaque cycle s'exécute de manière itérative et/ou incrémentale. Les avantages de l'approche itérative et incrémentale est qu'elle permet de s'adapter aux évolutions avec flexibilité en permettant une convergence progressive par l'essai/erreur vers la cible identifiée. Elle laisse également la possibilité de rediriger la trajectoire d'atteinte de la cible, si nécessaire. De plus, elle aide à réduire le risque d'échec en facilitant l'identification rapide des enjeux potentiels tout en apportant la flexibilité suffisante pour les adresser progressivement. Enfin ce type d'approche favorise le dialogue entre les différentes parties prenantes au projet à travers les différentes phases du cycle (Brasseur 2013).

Reprenons la lecture de ce diagramme et illustrons brièvement la démarche dans le cas de l'iA. Au démarrage du projet en septembre 2014, iA s'interrogeait sur l'exploitation actuelle de son grand volume de données internes et sur les possibilités de les exploiter davantage pour servir les objectifs stratégiques de la compagnie. Le projet de réflexion stratégique sur le changement à adopter (correspondant au cycle 1) s'est déroulé sur une période de deux ans et a débuté par une analyse du niveau de maturité actuel d'iA ainsi que du niveau de maturité visé. De cette réflexion a émergé une volonté de l'entreprise d'orienter davantage ses activités et processus d'affaires autour des besoins du client en exploitant les données relatives à leur clientèle et disponible à l'interne. Des recommandations, tirées de la mise en œuvre de cette stratégie de gestion du changement, ont permis d'enclencher la phase d'opérationnalisation (cycle 2). Dans le cadre de cette opérationnalisation, des initiatives ont été mises en place pour déployer de nouveaux processus d'affaires visant à outiller davantage certains secteurs d'affaires en matière d'analyse de données et l'exploitation des nouvelles technologies, ainsi que le réseau de distribution pour tenter de mieux répondre aux besoins du marché.

2.2.2 Détails sur le déroulement du cycle de gestion stratégique du changement

Pour commencer, la phase de lancement du cycle (P1.1 de la Figure 2.1) consiste à identifier les enjeux spécifiques liés aux données internes et sur lesquels la stratégie devrait se concentrer. Elle vise également à identifier et réunir les parties prenantes du projet de changement. Cette phase ne sera pas présentée en détails dans ce mémoire. Une fois la décision de lancer le projet de développement d'une stratégie de changement, la préparation de la mise en œuvre de la stratégie (P1. 2) est l'étape qui suit et aide à dresser le portrait actuel de l'organisation en tenant compte des enjeux soulevés lors de la phase de lancement du cycle (P1.1). Dans le cadre de ce mémoire, on s'intéresse à la problématique de l'exploitation des données internes de l'entreprise. Cette phase de préparation fera l'objet d'un raffinement dans la section 2.3. L'expérience tirée de la réalisation de cette phase, notamment la connaissance de la maturité initiale de l'organisation en matière de gestion des données internes, alimente la phase de planification de la mise en œuvre (P1.3). En effet, une fois que l'on a évalué les niveaux de maturité actuel et ciblé, on s'attelle à identifier les moyens nécessaires à déployer et l'horizon sur lesquels ces moyens vont être développés pour établir la stratégie de changement qu'on vise à mettre en œuvre. Cette phase ne sera pas non plus raffinée en détails dans ce mémoire. La planification de la mise en œuvre, une fois achevée, aboutit à l'identification des actions à exécuter lors de la mise en œuvre : établir la vision stratégique du changement, élaborer le plan d'action permettant de bâtir cette vision et tester ce plan d'action par la mise en place de projets pilotes. L'expérience tirée de la mise en œuvre (P1.4) permet d'évaluer la réussite de la phase et oriente éventuellement les actions à réaliser pour faire des ajustements. Cette phase fera l'objet d'un raffinement dans la section 2.4. L'évaluation finale (P1.5) du cycle permet de faire le bilan du déroulement du cycle et d'identifier pour chaque phase les points de réussite et les éléments d'amélioration en vue d'un nouvel incrément du cycle, mais également pour passer au cycle d'opérationnalisation. Cette phase sera expliquée en détails dans la section 2.5.

Il est pertinent à présent d'illustrer l'enchaînement de ces phases avec l'exemple de l'iA. Au chapitre 1, le planning de notre projet de recherche a été présenté. Il est important pour assurer la bonne compréhension du lecteur de le replacer dans le contexte de la démarche proposée ici.

Tableau 2.1: Planning de l'application de la démarche proposée appliquée au cas de l'iA

	A14	H15	E15	A15	H16	E16	A16
Analyse préliminaire au projet (étude du milieu pratique)							
Cycle 1 - Lancement et planification du cycle							
Planification du projet							
Cycle 1 - Préparation à la mise en œuvre							
Cycle 1 - Planification de la mise en œuvre							
Analyse préliminaire projet expérimental							
Cycle 1 - Mise en œuvre et évaluation							
Réalisation du projet expérimental							
Cycle 1 - Mise en œuvre et évaluation							
Validation du projet expérimental							
Cycle 1 - Évaluation finale du cycle							
Cycle 2 - Opérationnalisation du projet expérimental							
Rédaction mémoire							

On retrouve les principales étapes de la démarche de recherche présentée au chapitre 1. Chacune d'entre elle correspond à une ou plusieurs phases des cycles P1 et P2 de la démarche présentée dans ce chapitre. Le tableau apporte une vision de haut niveau de l'enchaînement temporel des phases des cycles stratégique et opérationnel et met en évidence le parallélisme de certaines phases dans le temps. La phase de lancement et planification du cycle s'est étendue de septembre 2014 à décembre 2014. Les parties prenantes identifiées ont été: les gestionnaires TI, les gestionnaires de plusieurs secteurs d'affaires (Assurance Individuelle, Épargne Individuelle et Assurance Auto et Habitation), les architectes de l'architecture d'entreprise et l'équipe de recherche universitaire composée de mon directeur de recherche et moi-même. Les principaux enjeux soulevés ont été l'exploitation des données internes pour répondre aux objectifs d'affaires, ainsi que la compréhension de l'impact des nouvelles approches et techniques existantes sur le marché liées à la gestion des données massives. La phase de préparation à la mise en œuvre s'est déroulée de janvier 2015 à la fin de l'été 2015, et a permis de dresser le portrait initial de l'iA. Les conclusions qui en sont ressorties ont été les suivantes :

- Le modèle d'affaires actuel (en 2015) de l'iA est très orienté autour du produit
- Le modèle d'affaires est organisé en silos entre les différents secteurs d'affaires, laissant peu de place à la collaboration et à l'intégration de l'information transversale issue des activités des différentes lignes d'affaires, transactions, opérations de vente, du service à la clientèle
- L'iA a peu de connaissances sur lesquelles se baser pour mieux comprendre les attentes de ses clients, hormis celles transmises par le réseau de distribution, ce qui mène à :
 - Une vision parcellaire de l'activité du client
 - Peu de relations directes avec le client
- L'iA possède une clientèle actuelle qui est peu représentative des générations futures de consommateurs
- Les nouvelles technologies disponibles sur le marché sont peu exploitées en interne dans les processus de conception et de développement de produits
- Les activités de vente, de développement de nouveaux services et de suivi de la performance sont basées sur une perception très intuitive qui s'appuie peu sur des données concrètes

L'analyse du niveau de maturité initial (en 2015) de l'iA a conduit à la conclusion que l'iA est encore peu mature en matière d'exploitation des données internes, qu'elle en est encore au stade d'essayer de comprendre ce que les techniques d'exploitation de données peuvent lui apporter pour faire croître l'entreprise. L'expérience tirée de cette phase a permis de comprendre qu'iA n'était pas prête à exploiter directement tout la richesse des données qu'elle possède et que les prérequis essentiels pour être en mesure de le faire n'étaient pas atteignables dans l'immédiat. On y reviendra plus en détail en traitant de l'évaluation de la maturité d'iA dans la section 2.3.2.

La phase suivante, la planification de la mise en œuvre (P1.3) qui s'est étendue de juin 2015 au milieu de l'automne 2015 a servi à identifier les moyens à déployer pour mettre en œuvre la stratégie à établir. En effet, les gestionnaires des différents secteurs d'affaires ont été réunis et consultés régulièrement afin de s'assurer du maximum de consensus sur les décisions de mise en œuvre prises. Des comités de réflexion ont aussi été mis en place pour convenir d'une vision et d'un plan d'action communs pour adresser les grands enjeux identifiés : 1) réduire les silos, 2) adopter une orientation produit des activités, 3) tirer parti de nouvelles approches et techniques d'exploitation de données. Ajouté à cela, des équipes de projets composées de personnes ressources internes et externes ont été mises en place pour proposer diverses expérimentations adressant un ou plusieurs des enjeux identifiés et un budget a été alloué à cette fin. Enfin, des échéances de temps ont été proposées pour mener à bien les différentes initiatives sur un horizon de 1 à 2 années, le temps de stabiliser la vision et le plan d'action une fois établis.

En parallèle à la planification (P1.3), la phase de mise en œuvre et évaluation (P1.4) a débuté en juillet 2015 pour s'achever un an plus tard, au milieu de l'été 2016. Les parties prenantes se sont entendues sur une nouvelle vision stratégique globale : devenir une compagnie moins orientée-produits et plus centrée sur les besoins du client.

Les grandes lignes du plan d'action mis en place pour atteindre cette vision ont été de :

- Développer davantage l'exploitation des données internes pour accroître la connaissance client en rassemblant les données clients issues des différents secteurs et lignes d'affaires impliqués dans le projet
- Favoriser l'innovation en matière de création de nouveaux produits et services en développant des partenariats avec des entreprises digitales très actives sur le marché
- Suivre de plus près les évolutions technologiques du marché en testant de nouvelles technologies dans le cadre de projets pilotes

Trois projets pilotes majeurs ont été mis en place pour ce plan d'action : une expérimentation sur le lancement de 3 à 5 nouveaux produits et services innovants; un projet pilote d'exploitation des données internes provenant de différents secteurs d'affaires dans une optique de développement de la connaissance client (présenté plus en détails dans le chapitre 3 de ce mémoire) et un projet de lancement d'une cellule d'incubation.

L'expérience tirée de l'exécution des projets pilotes a permis de mettre en évidence la nécessité de faire des réajustements au niveau du plan d'action car, malgré le succès global des projets pilotes, notamment les apprentissages qui en ont été tirés, certains enjeux n'avaient pas pu être atteints et demeuraient des cibles à atteindre. Les sections 2.4.2 et 2.4.3 reviendront plus en détails sur le déroulement de cette phase dans le cas de l'iA.

Finalement, la phase d'évaluation globale du cycle qui s'est déroulée durant l'été 2016 a mis en avant les points de réussite et d'amélioration majeurs et essentiels pour tirer une conclusion globale sur le cycle de gestion stratégique du changement en cours. Divers éléments clés de réussite ont été retenus par les parties prenantes au projet : 1) les retours positifs des bassins de clients ayant testé les offres de produits et services qui ont été proposés dans le cadre du premier projet pilote; 2) la prise de conscience unanime au sein de l'organisation de l'importance et du potentiel de l'exploitation de données en soutien des activités de développement d'affaires ;3) la qualité des efforts déployés concernant l'analyse des données clients a été soulignée.

Deux points importants d'amélioration ont alors été retenus : 1) l'urgence d'opérer plusieurs changements dans le modèle d'affaires et la culture d'entreprise actuels afin d'être en mesure d'opérationnaliser la vision et le plan d'action établis; 2) la nécessité d'investir de manière importante pour augmenter la qualité des données détenues à l'interne afin d'obtenir de meilleurs résultats d'analyse a été fortement mis en avant.

2.2.3 Détails sur le déroulement du cycle d'opérationnalisation du changement

On rappellera tout d'abord que notre mémoire porte principalement sur le cycle de gestion stratégique du changement. Néanmoins, le cycle d'opérationnalisation sera brièvement présenté dans cette section et un exemple d'illustration de son application sera développé dans le chapitre 4.

Comme illustré dans la Figure 2.1, la première phase de ce cycle est la préparation à la planification (P2.1 dans la Figure 2.1). En effet, l'achèvement du cycle de stratégie de gestion du changement s'accompagne d'une actualisation du niveau de maturité de l'organisation qui devrait avoir évolué positivement afin de se poursuivre par l'opérationnalisation de certains résultats. Au démarrage de la phase d'opérationnalisation du changement, il est important d'identifier principalement les acteurs clés du projet en prenant en considération les recommandations issues de l'évaluation finale du cycle 1 (P1.5).

Par exemple, on peut se demander dans quelle mesure les équipes opérationnelles, d'infrastructure et d'exploitation vont être impliqués dans le déroulement de ce cycle d'opérationnalisation. Au plus tôt les bons acteurs sont identifiés au mieux le cycle avancera.

Ensuite, la phase de planification (P2.2) sert, entre autres, à développer une vision et un plan d'action opérationnel couplés à la vision et au plan d'action stratégique préalablement établis, à préparer la mise en place de projets, non plus pilotes mais opérationnels, qui vont avoir un impact direct sur certains processus d'affaires de l'organisation, et enfin d'estimer les efforts de réalisation. Ces efforts sont plus importants que ceux déployés dans le cadre du cycle 1(P1) car il s'agit de planifier des projets de plus grande envergure. Cela implique de tenir compte des systèmes actuellement en production, et cela requiert donc l'implication de plusieurs équipes opérationnelles.

La phase de réalisation opérationnelle (P2.3) consiste principalement en la conduite des projets planifiés lors de la phase précédente. Finalement, le cycle se termine par la phase de développement de nouveaux processus d'affaires (P2.4). Les résultats des projets lancés entraînent une révision de certains processus d'affaires, voire la création de nouveaux. Cela implique bien souvent de nombreux changements qui ont des impacts à tous les niveaux de l'organisation. C'est la raison pour laquelle le déploiement de ces processus nécessite une attention particulière. La conduite de ce cycle opérationnel accroît la maturité de l'entreprise au même titre que le cycle 1 de gestion stratégique du changement. Ce cycle d'opérationnalisation s'étend sur une durée beaucoup plus grande que le cycle 1. On peut effectivement compter raisonnablement entre 1 et 2 années pour le premier cycle stratégique et 3 à 5 ans pour son application opérationnelle. L'exécution de l'opérationnalisation peut, en outre, alimenter graduellement le lancement d'un ou de plusieurs nouveaux cycles de gestion stratégique du changement.

Dans le cas de l'iA, le cycle d'opérationnalisation a démarré en cours d'été 2016. En effet, davantage consciente du potentiel des données qu'elle possède, l'iA a établi une vision et un plan d'action stratégique qui vise à l'exploiter progressivement. Dès l'été 2016, des équipes de développement, d'exploitation, d'administration de systèmes ainsi que des fournisseurs de solution partenaires ont été intégrés aux discussions portant sur l'opérationnalisation des travaux pilotes réalisés au cours de l'année précédente. La vision du changement opérationnel qui en a résulté, a consisté à transformer les activités opérationnelles en lien avec les affaires dans une perspective plus orientée client que produit. Le plan d'action qui s'en est suivi consista en la mise en place de projets qui favorisent la réalisation de cette vision au fur et à mesure du développement de la capacité des équipes à intégrer le changement.

Voici quelques exemples de projets qui ont été lancés:

- Le lancement d'une mission avec les équipes marketing et TI pour déceler des bassins de clients potentiellement intéressés par l'achat de nouveaux produits et dont les références seraient transmises au réseau de distribution (ceci fera l'objet du chapitre 4)
- La mise en production de nouveaux produits issus de partenariats avec des compagnies dites « digitales »
- La mise en place d'équipes dédiées à la gestion de l'information client de la compagnie

Actuellement (début 2017), une réflexion est en cours au sein de l'iA sur la création de processus d'affaires permettant d'augmenter la qualité des données, notamment au niveau de la saisie de données faite par les agents du réseau de distribution et les agents de service à la clientèle. De plus, les processus de transmission des données des systèmes opérationnels vers le système client centralisé sont actuellement en cours d'évaluation afin d'être révisés et améliorés.

2.3 Raffinement de la préparation à la mise en œuvre de la gestion stratégique du changement

2.3.1 Présentation globale de cette phase

On a vu dans la section 2.2.2 que la phase de préparation à la mise en œuvre (P1.2 dans la Figure 2.1) se situe entre la phase de lancement et planification du cycle (P1.1) et celle de la planification de la mise en œuvre. Cette phase est particulièrement importante car elle apporte les fondements sur lesquels la vision et le plan d'action stratégique vont s'appuyer. En effet, la préparation à la mise en œuvre permet de dresser le portrait de la compagnie (dans notre cas pratique celui de l'iA) par rapport à son environnement externe. À ce titre on fait de la veille stratégique et technologique (P1.2.1 dans la Figure 2.2) et on étudie les transformations du marché et de la société (P1.2.2). Ces deux étapes de la phase peuvent être menées en parallèle. L'expérience tirée de ces deux étapes apporte une connaissance externe à l'entreprise en lien avec changement organisationnel qu'elle désire mettre en place. La dernière étape de cette phase, l'évaluation de la maturité (P1.2.3), peut également être démarrée parallèlement aux deux autres phases (P1.2.1 et P1.2.2). L'évaluation de la maturité sert de point de départ à l'élaboration la stratégie de changement de l'entreprise. On identifie d'où l'on part en termes de maturité et où on souhaite arriver. L'organisation développe ainsi de l'expérience qui s'ajoute à l'expérience acquise du cours des autres phases du cycle. L'expérience cumulée des 3 étapes de la phase alimente la dernière étape : bâtir une expérience progressive dans la préparation de la mise en œuvre stratégique du changement (P1.2.4).

La Figure 2.2 ci-dessous, illustre les différentes étapes de cette phase.

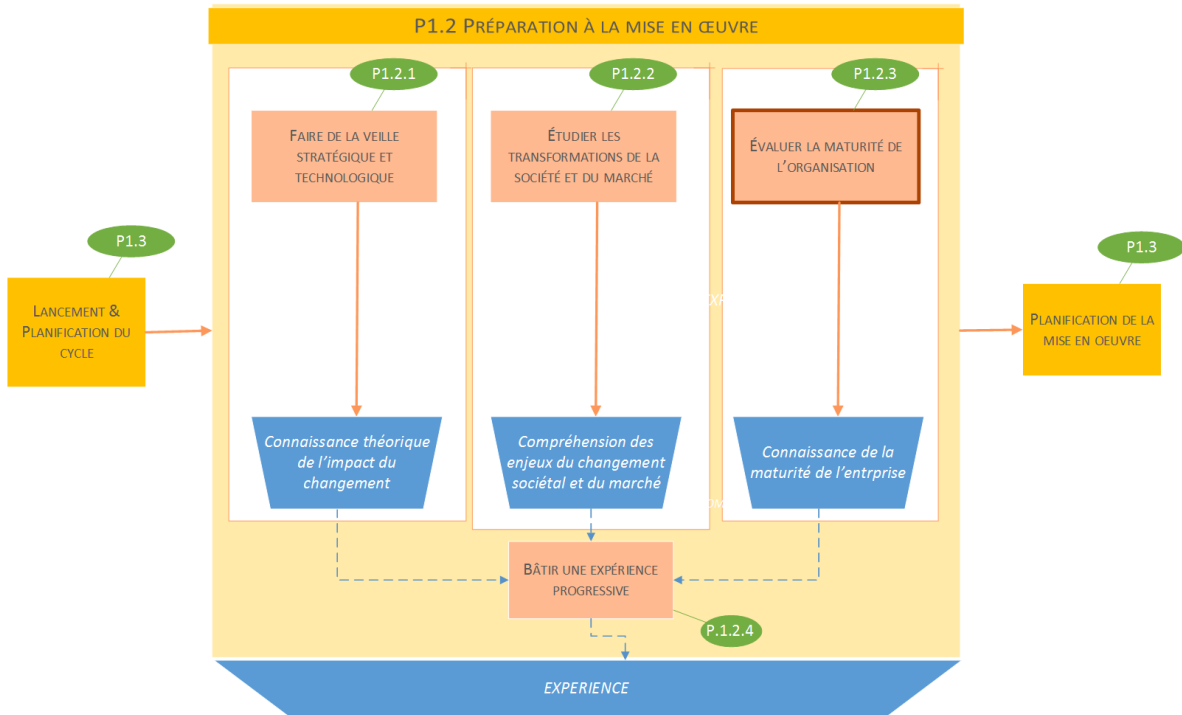


Figure 2.2: Diagramme de raffinement du cycle 1 - Préparation à la mise en œuvre

Les rectangles roses représentent les étapes clés de cette phase. On retrouve ensuite les trapèzes bleus qui correspondent à des accumulations de connaissance, comme dans la Figure 2.1, servant à bâtir l'expérience progressive acquise durant la phase. Cette expérience progressive conduit à la construction de l'expérience plus généralement acquise au travers de la phase matérialisée par le grand trapèze bleu à la base du rectangle global de la phase.

Pour bien comprendre le déroulement de cette phase nous allons préciser les notions de veille stratégique et technologique :

La veille est une activité continue et en grande partie itérative visant à faire une surveillance active de l'environnement technologique et commercial de l'entreprise [...afin d'] anticiper les évolutions [...]. La veille technologique regroupe l'ensemble des domaines susceptibles de donner à l'entreprise un avantage concurrentiel basé sur les techniques, sciences, méthodes et procédés liés à de nouvelles technologies [...]. La veille stratégique permet de balayer les principaux sujets d'évolution observables. La compréhension que l'on tire de ces deux processus permet de positionner l'entreprise en tenant compte de ces évolutions (Baumard 1998).

Pour dresser un portrait fidèle de la situation actuelle de la compagnie d'un point de vue stratégique et technologique il est important de tenir compte de l'évolution de son environnement externe. C'est pourquoi la veille stratégique est une étape importante de la préparation à la mise en œuvre de la gestion stratégique du changement.

Pour cela, il est important de balayer d'abord le plus largement possible les sujets ou domaines qui touchent de près ou de loin à l'activité de l'entreprise, puis de réduire le champ de recherche sur les sujets clés sur lesquels portera le changement envisagé. Une autre étape essentielle est l'étude ou la surveillance des tendances qui sont les plus actuelles sur le marché et dans le fonctionnement de la société de manière générale (P1.2.2 dans la Figure 2.2). Ce que l'on entend par transformation est la matérialisation d'un changement brutal en matière de conscience collective d'une société, ou d'un marché, (au niveau local, national ou mondial) à tel point que la réalité de cette société ou le marché en est redéfini (Wikipedia 2016). Les transformations de cette envergure apparaissent souvent par le biais d'un stimulus externe ou suite à une volonté forte d'une société ou d'un marché d'initier un changement. Par exemple, les découvertes scientifiques ont déclenché beaucoup de transformations sociales à travers l'histoire (Jensen 1993). Cette étape permet de comprendre en quoi notre société et le marché évoluent et orientent ainsi la direction du changement à suivre. La dernière grande étape qui compose la phase est l'évaluation de la maturité (P1.2.3 dans la Figure 2.2). Faire de la veille stratégique et technologique et faire état des évolutions de notre société et du marché en les comparant à la situation de l'entreprise ne suffit pas. Il est crucial de comprendre également dans quelle mesure l'entreprise est capable d'opérer elle-même les changements nécessaires pour suivre ces transformations. L'exécution de ces trois étapes aide à bâtir l'expérience progressive essentielle à la réussite du déroulement de la phase. En effet, on acquiert de l'expérience au cours de chacune de ces étapes. L'expérience qui se bâtit progressivement tient compte des éléments de complémentarité entre chaque étape : la veille technologique et stratégique explique partiellement certains changements de la société et du marché, et l'évaluation de la maturité permet d'envisager l'impact de ces changements sur l'organisation.

Voyons à présent comment les étapes de cette phase se sont déroulées dans le cas d'iA. Sur une période de 6 mois, nous avons réalisé une revue de littérature approfondie. Le choix des sujets à explorer a été fait en consultant les parties prenantes au projet.

Les principaux éléments de connaissance que nous avons recueillis au cours de de cette étape et qui sont d'intérêt pour l'iA sont présentés suivant différents volets (démographique, économique, social, technologie et organisationnel), tirés pour la plupart d'un rapport de la firme de consultation Pwc nommé « Project Blue » (Vooght, Jurczynski, et Nevin 2015).

- Études des tendances démographiques
 - Les générations de clients de demain seront les générations Y et Z. Ces générations sont à l'affût des dernières technologies (mobiles, tablettes, objets connectés, e-commerce) et ont des besoins différents des générations précédentes. En effet, par rapport à la génération de leurs parents et grands-parents, les générations Y et Z sont plus autonomes, plus exigeantes, plus volatiles, moins fidèles à une marque en particulier.
- Études des tendances économiques
 - Les pays émergents (Amérique du Sud et Afrique surtout) se développent et représentent des marchés très prometteurs pour le secteur de l'assurance
 - L'économie de partage bouleverse les règles de l'offre et de la demande du marché
 - Les entreprises issues du digital représentent une menace potentielle pour les entreprises traditionnelles car elles sont avant-gardistes technologiquement parlant et sont capables de répondre aux besoins du marché très rapidement
 - L'économie locale prend petit à petit le dessus sur l'économie globale (on met de plus en plus en valeur le savoir-faire local)
- Études des tendances sociales
 - La classe moyenne augmentera d'environ 180% d'ici 2040, le secteur de l'assurance doit donc s'adapter à la montée en puissance de cette classe sociale dont les besoins en matière d'assurance peuvent être différents
 - Les médias sociaux, par leur essor, peuvent fragiliser l'image de marque de certaines compagnies car ces dernières sont davantage exposées à l'opinion publique
- Études des tendances technologiques
 - D'ici 2020 il y aura dans le monde 2 fois plus d'objets connectés que d'individus
 - La demande pour des services bancaires et d'assurance sans intermédiaire augmente significativement, toute classe sociale confondue
 - L'informatique nuagique, l'ère du Big Data, l'apprentissage automatique, ainsi que l'internet des objets sont au cœur des développements futurs et deviennent de plus en plus pertinents à intégrer dans le secteur des assurances : les entreprises leaders du marché ont déjà commencé à le faire.
- Études des tendances organisationnelles
 - De nouveaux modèles d'affaires émergent et révolutionnent l'organisation du travail et les modes de collaboration à l'intérieur des entreprises, mais aussi avec les partenaires externes. On parle de plus en plus de l'entreprise entrepreneuriale, l'entreprise digitale, l'entreprise connectée ou encore de l'entreprise orientée-client.

Ces conclusions ont été présentées aux gestionnaires impliqués dans le projet. Des questions prioritaires pour l'iA ont été choisies: l'orientation client et l'exploration des nouvelles tendances technologiques pertinentes dans l'industrie des assurances (l'informatique nuagique, l'ère du Big Data, l'apprentissage automatique, l'internet des objets).

Afin d'évaluer la maturité de l'iA compte tenu de ces éléments, nous avons rencontré des experts et gestionnaires de différents secteurs d'affaires afin d'évaluer d'une part la situation actuelle de l'iA par rapport aux tendances observées, et d'autre part l'envergure des efforts à déployer pour opérer les changements nécessaires pour suivre les tendances priorisées. Par la suite, nous avons réalisé une nouvelle revue de littérature plus spécifique portant : 1) sur l'étude des modèles de maturité en lien avec l'orientation organisationnelle centrée – client; 2) sur l'exploration de nouvelles technologies et techniques dans le cadre de l'exploitation des données internes à l'iA. Cette nouvelle revue de littérature a fait l'objet de la section 1.3 du chapitre 1. En ce qui concerne la démarche détaillée que nous avons proposé pour évaluer la maturité, elle sera développée dans la sous-section suivante.

2.3.2 L'évaluation de la maturité

Soulignons en premier lieu que cette étape est déclenchée par le constat d'un décalage fort entre les changements externes observés et la situation actuelle de l'entreprise.

Les objectifs majeurs de cette étape sont tout d'abord de faire le bilan de la situation actuelle de l'entreprise sur les thématiques identifiées. Dans le cas de l'iA il s'agit de l'orientation client et de l'analyse avancée des données clients internes. Ensuite, l'objectif est d'identifier précisément l'ampleur de cet écart. Enfin, on s'attèle à examiner l'envergure des efforts à déployer pour réduire ces écarts (Morabito 2015).

La démarche à suivre que nous proposons est composée des étapes suivantes :

- 1. Identifier les domaines de l'entreprise impliqués dans ce processus d'évaluation de la maturité et faire un bilan de l'organisation par rapport à l'environnement externe en matière de :
 - Gestion de l'information
 - Quelles technologies sont exploitées pour supporter les activités de l'entreprise ?
 - Quelles démarches de collecte des données servent aux activités quotidiennes et de planification de l'entreprise ?
 - Quelles sont les interactions existantes entre le développement des produits et services et l'information générée par les données que possèdent la compagnie ?
 - Objectifs d'affaires
 - Quelle est la mission ou vocation de l'entreprise ?
 - Quelle est l'offre de produit et services ?
 - Avec quelle efficacité l'organisation crée-t-elle de la valeur d'affaires ?
 - Culture d'entreprise
 - Quelles sont les valeurs de travail et performance qui définissent la culture de l'entreprise?
 - Processus d'affaires
 - Quels processus d'affaires liés au client et à la gestion de ses informations sont en place ?

- 2. Choisir un modèle de maturité (qui est en général tiré d'un ou plusieurs modèles existant dans la littérature) qui servira d'étalon pour évaluer la compagnie. Ce modèle (ou combinaison de modèles) se décompose en niveaux croissants (appelés « stades ») et chaque nouveau stade franchi représente un degré de maturité supérieur atteint.
- 3. Identifier les conditions que doivent remplir l'organisation pour compléter chaque niveau

On identifie que l'étape est complétée lorsque :

- Le niveau de maturité a bien été identifié
- On est capable d'établir une cible d'évolution qui va servir de point de départ à l'établissement d'une vision

Sinon, on persiste dans l'évaluation de la maturité ou on abandonne le projet dans le cas où l'entreprise s'aperçoit qu'elle n'est pas prête à opérer des changements dans son fonctionnement.

Dans le cas de l'iA la combinaison des modèles de maturité choisie a déjà été présentée au chapitre 1. Les conclusions sur l'évaluation du stade de maturité de l'iA faite au démarrage du projet en tenant compte de ces modèles est présentée ici :

- Modèle Data Flux
 - Stade 1 (Indiscipliné) car la qualité globale des données détenues est de ce fait grandement affectée. En effet, les systèmes existants actuellement à l'iA n'ont pas été construits, pour la majorité d'entre eux, dans la perspective de faire de l'analytique mais uniquement de manière à répondre à des besoins opérationnels
- Modèle Delta de Thomas Davenport
 - Stade 1 (Analytiquement diminué) car :
 - L'organisation n'a pas un ou plusieurs des prérequis pour faire de l'analytique avancé
 - Les données sont de faible qualité, inconsistantes et peu organisées pour faire de l'analytique
 - Il y a peu d'intérêt ou de compréhension du potentiel de l'analytique
 - Il manque des compétences et ressources à l'interne dans le domaine de l'analytique
- Orientation client : modèle EY
 - Stade 1 : Non développé car
 - Les produits sont développés de manière « ad hoc » pour convenir au plus grand nombre de clients possibles
 - Le personnel en contact direct avec le client n'est pas suffisamment outillé pour accompagner celui-ci
 - Il n'y a pas de stratégie technologique commune de gestion de la clientèle
 - Il y a des disparités au niveau de la reconnaissance de la marque selon le secteur auquel le client a affaire
 - L'offre de service est indifférenciée; il n'y a pas de segment de client identifié par la compagnie
 - L'accent est mis sur les nouvelles ventes et pas suffisamment sur la rétention de clients
 - Les produits classiques de la marque suivent un cycle de vie traditionnel du secteur des assurances et sont donc complexes, peu flexibles, opaques, axés sur la maximisation du profit

- Il n'existe pas de processus d'identification de la valeur à long terme du client à l'interne de l'organisation
- Il n'y a pas de métrique ou indicateur clé produit sur la clientèle et partagé de manière transversale dans l'organisation

Le bilan global de maturité, compte tenu de ces trois modèles, est très faible mais offre un potentiel de croissance énorme sur lequel sont basés la vision et le plan d'action qui ont été établis par la suite.

2.4 Raffinement de la mise en œuvre et de l'évaluation intermédiaire de la gestion stratégique du changement

2.4.1 Présentation globale de cette phase

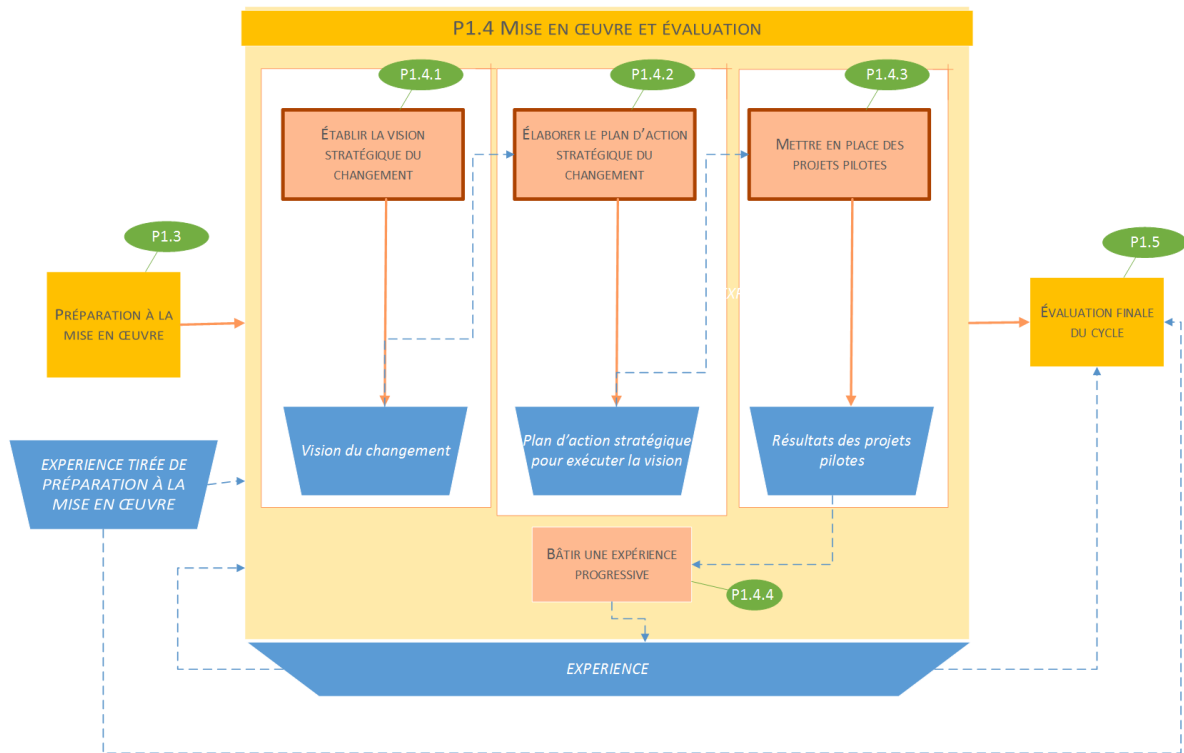


Figure 2.3 : Diagramme de raffinement du cycle 1 - Mise en œuvre et évaluation

La phase de préparation à la mise en œuvre (section 2.3) a abouti à l'identification du niveau de maturité actuelle de la compagnie. Si la volonté de l'organisation se confirme de vouloir opérer un changement pour évoluer par rapport à la maturité actuelle identifiée, la phase de mise en œuvre du changement (P1.4.1 dans la Figure 2.3) est enclenchée. Nous la présentons en détails dans cette section (voir Figure 2.3)

Le changement organisationnel étant un processus relativement lent, il est important de se projeter à long terme : c'est l'essence de l'établissement de la vision stratégique (processus P.4.1 dans la Figure 2.3). La question principale que cette étape soulève est : Étant donné le niveau de maturité actuel de l'entreprise, quel niveau supérieur veut-on atteindre de manière réaliste et sur quel horizon de temps ? Le plan d'action, quant à lui, sert à prévoir les actions nécessaires pour bâtir la vision dans l'horizon de temps fixé. On élabore un tel plan d'action à partir d'une vision à laquelle adhèrent les parties prenantes (phase P1.4.2). Enfin, en tenant compte de la vision et du plan d'action, on met en place des projets pilotes dits « expérimentaux » (phase P1.4.3 dans la Figure 2.3). Ceux-ci permettent d'ajuster ou de valider le plan d'action et la vision en testant la capacité de l'organisation à remplir les nouveaux objectifs qu'elle a adoptés lors de l'élaboration de la vision et du plan d'action. Ces étapes sont dépendantes les unes des autres. En effet, nous recommandons de construire le plan d'action à partir de la vision, une fois celle-ci établie. De la même façon, le plan d'action oriente la mise en place des projets pilotes, qui permettent de valider ou non la vision et le plan d'action lui-même.

L'expérience progressive acquise est le résultat de l'évaluation du bon déroulement des trois étapes et qui se matérialise souvent par l'analyse des résultats des projets pilotes (phase P1.4.4 dans la Figure 2.3). En effet, les conclusions globales des projets pilotes constituent en grande partie l'expérience de la phase. Cette phase a plus de chance d'être un succès lorsqu'elle s'exécute par itérations et c'est l'expérience progressive tirée de chacune des itérations successives qui aiguille l'itération suivante. Ceci est matérialisé sur la Figure 2.3 par la flèche bleue de l'expérience qui pointe à la fois à nouveau vers le rectangle global la phase de mise en œuvre et sur la phase suivante « Évaluation finale du cycle » (processus P1.5).

Dans le contexte de l'iA, sachant que l'entreprise a évalué son niveau de maturité sur les thématiques de l'exploitation des données internes et sur l'orientation-client de l'entreprise, voici les conclusions tirées lors du déroulement des étapes de la phase P1.4 :

- L'iA est rendue à un stade de maturité faible en matière de gestion des données internes pour servir les objectifs stratégiques de l'organisation. L'organisation vise à développer des capacités analytiques qui lui permettront d'avoir une approche d'exploitation de données plus transversale et capable de servir les besoins stratégiques de la compagnie
- Son modèle de vente est presque essentiellement orienté vers la vente de nouveaux produits et très peu vers la recherche de la satisfaction du client. Une connaissance accrue des profils et besoins des clients doit être visée par l'iA
- L'élaboration de la vision, du plan d'action et de la mise en place de projets pilotes dans le cas de l'iA seront développés dans les sous-sections 2.4.2, 2.4.3 et 2.4.4.
- Les résultats des projets pilotes ont été suffisamment probants pour valider la vision et le plan d'action, ainsi que pour envisager une opérationnalisation de cette stratégie de changement sur un plus court terme

2.4.2 L'élaboration de la vision stratégique du changement

Cette étape (processus P1.4.1 dans la Figure 2.3) est déclenchée lorsque la maturité actuelle est identifiée et que l'on cherche à déterminer le niveau de maturité supérieur que l'on vise.

L'objectif principal de l'étape d'élaboration de la vision stratégique du changement est de définir des orientations claires de changement que doit prendre l'organisation pour atteindre un niveau de maturité supérieur qui est à déterminer.

La démarche proposée pour cette étape est la suivante :

- 1. Identifier le niveau de maturité supérieur que l'on veut atteindre
 - Qu'est ce qui justifie l'état de maturité choisi, compte tenu du niveau de maturité actuel ?
- 2. Déterminer quelles alternatives d'évolution de vision sont envisageables pour atteindre ce nouveau degré de maturité
- 3. Prioriser une orientation d'évolution en se basant sur :
 - Les tendances du marché
 - Les tendances de la société
 - L'environnement externe / concurrentiel
 - Les perspectives d'évolution de l'industrie

On considère que l'étape est complétée lorsque :

- On est capable d'identifier des orientations d'évolution qui reflètent les besoins de l'entreprise, compte tenu du niveau à atteindre
- Une orientation spécifique a été choisie qui va servir de point de départ pour l'établissement de la vision

Si on constate que malgré une volonté de changement exprimée, le sens que doit prendre ce changement n'est pas clairement formulé ou que l'entreprise n'arrive pas à se projeter suffisamment dans le futur, alors il est souhaitable de revenir à l'étape d'évaluation du niveau de maturité.

Pour l'iA, les changements organisationnels envisagés en matière de d'exploitation de données, d'objectifs d'affaires et de culture d'entreprise ont été identifiés comme suit :

- Gestion de l'information
 - Faciliter l'accès à des données issues de sources d'information variées et internes à l'organisation
 - Augmenter la qualité des données collectées pour créer un climat de confiance au sein de l'entreprise un climat de confiance envers les données exploitées
 - Tirer parti des avancées technologiques pour faciliter le traitement et l'analyse de grands et complexes volumes de données
- Objectifs d'affaires
 - Marché visé
 - Dénicher de nouveaux marchés et/ou renforcer la présence de l'iA sur les marchés actuels grâce à des informations pertinentes sur les clients tirées de l'analyse des données collectées

- Vocation
 - Mettre le client (plutôt que le produit) plus au cœur des activités de l'entreprise
- Offre de produits et services visée
 - Examiner les opportunités de conception de nouveaux produits et services basées sur l'analyse avancée de données-clients
 - Soutenir les décisions tactiques et stratégiques grâce à l'exploitation des données internes
- Processus d'affaires
 - Établir des processus d'intégration d'analyses de données dans les activités d'affaires en lien avec la gestion de la clientèle
- Culture d'entreprise
 - Augmenter le niveau de collaboration et de partage de l'information client entre les secteurs d'affaires de l'organisation
 - Augmenter les aptitudes/compétences/expertises des équipes en ce qui concerne l'analyse des données

Compte tenu de l'évolution envisagée, le nouveau niveau de maturité visé par l'iA a été redéfini selon les orientations d'évolution proposées par différents modèles de maturité choisis et présentés dans le chapitre 1. Ainsi on vise un niveau de maturité caractérisé par :

- Modèle Data Flux :
 - Stade 3 (Pro active) : l'iA commence à envisager la mise en place de systèmes de gestion de l'information et les équipes TI et des divers secteurs d'affaires collaborent
- Modèle Delta de Thomas Davenport
 - Stade 3 (Aspirations analytiques) : l'entreprise cherche à bâtir des capacités analytiques communes à toute l'entreprise
- Orientation client : modèle EY
 - Stade 2 (Émergent) : l'iA veut développer des processus d'affaires adaptés aux différents profils de ses clients en développant une stratégie technologique de segmentation de sa clientèle. Cette stratégie pourrait être initiée par le développement d'activités de croisement de l'information détenue par les différents secteurs d'affaires. Enfin, la compagnie souhaite accompagner davantage son réseau de distribution chargé des ventes de produits aux clients en mettant à sa disposition des moyens de mieux comprendre la clientèle, notamment en commençant par produire des métriques et indicateurs clés liés aux clients

Une priorisation a été faite parmi ces orientations. On a choisi tout d'abord, de dresser le portrait de la clientèle afin de mieux comprendre la relation que celle-ci a avec l'iA actuellement. Ensuite, le développement de capacités analytiques communes à l'entreprise a été jugé comme prioritaire dans l'établissement de la vision. Ces deux axes sont les fondements de la nouvelle vision à bâtir : *développer des capacités analytiques pour optimiser la relation de l'iA avec ses clients afin de devenir progressivement une entreprise orientée - client*. L'horizon temporel fixé pour atteindre cette vision a été estimé à 9 mois.

2.4.3 L'élaboration du plan d'action

Lorsque la vision du changement est adoptée par les dirigeants et gestionnaires de l'entreprise, la réflexion sur le plan d'action à construire est enclenchée (processus P1.4.2 dans la Figure 2.3).

Dans cette étape on vise d'une part à définir des objectifs permettant de concrétiser la vision d'évolution identifiée, et d'autre part à déterminer les actions à mettre en place pour atteindre ces objectifs.

À ce titre, nous préconisons de suivre les étapes suivantes :

- 1. Analyser les orientations d'évolution proposées par les modèles de maturité choisis
- 2. Identifier les objectifs à atteindre compte tenu de ces orientations
- 3. Déterminer les résultats attendus
- 4. Identifier les activités à réaliser pour aboutir à ces résultats et atteindre les objectifs fixés
- 5. Formaliser le plan d'action
 - Qui sera chargé de chacune des activités ? Il est essentiel d'attribuer des responsabilités et de préciser les objectifs que doivent remplir les responsables
 - Quand ? Un calendrier prévisionnel des activités est nécessaire
 - Dans quel ordre les activités du plan d'action vont s'exécuter ? Il est important de définir la séquence des activités et d'identifier les interdépendances existantes
 - Comment le plan d'action va être exécuté ? Définir les besoins humains, techniques, matériels et financiers est indispensable
 - Comment suivre le bon déroulement du plan d'action ? Sélectionner les mesures et indicateurs pouvant être utilisés pour suivre l'avancement et la performance du plan d'action permet de s'assurer du bon déroulement du plan d'action
- 5. Évaluer la faisabilité du plan d'action
 - Replacer les activités du plan d'action par rapport au modèle de maturité choisi pour s'assurer qu'il s'applique au stade de maturité ciblé
 - Consulter les parties prenantes au projet pour valider le plan d'action stratégique

L'étape est complétée lorsqu'on considère que le plan d'action est adapté aux capacités de l'entreprise, compte tenu de son niveau de maturité actuel. Si on s'aperçoit que le plan d'action ne reflète pas la vision d'évolution au niveau désiré, ou que les moyens requis pour mettre en œuvre ce plan d'action sont insuffisants ou indisponibles, ou encore que le plan d'action n'est pas appuyé par les principaux acteurs du projet, il faut impérativement réévaluer la vision établie à l'étape précédente.

Les principales recommandations d'actions proposées par les modèles de maturité choisis dans le cas de l'iA sont de :

- Sensibiliser les équipes des différents secteurs d'affaires à l'importance de l'exploitation des données clients internes
- Lancer des projets pilotes pour mettre en avant le potentiel des données détenues à partir de l'extraction de données de différentes sources de données
- Construire un entrepôt de données centralisant toute l'information client, et à partir des données de production
- Intégrer des données externes démographiques à l'entrepôt de données
- Étendre le périmètre d'analyse aux données issues des filiales de l'iA

- Mettre en place une structure de gouvernance qui implique directement les secteurs d'affaires concernés

Quant aux résultats attendus par les parties prenantes du projet de changement, ils ont été principalement les suivants :

- Créer un engouement pour les approches analytiques chez les responsables exécutifs seniors de l'iA
- Créer un consensus entre différents secteurs d'affaires d'iA pour partager certaines de leurs données client afin de réaliser une première segmentation de la clientèle
- Avoir un portrait complet de la relation actuelle entre iA et ses différents segments de clientèle
- Mettre les techniques d'analyse et les préoccupations du client au cœur des processus d'affaires majeurs de l'entreprise

Plusieurs activités principales ont été identifiées pour atteindre ces résultats. La première activité a conduit à la constitution de comités composés de plusieurs ressources TI et des secteurs d'affaires avec pour mission de mener des expérimentations permettant de concrétiser le plan d'action établi. On a ainsi créé : *le comité de stratégie digitale, le comité de profilage client et le comité de la cellule d'innovation*. Le projet expérimental présenté au chapitre 5 a été supervisé par le comité de profilage client. Des rencontres fréquentes entre les comités de supervision et les équipes en charge de la mise en place des projets pilotes ont permis d'assurer la continuité du plan d'action, ainsi que de résoudre les problèmes et enjeux au besoin.

Pour finir, le plan d'action a été formalisé comme suit :

- Qui sera chargé de chacune des activités ?
 - Les parties prenantes du projet se répartissent dans les différents comités de supervision
 - Les projets pilotes sont menés par des équipes pluridisciplinaires : experts d'affaires, experts TI, équipe de recherche, consultants externes
- Quand ?
 - L'horizon temporel du projet dans sa globalité est de 2 ans
 - Les projets pilote se déroulent en parallèle, chacun ayant une durée de 6 mois à 1 an
- Dans quel ordre le plan d'action va s'exécuter ?
 - Étape 1 : créer l'adhésion
 - Étape 2 : impliquer des secteurs d'affaires et constituer les comités de supervision
 - Étape 3 : lancer les projets pilotes
- Comment le plan d'action va être exécuter ?
 - Identifier les besoins financiers pour chaque projet pilote
 - Cibler les besoins en ressources d'expertise technique et d'affaires pour atteindre les objectifs
- Comment suivre le bon déroulement du plan d'action ?
 - Les projets pilotes vont produire des métriques qui leur sont propres pour évaluer l'efficacité du plan d'action

2.4.4 Mise en place de projets pilotes

Cette étape (processus P1.4.3 dans la Figure 2.3) se déclenche lorsque la gestion stratégique du changement commence à se dessiner et que l'on veut s'assurer qu'elle est réaliste et opérationnalisable.

L'objectif est premièrement d'évaluer la validité du plan d'action et de la vision en les testant directement grâce aux projets pilotes, et deuxièmement d'identifier d'éventuels enjeux que seule la mise en œuvre du plan d'action peut amener à identifier.

Le diagramme de la Figure 2.4 ci-dessous illustre le processus de l'étape de mise en place de projets pilotes (P1.4.3 dans la Figure 2.3)

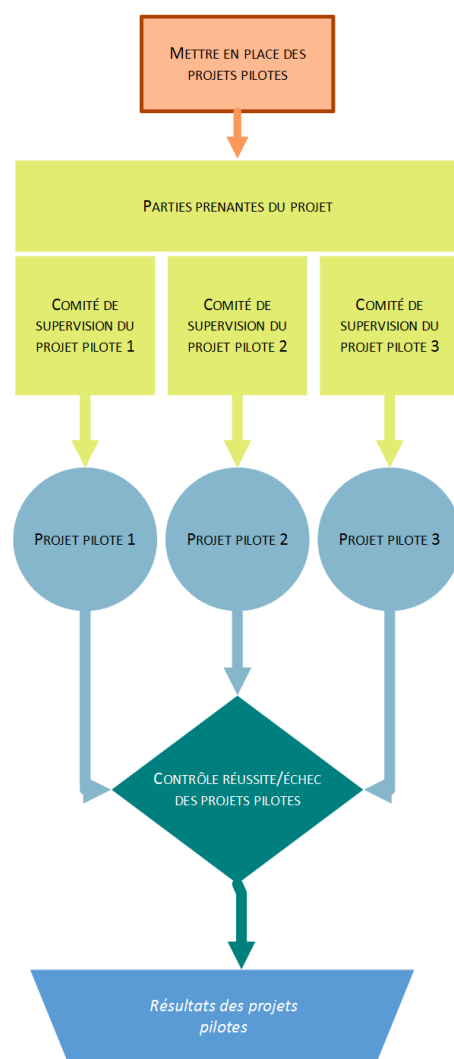


Figure 2.4: Diagramme de l'étape de mise en place de projets pilotes appuyant le plan d'action stratégique

La structure en vert clair représente le découpage de la structure de supervision de ces projets pilotes. Chaque comité supervise un projet en identifiant ses objectifs, le volet du plan d'action et de la vision qu'il permet de tester, ainsi que les résultats attendus pour juger de sa réussite.

Chacun des projets pilotes est ensuite exécuté par une équipe de réalisation dédiée (les projets sont exécutés en parallèle le plus souvent) selon une démarche fixée au préalable pour chaque projet et approuvée par le comité de supervision. Une fois que les premiers résultats obtenus de chaque projet sont analysés, une décision est émise pour chaque projet. Cette décision (losange vert dans la Figure 2.4 repose sur l'évaluation de la réussite ou de l'échec de la démarche poursuivie ainsi que sur la qualité des résultats obtenus, compte tenu des attentes exprimées (losange vert sur la Figure 2.4).

Cette décision est prise conjointement par le comité de supervision et l'équipe de réalisation de chaque projet. Ce processus décisionnel, répété pour chaque projet, permet d'aboutir à des conclusions plus globales sur les résultats du projet et sur l'expérience acquise à cette étape (l'expérience est matérialisée par le trapèze bleu de la Figure 2.4). À noter que la notion de résultat de projets peut prendre différentes natures en fonction du projet pilote considéré. En effet, l'échec de l'exécution de la démarche peut être vu, par exemple, comme un résultat du projet pilote. Tous ces résultats serviront d'intrants au moment de l'évaluation de l'efficacité du plan d'action et de la vision de la stratégie de changement adoptée (phase 1.5 de la Figure 2.1).

Pour réaliser cette étape, nous proposons la démarche suivante :

- 1. Mettre en place une structure de supervision propre à chaque projet pilote
- 2. Identifier les objectifs que doivent atteindre chaque projet pilote et quelle démarche sera appliquée pour chacun
- 3. Exécuter les projets pilotes (de manière séquentielle ou en parallèle)
 - À partir des objectifs identifiés, on fait une étude préliminaire afin d'identifier les phases principales du projet expérimental
 - Si une phase de l'expérimentation se déroule bien on passe à la suivante
 - Si une phase met en évidence des bloquants, on cherche à comprendre l'origine du bloquants et à les régler
 - Une fois les causes du bloquant ciblées, on redémarre la ou les phases ayant causé les bloquants
 - On contrôle le succès ou l'échec de chacune des phases :
 - Si on est capable de se rendre au bout de la démarche d'expérimentation avec des résultats satisfaisants on considère que le projet est un succès.
 - Si on n'arrive pas à résoudre les éléments bloquant le bon déroulement du projet, la continuité du projet pilote est à reconsidérer. Il en est de même, si on arrive à des résultats non satisfaisants.
 - Pour finir, on évalue la réussite globale des projets pilotes et on identifie ceux qui pourront faire l'objet d'une opérationnalisation.

Dans le cas de l'iA on prendra l'exemple du projet pilote de croisement de l'information client. La Figure 2.5 ci-dessous est une version appliquée à l'iA de la Figure 2.4. En effet, un comité de supervision composé de l'équipe de recherche et de gestionnaires TI a été constitué pour traiter la problématique du profilage des clients. La démarche expérimentale proposée pour ce projet est détaillée dans le chapitre 3.

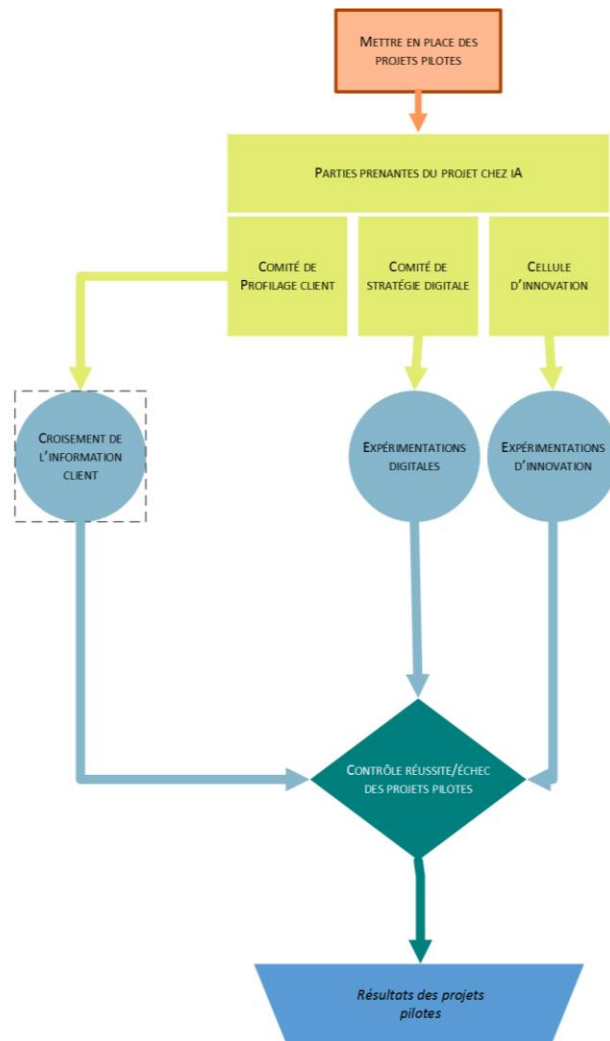


Figure 2.5 : Diagramme de la mise en place de projets pilotes - cas de l'iA

Une équipe composée de notre équipe de recherche, de consultants externes à l'iA et d'experts internes à l'iA a été formée pour réaliser ce projet sur une période de 6 mois. La démarche a été raffinée au fur et à mesure du déroulement du projet et plusieurs ajustements ont été réalisés. Ils seront discutés dans le [chapitre 3](#). Les résultats du projet pilote ont été suffisamment satisfaisants pour que les parties prenantes du projet envisagent de l'opérationnaliser dans les mois suivants. Le déroulement des autres projets pilotes de l'iA ne sera pas présenté dans le cadre de ce mémoire, pour des raisons de confidentialité.

Cette étape se décompose en trois étapes clés (qui peuvent être conduites en parallèle ou de manière couplée): mesurer l'efficacité de la préparation de la mise en œuvre (processus P1.5.1 de la Figure 2.6), mesurer la performance de la planification de la mise en œuvre (processus P1.5.2 de la Figure 2.6), et enfin évaluer la mise en œuvre de la stratégie élaborée dans la perspective d'une opérationnalisation d'un ou plusieurs projets pilotes (processus P1.5.3 de la Figure 2.6). Les étapes que l'on peut coupler, au lieu de les traiter séparément sont les étapes d'évaluation de la préparation et de la planification de la mise en œuvre (P1.5.1 et P1.5.2). Tout dépend du contexte de changement dans lequel on se trouve.

Les étapes de bilan P1.5.1 et P1.5.2 consistent à se poser les questions suivantes :

- La planification initiale a-t-elle été respectée ?
- Quels indices permettent de mesurer la performance de la planification ?
- Le niveau de préparation a-t-il été suffisant pour bâtir efficacement la vision et le plan d'action, ainsi que pour mettre en place les projets pilotes ?

Pour évaluer la mise en œuvre de la gestion stratégique du changement (processus P1.5.3 de la Figure 2.6) on cherche, dans un premier temps, à évaluer le plan d'action grâce aux résultats du projets pilotes puis à évaluer la vision bâtie en amont du plan d'action. Pour cela nous proposons de répondre aux questions suivantes :

- A-t-on réussi à atteindre les cibles identifiées pour les projets pilotes ? Ces cibles sont-elles alignées sur le plan d'action ?
- Quels indices permettent de mesurer la performance des projets pilotes ? Quelles valeurs prennent ces indices ?
- Quelles corrections du plan d'action sont à envisager compte tenu des résultats des projets pilotes ? Que nous apprend l'expérience de la coordination des projets pilotes ?
- Quels enjeux, et défis soulevés par les projets pilotes sont majeurs ? Lesquels sont mineurs ?
- Quels sont les enjeux de l'exécution du plan d'action en vue d'une opérationnalisation ?
- Le plan d'action permet-il d'atteindre l'objectif d'évolution du niveau de maturité visé au moment de la préparation à la mise en œuvre ?
- Comment la vision est-elle raffinée grâce au plan d'action ?
- Quels enjeux liés à l'exécution du plan d'action ont un impact majeur (ou mineur) pour bâtir la vision ?

L'exécution de ces trois étapes apporte à l'organisation une expérience supplémentaire dans la compréhension de son évolution par rapport au changement qu'elle souhaite accomplir. Cette expérience alimente la dernière étape de la phase : la préparation des recommandations en vue d'une éventuelle opérationnalisation du changement (P1.5.4 dans la Figure 2.6).

On peut considérer que cette phase est complétée lorsqu'on a identifié des recommandations pertinentes pour clôturer ce cycle ou pour démarrer un nouvel incrément du cycle. Ainsi, on peut soit déclencher le cycle d'opérationnalisation, soit lancer à nouveau le cycle de gestion de la stratégie de changement au cas où on voudrait continuer d'avancer dans l'évolution du degré de maturité avant d'opérationnaliser le changement.

À noter que si le bilan du cycle met en évidence un nombre trop important d'enjeux majeurs pour envisager une opérationnalisation directe, le cycle doit être relancé pour prendre en compte ces enjeux dans l'élaboration de la vision et du plan d'action.

Afin d'illustrer cette démarche, reprenons chacune des trois étapes clés et explicitons-les dans le cas de ce qui a été réalisé à l'iA :

- Quel bilan faire des phases de préparation et de planification de la mise en œuvre (processus P1.5.1 et P1.5.2) ?
 - La planification de 2 ans a globalement été respectée. Certains projets pilotes ont mené à une opérationnalisation plus vite que d'autres. C'est le cas du projet pilote de croisement de l'information
 - La préparation des données et la mise en place d'une gouvernance des données n'ont pas suffisamment fait l'objet d'une préparation et d'une planification. Elles seront à considérer avec la plus grande vigilance au moment de l'opérationnalisation
- Quelle évaluation de la mise en œuvre de la gestion stratégique du changement (processus P1.5.3) ?
 - Bilan du plan d'action mis en place grâce aux résultats du projet pilote de croisement de l'information client
 - Le potentiel de l'exploitation des données internes pour supporter les efforts de compréhension de la clientèle a été démontré
 - Les résultats attendus lors de la construction du plan d'action ont été remplis. En effet, on constate que l'on a :
 - Créé de l'engouement pour les approches analytiques chez les exécutifs seniors
 - Commencé à créer un consensus transversalement aux différents secteurs d'affaires d'iA pour partager certaines données client afin de réaliser une première segmentation de la clientèle
 - Mis l'analytique et les préoccupations du client au cœur des réflexions des projets de développement des processus majeurs de l'entreprise
 - Ces résultats bien qu'encourageants ne sont pas encore totalement des succès : l'organisation a progressé au niveau du profilage de sa clientèle, mais cela n'est pas suffisant pour être directement appliqué aux processus d'affaires clés de l'entreprise qui manipulent des données-clients.
 - Les indices permettant de mesurer la performance des projets pilotes choisis sont :
 - Le nombre d'objectifs atteints par rapport aux objectifs initialement fixés
 - Le degré de satisfaction des acteurs du projet quant aux résultats obtenus
 - Le ratio entre le budget initial et budget final de chaque projet pilote
 - Le ratio entre l'effort initial prévu et l'effort réalisé pour chaque phase du cycle
 - Les valeurs de ces indices seront présentées dans le chapitre 3 de ce niveau
 - Les corrections du plan d'action à envisager compte tenu des résultats du projet pilote touchent principalement à :
 - L'impact de la qualité des données qui a été sous-estimé lors des phases de planification et de la préparation à la mise en œuvre du présent cycle
 - La nécessité d'un engagement plus important des experts d'affaires car il s'agit d'un critère de succès important

- Bilan de la vision bâtie grâce au plan d'action mis en place
 - On constate une évolution de la maturité de l'iA à l'issu de ce cycle. Ceci est dû principalement à la prise de conscience du rôle majeur de l'exploitation de données-clients pour orienter la vision de l'évolution du modèle d'affaires vers une plus grande considération des besoins du client
 - Il est important de préciser que l'évolution réelle de cette maturité sera à réévaluer à l'issue du cycle d'opérationnalisation

2.6 Conclusion du chapitre

Rappelons que l'objectif de ce chapitre était de proposer une démarche pratique qui s'adresse à des gestionnaires, chef de projets ou dirigeants qui souhaitent intégrer et suivre les transformations du marché et de la société, quel que soit le domaine d'application de ces transformations. Par ailleurs, le domaine choisi dans le cadre de ce mémoire, et plus spécifiquement dans le cas pratique de l'Industrielle Alliance (iA), est l'exploitation des données internes pour développer une approche d'affaires orientée vers le client. Rappelons aussi que la démarche proposée visait à s'adresser à tout type d'organisation, mais plus particulièrement à celles dont le niveau de maturité est faible en regard de la gestion et de l'analyse des données internes.

Quel bilan peut-on faire des apports de cette démarche ? Tout d'abord, d'un point de vue théorique, la proposition d'une méthode de gestion du changement technologique d'un point de vue stratégique et opérationnel est une contribution importante de ce mémoire. En outre, l'approche présentée est suffisamment détaillée et appliquée pour être directement réutilisable par une organisation. De plus, elle apporte une réponse aux manques relevés dans la littérature actuelle, à savoir que beaucoup d'approches sont soit trop générales pour être applicables directement au cas pratique d'une entreprise, soit elles n'adressent pas suffisamment les besoins des entreprises moins matures.

D'un point de vue pratique, le cas de l'iA nous a permis d'illustrer chaque point de l'approche proposée et de montrer comment elle peut être concrètement appliquée. Ainsi, l'exemple de l'iA apporte une richesse d'illustration car le secteur des assurances est une industrie complexe et très diversifiée. Enfin, on a affaire à une compagnie dont le niveau initial de maturité a été identifié comme étant faible dans le cadre de l'exploitation des données client internes. Illustrer l'approche avec un tel exemple permet donc d'appuyer le fait que cette approche peut s'appliquer à des compagnies moins prêtes à effectuer des changements technologiques majeurs.

Il est à présent pertinent de soulever quelques points d'amélioration qui pourraient être envisagés pour enrichir la démarche proposée. En premier lieu, d'un point de vue théorique, le champ d'étude technique a été volontairement limité pour ne pas tomber dans une généralisation excessive. Cet aspect pourrait bénéficier d'un élargissement pour coller davantage avec les tendances les plus récentes que l'on observe sur le marché.

Ensuite, il serait intéressant de tester la démarche dans un autre domaine d'application ou dans une autre industrie à titre de comparaison. Cela permettrait de continuer à valider la généralité de la démarche proposée.

Enfin, le choix des techniques utilisées dans les projets expérimentaux pourrait être reconsidéré car le domaine de l'exploitation de données évolue très rapidement. Pour finir, il serait enrichissant d'appliquer la démarche au cas d'une entreprise plus mature, pour voir si les phases et étapes proposées s'appliquent de la même façon, ou si des ajustements de la démarche sont à considérer.

Pour conclure sur les aspects pratiques de notre recherche, les travaux réalisés pour élaborer cette démarche ayant été directement appliqués au cas de l'iA ; l'entreprise a pu bénéficier des apports que propose celle-ci. En effet, au démarrage du projet, l'iA a évalué sa maturité en exploitation de données internes comme étant faible et a identifié que son modèle d'affaires était peu orienté vers les préoccupations du client. La mise en œuvre d'une vision d'évolution de ce niveau de maturité, appuyé par un plan d'action adapté et validé partiellement par la mise en place de plusieurs projets pilotes, a permis à l'organisation de progresser dans l'identification de ses besoins en matière de compréhension de sa clientèle et des capacités en analyse des données clients qu'elle souhaite développer. L'iA a également démarré une réflexion plus approfondie sur la qualité des données qu'elle possède et sur la gouvernance nécessaire à mettre en place pour garantir le succès de cette nouvelle vision. Le chapitre suivant présente en détail la démarche expérimentale suivie dans le cadre du projet pilote de croisement de l'information client.

3 Chapitre 3 : Exemple de mise en œuvre de projets pilotes - cas du croisement de données clients

3.1 Introduction

L'objectif de ce chapitre est d'illustrer une partie de la phase de mise en œuvre de la gestion stratégique du changement (phase P1.4 de la Figure 2.1 du chapitre 2) et plus globalement de montrer l'intérêt de la démarche globale de recherche présentée dans ce mémoire à travers un exemple précis de projet pilote mené en milieu pratique. Rappelons que suite à l'évaluation de la maturité de l'iA actuelle et visée, une nouvelle vision (que l'on peut synthétiser par les termes « orientation client ») et un plan d'action ont été définis. Un volet de ce plan d'action vise à mettre place des projets pilotes pour valider la faisabilité des actions proposées et son alignement avec la vision du changement. Le projet pilote qui est présenté dans ce chapitre est celui du croisement des données-clients.

Au démarrage du projet, à la suite de diverses entrevues que nous avons réalisées avec les gestionnaires des différents secteurs d'affaires impliqués, il a été constaté que les informations sur lesquelles les secteurs d'affaires d'iA se basent pour prendre des décisions tactiques et stratégiques sont majoritairement fondées sur l'intuition. De plus, l'analyse du niveau de maturité de l'organisation avait conduit à la conclusion que l'iA ne connaît pas suffisamment sa clientèle et qu'elle s'appuie beaucoup sur son réseau de distribution pour gérer la relation avec la clientèle. Notre revue de littérature (voir chapitre 1) a aussi mis en évidence le fait que dans les entreprises plus traditionnelles, on sous-estime souvent le potentiel des données internes déjà collectées par les systèmes opérationnels/administratifs. Or ces systèmes renferment souvent de l'information pertinente sur les clients. C'est le cas particulièrement chez iA.

Dans ce chapitre, nous présentons la démarche que nous proposons pour analyser les données clients, et nous raffinons chacune des phases qui la constituent. Nous illustrons chaque étape avec le cas du projet pilote de croisement des données clients de l'iA.

3.2 Présentation de la démarche expérimentale généralisée de croisement de données clients

3.2.1 Présentation globale de la démarche

Comme l'illustre la Figure 3.1, nous proposons une démarche en 6 phases principales. Tout d'abord, l'approche consiste à **collecter** des données du périmètre d'étude de la clientèle choisie, puis à faire une première analyse de l'ensemble des données collectées pour **comprendre** les principales informations qu'elles renferment.

La compréhension de ces données permet de **cibler** l'étude qui va être réalisée. Une fois la cible définie, les données doivent être **préparées** afin de répondre aux objectifs d'exploitation ciblés. Les données préparées peuvent ensuite être **analysées** plus en profondeur selon les axes d'étude établis.

Enfin, grâce aux analyses, les données sont **interprétées**, ce qui aboutit à la production des résultats attendus au départ pour être en mesure de **faire le bilan** de l'expérimentation. Nous avons commencé à bâtir cette méthode à partir de l'expérience du cas spécifique de l'iA. Puis nous l'avons généralisée afin de l'adapter à plusieurs cas d'entreprises. À noter que l'on retrouve des démarches similaires dans la littérature, ce qui nous a conforté dans notre choix de formalisation de la démarche proposée.

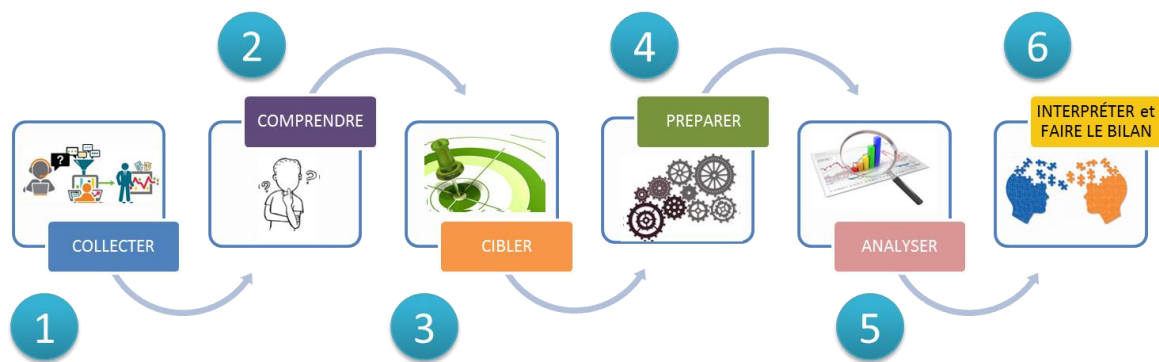


Figure 3.1 : Présentation synthétique de l'approche expérimentale proposée pour l'analyse de des données

Le lecteur pourrait se demander pourquoi la phase « cibler » se situe après les phases « collecter » et « comprendre ». C'est une phase que l'on retrouve peu dans les démarches existantes dans la littérature et rarement à ce stade du déroulement de la démarche. Nous avons fait ce choix car c'est effectivement après la collecte et la phase d'analyse préliminaire des données (la phase « comprendre ») que les intervenants de l'iA ont commencé à avoir une bonne vision de l'ensemble des données qu'ils possèdent. Cela leur a alors permis d'identifier clairement le but que le projet pilote devait atteindre afin de valider le plan d'action de la gestion stratégique du changement mis en place. C'est ce que nous appelons la « cible » (étape « cibler »).

3.2.2 Présentation schématique détaillées des étapes clés de chaque phase

Dans la Figure 3.2, nous présentons pour chaque phase de la démarche, les étapes que nous avons proposées et appliquées.

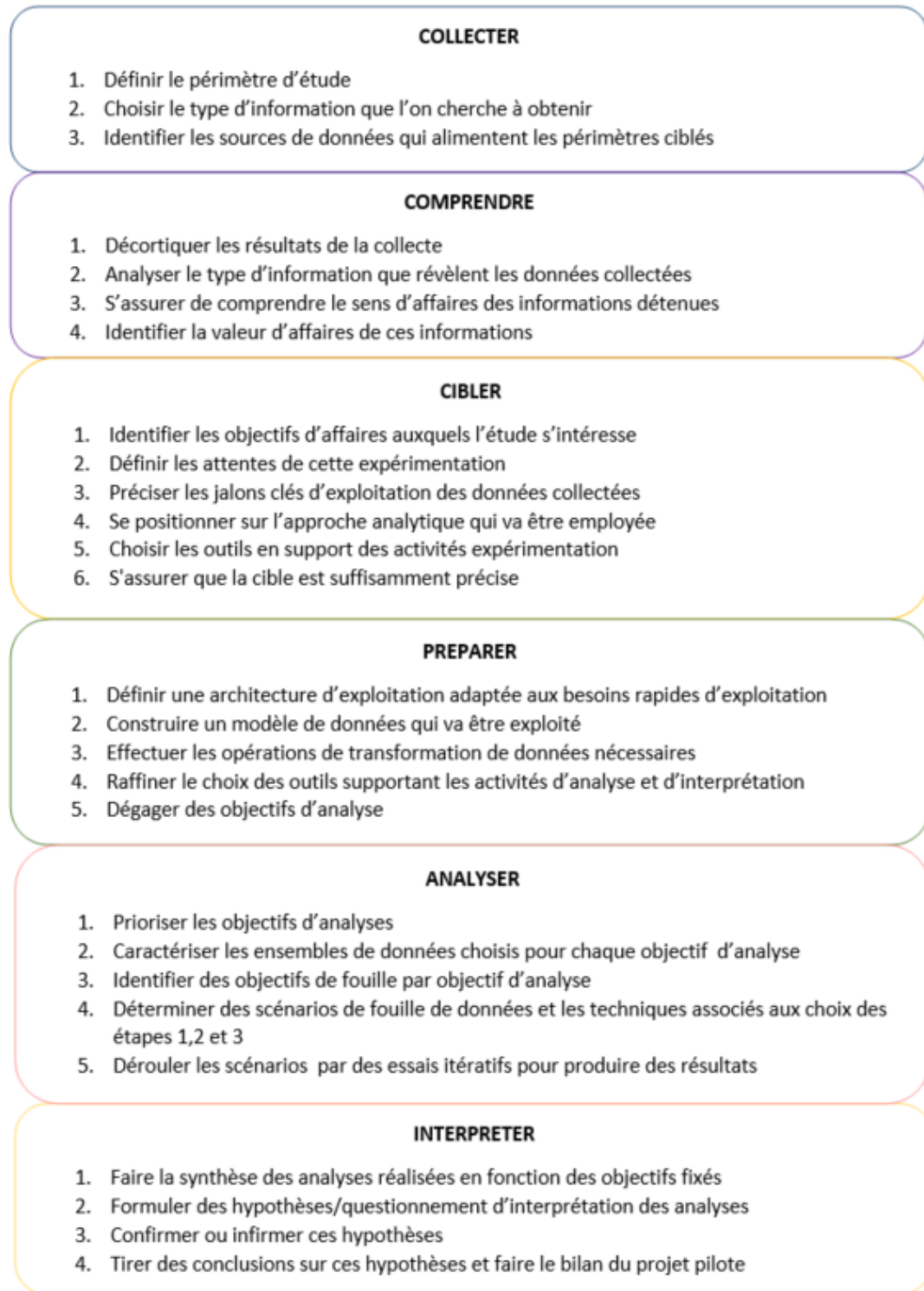


Figure 3.2 : Schéma détaillé de l'approche expérimentale proposée

Ces différentes phases vont être présentées en détails dans ce chapitre. Ainsi, les sections suivantes présentent chacune des phases, en justifiant les étapes que nous proposons et en illustrant systématiquement leur application suivant les 3 volets suivants :

- Les actions concrètes menées pour appliquer la démarche à l'iA
- Les résultats obtenus pour le cas de l'iA
- Les enjeux décelés chez iA lors de l'exécution de la phase

La démarche va être illustrée au travers d'un projet pilote expérimental que nous avons mené chez iA et qui a permis de valider, voire raffiner, la démarche proposée. Précisons que cette démarche a été pensée de manière à être suffisamment flexible pour s'adapter à différents contextes organisationnels. Ainsi, dans les phases proposées certaines étapes pouvaient être sautées ou ajustées selon les besoins de l'entreprise et du projet considéré.

3.3 Phase 1 : Collecter

Cette phase s'articule autour de la notion de périmètre de données internes au sein d'une entreprise. Il est donc pertinent d'explicitier ce terme dans ce contexte. Un périmètre est ensemble de données qui apportent des informations sur un aspect spécifique du sujet que l'on cherche à analyser. Si on prend l'exemple des données-clients, un périmètre peut-être les données démographiques des clients ou encore les données liées à leur profil financier (revenu, emploi, biens détenus, etc.).

3.3.1 La méthode proposée

1. Définir des périmètres en se ramenant aux objectifs stratégiques du projet, c'est-à-dire s'interroger sur le type d'information que l'on aimerait analyser du client. Il est important à cette étape de fixer des délimitations claires entre les périmètres.
2. Identifier les sources de données qui alimentent les périmètres ciblés. Les sources (systèmes, applications, bases de données, etc.) qui agrègent partiellement de l'information sont à privilégier surtout quand le système d'information est complexe, comme c'est le cas pour iA. Il faut également porter une attention particulière aux sources de données qui collectent de l'information à valeur ajoutée, c'est-à-dire qui récolte des informations qualitatives sur le client.
3. Pour chaque source, déterminer l'ampleur des efforts d'extraction de données nécessaires. Il faut, de plus, évaluer la nécessité réelle de récupérer l'ensemble des données traitées par la source comparativement à en extraire seulement une partie.

3.3.2 Les actions menées à l'iA

Le choix des périmètres finaux a été fait suite à des discussions avec les secteurs d'affaires en fonction de leur volonté de prendre part au projet. Afin de favoriser une élaboration par itérations de l'approche, il a été

volontairement choisi de limiter l'étendue des périmètres à 3 ou 4 secteurs d'affaires différentes sur une dizaine au total.

Nous avons réalisé une étude des modèles de données correspondant à chaque source identifiée et nous avons consulté tout document faisant référence à la nature des données détenues par la source afin de préparer l'extraction des données. Finalement, l'extraction des données a été réalisée par les experts de chaque périmètre impliqué et après que nous ayons formalisé les spécifications de l'extraction. Ces spécifications visaient notamment à préciser quels sous-ensembles de données extraire.

3.3.3 Les résultats obtenus

Dans le cas de l'iA on s'intéresse aux types d'informations suivantes sur le client : informations sociodémographiques, informations sur ses biens, profil financier et style de vie. La Figure 3.3 représente les différents périmètres choisis (notés P1, P2 et P3) ainsi que les sources de données qui y sont rattachées.

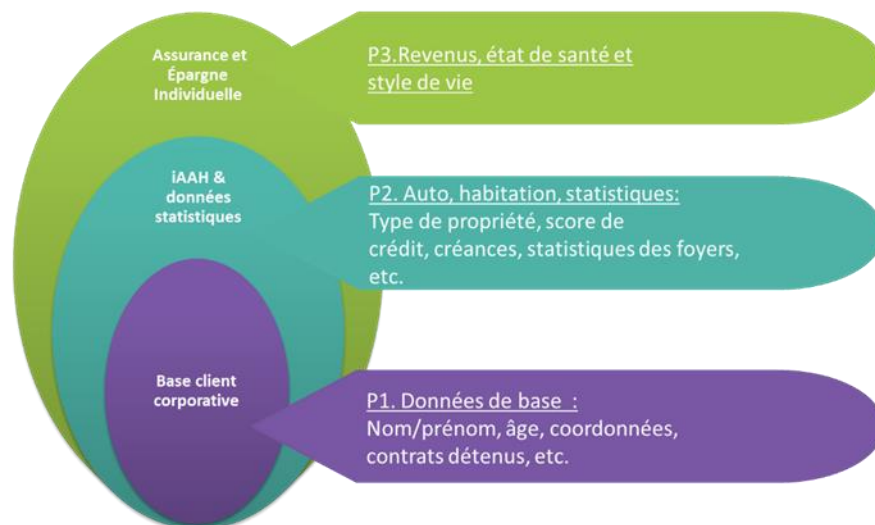


Figure 3.3: Périmètres de collecte des données d'iA choisies pour le projet expérimental

On trouve pour chaque périmètre les sources suivantes :

- P1 : Base client corporative, le système client de la compagnie qui collecte les données élémentaires des clients de plusieurs systèmes administratifs
- P2 :
 - Du côté de la ligne d'affaires Auto et Habitation (iAAH) : système client d'iAAH
 - Données statistiques externes fournies par un fournisseur spécialisé (KOREM)
- P3 :
 - Pour l'Assurance Individuelle (AssInd), il s'agit du système de tarification des produits d'assurance
 - Pour l'Épargne Individuelle (EI), la source est un entrepôt de données qui alimente une nouvelle application de gestion des ventes d'épargne individuelle.

Comme précisé plus haut, la base client corporative agrège les données issues des systèmes opérationnels de plusieurs lignes d'affaires (hormis celle des systèmes d'IAAH). Cependant, les données collectées actuellement sont très sommaires, ce qui limiterait rapidement l'étude si on s'en tient seulement à elles. Le système-client est alimenté par 23 systèmes, ce qui en fait une source de données intéressante dans un objectif de construction d'une vue consolidée du client. Un problème important lié à ce système est qu'il s'agit d'une ancienne base de données qui a été conçue à partir de règles d'alimentation qui ne sont pas uniformisées. Ainsi on ne peut pas garantir la qualité, l'intégrité et la complétude des données qui constituent le système client.

Le système client d'IAAH, quant à lui, est très récent et a été conçu pour faire de la gestion de clients au sens moderne du terme. Les données sont ainsi de meilleure qualité, intègres et rarement incomplètes, ce qui en fait une source de données plus facilement exploitable. De plus, il collecte, de par la nature des processus d'affaires en amont du système, un niveau important d'informations pertinentes sur les clients. Par exemple, la cote de propriété, le statut marital ou encore le type de propriété possédée.

En ce qui concerne les données statistiques que nous avons collectées, elles dénombrent le nombre de personnes et de foyers par code postaux selon une liste d'une centaine de critères regroupés en une quarantaine de catégories (exemple le nombre de ménages de la ville de Québec ayant un revenu supérieur à 50 000 \$). Ce dénombrement est a été réalisé pour les années 2010 et 2015.

Le système de tarification des produits d'assurance individuelle est, d'après les experts d'IA, l'une des sources de données internes collectant le plus d'information détaillées sur les clients. Mais ces données sont peu fiables et sont collectées à une étape de vente très critique : la tarification de la prime d'assurance. En effet, suite à cette étape du processus de vente, environ 20% des contrats d'assurances sont interrompus : le client est refusé par la compagnie ou bien la vente n'est pas finalisée.

Enfin, l'entrepôt de l'Épargne Individuelle (EI) donne accès à des données de qualité sur l'épargne obtenues à l'issu d'un projet de refonte des systèmes de vente de la ligne d'affaire. Cependant, l'entrepôt a été constitué pour répondre à un besoin spécifique. Ainsi certaines informations importantes sur le client n'y sont pas présentes.

3.3.4 Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA

- Problèmes pour lesquels nous avons réussi à trouver une solution :
 - Le problème majeur rencontré a été de clarifier minimalement le besoin de collecte pour orienter les experts en vue d'une extraction des données. En effet, l'efficacité de l'extraction par les experts responsables des sources de données dépendait beaucoup de notre capacité à spécifier clairement et précisément le besoin technique et fonctionnel. Or la démarche expérimentale que nous proposons (voir Figure 3.1) suggère de clarifier le besoin précis (la cible) après la phase de collecte. Cet écart entre notre vision à haut niveau de la cible de collecte et le besoin des équipes d'experts de l'iA d'avoir un niveau de détails important sur les données à extraire créa initialement des problèmes de communication importants avec les équipes d'experts. La solution qui a été retenue pour surmonter cette difficulté fut de s'en tenir à récupérer uniquement les données les plus simples à extraire, quitte à ratisser plus largement que nécessaire et à récupérer des données qu'on risque de peu utiliser en fin de compte. On a évité ainsi de trop limiter le champ de collecte et nous avons ainsi réussi à collaborer efficacement avec les experts responsables des sources de données dont nous avons besoin.
- Contraintes auxquelles il a fallu s'adapter car aucune réelle solution ne s'offrait à nous :
 - Avoir accès à des ressources disponibles (développeurs) pour partager leurs connaissances sur les sources de données choisies et effectuer l'opération d'extraction de données, alors qu'il ne s'agissait pas d'un besoin opérationnel. De ce fait, nous avons dû écarter certaines sources de données du périmètre du projet.
 - Gérer la complexité des sources de données auxquelles nous avons eu accès lorsque nous avons eu à spécifier des règles d'extraction. En effet, il nous a fallu tenir compte de : l'historique de changement non-préservé dans les sources, des formats de données difficilement interprétables ou ambigus et des spécificités structurelles des sources exploitant des technologies obsolètes. Il a fallu passer plus de temps à décortiquer la documentation disponible sur ces sources, ce qui a entraîné un retard de 2 mois par rapport à ce qui avait été estimé initialement pour l'exécution de cette phase

3.4 Phase 2 : Comprendre

3.4.1 La méthode proposée

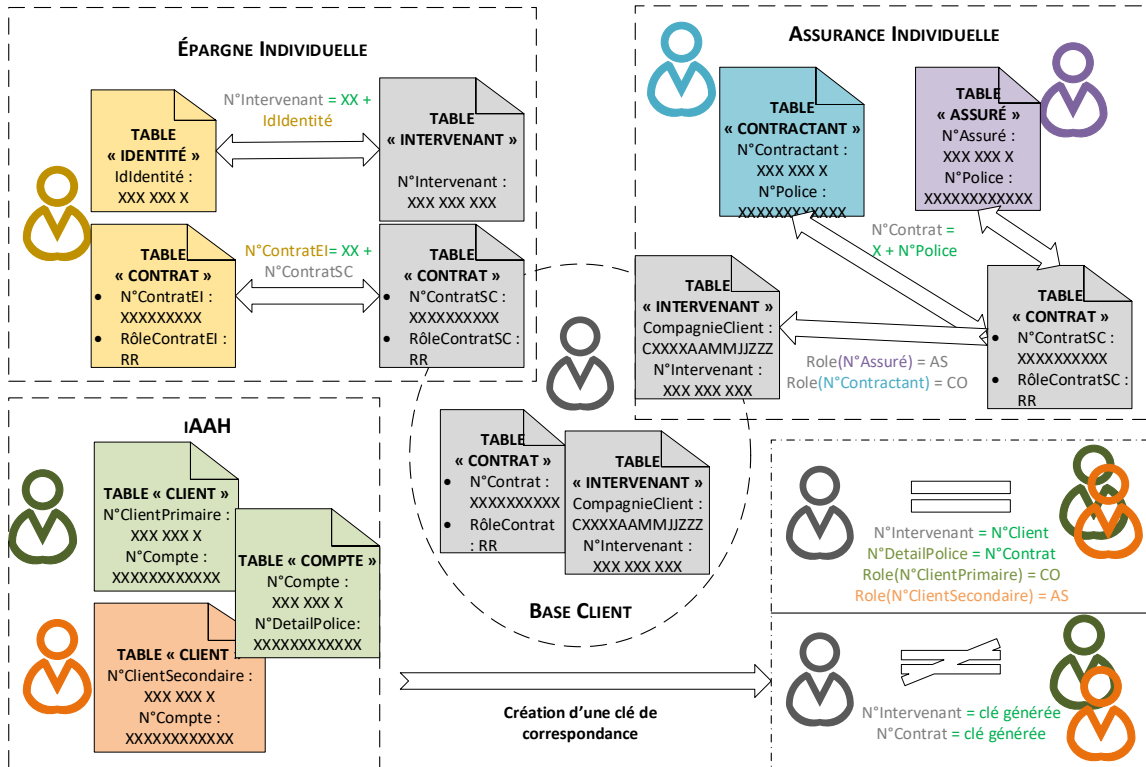
1. Analyser les résultats de la collecte de données. L'analyse implique de se poser minimalement les questions suivantes :
 - a. A-t-on assez des données suffisamment pertinentes ? Dans le cas de l'iA, oui on a eu accès à de l'information variée sur le client, mais le manque de complétude des données a affecté grandement l'exploitation de cette information
 - b. L'information est-elle complète ? Pour l'iA, on ne peut pas dire que l'information collectée était complète
 - c. Quelles données importantes supplémentaires peut-on aller chercher dans les périmètres déjà identifiés pour atteindre les objectifs ? Il existait chez iA d'autres sources intéressantes dans les périmètres choisis, mais l'effort pour les obtenir a été jugé trop élevé
 - d. Une vue parcellaire du client est-elle suffisante pour les besoins du projet ? Du fait du niveau de maturité faible d'iA, cette question fut difficile à répondre avant d'arriver à la phase d'analyse
 - e. Comment rendre compatibles ces données issues de différentes sources? Une analyse plus avancée des données est requise pour identifier la meilleure manière de concilier les sources de données
2. Analyser le type d'information que révèlent les données collectées
 - a. Rassembler les éléments de contexte nécessaires de chaque type de données pour permettre de l'interpréter (exemple, le profil financier d'un client à partir de sa cote de crédit)
 - b. Faire une synthèse claire des informations rassemblées et les classer afin de faciliter une compréhension rapide de l'étendue du portrait client que l'on est capable de faire
 - c. Commencer progressivement à identifier la valeur à tirer de ces informations et connaissances, c'est-à-dire s'interroger sur comment elles peuvent permettre à l'organisation d'atteindre les objectifs ciblés

3.4.2 Les actions menées à l'iA

Pour analyser au mieux les données collectées, nous avons consulté les experts d'affaires et de systèmes à plusieurs reprises. Par la suite, pour rendre compatibles les données maîtresses issues des systèmes, nous avons construit plusieurs modèles de correspondance autour de concepts clés que nous avons identifiés (le rôle du client, l'assuré, le contrat, etc.). Pour finir, afin de déterminer les informations ou connaissances tirées des données, l'analyse de dictionnaires de données et de documents référentiels a été nécessaire. L'intervention d'experts a également été d'une importance capitale à cette étape-ci pour valider l'interprétation des données analysées.

3.4.3 Les résultats obtenus

Tout d'abord, nous avons bâti un modèle de correspondance ayant pour but d'unifier l'information client, en tenant compte des différents rôles que celui-ci peut jouer en regard de son contrat.



e

Figure 3.4 Modèle conceptuel simplifié de croisement de l'information client

La Figure 3.4 met bien en évidence la complexité du croisement de l'information-client que nous cherchions à faire en présentant le modèle de correspondance que nous avons élaboré entre la Base Client (ellipse au centre de la figure) et les systèmes des secteurs d'affaires Épargne Individuelle, Assurance Individuelle et iAAH (rectangles en pointillés). L'objectif de ce modèle était de collecter les informations supplémentaires à celles de la Base Client que les sources de données des secteurs d'affaires détiennent au sujet du client. En effet, on retrouve pour chaque secteur des tables qui contiennent des données clés sur le client mais aussi sur les contrats qu'il détient. Ce modèle cherche à recouper ces informations avec celles de la Base Client en rendant « compatibles » ces données clés. Dans le cas des données d'iAAH, par exemple, nous avons été contraints de créer une nouvelle clé de correspondance, car les données clés étaient non compatibles avec celles de la base client et nous n'aurions pas pu croiser les informations de cette source avec celle des autres secteurs.

Nous avons aussi identifié avec l'aide d'experts de l'iA quels termes regroupaient le concept de « client » pour chaque secteur d'affaires. On retrouve notre synthèse de cette information dans le Tableau 3.1

Tableau 3.1: Le concept de "client" selon chaque secteur d'affaires d'iA

Information	Exemple	Assurance individuelle	Épargne et retraite individuelle	iAAH	Assurance collective	Épargne et retraite collective	Prêt hypothécaire
Qui considère-t-on "client"?	Assuré, contractant, créditrentier, emprunteur, etc.	Assuré et contractant, propriétaire (celui qui paye normalement 99% contractant mais il y a des cas...)	Contractant, co-contractant, créditrentier, emprunteur, co-emprunteur	Détenteur du compte client	Membre Intermédiaire Administrateur de groupe Détenteur de police Syndicat	Participant (celui qui apporte contribue financièrement) Il est identifié comme Contractant (et comme créditrentier pour la partie « assurance » du contrat)	Prospect, emprunteur, caution

3.4.4 Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA

- Problèmes pour lesquels nous avons réussi à trouver une solution
 - Trouver des plages de disponibilités suffisantes avec les différents experts du domaine (côté affaires et TI) pour organiser des ateliers d'analyse des données extraites. La solution que nous avons choisie pour résoudre ce problème a été : 1) de demander des accès supplémentaires à de la documentation existante sur les sources de données pour faciliter l'effort de compréhension du contenu informationnel celles-ci; 2) de solliciter au minimum les experts. Une autre solution que nous avons trouvée a été de travailler davantage avec des ressources humaines moins expertes (plus disponibles) mais suffisamment qualifiées pour répondre à nos questions de premier niveau au sujet des données.
- Contraintes auxquelles il a fallu s'adapter car aucune réelle solution ne s'offrait à nous :
 - Constituer des ensembles de données de qualité du fait des écarts de fiabilité, complétude et intégrité entre les sources étudiées. En effet, le système de l'iAAH et l'entrepôt de données EI sont les seuls systèmes choisis avec un niveau de qualité estimé supérieur à 80%. Ce degré de fiabilité est lié au fait qu'il s'agit de sources récentes et qui sont le fruit de projets de refonte de tout ou une partie du système d'information de ces secteurs d'affaires. Pour ce qui est des autres sources de données, à savoir la base client et le système de tarification de l'assurance individuelle, on estime environ de 50% de niveau de qualité, ce qui est faible. Il s'agit de systèmes qui n'ont pas été conçus pour répondre à des besoins de croisement d'information à des fins d'exploitation de données. Les règles à l'origine des traitements de données appliqués à ces sources étaient donc peu adaptées à nos besoins mais il s'agissait des seules sources disponibles. Nous avons donc augmenté nos estimés (budgets, temps) de la phase de « préparation » des données pour compenser ce manque de qualité des données.

Il est important de fournir quelques précisions sur les niveaux de qualité présentés ci-dessus. Les critères considérés lors de l'estimation de degré de qualité de chaque source sont les suivants : l'âge du système, les

règles de gestion de données qui le caractérisent, l'avis des experts sur le fonctionnement de ces sources et les processus d'affaires de gestion de l'information mis en place en amont de ces systèmes.

3.5 Phase 3 : Cibler

Rappelons que dans le contexte de notre projet la « cible » désigne le but que le projet pilote doit atteindre pour valider le plan d'action de la gestion stratégique du changement que l'on cherche à mettre en place.

3.5.1 La méthode proposée

1. Identifier les objectifs d'affaires auxquels l'étude s'intéresse
 - a. À quel aspect de l'activité d'affaires s'intéresse-t-on ? Dans le cas de l'iA, on veut développer une approche de vente centrée sur les besoins du client et moins sur la vente de nouveaux produits
 - b. Sur quel(s) processus en particulier concentrera-t-on l'étude ? Par exemple, pour avoir une approche de vente orientée vers les besoins de la clientèle il est important de développer une connaissance client adéquate.
 - c. À quels besoins d'affaires l'expérimentation cherche à répondre ? Formuler des objectifs précis qui synthétisent la réflexion sur les changements que l'on vise à opérer par le biais du projet pilote
2. Déterminer les attentes de cette expérimentation
 - a. À qui profitent les résultats de l'expérimentation ? Dans l'exemple de l'iA, les divisions marketing de chaque secteur d'affaires sont plus particulièrement concernées par cette démarche expérimentale. On cherche, en effet, à les doter de nouvelles approches d'analyse de l'information afin de favoriser le développement d'une meilleure connaissance de leur clientèle
 - b. Quels livrables sont attendus ? Des rapports ou tableaux de bords qui présente le portrait actuel de la clientèle consolidée et les principaux profils qui la composent a été l'un des principaux livrables attendus par le comité de supervision
3. Préciser les échéances clés à respecter pour les prochaines phases du projet pilote. Pour cela, il est essentiel de fixer des échéances réalistes par rapport à l'effort estimé et aux attentes formulées.
4. Se positionner sur l'approche analytique qui va être employée, c'est-à-dire cibler les approches disponibles sur le marché, qui peuvent être effectivement appliquées au contexte organisationnel de l'entreprise et à son niveau de maturité actuel.
5. Choisir les bons outils en support aux activités expérimentation
 - a. Pour chacune des activités exploitation, identifier les différentes solutions disponibles sur le marché qui répondraient potentiellement au besoin
 - b. Identifier les critères que l'on juge les plus importants, compte tenu des objectifs à atteindre et des activités d'expérimentation prévues
 - c. Réaliser ensuite une étude comparative afin de s'orienter vers une gamme d'outils qui respecte un maximum de critères
6. S'assurer que la cible est suffisamment précise. Pour cela, on peut se poser les questions suivantes :
 - a. La cible identifiée permet-elle d'apporter de la valeur ajoutée pour l'entreprise ? Dans l'exemple de l'iA, la cible est d'analyser les données-clients issues de différents secteurs

d'affaires afin de connaître les caractéristiques principales de la clientèle de l'entreprise et de mieux répondre à ses besoins (voir chapitre 1). Cette cible répond à une demande forte exprimée par la structure de supervision mise en place avant le lancement de notre projet d'expérimentation.

- b. La cible est-elle appuyée par la haute direction ? Cela a été le cas chez iA puisque le comité de supervision du projet était composé de plusieurs hauts gestionnaires.
- c. Comment la cible s'aligne avec les objectifs stratégiques de la compagnie ? Dans notre cas d'application la cible choisie a le potentiel de mettre en évidence des opportunités d'affaires qui sont actuellement non exploitées par l'organisation et permet de tester des approches qui sont innovantes dans l'industrie
- d. Est-on capable de mesurer le progrès, voire le succès, de la démarche ? Pour cela, il faut définir des indicateurs clés de performance (KPI) propres au contexte du projet pilote et il faut identifier les facteurs qui vont influencer les mesures réalisées.
- e. Est-ce que la cible est atteignable compte tenu de la maturité de l'organisation, de la disponibilité des ressources et des capacités de l'organisation ? Cela a été le cas pour iA, car les attentes étaient davantage axées sur la volonté de faire des apprentissages progressifs et de faire évoluer les réflexions en interne sur cette thématique.

3.5.2 Les actions menées à l'iA

En premier lieu, pour identifier les objectifs d'affaires sur lesquels l'étude porterait et définir les attentes de cette expérimentation, nous avons réalisé une trentaine d'entrevues entre novembre 2015 et juin 2016, aussi bien avec des gestionnaires des secteurs d'affaires que des TI. Ces entrevues nous ont aidés à dresser un premier portrait du fonctionnement des secteurs d'affaires et ont mis en évidence les besoins en matière d'exploitation des données internes. De plus, nous avons organisé 2 ateliers afin d'orienter les besoins vers des objectifs plus clairement formalisés. Le premier a eu lieu à la mi-janvier 2016 et le second en avril 2016.

En second lieu, en vue de préciser les échéances clés de l'expérimentation et de manière à faciliter le choix de l'approche analytique à suivre pour les activités d'expérimentation et des outils en support à ces activités, nous avons effectué des recherches approfondies dans la littérature durant les premiers mois du projet pilote, c'est-à-dire de novembre 2015 à février 2016. Nous avons en parallèle, durant cette même période, consulté des experts technologiques d'iA et organisé des rencontres avec divers fournisseurs de solutions afin de valider l'adéquation entre nos besoins et les capacités des outils qui étaient proposées.

Pour terminer, en ce qui concerne l'étape de raffinement de la cible, nous avons fixé des périodes régulières, de discussions, tout au long du projet, avec le comité de supervision du projet (tous les mois environ). Ces discussions servaient notamment à valider la compréhension commune des besoins entre les gestionnaires et notre équipe de réalisation.

3.5.3 Les résultats obtenus

Concernant les objectifs d'affaires fixés pour l'expérimentation menée à l'iA, suite aux entrevues avec les gestionnaires d'affaires qui ont accepté de participer au projet pilote, nous avons réalisé une synthèse des préoccupations soulevées. Elle est représentée par la figure ci-dessous.



Figure 3.5 : Synthèse des préoccupations soulevées par les gestionnaires de l'iA

Le schéma reprend les points clés mentionnés par les gestionnaires interrogés sur leur compréhension actuelle de leur clientèle, ainsi que de leur réseau de distribution. Les préoccupations plus spécifiques à un secteur d'affaires sont notées en couleur tandis que la préoccupation commune à toutes les lignes d'affaires est en gris à savoir : *Comment se comporte mon client ?* (en haut de la figure, en gris).

Pour tenter de répondre à cette préoccupation, nous avons proposé un volet du plan d'action stratégique auquel notre projet pilote tenterait de répondre, au moins en partie; ce qui permettrait de définir clairement la cible à atteindre pour ce projet. Nous avons synthétisé ce plan (Figure 3.6) en 2 volets principaux : renforcer le lien avec les clients actuels et faire croître les affaires en allant prospecter des clients de manière plus ciblée. Ce plan d'action est illustré par la Figure 3.6.

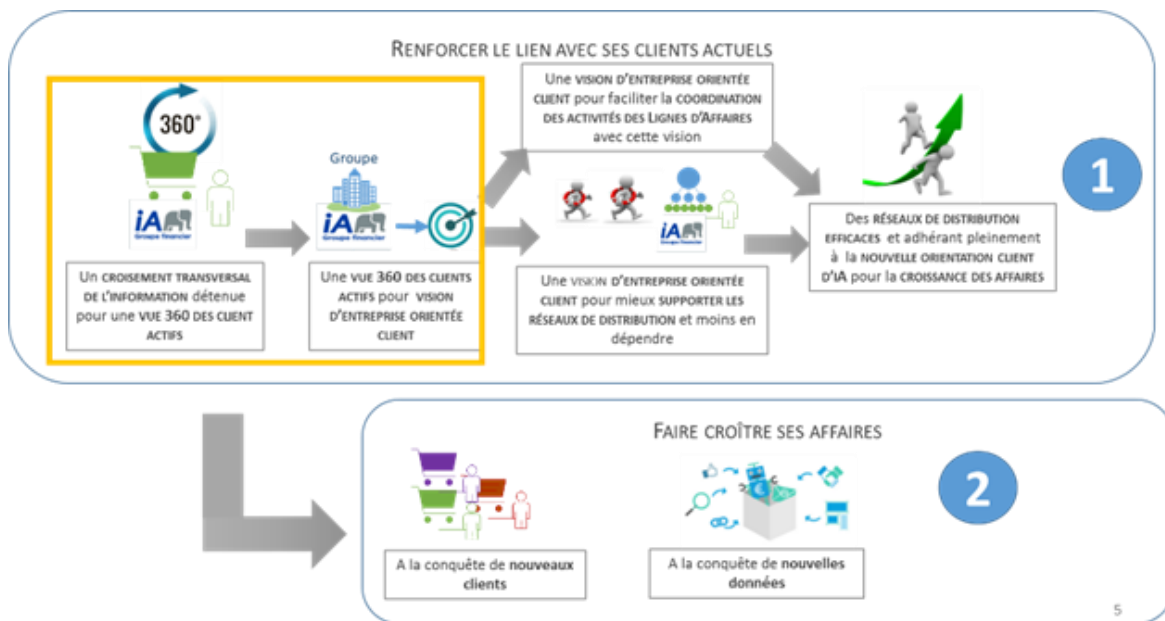


Figure 3.6: Schématisation du plan d'action en lien avec le croisement de données clients

Ces deux volets trouvent leur origine dans le développement d'une vision complète du client dite « vision 360 ». Cette vision serait bâtie à partir du croisement de l'information client détenue par les différents secteurs d'affaires d'iA. La vision 360 consiste à rassembler toutes les informations que les clients fournissent à l'entreprise à différents moments de leurs interactions avec elle afin d'obtenir un « portrait » unifié qui reflète la relation qu'ils entretiennent avec la compagnie. Pour cela, nous avons montré qu'il est nécessaire d'exploiter, pour commencer, les données internes disponibles au sein de l'entreprise.

Nous avons déterminé les orientations d'exploitation des données à partir de ce plan d'action. Elles servent d'intrants à la formalisation des objectifs de notre projet pilote, à savoir :

- Mieux cerner le potentiel concret des technologies émergentes en matière d'exploitation de données pour orienter les initiatives de changement en cours pour augmenter la connaissance client en interne
- Dégager des tendances servant à appuyer la nouvelle vision d'affaires orientée-client d'iA
- Explorer de nouvelles voies de développement de la compagnie en adéquation avec les progrès actuels de l'industrie.
- Montrer l'impact du croisement de données pour tirer profit des données en brisant les silos existant entre les divers secteurs d'affaires

À noter que, du fait que l'on se situe dans un cadre pratique expérimental et dans le cas d'une organisation au niveau de maturité relativement faible, il n'est pas surprenant de ne pas voir énoncé des objectifs ou des préoccupations plus opérationnels dans cette liste. De ce fait, les objectifs qui sont ressortis de l'exécution de cette phase ont été formulés ainsi :

- Objectif 1 : Entreprendre une exploration des données internes pour identifier les meilleures opportunités de développement des affaires supportées par des approches analytiques avancées

- Objectif 2 : Apporter la preuve que l'iA est capable de tirer de la valeur d'affaires des données qu'elle possède
- Objectif 3 : Avoir une vision globale de la segmentation de la clientèle pour mieux comprendre le client et ses attentes vis-à-vis d'iA

Compte tenu de ces objectifs, les attentes sont également moins axées sur l'opérationnel. En effet, on est en phase d'exploration, et donc on ne s'attend pas à des résultats précis, d'une excellente qualité et exploitables directement. Les attentes qui ont été fixées par la structure de supervision du projet pilote en février 2016 sont :

- Avoir une vision plus claire des capacités analytiques à développer au sein de l'organisation
- Conscientiser les secteurs d'affaires sur l'importance d'exploiter le potentiel des données détenues par la compagnie
- Dresser un premier portrait de la clientèle de la compagnie, qui soit le plus complet possible (en tenant compte des différentes informations détenues par les secteurs d'affaires)
- Dégager des tendances sur la consommation actuelle de produits et services par la clientèle d'iA
- Faire avancer la réflexion sur l'identification de ce qu'est « un client de valeur »

Étant donné les objectifs et attentes formulés ci-dessus, nous avons choisi le modèle d'exploitation de données-clients en échelle conçu par Thomas Davenport dans l'ouvrage « Analytics at work » (T. H. Davenport, Harris, et Morison 2010) comme référentiel pour identifier les principales étapes à franchir. Ce modèle est illustré par la Figure 3.7 et consiste en une succession de paliers ordonnés du bas vers le haut.



Figure 3.7: Modèle d'exploitation de données clients proposé par Thomas Davenport dans son ouvrage (T. H. Davenport, Harris, et Morison 2010)

Le premier palier de la figure consiste à détenir une base client de qualité, bien modélisée et à jour.

L'objectif étant d'avoir des données client « en ordre », c'est-à-dire bien définies, partagées, nettoyées et intégrées (« customer data in order »). Le palier suivant vise à développer un ensemble de segments de clients actuels ou ciblés (on parle de « prospects » dans le vocabulaire marketing). On cherche à déterminer quel a été le comportement passé du client et à rassembler les clients ayant des comportements similaires (étape « key targets / segments » dans Figure 3.7). La prochaine étape vise à étiqueter ces segments de clients en fonction des intérêts de la compagnie : Quel segment représente les clients les plus loyaux ? Quel segment représente les clients destructeurs de la marque ? Quel segment représente les clients qui rapportent le moins d'argent ? (étape « Differentiated action »). Une fois que l'on est capable de différencier les segments de clients, on veut être en mesure de prédire comment ils sont susceptibles de répondre à des stimuli spécifiques, qui ont été pensés pour cibler le besoin pressenti de ce segment. On est dans la prédiction de comportements clients (étape « Predictive action »). Si on est parvenu à se rendre jusque-là, on est prêt à concrétiser ces actions en créant des processus et des systèmes opérationnels qui automatisent les actions des paliers inférieurs (étape « Institutionalized action »). Enfin, la dernière étape, « Real-time optimization », prévoit une adaptation des opérations à chaque segment spécifique de clients spécifique dans un délai très court.

En ce qui concerne les méthodes d'analyse à étudier dans le cadre de cette phase, il a été choisi pour le cas particulier de l'iA de cumuler les approches suivantes :

- Approches dites Big Data pour orienter la démarche d'expérimentation dans un cadre de données volumineuses, variées et souvent de moindre qualité
- Techniques de collectes et de modélisation empruntées du BI (« Business Intelligence ») pour croiser les données
- Techniques d'apprentissage automatique pour approfondir les analyses
- Méthodes de visualisation de données traditionnelles et émergentes
- Approche de fouille de données (« Data Mining ») et de découverte de connaissances (« Knowledge Discovery »), également pour approfondir les analyses

En ce qui concerne les outils logiciels de support aux analyses, nous avons décidé de nous orienter vers :

- Une Base de Données (BD SQL Server) pour le stockage de l'information,
- Le système SQL Server Integration Services (SSIS) pour la gestion des flux d'ETL (voir Figure 3.8)
- Le logiciel Power Bi Desktop en tant qu'outil de visualisation (Figure 3.9)

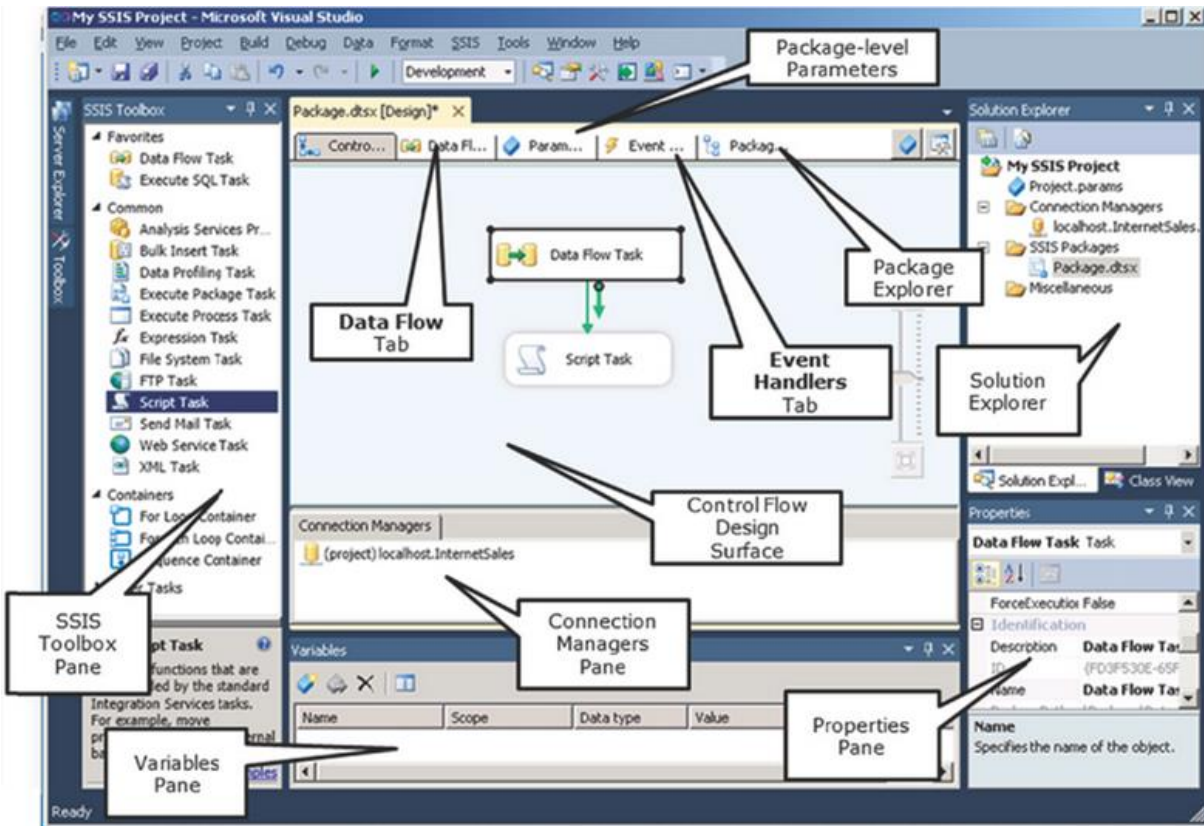


Figure 3.8 : Interface SSIS

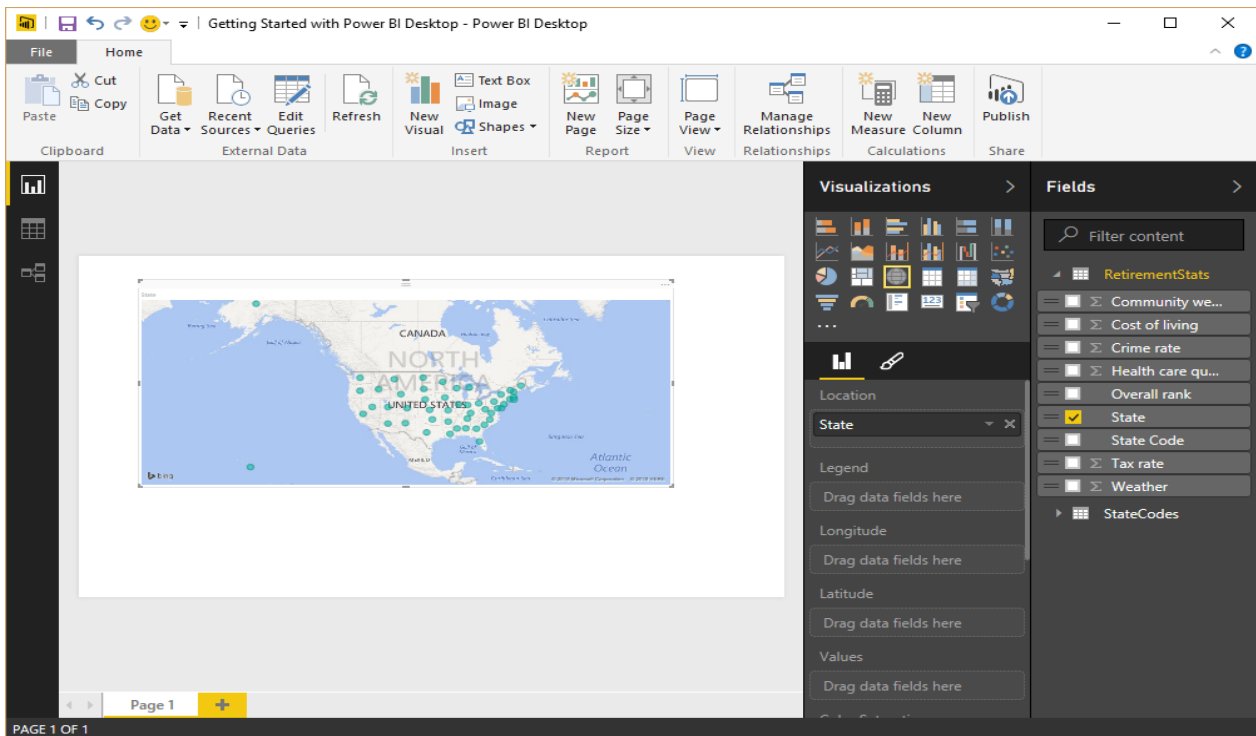


Figure 3.9: Interface Power Bi Desktop

Enfin les indicateurs clés de performance (KPI) identifiés pour suivre l'évolution de l'atteinte de la cible sont les suivants :

- Le nombre d'objectifs atteints par rapport aux objectifs fixés précédemment
- Le degré de satisfaction des acteurs impliqués dans le projet quant aux résultats prévus
- Le ratio entre le budget final et le budget initial
- Le ratio entre l'effort réalisé pour chaque phase et l'effort initial estimé

Les valeurs de ces KPI seront présentées dans la section de conclusion (3.10) de ce chapitre.

3.5.4 Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA

- Problèmes pour lesquels nous avons réussi à trouver une solution :
 - Les gestionnaires des secteurs d'affaires sont en général très occupés. Un problème récurrent pour cette phase a été d'avoir suffisamment accès à ces personnes pour être en mesure de raffiner le plus rapidement possible les objectifs d'expérimentation (la cible). La solution pour laquelle nous avons opté a été de fixer des rencontres aux deux semaines avec quelques gestionnaires parmi les plus actifs dans le projet et de mettre en place des ateliers de plusieurs jours (aux 2 mois) réunissant, cette fois, tous les gestionnaires parties prenantes du projet
 - Définir des KPI qui cadrent avec un contexte d'exploration et non pas opérationnel. En effet, le concept de KPI n'est pas totalement adapté pour mesurer des performances liées à des activités d'expérimentation. Nous avons décidé de créer des KPI qui mesurent non pas les performances des analyses en tant que telles, mais plutôt des indicateurs mesurant l'efficacité de la gestion du projet
- Contraintes auxquelles il a fallu s'adapter car aucune réelle solution ne s'offrait à nous :
 - Difficultés pour valider rapidement la faisabilité technique des objectifs fixés (infrastructure, outils, données) avec les experts TI. Ce fut une contrainte forte car leur compréhension des enjeux du projet était plus restreinte.
 - Avoir le recul nécessaire pour évaluer la précision des objectifs fixés. Il n'y avait pas de solution garantissant que nous avons atteint le niveau de recul suffisant pour passer aux phases suivantes avec des garanties de succès. Il a donc fallu redoubler de vigilance au moment de valider les objectifs et les résultats attendus.

3.6 Phase 4 : Préparer

3.6.1 La méthode

1. Définir une architecture d'exploitation adaptée aux besoins d'évolution rapide caractéristique d'une expérimentation. Pour cela, il est proposé de :
 - a. Identifier les ressources minimales nécessaires pour mener à bien le projet
 - b. Planifier la réalisation des étapes préalables à l'exploitation en termes :
 - i. D'infrastructures (BD, serveurs, puissance de calcul des ressources matérielles requises)
 - ii. D'intégration des données (flux ETL)
 - iii. De modélisation des données à analyser
 - iv. De gestion de la qualité des données
 - v. De production de la documentation minimalement nécessaire
2. Identifier les niveaux d'analyses préliminaires
 - a. Quel découpage des données collectées permet de remplir les objectifs fixés à l'étape « cibler » ?
 - b. Quelles réponses cherchent-on à obtenir pour atteindre les objectifs et attentes du projet pilotes grâce à ce découpage des données ?
3. Construire un modèle de données qui pourra être exploité facilement et qui sera suffisamment flexible. Des questions importantes à poser pour élaborer un tel modèle sont les suivantes :
 - a. Quel type de modèle est le plus adapté aux besoins d'exploitation ? Dans le cadre du projet à l'iA le choix se concentrait autour des 3 options suivantes :
 - i. La constitution d'un cube OLAP (voir chapitre 1)
 - ii. La construction d'un modèle multidimensionnel (voir chapitre 1)
 - iii. L'exploitation d'une table analytique (voir chapitre 1)
 - b. Comment appliquer le modèle théorique construit aux données que l'on possède ?
 - c. Quelles précautions (choix de clé des tables, choix des dimensions, logique des relations entre les différentes tables, etc.) sont à prendre au moment de la modélisation ?
 - d. Quel niveau de complexité du modèle reste acceptable pour son implantation dans les délais requis ?
4. Effectuer les opérations de transformation de données nécessaires pour alimenter le modèle, c'est-à-dire :
 - a. Simplifier le domaine de valeur des données les plus critiques aux besoins d'analyse. Par exemple, s'il existe plusieurs codes de sexe pour un même genre on réduit le domaine de valeur à 3 codes : un code pour le genre féminin, un code pour le genre masculin et un code « indéterminé » lorsque le sexe n'est pas renseigné.
 - b. Valider les formats de données que l'on utilise et s'assurer de leur conformité avec les besoins d'analyse (par exemple que les champs faisant référence à des dates ne sont pas au format numérique mais bien au format date)
 - c. Normaliser les données provenant de sources hétérogènes pour favoriser leur consolidation en informations communes et pour uniformiser les domaines de valeurs
 - d. Nettoyer les données après normalisation
 - i. Pour limiter les possibilités d'erreurs d'analyse
 - ii. Pour réduire les efforts d'analyse
 - iii. Pour rehausser le niveau de qualité des données

5. Raffiner le choix des outils qui supportent les activités d'exploitation. Bien que ce choix ait été amorcé lors de la phase de ciblage, c'est un choix qui a été fait a priori car on était à un stade où on ne comprenait pas encore suffisamment les caractéristiques des données à exploiter. Il est donc important de se poser les questions suivantes pour pouvoir réévaluer le choix des outils :
 - i. L'outil permet-il de faire les analyses prévues, compte tenu du modèle construit et des transformations de données réalisées ?
 - ii. L'outil permet-il de gérer facilement un grand ensemble de données ?
 - iii. L'outil gère-t-il la visualisation de données de manière suffisamment intuitive pour les utilisateurs ?
 - iv. Le paramétrage de chacun des outils est-il suffisamment facile à maîtriser en un temps réduit ?

Répondre à ces questions peut amener soit, à changer d'outil, soit à compléter la gamme d'outils en mettant en évidence des besoins non identifiés de la phase d'identification de la cible, ou à l'inverse de confirmer certains choix d'outils faits à ce moment-là.

3.6.2 Les actions menées à l'iA

Afin de déterminer l'architecture d'exploitation la plus adaptée aux besoins, en plus de nos recherches bibliographiques, nous avons sollicité les acteurs TI du projet pour obtenir les ressources et budgets nécessaires à la mise en place d'une équipe de réalisation. Cette équipe d'une dizaine de personnes était composée d'administrateurs de base de données, de chargés de projet, de développeurs et d'analystes (principalement notre équipe de recherche). Pour construire le modèle de données, nous en avons testé rapidement entre décembre et février 2016 plusieurs modèles différents pour parvenir à faire le meilleur choix de modèle. Cette démarche a été faite conjointement avec des experts en architecture de données disponibles à l'iA. Les opérations de transformation et de chargement des données à partir du modèle ont été réalisées en mars 2016 par une équipe de développeurs internes. De plus, l'identification des enjeux de qualité que nous avons réalisée lors de la phase de définition de la cible a été complétée par un profilage des données qui s'est également déroulé en mars 2016. Enfin, pour aider à valider le choix des outils, nous avons recontacté les fournisseurs d'outils de développement et d'analyse avec des questions plus spécifiques sur les capacités de leurs outils.

3.6.3 Les résultats obtenus

La Figure 3.10 présente de manière schématique l'architecture qui a été retenue pour ce projet pilote. Cette architecture est basée sur un modèle dit en « Spokes and Hub » (« rayons et concentrateurs ») (Godin et Desrosiers 2011).

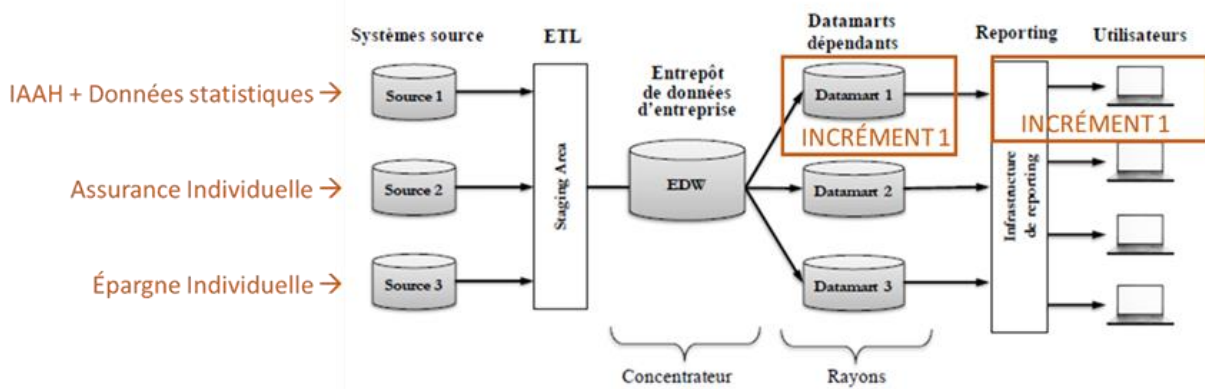


Figure 3.10: Architecture d'exploitation de données choisie pour le projet expérimental

Il s'agit d'une approche top-down favorisant l'intégration et la consolidation complète des données de l'entreprise. On part des sources de données auxquelles on a accès (en rouge dans la partie gauche de la figure). Après chargement, l'entrepôt (hub) contient toutes les données atomiques, c'est-à-dire au niveau de détail le plus fin, et provenant des différentes sources des périmètres ciblés. Les magasins de données (« spokes ») contiennent les données agrégées suivant le modèle dimensionnel choisi. Chaque magasin de données répond à un besoin d'analyse différent. Grâce à ces magasins de données, on est capable d'extraire des ensembles de données qui servent à l'analyse et à la production des résultats attendus (reporting). Les utilisateurs dans notre cas sont les gestionnaires d'affaires qui interviennent dans le cadre du projet pilote. Le modèle choisi pour le magasin de données dans ce cadre expérimental (incrément 1 de la Figure 3.10) est une table analytique qui s'intéresse aux niveaux d'analyse suivants :

- Données démographiques du client
- Relations contractuelles entre le client et iA
- Liens entre les individus ayant un même contrat

Le modèle de la table analytique du Tableau 3.2 regroupe 24 types de données « cibles » en lien avec ces niveaux concernant les axes d'analyses suivants :

- La segmentation sociodémographique de la clientèle actuelle
- La segmentation de la clientèle par type de produits détenus
- La construction du portrait global du bassin de clients actuel en tenant compte de sa géolocalisation
- L'identification de la rentabilité des produits en fonction de la clientèle qui achète les produits
- Le calcul de la valeur du client
- La validation des intuitions d'affaires grâce aux données à disposition

Tableau 3.2: Modèle analytique exploité pour l'expérimentation du projet croisement de l'information

Cible			Source			Cible			Source		
Nom du champs	Système Source	Nom de la colonne	Nom du champs	Système Source	Nom de la colonne	Nom du champs	Système Source	Nom de la colonne	Nom du champs	Système Source	Nom de la colonne
AGE_ACTUEL	IAAH	**Valeur calculée**	NOM_COMPAGNIE_EMMETRICE	AssInd	INFCO_CIE_NAME	SOURCE_CONTRAT	IAAH	**Valeur calculée**	STATUT_MARITAL	IAAH	MARITALSTATUSCD
	AssInd	**Valeur calculée**		EI	N/A		AssInd	**Valeur calculée**		EI	N/A
	EI	**Valeur calculée**		IAAH	CARRIERCD		EI	**Valeur calculée**		EI	N/A
AGE_SIGNATURE	IAAH	**Valeur calculée**	NUMERO_CIVIQUE	AssInd	NUMERO_CIVIQUE	TYPE_PRODUIT	IAAH	**Valeur calculée**	TYPE_PRODUIT	IAAH	LOB
	AssInd	**Valeur calculée**		EI	NUMERO_CIVIQUE		AssInd	**Valeur calculée**		EI	N/A
	EI	**Valeur calculée**		IAAH	STREETNUMBER		EI	**Valeur calculée**		EI	N/A
DATE_ADHESION	AssInd	APPSIGN_DT_AC	PRENOM	IAAH	FIRSTNAME	PROFESSION	AssInd	PRENOM	RAISON_ACHAT	AssInd	APPLRSN_TXT
	EI	DTAdhesion		AssInd	PRENOM		EI	N/A		AssInd	PRORSN_TXT_IN
	IAAH	INCEPTIONDATE		EI	PRENOM		EI	N/A		EI	N/A
DATE_DE_NAISSANCE	IAAH	dateOFBIRTH	PROFESSION	IAAH	OCCUPATIONCD	RECOMMANDE_IA	IAAH	N/A	RELATION_ENTRE_COAS_BENEF	IAAH	N/A
	AssInd	Date_de_naissance		AssInd	N/A		AssInd	N/A		AssInd	BENERLTN-TXT
	EI	Date_de_naissance		EI	N/A		EI	**Valeur calculée**		EI	**Valeur calculée**
EMPLOI	AssInd	APPLOCCUP_TXT	RAISON_ACHAT	AssInd	APPLRSN_TXT	RELATION_ENTRE_AS_CO	IAAH	N/A	REPLACEMENT_COMPAGNIE_PAR_IA	AssInd	APPLRLTN_TXT
	EI	N/A		AssInd	PRORSN_TXT_IN		EI	**Valeur calculée**		AssInd	INFREPLM_CD
	IAAH	OCCUPATIONDESC		EI	N/A		EI	N/A		EI	N/A
FSA	IAAH	**Valeur calculée**	RECOMMANDE_IA	IAAH	N/A	ROLE_CONTRAT	IAAH	**Valeur calculée**	SEXE	IAAH	GENDERCD
	AssInd	**Valeur calculée**		AssInd	N/A		AssInd	**Valeur calculée**		AssInd	**Valeur calculée**
	EI	**Valeur calculée**		EI	N/A		EI	CdRole		EI	SEXE
IDENTIFIANT_CLIENT	IAAH	CUSTOMERACCOUNT_ID	RELATION_ENTRE_COAS_BENEF	IAAH	N/A	SEXE	IAAH	N/A	SEXE	IAAH	SEXE
	AssInd	Compagnie_client		AssInd	N/A		AssInd	**Valeur calculée**		AssInd	**Valeur calculée**
	EI	Compagnie_client		EI	**Valeur calculée**		EI	**Valeur calculée**		EI	SEXE
LANGUE	AssInd	LANGUE	RELATION_ENTRE_AS_CO	AssInd	APPLRLTN_TXT	SEXE	AssInd	**Valeur calculée**	SEXE	AssInd	**Valeur calculée**
	EI	LANGUE		EI	**Valeur calculée**		EI	**Valeur calculée**		EI	**Valeur calculée**
	IAAH	LANGUECD		IAAH	TYPEOFRELATIONSHIP		IAAH	**Valeur calculée**		IAAH	**Valeur calculée**
LDU	AssInd	CODE_POSTAL	REPLACEMENT_COMPAGNIE_PAR_IA	AssInd	INFREPLM_CD	SEXE	AssInd	**Valeur calculée**	SEXE	AssInd	**Valeur calculée**
	EI	CODE_POSTAL		EI	N/A		EI	**Valeur calculée**		EI	**Valeur calculée**
	IAAH	POSTAL_CODE		IAAH	**Valeur calculée**		IAAH	**Valeur calculée**		IAAH	**Valeur calculée**
NO_CONTRAT	AssInd	App_No_Police	ROLE_CONTRAT	IAAH	**Valeur calculée**	SEXE	IAAH	**Valeur calculée**	SEXE	IAAH	**Valeur calculée**
	AssInd	App_No_Police		AssInd	**Valeur calculée**		AssInd	**Valeur calculée**		AssInd	**Valeur calculée**
	EI	NoContrat		EI	CdRole		EI	CdRole		EI	CdRole
NOM	IAAH	POLICYDETAIL_ID	SEXE	IAAH	GENDERCD	SEXE	IAAH	GENDERCD	SEXE	IAAH	GENDERCD
	AssInd	NOM		AssInd	SEXE		AssInd	SEXE		AssInd	SEXE
	EI	NOM		EI	SEXE		EI	SEXE		EI	SEXE
	IAAH	LASTNAME									

La colonne « cible » identifie le type d'information agrégée qui est exploité dans la table analytique, tandis que la colonne « source » identifie le champ du système source qui alimente le modèle cible.

Les opérations de transformation de données qui ont été utilisées en vue de charger le modèle analytique ont été : 1) le pré-calcul de valeurs dérivées, par exemple l'âge à partir de la date de naissance; 2) la suppression de doublons; 3) la création d'identifiants clients uniques; 4) la conversion des champs de type texte en date lorsque nécessaire; 5) la suppression des occurrences de clients hors périmètre (exemple les clients représentant une personne morale).

Dans l'étape de raffinement du choix des outils, nous avons finalement retenu la solution Rapidminer pour exploiter des algorithmes d'apprentissage automatique. Un exemple d'interface est présenté à la Figure 3.11. En effet, au moment de la phase précédente d'identification de la cible (voir section 3.5), nous n'avions pas décidé définitivement d'intégrer à l'expérimentation un volet analytique exploitant des techniques d'apprentissage automatique. C'est dans l'étape courante qu'il a été jugé nécessaire de se doter d'un tel outil.

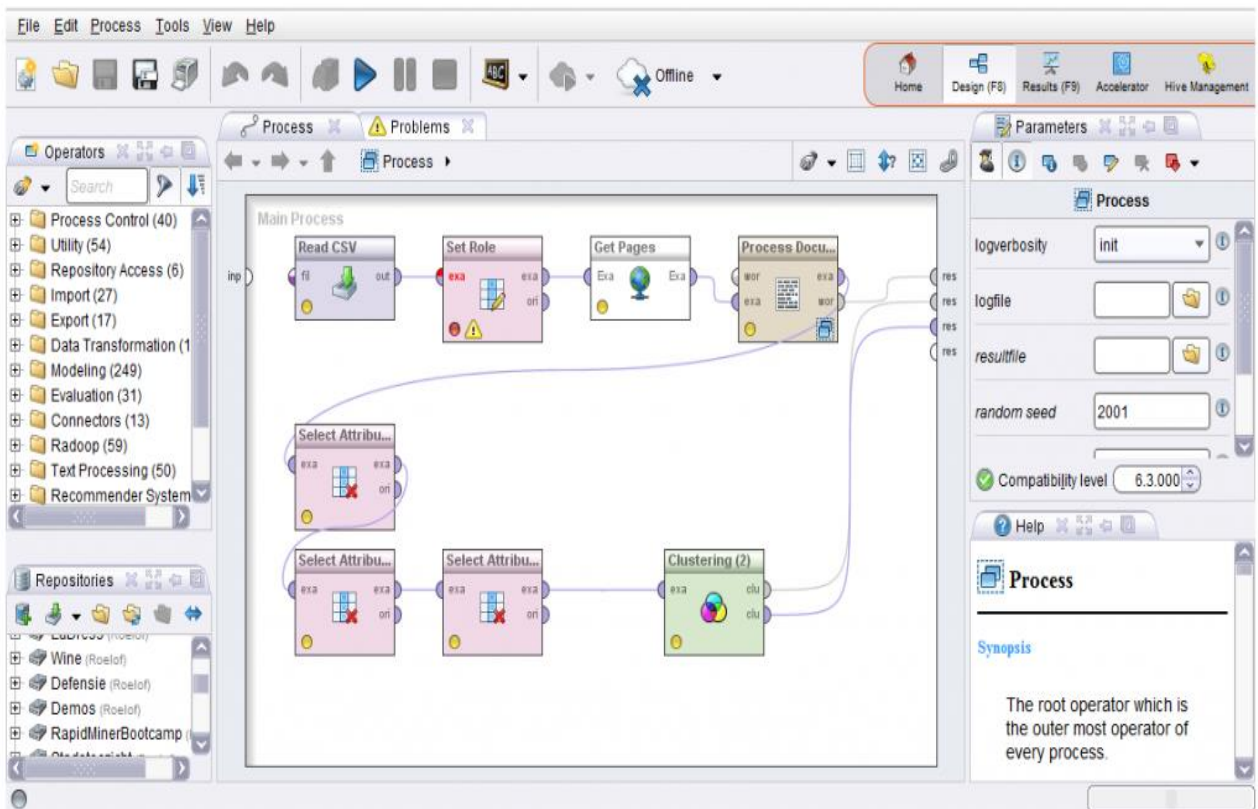


Figure 3.11: Interface de Rapidminer

3.6.4 Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA

- Problèmes pour lesquels nous avons réussi à trouver une solution :
 - Limiter les erreurs de modélisation car elles peuvent avoir de fortes répercussions sur les phases suivantes du projet. Pour remédier à ce problème, nous avons effectué plusieurs validations du modèle auprès de différents experts à l'iA. En effet, l'intervention de spécialistes des secteurs d'affaires lors de la conception et de la validation du modèle est requise pour s'assurer de créer un modèle en adéquation avec les besoins identifiés.
 - Il y a un risque non négligeable qu'on s'aperçoive plus tard dans l'analyse que l'outil s'avère moins propice à l'exploitation prévue, malgré la réévaluation des choix d'outils faite à cette étape. Ceci peut s'expliquer en partie par un manque de transparence des fournisseurs de solution au sujet des limites des outils qu'ils proposent. Il n'y a pas de solution parfaite pour éliminer ce problème. Nous avons, cependant tenté de limiter ce risque en testant le plus tôt possible les outils avec des jeux de données de test que nous avons constitués pour nous assurer que les outils répondaient bien aux critères attendus.

- Contraintes auxquelles il a fallu s'adapter car aucune réelle solution ne s'offrait à nous :
 - Stabiliser rapidement le modèle. Soit on veut trop le raffiner, soit on veut trop souvent changer de direction. Il faut trouver un équilibre entre ces deux aspects, le plus efficacement possible.
 - Pendant l'étape de transformation des données, il a été nécessaire de concentrer les efforts d'ETL uniquement sur l'essentiel, sans nécessairement viser la perfection car nous risquions de rallonger considérablement la durée du projet sans garantie de succès à la fin. La vitesse à laquelle les décisions sont prises à cette étape de préparation peut mettre en péril l'atteinte des échéances fixées si on met trop d'efforts à certaines étapes, notamment lorsque l'on souhaite rehausser la qualité des données.

3.7 Phase 5 : Analyser

3.7.1 La méthode

Le processus d'analyse que nous avons appliqué pour cette phase a été construit au fur et à mesure du déroulement du projet pilote à l'iA et de notre compréhension des bonnes pratiques d'analyse à mettre en place. Nous avons ensuite raffiné cette phase suite au bilan que nous avons fait de l'expérience acquise grâce à son application au cas spécifique de l'iA. De plus, tout comme pour la démarche expérimentale que nous proposons dans ce chapitre, ce processus vise à être suffisamment général pour s'adapter au cas d'autres projets pilotes, à d'autres contextes d'exploitation de données ou encore à d'autres entreprises. La Figure 3.12 illustre les grandes étapes du processus que nous proposons.

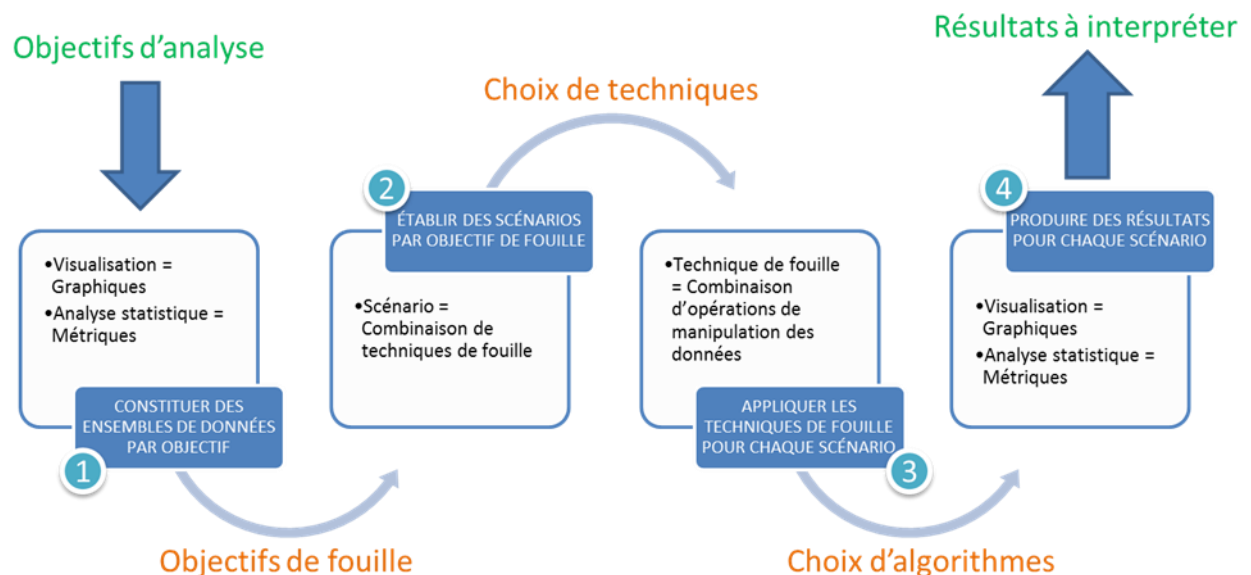


Figure 3.12 : Démarche d'analyse de données

L'étape 1 du processus d'analyse illustré par la Figure 3.12 a comme intrants les objectifs d'analyse qui ont été définis lors de la phase « Préparer » (section 3.6).

Grâce aux objectifs d'analyse identifiés en amont et à la production d'un modèle de données en adéquation avec ces objectifs d'analyse, on constitue des sous-ensembles de données (étape 1 dans la Figure 3.12) qui vont permettre de répondre à un ou plusieurs objectifs d'analyse spécifiques. Ensuite, chaque ensemble est visualisé globalement et des métriques statistiques sont produites pour mieux orienter la prochaine étape (étape 2 dans la Figure 3.12) : l'approfondissement de l'analyse. Pour cela, on définit des objectifs de fouille qui sont des objectifs qui précisent les objectifs d'analyse dans le contexte des ensembles de données choisis. Ces objectifs de fouille permettent de mettre au point un protocole de fouille, appelés scénarios de fouille qui prévoit l'utilisation d'une ou plusieurs des techniques d'analyse présentées dans le chapitre 1. Une fois ces scénarios conçus, on les déroule en utilisant les algorithmes propres à chaque technique (étape 3 dans la Figure 3.12). Les résultats obtenus sont visualisés et analysés (étape 4 dans la Figure 3.12). Ces résultats produits vont alimenter la phase 6 « interpréter » du projet pilote (voir Figure 3.1)

Afin de mieux comprendre comment ces différentes étapes s'imbriquent, nous illustrons à la Figure 3.13 la façon dont un axe ou objectif d'analyse spécifique peut être découpé en ensembles de données, objectifs de fouille, et scénarios. En effet, pour chaque objectif on identifie des ensembles de données tirés du modèle de données qui vont permettre de répondre au besoin. Pour chacun de ces ensembles de données on détermine ensuite, à un ou plusieurs objectifs de fouille qui vont permettre de répondre à l'objectif d'analyse (objectif d'analyse 1 de la Figure 3.13). Par la suite chaque objectif de fouille va donner lieu à la mise en place d'un ou plusieurs scénarios d'exploration qui seront déroulés afin d'obtenir un résultat répondant à l'objectif de fouille. C'est la combinaison des résultats obtenus à l'issue de chaque scénario pour tous les objectifs de fouilles et tous les ensembles de données qui s'y rattachent qui permettent de conclure sur l'atteinte de l'objectif d'analyse .

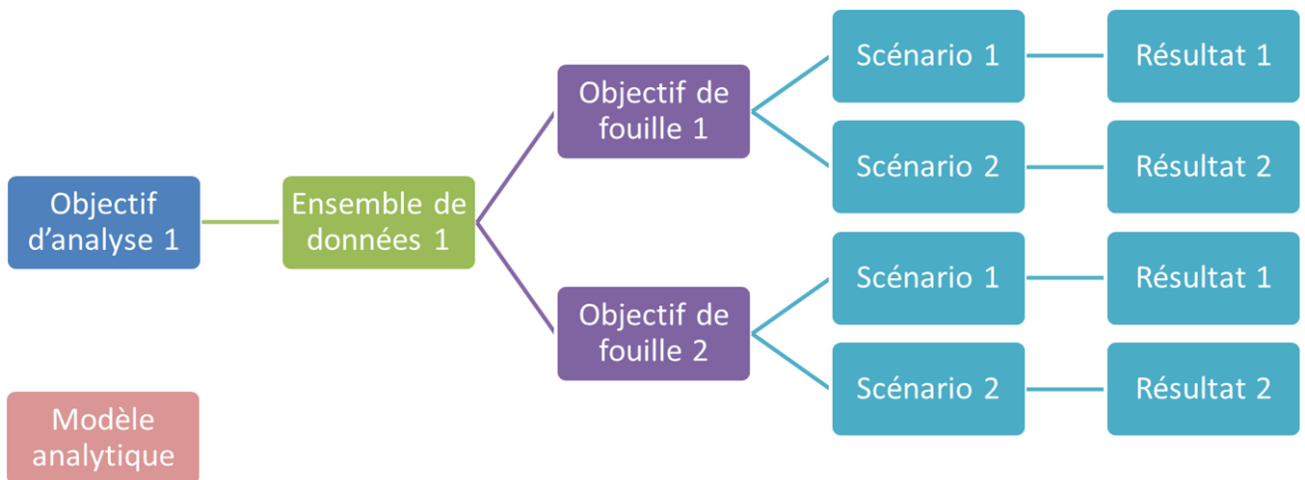


Figure 3.13: Représentation du découpage d'un objectif d'analyse

3.7.2 Les actions menées à l'iA

Pour prioriser les objectifs d'analyse et orienter les objectifs de fouille, Nous avons organisé 2 ateliers avec les analystes d'affaires impliqués dans le projet entre avril et juin 2016. De plus, nous avons suivi des formations sur les outils dont nous nous sommes servis pour mener à bien nos activités d'analyse et de fouille des données. Ces formations ont eu lieu à l'iA en faisant appel à des intervenants externes. Nous avons complété notre formation en consultant les documents disponibles en ligne sur ces outils.

En ce qui concerne les activités d'analyse, elles ont servi notamment à profiler les données. Dans le cas deux types d'analyses ont été réalisés : des analyses dites uni-variables et des analyses dites multi-variables. Parmi les analyses uni-variables on compte le calcul de tendance centrale (moyenne, valeur dominante, médiane) et les analyses de mesures de dispersion telles que le calcul de l'amplitude des données, le calcul de quartile ou encore le calcul de l'écart-type. Avec les analyses multi-variables on s'intéresse davantage à obtenir un profil central représentatif de l'ensemble des données considérées (ex : le profil client) en analysant plusieurs types de données en même temps et en cherchant à faire des corrélations entre les données. Par exemple, on peut mettre en évidence l'influence de l'âge sur l'achat de certains produits de l'entreprise.

Les activités de visualisation sont un autre moyen efficace de comprendre le comportement des données et de valider des hypothèses. Les moyens de visualisation utilisés dans notre cas ont consisté principalement à exploiter des graphiques et des visualisations de cartes grâce à l'outil Power Bi (voir l'exemple de l'interface de l'outil à la Figure 3.9 à la section Phase 3 : Cibler). Ces activités d'analyse et de visualisation se sont déroulées sur plusieurs semaines de mars à avril 2016.

Enfin, les techniques de fouille ont visé principalement à segmenter la clientèle actuelle d'iA et à identifier des pistes de calcul de la valeur client. Les principaux algorithmes utilisés ont été :

- Les algorithmes de clustering dits « Kmean » permettant de classer les données selon leur « distances » moyennes réciproques
- Les algorithmes de classification qui sont capables classer « naturellement » des points en catégories hiérarchiques. Cette classification est obtenue en répartissant progressivement les données en échantillons constitués de données homogènes et bien différenciés entre eux
- Les algorithmes de clustering dits “Kmedoid” qui sont basés sur le calcul de la valeur dominante de chaque type de données et non sur leur moyenne.
- Les algorithmes de régression linéaire servant à déceler le niveau de corrélation de deux variables

3.7.3 Les résultats obtenus

Les objectifs d'analyse qui ont été priorisés sont les suivants :

- Objectif d'analyse 1 : dresser un portrait global du bassin de clients avec la composante de géolocalisation
 - Segmentation sociodémographique de la clientèle actuelle
 - Segmentation de la clientèle par types de produits détenus
- Objectif d'analyse 2 : Valider des intuitions d'affaires grâce aux données comme par exemple :
 - Quelles ventes sont manquées par la non identification /exploitation des liens conjoints/parents-enfants ?

Le processus d'analyse que nous avons suivi va être présenté plus en détails en partant de l'objectif 1. Pour commencer, nous avons constitué deux ensembles de données : un premier sous-ensemble qui concerne uniquement l'information des clients d'IAAH (cet ensemble qui sera développé davantage dans notre exemple) et un deuxième sous-ensemble qui concerne l'information client issue de toutes les sources de données, mais en isolant les personnes qui ont contracté de l'assurance (on parle du payeur et non de l'assuré si ce sont deux personnes distinctes). On s'intéresse ici plus particulièrement au profil démographique et à la localisation du client, ainsi qu'à son profil de consommation des produits et services d'IAAH. La Figure 3.14 illustre quelles données ont été retenues pour faire cette analyse.

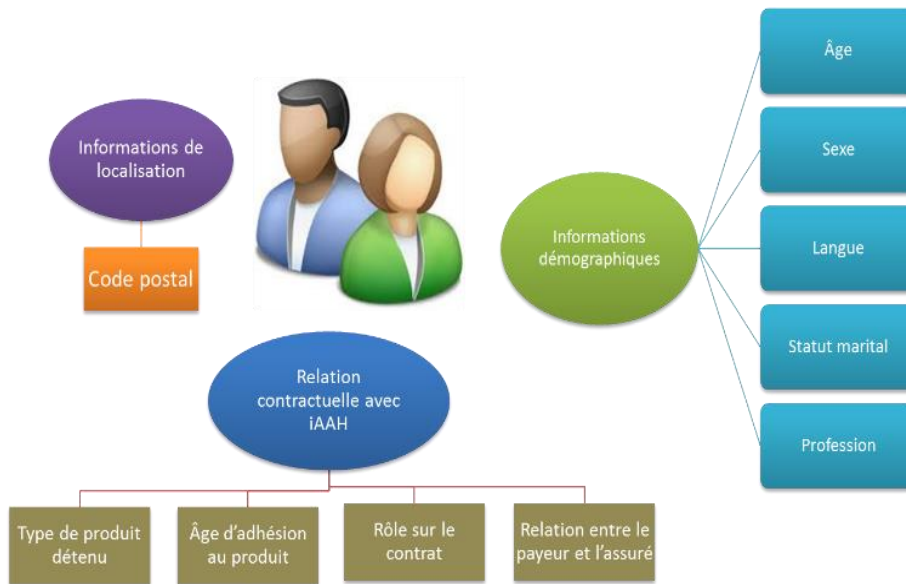


Figure 3.14: Représentation conceptuelle de l'ensemble de données choisi

À partir de cet ensemble de données, nous avons produit des visuels graphiques et des métriques statistiques avec Power BI. Nous avons créé un rapport avec l'outil rassemblant ces informations, ce qui a permis de créer une première vision globale du profil de la clientèle d'iAAH (voir Figure 3.15). Il est important de préciser que dans le secteur de l'assurance, l'âge est une donnée fondamentale. C'est pour cette raison qu'il s'agit de la donnée cible des métriques qui ont été produites.

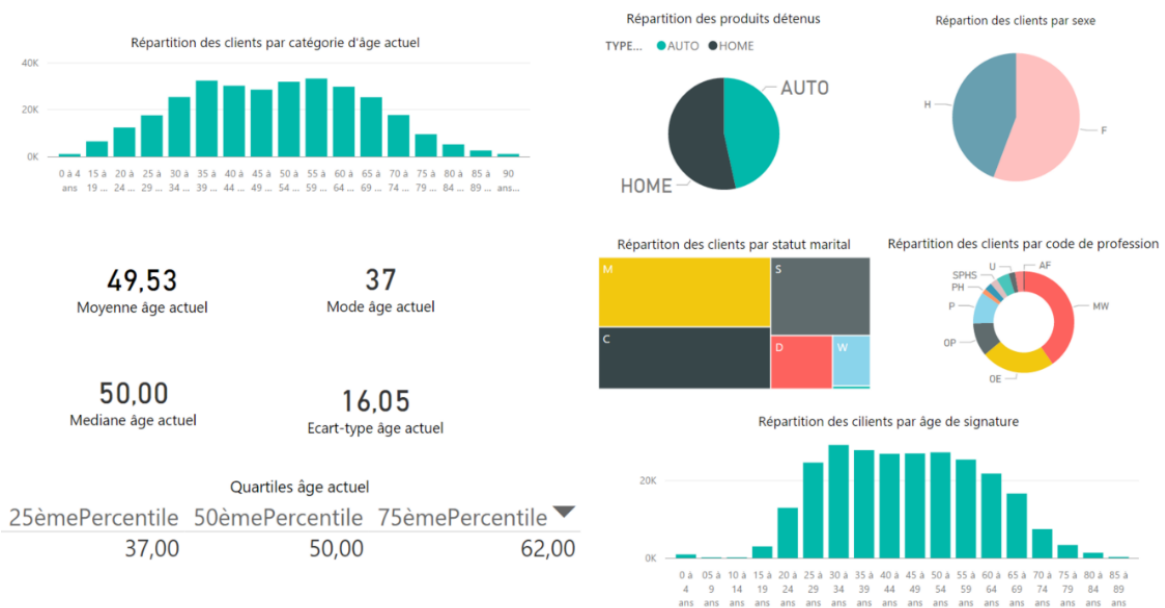


Figure 3.15: Visuels graphiques et métriques profilant la clientèle d'iAAH

Grâce à ce premier portrait, on a été en mesure d'orienter la fouille de données. On a dégagé les objectifs de fouille suivants :

- Objectif de fouille 1 : Établir 4 profils démographiques de clients distincts
- Objectif de fouille 2 : Établir 3 séquences de consommation de produits

On a ensuite élaboré des scénarios de fouille pour chacun de ces objectifs. Dans le cadre de cet exemple nous présenterons seulement le processus d'analyse suivi pour l'objectif de fouille 1. Les scénarios qui ont été sélectionnés sont :

- Scénario 1 : on exploite un algorithme de clustering de type K-mean
- Scénario 2 : on exploite un algorithme de clustering de type K-medoid
- Scénario 3 : on exploite un algorithme de clustering basé sur la densité (DSCAN)

On applique par la suite les techniques de fouille choisies pour chaque scénario. Seul le scénario 2 va être explicité à titre d'exemple. On commence par transformer les données avec les fonctions disponibles dans Rapidminer. Par exemple, les algorithmes de clustering fonctionnent mieux et plus rapidement avec des champs contenant uniquement des valeurs numériques et ne peuvent fonctionner avec des valeurs manquantes. On a donc remplacé les valeurs nominales d'un champ par un code numérique, les valeurs manquantes par des valeurs arbitraires telles que « inconnu » ou encore « 999 » pour les champs numériques. On a également exclu les champs contenant des identifiants de clients pour éviter qu'ils soient interprétés par les algorithmes comme des valeurs à prendre en compte dans le calcul de distance, ce qui aurait faussé les résultats.

On a ensuite appliqué la fonction « K-médoïd operator » de Rapidminer correspondant à l'exécution de l'algorithme et on a vérifié que les résultats produits sont cohérents en interrogeant les experts affaires qui ont une bonne idée (même intuitive) des profils qui composent leur clientèle. Plusieurs essais et ajustements ont été nécessaires à cette étape.

Enfin pour produire les résultats, on a réaffecté les données nominales aux codes numériques utilisés pour dérouler le scénario, et on a représenté synthétiquement les résultats de manière à alimenter la prochaine phase convenablement (la phase « interpréter et faire le bilan »).

Dans le Tableau 3.3 on voit que l'on obtient 4 profils différents de clients (colonnes cluster i) : 2 profils payeurs et 2 profils d'assurés. Il s'agit de profils suffisamment distincts pour être intéressants à interpréter dans la phase suivante de notre démarche expérimentale (voir section suivante).

Tableau 3.3 : Résultat du processus d'analyse 1

Cluster 0	Cluster 1	Cluster 2	Cluster 3
Homme	Homme	Femme	Homme
Payeur	Assuré	Payeur	Assuré
Professionnel	Étudiant post secondaire	Professionnel	Travailleur manuel
Marié	Célibataire	Mariée	Célibataire
Conjoint assuré	Relation avec le payeur indéterminé (probablement enfant)	Conjoint assuré	Relation avec le payeur indéterminé (probablement enfant)
Age actuel entre 60 et 64 ans	Age actuel entre 15 et 19	Age actuel entre 55 et 59 ans	Age actuel entre 30 et 34 ans
Age de signature entre 55 et 59 ans	Age de signature entre 10 et 14 ans	Age de signature entre 50 et 54 ans	Age de signature entre 25 et 29 ans
Achète un produit d'habitation	Achète un produit auto	Achète un produit auto	Achète un produit auto

3.7.4 Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA

- Problèmes pour lesquels nous avons réussi à trouver une solution :
 - Un problème important rencontré est qu'il nous arrivait de nous éparpiller dans la fouille de données car les objectifs d'analyse exprimés au départ par les experts d'affaires n'étaient pas suffisamment précis, découpés et priorisés pour faciliter le processus d'analyse. Pour essayer de remédier à ce problème nous avons redoublé de vigilance lors des activités de fouille en revenant le plus régulièrement possible aux axes d'analyse initiaux fixés et en nous recentrant sur les objectifs d'affaires poursuivis
- Contraintes auxquelles il a fallu s'adapter car aucune réelle solution ne s'offrait à nous :
 - Les outils de visualisation et de fouille de données demandent beaucoup de temps de prise en main et il n'a pas toujours été évident de trouver les techniques d'analyse adaptées à notre contexte d'expérimentation peu opérationnel
 - Certaines analyses peuvent mener à des résultats non concluants surtout du fait de la faible qualité des données. Il est donc important de justifier correctement les résultats obtenus lorsqu'ils sont jugés non concluants.
 - L'utilisation d'algorithmes avancés de fouille de données présente les inconvénients suivants :
 - Ils demandent beaucoup de temps, d'essais/erreurs
 - Ils nécessitent une grande puissance de traitement de calcul
 - Ils ne s'adaptent pas à tous les contextes d'analyse
 - Ils demandent parfois de connaître a priori la forme du résultat attendu (exemple: le nombre de clusters que l'on vise à obtenir en sortie)

3.8 Phase 6 : Interpréter

3.8.1 La méthode

Suite à l'application de la démarche d'analyse présentée dans la phase précédente et à partir de chaque objectif d'analyse, on obtient des résultats pour chacun des scénarios de fouilles identifiés. Chaque résultat doit être interprété pour arriver à des conclusions qui répondent aux objectifs fixés dans la phase 3 de l'établissement de la cible (voir Figure 3.1).

Pour cela, nous recommandons de :

1. Faire la synthèse des résultats des analyses réalisées en reprenant les résultats de chaque scénario et en identifiant s'il y a des liens à faire avec d'autres résultats de scénarios
2. Formuler des hypothèses ou des questionnements sur différentes interprétations possibles de ces résultats, en tenant compte des éléments de contexte liés aux données qui alimentent ces résultats et que l'on a identifiés dans les phases précédentes du projet pilote
3. Confirmer ou infirmer ces hypothèses avec les membres du comité de supervision du projet
4. Conclure sur les informations obtenues :
 - a. Les objectifs d'analyse ont-ils été atteints ?
 - b. Pourquoi ?
5. Faire le bilan du projet pilote au complet et préparer les recommandations pour la suite des activités de mise en place de la gestion stratégique du changement (voir processus P1.5 de la Figure 2.1 du chapitre 2 ; recommandations sur le plan d'action et la vision stratégique du changement notamment)

3.8.2 Les actions menées à l'iA

Les actions menées à l'iA pour mener à bien cette phase d'interprétation ont consisté, d'une part, en la réalisation d'une dizaine de livrables visuels, entre avril et juillet 2016, présentant la synthèse des résultats des analyses. De plus, nous avons organisé 3 ateliers, avec les intervenants TI faisant partie de l'équipe de réalisation et avec les analystes et gestionnaires des secteurs d'affaires qui avaient déjà pris part au projet auparavant. Ces ateliers visaient à valider la bonne compréhension des résultats des scénarios, approfondir la réflexion sur les hypothèses d'interprétation et confirmer les hypothèses d'interprétation. Enfin pour faire le bilan sur le projet pilote par rapport au plan d'action stratégique dans lequel le projet s'inscrivait, nous avons organisé une dernière rencontre à la mi-juillet avec les personnes faisant partie de la structure de supervision des projets pilotes. Lors de cette rencontre les résultats des autres projets pilotes menés dans le cadre du plan d'action de gestion du changement ont aussi été présentés. Ceci a permis de faire le point, plus globalement, sur le succès ou l'échec de la démarche de gestion du changement.

3.8.3 Résultats obtenus

Le Tableau 3.4 présente une vue synthétique de l'ensemble des résultats obtenus pour tous les scénarios déroulés dans le cadre de notre expérimentation. La colonne « objectif d'analyse initial » reprend les objectifs d'analyse formulés à la section 3.5 (voir Phase 3 : Cibler). La colonne « objectif de fouille » correspond au volet d'analyse que l'on a choisi d'approfondir pour un objectif d'analyse donné.

Le « scénario de fouille » explicite la technique qui a été employée pour chaque objectif de fouille. La colonne « résultat de l'analyse » présente les conclusions auxquelles on a abouti à la fin du processus d'analyse. Le « livrable fourni au comité de supervision du projet » précise la forme sous laquelle chaque résultat a été présenté au comité en charge de superviser le projet. La colonne « interprétation principale du résultat » dresse un premier bilan de la pertinence des résultats obtenus compte tenu de l'objectif initialement fixé (première colonne du tableau). La dernière colonne met en évidence quels objectifs ont été atteints, non atteints ou partiellement atteints (c'est-à-dire où l'on obtient des résultats intéressants, mais pas suffisamment concluants pour remplir l'objectif d'analyse fixé). La justification du succès ou non dans l'atteinte des objectifs se retrouve également dans la colonne « Résultat de l'analyse »

À noter que ce tableau offre une vision synthétique de ce qui a été réalisé pour servir d'intrant au bilan final du projet pilote. De plus, certains champs portent la mention « confidentiel » car nous jugeons qu'il s'agit d'informations trop sensibles pour être divulguées.

Tableau 3.4 : Résultats et interprétations des activités d'analyse et de fouilles de données

	Objectif d'analyse initial	Objectif de fouille	Scénario de fouille	Résultat de l'analyse	Livrable fourni au comité de supervision du projet	Interprétation principale du résultat	Objectif atteint ?
1	Retrouver les clients d'IAAH qui sont des clients d'autres secteurs d'affaires et vice-versa	Fusionner les clients ayant des produits de différents secteurs d'affaires	Créer et appliquer des règles de fusion	Un effort supplémentaire est nécessaire fusionner plus de clients	Estimation de l'effort nécessaire	Il faudrait fusionner autrement les clients pour obtenir de meilleurs résultats	partiellement
2	Segmenter la clientèle	Étudier la localisation des clients IAAH	Périmètre IAAH avec des index régionaux	Une forte tendance se dégage sur la région Est, résultat à interpréter à l'aide des gens d'affaires	Production d'un rapport	Il s'agirait de la conséquence d'un effort marketing important dans cette région qui aurait porté ses fruits	oui
3	Segmenter la clientèle	Profilier les clients IAAH selon des critères Socio-démographique	Choix du modèle de clustering	4 profils distincts identifiés	Tableau présentant les 4 profils	Les profils identifiés sont cohérents avec l'activité d'affaire d'IAAH	oui
4	Segmenter la clientèle	Profil de consommation du client IAAH	Choix du modèle de classification	Pas de résultat final obtenu	Estimation de l'effort nécessaire	On dénote un manque de puissance de l'outil de data mining Rapidminer et un besoin important de préparation de données pour exploiter pleinement l'outil	partiellement
5	Segmenter la clientèle	Profilier des comportements de consommation de produits du payeur communs à plusieurs secteurs d'affaires	Choix du modèle intuitif donné par un agent	L'âge des clients est identifié comme étant une variable explicative forte des caractéristiques du cycle de vie d'un client chez iA	Un cycle de vie "typique" d'un client chez iA en fonction de son âge	Le modèle est jugé trop subjectif (il s'agit uniquement de la perception de l'agent interrogé)	partiellement
6	Segmenter la clientèle	Profilier des comportements de consommation de produits du payeur communs à plusieurs secteurs d'affaires	Choix du modèle de clustering	Une erreur d'identification des clients a rendu impossible l'isolation de manière fiable les payeurs des assurés afin d'analyser leur profil de consommation	Aucun	Aucune	non
7	Segmenter la clientèle	Déterminer des profils de clients en fonction de leur "valeur"	Définir ce que veut dire un "client de valeur" pour iA	Confidentiel	Confidentiel	Confidentiel	non
8	Comprendre le comportement d'achat d'un produit en particulier	Dégager des profils-clients d'achat	Choix du produit de Maladie Grave	Confidentiel	Confidentiel	Confidentiel	oui
9	Identifier les clients loyaux	Dégager des profils de clients loyaux	Définir ce que veut dire un "client loyal" pour iA	Les critères obtenus ne sont pas suffisamment précis pour être satisfaisant	Liste des critères identifiés	Un travail d'analyse plus approfondi est nécessaire pour aboutir à des résultats plus concluants	non

Sur les 8 objectifs d'analyses fixés au départ, on peut voir sur le Tableau 3.4 que 3 d'entre eux ont pu être atteints (les objectifs 2,3 et 8), 3 ont été partiellement atteints (les objectifs 1, 4 et 5) et 3 n'ont pas été atteints (les objectifs 6,7 et 8).

Les principales recommandations que nous avons proposées au comité de supervision du projet pilote à partir de cette synthèse des résultats l'expérimentation sont les suivantes :

- En lien avec les objectifs atteints :
 - Bien que l'objectif de dresser un premier portrait de la clientèle d'iA ait été atteint grâce aux résultats obtenus par segmentation, ces profils sont à raffiner en travaillant davantage en étroite collaboration avec les équipes d'experts des différents secteurs d'affaires.
 - Un travail plus approfondi de validation des hypothèses aurait été bénéfique afin d'augmenter la qualité des résultats obtenus
- En lien avec les objectifs partiellement atteints :
 - Il est essentiel de se doter d'une solution plus performante de gestion des données internes d'iA afin d'en augmenter la qualité
 - Rapidminer est un outil de fouille de données intéressant et facile de manipulation mais il ne s'agit pas d'un outil suffisamment puissant. De plus, son utilisation requiert d'anticiper davantage la préparation des données pour les rendre compatibles aux algorithmes disponibles dans l'outil
- En lien avec les objectifs non atteints :
 - La maturité de l'iA, bien qu'ayant évolué positivement tout au long du projet pilote, reste à un niveau encore trop faible pour permettre de faire des analyses de segments de clients plus orientées sur les notions de rentabilité ou de loyauté. Or ces analyses auraient beaucoup de valeur pour répondre aux objectifs d'affaires visés par la gestion stratégique du changement que l'iA cherche à mettre actuellement en place

3.8.4 Les problèmes et contraintes rencontrés lors de l'application de la phase au cas de l'iA

- Problèmes pour lesquels nous avons réussi à trouver une solution :
 - Un des problèmes rencontrés portait sur les résultats des scénarios : certains résultats n'apportaient pas suffisamment d'éléments précis pour aboutir à une interprétation fiable. Pour résoudre ce problème, nous avons suggéré de se concentrer sur la vision d'ensemble des résultats plutôt que de s'attarder aux détails des résultats de chaque scénario
- Contraintes auxquelles il a fallu s'adapter car aucune réelle solution ne s'offrait à nous :
 - L'implication des personnes d'affaires étaient fortement limitée, de par leur manque de disponibilité. Ceci a complexifié nos tâches d'analyse des résultats des scénarios, et nos conclusions furent parfois laborieuses à obtenir, voire incomplètes
 - Il nous été difficile de savoir, à certains moments, si nous avions retenu la bonne hypothèse d'interprétation.

Nous nous sommes alors ramenés à nos objectifs d'affaires initiaux : mettre en évidence des tendances et non des vérités absolues. Ceci nous a permis de relativiser suffisamment les résultats que nous obtenions

3.9 Les itérations à l'origine de l'élaboration de la démarche expérimentale proposée

Comme cela a été précisé au chapitre 1, nous recommandons dans ce mémoire de développer une démarche de mise en œuvre de la gestion stratégique du changement et de mettre en place des projets pilotes qui suivent une approche itérative et incrémentale. En effet, adopter ce type d'approche facilite l'intégration graduelle d'évolutions et d'ajustements à la démarche et aide à atteindre les objectifs ciblés de manière plus certaine grâce à une approche par l'essai/erreur. Dans le cas du projet pilote de croisement de l'information qui fait l'objet de ce chapitre, nous avons réalisé plusieurs itérations au cours de chaque phase de la démarche que nous avons élaborée. Ces itérations nous ont permis d'obtenir la version de la démarche qui a été présentée dans les sections précédentes. La figure ci-dessous reprend les éléments de la Figure 3.1 mais en illustrant, cette fois-ci, l'enchaînement des phases avec des itérations qui sont symbolisées par les doubles flèches noires.

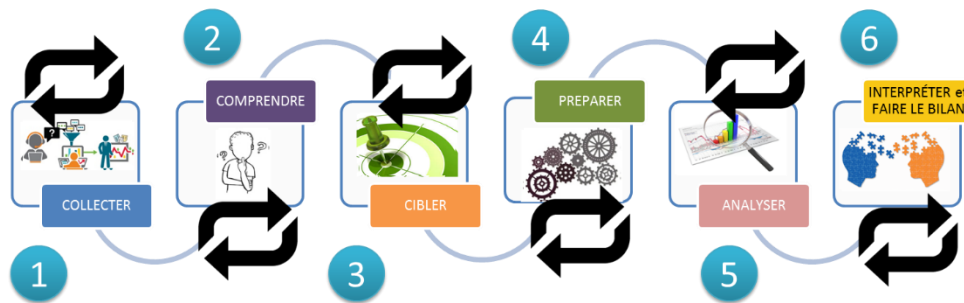


Figure 3.16 : Approche itérative de la démarche du projet pilote de croisement de l'information client

Dans cette section, nous montrons comment la démarche expérimentale a été appliquée par itérations à chaque phase du projet.

3.9.1 Itération et leçons tirées de la phase de collecte

Nous référons le lecteur à la section 3.3.3 pour comprendre le contexte de cette itération.

Un exemple d'itération pour la phase collecte est lié aux sources de données exploitées. L'entrepôt de l'Épargne Individuelle contient en plus des données d'épargne des données d'assurance vie.

Ces données ont été récupérées pour compléter les données du système de tarification afin de réduire les enjeux mentionnés (section 3.3.4) de qualité de ce système. On comprend ainsi que la méthode d'analyse préalable des sources de données mises à disposition doit être régulièrement révisée et que la démarche doit pouvoir s'adapter rapidement à des changements de sources de données.

La principale leçon tirée de l'exécution de cette phase chez iA est qu'il est essentiel d'impliquer les personnes des secteurs d'affaires dans la définition des besoins de collecte. Sinon, on s'expose au risque de collecter de l'information sur des périmètres trop larges et d'augmenter considérablement l'effort fourni pendant la phase de préparation.

3.9.2 Itération et leçons tirées de la phase de compréhension des données

Nous référons le lecteur à la section 3.3.4 pour comprendre le contexte de cette itération.

Au démarrage de la première itération de cette phase, on pensait que les données du système de tarification de l'assurance individuelle restituaient une représentation fidèle de la clientèle d'assurance individuelle. Suite aux travaux de décorticage des données et à aux indications des experts du système (affaires et TI), on a fait le constat que c'était en réalité une représentation biaisée. En effet, le système rassemble des données sur des individus, mais ces individus ne sont pas nécessairement tous des clients. Ils ont simplement fait l'objet d'un examen de tarification sans que cela ait mené obligatoirement à l'ouverture d'un contrat chez iA. Il a donc fallu en tenir compte dans l'analyse de la nature des données collectées et accroître les efforts de validation auprès des experts du système.

Cela nous a permis de comprendre que cette phase peut s'avérer laborieuse surtout lorsqu'il y a peu de documentation accessible sur les sources de données traitées. De plus, cela démontre à nouveau l'importance cruciale de collaborer étroitement avec les équipes responsables de ces sources.

3.9.3 Itération et leçons tirées de la phase de définition de la cible

Nous référons le lecteur à la section 3.5.3 pour comprendre le contexte de cette itération.

Lors de la première exécution de cette phase, on a obtenu des résultats mitigés lors de l'analyse de la segmentation de la clientèle faite en croisant les données de plusieurs sources et d'avoir l'appui de leurs gestionnaires pour qu'elles puissent dégager du temps pour notre projet.

Nous avons révisé notre cible d'exploitation en nous focalisant sur l'analyse des données de l'iAAH. Cependant, le fait de changer de cible implique de faire des ajustements sur les phases qui suivent celles-ci, à savoir: la préparation de données, l'analyse et l'interprétation.

On s'est ainsi aperçu que définir une cible sans pouvoir suffisamment bénéficier de la vision et des orientations des experts et gestionnaires d'affaires (par manque d'implication ou de disponibilité de leur part) rend la tâche plus difficile, car les choix qui en découlent ne sont pas nécessairement les bons.

3.9.4 Itération et leçons tirées de la phase de préparation

Nous référons le lecteur à la section 3.6.3 pour comprendre le contexte de cette itération.

Cette phase a fait l'objet de plusieurs itérations notamment au niveau de la modélisation des données collectées. En effet, du fait de la découverte progressive des enjeux de disponibilité et de qualité des données sources, le modèle de données a dû être révisé à plusieurs reprises et plusieurs versions de celui ont été générées :

- Version 1 (début février 2016) : Modèle selon le profil client /consommateur envisagé
- Version 2 (fin février 2016) : Cube OLAP envisagé
- Version 3 (début mars 2016) : Modèle basé sur la relation contrat-client
- Version 4 (fin mars 2016) : Modèle de représentation par localisation
- Version 5 (début avril 2016) : Modèle analytique simple avec ajout de champs calculés

Les problèmes rencontrés à l'origine de ces multiples itérations du modèle de données nous ont permis de prendre conscience de la nécessité d'anticiper davantage la possibilité d'être confronté à des difficultés de modélisation au cours des phases de compréhension et de définition de la cible et d'en tenir compte dans l'estimation des efforts.

Diverses leçons ont été tirées. En premier lieu, un modèle de données trop restreint dans la variété des informations qu'il représente ne permet pas d'aller suffisamment loin dans l'analyse, ce qui restreint l'interprétation par la suite. En second lieu, lorsqu'on est face à des problèmes importants de qualité de données, la priorisation des efforts de mise en conformité et de nettoyage des données devient cruciale. Sans cette priorisation on pourrait se lancer dans de grands travaux de nettoyage et passer peut-être à côté des objectifs de l'analyse, comme par exemple traiter des faits absolus plutôt que d'essayer de découvrir des tendances.

3.9.5 Itération et leçons tirées de la phase d'analyse

Nous référons le lecteur à la section 3.7.3 pour comprendre le contexte de cette itération.

Pour cette phase, c'est le choix des techniques de fouille qui a fait l'objet d'une itération. En effet, la fouille de données a posé un certain nombre de difficultés, notamment en ce qui concerne l'exploitation d'algorithmes de segmentation.

Les algorithmes de K-mean, par exemple, conduisent à des résultats qui sont difficilement interprétables directement et qui demandent beaucoup de transformation de données. Les algorithmes de K-medoid présentent plus de souplesse et permettent d'aboutir directement à des résultats interprétables.

On a choisi de se recentrer sur l'analyse descriptive et la visualisation, le temps de trouver une solution aux bloquants identifiés. Soulignons qu'il faut s'assurer que les scénarios sont réalisables. Aussi, porter une attention particulière aux choix des techniques d'analyse à l'étape de Préparation est essentiel. On retiendra que la plupart des techniques de fouilles employées dans ce genre de contexte nécessitent un degré supérieur de qualité de données pour aboutir à des résultats exploitables.

De plus, il est important de préciser que dans le processus d'analyse, la recherche de corrélations tôt dans le processus de fouille peut faciliter l'élaboration de scénarios en identifiant des liens pertinents entre les données.

3.9.6 Itération et leçons tirées de la phase d'analyse d'interprétation

Nous référons le lecteur à la section 3.8.3 pour comprendre le contexte de cette itération.

Les itérations qui ont été réalisées pour cette phase ont été liées principalement à l'interprétation des visualisations obtenues à partir des résultats d'analyse. En effet, si on prend l'exemple présenté à la Figure 3.17, la forme que prend le graphique peut donner conduire à deux interprétations différentes.

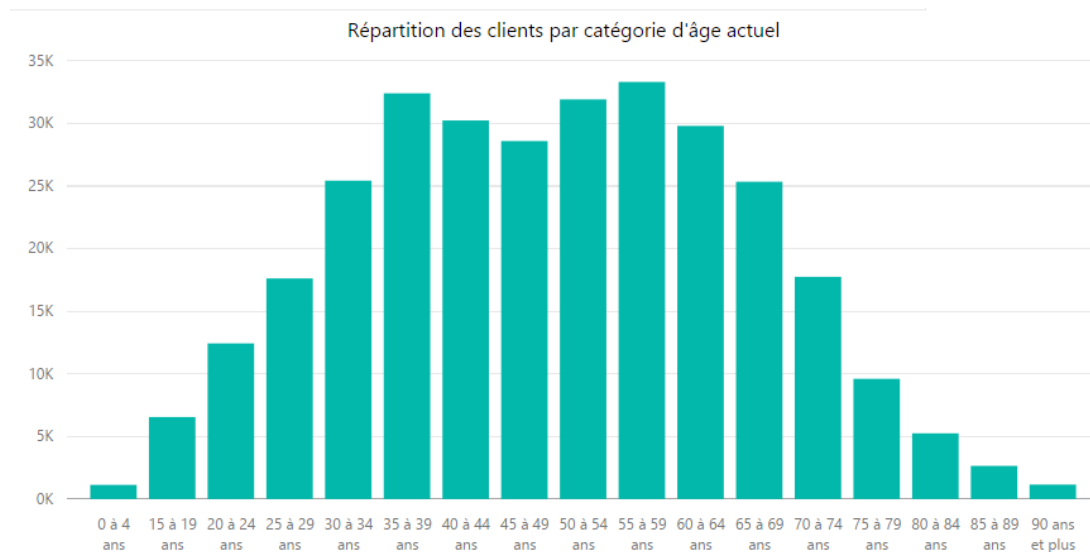


Figure 3.17 : Histogramme représentant la répartition des clients (tous secteurs confondus) par catégorie d'âge

Une première interprétation de ce graphique est la suivante : il y a une valeur seuil pour l'âge à partir de laquelle l'âge des clients concorde avec un âge que les vendeurs estiment « normal » pour la possession d'un produit d'assurance ou d'épargne. Cette interprétation est illustrée par la Figure 3.18.

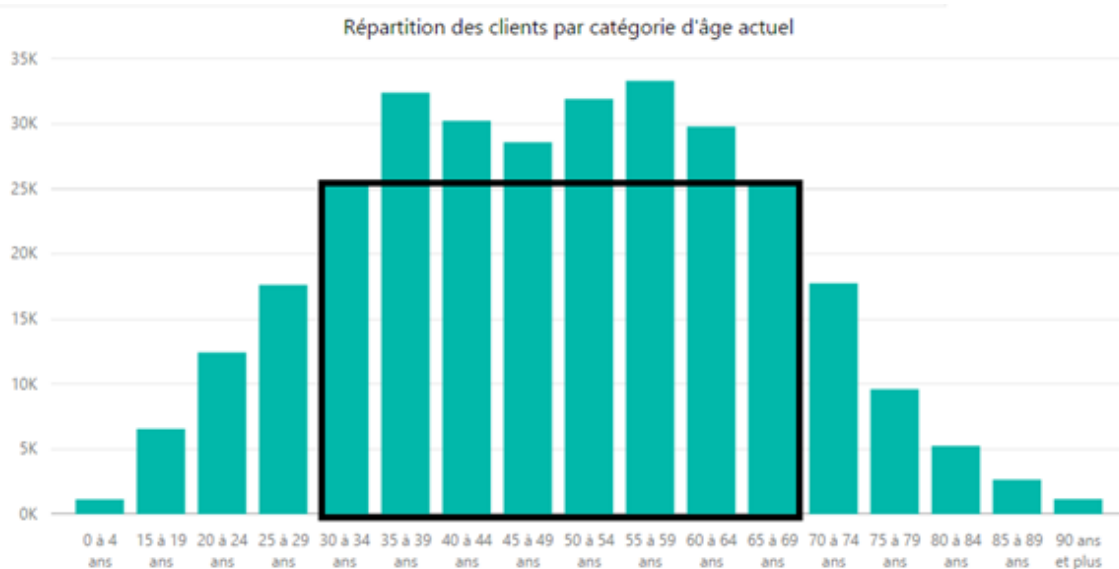


Figure 3.18 : Représentation d'une interprétation possible de la figure 3.16

Les valeurs au-dessus de ce seuil s'expliquent, dans cette logique d'interprétation, par des singularités pouvant être interprétées de plusieurs manières. Il pourrait s'agir des effets d'une campagne marketing particulièrement réussie à une période donnée et qui a attiré un nombre important de nouveaux clients. Il pourrait également s'agir de variations liées à des changements de situation personnelle des clients se produisant à la même époque.

Par exemple, les cas de divorces qui surviennent au cours d'une plage d'âges identique pour la plupart des clients et qui entraînent des changements importants au niveau des contrats d'assurance possédés.

Une autre interprétation possible du graphique de la Figure 3.17 est présentée à la Figure 3.19. Dans ce cas on fait l'hypothèse que la distribution des âges suit une loi normale, ce qui nous permet d'appliquer les propriétés d'interprétation propres à une loi normale de distribution. Par exemple, il est possible de calculer une valeur approchée de la probabilité que l'âge des clients soit dans un intervalle $[m - a, m + a]$ autour de la moyenne m .

Pour choisir l'interprétation la plus plausible, il faut consulter les experts du domaine qui sont capables d'amener des éléments de contexte pour interpréter les résultats obtenus.

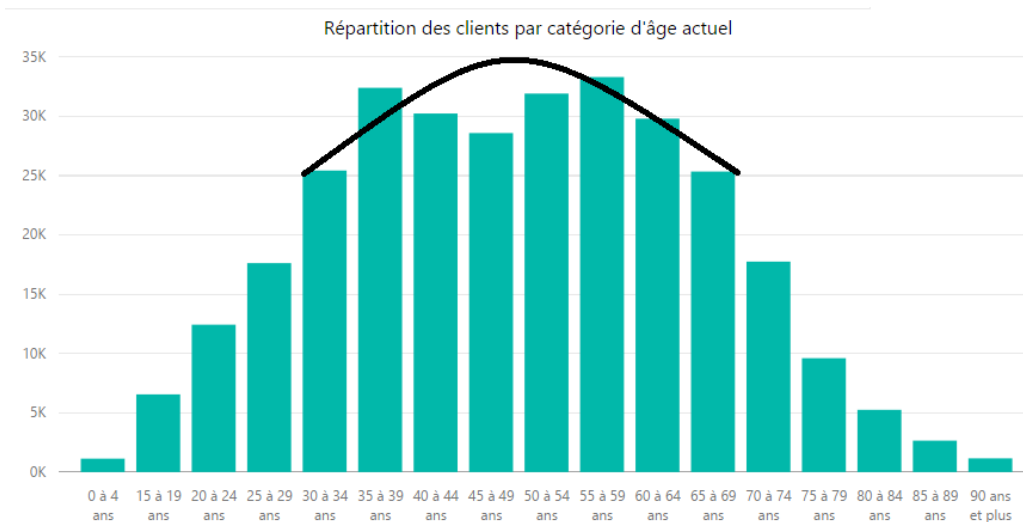


Figure 3.19: Représentation d'une autre interprétation possible de la figure 3.16

Ce que l'on retient de l'exécution de cette phase dans le cadre du projet pilote mené à l'iA est, qu'une fois de plus, sans avoir l'expertise d'affaires il est très difficile de confirmer ou d'infirmer les hypothèses tirées des analyses, et que même avec l'intervention des experts on ne peut pas avoir de certitude sur les tendances dégagées. D'où l'importance de se doter de moyens pour pouvoir tester la véracité des tendances observées.

3.10 Conclusion du chapitre

Nous présentons ici pour chacune des attentes initiales de ce projet pilote (notées Ai ci-dessous) les réponses que nous avons apportées (points Ri) :

- A1. Avoir une vision plus claire des capacités analytiques à développer au sein de l'organisation.
 - R1. Il a été identifié que deux types de capacités sont à développer : 1) l'accès à des ressources technologiques (infrastructure, outils, solutions) pour expérimenter de nouvelles approches en déployant moins d'effort; 2) l'approfondissement de la réflexion sur l'usage des approches analytiques à des fins stratégiques.
- A2. Conscientiser les gestionnaires des secteurs d'affaires sur l'importance du potentiel des données détenues par iA.
 - R2. Les résultats obtenus ont permis d'attirer l'attention des secteurs d'affaires sur les possibilités d'amélioration de leur processus d'affaires grâce à une meilleure compréhension et une vision globale des caractéristiques clés de leur clientèle
- A3. Dresser un premier portrait de la clientèle de la compagnie le plus transversal possible en tenant compte des différentes informations détenues par les secteurs d'affaires
 - R3. Le croisement transversal approfondi n'a pas mené à des résultats concluants dans le cadre de l'itération de la démarche du projet pilote illustrée mais met en évidence l'importance de porter une attention particulière à la phase de préparation de données.

Malheureusement cet aspect a été identifié seulement au cours de l'étape de la fouille dans le cas de ce projet pilote

- A4. Dégager des tendances sur la consommation de produits et services par la clientèle d'iA
 - R4.1 Des premières tendances globales ont été identifiées :
 - La clientèle d'iA est très variée du fait de la diversité des rôles existants pour un contrat donné (acheteur, assurés)
 - La clientèle est globalement plus jeune qu'imaginée a priori, ce qui amène à penser que le taux de croissance de la clientèle est bon
 - La consommation de produits s'explique plus facilement en considérant l'étape de vie dans laquelle se trouve le client à un moment donné
 - R4.2 Dans le cas d'iAAH
 - Les profils majeurs existants sont moins différenciés qu'attendu et correspondent globalement aux cibles-clients marketing choisies dans les dernières années par la compagnie
 - La variété des informations détenues sur le client permet d'approfondir davantage l'analyse de la segmentation
- A5. Faire avancer la réflexion sur l'identification de la valeur/rentabilité client
 - R5. La réflexion a été amorcée mais demandent davantage de temps car elle nécessite beaucoup d'efforts à la fois du côté de l'identification des besoins affaires, mais également du côté de la fouille de données

En ce qui concerne la performance du projet expérimental, voici les conclusions et valeurs de KPI auxquelles nous sommes arrivés :

- Sur les 8 objectifs d'analyse 3 objectifs ont pu être atteints, 3 objectifs ont été partiellement atteints et 3 objectifs n'ont pas été atteints
- La plupart des attentes ont été satisfaites
- Le budget établi initialement a été respecté
- Le délai de réalisation prévu pour ce projet a été dépassé d'environ 3 mois (8 mois étaient prévus au départ)

Les facteurs qui ont influencé cette performance sont :

- La petite taille de l'équipe de réalisation (une dizaine de personnes en tout) a limité la progression rapide du projet compte tenu de la charge de travail estimée
- L'intervention de davantage d'experts TI aurait permis de gagner du temps à certaines étapes, cruciales par exemple lors de la modélisation et du nettoyage des données, mais également lors de la phase « comprendre » où leur rôle était essentiel pour être en mesure de poursuivre le projet pilote
- Une implication plus fréquente des responsables d'affaires aurait également aidé à respecter davantage les échéances du projet, au moment de l'analyse et l'interprétation des résultats du projet pilote car ce sont eux qui possédaient toute la connaissance « métier »

De plus, certaines contraintes liées à l'historique du secteur, à l'historique d'iA et à son modèle d'affaires ont impacté à plusieurs reprises le projet, et ce à pratiquement chaque phase :

- La résistance au changement au sein de l'organisation
- Le manque de partage de l'information entre les divers secteurs
- Le manque de travail d'équipe au-delà des silos
- Les limites imposées par la réglementation sur l'exploitation de données-clients

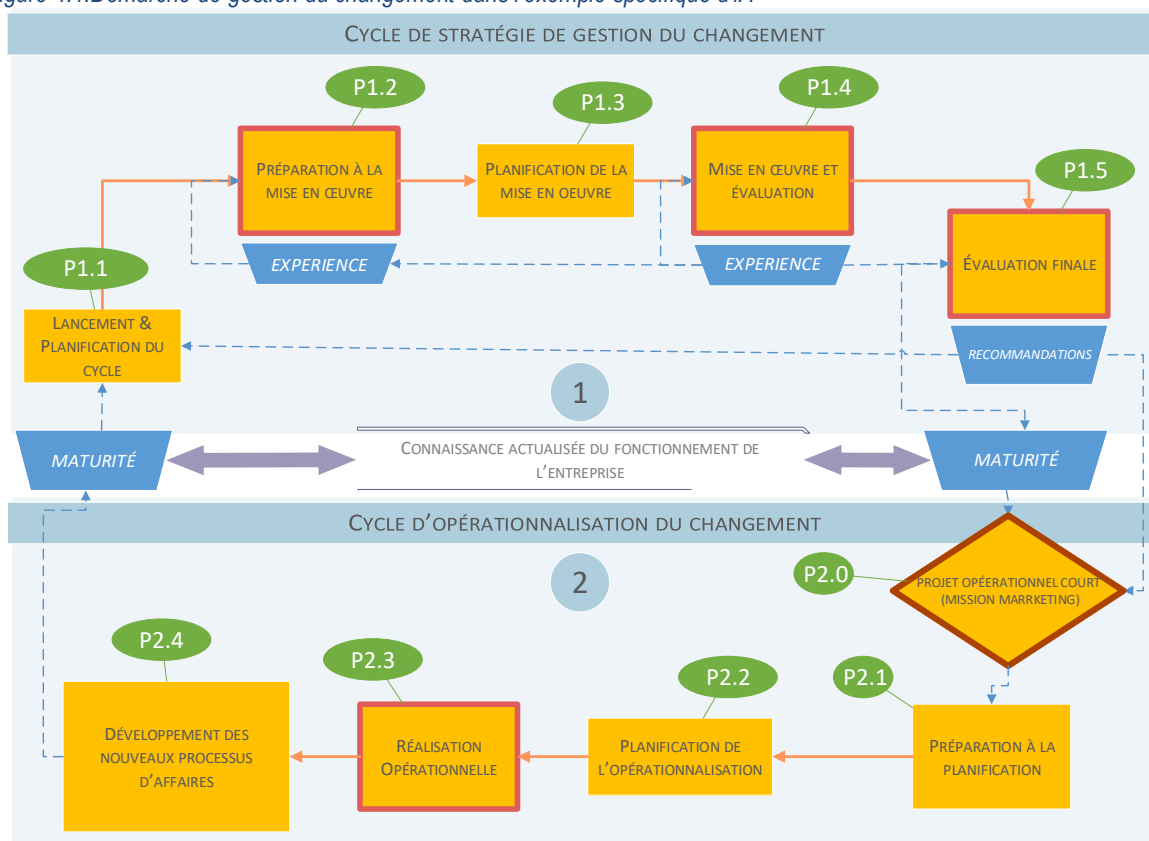
- Le manque d'accès à des ressources d'expertise en analyse de données avancée
- Le niveau faible de qualité des données collectées

4 Chapitre 4 : Opérationnalisation du projet pilote et leçons apprises

4.1 Introduction

Rappelons que, dans le chapitre 2 présentant la démarche de gestion du changement que nous proposons dans ce mémoire, deux cycles de gestion du changement ont été présentés : le cycle stratégique de gestion du changement (P1) et le cycle d'opérationnalisation du changement (P2). La vue d'ensemble de la démarche du chapitre 2 représentant ces deux cycles est replacée dans la **Erreur ! Source du renvoi introuvable.**

Figure 4.1: Démarche de gestion du changement dans l'exemple spécifique d'iA



Le chapitre précédent illustrait la mise en œuvre de gestion du changement (phase P1.4 dans la **Erreur ! Source du renvoi introuvable.**) par le biais de la mise en place d'un projet pilote visant à valider la vision et le plan d'action stratégique axé sur l'exploitation des données internes; ceci dans un but d'orienter les approches de ventes et de prospection pour mieux tenir compte des besoins du client.

Le chapitre précédent a également présenté les différentes itérations des phases de la démarche expérimentale que nous avons réalisées afin de raffiner la démarche du projet pilote. Nous allons voir dans la section 4.2 ce chapitre, d'une part, comment le démarrage du cycle opérationnel (P2 dans la **Erreur ! Source du renvoi introuvable.**) tient compte à la fois de l'application de la démarche du chapitre 3, de l'expérience tirée de ses itérations et des recommandations émises à l'issue de la phase d'évaluation finale du cycle stratégique de gestion du changement (processus P1.5 dans la **Erreur ! Source du renvoi introuvable.**). On propose ensuite d'introduire dans la section 4.3 de ce chapitre comment le démarrage du cycle opérationnel de gestion du changement chez iA en illustrant l'application d'une phase intermédiaire entre la fin du cycle de gestion du changement stratégique et le début du cycle de gestion du changement opérationnel : le lancement d'un projet opérationnel court appelé « mission marketing » (Phase 2.0 représentée par un losange dans la Figure 4.1**Erreur ! Source du renvoi introuvable.**). Cette mission a consisté en l'identification d'opportunités marketing entre l'iAAH et le secteur de l'Assurance Individuelle en exploitant des données internes sur le client et sur le réseau de distribution. La présentation de cette mission s'appuie sur le cadre méthodologique du chapitre 3 et permet de montrer comment la démarche expérimentale d'exploitation de données peut s'appliquer à un cas opérationnel concret.

À noter que seuls les éléments les plus illustratifs de l'application de la démarche seront présentés dans ce chapitre. De ce fait, certaines phases de la démarche du chapitre 3 n'y seront pas présentées.

4.2 Retour sur la fin du cycle de gestion stratégique du changement à l'iA : l'application de l'évaluation finale du cycle

4.2.1 Évaluation finale du plan d'action et de la vision du changement compte tenu des résultats du projet pilote

La Figure 4.2 reprend l'illustration de l'étape P1.4.3 (Figure 2.4 présentée à la section 2.4.3 L'élaboration du plan d'action). Les parties prenantes du projet à l'iA ont défini un plan d'action stratégique du changement dont un des volets consiste à mettre en place des projets pilotes liés à un ou plusieurs des objectifs du plan d'action afin d'en tester la validité. Chaque projet pilote a été supervisé par un comité de supervision qui a décidé à la fin du projet pilote et de ses itérations, de l'adéquation des résultats obtenus avec les objectifs du plan d'action et de la vision. C'est suite à cette évaluation que nous avons établi le bilan final du cycle stratégique de gestion du changement.

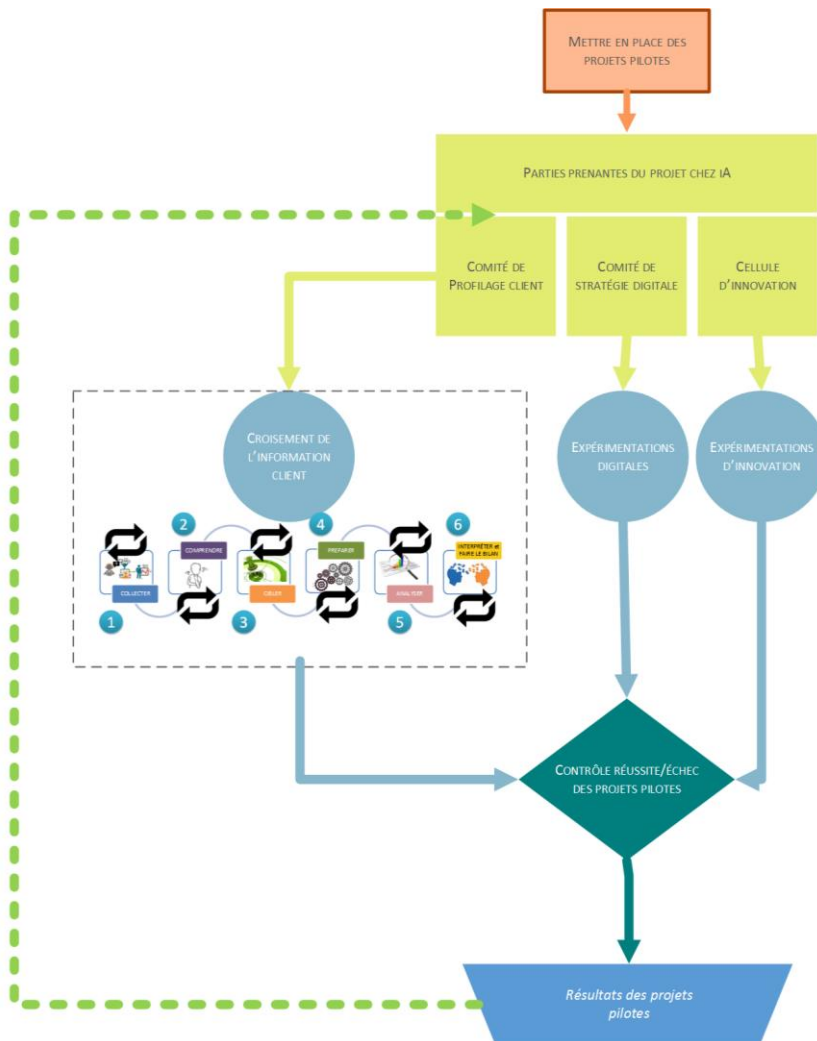


Figure 4.2: L'étape de mise en place de projet pilote à l'iA

Les résultats de notre projet pilote de croisement de l'information (voir chapitre 3) ont montré que l'iA était capable de collecter et exploiter des données-clients pour dresser un premier portrait de sa clientèle afin de mieux comprendre quelle relation celle-ci entretient avec iA. En effet, des segments de clients ont été découverts et analysés, ce qui a permis à l'iA de progresser en maturité et d'appliquer certains principes d'exploitation de données reliés au niveau de maturité visé par la gestion stratégique du changement (voir chapitre 2).

Cependant, certains résultats du projet pilote montrent que le niveau de maturité atteint à la suite de ce projet n'est pas suffisant pour remplir complètement les objectifs du plan d'action. En effet, la segmentation des clients qui a été obtenue n'est pas assez fiable pour exploiter les résultats des analyses en les intégrant aux processus d'affaires opérationnels et tactiques directement en lien avec la gestion de la clientèle.

Néanmoins, la mise en œuvre du projet pilote a permis de valider l'approche de gestion stratégique du changement consistant à mettre en place des projets pilotes pour valider le plan d'action et la vision du changement avant de prendre la décision de les opérationnaliser. Dans le cas de l'iA, la vision du changement et le plan d'action associé ont été révisés de la manière suivante :

- Révision de la vision :
 - Au lieu de chercher à optimiser la relation qu'iA entretient avec ses clients, l'iA désire renforcer progressivement sa compréhension des caractéristiques de sa clientèle.
- Révision du plan d'action :
 - Au lieu d'étendre le périmètre d'analyse aux données provenant des systèmes opérationnels des filiales de l'iA, la compagnie souhaite investir dans des travaux plus conséquents de rehaussement de la qualité des données qu'elle manipule actuellement en interne
 - Au lieu de chercher à mettre rapidement en place une structure de gouvernance des données qui implique directement les secteurs d'affaires concernés par le projet, les gestionnaires impliqués ont décidé de se concentrer davantage sur les modes de collaboration entre les secteurs pour faire de l'analyse de données croisées

Du fait de ces révisions du plan d'action et de la vision, il a été décidé par les gestionnaires d'iA parties prenantes du projet d'ajouter une phase P.2.0 (voir **Erreur ! Source du renvoi introuvable.**) au cycle opérationnel consistant à mettre en place un nouveau projet court pour tester la validité du plan révisé : il s'agit de la « mission marketing » qui sera présentée dans les sections suivantes.

4.2.2 Recommandations pour le lancement du cycle opérationnel

L'évaluation finale du cycle P1 (voir **Erreur ! Source du renvoi introuvable.**) a donné lieu à 3 recommandations majeures qui ont servi d'intrants au démarrage de la mission :

- Poursuivre l'approche de validation de la vision et du plan d'action en mettant en place des projets qui suivent la démarche expérimentale que nous avons mise au point (voir chapitre 3)
- Poursuivre les efforts de segmentation pour parfaire la connaissance de la clientèle, mais en appliquant la démarche expérimentale dans un contexte plus opérationnel
- Mettre davantage l'accent sur les aspects de sécurité, de performance des analyses et de confidentialité des données-clients en prévision d'une future intégration de ces analyses aux processus d'affaires de l'entreprise

4.3 Contexte de l'opérationnalisation du projet pilote : présentation de la mission réalisée avec une équipe marketing d'iA

4.3.1 Contexte de la mission

Comme précisé dans le chapitre 3, une des mise-en-œuvres de la gestion stratégique du changement a consisté à mettre en place une structure de supervision en charge d'orienter et superviser la stratégie digitale de l'organisation, en lançant des projets axés sur la digitalisation des activités de l'entreprise. Le Comité de stratégie digitale réunit aujourd'hui les principaux dirigeants de l'organisation.

Les orientations stratégiques du comité de stratégie digitale sont étroitement liées à l'exploitation de données internes puisque l'un de ses objectifs est de développer l'intelligence d'affaire au sein de la compagnie afin de trouver des opportunités d'affaires. Il a été décidé par le comité de stratégie digitale de lancer un projet court (la « mission ») qui vise à opérationnaliser certaines techniques testées pendant le projet pilote. Cette mission qui va être développée dans les sections suivantes est venue du constat qu'iAAH possède un segment de clients qui n'ont pas de conseillers financiers chargés de les accompagner dans le choix des autres produits offerts par la compagnie. En effet, il est important de préciser que le secteur iAAH a un modèle de vente un peu différent des autres secteurs de l'organisation dans la mesure où il n'y a pas de conseillers qui font l'intermédiaire entre les clients finaux et l'entreprise. Il n'y a donc pas de réseau de distribution qui soutient les activités de vente de produits de protection automobile et d'habitation. D'un autre côté, il y a le secteur de l'assurance et de l'épargne individuelle qui lui repose entièrement sur son réseau de conseillers pour réaliser des ventes. Or, la prospection de nouveaux clients et la rétention des conseillers de vente de l'iA sont des enjeux forts pour les secteurs d'assurance et d'épargne individuelle. L'objectif de la mission proposée était donc de tester des approches marketing incitant le client d'iAAH sans produit d'assurance et d'épargne à en acquérir.

4.3.2 Déroulement de la mission

La mission en tant que telle a duré 2 semaines mais la planification et la préparation de sa mise en œuvre s'est étalée sur 5 mois. Les parties prenantes de ce projet ont été :

- Les équipes marketing de l'assurance et de l'épargne individuelle
- Les équipes marketing d'iAAH
- L'équipe de Business Intelligence du côté TI composée de 3 architectes, 2 analystes et 3 développeurs

Les étapes de planification et de préparation de la mise en œuvre correspondent aux 4 premières phases de la démarche expérimentale du chapitre précédent : Collecter, Comprendre, Cibler, Préparer.

La mission concerne principalement les phases de ciblage, d'analyse et d'interprétation. Ainsi, les principales étapes du déroulement de la mission ont été :

1. Cibler les axes d'analyse du portrait de la clientèle visée (bassin d'étude)
2. Identifier des opportunités marketing possibles pour adresser cette clientèle
3. Identifier les clients actuels ayant des profils similaires au bassin de clients visés pour la prospection
4. Mettre en place des actions marketing clairement définies en se basant sur l'analyse des profils de la clientèle existante
5. Traiter les retours clients par rapport aux actions menées

4.4 Opérationnalisation de la phase « Cibler » dans la mission

4.4.1 La démarche opérationnelle et les opérations menées

Les gestionnaires des parties prenantes de la mission ont été rencontrés préalablement au démarrage de la mission pour fixer des objectifs d'affaires et des attentes réalistes, compte tenu de la situation actuelle de l'IA et de son niveau de maturité constatés à la suite à l'exécution du cycle stratégique de gestion du changement et des résultats du projet pilote.

Le périmètre de données choisi pour la mission est le même que celui identifié dans le projet expérimental durant le cycle stratégique de gestion du changement car on exploite les mêmes sources de données disponibles actuellement au sein de l'organisation. Néanmoins, au cours de la mission le périmètre de données à exploiter a été modifié : on a enrichi l'ensemble initial de données avec des données de rémunération des agents du réseau car elles permettaient de mieux comprendre la relation contractuelle qui lie un agent et son client.

L'approche analytique choisie suit les bonnes pratiques suggérées dans la littérature sur les projets opérationnels en intelligence d'affaires (BI). On n'est plus dans le cadre expérimental ou on a une BD constituée pour un besoin unique, on vise l'industrialisation de la démarche. Ainsi, on peut compter sur des efforts plus importants en termes d'architecture, de ressources et de moyens financiers.

En ce qui concerne les outils choisis pour la mission, l'équipe a considéré une gamme d'outils plus complète et orientée vers des technologies d'informatique nuagique (cloud) notamment pour des soucis de performance. Un nouvel outil supportant les activités de fouille de données a été choisi pour s'aligner sur les orientations corporatives de l'entreprise en regard des fournisseurs de solutions. Il a été également nécessaire dans ce cycle opérationnel de s'arrimer davantage avec les équipes d'exploitation de l'entreprise, en vue de tirer parti des données de production.

L'étape de raffinement de la cible a été facilité par les éléments suivants :

- Le mandat de la mission a été dressé directement par les gestionnaires d'affaires
- Le périmètre d'exploitation de données est plus restreint (il s'agit d'un segment spécifique de clients)
- Faire le bilan des opportunités réalisées sera un bon moyen de mesurer du succès
- Les attentes ne sont pas trop élevées dans la mesure où la mission est axée sur l'apprentissage et l'application de nouvelles approches
- La maturité de l'organisation a évolué positivement depuis la phase d'expérimentation : il y a une ouverture plus grande des équipes d'affaires à participer à des projets d'exploitation de données

4.4.2 Résultats obtenus

Les objectifs formulés dans le cadre de cette mission sont les suivants :

- Créer des opportunités d'affaires à partir du bassin de clients d'IAAH pour les conseillers du réseau de l'assurance et de l'épargne Individuelle tout en tenant compte des attentes des clients
- Favoriser l'intérêt des clients à élargir leur portefeuille de produits avec iA
- Aider à bâtir une relation entre le client d'IAAH et le conseiller du réseau afin que le client puisse élargir son portefeuille de produits
- Développer un apprentissage dans l'exploitation du BI pour soutenir les efforts en marketing
- Accroître les initiatives de convergence entre les activités marketing des deux secteurs (IAAH et l'assurance et épargne Individuelle)

Les livrables attendus à la fin de cette mission de deux semaines sont d'être en mesure de :

- Profiler le conseiller du réseau et ses clients
- Profiler les clients d'IAAH afin d'identifier ceux qui sont semblables aux clients des conseillers du côté de l'assurance et de l'épargne Individuelle afin d'augmenter le taux de réussite d'achats de nouveaux produits
- Définir un indicateur qui mesure le niveau d'intérêt des clients à élargir leur portefeuille de produits selon l'événement de vie qui déclenche le besoin de contracter un produit d'assurance ou d'épargne (ex. naissance d'un enfant, changement de style de vie, achat d'une maison, changement de voiture).
- Poser des actions marketing afin d'identifier un prospect qui a émis un signe d'intérêt fort pour un produit donné afin de soumettre ses informations à un conseiller qui pourra ensuite le contacter pour lui proposer d'acheter ce produit: ce type de prospect est appelé un « lead qualifié »
- Définir des indicateurs qui permettent de mesurer le nombre de leads qualifiés générés et le taux de conversion de ces leads en clients.

Comme cela a été précisé dans la section précédente le périmètre de données est sensiblement identique à celui défini pour le projet pilote au cycle stratégique. Enfin les indicateurs clés de performance choisis pour mesurer le succès du projet sont en partie les mêmes que ceux identifiés pour le projet pilote de croisement de l'information :

- Le nombre d'objectifs atteints par rapport aux objectifs initialement fixés
- Le degré de satisfaction des acteurs impliqués dans le projet quant aux résultats obtenus
- Le ratio entre le budget final et le budget initial
- Le ratio entre l'effort réalisé et l'effort initial estimé pour chaque phase

On ajoute à ces indicateurs trois autres mesures plus spécifiques au contexte de la mission à savoir :

- Le nombre d'opportunités réalisées
- Le taux de réussite des opportunités lancées 2 semaines après la fin de la mission
- La proportion de leads qualifiés parmi les prospects potentiels identifiés

4.4.3 Les enjeux opérationnels liés à cette phase

Le premier enjeu identifié est qu'il faut prévoir une durée suffisante pour la préparation de la mission avec les personnes des secteurs d'affaires directement impliquées dans la mission afin de s'assurer que la cible d'analyse soit claire pour tous les intervenants. Ensuite, compte tenu que la mission devait être réalisée sur une durée courte de deux semaines et en continu en présence des équipes d'affaires, il a été parfois difficile de prendre le recul nécessaire pour s'apercevoir qu'on s'écartait de la cible. Finalement, la nature et les attentes concernant les opportunités à déceler n'ayant pas été identifiées préalablement à la mission, il a été parfois difficile de mesurer efficacement si les attentes avaient bien été adressées.

4.5 Opérationnalisation de la phase « Préparer » dans la mission

4.5.1 La démarche opérationnelle et les opérations menées

Voici l'architecture qui a été proposée pour supporter des activités opérationnelles de cette mission :

- Une infrastructure complète a été mise en place pour charger des données dans un environnement de production, c'est-à-dire :
 - Accéder aux sources de données des systèmes de production
 - Modéliser les données dans un entrepôt
 - Constituer un comptoir de données et obtenir les bons accès pour faire l'exploration des données (accès en lecture/écriture minimalement)
 - Gérer de manière automatisée les flux de chargement des données
 - Respecter la sécurité des données
 - Limiter l'accès au comptoir de données aux analystes BI exclusivement
 - Accéder aux résultats d'analyse de manière sécurisée
 - Constituer un répertoire sécurisé de fichiers
 - Garantir la performance d'analyse
 - Mettre en place des plans de maintenance pour la fragmentation des index de la BD

L'architecture mise en place est schématisée par la figure ci-dessous :

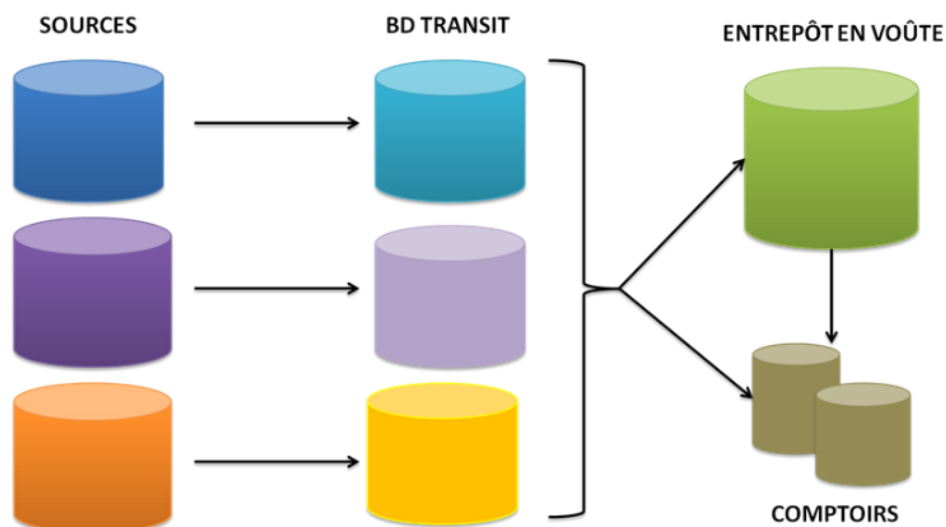


Figure 4.3: Architecture opérationnelle d'exploitation de données

On retrouve dans l'ordre sur la Figure 4.3 les sources de données en entrée, une base de données de transit qui stocke temporairement les données récupérées des sources, l'entrepôt de données modélisé en voûte et les comptoirs d'exploration qui sont des extractions de l'entrepôt pour réaliser des analyses spécifiques. En ce qui concerne la qualité des données, les données ne sont transformées que dans des vues au niveau des transits et dans les comptoirs d'exploration, contrairement à ce qui avait été fait dans le projet pilote. Cette pratique a été recommandée par les experts BI, lorsqu'on travaille dans un contexte opérationnel. Notons aussi que la documentation a été produite à chaque étape clé de préparation en respectant les normes d'entreprise.

Enfin, voici les outils choisis dans le cadre de cette mission opérationnelle, à chaque stade de l'exploitation de données :

- Stockage des données
 - BD SQL Server 2012
 - SQL Server Integration Services (SSIS)
- Modélisation et analyse de données
 - Plateforme Azure
- Visualisation et production de rapport de données
 - SQL Server
 - Excel 2016
 - SharePoint Server
 - Plateforme Azure

Les composantes de la plateforme Azure qui ont été utilisées sont les suivantes :

- Azure Data Lake pour avoir de l'agilité dans l'intégration et l'analyse des données
- Azure Machine Learning pour utiliser des techniques d'apprentissage automatique
- Power BI Desktop pour la visualisation des données
- Azure SQL Server pour stocker des données en cours de mission car il assure la compatibilité avec Azure Machine Learning

4.5.2 Résultats obtenus

À l'issue de cette phase de préparation de l'environnement d'exploitation des données on a obtenu les résultats suivants :

- Les comptoirs d'exploration stockent un ensemble de tables analytiques qui modélisent les données que les intervenants des secteurs d'affaires de la mission ont jugé pertinent d'analyser, à savoir les données sur :
 - Les agents et leurs clients rattachés
 - Les clients communs à l'Individuelle et iAAH
 - Les caractéristiques démographiques rattachées au code postal d'un client
 - Les résultats des requêtes de fouille de données
- Un nouvel entrepôt de données basé sur un modèle en voute a été construit.

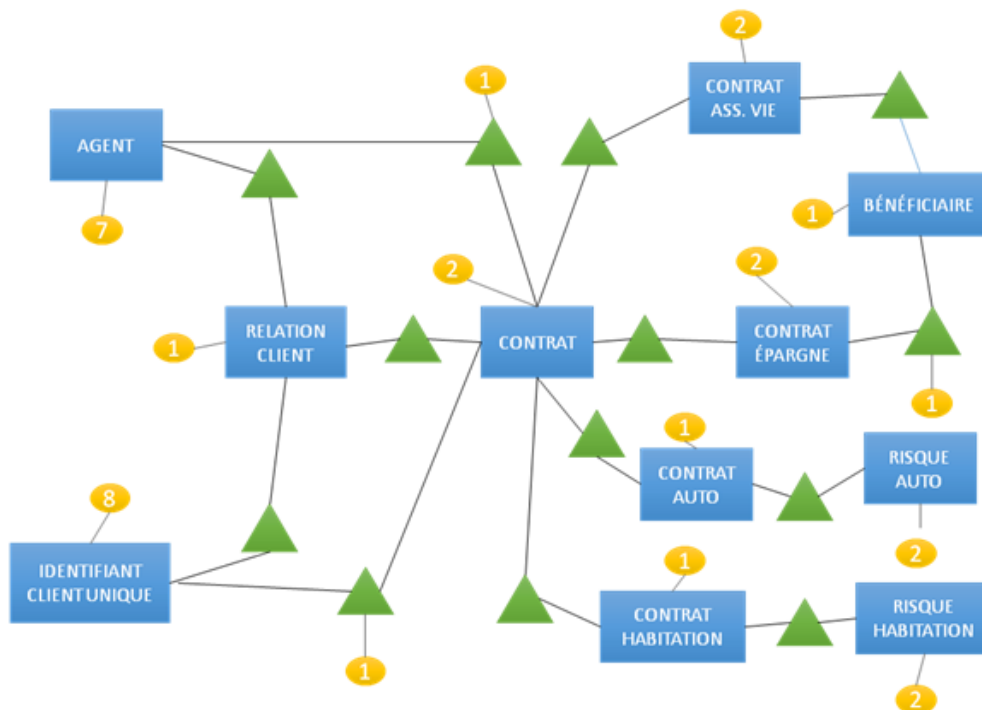


Figure 4.4: Modèle en voute de l'entrepôt de données opérationnel mis en place pour la mission

- Selon le lexique propre aux modèles de données en voûte, on retrouve sur la

Figure 4.4 les rectangles pour représenter les « hubs » (les tables contenant la liste des clés primaires), les triangles pour les « links » (les tables listant les relations entre les clés primaires) et les cercles pour le nombre de « satellites » (les tables contenant les données historiques associées aux hubs et links) propres à chaque hub ou link.

- Ce modèle de voûte a été réalisé pour supporter les besoins opérationnels d'exploitation de données-clients et plus particulièrement ceux de la mission marketing (processus P2.0 de la **Erreur ! Source du renvoi introuvable.**). Ce modèle représente les informations essentielles collectées sur le client : ses informations personnelles, les informations sur les contrats qu'il possède, les informations du conseiller qui gère ses contrats, ainsi que les informations concernant les autres personnes de son entourage qui sont concernées par ses contrats. Ce modèle présente également l'avantage de colliger de manière efficace les données provenant des systèmes de l'IAAH et de l'Assurance et Épargne.
- À noter qu'il s'agit d'une progression importante par rapport à ce qui avait été réalisé dans le cadre du projet pilote (chapitre 3) où il avait été difficile de fusionner ces données en raison des choix de modèles de données faits à ce moment-là.
- Ce choix d'un modèle de données en voûte pour structurer les données de l'entrepôt est particulièrement bien adapté à la réalité opérationnelle d'iA car il permet :
 - La modélisation de données issues de plusieurs sources avec un objectif d'historisation des données source
 - La traçabilité de l'information car on est capable de comprendre rapidement comment le modèle structure l'information

- La robustesse du modèle à d'éventuels changements de structure des données. En effet, les données sont actuellement de mauvaise qualité, mais le plan d'action opérationnel vise à une amélioration progressive de cette qualité. Cela peut donc entraîner des changements de structures de données. C'est pour cette raison qu'il est important d'avoir un modèle de données flexible.

4.5.3 Les enjeux opérationnels liés à cette phase

Les enjeux observés à ce stade sont les suivants :

- La construction du modèle en voûte a nécessité beaucoup de temps pour garantir qu'il supporte les modifications futures. En effet, plusieurs mois de travail de développement, d'intégration et d'analyse ont été nécessaires. Pour remédier à ces contraintes de temps fortes, il a été envisagé d'exploiter temporairement les BD de transit et les comptoirs (ce qui a été le cas pendant la mission)
- Pour permettre de croiser un grand volume de données depuis des bases de données en production et pour tenir compte de la sensibilité de l'information qu'elles renferment, les comptoirs d'exploration doivent être situés sur le même serveur que l'entrepôt. De plus, la mise en place de répertoires d'analyse permet de ne pas stocker des informations sensibles sur les postes de travail des analystes et facilite le partage des travaux
- La performance limitée du serveur dédié au projet a entraîné des blocages à plusieurs niveaux d'analyse et plus particulièrement lors des activités de fouille de données
- Pour intégrer de manière sécurisée les données dans Azure Data Lake il faut :
 - Prévoir le temps de prise en main de cette application
 - S'assurer de réussir à gérer la connexion entre les différents réseaux
- Les outils qui exploitent des technologies peu utilisées au niveau opérationnel dans l'entreprise soulèvent un enjeu important de maturité technologique. Aussi, on peut se demander si, à ce stade, l'entreprise est réellement prête à les exploiter au maximum.

4.6 Opérationnalisation de la phase « Analyser et interpréter » dans la mission

4.6.1 La démarche opérationnelle et les opérations menées

Dans le cadre de la mission présentée dans ce chapitre les objectifs d'affaires sont dès le départ plus clairs que lors du projet pilote (chapitre 2). En effet, on cherche à faire de la vente croisée entre les secteurs d'assurance et épargne Individuelle et le secteur de l'automobile et habitation (iAAH). On veut également dresser le portrait des agences du réseau de distribution susceptibles de vendre de nouveaux produits à des clients d'iAAH. Enfin dans le cadre du développement d'une approche de vente plus orientée vers les besoins du client, on souhaite répondre aux questions suivantes : Quels sont les clients qui consomment un produit donné? Quels sont les caractéristiques que possèdent ces clients ? Parmi les clients d'iAAH, quels critères peut-on exploiter pour les référer au bon conseiller dans les agences ?

De ce fait voici, les objectifs d'analyse qui ont été retenus :

- Dégager des profils de clients iAAH les plus susceptibles d'être intéressés par des produits ciblés de l'Individuelle. 4 produits ont été ciblés spécifiquement.
- Dresser le portrait des clients IAAH actuels qui possèdent ces produits pour confirmer que les hypothèses formulées sont valides
- Profiler les conseillers du réseau de distribution pour identifier ceux qui sont les plus aptes à solliciter ces clients

Mentionnons que dans l'ensemble les mêmes techniques de fouille et de visualisation ont été utilisées par rapport à ce qui a été fait au cours du projet pilote, bien que les outils soient différents. En fait, l'accent a davantage été mis sur la visualisation spatiale des données. Quant aux objectifs de fouille, ils sont orientés vers l'identification de profil de client intéressant à prospecter.

Pour cela, un des scénarios de fouille a été la recherche approfondie de corrélations pour faciliter l'identification de profils correspondant à des leads qualifiés et en exploitant l'outil Azure Machine Learning.

4.6.2 Résultats obtenus

Les résultats d'analyse obtenus sont présentés en prenant l'exemple du produit d'épargne pour les études des enfants vendu par iA (REEE ou R3E). On cherche à visualiser le portrait des clients possédant actuellement ce produit et qui sont des clients d'iAAH. Les résultats de ces visualisations est présenté sur la rapport dans la figure ci-dessous :

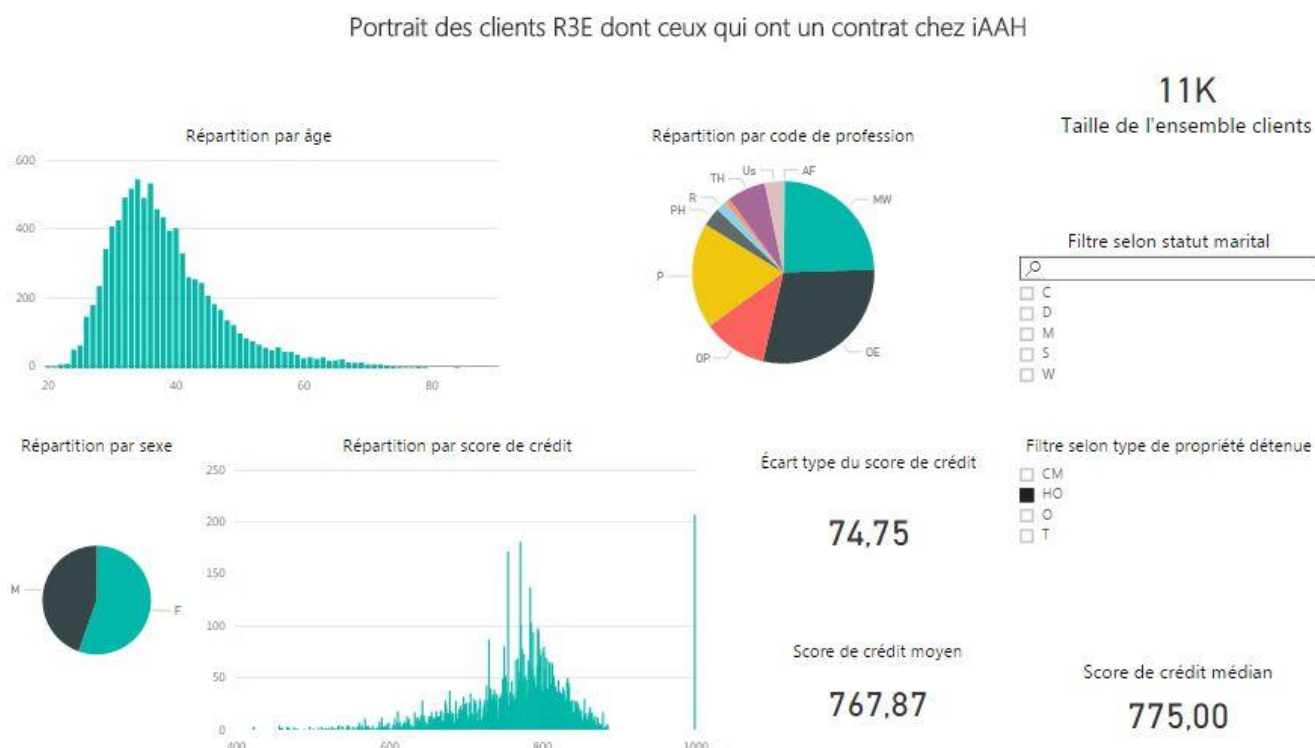


Figure 4.5 : Portrait des clients possédant un produit REEE et au moins un produit d'IAAH

Ainsi on observe dans la Figure 4.6 qu'entre 30 et 37 ans les employés de bureau et les professionnels ont davantage tendance à contracter un R3E.

Grâce à l'exécution de ce genre de scénario de fouille nous avons réussi à confirmer notre profil de client type et nous nous sommes servis de ce profil pour constituer un autre échantillon : l'échantillon de prospects ayant le profil « André(e) » mais qui ne possèdent pas (encore) de R3E chez iA.

Une offensive marketing a été lancée durant la mission et a consisté à envoyer un courriel à un échantillon test de 200 personnes de profil « André(e) » sans R3E.

Ce courriel contenait un message faisant la promotion des avantages du REEE et proposait un rabais en cas d'adhésion. Le courriel avait pour objectif d'inciter les destinataires à répondre au courriel en exprimant leur intérêt pour l'offre si tel était le cas.

Les personnes ayant répondu à ce courriel ont alors été considérées comme étant des leads qualifiés. Ces leads doivent être ensuite analysés pour être redirigés vers le conseiller adéquat qui se chargera de rencontrer le client en vue de lui vendre le produit. Pour cela, une analyse des agences à proximité des leads retenus a été réalisée, toujours au cours de la mission.

4.6.3 Les enjeux opérationnels liés à cette phase

Plusieurs enjeux majeurs sont apparus durant cette phase:

- Des difficultés d'accès à des sources d'information complémentaires pour raffiner la qualité des analyses
- Des enjeux liés au respect de la confidentialité pour obtenir le consentement des clients. Pour traiter ces enjeux, il a fallu remonter le plus haut possible dans la hiérarchie de gestionnaires. Il a été convenu que :
 - Seuls seraient considérés les clients ayant donné leur consentement pour l'utilisation de leurs renseignements personnels
 - Avant de réaliser une action marketing, les informations personnelles des prospects ont dû être validées à plusieurs reprises
 - L'action marketing qui s'en est suivie serait conduite sur un faible nombre de prospects
 - Les travaux durant la mission ont été supervisés par une personne des services juridiques de la compagnie
- Des enjeux liés à l'exploitation des outils :
 - Azure Machine Learning a permis de confirmer plusieurs intuitions lors des analyses, mais il n'a pas permis de faire ressortir des tendances nouvelles. En effet, les travaux de fouille de données nécessitent un périmètre de données plus large et précis que ce qui était disponible, ainsi qu'une qualité de données plus grande. De plus, l'outil ne semble pas idéal

pour être utilisé dans une mission de très courte durée car on ne dispose pas de suffisamment de temps pour réaliser des analyses en profondeur.

- Power BI n'est pas assez puissant pour pousser suffisamment loin les analyses spatiales. Par exemple, il est difficile de visualiser plusieurs informations sur une même carte

4.7 Bilan de la mission réalisée

4.7.1 Retour sur les attentes initiales

À l'issue de la mission, 4 opportunités ont été identifiées et ont mené à des actions marketing concrètes en s'appuyant sur les données qui avaient été mises à disposition. Voici les attentes qui ont été rencontrées :

- Profiler le conseiller du réseau de distribution et de ses clients
- Profiler les clients d'IAAH afin d'identifier ceux qui sont semblables aux clients actuels du conseiller
- Poser des actions marketing afin d'identifier un prospect qui a émis un signal d'intérêt afin de supporter les conseillers pour acquérir de nouvelles clientèles
- Définir des indicateurs qui permettent de mesurer le nombre de leads qualifiés générés et le taux de conversion de ces leads en clients

L'attente qui visait à définir un indicateur pour mesurer le niveau d'intérêt des clients à élargir leur portefeuille de produits n'a pas été réalisable car la maturité de l'organisation en matière de définition de besoins précis d'exploitation de données était encore insuffisante au moment où cette mission a été réalisée.

4.7.2 Retour sur les KPIs

En ce qui concerne les objectifs d'affaires initiaux présentés à la section 4.4.2, ils ont tous été atteints au cours de la mission. Un haut degré de satisfaction des gestionnaires mandataires de la mission a donc été atteint. En ce qui concerne les KPIs liés au budget du projet, ils ne seront pas présentés pour des raisons de confidentialité. En ce qui a trait aux efforts de réalisation, nous avons fait le constat que la durée de certaines phases a été sous-estimée, plus particulièrement les phases de ciblage et de préparation de la mission.

Pour ce qui est des résultats liés aux opportunités, le nombre de leads générés a donné satisfaction, étant donné la taille des échantillons de prospection : en moyenne un lead a été généré par opportunité. À noter que le taux de réussite des opportunités est à mettre en perspective avec le fait qu'aucun des leads qualifiés n'a abouti à une vente dans le mois suivant la mission.

4.8 Conclusion du chapitre

Dans le cadre du projet pilote de croisement de l'information présenté au chapitre 3, on cherchait à valider la démarche d'exploration des données internes à l'IA. Cette démarche expérimentale a été raffinée et réappliquée dans un contexte d'opérationnalisation du changement : améliorer le processus de prospection de nouveaux clients par le réseau de vente par l'identification de « leads qualifiés » grâce à des analyses avancées des données des clients communs à l'IAAH et du secteur d'affaire d'assurance et épargne individuelle.

Cette nouvelle expérimentation (opérationnelle cette fois-ci) a été faite sous la forme ateliers intensifs réunissant les experts d'affaires marketing d'IAAH, les experts d'affaires marketing du secteur d'affaire de l'assurance et épargne individuelle et d'experts en BI ainsi que notre équipe de recherche sur une période de deux semaines consécutives. Ces ateliers ont conduit à des résultats favorables à la poursuite du cycle opérationnel de gestion du changement. Ce projet a également permis de faire progresser davantage le niveau de maturité de l'organisation.

Quel bilan peut-on tirer du déroulement de la mission? En matière d'infrastructure, le bilan est que la performance et la disponibilité des infrastructures doivent être vérifiées plusieurs fois en prévision de demandes de traitement importantes et fréquentes. De plus, il est important de mentionner les limites des fonctionnalités des outils choisis lorsqu'on a voulu approfondir davantage certains axes d'analyse (par exemple la dimension spatiale). Un autre élément important concerne la sécurité de l'information que l'on manipule. Effectivement, l'utilisation de ressources matérielles adaptées (dossiers sécurisés, cryptage des données sensibles, etc.) ainsi que le recours à des experts en sécurité informatique est un aspect fondamental pour assurer une étape cruciale pour le bon déroulement de projets de ce type.

D'un point de vue méthodologique les constats que l'on peut faire sont les suivants :

- La composante géomatique de l'analyse est très pertinente pour analyser un marché dans un contexte marketing
- La validation des listes d'envoi pour une offensive marketing basée sur une analyse de tendance est importante pour ne pas prendre le risque de commettre des erreurs pouvant affecter l'image de marque de l'organisation (par exemple : recommandé à tort un produit)
- Il est nécessaire de s'assurer de formaliser suffisamment les attentes, engagements et objectifs des parties prenantes avant de démarrer de telles missions pour assurer par la suite qu'ils sont alignés avec les actions prises
- Insérer dans les systèmes opérationnels des processus de collecte de retour sur expérience des clients permettrait de mieux mesurer l'efficacité des actions portées et valider les tendances dégagées

- Lorsqu'une cible de prospection est identifiée, il est important de valider la pertinence des actions de prospection réalisées, en répétant les mêmes actions sur un bassin non ciblé, de manière à s'assurer que les résultats obtenus sont bien dus au choix du segment et non au hasard
- Les scénarios de fouille sont longs à dérouler; une mission de 2 semaines avec des interactions constantes avec les intervenants d'affaires nécessite que l'organisation ait un niveau de maturité technologique suffisant

Ces constats nous amènent à faire un bilan à plus haut niveau en lien avec l'amorce du cycle opérationnel préconisé par notre démarche. En effet, au cours du cycle stratégique de notre démarche, l'iA a élaboré sa propre vision stratégique du changement face à la transformation de la société et du marché. Cette vision se concentre en partie sur l'exploitation avancées de données. Le plan d'action qui en découle vise à explorer les données que l'entreprise possède à l'interne pour être en mesure d'identifier des tendances au niveau des profils de sa clientèle. Le but de cette exploration est de développer une meilleure connaissance de la clientèle et de transformer progressivement le modèle d'affaires actuellement orienté-produit vers un modèle orienté-client.

Afin de valider cette orientation stratégique, plusieurs projets pilotes ont été mis en place dont le projet que nous avons réalisé et qui visait à croiser et fusionner les données-clients provenant de plusieurs secteurs d'affaires. Ce projet pilote a apporté des résultats encourageants, mais non suffisants pour enclencher directement l'opérationnalisation du changement initié à l'iA. Une alternative qui a été choisie fut de réaliser un nouveau projet pilote mais, cette fois, directement lié aux activités opérationnelle et tactiques de l'entreprise. Les résultats obtenus à l'issue de cette mission ont cette fois-ci apporté suffisamment de résultats satisfaisants pour que l'iA décide d'enclencher réellement un changement opérationnel basé sur la stratégie construite au fur et à mesure de l'application de la démarche. Aujourd'hui l'iA a mis en place une équipe entièrement dédiée au développement de la connaissance du client en exploitant le plus de données-clients possibles disponibles dans l'entreprise. Cette équipe réalise actuellement (en 2017) plusieurs projets opérationnels visant à développer de nouveaux processus d'affaires plus matures dans leur intégration et leur exploitation des données clients.

5 Chapitre 5 : Conclusion du mémoire et perspectives

Ce chapitre de conclusion se décline en deux parties. Une première partie fera brièvement la synthèse du contexte dans lequel ce mémoire s'inscrit, ce que nous avons réalisé et qui constitue nos contributions et dans quelle mesure les objectifs du mémoire ont été atteints d'un point de vue théorique et pratique. Cette première partie mettra également en évidence quelques éléments qui n'étaient pas prévus au départ et qui ont permis d'enrichir la démarche proposée. Dans une deuxième partie nous mettrons en perspective ce qui a été réalisé en présentant les principaux points forts de notre travail de recherche, mais aussi les points d'amélioration que nous avons envisagés. Enfin, nous présenterons une courte discussion sur les suites à plus long terme de la problématique adressée.

5.1 Bilan sur le travail de recherche réalisé

Dans le chapitre 1 plusieurs volets du contexte dans lequel s'inscrit notre projet de recherche ont été mis en évidence car ils sont pertinents pour comprendre la problématique et les contributions de ce mémoire : le volet historique, le volet théorique et le volet pratique de l'exploitation avancée de données.

En ce qui concerne le volet historique nous avons présenté la transformation de notre société actuelle vers une nouvelle ère industrielle qui se caractérise notamment par l'apparition de nouvelles technologies dites 'disruptives' (l'internet mobile, l'automatisation de la collecte des connaissances, l'internet des objets, etc.) qui créent un phénomène de génération massive de données, à grande vitesse et sous une multitude de formats, ce qu'on appelle le phénomène Big Data. Au cours des dernières années, une certaine emphase a été mise, tant dans la recherche que dans les industries de pointe, sur l'importance de trouver de nouvelles façons d'exploiter ces données, et surtout d'en tirer de l'information différemment.

Cette transformation de notre société que certains qualifient de 'nouvelle révolution industrielle' offre de nouvelles possibilités de développement pour l'évolution des marchés et des compagnies qui y interviennent. On a cité notamment l'exemple de l'essor de l'économie collaborative qui est basée sur l'exploitation de ces nouvelles technologies pour créer une nouvelle économie de consommation basée sur l'échange de biens et services entre particuliers. L'argument que nous avons défendu en lien avec cet exemple est que ces changements ou développements dans notre société sont portés par des organisations qui ont adopté des modèles d'affaires innovants et qui sont devenus rapidement des leaders du marché. C'est le cas de compagnies comme Google, Facebook et Amazon, ou encore Airbnb et Uber.

Ces compagnies sont nées avec l'apparition de nouvelles technologies dites « digitales » et forcent les entreprises plus traditionnelles, qui étaient auparavant en position de force sur le marché, à remettre en question leurs paradigmes fondamentaux (façon de considérer le monde, les marchés et les clients), leurs façons de faire et leurs modèles d'affaires actuels.

L'un de ces nouveaux modèles d'affaires qui tend à se généraliser sur le marché est le modèle d'affaires orienté-client. Comme nous l'avons présenté dans le chapitre 1, ce modèle s'appuie sur l'utilisation de grandes quantités de données concernant les clients (actuels ou potentiels) d'une compagnie afin de leur offrir une 'expérience client personnalisée'.

Cependant, pour que les organisations soient en mesure d'appliquer ce type de modèles, elles doivent utiliser de nouvelles techniques d'exploitation des données.

Ces techniques d'exploitation des données peuvent être réparties en 4 grandes catégories : 1) Les techniques de stockage comme le NoSQL, l'informatique nuagique (Cloud) ou le lac de données (Data Lake); 2) Les techniques de chargement, de traitement et de modélisation comme Hadoop, Map Reduce ou Spark; 3) Les techniques de visualisation comme la visualisation interactive sous forme de tableaux de bord ou encore les systèmes d'information géographiques (SIG); 4) Les techniques d'analyse de données comme les techniques d'apprentissage automatique ou d'apprentissage profond.

Néanmoins, l'application de ces nouvelles techniques nécessite que ces organisations soient non seulement suffisamment matures technologiquement, mais aussi en mesure d'adopter un nouveau point de vue organisationnel pour être capables de les intégrer à leurs activités et processus d'affaires. Or il s'avère que cette maturité est encore rare dans un grand nombre d'organisations et d'entreprises actuelles, toutes industries confondues. Ainsi, peu de ces organisations ont les prérequis (en termes de vision et modèles d'affaires, de culture d'entreprise, d'intégration des systèmes informatiques et des données associées, de disponibilités de ressources humaines entraînées à l'exploitation avancée des données) pour effectivement et efficacement effectuer les transformations requises pour participer activement à cette révolution.

C'est l'essence de la problématique que nous adressons dans ce mémoire. En effet, notre recherche nous a conduit à proposer une démarche de gestion du changement destinée aux organisations qui souhaitent mettre en place les transformations nécessaires pour tirer parti des techniques d'exploitation avancée de données massives. Ce projet de recherche a été financé par le Fonds de recherche du Québec - Nature et technologies (FRQNT), le Conseil de recherche en sciences naturelles et en génie (CRSNG) et par l'entreprise partenaire de ce projet, la compagnie de services financiers canadienne Industrielle Alliance (iA). L'observation du fonctionnement de l'iA et une revue de littérature approfondie sur les démarches proposées actuellement aux organisations pour effectuer le 'virage des données massives', nous ont permis de mettre en évidence un élément manquant qui nous a semblé crucial.

En effet, La grande majorité de ces approches présuppose que les organisations ont les prérequis nécessaires à l'adoption des techniques avancées de gestion des données : une vision et un modèle d'affaires appropriés, une maturité technologique suffisante, des systèmes informatiques et des entrepôts de données intégrés (minimalement au niveau opérationnel, et idéalement aux niveaux tactique et stratégique) ainsi que la disponibilité de ressources humaines entraînées à l'exploitation avancée des données. Si tel est le cas pour les compagnies leaders du marché, c'est loin d'être le cas de la majorité des autres compagnies, même parmi les entreprises piliers de ce que certains appellent 'l'ancienne économie'.

Partant de ce constat et grâce à la participation de l'iA à ce projet de recherche, nous avons mis au point une démarche de gestion du changement technologique et organisationnel. Cette démarche a été conçue à partir des objectifs suivants : 1) Apporter aux organisations dont la maturité technologique et organisationnelle est éloignée de celle des compagnies leaders de la transformation « digitale » actuelle de notre société, un certain nombre de connaissances sur ce qu'il est nécessaire de faire pour rendre effective cette transformation; 2) Préciser les changements à effectuer pour permettre à l'organisation d'exploiter ses données internes (dont plusieurs caractéristiques relèvent du phénomène 'big data' - masses considérables de données, non intégrées, de formats différents, de qualités disparates); 3) Proposer une démarche systématique pour gérer ce changement en tenant compte de l'évolution du niveau de maturité de la compagnie; 4) Appliquer cette démarche à un cas concret pour une entreprise (développer l'orientation-client de son modèle d'affaires); 5) Montrer comment cette démarche peut s'appliquer concrètement au cas d'une entreprise.

Cette démarche, qui est l'apport principal de ce mémoire, a été développée en détails au chapitre 2 et se décline en deux cycles : un cycle de mise en œuvre d'une gestion stratégique du changement et un cycle d'opérationnalisation du changement. Chaque cycle est découpé en phases, elles-mêmes découpées en étapes. Nous avons élaboré cette démarche en nous appuyant sur une seconde revue littérature. Puis nous l'avons raffinée progressivement grâce à son application directe et de manière itérative en milieu pratique, en collaboration avec plusieurs intervenants de l'iA jusqu'à ce qu'elle soit suffisamment généralisée pour être présentée dans sa forme actuelle. Cette démarche a donc tiré parti de la complémentarité de nos recherches académiques (le milieu de recherche) et de l'expérience en milieu pratique (l'iA).

Un autre apport de ce mémoire a été de tester plus en profondeur, dans le cadre d'un projet pilote réalisé avec des équipes de l'iA, une phase essentielle de la gestion stratégique du changement que nous proposons. Ce projet pilote s'inscrit dans la phase de *mise en œuvre de la gestion stratégique du changement* de notre démarche. Il nous a permis de valider une partie du plan d'action stratégique adopté plus tôt par nos partenaires à l'iA (correspondant à la phase de *Préparation à la mise en œuvre* de notre démarche).

Pour cette phase d'expérimentation d'analyse de données (projet pilote), nous avons mis au point une deuxième démarche suivant le même principe d'élaboration que pour notre démarche principale de gestion stratégique du changement, à savoir : 1) Faire des recherches bibliographiques sur les démarches existantes; 2) Construire une première version de la démarche en s'appuyant sur les principes fondamentaux de ces démarches, mais en adressant spécifiquement le cas de l'iA; 3) Appliquer, tester et itérer chaque phase de la démarche proposée afin de la stabiliser 4) Généraliser notre approche pour qu'elle soit plus largement applicable. Cette démarche expérimentale est composée de 5 phases et a été appliquée pour guider un projet de croisement de données-clients internes de l'iA pour dégager des tendances sur les caractéristiques des clients et le type de relation que la clientèle entretient avec la compagnie.

Les responsables de l'iA ont jugé satisfaisants les résultats de la démarche expérimentale et du projet pilote en estimant qu'il avait permis d'atteindre l'objectif du plan d'action stratégique qui visait à mettre en évidence le potentiel de l'exploitation des données-clients et l'intérêt de l'analyse avancée des données de l'organisation. Cela contribuait aussi à renforcer la vision d'évolution stratégique adoptée antérieurement pour favoriser une plus grande maturité technologique et organisationnelle.

Cependant, les gestionnaires de l'iA qui étaient parties prenantes du projet ont estimé que ces résultats n'étaient pas encore suffisamment aboutis pour enclencher directement l'opérationnalisation du changement envisagé. En effet, un des volets de la vision stratégique était de développer au sein de l'organisation un niveau de maturité suffisant pour pouvoir exploiter les résultats des analyses en les intégrant aux processus d'affaires opérationnels et tactiques directement en lien avec la gestion de la clientèle. Or les gestionnaires ont estimé que le niveau de connaissances des secteurs d'affaires concernés en matière de segmentation de la clientèle n'était suffisant pour atteindre cet objectif à ce moment-là.

Ce bilan survenant à la fin du cycle stratégique prôné par notre démarche, a conduit à un léger changement dans notre démarche et que nous n'avions pas anticipé au départ. En effet, les responsables de l'iA ont décidé de lancer un second projet expérimental court et de nature plus opérationnelle. Dans notre démarche, ce deuxième projet pilote s'inscrit au début du cycle opérationnel et vise à tester les ajustements proposés pour le plan d'action en tenant compte des recommandations formulées au moment de l'évaluation finale du cycle stratégique. Ce second projet expérimental a été mené sur une période de deux semaines consécutives à la fin du projet de recherche et a été le fruit d'une étroite collaboration entre les équipes affaires des différents secteurs de l'iA et les équipes TI. A l'issue de cette mission, les résultats produits ont été à nouveau évalués et un autre bilan a été fait, qui a conduit cette fois-ci à déclencher officiellement le cycle d'opérationnalisation du changement. Ainsi, à la fin 2016 quelques semaines après l'achèvement du deuxième projet pilote, les gestionnaires exécutifs de l'iA ont mis en place une équipe entièrement dédiée à la réalisation de projets opérationnels pour le développement de nouveaux processus d'affaires exploitant les données-clients, en favorisant la collaboration entre les divers secteurs d'affaires de la compagnie.

Cette équipe a également eu pour mandat de continuer à faire évoluer la maturité technologique de l'iA en analyse avancée de données.

5.2 Perspectives sur la démarche et la problématique du mémoire

Tout d'abord, il est important de préciser que les principaux objectifs de ce mémoire ont été atteints. En effet, en premier lieu, les démarches de gestion du changement et de conduite du projet pilote que nous proposons, et qui constituent l'apport majeur de ce mémoire, présentent les principaux avantages suivants :

- Il s'agit d'approches qui rejoignent les principes et bonnes pratiques que l'on trouve dans la littérature
- Ces démarches ont été élaborées, raffinées et testées en plusieurs itérations afin de proposer des méthodes détaillées, illustrées par des actions qui conduisent à des résultats concrets et qui mettent en évidence certains enjeux ou difficultés qui nécessitent une attention particulière pour le lecteur qui souhaiterait appliquer ces méthodes
- Ces approches proposées ont été conçues pour s'adapter à la réalité des projets d'entreprises, que l'on pourrait qualifier de traditionnelles, qui sont similaires à l'iA en termes de maturité technologique et organisationnelle vis-à-vis de l'exploitation avancée de leurs données internes, tout en permettant de modifier, sauter ou répéter certaines étapes ou phases, sans remettre en question la validité de la démarche

Ensuite, dans le cas de l'iA, cette démarche a été transférée avec succès puisque qu'une équipe opérationnelle dédiée a été constituée et que plusieurs projets similaires ont été menés depuis la fin de notre projet de recherche. De plus, ce projet a contribué à entamer le cycle l'opérationnalisation du changement qu'iA souhaitait mettre en place au début de notre projet de recherche.

Néanmoins, la démarche proposée pourrait bénéficier de plusieurs améliorations pour adresser encore mieux la problématique de transformation digitale des organisations traditionnelles, notamment en ce qui concerne :

- La démarche théorique
- L'élaboration de la démarche expérimentale
- La façon dont elle a été appliquée

En effet, parmi les pistes d'améliorations possibles on pourrait envisager de tester à nouveau cette démarche de gestion du changement au cas d'autres entreprises dans les contextes suivants :

- Le contexte d'une entreprise plus avancée en termes de maturité technologique et organisationnelle que l'iA, mais qui n'est pas encore rendue au stade de concurrencer les compagnies entièrement digitalisées
- Le contexte d'une entreprise qui œuvre dans un autre secteur d'activité que l'iA

Pour conclure, ce mémoire nous tenons à préciser que la problématique de l'exploitation des données massives risque fortement d'évoluer d'ici quelques années. Deux principales visions de cette évolution peuvent alimenter de nouvelles réflexions.

La première est que d'ici quelques années, la plupart des compagnies traditionnelles auront été en mesure de transformer leurs modèles d'affaires en suivant les pas des leaders du marché. Dans cette perspective, il est fort à parier de nouvelles problématiques organisationnelles apparaîtront, notamment la disparition de certains métiers considérés par le passé comme non automatisables, et ceci grâce aux avancées technologiques en intelligence artificielle notamment. Par exemple, le métier de conseiller (financier, de service à la clientèle, de vente, etc.) pourrait fortement évoluer.

Une autre alternative plausible d'évolution due à cette révolution industrielle pourrait être que d'ici quelques années on voit émerger les limites du phénomène Big Data, de l'explosion de la production de données ainsi que de l'utilisation des renseignements personnels des individus à des fins de profits. On voit d'ailleurs déjà apparaître diverses initiatives soulevant des préoccupations éthiques liées à la « surveillance programmée » des individus par les compagnies privées. En effet, de nouvelles réglementations sont en cours de réflexion visant à limiter l'exploitation des données personnelles, qui pourraient remettre en question les promesses de bénéfiques du Big Data.

Bibliographie

- Alam, Md Imran, Manjusha Pandey, et Siddharth S. Rautaray. 2015. « A comprehensive survey on cloud Computing ». *International Journal of Information Technology and Computer Science (IJITCS)* 7 (2): 68.
- « Apprentissage automatique ». 2016. *Wikipédia*.
https://fr.wikipedia.org/w/index.php?title=Apprentissage_automatique&oldid=133154484.
- Baumard, Philippe. 1998. « Organiser la veille sur les nouvelles technologies de l'information ». CIGREF.
- Begoli, E., et J. Horey. 2012. « Design Principles for Effective Knowledge Discovery from Big Data ». Dans *2012 Joint Working IEEE/IFIP Conference on Software Architecture and European Conference on Software Architecture*, 215-18. doi:10.1109/WICSA-ECSA.212.32.
- Belk, Russell. 2014. « You are what you can access: Sharing and collaborative consumption online ». *Journal of Business Research* 67 (8): 1595-1600. doi:10.1016/j.jbusres.2013.10.001.
- Brashear, Thomas, Heiko Gebauer, et Christian Kowalkowski. 2012. « Customer-focused and service-focused orientation in organizational structures ». *Journal of Business & Industrial Marketing* 27 (7): 527–537.
- Brasseur, Christophe. 2013. *Enjeux et usages du Big Data: Technologies, méthodes et mise en oeuvre*. Lavoisier.
<https://books.google.ca/books?hl=fr&lr=&id=yPLAAgAAQBAJ&oi=fnd&pg=PA5&dq=Enjeux+et+Usages+Du+Big+Data+BRASSEUR+Christophe&ots=KNNI-jTX5Y&sig=Vs5sQH79mflGhxwGB7bJRpbwl4>.
- Butterfield, Andrew, et Gerard Ngondi Ekembe. 2016. « Dictionary of Computing ». <http://www.oxfordreference.com/acces.bibl.ulaval.ca/view/10.1093/acref/9780199234004.001.0001/acref-9780199234004>.
- Carlsson, Bo. 2004. « The Digital Economy: what is new and what is not? » *Structural change and economic dynamics* 15 (3): 245–264.
- Chugunov, Andrei V., Radomir Bulgov, Yuri Kabanov, George Kampis, et Maria Wimmer. 2016. *Digital Transformation and Global Society: First International Conference, DTGS 2016, St. Petersburg, Russia, June 22-24, 2016, Revised Selected Papers*. Vol. 674. Springer. <https://books-google-ca.acces.bibl.ulaval.ca/books?hl=fr&lr=&id=WteSDQAAQBAJ&oi=fnd&pg=PR5&dq=digital+era+society+transformation&ots=UdoeaZRfCP&sig=2Jqylhock7eUnRF39xSLfGjys04>.
- Davenport, Thomas. 2014. *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*. Harvard Business Review Press.
- Davenport, Thomas H., Jeanne G. Harris, et Robert Morison. 2010. *Analytics at Work: Smarter Decisions, Better Results*. Boston, Mass: Harvard Business Review Press.
- Fang, Shifeng, Yunqiang Zhu, Lida Xu, Jinqiu Zhang, Peiji Zhou, Kan Luo, et Jie Yang. 2017. « An integrated system for land resources supervision based on the IoT and cloud computing ». *Enterprise Information Systems* 11 (1): 105-21. doi:10.1080/17517575.2015.1086816.
- Fogelman-Soulié, Françoise, et Wenhuan Lu. 2016. « Implementing Big Data Analytics Projects in Business ». Dans *Big Data Analysis: New Algorithms for a New Society*, édité par Nathalie Japkowicz et Jerzy Stefanowski, 141-58. *Studies in Big Data* 16. Springer International Publishing. http://link.springer.com/acces.bibl.ulaval.ca/chapter/10.1007/978-3-319-26989-4_6.
- Fullerton, Gordon. 2014. « The moderating effect of normative commitment on the service quality-customer retention relationship ». *European Journal of Marketing* 48 (3/4): 657–673.
- Harmon, Robert R. 2003. « Marketing Information Systems A2 - Bidgoli, Hossein ». Dans *Encyclopedia of Information Systems*, 137-51. New York: Elsevier. <http://www.sciencedirect.com/science/article/pii/B0122272404001106>.
- Hollander, D., K. Hertz, et B. H. Wassink. 2013. *The Journey Toward Greater Customer Centricity*. Ernst&Young.
- Hurwitz, Judith, Alan Nugent, Fern Halper, et Marcia Kaufman. 2013. *Big Data For Dummies*. Hoboken, NJ: For Dummies.

- Jensen, Michael C. 1993. « The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems ». *The Journal of Finance* 48 (3): 831-80. doi:10.1111/j.1540-6261.1993.tb04022.x.
- Keim, D. A. 2002. « Information visualization and visual data mining ». *IEEE Transactions on Visualization and Computer Graphics* 8 (1): 1-8. doi:10.1109/2945.981847.
- Kwakkel, Jan H., Stephen Carley, John Chase, et Scott W. Cunningham. 2014. « Visualizing geo-spatial data in science, technology and innovation ». *Technological Forecasting and Social Change* 81 (janvier): 67-81. doi:10.1016/j.techfore.2012.09.007.
- Lake, Peter, et Robert Drake. 2014. « The Future of IS in the Era of Big Data Big Data ». Dans *Information Systems Management in the Big Data Era*, 267-88. Advanced Information and Knowledge Processing. Springer International Publishing. http://link.springer.com/acces.bibl.ulaval.ca/chapter/10.1007/978-3-319-13503-8_12.
- Lamberti, Lucio. 2013. « Customer centricity: the construct and the operational antecedents ». *Journal of Strategic Marketing* 21 (7): 588-612. doi:10.1080/0965254X.2013.817476.
- Liu, LING, et M. TAMER Özsu, éd. 2009. « Encyclopedia of Database Systems ». *Encyclopedia of Database Systems*. Springer US. http://link.springer.com/acces.bibl.ulaval.ca/referenceworkentry/10.1007/978-0-387-39940-9_2557.
- Manyika, James, Michael Chui, Jacques Bughin, Richard Dobbs, Peter Bisson, et Alex Marrs. 2016. « Disruptive technologies: Advances that will transform life, business, and the global economy | McKinsey & Company ». Consulté le décembre 20. <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/disruptive-technologies>.
- Marketline. 2016. « Insurance in Canada ». 0070-2087.
- Mayer-Schönberger, Viktor, et Kenneth Cukier. 2013. *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt. <https://books-google-ca.acces.bibl.ulaval.ca/books?hl=fr&lr=&id=uy4lh-WEhhIC&oi=fnd&pg=PP1&dq=Big+data+&ots=Jsk5bhHPNP&sig=v6XEszMr2BkOi4R5PFTWal-8GfQ>.
- Misra, Sumit, Sanjoy Kumar Saha, et Chandan Mazumdar. 2013. « Performance comparison of hadoop based tools with commercial etl tools—a case study ». Dans *International Conference on Big Data Analytics*, 176–184. Springer. http://link.springer.com/acces.bibl.ulaval.ca/chapter/10.1007/978-3-319-03689-2_12.
- Morabito, Vincenzo. 2015. « Big Data Governance ». Dans *Big Data and Analytics*, 83-104. Springer International Publishing. http://link.springer.com/acces.bibl.ulaval.ca/chapter/10.1007/978-3-319-10665-6_5.
- « Online Analytical Processing ». 2016. *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Online_analytical_processing&oldid=757186245.
- Pappas, Lisa, et Lisa Whitman. 2011. « Riding the Technology Wave: Effective Dashboard Data Visualization ». Dans *Human Interface and the Management of Information. Interacting with Information*, 249-58. Springer, Berlin, Heidelberg. http://link.springer.com/acces.bibl.ulaval.ca/chapter/10.1007/978-3-642-21793-7_29.
- Rocard, Michel. 2014. « Au secours, ma boîte se digitalise ! » *lesechos.fr*. http://archives.lesechos.fr/archives/cercle/2014/07/17/cercle_106182.htm.
- Rota, Véronique. 2009. *Gestion de projet*. Eyrolles. http://www.eyrolles.com/Chapitres/9782212121650/pp-37-55_Rota.pdf.
- Schmidhuber, Jürgen. 2013. « My first deep learning system of 1991+ deep learning timeline 1962-2013 ». *arXiv preprint arXiv:1312.5548*. <http://arxiv.org/abs/1312.5548>.
- Shmueli, Galit, Nitin R. Patel, et Peter C. Bruce. 2016. *Data Mining for Business Analytics: Concepts, Techniques, and Applications in XLMiner*. John Wiley & Sons. https://books-google-ca.acces.bibl.ulaval.ca/books?hl=fr&lr=&id=PlvWCgAAQBAJ&oi=fnd&pg=PR21&dq=data+mining+concepts+and+techniques&ots=anldTo4vc1&sig=84OFq2ymxXliFCd9niwB6X_vRGw.
- Solove, Daniel J. 2004. *The digital person: Technology and privacy in the information age*. NyU Press. <https://books-google->

- ca.acces.bibl.ulaval.ca/books?hl=fr&lr=&id=CuugBwAAQBAJ&oi=fnd&pg=PR11&dq=digital+era++privacy+protection&ots=tWZjTBR9_h&sig=pHK0dbQn71N1HbH_6lk3F7qa6PQ.
- Stonebraker. 1999. *Readings in Database Systems*. 3 edition. San Francisco: Morgan Kaufmann.
- Tapscott, Don. 1996. *The digital economy: Promise and peril in the age of networked intelligence*. Vol. 1. McGraw-Hill New York. <http://http-server.carleton.ca/~aramirez/5401/BookReviews/Babu.pdf>.
- Telea, Alexandru C. 2014. *Data visualization: principles and practice*. CRC Press. https://books-google-com.acces.bibl.ulaval.ca/books?hl=fr&lr=&id=AGjOBQAAQBAJ&oi=fnd&pg=PP1&dq=data+visualizati on+principles&ots=NjCru1-Gk&sig=F-mpocAqgBokrmQC25_rT4Obfjs.
- Teste, Olivier. 2009. « Modélisation et manipulation des systèmes OLAP: de l'intégration des documents à l'utilisateur ». Université Paul Sabatier-Toulouse III. <https://tel.archives-ouvertes.fr/tel-00479460/>.
- Trestian, R., P. Shah, H. Nguyen, Q. -T. Vien, O. Gemikonakli, et B. Barn. 2017. « Towards connecting people, locations and real-world events in a cellular network ». *Telematics and Informatics* 34 (1): 244-71. doi:10.1016/j.tele.2016.05.006.
- Vooght, Niguel, Andrew Jurczynski, et Andrew Nevin. 2015. « Project Blue 2.0: Redefining financial services ».
- Wiederschain, G. Ya. 2011. « Data Mining Techniques for the Life Sciences ». *Biochemistry (Moscow)* 76 (4): 494-494. doi:10.1134/S0006297911040158.
- Wikipedia. 2016. « Social Transformation ». *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Social_transformation&oldid=724135974.
- Wikipédia. 2016. « Appareil mobile ». *Wikipédia*. https://fr.wikipedia.org/w/index.php?title=Appareil_mobile&oldid=132833475.
- Wu, X., X. Zhu, G. Q. Wu, et W. Ding. 2014. « Data mining with big data ». *IEEE Transactions on Knowledge and Data Engineering* 26 (1): 97-107. doi:10.1109/TKDE.2013.109.
- Xu, B., L. D. Xu, H. Cai, C. Xie, J. Hu, et F. Bu. 2014. « Ubiquitous Data Accessing Method in IoT-Based Information System for Emergency Medical Services ». *IEEE Transactions on Industrial Informatics* 10 (2): 1578-86. doi:10.1109/TII.2014.2306382.
- Yana. 2016. « Middle Tennessee Natural Gas Utility District Maps, Manages Network with Esri ArcGIS ». Consulté le décembre 31. <http://geospatial-news.blogspot.com/2011/10/middle-tennessee-natural-gas-utility.html>.
- Zaharia, Matei, Reynold S. Xin, Patrick Wendell, Tathagata Das, Michael Armbrust, Ankur Dave, Xiangrui Meng, et al. 2016. « Apache Spark: a unified engine for big data processing ». *Communications of the ACM* 59 (11): 56-65.
- Zhang, Shuai, Shufen Zhang, Xuebin Chen, et Xiuzhen Huo. 2010. « Cloud computing research and development trend ». Dans *Future Networks, 2010. ICFN'10. Second International Conference on*, 93-97. Ieee. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5431874.
- Zudilova-Seinstra, Elena, Tony Adriaansen, et Robert van Liere. 2014. *Trends in interactive Visualization*. Springer. <http://link.springer.com.acces.bibl.ulaval.ca/content/pdf/10.1007/978-1-84800-269-2.pdf>.