



Elen Grigoryan

L'efficacité d'usage du big data en entreprise en contexte arménien

GRIGORYAN, Elen. *L'efficacité d'usage du big data en entreprise en contexte arménien* **[en ligne]**. Sous la codirection de Katia LOBRE-LEBRATY et de Jean-Fabrice LEBRATY. Thèse de doctorat : Sciences de gestion. Lyon : Université Jean Moulin Lyon 3, 2022.

Disponible sur : <https://www.theses.fr/2022LYO30028>



Document diffusé sous le contrat *Creative Commons* « **Attribution – Pas d'utilisation commerciale - Pas de modification** »

Vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.



N° d'ordre NNT : 2022LYO30028

THÈSE DE DOCTORAT DE L'UNIVERSITÉ JEAN MOULIN LYON 3

Membre de l'université de Lyon

École doctorale n° 486 - Sciences économiques et de gestion

Discipline :
Sciences de gestion

Soutenue publiquement le 20/10/2022, par

Elen GRIGORYAN

L'efficacité d'usage du big data en entreprise en contexte arménien

Laboratoire de recherche : **Centre de recherche de l'IAE de Lyon - Magellan**

Codirecteurs de thèse : **Mme Katia LOBRE-LEBRATY** et **M. Jean-Fabrice LEBRATY**

Devant le jury composé de :

Mme Katia LOBRE-LEBRATY

Maîtresse de conférences HDR, université Jean Moulin Lyon 3. Codirectrice de thèse

M. Jean-Fabrice LEBRATY

Professeur des universités, université Jean Moulin Lyon 3. Codirecteur de thèse

Mme Marie-Christine CHALUS-SAUVANNET

Professeure des universités, université Jean Moulin Lyon 3. Présidente du jury

M. Marc BIDAN

Professeur des universités, Nantes université. Rapporteur

M. Ivan PASTORELLI

Maître de conférences HDR, université Côte d'Azur, Nice. Rapporteur

M. Aram YESAYAN

Maître de conférences, université française en Arménie, Erevan. Examineur

Résumé

Ce travail de recherche s'intéresse aux aspects managériaux du phénomène de big data. Il traite la question suivante : quels sont les facteurs qui contribuent le plus à l'efficacité d'usage du big data en entreprise en contexte arménien ? Cette question est examinée sur le terrain à partir du modèle conceptuel d'efficacité d'usage du big data proposé par Surbakti et ses coauteurs (Surbakti, Wang, Indulska, & Sadiq, 2019). Ce travail présente un corpus méthodologique pour mener une recherche contextualisée en sciences de gestion à partir des trois pôles de contextualisation, proposés par Livian (2020). La démarche principale prend appui sur des données quantitatives, issues d'une enquête auprès de 211 professionnels du domaine. Elle mobilise aussi des données de type qualitatif à partir de 10 entretiens réalisés. L'analyse statistique et l'analyse thématique des données met en évidence l'existence de deux groupes différents d'entreprises parmi les utilisateurs du big data : « utilisateur novice » et « utilisateur avancé ». Ainsi, la thèse présente l'ensemble des facteurs qui contribuent à l'efficacité d'usage du big data pour chacun de ces deux types d'utilisateurs et discute les particularités contextuelles arméniennes liés à ces résultats.

Mots clés : Big data, efficacité d'usage, contexte arménien, efficacité d'usage du big data, utilisation, usage, utilisateur novice, utilisateur avancé, recherche contextualisée.

Abstract

This research focuses on the managerial aspects of big data phenomenon. The research question is therefore: what factors contribute the most to effective use of big data in companies in the Armenian context? This question is examined through a conceptual model of effective use of big data (Surbakti et al., 2019). Accordingly, this work proposes a methodological basis for conducting contextualized research in management sciences through the three poles of contextualization, proposed by Livian (2020). The main approach is based on quantitative data, from 211 surveys, completed by the professionals in the field. It also mobilizes qualitative data from 10 interviews conducted. Statistical analysis and thematic analysis of the data highlights the existence of two different groups of companies among big data users: “novice user” and “advanced user”. Thus, the thesis presents all the factors that contribute to the effective use of big data for each of these two types of users and discusses the Armenian contextual particularities linked to these results.

Keywords: Big data, effective use, Armenian context, effective use of big data, utilization, use, novice user, advanced user, contextualized research.

Remerciements

Je tiens à remercier l'ensemble des personnes qui m'ont aidée à aller jusqu'au bout de ce parcours passionnant qu'est la thèse.

En tout premier lieu, mes sincères remerciements sont adressés à mes directeurs de thèse, Madame Katia Lobre-Lebraty, Maître de Conférences HDR et le Professeur Jean-Fabrice Lebraty. Leur confiance, leur intérêt, leur disponibilité, leur patience mais aussi leurs conseils éclairés au fil de ces quatre années de thèse, m'ont permis d'achever ce travail. Je les remercie ainsi pour leurs qualités humaines qui ont largement dépassé l'encadrement doctoral. Je garderai à jamais un tout profond respect pour ces personnes.

Ce travail n'aurait pas été possible sans le soutien de l'Université française en Arménie (UFAR), qui a financé ce projet doctoral. J'exprime ma sincère reconnaissance au Professeur Jean-Marc Lavest, l'ex-Recteur de l'UFAR, pour son dévouement au lancement du partenariat entre l'Université Jean Moulin Lyon 3 et l'Université française en Arménie au niveau de cette formation doctorale.

Je tiens à remercier également le Professeur Bertrand Venard, le Recteur de l'UFAR, pour ses précieux conseils, son soutien continu, son exigence et son engagement.

Mes sincères remerciements vont au Professeur Marc Bidan pour ses conseils constructifs lors de la réunion du comité de suivi de thèse. Je le remercie également d'avoir accepté d'être rapporteur de ce manuscrit.

Je témoigne également ma gratitude à Monsieur Ivan Pastorelli, Maître de conférences habilité à diriger des recherches, d'avoir accepté d'être rapporteur de cette thèse.

J'adresse ainsi mes remerciements au Professeur Marie-Christine Chalus-Sauvannet et à Monsieur Aram Yesayan, Maître de conférences, pour avoir accepté de participer au jury de thèse.

Je souhaite remercier Madame Barbara Gandriaux, l'ex-secrétaire générale de l'UFAR, pour son accompagnement professionnel, pour son aide, ses encouragements et sa confiance. Je lui suis ainsi reconnaissante pour la relecture du manuscrit de la thèse.

Mes pensées vont à l'ensemble des professionnels d'avoir accepté de prendre de leur temps pour participer à nos études et aux entretiens.

Je remercie également les membres de mon laboratoire de recherche, Magellan, dirigé par Monsieur Jean-François Gajewski, pour leur soutien, leur accompagnement et leur accueil pendant mes séjours doctoraux à Lyon.

Je remercie ainsi les membres de l'École Doctorale Sciences Économiques et de Gestion, dirigée par Madame Isabelle Royer, pour leur soutien professionnel et leur accompagnement au fil de ce parcours doctoral.

J'exprime ma reconnaissance à Madame Anaida Gasparyan, le responsable de la chaire de français à l'UFAR, ainsi qu'à Roza, Anna et Raphaël pour la relecture de la thèse.

Je tiens aussi à remercier mes proches, tout particulièrement Mariné et Emmanuelle, pour leur soutien et leur accompagnement pendant cette expérience.

A ma mère, Nounée et mon père Artsruni, merci pour votre soutien sans faille, votre bienveillance et votre confiance qui m'ont permis de m'engager dans ce parcours et de le mener à bien.

A mon mari, Gor, merci pour ta patience, tes encouragements et ton soutien qui ont été constants pendant ces années.

A ma sœur, Anahit, merci pour ton écoute et ta présence quand j'en avais besoin.

Enfin, je veux remercier mes enfants, ma fille Yéva et mon fils Tigran. Leur amour, leurs sourires et leur tendresse ont été ma plus forte source de motivation et d'énergie tout au long de ce parcours.

Liste des abréviations

AST : Adaptive Structuration Theory

AWS : Amazon Web Services

BD : Big data

BI : Business Intelligence

BPO : Business Process Orientation

CSP : Catégories socioprofessionnelles

DBMS : Database Management Systems

EU : Efficacité d'usage

GPS : Global Positioning System

IA : Intelligence Artificielle

ISURA : IS use-related activity

MG : Mécanismes Générateurs

ML : Machine Learning (Apprentissage Machine)

RBMS : Relational Database Management System

RH : Ressources humaines

SI : Système d'information

TAM : Technology Acceptance Model

TEU : Théorie de l'efficacité d'usage

TI : Technologie de l'information

TR : Théorie de Représentation

TTF : Task Technology Fit

URRS : Union des républiques socialistes soviétiques

UTAUT : Unified Theory of Acceptance and Use of Technology

Sommaire

Introduction générale	17
Première partie : Le choix d'un cadre théorique au big data : L'efficacité d'usage.....	25
Chapitre 1 Utilisation du big data par les entreprises	27
1.1 La genèse et les définitions du big data.....	28
1.1.1 Définition du big data à partir de ses caractéristiques spécifiques	30
1.1.2 Définition du big data par sa dimension technologique.....	34
1.1.3 Autres dimensions pour définir le big data	35
1.2 Les principales sources du big data.....	37
1.3 L'usage et le traitement du big data	39
1.3.1 Le traitement du big data	40
1.3.2 Les applications du big data.....	42
1.3.3 Les systèmes de stockage du big data.....	50
1.3.4 Traitement du big data par la société JKL Armenia	52
Chapitre 2 Efficacité d'usage: Fondements théoriques	55
2.1 Travaux fondateurs.....	55
2.2 Efforts de définition d'usage du système	56
2.3 Théorie de l'efficacité d'usage (TEU)	57
2.3.1 Théorie générale du SI.....	57
2.3.2 Visions de développement de TEU.....	60
2.3.3 Efficacité d'usage d'un SI : Développement de théorie à 3 étapes.....	61
2.3.4 Développement de la discussion sur la théorie d'efficacité d'usage	66
2.3.5 La place de la théorie d'efficacité d'usage dans les théories existantes	70
2.3.6 Contribution de Burton-Jones et Grange (Burton-Jones & Grange, 2013)	76
Chapitre 3 Efficacité d'usage du big data : Modèle conceptuel mobilisé.....	79
3.1 Thème 1 : Intérêt organisationnel perçu.....	81
3.1.1 Valeur perçue	81
3.1.2 Risques perçus	82
3.1.3 Efficacité perçue	84

3.1.4	Facilité perçue de l'utilisation.....	85
3.1.5	Intention à utiliser	85
3.1.6	Observabilité perçue	86
3.1.7	Analyse coût-avantage	87
3.2	Thème 2: Management des processus	88
3.2.1	Orientation processus.....	88
3.2.2	Intégration du business processus en TI	89
3.2.3	Management des données	90
3.3	Thème 3: Confidentialité, sécurité et gouvernance des données	91
3.3.1	Confidentialité et sécurité	91
3.3.2	Gouvernance des données.....	92
3.4	Thème 4: Qualité des données	93
3.4.1	Qualité des données et qualité de l'information	93
3.4.2	Complétude des données.....	94
3.4.3	Actualité des données	94
3.4.4	Accès aux données	95
3.4.5	Pertinence des données	96
3.4.6	Exactitude des données	96
3.4.7	Cohérence des données	96
3.5	Thème 5 : Aspect Humain.....	97
3.5.1	Savoirs et compétences des individus.....	97
3.5.2	Confiance	99
3.5.3	Champions	99
3.5.4	Engagement du personnel	100
3.5.5	Participation des utilisateurs	100
3.5.6	Caractéristiques individuelles	101
3.6	Thème 6 : Aspect organisationnel.....	101
3.6.1	Compétence organisationnelle et culturelle	102
3.6.2	Gestion du talent	103
3.6.3	Programme de gestion du changement	103
3.6.4	Alignement stratégique	104

3.6.5	Management du projet	105
3.6.6	Management de la performance.....	105
3.6.7	Structure organisationnelle et taille	106
3.6.8	Collaboration interdépartementale.....	106
3.6.9	Communication.....	107
3.6.10	Soutien de la haute direction.....	108
3.6.11	Effet environnemental.....	108
3.6.12	Objectifs précis	109
3.6.13	Accent sur l'innovation.....	109
3.7	Thème 7 : Systèmes, outils et technologies.....	110
3.7.1	Qualité du système.....	110
3.7.2	Infrastructure TI.....	112
3.7.3	Support Fournisseurs	113
Seconde partie : Les enseignements de l'efficacité d'usage du big data dans les entreprises arméniennes		115
Chapitre 4 Positionnement épistémologique et méthodologie de la recherche		117
4.1	Positionnement épistémologique et paradigme.....	117
4.1.1	Positionnements épistémologiques et trois principaux paradigmes en SI	117
4.1.2	Paradigme de notre travail de recherche.....	119
4.2	Méthodologie de la recherche	127
4.2.1	Présentation de la question de recherche à partir de six paramètres de Noël (2012) 128	
4.2.2	Deux volets de la méthodologie de recherche	131
Chapitre 5 Terrain de recherche.....		153
5.1	Pourquoi faut-il contextualiser une recherche en science de gestion ?	153
5.1.1.	Contextualisation théorique	154
5.1.2.	Contextualisation empirique	155
5.2.	Profil de l'Arménie : Contextualisation à 3 pôles de Livian (2020)	156
5.2.1.	Arménie : une brève présentation	156
5.2.2.	Trois pôles de contextualisation de Livian (2020).....	157

5.2.3.	Environnement technologique	168
5.3.	Présentation détaillée des acteurs du terrain	174
5.3.1.	Première étude exploratoire : Étude d'un expert en utilisation du big data.....	175
5.3.2.	Étude quantitative : profils des participants.....	177
5.3.3.	Deux cas bien précis : « utilisateur novice » et « utilisateur avancé ».....	184
Chapitre 6 Résultats		191
6.1	Résultat 1: « Utilisateur novice » et « utilisateur avancé »	192
6.2	Résultat 1 bis : Société ABC, « utilisateur avancé » du big data	193
6.2.1	F1 : Intérêt organisationnel perçu	193
6.2.2	F2 : Management des processus	196
6.2.3	F3 : Confidentialité, sécurité et gouvernance des données	196
6.2.4	F4 : Qualité des données	197
6.2.5	F5 : Aspect humain	198
6.2.6	F6 : Aspect organisationnel	199
6.2.7	F7 : Systèmes, outils et technologies	200
6.3	Résultat 2 : Deux modèles différents de l'efficacité d'usage du big data en entreprise en contexte arménien	201
6.3.1	Modèle d'efficacité d'usage du big data pour un « utilisateur novice ».....	201
6.3.2	Modèle d'efficacité d'usage du big data pour un « utilisateur avancé »	203
6.4	Résultat 2 bis : Efficacité d'usage du big data chez une entreprise « utilisateur avancé » et « utilisateur novice »	204
6.4.1	CDE, « utilisateur avancé »	204
6.4.2	BCD Arménie, « utilisateur novice ».....	208
6.5	Résultats 3 : L'existence de deux types d'utilisateurs ne dépend pas du contexte d'examen	210
6.6	Résultat 3 bis : Particularités contextuelles arméniennes.....	211
6.7	Conclusion des résultats	213
Chapitre 7 Conclusion et discussion		215
7.1	Réponse à la question de recherche.....	215
7.1.1	Le cas de l'entreprise « utilisateur novice ».....	216
7.1.2	Le cas de l'entreprise « utilisateur avancé »	218

7.2	Trois axes de discussion.....	220
7.2.1	Prolongement des résultats	220
7.2.2	Place du contexte arménien dans cette étude	223
7.2.3	Mise en perspective de notre recherche	226
7.3	Contributions théoriques et managériales de la thèse	238
7.3.1	Contributions théoriques.....	238
7.3.2	Contributions managériales	240
7.4	Limites et pistes de recherche	242
	Table des illustrations	245
	Bibliographie.....	247
	Annexes.....	285

Introduction générale

« Vous connaissez évidemment la fameuse histoire publiée dans le New York Times, qui racontait comment les algorithmes de Target, un détaillant américain bien reconnu, avaient repéré quasiment avant tout le monde, le fait qu'une adolescente de 16 ans soit enceinte... ». Relevé par A.M., directeur recherche et développement chez Technologies X5¹, cet exemple illustre parfaitement un cas classique d'usage du big data dans le monde du commerce prédictif.

Dans cet exemple bien précis, nous voulons souligner quelques aspects importants :

- Utilisation par l'entreprise de données détenues sur les consommateurs pour améliorer son offre (la valeur pour Target) (Batistič & van der Laken, 2019) ;
- Information issue des données de quantité et de nature très hétérogène (Ghasemaghaei, 2021);
- Outillage nécessaire en matière d'infrastructure technologique, de techniques et de méthodes mises en place pour extraire du big data une information utile. Autrement dit, il s'agit d'information actionnable qui permettra d'apporter une solution efficace au problème posé (Technologie et Méthodes) (Fotaki, Voudouris, Lioukas, & Zyglidopoulos, 2021), (Hashem, Chang, et al., 2016).

La réflexion sur ces trois aspects nous renvoie à une conclusion intermédiaire : la production de la valeur au moyen du traitement des données sur les consommateurs n'est pas récente. Depuis les années 1980, elle est même au cœur des outils de marketing (étude, segmentation, ciblage...) (Giannelloni & Le Nagard, 2016).

Par conséquent, en quoi consiste le changement induit par le big data ?

Premièrement, il s'agit de la variété des caractéristiques spécifiques de ces données (Volume, Vitesse, Variété, Vérité, Valeur, Variabilité et Visualisation (Seddon & Currie, 2017). Deuxièmement c'est l'outillage nécessaire (en matière d'infrastructure technologique, de

¹ Société « X5 Technologies », localisée dans la Fédération de Russie

techniques, de méthodes et de ressources) permettant une agrégation efficace et le rendement intelligible de ces données (Ghasemaghaei, 2019a).

On peut aisément imaginer que ces particularités peuvent modifier potentiellement les règles du jeu liées au management et l'utilisation efficace du big data par les entreprises (Jha, Agi, & Ngai, 2020).

Ainsi, pour une utilisation efficace du big data, les entreprises font face aujourd'hui à de multiples enjeux technologiques, managériaux et économiques (Liu, Wang, & Zhang, 2022), (Calic & Ghasemaghaei, 2021), (Maroufkhani, Tseng, Iranmanesh, Ismail, & Khalid, 2020), (Mikalef, Pappas, Krogstie, & Pavlou, 2020), (Elisabetta Raguseo, Vitari, & Pigni, 2020), (Wamba, Dubey, Gunasekaran, & Akter, 2020), (Karoui, Davauchelle, & Duzert, 2014), (Gantz, Reinsel, & Shadows, 2012).

Ce travail de thèse s'intéresse principalement aux aspects managériaux du phénomène de big data au sein d'entreprises, dans un contexte arménien.

Plus particulièrement, il s'agit d'étudier les facteurs qui contribuent le plus à l'efficacité d'usage du big data en entreprise dans ce contexte particulier. Pour ce faire, nous proposons d'y appliquer un modèle conceptuel de l'efficacité d'usage du big data, développé par Surbakti et ses coauteurs (2019) (Annexe 1).

Le modèle traite soigneusement 41 facteurs regroupés dans les 7 thématiques ; « intérêt organisationnel perçu », « management des processus », « confidentialité, sécurité et gouvernance des données », « qualité des données », « aspect humain », « aspect organisationnel » et « systèmes, outils et technologies ». Dans le cadre de ce travail de thèse, notre objectif principal est de construire et appliquer une démarche méthodologique cohérente pour faire tester ce modèle. À notre connaissance, il s'agit du premier test empirique de ce modèle conceptuel.

Dans cette thèse, nous examinerons dans un premier temps les fondements théoriques du concept de l'efficacité d'usage d'un système d'information (SI) (Burton-Jones & Grange, 2013). Ces fondements nous permettront ensuite de construire une compréhension complète du modèle conceptuel de l'efficacité d'usage dans le contexte du big data (Surbakti et al., 2019).

Au plan académique, cette problématique tend à développer la théorie d'efficacité d'usage au niveau organisationnel, plus particulièrement il démontre comment l'efficacité d'usage peut être empiriquement explorée et contextualisée dans un domaine particulier, celui du big data. Pour ce faire, nous viserons :

- Une opérationnalisation du construit « d'efficacité d'usage » dans un contexte particulier du big data, ce qui est aussi une des recommandations pour les futures recherches proposées par Surbakti et ses coauteurs (2019);
- Une démarche méthodologique détaillée pour mener une recherche contextualisée, qui utilise les axes proposés et adaptés de Livian (2020) ;
- Un outillage nécessaire pour étudier l'efficacité d'usage au niveau organisationnel (instrument de mesure, guides d'entretien) ;
- Une amélioration d'état de la connaissance sur le sujet de l'usage du big data en entreprise à travers le modèle proposé par (Surbakti et al., 2019) qui permet deux adaptations importantes.

Au plan managérial, la problématique de ce travail permet une compréhension plus fine et pluridimensionnelle (non seulement axée sur les aspects technologiques, mais aussi sur ceux de nature organisationnelle et motivationnelle) de l'utilisation efficace du big data par les entreprises dans un contexte arménien. L'évaluation de l'impact des facteurs du modèle propose une image claire sur leur potentiel à contribuer à l'efficacité d'usage du big data en entreprise en contexte étudié. Son apport principal ainsi consiste dans la mise en évidence de l'importance de prise en compte des facteurs, qui contribuent le plus à l'efficacité d'usage du big data en entreprise dans ce contexte particulier.

Selon Bachelard « *Pour un scientifique toute connaissance est une réponse à une question* » (Benoit-Smullyan & Bachelard, 1940). Nous suivons cette philosophie, et proposons de traiter la problématique présentée au travers de la question de recherche suivante :

« Quels sont les facteurs qui contribuent le plus à l'efficacité d'usage du big data en entreprise en contexte arménien » ?

Nous avons formulé cette question de recherche à partir des six paramètres conseillés de Noël (2012) : clarté ; faisabilité ; pertinence ; interrogation ; auditoire pour les résultats et intelligence. Ces aspects sont soigneusement interprétés dans le chapitre 4 de la thèse.

Du point de vue académique, cette question de recherche a pour intérêt de mettre à l'épreuve le modèle conceptuel de Surbakti et al. (2019) et ajouter à la littérature existante plusieurs volets :

- Mettre en évidence, de manière empirique, une valorisation des facteurs composants du modèle conceptuel (Surbakti et al., 2019) et proposer une classification hiérarchique des facteurs par leur importance à contribuer à l'efficacité d'usage du big data en entreprise en contexte arménien ;
- Permettre et explorer un contexte technologique particulier arménien à partir des 3 pôles de contextualisation proposés par Livian (2020) ;
- Soumettre les résultats issus des études réalisées pour une analyse comparative avec des études précédentes sur le sujet.

La démarche méthodologique adoptée dans ce travail de thèse a pour but d'approcher au plus près la réalité en multipliant les méthodes (Allard-Poesi & Perret, 2014). L'évidence qualitative, basée sur les résultats issus principalement des entretiens, nous sert comme *une source de richesse* (Weber, 2004) tout au long de notre travail de recherche. La démarche principale prend appui sur des données quantitatives issues des enquêtes auprès des professionnels du domaine. L'analyse statistique de ces données nous permet de proposer une réponse complète à la question de recherche.

Du point de vue managérial, la question de recherche envisage plusieurs implications importantes :

- Quelle que soit la granularité de compréhension du sujet, les praticiens du domaine et les managers de différents niveaux peuvent profiter de ces connaissances. Elle constitue donc un levier pour une meilleure maîtrise de la gestion d'un projet du big data, aux différents stades de son développement ;

- L'outil principal de l'enquête peut être utilisé par les gestionnaires comme un instrument d'évaluation de leur niveau d'efficacité d'usage du big data ;
- Pour les gestionnaires en Arménie, en particulier, les résultats des études réalisées fournissent une information riche sur les deux types d'utilisateur du big data ; compte tenu des particularités contextuelles.

Notre travail de recherche se focalise principalement sur les praticiens qui travaillent dans les 16 entreprises localisées en Arménie, qui exploitent potentiellement du big data. Parmi ces entreprises, sept sont de grandes entreprises (plus de 250 salariés) qui ont adopté le big data au cours de leur activité principale ; les autres sont des utilisateurs avancés du big data : l'activité de ces sociétés est basée principalement sur l'exploitation du big data.

Dans le cadre de ce travail de recherche, nous partageons la vision de Louitri et Sahraoui (2014) qui notent bien qu'une idée comme « toute chose est égale par ailleurs » peut biaiser les résultats des recherches. Nous sommes aussi d'accord dans l'idée que l'apport majeur managérial d'un travail consiste dans son actionnabilité pour le contexte étudié. Ainsi, nous trouvons primordial de prendre en compte des particularités contextuelles dans cette étude en science de gestion où la production des connaissances contextualisées est souvent liée aux spécificités du contexte d'examen.

Mener une étude dans la thématique technologique au contexte arménien porte un intérêt particulier :

L'Arménie, dotée d'une recherche scientifique dont les niveaux de structuration et de maturité peuvent être considérés comme insuffisants, connaît un essor dans l'industrie technologique, dont le potentiel de développement reste prometteur et affiche des avancées majeures.

Ce paradoxe intéressant nous renvoie au fait de subdiviser la contextualisation en deux parties ; la contextualisation théorique et la contextualisation empirique, proposés par D.S. Bentaleb et C. Bentaleb (2020). Par la suite, nous traiterons le contexte à partir des trois pôles de contextualisation conseillés par Livian (2020). Enfin, nous présenterons l'environnement technologique arménien qui complétera la description du contexte étudié.

Ce travail de thèse est composé de deux parties principales. La première partie détaille le sujet de la thèse et traite du cadre conceptuel. La deuxième partie traite les enseignements de l'efficacité d'usage du big data dans les entreprises arméniennes.

La première partie du travail se compose de trois chapitres portant principalement sur les fondements théoriques de l'efficacité d'usage du big data. Nous partons des bases d'une théorie générale (Burton-Jones & Grange, 2013) et mettons en évidence son application dans un contexte spécifique, celui du big data (Surbakti et al., 2019). Le phénomène du big data et son usage par les entreprises sont détaillés en particulier.

Le premier chapitre traite de la genèse et des définitions du big data afin de construire une compréhension fine de son utilisation par les entreprises. Les aspects techniques sont détaillés, en mettant l'accent sur les enjeux managériaux. Enfin un exemple est exposé pour illustration.

Le deuxième chapitre présente les fondements théoriques du concept de l'efficacité d'usage, définis et développés préalablement pour un système d'information général. Nous traitons soigneusement les principaux modèles et les cadres théoriques sur lesquels a été construite la théorie de l'efficacité d'usage.

Le troisième chapitre se concentre sur la présentation du construit de l'efficacité d'usage du big data et détaille le modèle conceptuel de Surbakti et ses coauteurs (2019). Une revue de la littérature sur les 41 facteurs composants du modèle conceptuel a été consolidée.

La deuxième partie du travail comprend quatre chapitres, présentant les éléments du cadre méthodologique appliqués afin d'apporter une réponse à la question de recherche proposée. Enfin, une discussion sur les interprétations des résultats et la réponse à la question de recherche sont exposées.

Le quatrième chapitre explicite les présupposés épistémologiques essentiels qui permettent de légitimer notre démarche de recherche. Nous justifions le positionnement épistémologique en proposant un cadre méthodologique précis et cohérent avec l'objectif de notre projet de recherche.

Notre recherche s'inscrit dans une démarche positiviste et adopte un raisonnement de déduction (de lois et théories universelles aux explications et prédictions) et donc une démarche hypothético-déductive est utilisée (Thiéart, 2014). Cette démarche permet le cheminement suivant : l'on part de la théorie de l'efficacité d'usage du big data, à partir d'un modèle préalablement conçu. Nous examinons ce modèle sur le terrain afin d'arriver à une conclusion qui va nous permettre de répondre à la question principale de recherche. Donc, il s'agit bien de tester un modèle, développé préalablement et le mettre à l'épreuve dans les situations particulières.

Ce choix est résumé à partir des quatre axes proposés par Mbengue et Vandangeon-Derumez (1999). Les quatre dimensions discutées sont : le statut accordé aux données (1) ; le mode de collecte des données (2) ; la relation entre la théorie et les observations empiriques (3) et les critères de scientificité de la recherche (4).

Les sections concernées détaillent bien l'ensemble des étapes d'études réalisées : étude qualitative exploratoire, étude quantitative, étude qualitative confirmatoire.

Le cinquième chapitre met en évidence l'importance de la contextualisation d'une recherche dans la science en gestion. Il décrit le profil de l'Arménie au regard d'une contextualisation à trois pôles : culturel, institutionnel, post-soviétique. Cette présentation met également en évidence l'intérêt de mener une recherche sur le sujet de gestion de technologies de l'information dans un pays comme l'Arménie, où l'on peut rencontrer un paradoxe : comment se fait-il qu'en Arménie, dont l'économie est faible et les secteurs peu développés, l'industrie technologique prospère bien ? La deuxième section de ce chapitre propose une description complète du terrain mobilisé dans ce travail de thèse. Pour une première approche exploratoire, nous étudions un cas bien précis – un service de taxi. Ce terrain nous permet de vérifier que les composants du modèle sont bien présents dans les préoccupations des acteurs. Ensuite, nous proposons une représentation des 16 entreprises arméniennes qui exploitent du big data. Ce sont les entreprises qui font partie de notre enquête quantitative. Enfin, nous présentons les deux cas bien précis d'efficacité d'usage du big data chez les deux acteurs participants à l'étude quantitative, qui nous permet un prolongement des résultats, que nous discutons dans le chapitre 7.

Le sixième chapitre détaille bien les résultats issus de trois études réalisées.

Et le dernier et septième chapitre permet de répondre à la question principale de la thèse et de faire une conclusion sur notre travail de thèse. Pour ce faire, nous développons une discussion autour de trois axes en interpellant les résultats issus des études réalisées. Enfin, les contributions théoriques et managériales, ainsi que les limites et perspectives de la recherche sont exposées.

**Première partie : Le choix d'un cadre théorique
au big data : L'efficacité d'usage**

Chapitre 1 Utilisation du big data par les entreprises

On a coutume de parler des opportunités qu’offre le big data : l’amélioration de la performance opérationnelle (Dubey, Gunasekaran, Childe, Blome, & Papadopoulos, 2019) ; l’amélioration de la performance dans un domaine ou dans un secteur spécifique (J. Wang, Xu, Zhang, & Zhong, 2022), (Batko & Ślęzak, 2022), (Munawar, Ullah, Qayyum, & Shahzad, 2022), (Papadopoulos & Balta, 2022), (Hamdam, Jusoh, Yahya, Abdul Jalil, & Zainal Abidin, 2022), (Okwechime, Duncan, & Edgar, 2018) , (Golightly, Kefalidou, & Sharples, 2018) ; l’amélioration de la performance et de la stratégie (Sena, Bhaumik, Sengupta, & Demirbag, 2019); l’innovation dans les modèles d’affaires des produits et des services (Mikalef, Boura, Lekakos, & Krogstie, 2019b) ; l’amélioration de la qualité du service client (De Luca, Herhausen, Troilo, & Rossi, 2021), (Batistič & van der Laken, 2019), (He, Tian, Hung, Akula, & Zhang, 2018); l’outil de gestion de la fraude et des risques (Leavitt, 2013); le support d’aide à la décision (Akhtar, Frynas, Mellahi, & Ullah, 2019), (Frisk & Bannister, 2017), (Xu, Frankwick, & Ramirez, 2016). Ici, il convient de noter que le principal apport du big data est le changement du modèle de prise de décision de « ce que je pense » vers « ce que je sais » (Ranjan & Foropon, 2021), (McAfee & Brynjolfsson, 2012a), la découverte de connaissances cachées, l’information issue de l’exploitation du big data et la génération de nouvelles connaissances (Z. Khan & Vorley, 2017), (Pauleen & Wang, 2017). Mais des interrogations demeurent : notamment, sur la façon dont les entreprises peuvent exploiter ces avantages (Wu, Liang, & Chen, 2022), (Mithas, Lee, Earley, Murugesan, & Djavanshir, 2013) et la manière dont elles peuvent appréhender les particularités d’une technologie relativement nouvelle et complexe (Dossa, 2019), (Boudreau & Seligman, 2011) ainsi que les raisons pour lesquelles les initiatives du big data mènent souvent les entreprises à la faillite (Sheng, Amankwah-Amoah, & Wang, 2017). Ces questions n’ont pas de réponses claires.

Une problématique assez répandue dans le monde des affaires aujourd’hui est le suivant: d’un côté l’accès des entreprises au big data semble virtuellement possible, mais d’un autre côté l’on constate une méconnaissance des entreprises quant à l’usage pratique du big data (Jha et al., 2020).

Dans ce contexte, les entreprises sont souvent exposées à la « zone aveugle » (blind zone). Quand le volume des données disponibles aux organisations est en croissance, un effet contraire se produit : le pourcentage des données qu'elles sont en capacité de traiter, de comprendre et d'analyser est en déclin, bien illustré sur la figure 1 (Zikopoulos & Chris Eaton, 2011).

Par conséquent, l'intérêt principal de ce travail est d'étudier en profondeur l'usage du big data en entreprise et les facteurs les plus importants qui maximisent l'efficacité de son usage.



Figure 1 Volume des données accessibles aux entreprises par rapport aux volumes qu'elles sont en capacité de traiter (Zikopoulos & Chris Eaton, 2011)

1.1 La genèse et les définitions du big data

Le big data est un concept ancien, dont l'usage demeure un phénomène émergent et non normé dans les entreprises.

L'idée que, « plus, il y a des données collectées, meilleure est la prise de décision » est trop simpliste. Ces données doivent être converties de manière robuste et fiable afin d'être traduites en connaissances applicables (Brinch, Gunasekaran, & Fosso Wamba, 2021). Et donc la valeur du big data c'est l'information utile d'y « retenue ».

Pour comprendre les fondements du big data, ou des données massives, l'on peut s'appuyer sur les relations qui s'établissent entre les trois catégories : « Données – Informations – Connaissances ». Les données sont des faits représentant une certaine réalité, elles sont

indépendantes de celui qui les utilise (Lépinard, Lebraty, & Lobre, 2013). Les informations sont des données auxquelles un individu ou un groupe a ajouté du sens; et le passage de la notion d'information à celle de connaissance s'opère au travers d'un mécanisme de cognition sur les informations (Lebraty, 2011).

Il faut bien noter ici que l'information possède une structure temporelle. Dit autrement, l'apport de l'information n'est pas le même selon son positionnement temporel (Chalus-Sauvannet, 2021).

Le terme « big data » a été diffusé pour la première fois, en 2011 (Gandomi & Haider, 2015), la figure 2 illustre la fréquence d'apparition du terme dans les 2 ans qui suivent.

Karoui et ses coauteurs (Karoui et al., 2014) définissent l'expression « big data » comme *les technologies, les processus et les techniques qui permettent à une organisation de créer, de manipuler et de gérer des données à grande échelle pour en extraire de nouvelles connaissances et créer une valeur économique.*

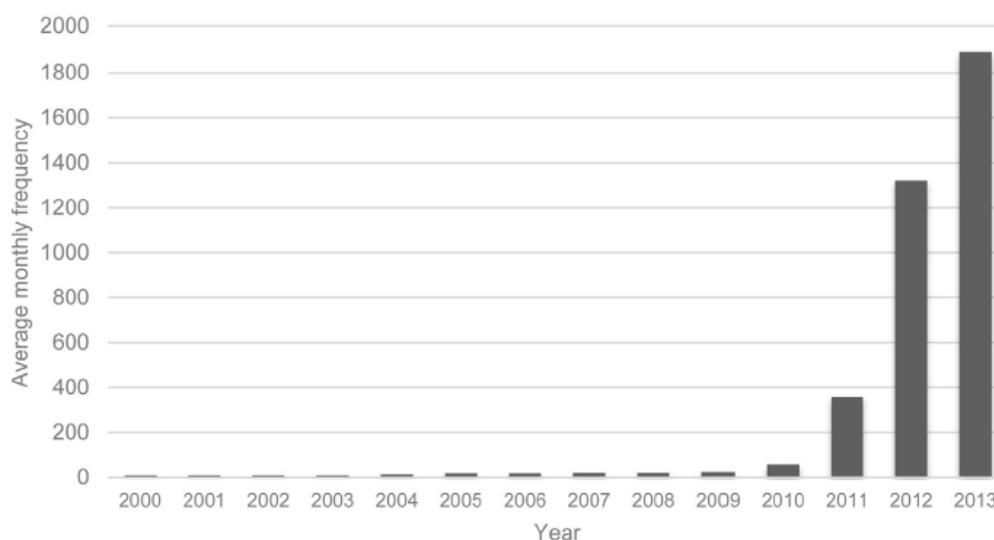


Figure 2 Fréquence d'apparition du terme "big data" (Gandomi & Haider, 2015)

Gartner propose de définir le big data comme : « ...des données d'un grand volume et d'une grande variété qui se génèrent à haute vitesse, ce qui nécessite des formes de traitement

particulières pour une meilleure compréhension afin d'avoir une approche innovante de prise de décision, et l'automatisation des processus tout en ayant un coût optimisé (cost-effective) »².

Philip Chen et Zhang proposent qu'un ensemble de données puisse être appelé « big data » *s'il est parfaitement adapté pour effectuer la capture, la conservation, l'analyse et la visualisation au travers des technologies existantes* (Chen & Zhang, 2014).

D'autres, comme Hashem et ses coauteurs (Hashem et al., 2015) caractérisent le big data par trois aspects : les données sont nombreuses (1), les données ne peuvent pas être classées dans une base de données relationnelle traditionnelle (2), les données sont générées, capturées et traitées rapidement (3).

Pour une compréhension structurée sur les processus du management du big data, on distingue deux catégories : la science du big data (big data science) et l'infrastructure du big data (Siddiqa et al., 2016). La science du big data étudie les techniques et les technologies qui vont permettre l'analyse, la visualisation et l'exploitation des connaissances des données, alors que l'infrastructure du big data est une représentation de plusieurs systèmes (framework)³ du big data.

1.1.1 Définition du big data à partir de ses caractéristiques spécifiques

Des auteurs définissent le big data en associant des caractéristiques spécifiques, dites les Vs: 3Vs (Volume, Vitesse, Variété) - (Chen & Zhang, 2014); 4Vs (Volume, Vitesse, Variété, Véracité) - (Sivarajah, Kamal, Irani, & Weerakkody, 2017) ; (De Mauro, Greco, & Grimaldi, 2016) ; (Abbasi, Sarker, & Chiang, 2016); 5Vs (Volume, Vitesse, Variété, Véracité, Valeur)- (Sharda, Delen, Turban, & King, 2014) et 7Vs (Volume, Vitesse, Variété, Véracité, Valeur, Variabilité et Visualisation) - (Seddon & Currie, 2017).

Le volume se réfère à la taille des données ; *la taille des données augmente de manière très significative* (Ghasemaghaei, Ebrahimi, & Hassanein, 2018). *Une estimation est donnée dans le « The Economist », publiée en 2017 qui prévoit que le volume des données atteindra*

² <https://www.gartner.com/en/information-technology/glossary/big-data>

³ Big data framework présente un ensemble de logiciels avec les algorithmes associés qui permettent le traitement et analyse de problèmes du big data dans les clusters des unités de machines (computer units)(Hu et al., 2014).

180ZB en 2025 (Vassakis, Petrakis, & Kopanakis, 2018) ; notons que 1 zettaoctet est, en données, l'équivalent de 250 milliards de DVD⁴.

Leur variété due aux nombreuses sources et types de données :

- Structurées, comme les données du scanner ou du capteur, enregistrements, fichiers et bases de données) ;
- Semi-structurées, ces données intègrent les différents types de logiciels qui ordonnent les données non-structurées ; par exemple, le logiciel SGML⁵ permet de déterminer, dans des vidéos, les éléments intéressants l'entreprise. Cela peut être une vidéo publiée sur YouTube⁶ montrant des personnes utilisant le produit ou le service de l'entreprise (combien d'entre eux semblent heureux, contents, etc.) ;
- Non-structurées, qui comprennent des données textuelles (blogs, messages-textuels) et des données non-textuelles (vidéos, images, enregistrements audio) (Erevelles, Fukawa, & Swayne, 2016).

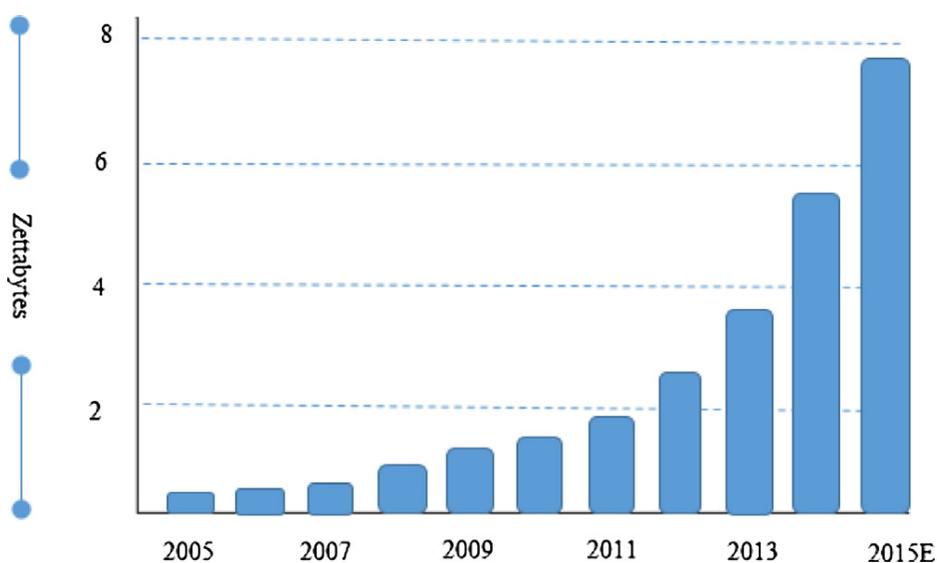


Figure 3 Évolution des données générées (Yaqoob et al., 2016c)

⁴https://www.cisco.com/cdc_content_elements/networking_solutions/service_provider/visual_networking_ip_traffic_chart.html

⁵ Standard Generalized Mark-up Language

⁶ <https://www.youtube.com/>

De plus, on peut noter que les données non structurées croissent plus vite que les données structurées. Elles occupent 90% de toutes les données (Gantz et al., 2012).

La vélocité se réfère à la vitesse et à la fréquence du traitement et de l'intégration des données (Ghasemaghaei & Calic, 2020). Notons que sont réalisés un milliard de nouveaux tweets tous les trois jours, 5 milliards de recherches quotidiennes, et que le nombre des messages quotidiens sur les forums et des entrées de blogs dépasse également les quelques millions. (Abbasi et al., 2014). En 2016 ScienceDaily⁷ a annoncé que les 90% des données du jour ont été générés au cours des deux dernières années (ScienceDaily, 2016). La figure 3 illustre clairement l'exponentialité de la génération des données (Yaqoob et al., 2016c).

La véracité, c'est la crédibilité et la fiabilité des données, dont les sources peuvent varier (Abbasi et al., 2016) ; ça montre aussi le niveau de bruit dans les données (Goes, 2014). Notons qu'on estime que 20% de toutes les pages sur World Wide Web sont du spam (Abbasi et al., 2014).

La valeur fait référence à la capacité d'utiliser les données pour en extraire des informations ayant une valeur pour l'organisation (« quelle valeur » ?) qui se traduira par une intelligence économique accrue et une aide à la prise de décision (Aljumah, Nuseir, & Alam, 2021), (Shim, French, Guo, & Jablonski, 2015), afin de créer une valeur économique (Ylijoki & Porras, 2016).

La variabilité se réfère aux opportunités dynamiques disponibles pour interpréter des données non structurées ; et la visualisation est atteinte grâce aux outils qui assurent une représentation claire, par exemple, la génération des modèles à l'aide de l'intelligence artificielle (Seddon & Currie, 2017).

Seddon et Currie (2017) ont développé ainsi un modèle conceptuel (Figure 4) du big data pour les entreprises du domaine de « High-Frequency Trading » (Trading Haute Fréquence). Dans ce modèle les 7 caractéristiques du big data sont regroupées dans les trois catégories, dites BC (1), BD (2) et FD (3). La première se réfère à la capacité d'un grand pouvoir de traitement (Big Compute Capability). Pour le big data (BD), il s'agit du stockage et du traitement des

⁷ <https://www.sciencedaily.com/>

données et la troisième, c'est le « Fast Data access », l'accès rapide aux données pour le commerce en temps réel.

(BD = Big Data; FD = Fast Data and BC = Big Compute)

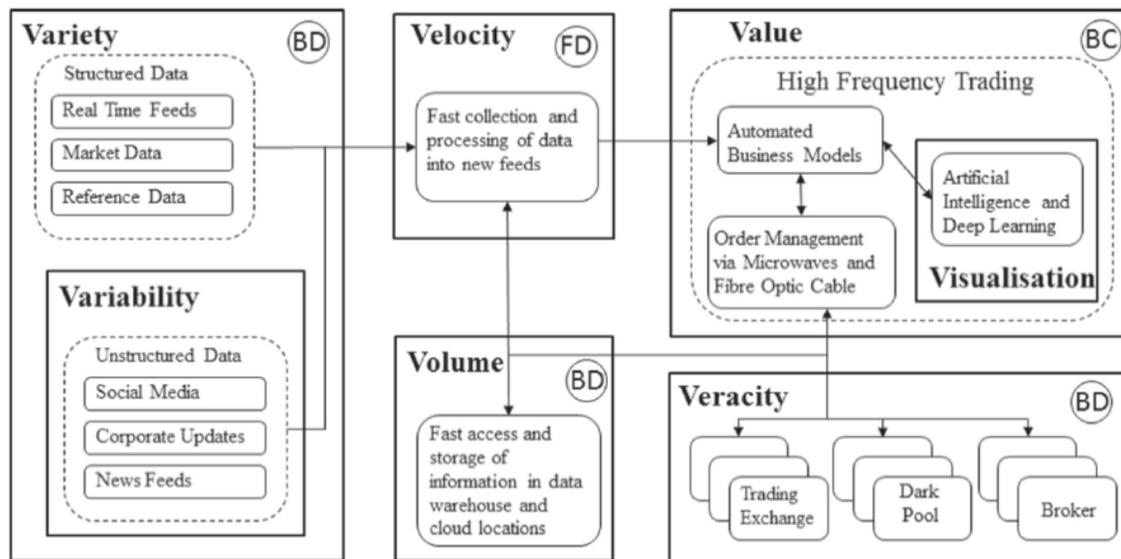


Figure 4 Modèle conceptuel du big data pour les entreprises du domaine de « High-Frequency Trading » (Seddon & Currie, 2017)

Ces caractéristiques spécifiques que possède le big data, demandent une attention et des mécanismes particuliers pour leurs traitements (Janssen, van der Voort, & Wahyudi, 2017).

	Traditional Data	Big Data
Volume	GB	constantly updated (TB or PB currently)
Generated Rate	per hour, day, ...	more rapid
Structure	structured	semi-structured or un-structured
Data Source	centralized	fully distributed
Data Integration	easy	difficult
Data Store	RDBMS	HDFS, NoSQL
Access	interactive	batch or near real-time

Tableau 1 Comparaison entre les données traditionnelles et le big data (Hu, Wen, Chua, & Li, 2014)

Cette brève description des caractéristiques du big data montre non seulement les différents aspects par lesquels le big data diffère des données traditionnelles, mais aussi rend leur

d'une sorte de seuil qui dépasse les capacités de traitement des systèmes de base de données traditionnelles.

Les aspects technologiques du big data vont être décrits plus en détail dans une section dédiée de ce chapitre.

1.1.3 Autres dimensions pour définir le big data

On présente aussi le big data comme un état transformatif aux niveaux non seulement des organisations, mais aussi des individus, de l'état et des sociétés (Markus & Topi, 2015).

Il est aussi défini comme un phénomène culturel, technologique et scientifique qui repose sur l'interaction de la technologie, de l'analyse et de la mythologie, provoquant encore une vaste rhétorique utopique et dystopique (Boyd & Crawford, 2012).

La révolution provoquée par le big data transforme potentiellement non seulement la façon dont on vit, travaille et réfléchit (Mayer-Schönberge et al., 2013), mais en plus, elle change la manière de conduire une recherche par l'émergence d'un nouveau et 4^e paradigme scientifique, appelé Data-intensive science (Tolle, Tansley, Hey, 2011). Le but de ce dernier est de fournir les outils nécessaires pour traiter de grands problèmes du big data.

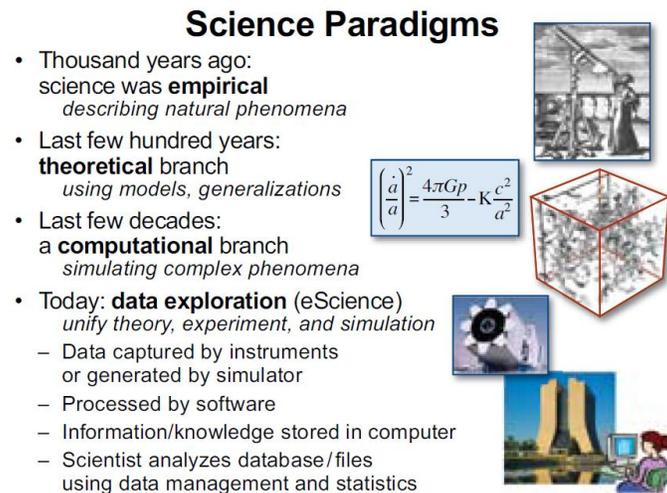


Figure 6 Les paradigmes de la science (Hey, Tansley, & Tolle, 2009)

Ce nouveau paradigme vient en complément des trois paradigmes précédents :

- Science empirique, émergée, il y a mille ans, quand l'expérience empirique humaine essaie de décrire des phénomènes naturels (1^e paradigme, aussi décrit comme le début de la science) ;
- Science théorique, émergée, il y a quelques centaines d'années, où l'on essaie de décrire les phénomènes par l'utilisation de modèles, des généralisations avec la loi des mouvements de Newton, la loi de Kepler, etc. (2^e paradigme) ;
- Par la suite, le monde de la science a évolué et fait naître un troisième paradigme (durant les dernières décennies) qui apparaît comme une branche informatique où l'on décrit un phénomène complexe par des simulations.

Aujourd'hui, nous vivons dans une ère d'exploration des données (eScience), où l'on mobilise à la fois la théorie, l'expérimentation et la simulation, voir la figure 6 pour l'illustration (Hey et al., 2009) ; (Philip Chen & Zhang, 2014).

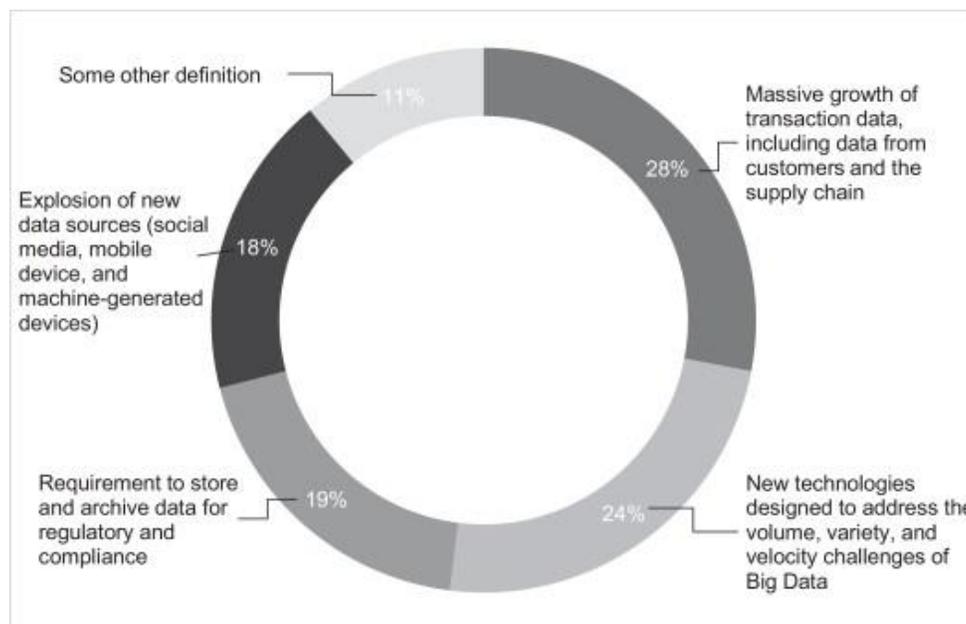


Figure 7 Résultats de l'enquête: Définitions du big data proposées par les cadres exécutifs (Gandomi & Haider, 2015)

Les résultats d'une enquête montrent que la rapidité d'évolution des définitions du « big data » a causé une confusion dans la compréhension de ce terme par les cadres exécutifs. Cette enquête a montré que parmi les 154 cadres exécutifs dans les entreprises internationales de

différents secteurs, seuls les 25% des interrogés s'accordent sur une définition similaire du big data, tandis que près de 75% des répondants voient le big data juste comme une opportunité⁸.

La figure 7 montre les résultats de l'enquête sur les définitions proposées par les cadres exécutifs (Gandomi & Haider, 2015).

1.2 Les principales sources du big data

Le big data est un ensemble de différents types des données. Baesens et ses coauteurs (Baesens, Bapna, Marsden, Vanthienen, & Zhao, 2016b) proposent de distinguer les cinq principales sources du big data dans les économies contemporaines :

Source	Production
YouTube [15]	(i) Users upload 100 hours of new videos per minute (ii) Each month, more than 1 billion unique users access YouTube (iii) Over 6 billion hours of video are watched each month, which corresponds to almost an hour for every person on Earth. This figure is 50% higher than that generated in the previous year
Facebook [16]	(i) Every minute, 34,722 Likes are registered (ii) 100 terabytes (TB) of data are uploaded daily (iii) Currently, the site has 1.4 billion users (iv) The site has been translated into 70 languages
Twitter [17]	(i) The site has over 645 million users (ii) The site generates 175 million tweets per day
Foursquare [18]	(i) This site is used by 45 million people worldwide (ii) This site gets over 5 billion check-ins per day (iii) Every minute, 571 new websites are launched
Google+ [19]	1 billion accounts have been created
Google [20]	The site gets over 2 million search queries per minute Every day, 25 petabytes (PB) are processed
Apple [20]	Approximately 47,000 applications are downloaded per minute
Brands [20]	More than 34,000 Likes are registered per minute
Tumblr [20]	Blog owners publish 27,000 new posts per minute
Instagram [20]	Users share 40 million photos per day
Flickr [20]	Users upload 3,125 new photos per minute
LinkedIn [20]	2.1 million groups have been created
WordPress [20]	Bloggers publish near 350 new blogs per minute

Tableau 2 Rapidité de génération des données non-structurées (N. Khan et al., 2014)

- Systèmes d'entreprise à grande échelle (large-scale) qui peuvent comprendre le ERM – entreprise resource planning, le CRM – customer relationship management, le SCM – supply chain management, etc. ;
- Plateformes sociales en ligne comme Facebook, Twitter, Weibo, WeChat, etc. pour mieux illustrer, voir le tableau 2 proposé par Khan et ses coauteurs (2014);
- Appareils mobiles – avec pratiquement 5 milliards d'appareils dans le monde ;
- IoT – « Internet of Things » qui représente un écosystème physique intelligent avec les objets connectés utilisant les capteurs qui génèrent du big data ;

⁸ Small and midsize companies look to make big gains with “bigdata” récupéré : <https://news.sap.com/2012/06/>

- Les données ouvertes : « Open Data » ou « Data Public » – comme les données sur la météo, le trafic, les cartes, l’environnement, etc. Ici il faut noter que dans le cas des données ouvertes, les deux types d’acteurs existent : *les émetteurs de données qui sont les entités donnant libre accès à leurs données via le web et les ré-utilisateurs qui sont les personnes physiques ou morales, qui développent de nouveaux usages à partir des données ouvertes* (Lobre & Lebraty, 2012).

Le tableau 2 propose une illustration très claire de la rapidité de génération des données non-structurées avec leurs sources et les éléments-clés de leur production.

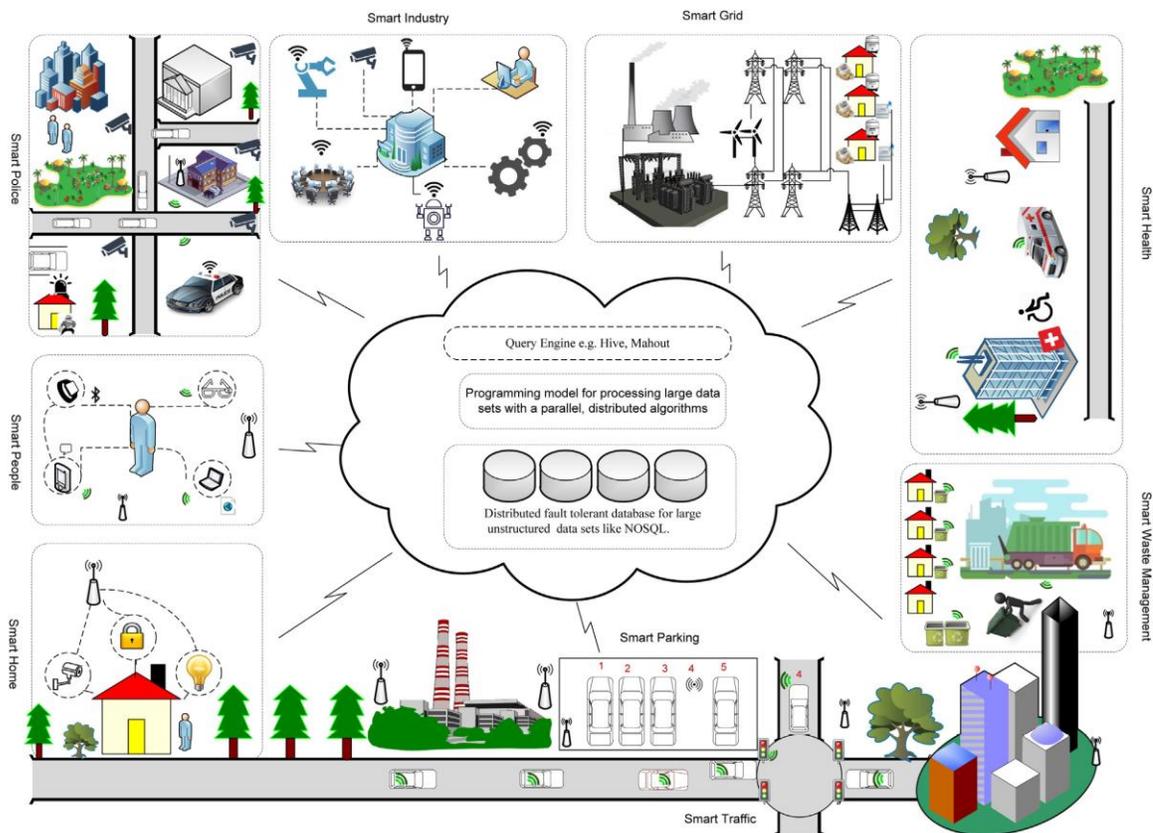


Figure 8 Un exemple de « smart city » (Yaqoob et al., 2016a)

Yaqoob et ses coauteurs (2016 a) pour leur part, proposent les quatre types de sources du big data :

- IoT – Internet of Things, avec les données générées par appareil GPS, des objets intelligents – smart cars, vêtements intelligents, etc. ;

- Données auto-quantifiées qui sont générées par les individus pour la quantification de leur comportement, comme les données des bracelets utilisés pour surveiller les mouvements ou les sphymomanomètres utilisés pour mesurer la pression artérielle ;
- Données multimédias, issues de sources diverses, comme les textes, images, audio, vidéo : chaque individu connecté à Internet génère les données multimédias ;
- Enfin, les médias sociaux comme Facebook, Twitter, LinkedIn, YouTube, Google, Apple, etc.

Pour une compréhension plus claire, on peut s'appuyer sur la figure 8, qui illustre un exemple de « smart city » avec les technologies de big data appliquées, proposée par Yaqoob et ses coauteurs (Yaqoob et al., 2016a).

1.3 L'usage et le traitement du big data

Les entreprises font des efforts considérables afin de renforcer leurs capacités organisationnelles pour pouvoir traiter le big data aujourd'hui (Ghasemaghahi, 2021). Pour le réussir, il convient d'abord de construire une compréhension claire sur l'usage du big data.

Les technologies du big data représentent une nouvelle génération de technologies et d'architectures, conçues pour, à partir de très gros volumes d'une grande variété de données, capturer à haute vitesse, analyser et extraire de l'information d'une manière économiquement rentable (Phillips-Wren & Hoskisson, 2015).

Pour tirer de la valeur du big data, diverses techniques ont été développées. Ces techniques sont pluridisciplinaires : mathématiques, statistiques, informatiques, économiques. Par exemple, Walmart traite ses données des transactions en appliquant la ML (Machine Learning) et des méthodes statistiques (Chen & Zhang, 2014).

Les trois motifs majeurs pour l'adoption d'une technologie big data par une entreprise sont : minimaliser les coûts des matériaux « hardware » (1), vérifier la valeur des données avant d'engager d'importantes ressources de l'entreprise (2) et réduire les coûts de traitement (3) (N. Khan et al., 2014), (Yaqoob et al., 2016 b).

1.3.1 Le traitement du big data

Pour mieux imaginer le traitement du big data, on propose souvent de le décrire par une chaîne d'étapes, une « chaîne de big data » qui comprend les quatre phases du processus : collecte, préparation et analyse des données puis la prise de décision (Janssen et al., 2017) ; ou encore la génération, l'acquisition, le stockage et l'analyse des données (M. Chen, Mao, & Liu, 2014). Il en résulte que les principaux défis (challenges) qui sont liés au big data, concernent les quatre processus dans la chaîne décrite ci-dessus.

On comprend qu'une technologie big data sert à accomplir un ou plusieurs processus décrits ci-dessus avec des données massives (big data). Dans cette logique, on fait souvent une distinction des technologies de big data en quatre groupes : les technologies de stockage (1) ; d'analyse (2) ; d'exploration des données ou data mining (3) et les technologies de visualisation des données (4).

Comme l'objectif de la thèse est d'étudier l'efficacité d'usage du big data en entreprise, nous trouvons d'abord important de décrire clairement en quoi consiste l'usage du big data à partir de deux paradigmes ou modèles principaux de traitement utilisés par les entreprises.

Ces deux paradigmes sont : le traitement en continu (Streaming Processing) et le traitement par lots (Batch Processing) (Yaqoob et al., 2016 c) ; (Hu et al., 2014) ; (Hashem et al., 2015). Dans le paradigme du traitement en continu, le but de traitement des données, c'est leur « fraîcheur », et dans ce cas-là, il est souhaitable de traiter les données au plus vite ou *en temps réel* (Yaqoob et al., 2016 a). Les données arrivent en continu (stream), et au vu de leur gigantesque volume, seule une petite partie est stockée en mémoire, car cette dernière est limitée (Hu et al., 2014).

Le Storm (Chen & Zhang, 2014), *le S4* (Neumeyer, Robbins, Nair, & Kesari, 2010) et *le Kafka* (Kreps, Narkhede, & Rao, 2011) sont exemples des technologies basées sur le stream.

Le deuxième paradigme est le traitement par lots (batch), c'est-à-dire que l'on stocke d'abord les données pour les traiter après (Hu et al., 2014), qui est souvent appliqué sur les données historiques (Landset, Khoshgoftaar, Richter, & Hasanin, 2015). L'exemple le plus cité

de ce modèle de traitement est *MapReduce* (Hashem, Anuar, et al., 2016) ; (Dean & Ghemawat, 2008).

Dans la figure 9, l'on propose une illustration sur ces deux paradigmes de traitement, développés par Hu et ses coauteurs (Hu et al., 2014).

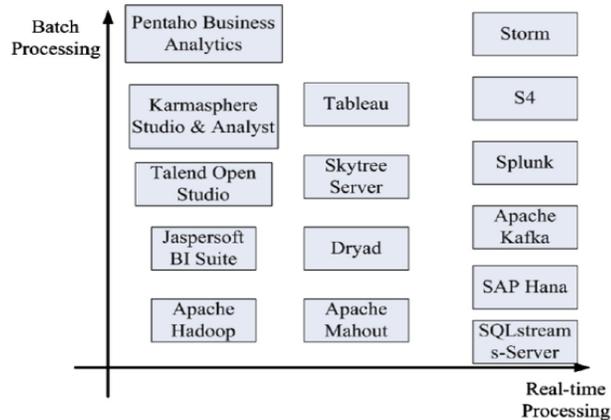


Figure 9 Plateformes du big data (Hu et al., 2014)

Le tableau 3 en plus, illustre des exemples pour les deux plate-formes de traitement du big data, développé par Chen et Zhang (2014). Comme on peut voir, la capacité de traitement proche du temps réel augmente (le temps de réponse diminue) de gauche à droite et la capacité de traitement par lots augmente du bas vers le haut.

	streaming processing	batch processing
Input	stream of new data or updates	data chunks
Data size	infinite or unknown in advance	known & finite
Storage	not store or store non-trivial portion in memory	store
Hardware	typical single limited amount of memory	multiple CPUs, memories
Processing	a single or few pass(es) over data	processed in multiple rounds
Time	a few seconds or even milliseconds	much longer
Applications	web mining, sensor networks, traffic monitoring	widely adopted in almost every domain

Tableau 3 Comparaison entre les 2 plateformes de traitement du big data (Philip Chen & Zhang, 2014)

1.3.2 Les applications du big data

Pour approfondir les connaissances sur l'usage du big data et son traitement, on passe à une brève présentation des applications principales.

Si l'on jette un regard sur le développement des modèles des applications (logiciels), on peut constater les cinq principales (Abolfazli, Sanaei, Gani, Xia, & Yang, 2014), parmi elles :

- *Autonome (standalone)*, qui emploie une seule unité de traitement pour refléter les actions des utilisateurs en fonction de la vitesse de calcul de l'hôte machine, par exemple, Notepad, Calculator, Microsoft Word...
- *De bureau (desktop)*, ce sont des applications autonomes qui s'exécutent sur un ordinateur de bureau sans accéder à Internet, par exemple, les applications de messagerie instantanée, *le Web* qui a rendu possible l'accès aux services locaux et les données via Internet (Google Docs, Meebo, Jaycut, etc.) ;
- *Internet Riche (Rich Internet)* qui combine les applications Web et de bureau ;
- *Big data*.

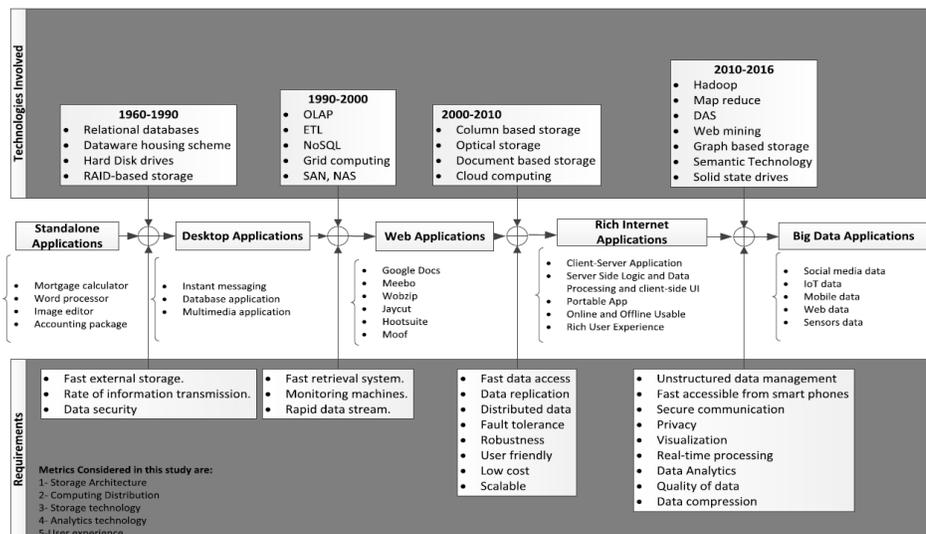


Figure 10 Genèse des applications du big data (Abolfazli et al., 2014)

Pour leur évaluation, les cinq critères suivants sont proposés par Abolfazli et ses coauteurs (Abolfazli et al., 2014) :

- *Architecture de stockage*, qui fait référence aux données stockées dans un environnement informatique ;
- *Distribution informatique (computing distribution)*, qui fait référence à de nombreuses composantes de logiciels dans des ordinateurs en réseau qui fonctionnent comme un système unique ;
- *Technologie de stockage*, qui fait référence à l'emplacement où les données sont stockées sous une forme électromagnétique ou optique ;
- *Technologie analytique*, qui se réfère à l'analyse informatique systématique afin de transformer les données en information ;
- *Expérience utilisateur (user experience)*, qui fait référence à la qualité globale de l'interaction d'un utilisateur avec le système.

Une évolution de ces développements est illustrée sur la figure 10.

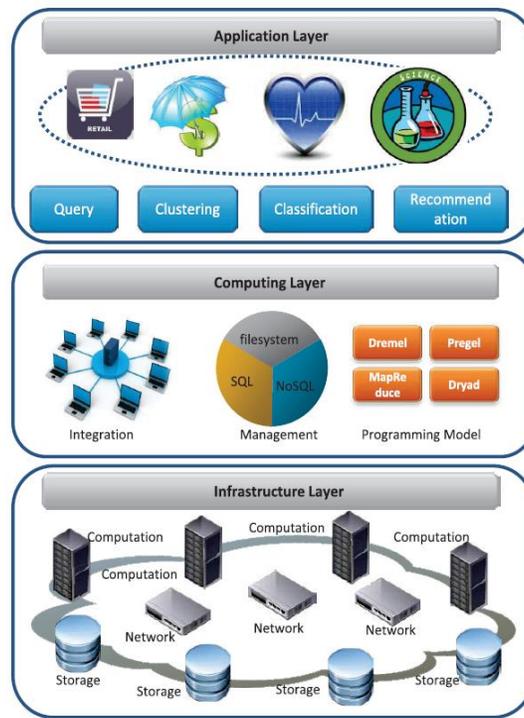


Figure 11 Système du big data à 3 catégories (Hu et al., 2014)

En analysant les réflexions de plusieurs auteurs qui ont essayé de décrire cet usage complexe du big data, on a souvent l'impression qu'on se promène autour du phénomène et l'on n'arrive pas à l'appréhender ni à le comprendre en profondeur.

Pour combler cet écart, nous proposons de nous appuyer sur la figure 11, qui décrit un système du big data à 3 catégories.

Développée par Hu et ses coauteurs (Hu et al., 2014), la figure 11 montre qu'un système du big data a trois catégories : l'infrastructure, l'informatique (computing) et le logiciel (application). L'infrastructure consiste en des éléments de technologies informatiques, où l'on voit les trois éléments-clés : le stockage, le réseau et la force de calcul (computation). Ces derniers peuvent être organisés par l'infrastructure du « cloud computing », activé par la technologie de virtualisation.

Dans la catégorie de l'informatique ou « computing », nous voyons trois ressources : l'intégration, le management et le modèle de programmation. Ces ressources doivent être allouées de façon à augmenter l'utilisation du système et la simplification opérationnelle. L'intégration des données signifie l'acquisition des données de différentes sources et leur intégration sous une forme unifiée avec les opérations nécessaires de prétraitement.

Le management des données se réfère aux mécanismes et outils de stockage des données et une gestion efficace comme les systèmes de fichiers distribués et les bases de données SQL ou NoSQL. Les modèles de programmation facilitent le fonctionnement des applications des données : MapReduce, Dryad, Pregel, Dremel exemplifient les modèles de programmation. Et la dernière catégorie, est l'interface fournie par des modèles de programmation pour implémenter de diverses fonctions d'analyse des données y compris Query (interrogation), Clustering (regroupement), Classification, Recommandation.

Pour développer la compréhension plus fine sur l'usage du big data, nous allons encore, détailler les processus concernés. Pour le faire, nous proposons d'interpréter la figure 12 où est détaillé un cycle de vie des données. C'est une illustration suffisamment complète et facile à comprendre pour constituer une vision claire sur le sujet.

On peut noter que le cycle de vie des données comporte les étapes suivantes : la collecte, le filtrage et la classification, l'analyse des données, le stockage, le partage et l'édition, la récupération et la découverte des données.

Le cycle de vie des données démarre par la transformation des données matières (raw data) et leur stockage. Par la suite, on passe à la collecte des données au travers de diverses techniques d'acquisition des données d'un domaine spécifique.

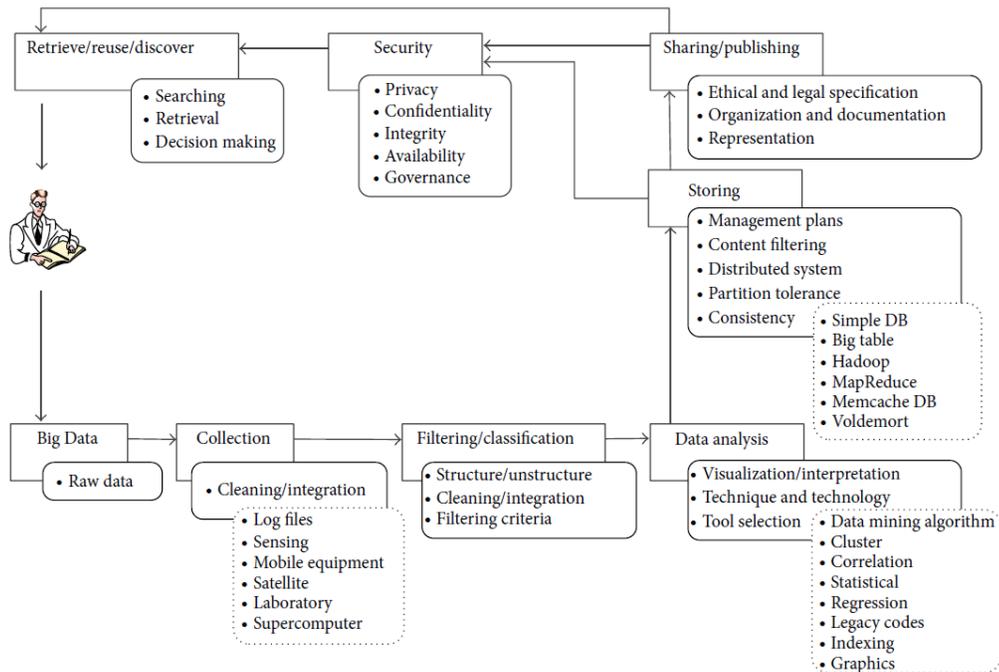


Figure 12 Cycle de vie des données (N. Khan et al., 2014)

Les méthodes les plus utilisées de collecte des données sont :

- Les fichiers journaux (log files)(Abd Wahab, Mohd, Hanafi, & Mohamad Mohsin, 2008), qui sont utilisés pour collecter les données au travers des fichiers automatiquement enregistrés (le nombre des visites, de clics notés par le serveur Web et l'historique d'accès des utilisateurs Web) (Nanopoulos, Zakrzewicz, Morzy, & Manolopoulos, 2002), l'intérêt des utilisateurs Web (Loia, Pedrycz, Senatore, & Sessa, 2006);
- Les capteurs, souvent utilisés pour mesurer des quantités physiques qui sont converties dans les signaux digitaux pour leur stockage et le traitement (les ondes sonores, la vibration, la voix, la température, etc.), ces données détectées sont transférées au point de stockage au travers des réseaux filaires ou sans fil (wired or wireless networks) pour le traitement (F. Wang & Liu, 2011);

- La capture des données de réseaux (Network Data Capture) par une combinaison des méthodes du Web Clower (un programme qui télécharge et stocke des pages Web (Cho & Garcia-Molina, 2000)), de segmentation des tâches et mots, l'index; la technologie de capture de paquets ZC (Zero-Copy) (B. L. Wang, Fang, & Yun, 2005);
- L'équipement mobile, les appareils mobiles qui sont une autre source de données, fournissent différents types d'informations, comme la géolocalisation, l'information audio au travers des micros, autre type d'information multimédia ; pour illustrer le propos, nous avons les iPhones qui collectent les données sans fil (Wireless data) et l'information sur la géo location et les transmet à Apple Inc. pour un futur traitement (sans que l'utilisateur en soit conscient).

Dans la phase du filtrage et la classification, les données passent par le filtrage et sont catégorisées par leur structuration (structurée / non structurée). L'objectif ici est de construire une sortie (output) intermédiaires des données (Ahmed Oussous a, Fatima-Zahra Benjelloun a, Ayoub Ait Lahcen 2018).

Par la suite, les données passent par une analyse. On note les deux principaux objectifs d'analyse des données : d'abord, celui de comprendre les relations entre les caractéristiques différentes et deuxièmement, celui de développer des méthodes efficaces pour l'exploration des données qui peuvent prédire les futures observations avec précision. Les techniques analytiques utilisées sont l'exploration des données. Le data mining est largement utilisé dans les domaines de la science, l'ingénierie, la médecine et le business. Cette technique est utilisée pour obtenir divers motifs cachés de données (Rao, Mitra, Bhatt, & Goswami, 2019). Dans cette technique analytique l'information est extraite par un gros volume de données qui sont incomplètes, confuses et bruyantes (noisy).

La visualisation, c'est-à-dire l'utilisation des techniques de représentations graphiques des données ou techniques d'interaction se réfèrent aux formes visuelles dans lesquelles les données ou les modèles sont affichés ; par exemple, un graphique à barres ou un graphique linéaire (Zhang et al., 2012)).

L'analyse statistique repose sur le choix de méthodes statistiques cohérentes qui peuvent assurer la construction de grandes matrices de covariance qui déterminent la structure de

corrélation et l'apprentissage machine (ML-Machine Learning) avec ses trois sous-catégories de l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement, utilisés pour divers applications des moteurs de recommandation, des systèmes de reconnaissance, des systèmes de contrôle autonomes, etc. (Bishop, 2014), (Qiu, Wu, Ding, Xu, & Feng, 2016)).

Plusieurs techniques traditionnelles d'analyse des données sont aussi applicables pour le traitement du big data. Les algorithmes de l'exploration des données, comme *SVM*, *C4.5*, *Apriori*, *k-means*, *Cart*, *EM*, and *Naive Bayes* sont utilisés (J. Chen, Wang, & Wang, 2009). Ils assurent bien les capacités suivantes :

- *Classification* ;
- *Régression* ;
- *Regroupement (clustering) où il s'agit d'un regroupement des objets en leur attribuant statistiquement certaines règles ou caractéristiques* ;
- *Analyse des corrélations (pour déterminer des relations entre les phénomènes pratiques, y compris la restriction mutuelle, la corrélation, la dépendance corrélative)* ;
- *Apprentissage statistique* ;
- *Extraction des liens*.

Pour le traitement des données structurées, l'on peut utiliser les langages de requêtes (query languages) tels que SQL (Bjeladinovic, 2018). Les données non-structurées peuvent être traitées à l'aide d'Hadoop, qui est capable de traiter de grands volumes de données non-structurées dans un court laps de temps utilisant des techniques de regroupement (clustering) (Ghasemaghaei, 2019a).

Hadoop (High availability distributed object oriented platform : <http://hadoop.apache.org/>) est un projet développé par Apache Software Foundation, écrit en Java. Hadoop Il est composé de trois objets : HDFS (Hadoop Distributed File System), le modèle Hadoop MapReduce et Hadoop Common. Les données dans le système Hadoop sont décomposées en petites parties, nommées des blocs, et sont distribuées au travers du cluster. Ainsi, les fonctions « map » et « reduce » sont appliquées sur les petits sous-ensembles de la masse volumineuse des données, ce dernier assure l'évolutivité (scalability), nécessaire pour le

traitement du big data (Zikopoulos & Chris Eaton, 2011). Hadoop surmonte la limite des DBMS (Database Management Systems) traditionnels qui traitent seulement des données structurées. Hadoop déconstruit, regroupe et puis traite les données non-structurées ou semi-structurées au travers de MapReduce (voir la figure 12 ci-dessus pour l'illustration, développée par Khan et ses coauteurs (N. Khan et al., 2014)).

La plupart des entreprises comme Microsoft, IBM, Intel, Cloudera, etc. qui traitent de grandes quantités de données, utilisent Apache Hadoop. Elles utilisent le Map-Reduce, le modèle de programmation pour traiter ces énormes données (Thusoo et al., 2009). Map-Reduce est adopté au travers de l'implémentation Hadoop qui est rapidement diffusé dans le domaine du big data (Hashem, Anuar et al., 2016). Ce modèle de programmation fonctionne selon la méthode de division et de conquête en découpage d'un problème en plusieurs petites parties. Deux types de nœuds existent dans l'infrastructure Hadoop maitresse ou « Master » (1) ; responsable du découpage d'une tâche en plus petites parties pour les distribuer aux nœuds travailleurs « Worker » (2). Une fois que tous les « Worker » nœuds accomplissent leurs tâches, ils renvoient les petites parties au nœud maitre. Puis, le nœud maitre combine toutes les petites parties afin de fournir une solution (dans le langage professionnel - un output) à une question (problématique) posée. Les principaux avantages d'Hadoop sont le traitement des données distribuées, les tâches indépendantes, la facilité à gérer un échec (failure) partiel, le redimensionnement linéaire (scaling), le modèle de programmation simplifié. Les désavantages de cette infrastructure sont le modèle de programmation limitée, l'assemblage des ensembles de données (data set) multiples qui rend difficile et lent le traitement (Yaqoob et al., 2016 b). Un autre inconvénient majeur, c'est que Hadoop a été développé pour le traitement basé sur le lot (batch processing) et ne peut pas traiter les données en temps réel (Philip Chen & Zhang, 2014).

Les systèmes analytiques et les systèmes de management des données traditionnelles ont été développés sur les systèmes de gestion des bases de données relationnelles, dit RDBMS (Relational Database Management System) qui ne peuvent pas s'appliquer aux données semi-structurées et non structurées. De plus, RDBMS demande de l'équipement (hardware) de plus en plus coûteux.

Par exemple, le Cloud Data Services, dont les solutions autorisent un serveur à être hébergé dans le « cloud », dans une machine virtuelle, permettant une consolidation des serveurs

grâce à la meilleure utilisation des ressources matérielles hardware (Chaudhuri, Dayal, & Narasayya, 2011) qui est plus flexible et moins couteux pour les exigences de l'infrastructure du big data.

Ainsi, Hashem et ses coauteurs (Hashem et al., 2015), proposent les scénarios de traitement du big data, selon leur structuration (structuré ou non). La figure ci-dessous (Figure 13) illustre ces deux scénarios. On voit que les données passent une sorte de transformation avant d'être analysées. Dans le cas où les données seraient structurées, elles passent par un processus de prétraitement avant d'être stockées dans des bases de données relationnelles et puis, les données sont récupérées pour l'analyse.

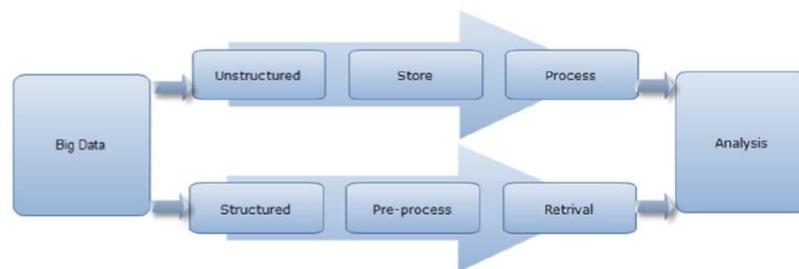


Figure 13 2 Scénarios de traitement du big data (Hashem et al., 2015)

Dans le cas des données non-structurées, elles sont d'abord stockées dans des bases de données distribuées, comme Hbase, Cassandra ou simpleDB, puis sont analysées. Le tableau 4 illustre bien la comparaison entre ces deux types de données.

	Structured data	Unstructured data
Format	Row and columns	Binary large objects
Storage	Database Management Systems (DBMS)	Unmanaged documents and unstructured files
Metadata	Syntax	Semantics
Integration tools	Traditional Data Mining (ETL)	Batch processing

Tableau 4 Comparaison entre les données structurées et non structurées (N. Khan et al., 2014)

Pour être finalement analysé, le big data non structuré passe par les étapes consécutives, illustrées dans la figure 14 ci-dessous, développée par Siddiqa et ses coauteurs (Siddiqa et al., 2016) : l'extraction (mining) au travers des techniques et algorithmes de « data mining », le nettoyage (cleaning) des données des erreurs et des redondances, l'intégration et la transformation.

Les deux processus suivants du modèle du cycle sont le stockage et le partage. Une fois traitées, les données analysées sont stockées et distribuées afin d'arriver à leur but final. De même, il faut que les infrastructures de stockage assurent un espace fiable pour le stockage et une interface solide d'accès aux données.

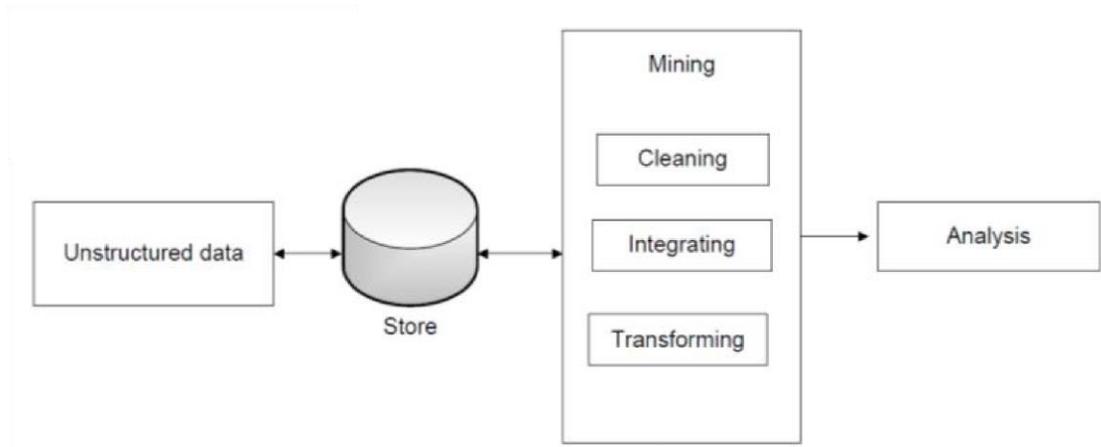


Figure 14 Processus du traitement des données non-structurées (Siddiq et al., 2016)

1.3.3 Les systèmes de stockage du big data

Les systèmes de stockage qui sont capables d'assurer les demandes et les exigences du big data sont catégorisés en deux catégories :

Première catégorie : Les DAS (direct attached storage – stockage attaché directement) où les divers disques durs sont connectés aux serveurs. Chaque disque dur reçoit une quantité limitée de ressources entrée/sortie (I/O-input/output) qui est gérée par les applications individuelles. Ainsi, le DAS est applicable pour les serveurs qui sont interconnectés à petite échelle. Cette évolutivité (scalability) faible assure une forte capacité de stockage, mais limite l'extensibilité (expandability), la capacité de mise à jour (upgradeability) et les NS (Network Storage – stockage en réseau). Le stockage en réseau est classifié en NAS (Network Attached Storage – Stockage attaché au réseau) et SAN (Storage Area Network – Stockage aux zones de réseau). Le NAS est un dispositif de stockage qui soutient le réseau. Il est connecté directement au réseau au travers d'un commutateur ou un concentrateur, un switch ou un hub, via les protocoles TCP / IP. Dans le NAS, les données sont transférées en fichiers. Le serveur NAS

peut indirectement avoir accès au dispositif de stockage au travers des réseaux, ainsi la charge d'entrée/sortie (Input/Output load) sur le serveur NAS est considérablement plus légère que celle sur un serveur DAS.

Le système de stockage des données SAN est indépendant quant au stockage sur le réseau local (LAN –Local Area Network).

Ainsi, l'organisation des systèmes de stockage des données (DAS, NAS et SAN) peut être divisée en trois parties :

- *L'ensemble de disques (disk array)* où la fondation d'un système de stockage fournit une garantie fondamentale ;
- *La connexion et les sous-systèmes de réseau*, qui connecte une ou plusieurs baies de disques et serveurs ;
- *Le logiciel de la gestion de stockage*, qui supervise le partage des données, la gestion de stockage et les tâches de reprise après une erreur pour les serveurs multiples.

Deuxième catégorie : La deuxième catégorie de systèmes qui sont capables de stocker le big data, sont les DSS (Distributed Storage System – Système distribué de stockage). Les trois principales caractéristiques de DSS pour assurer le stockage, un traitement efficace et l'analyse des données de gros volumes, sont la cohérence, la disponibilité et la tolérance. La question de la cohérence survient lorsque pour le stockage du big data par les multiples serveurs demandent un système distribué de stockage. Par conséquent, les cas de défaillance du système (server failure) peuvent arriver et pour assurer la disponibilité des données pendant la panne du serveur, les données sont distribuées en diverses pièces qui sont stockées sur plusieurs serveurs. Et donc, il en résulte des pannes du serveur et du stockage parallèle que les copies des données générées sont incohérentes dans de multiples endroits. Pourtant, le principe de la cohérence exige que les multiples copies des données soient identiques dans un environnement du big data.

Un système distribué de stockage fonctionne dans les multiples serveurs de divers endroits : quand le nombre des serveurs augmente, la probabilité de défaillance augmente aussi. Pourtant, l'ensemble du système doit répondre aux exigences de l'utilisateur en matière d'opérations de lecture/écriture (reading and writing operations). Ainsi, dans un système distribué du big data, la qualité du système (QoS –Quality of System) est indiquée par la disponibilité.

Dans un système distribué, de multiples serveurs sont liés au travers du réseau. Il faut que ce système soit capable de supporter les problèmes induits de défaillances des réseaux et le stockage distribué doit être efficace même quand le réseau est partitionné (partitioned). Et donc les défaillances ou les congestions doivent être anticipées.

La phase suivante du cycle est la sécurité : cette étape décrit la sécurité des données, les structures de gouvernance, les organisations, les agendas. Une forte attention doit être contribué au développement des systèmes, des outils, des procédures de l'usage et de la distribution des données pour légitimer leur confidentialité, sécurité et protection de la propriété intellectuelle.

Et la dernière phase du cycle de vie des données, c'est la récupération, la réutilisation et la découverte des données. La récupération des données doit assurer la qualité des données, la valeur ajoutée et la préservation des données par leur réutilisation afin de découvrir des informations nouvelles et précieuses.

1.3.4 Traitement du big data par la société JKL Armenia

La société JKL Armenia⁹, positionnée comme un fournisseur des informations innovantes et fiables aux acteurs de l'industrie du divertissement ; spécialisée dans les jeux vidéo et le merchandising, capte des données massives sur les ventes et le streaming. Elle fonctionne sur le marché arménien et a pris part dans notre étude quantitative. Il est présenté le processus de production d'un insight au travers du big data ; que nous allons détailler dans les lignes qui suivent.

⁹ <https://www.jklarmenia.com/>

JKL Armenia utilise, une plateforme EMBER (proposée par une association Cloudera et MetiStream) pour traiter le big data en temps réel.

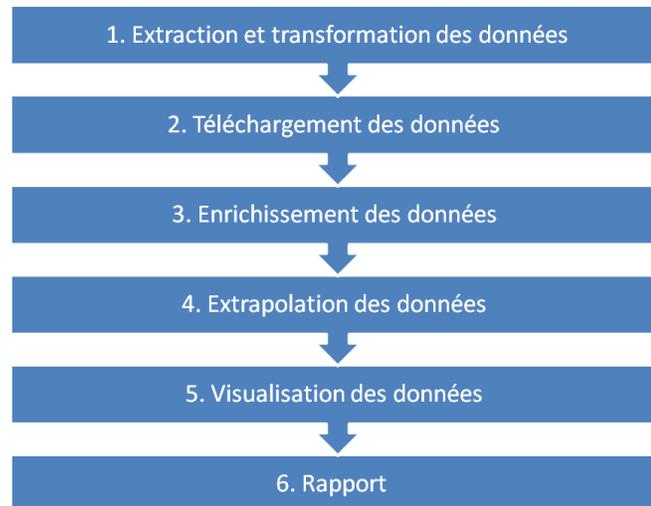


Figure 15 Processus de production d'un insight par l'exploitation du big data

Le processus de production d'un insight au travers du big data, présenté dans la figure 15 passe par les 6 étapes suivantes :

(1) Extraction et transformation des données : Ember « mange » les fichiers des données brutes dans leur format original et les transforme à un format unique. Ce premier processus est complètement automatisé et basé sur des règles prédéfinies ;

(2) Téléchargement des données : tous les fichiers sont téléchargés dans un Ember lac de données (data lake), prêts à se transformer aux « insights » ;

(3) Enrichissement des données : la société construit une base de données la plus complète et la plus fiable de l'industrie avec plus de 240 000 articles et références de produits existants. L'enrichissement automatisé se fait en continuité au travers des produits basés sur l'IA, la cartographie des ventes et de l'audience ;

(4) Extrapolation des données : au travers des informations sur les tendances, spécificités et dynamiques des marchés qui permettent d'avoir une image plus complète sur les choses ;

(5) Visualisation des données : les ventes et l'information sur « streams » se font sur une plateforme en ligne ;

(6) Rapport : d'une façon automatisée, le client peut souhaiter, recevoir sur son courriel un rapport quotidien, hebdomadaire, mensuel sur les aspects préférés.

Par cette très brève présentation, nous avons essayé de construire une vision claire sur un cycle détaillé des processus par lesquels passe le big data pour être utilisé par une entreprise, détaillée dans la littérature et par une entreprise concrète.

Chapitre 2 Efficacité d'usage: Fondements théoriques

Ce travail de recherche est dédié à l'étude de l'efficacité d'usage dans le contexte des caractéristiques spécifiques du big data. L'objectif principal de ce chapitre est de présenter les fondements théoriques du concept de l'efficacité d'usage, définie et développée préalablement pour un système d'information. Cette présentation des fondements théoriques facilitera la compréhension du phénomène dans le contexte du big data.

2.1 Travaux fondateurs

Depuis les années 1970, les études sur le construit de l'usage du système jouent un rôle central dans les recherches du SI (Barkin & Dickson, 1977). Les études sur ce construit restent dominer encore de nos jours les recherches en SI (Córdoba, Pilkington, & Bernroider, 2012), (Barki, Titah, & Boffo, 2007), (Agarwal R., 2000).

En 1997, P.B. Seddon (1997) publie un papier où il propose une version spécifiée et légèrement étendue du modèle de succès de SI de DeLone et McLean's (DeLone & McLean, 1992). L'objectif était de fournir une clarification des ambiguïtés et confusions contenues dans le modèle, en plus de fournir une justification. Dans sa critique P.B. Seddon touche l'idée que *le seul usage du système n'est pas suffisant* pour obtenir un avantage ; ainsi s'approchant à l'idée que *l'usage doit être efficace*. Quelques ans plus tard, Marcolin et ses coauteurs (Marcolin, Compeau, Munro, & Huff, 2000) dans leur papier traitent le problème de maximisation des avantages des investissements en TI, où ils abordent pour la première fois la question de *l'efficacité d'usage*. Dans leur papier, ils proposent ainsi de passer de l'étude de l'usage à l'étude de l'efficacité d'usage.

Boudreau et Seligman (2005) soulignent que les avancements dans les recherches menées sur le construit de l'usage du SI ne couvrent pas les questions pour une technologie complexe. Ils ont étudié la qualité d'usage d'une technologie complexe en faisant une distinction entre l'usage et la qualité d'usage. Dans ce travail, ils ont étudié les facteurs qui influencent la qualité d'usage d'une TI complexe.

Le construit de l'usage d'un SI a été ainsi largement étudié. Elle est étudiée surtout dans la littérature sur la théorie de l'adoption de la technologie et la diffusion de l'innovation (DIT) (Rogers, 1983), la théorie de l'action raisonnable (TRA) (Fishbein & Ajzen, 1975), la théorie de comportement planifié (TPB) (Ajzen, 2002), la théorie décomposée du comportement planifié (Taylor & Todd, 1995), le modèle d'acceptation de la technologie (TAM) (F. D. Davis, Bagozzi, & Warshaw, 1989), le modèle d'acceptation de la technologie 2 (TAM2) (Venkatesh & Davis, 2000), le modèle d'acceptation de la technologie 3 (TAM3) (Venkatesh & Bala, 2008).

Dans ces travaux les auteurs ont fait des efforts pour conceptualiser la qualité de l'usage par un modèle qui est basé sur les 3 éléments principaux : *la formation, l'apprentissage et leur influence sur les cognitions*. Le modèle présente les changements dans les croyances (beliefs) des utilisateurs finals de manière qu'ils ont approprié la nouvelle technologie complexe.

2.2 Efforts de définition d'usage du système

La revue de littérature constate que dans le domaine du SI les efforts de définition d'usage du système n'ont pas donné un résultat consensuel. En 1977, Barkin et Dickson (1977) ont proposé leur définition à l'usage : «*L'utilisation, l'usage ou l'usage d'un système se réfère à l'inclusion des données générées par le système de la part du décideur dans leur système de HIP (Human Information Processing¹⁰)* ». Dans cette logique on ne considère qu'un SI est utilisé quand la sortie du SI est organisée et/ou manipulée et/ou intégrée dans le processus de prise de décision.

Burton-Jones et Straub (2006) de leur part, proposent de considérer l'usage du système comme une activité qui contient trois éléments principaux: utilisateur (user), système et tâche (task). Dans cette activité, *l'utilisateur (user)* est une personne physique qui emploie un SI dans une tâche. Ici on suppose que son comportement soit possible d'étudier au niveau individuel. Ce construit a été ainsi bien conceptualisé par Lamb et Kling (2003) en tant qu'un acteur social ; *le système d'information* est perçu comme un artéfact qui fournit des représentations d'un ou plusieurs domaines de tâches. Et enfin, *la tâche* est une activité dirigée vers un but qui est menée par l'utilisateur.

¹⁰ Un système cognitif qui a une capacité d'organiser, manipuler et intégrer des données pour les décideurs

2.3 Théorie de l'efficacité d'usage (TEU)

Ainsi, dans ce passage de l'étude de l'usage à l'étude de l'efficacité d'usage, Burton-Jones et Grange (2013) ont fait de premiers efforts de proposer une théorie de l'efficacité d'usage.

Pour définir l'efficacité d'usage, les auteurs (Burton-Jones & Grange, 2013) mettent en avant les quatre conditions :

- On reconnaît que *l'usage* peut survenir à tout niveau d'analyse, mais le focus s'est mis au niveau individuel dans cette étape initiale de développement de la théorie ;
- On considère que les systèmes ne sont jamais utilisés juste pour « être utilisés ». Ils sont utilisés pour atteindre à certains objectifs (pour l'efficacité d'usage, on suppose que l'objectif soit n'importe quel point final pour lequel le système est utilisé) ;
- Achèvement (accomplissement) d'un but a des qualités objectives, peut-être difficile de l'évaluer dans certains cas, mais il n'est pas complètement subjectif ;
- On suppose que les différentes parties prenantes peuvent avoir des visions différentes sur les buts ou les objectifs pour lesquels le système est utilisé.

Pour ces auteurs, l'efficacité d'usage concerne l'usage d'un système de manière qu'il contribue à atteindre aux objectifs fixés pour le(s)quels le système est utilisé.

Pour développer la théorie de l'efficacité d'usage, Burton-Jones et Grange se sont appuyés sur une théorie générale du SI, la théorie de représentation (RT). Le point de départ de cette théorie c'est qu'un SI fournit des représentations ; et pour en obtenir de la valeur, l'utilisateur doit pouvoir accéder à ces représentations par la surface du système (apparence d'un système à ses utilisateurs) et sa structure physique ; et les utilisateurs de leur part, ont besoin des représentations fidèles et que ces dernières les permettent d'agir (Gregor, 2006).

2.3.1 Théorie générale du SI

Ainsi, pour fournir clairement les fondations de cette théorie, nous allons proposer d'abord une brève revue de la littérature sur la théorie générale du SI. Nous allons nous

concentrer sur la théorie de représentation (RT) qui explique au mieux l'objectif et la nature du système d'information en général.

2.3.1.1 Deux vues de représentation du SI

Le système d'information est conçu comme une représentation par les « artefacts » (artifactual) des systèmes du monde réel ; ainsi l'objectif est de modeler le point de vue de quelqu'un ou d'un groupe sur les états ou le comportement des systèmes de ce monde réel (Wand & Weber, 1995). Ces auteurs s'appuient sur les deux vues pour représenter le phénomène du SI :

- **Vue externe** traitant le système d'information comme un « black box » qui livre des services particuliers à l'organisation, affecte les vies de ses utilisateurs et évoque de certains types de préoccupations managériales. Dans cette perspective, on s'intéresse aux processus utilisés par les parties prenantes pour définir leurs besoins en SI, aux changements des pouvoirs formels ou informels qui surviennent parmi les utilisateurs, une fois l'organisation met en œuvre un SI, en plus aux façons dont l'organisation peut utiliser un SI pour obtenir un avantage compétitif.
- **Vue interne**, s'intéresse aux caractéristiques dont un SI doit posséder afin de réaliser les objectifs qu'y sont attribués. Dans cette perspective, on se focalise plutôt sur les attentes et besoins des utilisateurs, sur les aspects de bon fonctionnement du système, y compris les aspects techniques en matière de vitesse et du temps de réponse nécessaire.

Wand et Weber se focalisent sur la perspective de vue interne du SI dans leurs travaux de recherches, et s'intéressent surtout au système d'information, comme un artefact indépendant qui est intégré pour atteindre à certains objectifs (Wand & Weber, 1988), (Wand & Weber, 1990), (Wand & Weber, 1995), (Ron Weber, 2003). Il faut noter, que ces auteurs ne dénigrent pas le rôle majeur des aspects sociaux et organisationnels dans le développement, la mise en œuvre et l'usage d'un SI, mais juste croient que par un découpage des vues en interne et externe, et une concentration des efforts sur un sujet d'étude concret, peuvent assurer le succès dans les recherches du domaine.

2.3.1.2 Trois structures du SI

Par sa nature, un système d'information est consisté de 3 structures : *la structure de surface, la structure profonde et la structure physique*, figure 16 (Wand & Weber, 1995).

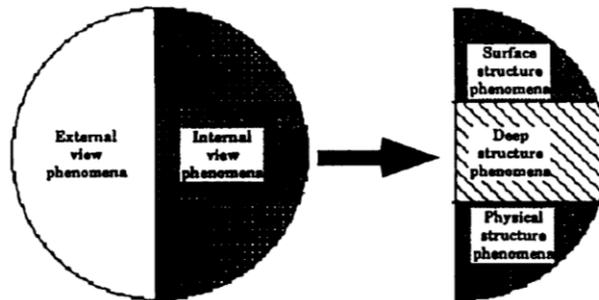


Figure 16 Structures d'un SI (Wand & Weber, 1995)

L'apparence d'un système à ses utilisateurs s'appelle la **structure de surface** d'un SI. La **structure profonde** reflète le sens du système d'un monde réel auquel un SI est destiné à modeler. Et la **structure physique** reflète la nature et la forme de la technologie utilisée pour la mise en œuvre du système.

Types of information system		Examples of system components		
Category	Example	Representations	Surface structure	Physical structure
Function IT ^a	Word processing	Representations of thoughts in a user's mind	Interfaces through which a user can read, edit, and reorganize text	Computer(s) on which the program and document are stored and can be accessed
Network IT ^a	Email	Representations of messages sent and received among individuals over time	Interfaces through which a user can read and write emails and organize past emails, such as into folders	Clients, servers, and networks through which messages are copied, exchanged, and stored
Enterprise IT ^a	Enterprise-wide software	Representations of the state of a work process in a business (such as the state of an inventory supply chain)	Forms through which users enter data and the menus and screens that users access to perform tasks	Clients, servers, and networks on which data is stored, programs are executed, and messages are sent
About reality ^b	Accounting system	Representations of the financial situation of a person or business	Forms through which users enter data and menus and screens that users access to perform tasks	The computer(s) on which the program and data are stored and can be accessed
For reality ^b	Decision support system	Representations of initial conditions (such as the decision situation and decision criteria), and suggested decisions and their explanations	Forms through which users enter data and output views provided to users to convey recommended decisions	The computer(s) on which the program and data are stored and can be accessed
As reality ^b	Video conferencing system	Representations of two or more people participating in a meeting	Viewing window through which users can view participants and icons they can click to change viewing settings	The physical screens, communication devices, and networks through which calls are made and viewed

Tableau 5 Exemples des composants du système suite les types du SI (Burton-Jones & Grange, 2013)

Pour une meilleure illustration de ces concepts, on peut s'appuyer sur le tableau 5 qui montrent des exemples clairs pour divers systèmes d'information (Burton-Jones & Grange, 2013).

Ainsi, on voit qu'un SI est lié à la compréhension et la représentation du monde réel. On comprend bien que les systèmes de représentation peuvent être différents, mais leur but est le même, d'aider à comprendre au mieux un état réel du monde ou de la situation.

Ainsi, pour un bref résumé sur la théorie de la représentation nous pouvons rappeler ici les quatre points importants :

- Le système d'information est représenté par deux vues – interne et externe ;
- Chaque SI est composé de trois éléments – ses représentations ou sa structure profonde, la structure de surface et la structure physique ;
- La structure de surface et la structure physique d'un SI assure l'accès à sa structure profonde, autrement dit aux représentations du système d'information ;
- Le niveau de la fidélité des représentations est toujours mis en question. La structure de surface et la structure physique peuvent être modifiées en fonction des développements technologiques et les changements des conditions sociales pour assurer des représentations améliorées.

2.3.2 Visions de développement de TEU

Pour développer leur théorie d'efficacité d'usage Burton-Jones et Grange (2013), s'appuient sur une vue qui combine la sémantique et la pragmatique. Toujours en s'appuyant sur la caractéristique de fidélité, on se pose deux types de questions :

- (1) Si le système représente tous les éléments du domaine dans tous leurs détails ? (un jugement plutôt universel ; au sens sémantique)
- (2) Si le système représente tous les éléments dans le domaine qui sont pertinents aux besoins d'utilisateurs ? (un jugement spécifique, contextuel ; au sens pragmatique).

Au cœur de la théorie de la représentation, l'objectif est de fournir des représentations qui répondent aux besoins des utilisateurs. Ainsi, la théorie de l'efficacité d'usage, développée par

ces auteurs, adopte une combinaison de visions sémantiques et pragmatiques. Il est utile de prendre en compte les besoins de différentes parties prenantes – utilisateurs qui ont de différents besoins en représentation dans les contextes différents.

Une autre vision importante dans le développement de la théorie d’efficacité d’usage, c’est que l’on doit considérer le système dans son ensemble et à tous les niveaux de ces structures pour bien comprendre son efficacité d’usage (Figure 17). Tandis que dans la théorie de la représentation, l’on traite le système comme il est perçu au niveau de sa structure profonde. Et dans la théorie de l’efficacité d’usage, les questions sur la réalité et comment on l’aperçoit au niveau de sa structure de surface et de physique sont les aspects importants à traiter.

2.3.3 Efficacité d’usage d’un SI : Développement de théorie à 3 étapes

Nous allons présenter la théorie de l’efficacité d’usage par les 3 étapes de construction, proposées par Ostrom (2009). Ces 3 étapes de construction de la théorie sont : le cadre théorique (Framework), l’application du cadre théorique et le modèle proposé à tester.

2.3.3.1 Cadre théorique

Les hypothèses métathéoriques issues de la TR (Théorie de Représentation) qui vont contribuer à étudier l’efficacité d’usage, c’est le réalisme critique (utilisé pour expliquer la nature des représentations) et la téléologie (utilisée pour expliquer leur finalité).

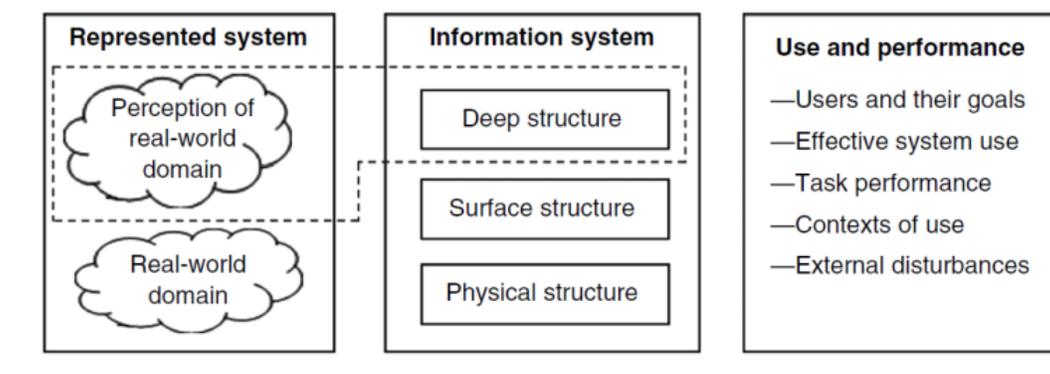


Figure 17 Niveaux de structures du SI (Burton-Jones & Grange, 2013)

L’on suppose que le réel existe indépendamment de nos perceptions : *il existe un réel en soi indépendant de, et antérieur à l’attention que peut lui porter un humain qui l’observe* (Bhaskar, 2013). Nous ne pouvons connaître le réel que par nos perceptions qui sont subjectives

et pas parfaites (le réalisme critique) et le comportement humain est motivé par le but (la téléologie). Ainsi, l'on peut supposer que l'homme est motivé de créer des systèmes plus parfaits (goal directed action) afin d'avoir des représentations améliorées du réel, un processus qui ne peut être jamais atteint.

Pour étudier l'efficacité d'usage, Burton Jones et Grange (2013) s'appuient sur un cadre général de « basic cybernetic control » (Johnson & Wiener, 1949), qui reflète bien les réflexions ci développées (Figure 18).

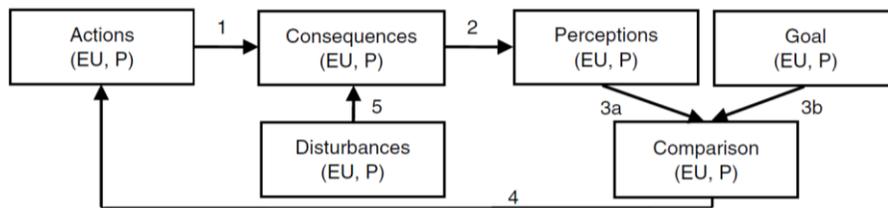


Figure 18 Cadre général pour étudier l'efficacité d'usage (Burton-Jones & Grange, 2013)

Pour dresser le schéma (Figure 18), les auteurs se basent sur la déduction faite de la théorie de la représentation : pour atteindre à l'efficacité d'usage l'on passe par un processus continu, sujet aux erreurs. L'efficacité d'usage doit comprendre l'accès aux représentations du système au travers de sa structure de surface et de physique, comme la base de l'usage d'un SI est d'obtenir l'accès à ces représentations. Ce dernier induit à considérer l'interaction transparente comme une dimension d'efficacité d'usage. De leurs parts, les représentations obtenues du système doivent être fidèles où l'utilisateur désire obtenir les représentations plus fidèles du domaine. Et donc la fidélité représentationnelle est considérée comme une autre dimension de l'efficacité d'usage. L'on comprend aussi que les représentations engendrent les actions (qui font agir). Les utilisateurs désirent « recevoir » les représentations fidèles, car elles les permettent d'agir. Par conséquent, l'action informée (aisée) est la dimension suivante de l'efficacité d'usage. Comme un SI est composé d'un ensemble des structures complexes et cette complexité induit à les apprendre ; l'apprentissage des composants d'un SI peut aider l'accès aux représentations. Dans cette logique, l'apprentissage est considéré comme un facteur influençant (driver) l'interaction transparente.

L'on suppose que les représentations d'un SI sont faillibles et doivent être accédées au travers d'un ensemble des structures complexes. Les systèmes d'information sont

intrinsèquement malléables ; par conséquent, un utilisateur peut améliorer son accès aux représentations fidèles au travers des actions d'adaptation. Ainsi, l'adaptation ou les actions d'adaptation est l'autre facteur influençant l'interaction transparente et la fidélité représentationnelle.

Et enfin, comme les représentations d'un système d'information sont faillibles, l'on suppose que les représentations plus fidèles prévoient une meilleure base pour l'action. Ainsi, un utilisateur doit apprendre le domaine et connaître à quel point un SI le représente fidèlement, pour comprendre ce qu'il faut accéder, comment y appliquer les adaptations et exploiter. Ainsi, l'on comprend que l'adaptation et l'apprentissage sont un double facteur influençant la fidélité représentative, y compris l'apprentissage comme un facilitateur (enabler) de l'action informée (aisée).

Cette brève introduction des éléments-clés nous permet d'interpréter la figure 18 comme suit :

- L'homme prend des actions (agit) pour créer et utiliser le système d'information afin d'obtenir de meilleures représentations (l'action humaine a des conséquences) ;
- L'on aperçoit les conséquences de la création et de l'usage de ces actions (l'homme aperçoit les conséquences de ses actions) ;
- Si les conséquences perçues ne correspondent pas au but initial (pour lequel l'action a été mise en place) (3a, b) et donc l'homme aperçoit, ses perceptions avec ses buts initiaux, et de nouvelles actions sont entreprises pour l'amélioration ;
- L'homme met en œuvre les actions correctives si les objectifs n'ont pas été atteints ;
- Les perturbations peuvent arriver à tout moment.

EU et P dans la figure consignent l'efficacité d'usage (EU) et la performance (P). Ils sont inclus dans chaque élément ce qui signifie que l'homme peut agir pour améliorer l'efficacité d'usage, également, il peut mettre en place des actions pour l'amélioration de la performance.

2.3.3.2 Application du cadre théorique

Ce cadre a été appliqué en système d'information par Carver et Scheier (1998) et en psychologie par Liang et Xue (2009).

Dans une deuxième étape de construction de théorie, l'on a appliqué ce cadre sur le contexte d'efficacité d'usage et la performance. La figure montre que l'action humaine à chaque point a un but qui le conduit pour faire une action/agir. Cette action est mise en place pour atteindre un objectif initialement fixé. L'homme peut mettre en place aussi des sous actions pour atteindre un certain nombre de sous-objectifs afin d'arriver à un niveau de la satisfaction de réalisation de l'objectif global. Un objectif global (principal) a été fixé pour lequel des actions principales ont été mises en œuvre ; la panne de machine a évoqué une perturbation, et donc un sous-objectif a été fixé de régler cet inconvénient, pour lequel des sous actions ont été mises en place (Carver & Scheier, 1998), (Gasser, 1986).

Ainsi, les auteurs (Burton-Jones & Grange, 2013) ont défini deux types d'objectifs utilisés dans leur théorie :

- Efficacité d'usage, qui est considérée comme un objectif moyen ;
- Performance, qui est un objectif global principal souhaité.

Il faut noter que dans la théorie d'efficacité d'usage, l'on ne concentre que sur les actions qui améliorent l'efficacité d'usage (1a) : nous illustrons un petit zoom sur le lien 1 (Figure 19) pour une meilleure compréhension.

2.3.3.3 Modèles proposés à tester

Pour la suite, nous allons traiter la phase de la construction de la théorie d'efficacité d'usage, et présenter les modèles testables qui en sont issus. Pour cette étape, les auteurs ont présenté les deux modèles : l'un est concentré sur la nature de l'efficacité d'usage et l'autre, sur les facteurs qui l'influencent.

Ce premier modèle testable traite le lien 1b (Figure 19) qui montre comment l'efficacité d'usage peut améliorer la performance. La figure ci-après qui illustre ce premier modèle testable, recense certaines dimensions des construits.

Les dimensions dans la figure 20 reflètent un état du haut à bas. Comme c'était déjà noté, la surface du système et sa structure physique doivent assurer l'accès de l'utilisateur aux représentations (prémisse de la théorie de représentation).

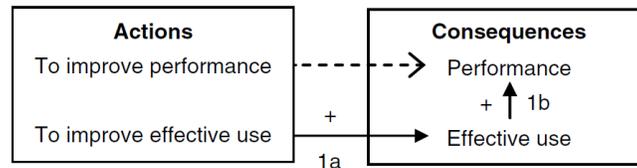


Figure 19 Zoom sur le lien 1 de figure 18 (Burton-Jones & Grange, 2013)

Ainsi, on définit la première dimension de l'efficacité d'usage, comme *l'interaction transparente* (transparent interaction), qui améliore la performance grâce au gain du temps des utilisateurs, avec une croissance d'efficience.

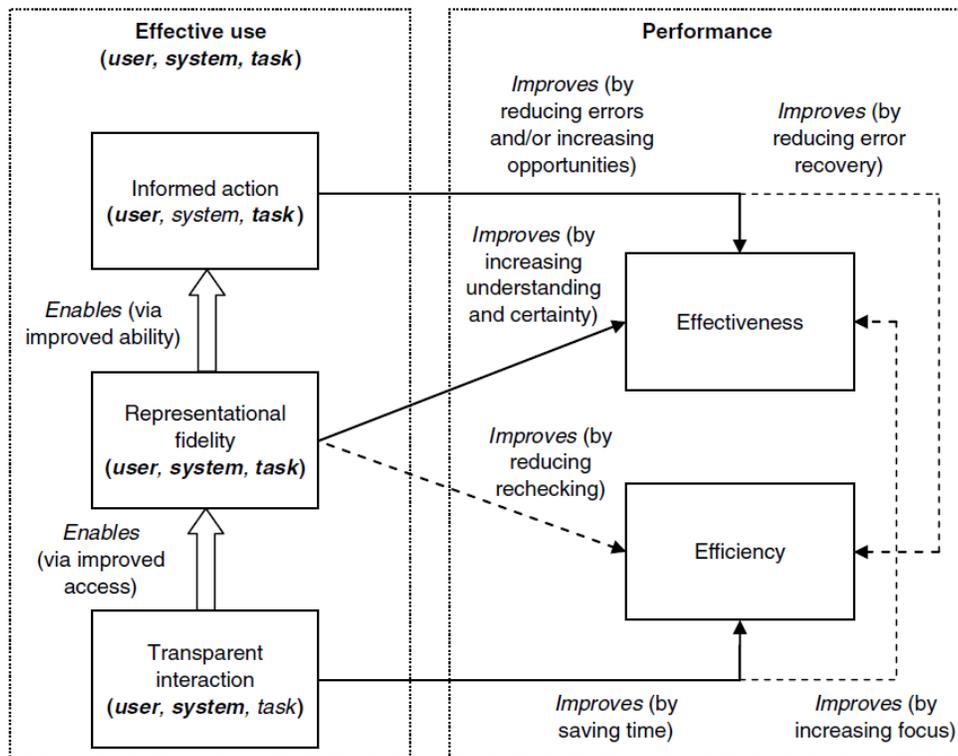


Figure 20 Efficacité d'usage et ses effets sur la performance (Burton-Jones & Grange, 2013)

La deuxième dimension est *la fidélité représentative* (representational fidelity) qui consigne que les représentations obtenues du système par l'utilisateur reflètent fidèlement le domaine représenté.

Et la dernière dimension de la figure est une action informée (aisée – informed action) dans la logique que les représentations fidèles obtenues du système, par l'utilisateur, permettent d'améliorer ses actions. Le terme d'action informée ou aisée, l'expression anglo-saxonne « informed action » a été proposée par les auteurs Burton-Jones et Grange. Les termes pour les 2 autres dimensions ont été adoptés. L'interaction transparente (« transparent interaction ») a été adoptée des études menées de Lave and Wenger (1991), de Winograd et Flores (1987). Le terme de la fidélité représentationnelle (« representational fidelity ») a été adopté des études de Burton-Jones et ses coauteurs (2009) et de Parsons et Cole (2005).

Dans la figure (Figure 20), on peut noter aussi une hiérarchie des dimensions du bas vers le haut : chaque dimension en bas est nécessaire mais pas suffisante pour une dimension au-dessus. Comme la figure montre, l'utilisateur doit pouvoir accéder aux représentations les plus transparentes ; et doit pouvoir obtenir des représentations fidèles pour prendre des actions informées (aisées). Ainsi, l'on comprend que l'efficacité d'usage prend en compte les deux aspects à la fois : l'usage du système au niveau individuel, plus l'usage de l'information par l'utilisateur, fourni du système (information).

2.3.4 Développement de la discussion sur la théorie d'efficacité d'usage

Comme il était noté précédemment (Burton-Jones & Straub, 2006), l'usage du système consiste de 3 éléments-clés : l'utilisateur, le système et la tâche.

Ici, nous trouvons important d'intervenir sur ces éléments clés par de brèves explications fournies par Klein et ses coauteurs (1997) afin de clarifier les développements de discussion sur la théorie d'efficacité d'usage (TEU).

Plus la personne est motivée à bien performer dans sa tâche et possède des connaissances sur cette tâche et son système, plus elle perçoit précisément le niveau de l'efficacité d'usage et de performance. Dans la figure 21, la performance personnelle liée à une tâche concrète l'on peut

présenter comme une fonction de facteurs déterminants fournis de Campbell (1990), (Campbell & Wiernik, 2015).

$$\text{Performance} = f(\text{declarative knowledge, procedural knowledge and skills, choice to expend effort, choice of degree of effort to expend, choice to persist})$$

Figure 21 Facteurs déterminants la performance individuelle (Campbell, 1990)

La théorie de la performance personnelle liée à une tâche concrète suggère que l'expérience, la connaissance et l'effort peuvent potentiellement affecter les capacités individuelles à accomplir la tâche concernée. Dans cette théorie, Campbell (1990) fait l'accent sur la relation entre l'effort et la performance en tâche.

Plus le système est nouveau et complexe pour l'utilisateur, plus il est difficile à déterminer le niveau d'efficacité de son usage (système). Ce jugement concerne aussi la performance des utilisateurs : plus les tâches sont nouvelles et complexes pour eux, plus il est difficile de juger le niveau de leur performance dans l'accomplissement de celles-ci.

Klein et ses coauteurs (1997) soulignent encore l'importance de 2 éléments – la fréquence et la régularité des retours (commentaires) / feedback sur les niveaux de la performance et l'efficacité d'usage.

Ainsi, pour évaluer *l'usage*, l'on doit considérer dans quelle mesure l'utilisateur utilise le système dans une tâche concrète.

Pour évaluer *la fidélité représentative*, l'on considère dans quelle mesure les représentations obtenues du système par l'utilisateur reflètent fidèlement le domaine. Ici l'on comprend que les différents utilisateurs peuvent recevoir de même système les représentations de différents niveaux de fidélité. Ça dépend tout d'abord de leur efficacité d'usage.

Pour *l'action informée ou aisée*, l'on évalue le niveau de manière que l'utilisateur peut gagner de profit des représentations fidèles pour son action informée ou aisée (comment l'utilisateur exploite l'information fournie du système dans une tâche concrète).

Les facteurs clés comme les dimensions de l'efficacité d'usage sont issus de la théorie de représentation. Ils représentent deux types d'actions à prendre pour améliorer l'efficacité d'usage :

- Actions d'adaptation ; ce sont des actions qu'un utilisateur met en œuvre pour améliorer la représentation du système du domaine d'intérêt ou bien son accès au système (par sa surface ou la structure physique) ;
- Actions d'apprentissage ; ce sont des actions qu'un utilisateur met en œuvre pour connaître le système, ses représentations, sa surface ou la structure physique (ou encore le domaine représenté ou la mesure dans laquelle il représente fidèlement le domaine, ou encore la manière de comment s'engager dans les actions plus informées (aisées) à travers des représentations obtenues du système).

Comme les représentations sont faillibles, les utilisateurs prennent les actions d'adaptation afin d'améliorer d'abord leur fidélité. Dans ce sens-là, l'adaptation de la structure de surface et la structure physique ont pour l'objectif d'assurer l'accès aux représentations. Ainsi, l'on cherche les adaptations qui facilitent et contribuent à l'amélioration de l'accès.

Les actions d'apprentissage sont prises par l'utilisateur encore de raison de la faillibilité des représentations : il doit pouvoir mesurer le niveau de faillibilité de ces représentations afin d'intervenir par les actions d'apprentissage pour l'améliorer.

Dans la figure 22, l'on va discuter l'impact (effect) des actions d'adaptation et d'apprentissage sur l'efficacité d'usage.

Dans cette figure l'on voit qu'un utilisateur peut améliorer l'interaction transparente par 2 moyens : la première opportunité est de considérer le système comme il est (as is) et apprendre de manière comme il offre des représentations par sa structure de surface et de structure physique (lien 1 dans la figure 22). La deuxième opportunité d'améliorer l'interaction transparente, c'est

adapter les structures de surface et de physique de manière instruite (instruit par l'apprentissage de ces structures (lien 2).

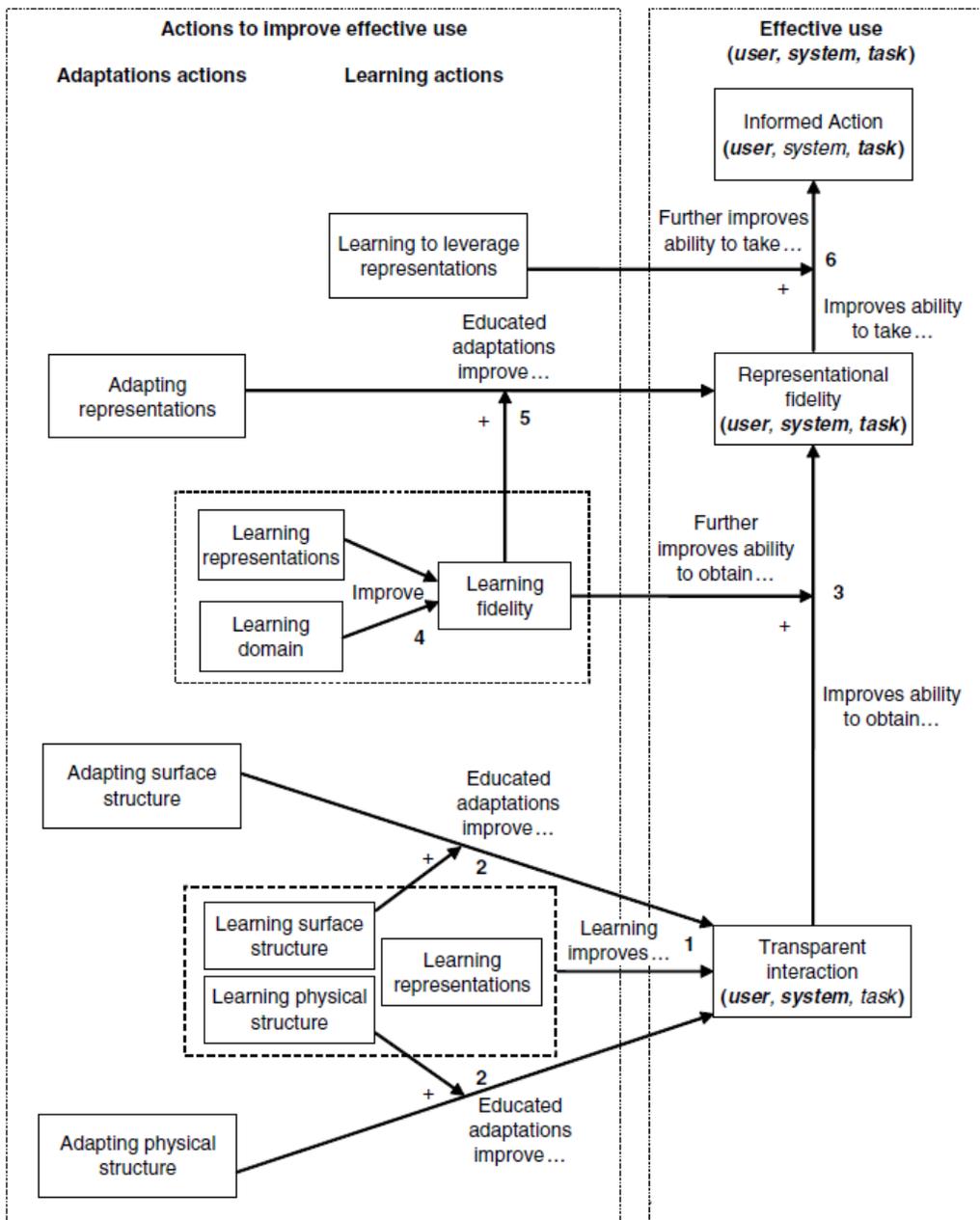


Figure 22 Facteurs influençant l'efficacité d'usage (Burton-Jones & Grange, 2013)

On peut supposer que les actions d'adaptation pas instruites ou mal instruites peuvent induire aux problèmes ou bien ne pas avoir des effets attendus. La logique est la même pour la fidélité représentative : l'on peut améliorer cette dimension en se focalisant sur l'apprentissage à quoi accéder et comment y accéder. Quand l'interaction transparente est associée avec la

connaissance de la fidélité, le lien devient plus certain, car les utilisateurs savent déjà ce qu'il faut chercher et comment y accéder (lien 3). Cette figure montre que les connaissances sur les représentations et les connaissances sur le domaine peuvent accroître la fidélité (lien 4).

L'autre méthode pour améliorer la fidélité représentative est d'adapter les représentations. Comme l'on voit dans la figure, les adaptations instruites peuvent améliorer la fidélité représentative (lien 5). Au contraire, la fidélité représentative est mauvaise, et les utilisateurs n'ont pas appris comment utiliser les représentations, donc la prise des actions mal informées va arriver (lien 6).

2.3.5 La place de la théorie d'efficacité d'usage dans les théories existantes

Ici, l'on trouve aussi important d'aborder une discussion sur la place de cette nouvelle théorie d'efficacité d'usage dans les modèles ou théories existants liées au sujet. Nous en allons résumer ci-après les plus reconnus dans le domaine.

2.3.5.1 TEU et TAM

Le TAM (Technology Acceptance Model), initialement proposé par Fred Davis, en 1985 (Legris, Ingham, & Colletette, 2003) est un modèle bien connu qui porte sur l'usage des technologies informatiques (Davis, 1989). Adapté de la théorie socio psychologique / comportemental, ce modèle a été développé par de nombreux chercheurs.

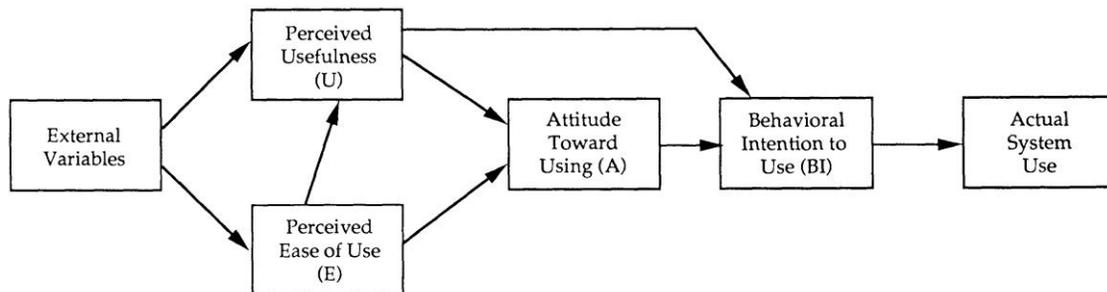


Figure 23 TAM original (Legris et al., 2003)

Etant donné l'écart entre, d'un côté des énormes investissements réalisés pour l'implantation du SI au sein des organisations, et d'un autre côté des taux de réussite relativement faibles, les chercheurs ont mobilisé leurs efforts pour identifier les facteurs qui facilitent l'acceptation d'une technologie par ses utilisateurs. Les auteurs recensent les 2 déterminants principaux de l'acceptation par l'utilisateur : l'utilité perçue et la facilité perçue d'utilisation (le modèle TAM illustré dans la figure 23).

Il faut noter que le construit de la théorie de l'efficacité d'usage proposé de Burton-Jonse et Grange, présente des similitudes avec ceux du TAM. D'abord, la définition de l'efficacité d'usage est proche de la définition de la facilité perçue d'utilisation, en plus la définition de l'interaction transparente porte des similitudes avec la définition de l'utilité perçue. Mais là, il faut aussi bien noter la différence majeure : les construits de la théorie de l'efficacité d'usage reflètent des comportements observables et le TAM est basé sur les perceptions des utilisateurs. En plus, l'objectif du TAM est d'expliquer l'acceptation de TI, pourtant la théorie de l'efficacité d'usage explique ce que les utilisateurs doivent faire pour utiliser les systèmes efficacement et augmenter leur performance (Burton-Jones & Grange, 2013).

2.3.5.2 TEU et UTAUT

Un autre modèle, UTAUT (Unified Theory of Acceptance and Use of Technology), se propose d'unifier les huit modèles déclinant le TAM. Selon ce modèle (Venkatesh, Morris, Davis, & Davis, 2003), les quatre principaux déterminants de l'acceptation par l'utilisateur et le comportement dans l'usage, qui sont considérés comme les déterminants directs, sont : l'espérance de performance, l'espérance d'effort, l'influence sociale et les conditions facilitatrices. En dehors des déterminants directs, le modèle inclut également des éléments modérateurs clés tels que : le sexe, l'âge, le volontarisme de l'usager et l'expérience. L'espérance de la performance traduit le degré de croyance d'un individu dans le système, c'est-à-dire la capacité du système à soutenir son travail, à améliorer ses performances dans son travail. L'espérance d'effort souligne le degré de facilité d'usage du système (Figure 24 illustre leur modèle proposé).

Comme noté par Surbakti et ses coauteurs (2019), (Surbakti, 2020), la Théorie de l'Efficacité d'usage (TEU) est une évolution naturelle des modèles TAM et UTAUT.

Elle représente ainsi un passage nécessaire « de l'étude de l'usage » à « l'étude de l'efficacité d'usage ».

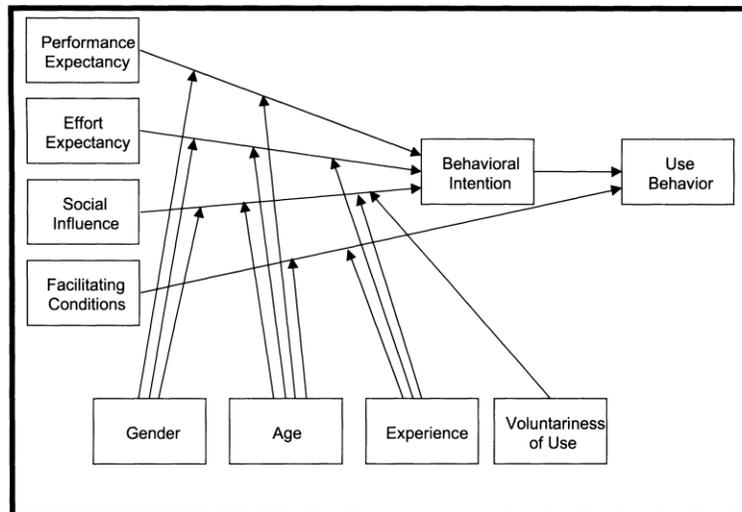


Figure 24 Modèle UTAUT (Venkatesh et al., 2003)

2.3.5.3 TEU et le modèle de DeLone et McLean

L'autre modèle à souligner, c'est le modèle de succès de SI proposé de DeLone et McLean (1992). Ses auteurs ont essayé de proposer une vue intégrée du concept du succès d'un système d'information.

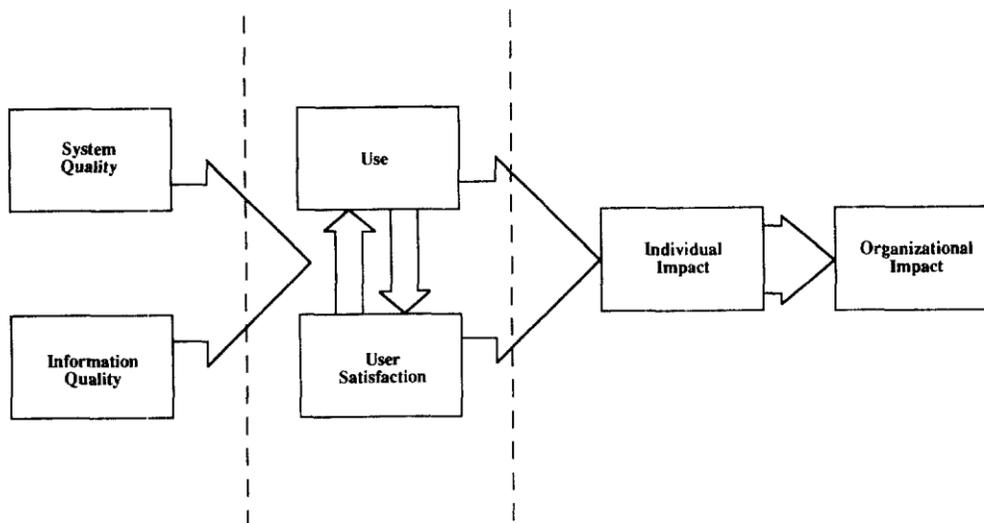


Figure 25 Modèle initial du succès d'un SI (DeLone & McLean, 1992)

Les 6 dimensions importantes du succès d'un SI sont proposées : la qualité du système, la qualité de l'information (retenu du système), l'usage, la satisfaction de l'utilisateur, l'impact individuel et l'impact organisationnel (la figure 25 illustre le modèle initial du succès proposé en 1992).

Le TEU propose une définition de la fidélité représentative qui semble d'une première vue avoir des similitudes avec la définition de « la qualité de l'information » du modèle de DeLone et McLean. Pourtant, il faut bien souligner que ces deux construits n'est pas le même. Il faut bien noter que la qualité de l'information chez DeLone et McLean, c'est une propriété du système, alors que le concept de la fidélité représentative chez Burton-Jones et Grange (2013), c'est la propriété de l'usage. En plus, le modèle du succès du SI traite aussi le lien entre l'usage et la performance, mais laisse la question sans une réponse claire en faisant un appel à le traiter dans les futures recherches (DeLone & McLean, 2003).

2.3.5.4 TEU et TTF

Une question qui était toujours dans le centre des recherches du domaine, concernait le rapport entre la performance individuelle et le système d'information.

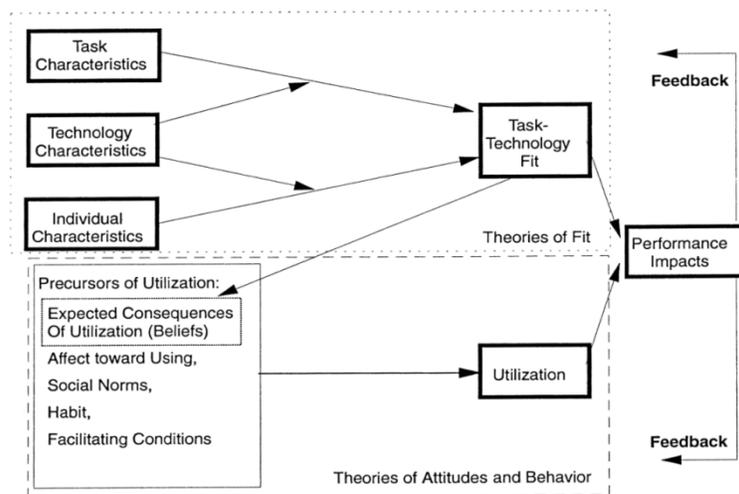


Figure 26 Modèle proposé de Goodhue et Thompson –TTF (Goodhue & Thompson, 1995)

Un modèle proposé de Goodhue et Thompson (Figure 26) (Goodhue & Thompson, 1995), affirme qu'une technologie peut avoir un impact positif sur la performance individuelle si elle est utilisée (1) et bien adaptée aux tâches qu'elle soutient (2). Ce modèle est cohérent avec le

modèle de Delone et McLean (1992), de manière que tous les deux soutiennent l'avis que l'usage et les attitudes des utilisateurs de la technologie impactent la performance individuelle. Ainsi, la théorie de TTF (Task-Technology-Fit) semble avoir un objectif similaire aussi avec la TEU, expliquant comment une TI conduit à des résultats de performances différents. La notion de compatibilité (fit) en TTF est incluse dans la notion de fidélité représentative. Bien que la théorie TTF pose un lien entre l'usage et la performance, elle ne traite pas la question d'impact, si ce lien est positif ou quelles sont les actions à entreprendre pour une amélioration. Dans ce contexte, la TEU propose une extension de TTF en se focalisant sur les types d'actions d'utilisateurs qui peuvent être mises en œuvre par l'entreprise pour améliorer l'usage et contribuer à l'efficacité d'usage.

2.3.5.5 TEU et AST

DeSanctis et Poole (1994) ont proposé une théorie AST (Adaptive Structuration Theory) pour étudier le rôle des TIs avancées dans le processus du changement au sein de l'entreprise (Figure 27).

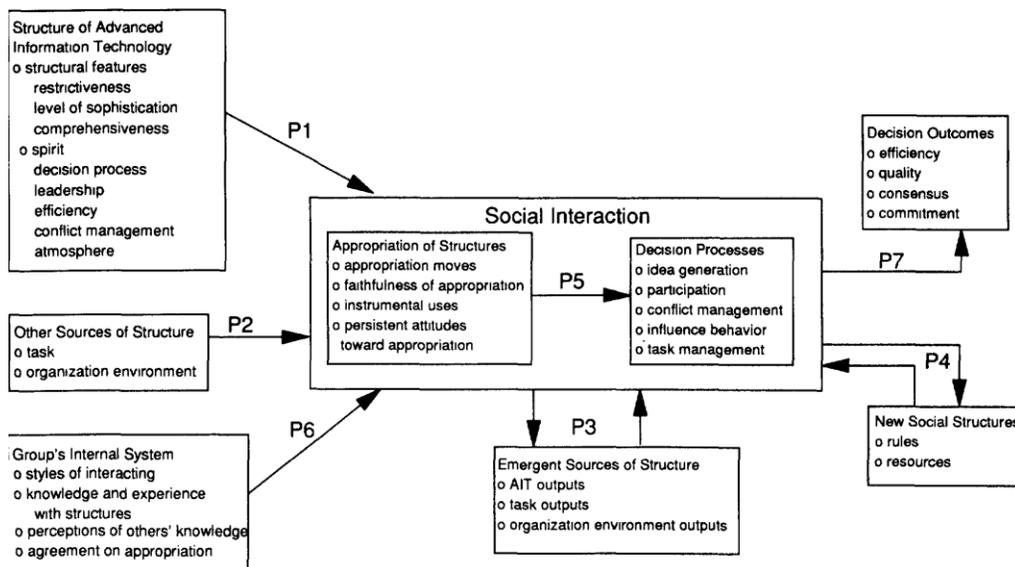


Figure 27 Construits Majeurs et les contributions de la théorie AST (DeSanctis & Poole, 1994)

Cette théorie étudie le processus du changement de 2 points de vue : les types de structures qui sont fournis par les technologies avancées (1) et les structures qui émergent dans l'action humaine quand les utilisateurs s'échangent avec ces technologies.

La théorie AST, similaire à la théorie TEU, traite les variations des impacts de TI. AST partage aussi l'avis proposé par TEU, qui suggère que l'impact de TI dépend de la manière dont elle est utilisée plutôt que du système lui-même. L'idée majeure qui contraste avec la suggestion de TEU, c'est que la théorie AST suggère que la performance est améliorée si les utilisateurs utilisent le système d'une manière cohérente (appropriation fidèle). Au contraire, TEU propose que l'efficacité d'usage soit plutôt une question d'obtention et de valorisation des représentations nécessaires pour accomplir une tâche quel que soit le niveau d'interaction fidèle avec le système (usage/système).

2.3.5.6 TEU et ISURA

Une conceptualisation de l'usage du SI au niveau individuel a été aussi proposée par Barki et ses coauteurs (2007). Elle comprend à la fois des comportements de l'interaction avec TI et les activités maintenues par les utilisateurs à s'adapter au système de chaîne « tâche-technologie-individu ». Le nouveau construit proposé par ces auteurs est appelé ISURA (IS use-related activity - activité liée à l'usage du SI) ; il est basé sur un principe de fournir une représentation fidèle de manière comment les utilisateurs utilisent le SI dans leur travail. Dans ce sens-là, l'activité liée à l'usage du SI est définie comme un ensemble des comportements entrepris par les utilisateurs liés à un contexte spécifique de « tâche-technologie-individu ». D'abord ces comportements incluent les actions que les utilisateurs entreprennent pour améliorer l'accomplissement d'une tâche.

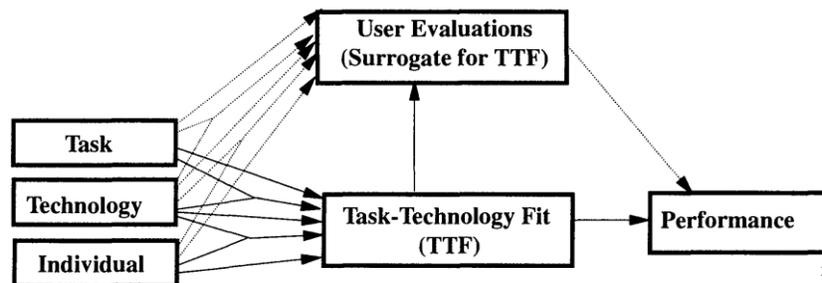


Figure 28 Modèle de TTF (Task-Technology Fit) avec les évaluations des utilisateurs (Goodhue, 1995)

Goodhue (1995) de sa part constate que les évaluations du SI fournies par les utilisateurs peuvent être utiles dans le concept de compatibilité tâche-technologie (Task-Technology-Fit / TTF (Figure 28).

Ces évaluations, bien évidemment, sont impactées de caractéristiques du système, de la tâche et de l'individu : les utilisateurs acceptent les SIs comme les outils qui les aident ou les empêchent dans l'accomplissement de leurs tâches.

ISURA considère non seulement les interactions des utilisateurs avec la TI dans le processus de l'accomplissement des tâches, mais aussi les activités des individus qui adaptent, changent ou modifient les éléments dans le contexte du tâche-technologie-individu. Barki et ses coauteurs (2007) soulignent également l'importance des actions d'adaptation, plus les actions d'apprentissage dans l'amélioration de la performance.

Si l'on compare la théorie d'efficacité d'usage avec le modèle de l'activité d'usage du SI (ISURA), la première chose que l'on remarque, c'est que Burton-Jones et Grange (2013) considèrent le construit de l'usage au sens plus strict. Si ISURA traite les activités d'amélioration de l'usage comme une intégrale de l'activité liée à l'usage, la TEU fait une distinction entre les activités prises à améliorer l'usage et les évaluations de l'usage lui-même.

Deuxièmement, ISURA propose qu'une interaction avec TI seule puisse entraîner des résultats positifs. Et la TEU confirme que l'usage doit être efficace afin d'assurer les résultats positifs.

2.3.6 Contribution de Burton-Jones et Grange (Burton-Jones & Grange, 2013)

Ces auteurs ont proposé une théorie de l'efficacité d'usage en déterminant ce que comprend l'efficacité d'usage (ses dimensions) et ce qui le motive (ses antécédents). Dans les sections précédentes, nous avons interprété en profondeur les 2 modèles que ces auteurs ont proposés. Dans ces modèles, il est considéré que les actions prises par les utilisateurs sont susceptibles d'avoir du succès dans les circonstances où les utilisateurs sont plus compétents, expérimentés, motivés avec le système ou la tâche (Sonntag, 2002):

- Les systèmes et les tâches sont simples, flexibles, familiers et indépendants des autres systèmes / tâches ;
- Les utilisateurs peuvent prendre des mesures ;
- Ils peuvent voir leurs conséquences rapidement.

Il est supposé que non seulement les actions d'apprentissage et d'adaptation prises par les utilisateurs ont une influence immédiate sur l'efficacité d'usage, mais elles peuvent impacter les contextes d'utilisateurs (ils peuvent devenir plus compétents, connaître beaucoup plus le système) ou du système (rendre le système plus ou moins complexe). Ces actions affectent l'ensemble du processus d'amélioration de l'efficacité d'usage et de la performance. Et donc amélioration de l'efficacité d'usage peut être atteinte par un processus continu, sujet à l'erreur, et quelque part imprévisible (Burton-Jones & Grange, 2013).

Nous avons essayé de présenter dans ce chapitre les principales bases théoriques de la théorie de l'efficacité d'usage d'un SI. Cette nouvelle théorie (Burton-Jones & Grange, 2013) a été détaillée par les 3 étapes de construction, proposées par d'Ostrom (2009). En fin, nous avons développé une discussion sur la place de la théorie d'efficacité d'usage dans les théories et modèles existants. L'objectif principal était de construire une compréhension complète sur les fondements théoriques de l'efficacité d'usage, pour pouvoir les traiter dans le contexte spécifique du big data.

Pour pouvoir traiter la question de notre recherche, nous proposons de développer dans le chapitre suivant, la théorie de l'efficacité d'usage dans le contexte du big data.

Chapitre 3 Efficacité d'usage du big data : Modèle conceptuel mobilisé

Ce chapitre vise à présenter les fondements théoriques de l'efficacité d'usage du big data. Pour ce faire, nous allons interpréter un modèle conceptuel de l'efficacité d'usage du big data, développé par Surbakti et ses coauteurs (Surbakti et al., 2019). Selon ses auteurs et à notre connaissance, il s'agit de la première tentative d'élaboration d'un modèle complet des facteurs qui influencent l'efficacité d'usage du big data.

Ayant un objectif ultime de contribuer au développement de la théorie de l'efficacité d'usage dans le contexte du big data, dans cette interprétation nous avons essayé de développer les bases théoriques pour chacun de 41 facteurs du modèle. En se focalisant sur son contexte du big data, nous avons interprété son ontologie qui découle du contexte général du système d'information.

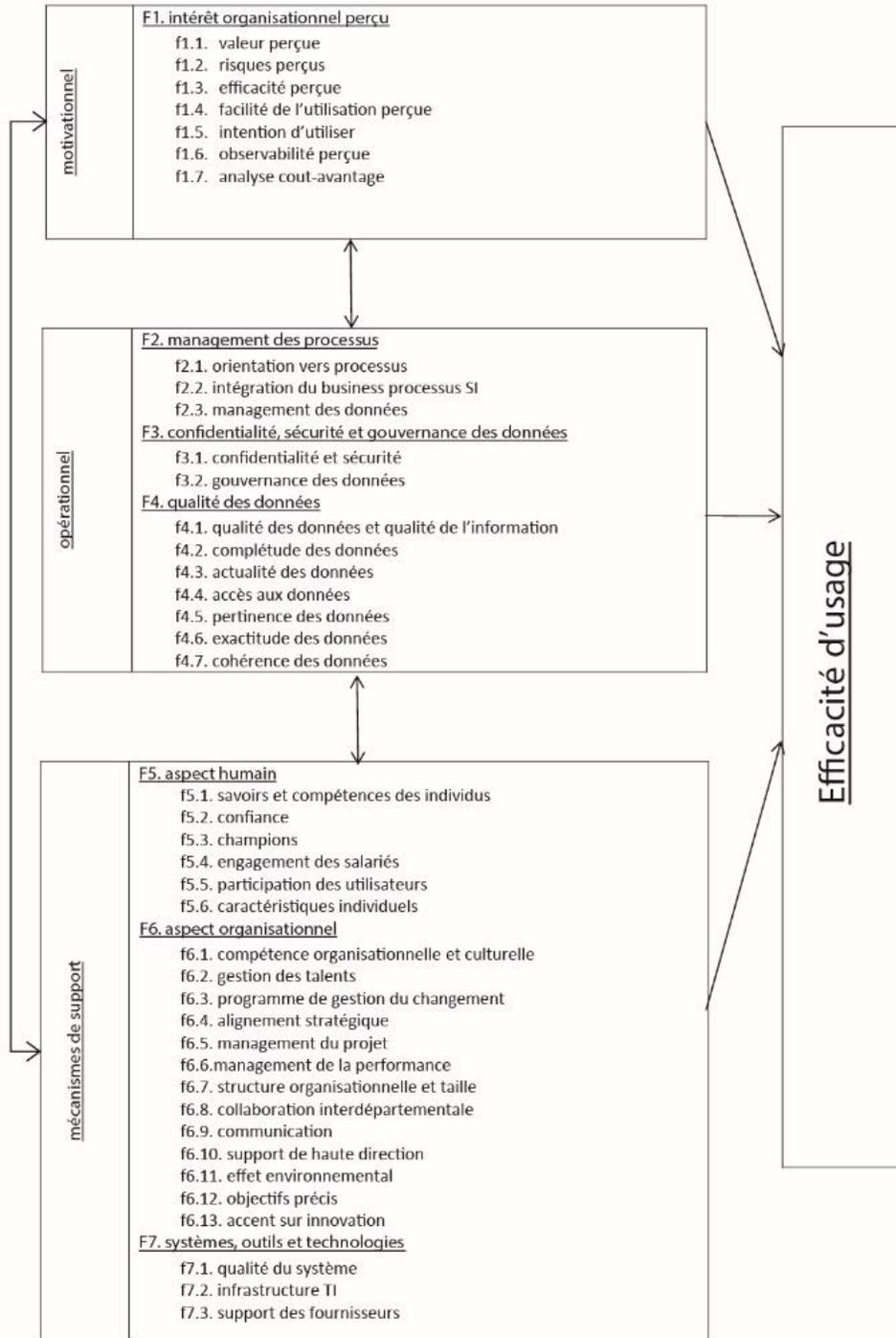
Pour démarrer le développement de ce chapitre, nous devons noter que la théorie de l'efficacité d'usage (TEU), détaillée dans le chapitre précédent, a été construite sur la base d'un SI traditionnel qui ne tient pas compte les caractéristiques spécifiques du big data (les Vs).

Ainsi, la présentation qui suit, prend bien en compte les spécificités particulières du big data.

Dans le modèle (Figure 29), les auteurs (Surbakti et al., 2019) décrivent les 7 thèmes qui comprennent les 41 facteurs qui peuvent potentiellement influencer l'efficacité d'usage du big data.

Les 7 thèmes dans lesquels sont contenus les 41 facteurs, de leur part, sont regroupés dans les 3 catégories.

Figure 29 Modèle de l'efficacité d'usage, traduit de (Surbakti et al., 2019)



3.1 Thème 1 : Intérêt organisationnel perçu

Ce thème a été bien étudié dans le domaine du système d'information. Dans le contexte du big data, il se réfère à la valeur commerciale que l'entreprise attend bénéficier de l'exploitation du big data (Ji-fan Ren, Fosso Wamba, Akter, Dubey, & Childe, 2017). Suite à une étude réalisée par Russom (2013), cet intérêt peut aussi être interprété par l'amélioration des « insights » analytiques, des supports des fonctions de vente et de marketing.

Ce thème consiste de 7 facteurs suivants : la valeur perçue, les risques perçus, l'efficacité perçue, la facilité de l'utilisation perçue, l'intention d'utiliser, l'observabilité perçue et l'analyse coût-avantage. Observons-les plus en détail.

3.1.1 Valeur perçue

L'intérêt des entreprises de tirer une valeur commerciale des ressources technologiques est en forte croissance (Kiron, Prentice, & Ferguson, 2014). Suite Gao et ses coauteurs (2015), la valeur commerciale d'un projet du big data doit être clairement identifiable pour l'entreprise, autrement dit, ce projet doit répondre aux besoins de l'entreprise.

Ainsi, au niveau de l'usage du big data, la valeur perçue est considérée avant tout, comme une valeur commerciale.

Ici, il nous semble important de faire une annonce d'un concept de « paradoxe de la productivité », qui représente l'échec apparent des investissements importants dans les technologies informatiques (TI) (Carlaw & Oxley, 2008). De ce point de vue, on comprend que les investissements dans les technologies du big data n'atteindront pas à la réalisation d'une valeur commerciale, *s'ils sont réalisés juste pour que l'entreprise paraisse innovante (Jing Gao et al., 2015).*

Ainsi, pour pouvoir réaliser une valeur commerciale, l'entreprise doit être formée sur les équipes transversales et pluridisciplinaires. Ces équipes doivent assurer une coopération forte entre les départements et les ressources humaines impliqués dans le projet du big data (Akhtar et al., 2019), (Akhtar, Khan, Frynas, Tse, & Rao-Nicholson, 2018).

Une forte connaissance non seulement des opportunités que le big data peut apporter, mais aussi et surtout des risques, doit être bien prise en considération afin de pouvoir réaliser une valeur de l'usage du big data (Elisabetta Raguseo, 2018).

Melville et ses coauteurs (2004) ont étudié les composants du construit de la valeur commerciale et ont proposé leur modèle. Ces auteurs décrivent le processus de la réalisation de la valeur commerciale de TI à 3 niveaux :

- D'abord ils définissent que le lieu de la réalisation de cette valeur est l'entreprise qui investit et déploie des ressources technologiques, appelé « la firme focale » ;
- La réalisation de la valeur commerciale de TI est impactée des facteurs externes tels que l'environnement concurrentiel (caractéristiques industriels, partenaires commerciaux) ;
- En fin, c'est l'impact de l'environnement macro.

Ainsi, pour étudier ce construit, un ensemble des éléments suivants doit être pris en considération (Ji-fan Ren et al., 2017), (Gregor, Martin, Fernandez, Stern, & Vitale, 2006), (Elisabetta Raguseo, 2018) :

- *Création d'un avantage concurrentiel ;*
- *Alignement de l'analyse avec la stratégie commerciale ;*
- *Etablissement des liens avec d'autres organisations comme des fournisseurs ;*
- *Habilitation d'une organisation à répondre rapidement au changement ;*
- *Amélioration du service client ;*
- *Fourniture de meilleurs produits et services.*

3.1.2 Risques perçus

On comprend bien que l'usage du big data peut potentiellement être lié aux risques. Dans ce contexte, les risques sont considérés comme une perception de la possibilité de production des résultats inattendus avec les conséquences indésirables provoqués par l'usage du big data (Surbakti et al., 2019).

Cunningham (1966) sépare le risque perçu en 2 déterminants : l'incertitude et la probabilité subjective. Bauer (1960) pour sa part, définit le risque perçu comme un mélange d'incertitude et de gravité des résultats impliqués (Soon, Lee, & Boursier, 2016).

Bien évidemment, la littérature sur le big data se focalise surtout sur les opportunités et les avantages, tandis que le sujet des risques provoqués du big data reste encore très peu étudié.

Clarke (2016) essaie de mettre en évidence les « big » risques liés au big data quand les données sont réutilisées, consolidées de sources différentes, analysées et exploitées. Il étudie ce sujet sous l'angle de la responsabilité morale et légale. Dans son étude, il démontre que ces points sont touchés par les disciplines de management, mais sans un résultat significatif : les managers de haut niveau « restent tout seul » avec leurs décisions à prendre des responsabilités sur *l'assurance de la qualité, l'évaluation et le management des risques*.

Ainsi, pour permettre une réalisation de la valeur des investissements en TI, tous les deux éléments, les risques et les avantages du big data doivent être pris en considération (Love, P. E. D., Irani, Z., Standing, C., Lin, C., & Burn, 2005).

Quelques exemples des situations qui peuvent potentiellement engendrer les risques, sont cités ci-dessus (Mcfarlan, 1981), (Willcocks & Margetts, 1994) :

- Les coûts de mise en œuvre de la technologie et son exploitation sont plus élevés que prévu ;
- La performance technique du système est considérablement au-dessus des résultats estimés ;
- L'incompatibilité du système avec les paramètres du matériel informatique.

Les éléments considérés dans la littérature comme les facteurs de risque sont (E. Raguseo, 2018), (Love, P. E. D., Irani, Z., Standing, C., Lin, C., & Burn, 2005):

- La réticence des employés à s'adapter aux changements ;
- Le manque de soutien à l'infrastructure du système d'information ;
- L'incertitude technique et manque de connaissances ;
- L'expertise informatique minimale ;
- L'incertitude sur la façon de mesurer les avantages potentiels ;

- L'incertitude sur la façon de mesurer les coûts qui en découlent ;
- Les dépenses en capital sans garantie de rendements probables ;
- Les problèmes de sécurité ;
- La réaffectation du personnel formé aux solutions big data ;
- Les problèmes de confidentialité.

3.1.3 Efficacité perçue

Ce facteur est considéré comme la perception de l'utilisateur de la manière dont l'exploitation du système du big data peut contribuer à l'accomplissement de son travail et des tâches quotidiennes (Surbakti et al., 2019).

Les résultats d'une étude sur l'usage et l'adoption du big data par les entreprises, mettent en évidence le fait que l'intention d'usage du big data est fortement influencée par l'efficacité perçue de l'utilisateur (Verma, 2017).

L'efficacité perçue est un élément extrinsèque clé de la motivation à utiliser un système d'information, valide pour les deux cas d'utilisation : routine et innovante (Bischoff, Aier, Haki, & Winter, 2015). L'efficacité perçue est aussi l'un de deux éléments-clés de l'intention à utiliser une technologie, prédit de Davis et ses coauteurs (F. Davis, 1989)(F. D. Davis et al., 1989) ; le 2^e élément étant la facilité perçue de l'utilisation.

Ainsi, Agarwal et Karahanna (2000) dans leur étude affirment que l'efficacité perçue par l'utilisateur a un effet positif sur l'utilisation de la technologie informatique. Barnes (2011) et Kim (2012) de leur part mettent en évidence le fait que l'efficacité perçue est un des facteurs influençant l'intention de la continuité d'utilisation.

Les éléments qui permettent d'étudier ce construit sont (Demoulin & Coussement, 2020), (Gefen, Karahanna, & Straub, 2003), (Gefen & Straub, 2000), (Moore & Benbasat, 1991), (B. Kim, 2012):

- *Amélioration de la performance dans le travail ;*
- *Augmentation de la productivité dans le travail ;*
- *Amélioration de l'efficacité dans le travail ;*

- *Utilité dans le travail.*

3.1.4 Facilité perçue de l'utilisation

Au contexte de l'usage du big data, ce facteur est considéré comme la perception de l'utilisateur de la manière de faciliter le travail et de l'accès aux fonctionnalités du système (Surbakti et al., 2019).

Ce construit est bien étudié dans la littérature du système d'information. C'est le 2^{ème} élément important (avec l'efficacité perçue) dans les recherches sur l'usage du SI (F. D. Davis et al., 1989), (F. Davis, 1989). Comme on a observé, noté dans le chapitre précédent, l'intention de l'acceptation volontaire, autrement dit, l'usage d'une nouvelle technologie informatique est déterminé par 2 croyances qui portent sur l'efficacité perçue et la facilité perçue de l'utilisation de la TI concernée. Ici, si l'efficacité perçue est le résultat d'une estimation subjective de l'individu de l'utilité offerte par cette nouvelle technologie informatique au contexte lié à *une tâche spécifique*¹¹ (Agarwal & Karahanna, 2000), (F. D. Davis, Bagozzi, & Warshaw, 1992), la facilité perçue de l'utilisation est l'indicateur de l'effort cognitif nécessaire à apprendre et utiliser une nouvelle TI (Gefen et al., 2003), (Gefen & Straub, 2000).

Ce construit a été étudié par les principaux éléments suivants (Demoulin & Coussement, 2020), (Gefen et al., 2003), (Gefen & Straub, 2000), (Moore & Benbasat, 1991), (B. Kim, 2012):

- *Clarté et l'intelligibilité du système à l'utilisateur;*
- *Facilité de l'utilisation du système;*
- *Disponibilité du système pour accomplir une tâche.*

3.1.5 Intention à utiliser

Ce facteur se réfère à l'intention d'adoption pour la première fois des technologies du big data dans l'entreprise ; ce sujet a été étudié par Verma (2017). Dans cette étude, on traite la question de l'adoption des services du big data par les entreprises de l'industrie manufacturière en Inde. Les résultats montrent que l'intention à adopter (utiliser) les services du big data est positivement influencée par leur qualité et leurs avantages perçus. Par contre, les coûts et les dépenses associés à l'implantation et leur maintien n'affectent pas l'intention d'adoption.

¹¹ Initialement, ce construit a été défini par rapport à la performance du travail d'un individu (F. Davis, 1989).

La qualité des services du big data et les avantages perçus ont été mis en évidence par une autre étude réalisée par Shin sur l'intention comportementale à utiliser les services du big data en Corée (Shin, 2016). Cet auteur met également en évidence l'influence des facteurs comme les caractéristiques du système du big data (qualité-sécurité-interopérabilité), l'efficacité perçue et la facilité perçue d'utilisation.

Esteves et Curto (2013) ont étudié l'intention à adopter le big data et son adoption actuelle en Espagne. Ces auteurs démontrent l'importance des conditions facilitatrices sur l'intention à adopter les technologies du big data :

- Les facteurs internes et externes peuvent gêner ou améliorer la performance du comportement (Triandis, 1977);
- Les coûts perçus, la compatibilité (qui se réfère au degré qui reflète l'adéquation de l'innovation avec les valeurs, l'expérience et les besoins actuels de l'« adopteur » (Rogers, 1983)) ;
- L'efficacité personnelle (Bandura, 1982).

Les éléments par lesquels ce facteur a été étudié sont (Demoulin & Coussement, 2020), (Kwahk & Lee, 2008), (Venkatesh et al., 2003):

- *Intention d'utiliser le système à l'avenir ;*
- *Prévision de l'utilisation ;*
- *Planning de l'utilisation.*

3.1.6 Observabilité perçue

Ce facteur représente le degré d'observabilité des résultats obtenus après l'usage du nouveau système. Initialement ce concept a été défini comme le degré où les résultats d'une innovation sont visibles et communicables aux autres (Rogers, 1983).

L'observabilité perçue est traitée comme un des 4 déterminants (les 3 autres sont : l'avantage relatif, la complexité et l'adéquation). L'ensemble de ces déterminants peuvent assurer le succès d'un système de BI (Business Intelligence) (Ahmad, Ahmad, & Hashim, 2016). Rogers (1995) y ajoute encore un élément complémentaire, celui de la sécurité et de la

confidentialité. Ce facteur est traité comme important surtout au stade de démarrage de l'adoption d'une nouvelle technologie ; où la minimisation des incertitudes sur la performance affecte le degré de la diffusion de l'innovation.

Au contexte de l'adoption du big data, l'observabilité est définie comme le processus par lequel les entreprises observent le facteur du succès des autres entreprises qui ont déjà adopté le système du big data : *les organisations observent les autres entreprises afin d'analyser les avantages de l'adoption* (Baig, Shuib, & Yadegaridehkordi, 2019).

Dans un contexte malaisien, on a traité les facteurs influençant l'adoption du CE (Commerce Electronique) en s'appuyant sur les caractéristiques de l'innovation proposées par Rogers (1995). Les résultats ont mis en évidence l'effet considérable de ce facteur dans l'adoption du CE (Alam, Alam, Khatibi, Ahmad, & Ismail, 2008).

Les éléments qui permettent l'étude de ce facteur sont (Maroufkhani, Wan Ismail, & Ghobakhloo, 2020), (Moore & Benbasat, 1991), (Huang, Liu, & Chang, 2012):

- *Tendance des acteurs sur le marché à utiliser la nouvelle solution ;*
- *Perspectives du développement sur l'arène national et international ;*
- *Accès aisé aux technologies ;*
- *Utilisation aisée ;*
- *Visibilité des résultats améliorés après l'adoption de nouvelle technologie.*

3.1.7 Analyse coût-avantage

Sous cette notion on comprend d'abord la technique qui permet d'identifier parmi les alternatives la meilleure solution de technologie du big data à adopter (Gopalkrishnan, Steier, Lewis, & Guszczka, 2012). Cette priorité fournit des outils qui permettent de mesurer les avantages du choix (Siddiqi et al., 2016).

Pour adopter une technologie du big data, l'entreprise choisit des solutions alternatives disponibles. Ce composant motivationnel conditionne vers un choix qui semble le plus avantageux en matière de ses avantages et coûts engagés (Jing Gao et al., 2015).

D'autres études mettent en évidence le fait que les entreprises sont de plus en plus méfiantes concernant les investissements en TI. D'un côté les coûts engagés sont prohibitifs (Ghobakhloo, Hong, Sabouri, & Zulkifli, 2012), de l'autre, le taux d'échec relativement fort (Markus & Topi, 2015). Madlberger (2009) met en évidence une association négative entre les coûts perçus et l'adoption de la technologie.

Les éléments qui permettent l'étude de ce facteur sont (Maduku, Mpinganjira, & Duh, 2016), (Wixom & Watson, 2001), (Polites & Karahanna, 2012):

- *Perception de rapport coûts et avantages ;*
- *Perception des coûts de maintien du système ;*
- *Perception des coûts de fournitures des mécanismes de support du système ;*
- *Perception des coûts engagés dans les formations des individus à utiliser le système.*

3.2 Thème 2: Management des processus

Le management des processus se réfère à la capacité d'une entreprise à prendre des mesures de restructuration du business en fonction des processus TI (Shuradze & Wagner, 2016). Ce thème regroupe de 3 facteurs : l'orientation processus, l'intégration du business processus en TI et le management des données.

3.2.1 Orientation processus

L'orientation processus ou BPO (Business Process Orientation) est le degré de l'attention que l'organisation a sur ses processus essentiels par rapport à ses concurrents (Willaert, Van Den Bergh, Willems, & Deschoolmeester, 2007), (Graupner, Berner, Maedche, & Jegadeesan, 2014). Cette définition signifie que les individus dans l'organisation développent une attitude orientée vers les processus.

L'on suppose que le plus haut degré de maturité de l'entreprise au niveau du BPO (qui peut être défini au niveau de 6 facteurs : l'alignement stratégique, la culture, l'aspect humain, la gouvernance et le SI/TI (Willaert et al., 2007)), peut plus positivement impacter l'efficacité d'usage du big data en entreprise (P. S. Fosso Wamba, 2017), (G. Kim, Shin, Kim, & Lee, 2011).

Les éléments qui permettent l'étude de ce facteur sont (P. S. Fosso Wamba, 2017), (Hong & Kim, 2002):

- *Comparaison avec les concurrents concernant la connexion des parties (communication, partage d'information) au sein d'un business processus ;*
- *Comparaison avec les concurrents au niveau de minimisation des coûts au sein d'un business processus ;*
- *Comparaison avec les concurrents au niveau d'application des méthodes analytiques complexes à un business processus ;*
- *Comparaison avec les concurrents au niveau de fourniture d'une information détaillée à un business processus.*

3.2.2 Intégration du business processus en TI

Ce facteur représente la capacité d'adaptation du business et les processus TI existants d'une manière efficace et efficiente en renforçant les capacités d'une TI émergente (Shuradze & Wagner, 2016).

Autrement dit, afin de profiter de l'usage du big data, les entreprises ont besoin de mettre en œuvre des changements et des améliorations non seulement au niveau de l'infrastructure technologique, mais aussi au niveau des aspects managériaux et surtout de leurs business processus (Buhl, Röglinger, Moser, & Heidemann, 2013), (Al Nuaimi, Al Neyadi, Mohamed, & Al-Jaroodi, 2015).

Dans une étude sur les facteurs influençant la qualité de la prise de décision, *l'intégration du processus et la standardisation* sont considérées comme des facteurs principaux. C'est aussi traité comme une condition importante dans la routinisation de l'usage du big data (Janssen et al., 2017).

Le degré de la capacité de l'intégration du business processus est considéré aussi comme un des facteurs influençant la capacité de la prise de décision sur big data (S. Shamim, Zeng, Shariq, & Khan, 2019), (Janssen et al., 2017).

Les éléments qui permettent l'étude de ce facteur sont (S. Shamim et al., 2019), (Bharadwaj, Sambamurthy, & Zmud, 1999):

- *Capacité de l'entreprise à intégrer les processus de chaîne du big data ;*
- *Rapport entre la minimisation de coût de l'usage du big data de l'intégration des processus de chaîne du big data ;*
- *Rapport entre l'intégration des processus de chaîne du big data et la minimisation des efforts à analyser le big data.*

3.2.3 Management des données

Cet élément à valeur stratégique est une fonction complexe de l'usage du big data. Selon Adrian et ses coauteurs (2017), les tâches liées au management des données comprennent les processus d'un ensemble de plusieurs éléments :

- Gouvernance des données (Y. Wang, Kung, & Byrd, 2018) ;
- Source, intégration et exécution des données (Popovič, Hackney, Tassabehji, & Castelli, 2018) ;
- Accès aux données (Gupta & George, 2016) ;
- Confidentialité des données (Halaweh & Ahmed El Massry, 2015).

Ce facteur a été bien étudié non seulement dans un contexte d'un SI avec les données traditionnelles (Ramamurthy, Sen, & Sinha, 2008), (Shaul & Tauber, 2013), mais aussi au contexte du big data (Verma, 2017), (Adrian, Abdullah, Atan, & Jusoh, 2016), (Wielki, 2013), (Kung, Kung, Jones-Farmer, & Wang, 2015).

Pour comprendre cet élément complexe, on prend en compte d'abord la capacité de l'entreprise d'ajustement de ses ressources à l'exécution des processus cités ci-dessous.

Ainsi, l'ensemble des processus peut être étudié par les éléments suivants (Ghasemaghaei, Hassanein, & Turel, 2017), (Gupta & George, 2016):

- *Degré de l'ajustement du potentiel en outils analytiques de l'entreprise pour traiter les données ;*

- *Cohérence entre l'accès de l'entreprise aux outils analytiques et le besoin de l'analyse des données.*

3.3 Thème 3: Confidentialité, sécurité et gouvernance des données

Ce deuxième thème de catégorie opérationnelle est composé de deux éléments suivants : la confidentialité ; la sécurité et la gouvernance des données. Ici, on peut noter que le facteur de la gouvernance des données était aussi inclus dans la thématique du management des processus. Il faut y faire une distinction : dans la thématique de « confidentialité, sécurité et gouvernance » le facteur est traité sous l'angle de la régulation juridique et éthique, de droits et des responsabilités liés au traitement et la prise de décision sur le big data. Tandis que, dans la thématique de « management des processus » ce facteur est traité sous l'angle opérationnel-technique-processuel, comme le management de l'architecture des données, la qualité, les opérations de base des données, etc.

3.3.1 Confidentialité et sécurité

Les données qui sont distribuées, partagées et stockées posent de problèmes de leur gestion sécurisée. Ce facteur est considéré fragile et en tant qu'élément-clé de succès du projet de big data. Dans le cas, où les données comprennent les informations personnellement identifiables, par exemple, les dossiers médicaux et les données bio (M. K. Kim & Park, 2017) la question de la confidentialité et de la sécurité est prioritaire. Pourtant, dans le domaine du trafic (données du trafic et de la météo, par exemple) où les données sont minimalement restreintes ou sans restriction, la question n'a pas de priorité (Dorr et al., 2015).

Comme au niveau du traitement de big data, au niveau de l'assurance de leur sécurité aussi, les mécanismes et les outils de confidentialité appliqués aux données traditionnelles sont inadéquats au contexte du big data.

Ainsi, dans le modèle (Figure 29), le facteur de confidentialité et de sécurité se réfère aux moyens par lesquels l'entreprise fournit les approches cohérentes avec les caractéristiques des données massives à assurer le management des données au niveau de leur distribution, partage, stockage et utilisation (Tai, Yu, Yang, & Chen, 2011), (Li, Aggarwal, & Wang, 2011).

Ce facteur a été étudié par les auteurs à travers des éléments suivants (Ghasemaghaei, 2020), (Zheng, Zhao, & Stylianou, 2013):

- *Risque de l'accès non autorisé aux données ;*
- *Niveau de la protection de l'information personnelle ;*
- *Niveau de l'utilisation sécurisée des données.*

3.3.2 Gouvernance des données

Comme souligné précédemment, cette thématique se réfère plutôt aux aspects juridiques et éthiques de la régulation des données.

Ylijoki et Porras (2016) suggèrent que cette assurance doit comprendre tout le cycle de vie des données (N. Khan et al., 2014).

Des entreprises qui exploitent le big data, possèdent normalement une fonction attachée de « gouvernance des données » qui représente un ensemble des processus et des procédures à assurer une régulation cohérente avec les lois et l'éthique.

Bien évidemment, les entreprises qui mettent un focus important à cette fonction, ont de plus de chance à réussir un projet de big data dans ses différents étapes de maturité (Brous, Janssen, Schraven, Spiegelner, & Duzgun, 2017), (Sivarajah et al., 2017), (Thompson, Ravindran, & Nicosia, 2015).

Les éléments et les conditions qui permettent de comprendre le niveau de maturité de cette fonction dans l'entreprise sont (Mikalef, Boura, Lekakos, & Krogstie, 2019a), (Byrd & Turner, 2000) :

- *Existence d'un ensemble des pratiques maîtrisées et régulières qui peuvent servir comme un contrôle de la gouvernance des données /en matière de définir des politiques de conservation (par exemple, la durée de vie) des données ;*
- *Routines de sauvegarde ;*

- *Etablissement / surveillance de l'accès (par exemple, l'accès utilisateur) aux données ;*
- *Classement des données en fonction de leur valeur ;*
- *Surveillance des coûts par rapport à la valeur des données.*

3.4 Thème 4: Qualité des données

Cette thématique est la dernière dans la catégorie opérationnelle. Ce thème est composé de 7 facteurs : la qualité de données et la qualité de l'information ; la complétude des données ; l'actualité des données ; l'accès aux données ; la pertinence des données ; l'exactitude des données et la cohérence des données.

Globalement, le thème de « qualité des données » se réfère à la conformité des données à leur usage prévu. Dans la catégorie opérationnelle, ce facteur est considéré comme un conducteur du succès de projet du big data (Saltz & Shamshurin, 2016). Ci-après, observons chaque facteur composant plus en détail.

3.4.1 Qualité des données et qualité de l'information

D'abord, rappelons les relations qui s'établissent entre « Données – Informations – Connaissances ». Les données sont des faits représentant une certaine réalité, elles sont indépendantes de celui qui les utilise (Lépinard et al., 2013). Les informations sont des données auxquelles un individu ou un groupe a ajouté du sens; et le passage de la notion d'information à celle de connaissance s'opère au travers d'un mécanisme de cognition sur les informations (Lebraty, 2011). Après cette définition importante, il faut qu'on comprenne que la qualité des données se détermine au moment de leur collecte, pourtant la qualité de l'information ne peut pas être « évaluée » avant son usage (Surbakti et al., 2019). Aujourd'hui presque, 30% des dirigeants ne font pas confiance à la qualité de l'information qu'ils utilisent (S. Fosso Wamba, Akter, Edwards, Chopin, & Gnazou, 2015). Ce facteur est noté comme l'un des 3 principaux facteurs influençant l'usage du big data, les 2 autres sont la confidentialité et la sécurité (Lynn Segarra et al., 2016).

La qualité des données et la qualité de l'information ont été étudiées à travers des éléments suivants (S. Shamim et al., 2019), (Nelson, Todd, & Wixom, 2005) :

- *Qualité des sources des données ;*
- *Réputation des fournisseurs du système de big data ;*
- *Qualité d'output fourni du système.*

3.4.2 Complétude des données

Selon Clarke (2016), la complétude des données se réfère à l'information contextuelle suffisante où le contenu des données n'est pas susceptible d'être mal interprété. Autrement dit, la complétude indique aussi la mesure dans laquelle l'utilisateur perçoit que le système du big data fournit toute information nécessaire (Wixom & Todd, 2005).

On comprend ainsi que la complétude des données, c'est le degré auquel l'on perçoit que les données sont entièrement complètes pour leur contexte d'usage, sans aucun élément manquant.

Selon Wand and Wang (1996), la complétude décrit l'état d'un volume minimal des données capturées qui permettent de fournir l'information nécessaire (Hazen, Boone, Ezell, & Jones-Farmer, 2014).

Pour étudier ce facteur dans le contexte du big data, on examine *le degré de complétude d'information fournie du système et le niveau de satisfaction d'utilisateurs en matière de son besoin d'information fournie du système* (Ji-fan Ren et al., 2017), (Nelson et al., 2005), (Lee, Strong, Kahn, & Wang, 2002).

3.4.3 Actualité des données

L'actualité se réfère au degré de la perception d'un utilisateur que les données sont à jour (Wixom & Todd, 2005), (Ji-fan Ren et al., 2017) ou le degré d'actualité des données avec le monde qu'elles modélisent (Jerry Gao, Xie, & Tao, 2016).

En cohérence avec son contexte spécifique, il est demandé normalement d'adopter une procédure de régularisation de la mise à jour des données. Ça peut se faire par la définition des règles pour noter les limites de « la durée de vie » des données, quand il est nécessaire qu'elles soient actualisées ou depuis quand elles peuvent se considérer comme « actuelles ».

On peut imaginer que ce facteur reçoit une attention spécifique dans le contexte du big data où les flux et la rapidité du partage de grosses masses des données rendent plus difficile son opérationnalisation (T. H. Davenport, Barth, & Bean, 2012).

Pour étudier ce facteur, on peut se focaliser sur les éléments suivants (Ji-fan Ren et al., 2017), (Nelson et al., 2005) :

- *Degré de l'actualité de l'information fournie du système ;*
- *Degré de l'actualité des données utilisées dans le système ;*
- *Niveau de l'échec du système de fournir une information actualisée.*

3.4.4 Accès aux données

L'accès et l'intégration de diverses sources de données est considéré comme un facteur clé de succès pour la réalisation des opportunités de transformation des projets en big data (S. Fosso Wamba et al., 2015). Ainsi, on traite le facteur d'un côté comme une capacité de rendre les données disponibles et d'autre côté comme un état où les données peuvent être facilement récupérées de diverses ressources des données.

Une autre mesure qui est utilisée dans la littérature avec l'accès aux données, appelée « rapidité de l'accessibilité des données », désigne l'attente du temps pour accéder aux données (Jerry Gao et al., 2016).

Et donc on voit comment ce facteur affecte l'efficacité d'usage du big data : l'infrastructure qui permet un accès efficace aux données nécessaires dans les délais acceptables contribue à la réalisation de l'objectif fixé pour cet usage.

Ce facteur a été étudié au travers des éléments suivants (Demoulin & Coussement, 2020), (Lee et al., 2002):

- *Degré de facilité pour récupérer les données ;*
- *Degré de facilité pour accéder aux données ;*
- *Degré de facilité de disponibilité des données ;*
- *Niveau de rapidité de l'accès aux données.*

3.4.5 Pertinence des données

La pertinence se réfère à la capacité des données à correspondre aux demandes des initiatives du big data. Il est important de noter la lacune entre ce que propose le système et le besoin réel de l'utilisateur en matière de l'information (Grublješič & Jaklič, 2015), (Hou, 2012), qui influence potentiellement aussi la qualité de résultat final / obtenu (output) (Gopalkrishnan et al., 2012).

Wang et Strong (1996) qui désignent ce facteur comme la mesure dans laquelle les données sont applicables et utiles pour une tâche à accomplir, proposent les 4 dimensions à décrire la pertinence des données. Ces dimensions sont la capacité à ajouter une valeur (une valeur opérationnelle), la pertinence contextuelle (conduite applicable), la cohérence avec le besoin de l'utilisateur et le volume nécessaire des données.

Ce facteur a été étudié à travers le degré de *l'utilité (informative) de l'information avec les besoins de la tâche concrète* (Zheng et al., 2013), (Lee et al., 2002).

3.4.6 Exactitude des données

Ce facteur se réfère au degré de la concordance entre le contenu des données et le phénomène réel destinés à être (Clarke, 2016). Autrement dit, l'exactitude se concentre sur la « justesse » perçue de l'information (Ji-fan Ren et al., 2017). Dans un contexte plus général, ça reflète la mesure dans laquelle les données sont correctes (Wixom & Todd, 2005), fiables et exemptes de toute erreur (Nelson et al., 2005), (Wang, R. Y., and Strong, 1996).

En contexte du big data, ce facteur joue un rôle majeur sur l'efficacité d'usage.

Les éléments pour étudier ce facteur sont *le degré de la justesse des données, le degré d'erreur dans les informations fournies du système et le niveau de l'exactitude des informations fournies* (Ji-fan Ren et al., 2017), (Nelson et al., 2005).

3.4.7 Cohérence des données

Le dernier facteur de la thématique se réfère au degré où toutes les données nécessaires pour l'activité courante et l'activité future de l'entreprise (par exemple, la prise de décision) sont

disponibles dans l'entreprise (Kwon, Lee, & Shin, 2014). Il est essentiel que cette cohérence favorise l'intégration des processus, notamment l'intégration de l'information sur les flux financiers, physiques, etc. (Rai, Patnayakuni, & Seth, 2006).

De plus, le manque de la cohérence des données peut causer de gros problèmes de traitement des données qui viennent des sources internes et externes des données (Pitoura & Bhargava, 1999). Ainsi, il est indispensable que l'infrastructure du big data de l'entreprise assure les mécanismes de conformité et des standards cohérents.

Dans le contexte du big data, on s'intéresse surtout au *degré d'amélioration de la cohérence des données dans les entrepôts de l'entreprise, fournie de fonctionnalité des systèmes* (Kwon et al., 2014), (Wixom & Watson, 2001).

3.5 Thème 5 : Aspect Humain

Cette thématique concerne l'ensemble des facteurs qui sont liés aux plusieurs éléments suivants :

- Savoirs et compétences des individus, demandés pour la gestion du big data ;
- Confiance des utilisateurs envers le système ;
- Champions qui soutiennent les initiatives du big data ;
- Engagement ;
- Contribution ;
- Caractéristiques individuelles des utilisateurs qui peuvent freiner ou contribuer à l'usage efficace du big data.

Cet ensemble de 6 éléments, nous allons l'observer ci-après plus en détail.

3.5.1 Savoirs et compétences des individus

Ce facteur se réfère à la disponibilité des professionnels en entreprise, qui possèdent les savoirs, compétences et spécialités nécessaires à gérer le big data (Cetindamar, Shdifat, & Erfani, 2022), (Halaweh & Ahmed El Massry, 2015). La figure 30 illustre bien la proportion de ces compétences pour la gestion du big data (Fogelman-Soulié & Lu, 2016).

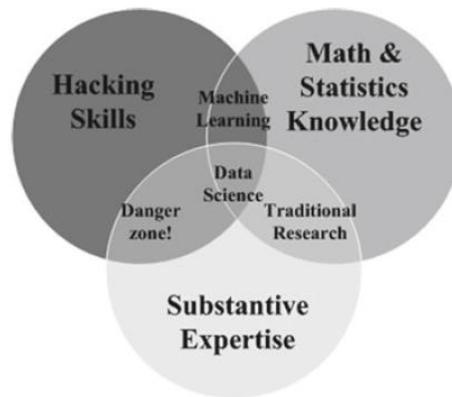


Figure 30 Compétences nécessaires pour la gestion du big data (Fogelman-Soulié & Lu, 2016)

Pourtant, le manque des établissements qui sont en capacité de former de vrais spécialistes avec une forte expertise en gestion du big data, pose de problèmes pour les entreprises (Davenport et al., 2012).

Une initiative du big data, qui comprend une innovation, provoque un changement technologique et culturel en entreprise ; de nouvelles compétences sont demandées pour réussir l'adoption et l'exécution efficace du big data (Halaweh & El Massry, 2017).

Si la base de l'apprentissage machine (ML) se pose sur les compétences et les connaissances en statistiques, la gestion du big data demande des compétences en science des données. Il faut bien noter que les spécialistes en sciences des données sont des experts techniques, pourtant le management d'un projet du big data demande aussi des compétences dans les aspects managériaux du big data. Ce travail de thèse tente de contribuer au développement de ces compétences.

Les éléments qui permettent d'étudier ce facteur sont (Mikalef et al., 2019a), (Wixom & Watson, 2001):

- Niveau des compétences du personnel pour accomplir les tâches avec succès ;
- Niveau de formation en projet du big data ;
- Facilité d'accès à la formation du personnel pour le nouveau projet du big data.

3.5.2 Confiance

Pour un contexte général du SI, aussi pour le contexte du big data, ce facteur joue un rôle important dans l'usage du système.

Ce facteur se traduit en confiance des utilisateurs d'abord pour le système qu'ils utilisent, et plus particulièrement, la confiance en information (output) du système (Bischoff et al., 2015).

Aberg et Shahmehri (2001) aussi mettent en évidence l'importance de la confiance des utilisateurs vers le système qu'ils utilisent et soulignent la nécessité des études des facteurs qui peuvent affecter la confiance des utilisateurs.

Dans un contexte général du SI, la confidentialité et la sécurité du système sont considérées comme les deux éléments importants qui peuvent affecter la confiance des utilisateurs (Molla & Licker, 2001).

La confiance des utilisateurs est étudié d'abord par *leur perception que le système possède des avantages ; la perception de l'utilisateur sur les perceptions des collègues ; la perception de la direction* (Lankton, Harrison Mcknight, & Tripp, 2015), (Amoako-Gyampah & Salam, 2004).

3.5.3 Champions

On suppose que les entreprises qui possèdent des "champions" des projets, ont plus de chances de promouvoir les initiatives. Il s'agit d'un aspect psychologique qui influence l'usage du système quand un champion au sein de l'entreprise le soutient et l'encourage.

La présence d'un champion au sein de l'entreprise affecte positivement au stade d'une implantation d'un nouveau système (Wixom & Watson, 2001). D'une part, les champions peuvent accélérer un comportement du leadership transformationnel, et, d'autre part, ils contribuent à surmonter la résistance et l'opposition en entreprise (Howell & Higgins, 1990).

Les champions du *haut niveau viennent souvent du domaine du SI, mais ils peuvent présenter un ou plusieurs domaines fonctionnels de l'entreprise* (Premkumar & Ramamurthy, 1995), (Wixom & Watson, 2001).

3.5.4 Engagement du personnel

L'engagement du personnel dans le projet du big data dès le début de l'implantation est considéré comme un élément-clé de la réussite. Leur contribution au projet est distinguée comme un des facteurs essentiels qui peuvent impacter le résultat tout au long du projet (Fosso Wamba et al., 2015).

Wixom et Watson (2001) considèrent l'engagement du personnel dans le projet comme un facteur – composant du succès d'une implantation organisationnelle. En plus, cet engagement est fortement associé au soutien de la part de la haute direction. Dans ce contexte, Rondeau et ses coauteurs (2006) mettent en évidence, que la haute direction peut assurer le soutien en créant un environnement coopératif et un travail en équipe au travers des objectifs organisationnels.

Ainsi, le degré de soutien du projet par le personnel de divers départements fonctionnels peut montrer le niveau d'efficacité d'usage du big data en entreprise.

3.5.5 Participation des utilisateurs

Ce facteur montre le degré d'implication de l'utilisateur dans une implantation du projet (Grublješić & Jaklič, 2015). Cette implication, de sa part est influencée par d'autres facteurs déjà traités dans cette section (F. D. Davis et al., 1989).

Yoon et ses coauteurs (1995) proposent que cette participation et l'implication, comme l'engagement (Rondeau et al., 2006) doivent être assurées de la part de chefs de projet. La même idée a été partagée aussi de White et Leifer (1986) pour un contexte général du SI.

Ainsi, pour étudier ce facteur, on peut s'appuyer sur les éléments comme le degré d'intégration dans un travail des utilisateurs concernés et le personnel d'autres départements fonctionnels ; le temps que les utilisateurs sont impliqués dans le projet et la nature des tâches qu'ils effectuent au sein du projet (Wixom & Watson, 2001).

3.5.6 Caractéristiques individuelles

Comme dans le cas du système d'information pour un contexte général, dans le cas d'un système du big data, l'usage peut être affecté par les caractéristiques individuelles de ses utilisateurs.

Ces caractéristiques peuvent être l'âge, les connaissances en informatique, l'expérience concernée, le comportement et l'attitude, l'anxiété informatique et autres (Grublješič & Jaklič, 2015).

L'anxiété est considérée comme l'état de l'anxiété et les réactions émotionnelles quand on utilise une nouvelle technologie.

Nous allons traiter ce facteur pour le contexte du big data en nous appuyant sur les éléments d'étude de l'anxiété fournis de Venkatesh et ses coauteurs (Venkatesh et al., 2003):

- *Degré de la peur à utiliser un nouveau système ;*
- *Degré de la peur de perte d'informations en cas d'une mauvaise action ;*
- *Degré d'hésitation d'usage du système pour ne pas faire des erreurs,*
- *Degré de menace du système pour l'utilisateur.*

3.6 Thème 6 : Aspect organisationnel

Cette thématique concerne tous les éléments qui peuvent affecter l'efficacité de l'usage du big data au niveau organisationnel.

On constate bien que *les facteurs organisationnels (qui sont souvent ignorés par les entreprises au stade de l'adoption des technologies du big data) peuvent prédire si l'entreprise est prête à adopter un tel projet ou non* (Mikalef, van de Wetering, & Krogstie, 2021), (Marshall, Mueck, & Shockley, 2015).

La thématique comprend les 13 facteurs suivants :

- Compétences organisationnelles et culturelles ;
- Gestion du talent ;

- Programme de gestion du changement ;
- Alignement stratégique ;
- Management du projet ;
- Management de la performance ;
- Structure organisationnelle et la taille ;
- Collaboration interdépartementale ;
- Communication ;
- Support de la haute direction ;
- Effet environnemental ;
- Objectifs précis ;
- Accent sur l'innovation.

Dans les paragraphes suivants, nous allons détailler chaque élément.

3.6.1 Compétence organisationnelle et culturelle

Ce facteur peut potentiellement influencer l'efficacité d'usage du big data en entreprise à travers aux normes et aux comportements organisés d'un groupe d'individus ou d'un département fonctionnel au sein de l'entreprise visée à atteindre certains objectifs ou les croyances liés au projet.

Dans ce contexte, une culture orientée vers les données signifie qu'un groupe d'individus peut se comporter de manière qu'ils croient, comprennent et utilisent le système du big data. Et ce dernier joue un rôle important dans leur travail et contribue à la réalisation des objectifs de l'entreprise (Kiron & Shockley, 2011). Ces auteurs détaillent les composants d'une culture orientée vers les données qui peuvent être traduites à un contexte du big data de manière suivante :

- Usage du système du big data comme un atout stratégique ;
- Soutien du projet par la haute direction au sein de toute entreprise ;
- Disponibilité des outputs du système pour tous ceux qui en ont besoin.

Ainsi, les conditions contextuelles sont nécessaires afin d'améliorer l'efficacité d'usage d'un nouveau système par une entreprise (Donate & Sánchez de Pablo, 2015).

Des éléments qui peuvent contribuer à l'étude de ce facteur sont : *la perception de l'organisation de la décision basée sur les données et la perception de l'usage des données dans les buts stratégiques et dans les tâches de routine* (Gupta & George, 2016), (Saqib Shamim, Cang, & Yu, 2019).

3.6.2 Gestion du talent

D'abord, il faut noter que pour un projet du big data, le talent concerne les scientifiques des données qui sont capables de travailler avec de grosses masses et volumes (Wielki, 2013).

Donc, la gestion du talent est considérée comme l'un des principaux défis liés à l'usage du big data. Le leadership, la gestion du talent, la technologie et la culture organisationnelle peuvent contribuer beaucoup plus à l'efficacité d'usage que les éléments techniques (McAfee & Brynjolfsson, 2012b).

Les recherches actuelles sur la gestion du talent sont dans leur stade de développement, avec de très peu d'implications dans la science (Krishnan & Scullion, 2017).

Les éléments qui montrent le niveau de la gestion du talent en entreprise sont : *le niveau d'expertise des professionnels dans le domaine du big data, le programme des ressources humaines (RH) pour l'embauche des spécialistes du domaine et le développement de leurs compétences, le programme RH pour l'encouragement et la gestion du talent au sein de l'entreprise* (S. Shamim et al., 2019), (Akter, Wamba, Gunasekaran, Dubey, & Childe, 2016).

3.6.3 Programme de gestion du changement

L'adoption des technologies du big data au sein de l'entreprise provoque des changements qui doivent être prévus et gérés efficacement comme un ensemble de la nouvelle infrastructure.

Et les raisons des échecs des projets du big data sont principalement liées plutôt au management qu'à la technologie elle-même (Comuzzi & Patel, 2016).

Ce facteur concerne le planning, l'organisation et la coordination des processus par lesquels le changement est implanté où la haute direction prend un rôle de « motivateur » et de « conducteur » des salariés. Et ce niveau de disposition du changement détermine la position sur le marché en matière de compétitivité (Grublješić & Jaklič, 2015), (Kwahk & Lee, 2008).

Ainsi, le premier indicateur de l'efficacité d'usage du big data en terme du changement provoqué est considéré comme le degré de l'efficacité de gestion de changement au sein de l'organisation (Wixom & Watson, 2001).

La considération de la haute direction en matière de changement peut se traduire par *sa position de fournir une vision claire sur le programme de gestion des changements provoqués de l'usage du big data* (S. Shamim et al., 2019).

3.6.4 Alignement stratégique

Les initiatives du big data doivent être parrainées par la haute direction et être alignées à tout niveau de la stratégie de l'entreprise. La maturité de l'alignement stratégique montre la place du big data dans la stratégie de l'entreprise, la considération de la prise de décision fondée sur les données comme un standard stratégique (Comuzzi & Patel, 2016).

Autrement dit, c'est la mesure dans laquelle la stratégie du SI est alignée avec la stratégie globale de l'entreprise (Setia & Patel, 2013). Ou bien, le degré où les priorités, les buts et les objectifs de la stratégie du SI sont alignés avec les priorités, les buts et les objectifs de la stratégie de l'entreprise (Oh & Pinsonneault, 2007).

Les indicateurs du niveau de l'alignement stratégique dans l'entreprise sont démontrés par les éléments suivants (Akter et al., 2016):

- *Degré d'alignement du plan du projet de big data avec les priorités, missions et objectifs de l'entreprise ;*
- *Niveau de quantification des objectifs et des buts du plan de projet de big data ;*
- *Niveau du soutien que reçoit le projet du big data ;*
- *Priorité des investissements dans le projet du big data au plan de la stratégie globale de l'entreprise.*

3.6.5 Management du projet

Le management du projet concerne d'abord l'application des processus, des méthodes, la mobilisation des connaissances, des compétences et de l'expérience afin de réaliser les objectifs fixés pour les initiatives du big data (Dutta & Bose, 2015).

Ici, il est intéressant de noter les résultats d'une étude réalisée sur les entreprises qui ont réussi le management du projet de big data. Les résultats montrent que l'alignement stratégique joue un rôle important dans une implantation du projet. L'autre aspect important, consiste dans une méthode de « voyage – journey » dans l'implantation du projet, c'est-à-dire, les entreprises démarrent par un projet pilote, puis elles le développent en tout contrôle avec leur planning fixé (Sanders, 2016).

Les éléments qui permettent d'étudier ce facteur sont : *le niveau de capacité de l'entreprise en management du projet ; la capacité de suivre un planning précis sur l'usage des opportunités développées sur le marché lié à l'usage du big data et la capacité d'adaptation du planning de projet du big data avec les changements de l'environnement* (Akter et al., 2016), (Wixom & Watson, 2001).

3.6.6 Management de la performance

Ce facteur se réfère à un ensemble des activités engagées par l'entreprise pour augmenter la performance d'une personne ou d'un groupe concerné par une amélioration de la performance organisationnelle (Surbakti et al., 2019), (Mawed & Aal-Hajj, 2017).

Pour pouvoir utiliser efficacement les résultats de l'évaluation de la performance¹², les entreprises doivent faire un passage de « l'évaluation de la performance » (un outil du management plus efficace) au « management de la performance » (Amaratunga & Baldry, 2002). Si l'évaluation de la performance répond à la question « ce qui est arrivé ? », le management de la performance tend à répondre aux questions « pourquoi c'est arrivé ? », « ce qu'il faut faire ? » ;

¹² Cette évaluation fournit à l'entreprise une base pour faire un état sur le niveau de la réalisation des objectifs prédéterminés, identifier les points forts et les faiblesses, faire une projection sur les futures initiatives afin d'améliorer la performance organisationnelle.

ainsi faire une anticipation des changements nécessaires alignés avec la stratégie globale de l'entreprise.

Quelques points-clés qui sont affectés du management de la performance par rapport aux principaux concurrents sont (Mikalef et al., 2019a), (Akter et al., 2016):

- *Rentabilité, les bénéfices en pourcentage des ventes ;*
- *Améliorations en temps du cycle de livraison ;*
- *Réponse rapide à la demande du marché ;*
- *Amélioration de la satisfaction client.*

3.6.7 Structure organisationnelle et taille

Ce facteur concerne la configuration formelle entre les individus et les groupes concernant la répartition des tâches, responsabilité au sein de l'entreprise, ainsi que le nombre des salariés (Surbakti et al., 2019).

Dans le développement d'un projet du big data au sein de l'entreprise, la revue de la structure organisationnelle est considérée comme indispensable, plus particulièrement, dans la section de SI/TI. Ainsi, les changements appliqués à la structure organisationnelle doivent être envisagés et mis en place afin de favoriser une culture digitale et l'alignement stratégique au sein de l'entreprise (Muniz Félix, Tavares, & Freitas Cavalcante, 2018).

Les éléments qui peuvent servir comme les indicateurs de ce facteur, sont : *la flexibilité à adapter la structure organisationnelle aux changements fonctionnels et la capacité à gérer efficacement des effets de changement organisationnel* (Srinivasan & Swink, 2018), (Mikalef et al., 2019a).

3.6.8 Collaboration interdépartementale

L'efficacité d'usage du big data est aussi influencée par le niveau de la collaboration interdépartementale. Elle concerne un processus où les départements de l'entreprise fonctionnent d'une manière coordonnée, avec une compréhension mutuelle, une vision commune et partagent les ressources afin de réaliser les objectifs communs (Surbakti et al., 2019).

Plus particulièrement, il faut souligner l'importance de l'implication élevée de l'équipe qui a implanté la nouvelle infrastructure du big data dans la collaboration avec des employés de différents départements (S. Fosso Wamba et al., 2015). Au stade de l'adoption des technologies du big data, aussi dans leur future exploitation, la collaboration interdépartementale peut contribuer à améliorer l'efficacité de la résolution des problèmes (Zhan, Tan, Ji, Chung, & Tseng, 2017).

L'étude de ce facteur peut se faire par des éléments suivants : *le degré de la collaboration entre les utilisateurs du big data et ses fournisseurs ; les décideurs et le degré de la collaboration entre les décideurs et les fournisseurs des solutions du big data* (S. Shamim et al., 2019), (Akter et al., 2016).

3.6.9 Communication

Etant cité parmi les facteurs potentiels clés de succès d'une implantation de projet SI, la communication se réfère à la capacité d'une entité (personne physique ou unité d'entreprise) de parler et de partager de l'information avec d'autres sans égard des barrières institutionnelles (Surbakti et al., 2019).

La communication efficace est considérée aussi comme un prédicteur d'un management du changement réussi (Hawley, 2016). En plus, comme la raison principale des fiascos dans les trois quarts de projets en BI (Business Intelligence) on distingue la communication pauvre (Saltz & Shamshurin, 2016).

Ce facteur peut être traité par *le degré de partage des connaissances de la part des individus - experts au sein de l'entreprise, le partage des connaissances sur les éléments-clés des fonctionnalités spécifiques du projet (la collecte du big data, l'usage et l'échange des données au sein de l'entreprise, les éléments-clés sur le traitement et la gestion des processus impliqués)* (S. Shamim et al., 2019), (B. Kim, 2012).

3.6.10 Soutien de la haute direction

Le facteur concerne la mesure dans laquelle la haute direction fournit de l'autorité et des ressources pendant et après l'acquisition de nouvelles solutions technologiques (Surbakti et al., 2019), (Popovič et al., 2018).

Le soutien de la haute direction est considéré comme un influenceur important pour de nombreux types d'implantations de TI (Guimaraes, Igbaria, & Lu, 1992), (Igbaria & Tan, 1997).

Selon Karahanna et ses coauteurs (1999), normalement, les utilisateurs tentent d'être conformes aux attentes de la haute direction ; et donc, si elle soutient le système, éventuellement son usage sera efficace.

Les éléments qui décrivent ce facteur au sein de l'entreprise sont : *le degré avec lequel la haute direction encourage et promeut l'usage du système ; le niveau de préoccupation de la haute direction par la satisfaction des utilisateurs et l'intérêt dans les perspectives de développement du projet* (Demoulin & Coussement, 2020), (Maroufkhani, Tseng, et al., 2020), (Wixom & Watson, 2001).

3.6.11 Effet environnemental

Ce facteur concerne l'influence des effets de pression de la concurrence commerciale sur l'entreprise. Ce facteur influence indirectement l'usage du big data (Chen, Preston, & Swink, 2015).

La haute direction joue un rôle de modérateur entre les effets de pressions environnementales et les réactions de l'entreprise (Hambrick & Mason, 1984).

Ces effets peuvent avoir une influence positive à stimuler les tendances imitatives de l'adoption de nouvelles TI par les entreprises, si leurs concurrents les ont déjà implantées (Liang, Saraf, Hu, & Xue, 2007).

Pour étudier le facteur d'effet environnemental, on peut examiner les éléments suivants : *le degré des turbulences technologique aux niveaux local et national ; les turbulences sur le*

marché affectées des concurrents aux niveaux local et national (Côte-Real, Oliveira, & Ruivo, 2017), (Zhu & Kraemer, 2005).

3.6.12 Objectifs précis

Au niveau du contexte général, on admet que le bon fonctionnement d'une entreprise suppose d'avoir une vision commerciale et les objectifs bien établis, claires et compréhensibles à toute entité (physique ou groupe des individus, unité) au sein de l'entreprise.

Pour le contexte du big data, un objectif bien établi sur l'implantation de tel projet fournit d'abord un contexte stratégique à la prise de décision sur coûts-avantages de l'usage des techniques du big data (Gopalkrishnan et al., 2012).

Robey et ses coauteurs (1979) soulignent que l'usage d'une TI au niveau du contexte général affecte positivement sur le rendement des objectifs plus précis. La même relation est affirmée de Shamim et ses coauteurs (2019) pour le contexte de l'usage du big data.

3.6.13 Accent sur l'innovation

Ce facteur concerne le degré avec lequel l'entreprise cherche à développer et à exploiter de nouvelles applications de la technologie du big data (Liu et al., 2022), (Muniz Félix et al., 2018).

Les entreprises portées sur l'innovation possèdent plus de capacités et de flexibilité à s'adapter aux changements provoqués par l'adoption d'un nouveau SI (G. Kim et al., 2011).

Bien évidemment, l'entreprise qui examine en permanence les opportunités innovantes pour l'utilisation stratégique du big data, a plus de chance d'une exploitation efficace du big data.

Les entreprises avec un accent sur l'innovation ont déjà des infrastructures efficaces d'adaptation et de l'exploitation des projets innovants, y compris les processus systématiques et formalisés ajustés à leur planning (Ghasemaghaei, 2019b).

3.7 Thème 7 : Systèmes, outils et technologies

Ce thème comprend les 3 éléments suivants : la qualité du système, l'infrastructure TI et le support des fournisseurs. Comme le souligne son nom, cette thématique représente la capacité et le support en système, mécanismes, outils et technologies nécessaires pour traiter le big data. Evidemment, une infrastructure avec les caractéristiques d'une qualité plus élevée assure un plus haut degré de l'efficacité d'usage.

3.7.1 Qualité du système

Selon Reeves et Bednar (1994), il existe 4 positions dominantes sur la notion de « qualité » : la qualité comme « excellence » (excellence), la qualité comme « valeur » (value), la qualité comme « conformité aux spécifications » (conformance with specifications) et la qualité comme « réponse aux attentes » (meeting expectations).

La vision de l'excellence suggère que la qualité est évaluée sur une norme absolue. La perspective de la valeur affine cette notion pour suggérer que les normes d'excellence doivent être évaluées par rapport aux coûts de leur réalisation. La vue de conformité suggère que la qualité soit évaluée en matière de livraison cohérente et quantifiable de valeur par rapport à la conception spécifique idéale. Enfin, la notion de qualité comme réponse aux attentes suggère que la qualité se définit par la conformité aux attentes des clients qui peuvent être liées à l'excellence, la valeur et d'autres attributs, essentiels aux consommateurs pour façonner leurs perceptions de la qualité (Nelson et al., 2005).

Selon de nombreux auteurs comme Shannon et Weaver (1949), Mason (1978), DeLone et McLean (1992), la qualité du système et la qualité de l'information sont les principaux antécédents initiaux pour le succès du SI. Bien évidemment, il existe une forte relation positive entre la qualité du système et la qualité de l'information (Gorla, Somers, & Wong, 2010).

Ainsi, les dimensions de la qualité du système représentent les perceptions des utilisateurs qui sont en interactions avec le système au fil du temps (Nelson et al., 2005). Dans ce sens-là, les systèmes de meilleure qualité sont perçus comme les systèmes qui sont faciles à utiliser et qui ont des niveaux d'utilité et d'utilisation plus élevées (F. D. Davis et al., 1989).

L'interaction avec le système se produit dans un milieu organisationnel avec l'objectif de compléter une tâche concrète, et pour cela, il est intéressant de considérer la qualité du système dans le spectre du cercle du système à la tâche. Les dimensions du système sont les caractéristiques du système qui sont largement invariantes selon les différentes utilisations et peuvent être évaluées indépendamment de la tâche, du contexte ou de l'application. Les dimensions de la tâche sont ceux pour lesquels une évaluation dépendra de la tâche et du cadre (Nelson et al., 2005).

Pour décrire la qualité du système dans un contexte de big data, on parle souvent des aspects techniques du système qui sont spécifiques, qui sont développés au cours du temps et sont difficiles à imiter (Ji-fan Ren et al., 2017).

La littérature propose les 6 notions suivantes à désigner la qualité du système dans un environnement du big data et au contexte général (S. Shamim et al., 2019), (Ji-fan Ren et al., 2017), (T. H. Davenport et al., 2012), (T. Davenport & Harris, 2007), (Wamba & Akter, 2015), (McAfee & Brynjolfsson, 2012a), (Nelson et al., 2005), (Parasuraman, Zeithaml, & Malhotra, 2005) :

- *Fiabilité du système (system reliability) ;*
- *Adaptabilité du système (system adaptability) ;*
- *Intégration du système (system integration) ;*
- *Accessibilité du système (system accessibility) ;*
- *Temps de réponse du système (system response time) ;*
- *Confidentialité du système (system privacy).*

Nelson et ses coauteurs (2005) décrivent la fiabilité du système comme son degré de disponibilité technique au fil du temps.

L'adaptabilité du système concerne le degré dans lequel la plate-forme de traitement peut être adaptée pour répondre aux divers besoins dans des situations changeantes (Kiron et al., 2014).

L'intégration du système fait référence à la capacité de la plateforme de traitement d'intégrer une variété de données (c.-à-d. transaction, flux de clics, voix et vidéo, etc.). L'accessibilité du système mesure le degré dans lequel une plate-forme de traitement est accessible aux managers, garantissant la commodité et l'évolutivité (T. H. Davenport et al., 2012).

Le temps de réponse du système mesure la rapidité et la promptitude de la plate-forme de traitement (McAfee & Brynjolfsson, 2012b).

Et finalement, la confidentialité du système fait référence à la mesure dans laquelle la plate-forme de traitement est sûre et il n'y a aucun risque de fuite d'informations privées (Court & Barton, 2012).

L'accessibilité et la fiabilité sont des dimensions du système. Elles représentent des propriétés définies qui sont largement indépendantes de l'usage. Le temps de réponse, la flexibilité et l'intégration sont des caractéristiques qui sont peut-être mieux évaluées dans le contexte des tâches concrètes et devront être considérées comme « liés à la tâche » (Nelson et al., 2005).

3.7.2 Infrastructure TI

Concernant le contexte technologique, on constate des améliorations liées aux logiciels dits « open-sources », obligatoires pour traiter la vitesse, le volume et la variété de big data (S. Shamim et al., 2019). L'outil le plus utilisé, c'est Hadoop qui combine le hardware avec les logiciels open-source (Brynjolfsson & McAfee, 2012).

Les autres outils populaires de big data sont Dryad, Apache Mahout, Jaspersoft BI Suit, Pentahu Business analyst, Skytree Server, Tableau, Karmashpere Studio and Analyse and Talend Open Studio, qui ont été détaillés dans le premier chapitre concerné. On constate que la capacité technologique de stockage des informations augmente de près de 100% tous les 3 ans. On a besoin des volumes de stockage plus importants, et les vitesses plus élevées sont nécessaires pour collecter, stocker et accéder aux données (Philip Chen & Zhang, 2014). Ainsi, on voit comment le big data a changé la façon dont les organisations gèrent les données (Oliveira, Furlinger, & Kranzlmüller, 2012).

On comprend que les entreprises ont besoin non seulement d'une capacité organisationnelle à utiliser le big data, mais d'abord d'une capacité technologique nécessaire, d'une infrastructure technologique à traiter ce gros volume des données.

La technologie influence bien le résultat de travail (W. Form, R.L. Kaufman, T.L. Parcel, 1988), en plus, elle peut contribuer à l'intégration des tâches (Altmann, Köhler, & Meil, 2017). McAfee et Brynjolfsson (2012b) suggèrent que la prise de décision fondée sur le big data demande l'utilisation de la technologie la plus efficace et la plus récente pour pouvoir collecter, stocker, analyser et visualiser les données.

Ainsi, l'infrastructure TI au contexte du big data fait référence à l'ensemble des technologies, outils, mécanismes de stockage et de traitement, l'architecture de gestion qui sont capables à assurer le management de chaque cycle de vie des données en entreprise (Shuradze & Wagner, 2016).

Ce facteur a été étudié au travers des éléments suivants : *le degré des technologies de pointe, utilisées en entreprise ; la compétence organisationnelle en gestion des technologies ; le niveau varié des outils technologiques utilisés et le niveau concurrentiel des mécanismes et outils utilisés par rapport aux concurrents* (S. Shamim et al., 2019), (Wixom & Watson, 2001).

3.7.3 Support Fournisseurs

Le support des fournisseurs se réfère au support technique et à la consultation au cours d'adoption et de l'utilisation de la nouvelle technologie et système (Shaul & Tauber, 2013).

Dans le contexte du big data, les études récentes considèrent le facteur du support externe des fournisseurs comme un élément-clé de l'adoption réussie des technologies du big data, et leur utilisation efficace (Gangwar, 2018). Mais là, il y a un risque potentiel quand l'entreprise perd le contrôle (en permettant à l'accès des fournisseurs) sur ses données / informations. Celui-ci doit être bien pris en compte (Maroufkhani, Tseng, et al., 2020).

Un autre élément qui permet d'étudier ce facteur est le niveau de support en formation adéquate fournie des fournisseurs et des développeurs du système (Robey, 1979).

La finalité de cette présentation détaillée sur chacun des 41 éléments composants du modèle conceptuel développé de Surbakti et al. (2019), est de fournir d'abord une compréhension importante sur les facteurs à tester, une compréhension basée sur une consolidation de la revue de la littérature du SI général et de la littérature du domaine spécifique du big data.

**Seconde partie : Les enseignements de l'efficacité
d'usage du big data dans les entreprises arméniennes**

Chapitre 4 Positionnement épistémologique et méthodologie de la recherche

Dans ce chapitre nous expliciterons les présupposés épistémologiques essentiels qui permettent de légitimer notre démarche de recherche : spécifier le positionnement épistémologique en lui proposant un cadre méthodologique précis et cohérent avec l'objectif de notre projet de recherche.

Plus particulièrement, nous allons présenter la conception de la réalité des phénomènes en sciences de gestion que nous nous intéressons à étudier. Nous partageons l'idée qu'une thèse est un seul sujet et une seule partie de sujet à traiter (Pierre Romelaer & Michel Kalika, 2016). Ainsi, nous admettons que la qualité d'un travail de recherche dépend globalement des éléments suivants :

- Cohérence entre le paradigme épistémologique adopté et l'objectif du travail de recherche ;
- Finalité de la recherche à élaborer et le test de connaissance ;
- Type de recherche menée (qualitatif, quantitatif ou mixte) ;
- Techniques de collecte et d'analyse mobilisées ;
- Critères de validité.

Ces éléments seront détaillés dans les sections qui suivent.

4.1 Positionnement épistémologique et paradigme

4.1.1 Positionnements épistémologiques et trois principaux paradigmes en SI

Depuis la fin du XXe siècle, il apparaît indispensable de clarifier le paradigme épistémologique dans lequel s'inscrit la recherche scientifique effectuée dans la plupart des disciplines.

Selon Kuhn (1962), le paradigme, c'est ce que partagent les membres d'une communauté scientifique et inversement, la communauté scientifique est constituée des individus qui partagent un paradigme. Donc, dans ce contexte, Kuhn considère que la communauté scientifique est constituée des praticiens de spécialité scientifique.

Le terme « épistémologie » est apparu au début du XXe siècle et désigne une branche de la philosophie spécialisée dans l'étude des théories de la connaissance (Gavard-Perret, Gotteland, Haon, & Jolibert, 2012). L'épistémologie « *s'interroge sur ce qu'est la science en discutant de la nature, de la méthode et de la valeur de la connaissance* » (Perret et Seville, 2007).

Selon Piaget, l'épistémologie est l'étude de la constitution des connaissances valables (Piaget, 1967) ; ce terme s'utilise comme le synonyme de philosophie des sciences. Dans cette branche de la philosophie, le chercheur occupe un rôle de « producteur de la connaissance », rôle qui lui impose de répondre aux trois questions suivantes :

- Qu'est-ce qui est, pour lui, la connaissance ;
- Quelles sont les hypothèses fondatrices sur lesquelles sa conception de la connaissance repose ;
- Quelle est la manière de justifier la validité des connaissances qu'il élabore (Gavard-Perret et al., 2012).

La réflexion sur ces questionnements épistémologiques conduit le chercheur à se positionner d'abord sur les méthodes qui permettent d'élaborer les connaissances valables. Ici, on parle déjà de la méthodologie, qui apparaît effectivement comme un des volets de l'épistémologie.

Deux questions peuvent éclairer l'ambiguïté qui existe entre les notions de « épistémologie » et « méthodologie », étant souvent une source des confusions dans les milieux scientifiques :

- « Qu'est-ce que la connaissance ? » représente la nature épistémologique.
- « Comment elle se produit ? » est de nature méthodologique.

On suggère qu'il existe trois principaux paradigmes épistémologiques en SI : le positivisme, l'interprétativisme et la théorie critique (Orlikowski & Baroudi, 1991). Une étude réalisée par ces deux auteurs montre que les publications dans les quatre majeures sources de communications du domaine de SI (Communications of the ACM, MIS Quarterly, Proceedings

of ICIS (International Conference on Information System) et Management Science)) s'inscrivent dans un courant positif, prédominant à 96.8%.

Selon Avenier et Thomas (2015), les 4 structures épistémologiques qui sont en forte diffusion dans les recherches en domaines de SI et Organisation et Management, sont :

- Positivisme et le post-positivisme ;
- Réalisme critique ;
- Constructivisme pragmatique ;
- Interprétativisme.

Le tableau 6 définit ces quatre structures.

P. E. Réaliste critique (transcendantal) (Bhaskar, 1978, 1998a,b,c,d; Mingers, 2004; Mingers et al., 2013; Smith, 2006)	
Hypothèses d'ordre ontologique	Il existe un réel en soi indépendant de, et antérieur à, l'attention que peut lui porter un humain qui l'observe. Le réel est organisé en trois domaines stratifiés : le réel profond, le réel actualisé et le réel empirique.
Hypothèses d'ordre épistémique	Le réel profond n'est pas observable. L'explication scientifique consiste à imaginer le fonctionnement des mécanismes générateurs (MG) qui sont à l'origine des événements perçus.
But de la connaissance	Mettre au jour les mécanismes générateurs et leurs modes d'activation. Conception représentationnelle des mécanismes générateurs.
Modes de justification spécifiques	Pouvoir explicatif des MG identifiés. Justification de la validité des MG via des mises à l'épreuve successives dans des recherches quantitatives ou qualitatives.

Tableau 6 Paradigme épistémologique réaliste critique (transcendantal) (Gavard-Perret et al., 2012)

4.1.2 Paradigme de notre travail de recherche

Le paradigme de cette recherche est le post-positivisme : nous partageons les constats du courant du réalisme critique (au sens du réalisme critique transcendantal) ; dans une mesure utilisée par Bhaskar (1998).

La philosophie positiviste, remplacée ensuite par le terme « positivisme », contrairement à d'autres philosophies, comme le christianisme ou le fouriérisme, était « la seule » qui n'était pas nommée d'après son fondateur, Auguste Comte (Pickering, 2011). A l'époque de sa mort (1857), ils existaient de multiples variantes de positivisme provoquant des désaccords et controverses (Weirich, 2005). Ces critiques ont font naître plus tard le paradigme post-positiviste (Boisot & Mckelvey, 2011), (Gephart, 2013), le réalisme scientifique (Hunt, 1990), (Hunt, 1993), (Hunt, 2005), (Bunge, 1993) et le réalisme critique (Bhaskar, 1988).

Le positivisme « logique » suppose que le monde est réel, relativement stable et objectif, rempli de phénomènes qui peuvent être rationnellement connus et rationnellement analysés par des observateurs (Boisot & Mckelvey, 2011).

Le post-positivisme prend ses racines (positivisme logique) dans les sciences de la nature, de la physique et de la mécanique, particulièrement. Selon Gephart (Cortina & Landis, 2013), depuis Popper (1959), le post-positivisme diffère du positivisme principalement en supposant que le monde est probabiliste et non déterministe, et, par conséquent, on ne peut que falsifier, et ne pas confirmer les hypothèses.

Nous croyons ainsi que le réel et la connaissance sont différents, selon Roy Bhaskar / A Realist Theory of Science, 1975/(Bhaskar, 2013). Le réel a trois niveaux : (1) le réel /profond ou bien tout ce qui existe ; (2) l'actuel ou l'occurrence d'événements et (3) l'empirique, qui est observable.

Ici, d'une part, on peut noter que le réalisme critique s'éloigne du positivisme logique, en considérant que le réel n'est pas directement connaissable et donc l'observateur ne peut pas avoir l'accès au réel profond. D'autre part, si l'on admet que les mécanismes générateurs du réel existent, il faut donc mettre au jour ces mécanismes et leurs modes d'activation (par exemple, contextes). Ainsi l'observateur doit prendre en compte le contexte des mécanismes et le poids des structures (Thiétart, 2014). Dans cette logique, les chercheurs post-positivistes contemporains reconnaissent les limites de la connaissance qu'ils tendent de produire : plusieurs facteurs, y compris la culture, l'expérience, l'histoire, etc. ont un impact sur le travail de recherche, ainsi que sur les résultats obtenus (Weber, 2004).

Ainsi, toute observation est imparfaite et sujette à des inexactitudes ; donc, même si la science et la recherche peuvent faire leur possible pour atteindre la réalité, ce but ne peut jamais être réalisé (Ricucci, 2010).

Les hypothèses d'ordre ontologique et épistémique, ainsi que le but de la connaissance et les modes de justification spécifiques sont détaillés dans le tableau 6.

Ainsi, nous ne réfutons pas un positionnement épistémologique par rapport à un autre, car nous trouvons que le positionnement épistémologique adapté dans un travail de recherche dépend tout d'abord de sa nature, ses caractéristiques et ses spécificités. Dans cette logique, chacun des positionnements épistémologiques peut être applicable et cohérent avec les buts et des objectifs du travail de recherche concerné. Comme cité par Weber (2004), le but d'un chercheur, dans cette tâche de choix d'un positionnement plus cohérent à son travail de recherche, doit être l'amélioration des connaissances sur un phénomène observé.

Nous nous opposons aussi au préjugé répandu que l'outillage et les techniques quantitativistes sont réservés aux postures positivistes, par contre, ceux de nature qualitative conviennent aux postures constructivistes et interprétativistes. Ainsi, nous partageons les considérations de David (2012) et dépassons les oppositions classiques entre démarche inductive et démarche hypothéticodéductive; entre positivisme et constructivisme; et intégrons les différentes approches au sein d'un même schéma conceptuel.

Pour pouvoir argumenter le choix du positionnement épistémologique et de la méthodologie adaptée, nous allons d'abord comprendre quelques éléments-clés sur la nature de cette recherche.

4.1.2.1 Justification du choix du positionnement épistémologique et mode de raisonnement adopté à partir de quatre axes de Mbengue et Vandangeon-Derumez (1999)

Nous essayons de justifier la pertinence et l'importance de l'explicitation du positionnement épistémologique adopté dans ce travail de recherche à partir des quatre axes proposés de Mbengue et Vandangeon-Derumez (1999). Pour ce faire, nous allons justifier notre choix du positionnement en décrivant aussi les autres paradigmes épistémologiques. Cette approche illustrera la cohérence du choix de paradigme avec notre projet de recherche.

Les quatre dimensions proposées par Mbengue et Vandangeon-Derumez sont : le statut accordé aux données (1) ; le mode de collecte des données (2) ; la relation entre la théorie et les observations empiriques (3) et les critères de scientificité de la recherche (valeur scientifique des résultats attendus (4). Dans les sections qui suivent, nous allons en détail traiter ces quatre éléments.

4.1.2.2 Statut accordé aux données

La première dimension proposée par Mbengue et Vandangeon-Derumez concerne le statut accordé aux données.

D’abord, nous nous posons deux questions concrètes : « quelles données sont disponibles au contexte de notre travail de recherche » et « quel type d’étude il faut mettre en place pour les recueillir ».

Afin de fournir des réponses à ces deux questions, nous devons dans un premier temps clarifier le statut des données. Si les données sont des mesures du réel considérées comme « objectives », *la réalité du réel, la naturalité de la nature, dans ce sens, toute proposition peut être considérée comme vrai qui décrit effectivement la réalité ; et le but de la science est de découvrir cette réalité ;* (Le Moigne, 1994), (David, 2012). Ou bien les données sont des interprétations du réel qui « peuvent être » considérées comme « objectives » ; par contre, les données peuvent être considérées comme « subjectives », les interprétations du réel et elles peuvent être le résultat d’une construction par interaction entre le chercheur et l’objet étudié (Figure 31). Variant en fonction du statut accordé à la subjectivité du chercheur, les démarches de recherche développées avec un positionnement constructiviste apportent un éclairage sur les mécanismes qui sous-tendent les stéréotypes, quand celles s’appuyant sur une posture positiviste tentent d’en établir le contenu, de les classer et d’identifier leurs déterminants (Bertereau, Marbot, & Chaudat, 2019).

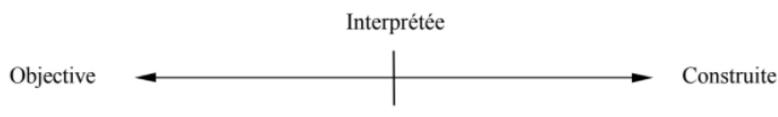


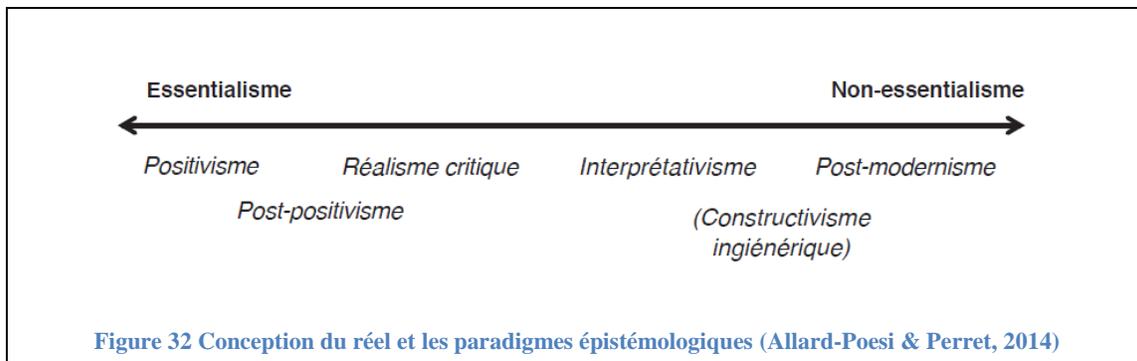
Figure 31 Statut accordé aux données (Mbengue & Vandangeon-Derumez, 1999)

Dans les deux premiers scénarios, la réalité est considérée comme ayant une essence propre, déterminée par des lois naturelles et « uniques », où les données existent indépendamment de l'attention que l'observateur peut y porter et précèdent son intervention. « *L'observation de l'objet réel par l'observant ne modifie ni l'objet réel ni l'observant. Si l'observant est modifié, cela ne concerne pas la science (l'esprit humain ne fait pas partie des objets réels sur lesquels il puisse lui-même s'exercer)* » ; (Le Moigne, 1994), (David, 2012). La connaissance du réel exige du chercheur une posture de neutralité par rapport au phénomène étudié (Avenier & Thomas, 2011). Dans les deux autres cas, la réalité peut être considérée comme « multiple », étant un résultat de l'expérience des acteurs sociaux.

Dans les futures interprétations, nous nous focaliserons sur les deux premiers scénarios, qui découlent de la nature de notre travail de recherche.

Pour proposer ainsi une interprétation plus détaillée, nous pouvons toutefois nous appuyer sur la figure 32 (Allard-Poesi & Perret, 2014). Cette figure illustre la conception du réel et les paradigmes épistémologiques. La nature essentialiste de la conception du réel suppose qu'il a une essence propre, unique ; indépendante de l'attention de son observateur. Ainsi, notre posture épistémologique, le réalisme critique est conforme à ces constats.

La nature non - essentialiste de sa part suppose que la réalité n'est pas donnée, mais construite et il y a une interdépendance entre la connaissance, l'observateur et le sujet. Les paradigmes interprétativistes, postmodernisme et constructivisme ingénierique s'inscrivent dans cette orientation.



Dans cette approche, on considère que le contexte dans lequel nous examinons l'efficacité d'usage du big data (la réalité en soi) existe indépendamment de notre observation et

intervention. Ici, il est important de souligner notre position (observateur) d'extériorité et de neutralité face au sujet étudié. Concomitamment, nous croyons que l'observateur ne peut pas toujours être extérieur, mais il *doit essayer* d'atteindre à un idéal d'objectivité (Avenier & Thomas, 2015).

Notre démarche de recherche prend appui principalement sur des données quantitatives ; mais elle mobilise aussi des données de type qualitatif. Cette pluralité méthodologique vise à améliorer la compréhension du phénomène émergent. Ainsi, dans ce contexte-là, l'évidence qualitative va nous servir comme une source de richesse à notre travail de recherche (R. Weber, 2004). Et donc nous essayons d'approcher au plus près la réalité en multipliant les méthodes (Allard-Poesi & Perret, 2014).

4.1.2.3 Mode de collecte des données

Le mode de collecte des données est influencé d'abord par l'essence du réel (Figure 32).

Dans le premier scénario décrit ci-dessus, la réalité est déterminée par des lois naturelles et les données sont des mesures de celle-ci, considérées comme objectives. Dans ce scénario, le chercheur est dans une position de recherche des faits mesurés par des données, assurant une indépendance forte de l'observateur et de l'objet étudié : les données existent en dehors de toute interprétation humaine, l'instrument de collecte des données est objectif. Dit autrement, les faits restent objectifs, quels que soient l'observateur du sujet et le mode de collecte des données utilisé.



Figure 33 Dépendance du chercheur et de l'objet de l'étude (Mbengue & Vandangeon-Derumez, 1999)

Dans un deuxième scénario, les données sont considérées comme des interprétations de la réalité. Cette réalité est donc considérée comme objective et l'observateur interprète des faits mesurés par des données objectives (Figure 33). Ici on tend à assurer l'objectivité des faits étudiés par une mise en œuvre des stratégies de collecte adéquates (Miles & Huberman, 2003).

4.1.2.4 Relation entre la théorie et les observations empiriques

Les trois méthodes principales qui permettent la validation des connaissances sont la déduction, l'induction et l'abduction (Charreire-Petit & Durieux, 2014). On peut toutefois s'appuyer sur le schéma en figure 34 pour une meilleure compréhension.

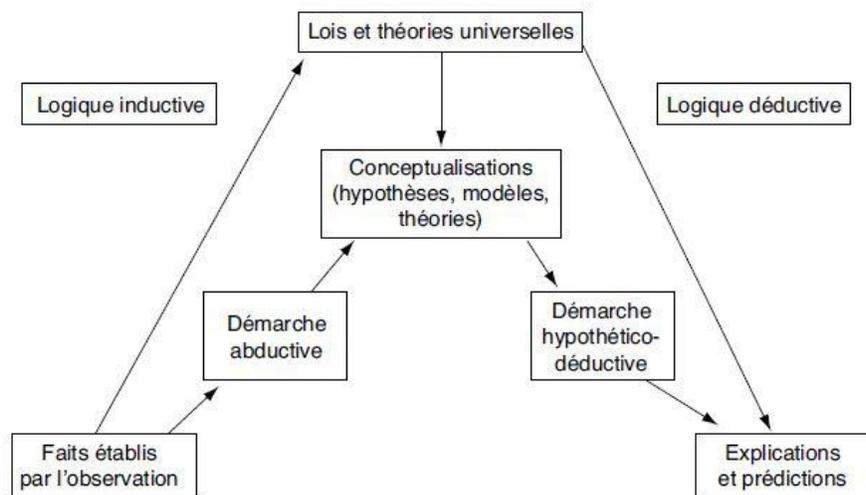


Figure 34 Modes de raisonnement et connaissance scientifique (Charreire-Petit & Durieux, 2014)

Quand l'on parle de la relation entre la théorie et les observations empiriques, deux scénarios possibles émergent :

- Dans un premier scénario, « la connaissance produite est le fruit des observations empiriques », où l'observateur remonte des faits à la loi.
- Dans un deuxième scénario, « la connaissance théorique précède les observations empiriques et donc sert à vérifier la connaissance initiale ». Dans ce scénario, l'observateur connaît à l'avance ce qu'il peut découvrir.

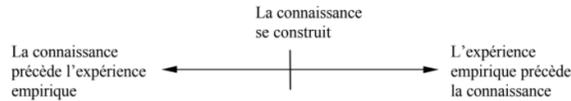


Figure 35 Construction des connaissances et les observations empiriques (Mbengue & Vandangeon-Derumez, 1999)

Sur la figure 35, on peut remarquer que la connaissance se construit au milieu de cette l'interaction entre la théorie et les données empiriques.

Nous avons deux scénarios à discuter : la première démarche hypothético-déductive vise à vérifier ; les observations sont soumises à la théorie, au cas contraire, nous sommes dans une démarche hypothético-déductive de falsification. Cette démarche de recherche est donc définie avant l'intervention de l'observateur sur le terrain.

Nous appliquons ainsi le raisonnement de déduction (de lois et théories universelles aux explications et prédictions) et donc une démarche hypothético-déductive est utilisée dans ce travail de thèse (R. Thiéart, 2014). Ce mode de raisonnement permet le cheminement suivant : sur la base de la théorie de l'efficacité d'usage du big data et à partir d'un modèle conceptuel, nous l'examinerons sur le terrain. Nous pourrons ainsi parvenir à une conclusion pour répondre à notre question principale de recherche. Il s'agit donc bien de tester un modèle, développé préalablement et le mettre à l'épreuve dans les situations particulières.

4.1.2.5 Critères de scientificité de la recherche

Un des points importants à aborder dans cette section est l'estimation de la validité de recherche qui permet d'évaluer les connaissances produites.

Pour accomplir le processus de création de la connaissance, nous allons expliciter les deux éléments importants cohérents avec le paradigme positiviste : la validité interne et la validité externe. Cette explicitation nous permettra de déterminer la qualité et la valeur des connaissances élaborées pour démontrer la valeur scientifique de notre travail de recherche (Creswell & Plano Clark, 2017), (Gavard-Perret et al., 2012).

Dans le paradigme épistémologique positiviste, la justification de la fiabilité des connaissances produites est particulièrement liée à l'outillage de la mesure des phénomènes étudiés (Gavard-Perret et al., 2012).

Ayant mobilisé dans ce travail de thèse une démarche méthodologique à dominante quantitative, nous avons développé l'instrument de l'enquête à l'aide d'une opérationnalisation des variables suivant une méthode bibliométrique (Tableau 9), préalablement validée. Les étapes de construction de l'instrument mettent en évidence la fiabilité globale de toutes les phases sous-jacentes, qui sont détaillées dans la section concernée et assurent une qualité globale de scientificité de la recherche.

Ainsi, ce chapitre vise à expliciter d'une manière complète et précise, toutes les étapes de la trajectoire fixée à notre démarche méthodologique. La finalité est donc de construire une réflexion cognitive fiable sur les processus mis en place dans ce travail de thèse.

Ce cheminement logique démontre les deux critères importants de la scientificité de la recherche :

- *Validité interne*, que nous assurons par la cohérence interne dans un ensemble de processus des travaux de recherche ;
- *Validité externe*, mise en évidence par les ressources empiriques impliquées dans le travail (l'échantillon mobilisé pour l'étude quantitative, les cas précis pour les études qualitatives) (Drucker-Godard, Ehlinger, & Grenier, 2014).

Tous ces aspects sont soigneusement explicités dans les sections concernées de ce chapitre.

4.2 Méthodologie de la recherche

Nous rappelons que ce projet de recherche vise à tester un modèle conceptuel développé par Surbakti et al. (2019) dans un contexte arménien. Dans le chapitre 5, nous allons détailler en profondeur la nature de cette recherche, ainsi que l'importance de mener une recherche contextualisée en sciences de gestion sera détaillée. Nous admettons que les spécificités contextuelles ne doivent pas être « dépassées » (Buchanan & Bryman, 2007). Pour proposer un

travail de recherche avec les résultats actionnables (Louitri & Sahraoui, 2014), nous réalisons une étude avec une application de trois pôles de contextualisation proposés et adaptés de Livian (2020). Nous allons aussi traiter un quatrième élément sur l'environnement technologique du contexte. Ce dernier découle bien de la nature de ce travail de thèse. Nous allons résumer ces aspects dans le chapitre suivant de cette thèse et dans les sections ultérieures : nous nous concentrons principalement sur le traitement des éléments méthodologiques de conduite de trois études mobilisées.

4.2.1 Présentation de la question de recherche à partir de six paramètres de Noël (2012)

Pour rappel, nous proposons de traiter dans cette thèse la question de recherche suivante :

« Quels sont les facteurs qui contribuent le plus à l'efficacité d'usage du big data en entreprise en contexte arménien ? ».

Tout d'abord, nous proposons de suivre les 6 paramètres conseillés de Noël (2012), pour détailler notre question de recherche.

Ces paramètres sont : la clarté ; la faisabilité ; la pertinence ; l'interrogation ; l'auditoire pour les résultats et l'intelligence.

Dans les lignes qui suivent, nous expliciterons la question de recherche par rapport à chacun de ces éléments.

4.2.1.1 Élément 1 : Clarté

Notre question de recherche vise à traiter un problème concret : recenser les facteurs importants qui contribuent le plus à l'efficacité d'usage du big data en entreprises, selon les praticiens. Ainsi, ce travail tend à apporter une réponse complète à cette question importante du point de vue académique et pratique.

4.2.1.2 Élément 2 : Faisabilité

Afin de répondre à la question de recherche posée, une trajectoire d'actions à 3 étapes, détaillées ci-dessous, a été définie. Ainsi, la démarche de production de la connaissance de ce

travail de recherche est axée sur trois volets : étude qualitative exploratoire, étude quantitative confirmatoire et étude qualitative confirmatoire.

Nous démarrons la recherche par une étude qualitative exploratoire. L'objectif est de mettre en évidence la présence des 7 thèmes du modèle conceptuel dans l'activité d'une entreprise ayant une forte expertise dans le domaine de l'exploitation du big data.

Nous continuons après par une étude quantitative, dont l'objectif principal est de tester quantitativement le modèle et y recenser la réponse à la question de recherche.

Et nous finalisons la recherche sur le terrain par une étude qualitative confirmatoire, qui tend d'apporter un éclairage sur les interprétations des résultats, issus de l'enquête quantitative.

Nous reviendrons sur chacune de cette étape et les détaillerons dans les sections concernées qui suivent.

4.2.1.3 Élément 3 : Pertinence

Comme la valeur des connaissances élaborées par un chercheur est appréciée d'abord par son utilité épistémique et pragmatique, il est nécessaire au préalable de justifier la pertinence de la question de recherche proposée, du point de vue pragmatique et du point de vue épistémique (apport académique).

La revue de littérature exposée dans les chapitres précédents récence des éléments importants permettant de justifier cette pertinence.

Du point de vue épistémique, ce travail :

- Développe la théorie de l'efficacité d'usage au contexte général du SI (Burton-Jones & Grange, 2013) ;
- Permet, une opérationnalisation du construit « d'efficacité d'usage » dans un contexte particulier du big data et propose un instrument et l'outillage complet pour l'étude au niveau d'entreprise ;
- Améliore l'état de la connaissance sur le sujet de l'usage du big data en entreprise au travers du modèle proposé de (Surbakti et al., 2019) et permet 2 adaptations du modèle.

Du point de vue pragmatique, ce travail permet :

- De comprendre profondément quels facteurs impactent le plus l'efficacité d'usage du big data au sein de l'entreprise, selon sa référence d'usage du big data, « utilisateur novice » ou « utilisateur avancé », au contexte arménien ;
- D'évaluer, au niveau de l'entreprise, l'efficacité d'usage du big data et de pouvoir recenser les points faibles ;
- Suite aux 2 scénarios possibles explicités dans les études de cas, présentés en chapitre 5, établir un plan d'actions.

Les principales contributions théoriques et managériales de la thèse sont détaillées dans la section concernée du chapitre 7.

4.2.1.4 Élément quatre : Interrogation

La démarche adaptée à ce travail de recherche a pour l'objectif, non seulement de décrire profondément le sujet étudié, mais aussi d'être ouvert à explorer en essayant de le comprendre et expliquer à partir des études de cas, détaillées dans le chapitre suivant.

4.2.1.5 Élément 5 : Auditoire pour les résultats

Il convient de souligner qu'il n'existe pas d'auditoire universel pour les résultats obtenus d'un travail de recherche. Ainsi, nous délimitons cet auditoire à 2 niveaux :

- Au macro-niveau, les connaissances produites dans ce travail de recherche intéressent la communauté scientifique des sciences de gestion ou sciences du management, plus particulièrement les champs disciplinaires en systèmes d'information au sein de la communauté francophone, ainsi que les personnes qui s'occupent de la mise en œuvre des systèmes d'information ;
- Au micro-niveau, nous ciblons les communautés scientifiques et pratiques du domaine en Arménie, suite la contextualisation de la recherche proposée dans le chapitre 5.

4.2.1.6 Élément 6 : Intelligence

Ce dernier paramètre proposé par Noel (2012) suppose qu'il faut faire aussi attention à l'état d'intelligence (et non purement l'intellectualisme) : l'intelligence humaine peut se

connaître, mais le principe de sa connaissance lui vient du dehors, sans image sensible, aucune connaissance n'est possible: et donc il faut être progressiste dans les sciences (Longwell & Rousselot, 1911).

Ainsi, notre démarche méthodologique adoptée prend en compte aussi cet aspect, résumé dans le dernier chapitre de ce travail.

Pour rappel, la problématique de la recherche consiste à étudier l'usage du big data en entreprise. Plus spécifiquement, il s'agit d'évaluer la réalisation de valeur du big data, traitée sous l'angle de l'efficacité d'usage.

Pour ce faire, nous examinerons d'abord les fondements théoriques du concept de l'efficacité d'usage d'un SI (Burton-Jones & Grange, 2013). Cela nous permettra de construire une compréhension complète basée sur un modèle conceptuel de l'efficacité d'usage dans le contexte du big data (Surbakti et al., 2019), (Surbakti, 2020), (F. P. Sejahtera, Wang, Indulska, & Sadiq, 2018).

La problématique découle effectivement de la revue de littérature du domaine et envisage une mise en œuvre des connaissances développées préalablement. Ainsi, il semble que les objectifs et la problématique fixés à ce travail nous conduisent à pleinement justifier le choix d'un positionnement épistémologique positiviste.

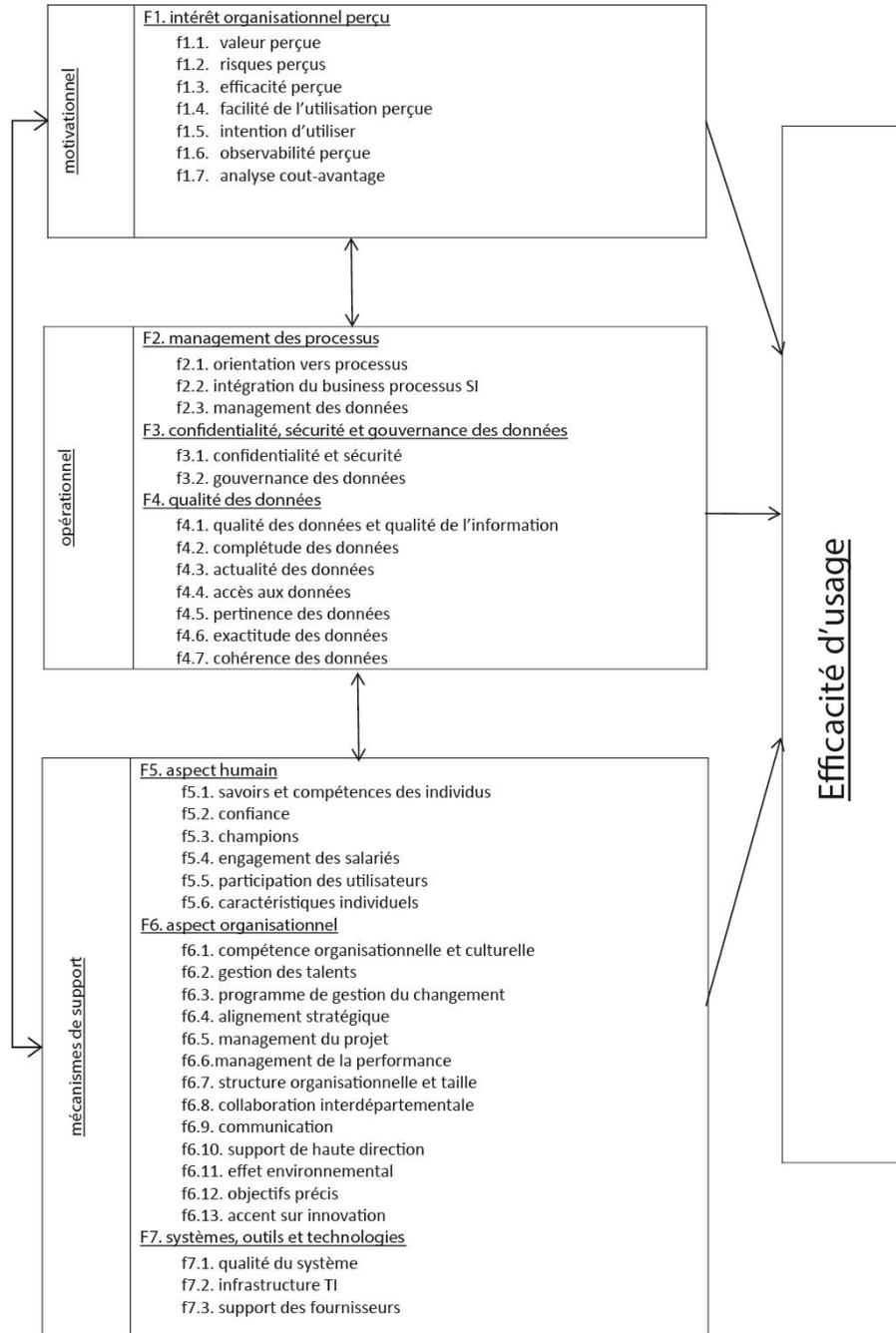
4.2.2 Deux volets de la méthodologie de recherche

Comme déjà noté, l'élément-clé de notre travail de recherche est le modèle de l'efficacité d'usage de Surbakti et al. (2019) (Figure 36). Etant donné que celui-ci a déjà été traité dans le chapitre 3, seul un rappel sera fait dans la présente section. Le modèle est composé de 7 thèmes, comprenant les 41 facteurs qui ont du potentiel pour influencer l'efficacité d'usage du big data en entreprise.

Les deux processus globaux de production de la connaissance, l'exploration et le test, ont été mobilisés dans ce travail de recherche. Ainsi, nous avons choisi d'appliquer une pluralité méthodologique pour mobiliser aussi bien les avantages du mode quantitatif que ceux du mode qualitatif. Nous croyons que cette conduite permet de maîtriser le phénomène dans ses meilleures dimensions et que ces deux approches se complètent aisément dans notre travail de recherche.

Ainsi, dans le cadre de ce travail de recherche, nous avons réalisé trois études que nous détaillerons dans la suite.

Figure 36 Modèle conceptuel de l'efficacité d'usage proposé de Surbakti et al. (2019)



4.2.2.1 Première étude : Étude qualitative exploratoire

Nous démarrerons par une étude qualitative exploratoire, fondée sur des entretiens non directifs. L'objectif de cette étude est de mettre en évidence empiriquement la présence des 7 thèmes du modèle conceptuel (Annexe 22).

En effet, avant de pouvoir tester quantitativement le modèle, il nous fallait d'abord vérifier que ses composants sont bien présents dans l'activité d'une entreprise ayant une forte expertise dans le domaine de l'exploitation du big data.

De plus, elle nous permet de récolter une masse critique de données sur un phénomène émergent du big data dans un objectif de progression de la recherche (Noël, 2012).

D'une part, nous admettons que le recueil de données primaires nous offre l'opportunité de se confronter directement à la « réalité » que nous avons choisi d'étudier. D'autre part, nous trouvons qu'un accès aux données secondaires nous donne le sentiment de complétude de notre recherche : nous aurons l'impression d'avoir eu accès « à tout ce qui était accessible » (Baumard & Ibert, 2014).

Pour cela, et de manière préliminaire, nous avons collecté des données primaires et secondaires. Les données primaires ont été recueillies via des entretiens non-directifs réalisés avec le directeur régional d'ABC Group. Le caractère non directif des entretiens a été privilégié pour laisser le maximum de latitude à l'interviewé dans ses descriptions et explications (Gavard-Perret et al., 2012), (Blanc, Drucker-Godard, & Ehlinger, 2014). Cette méthode d'entretien facilite effectivement la qualification des idées exposées. Ainsi, ce choix répond au caractère exploratoire de la première partie de l'étude. Différents entretiens avec le même interviewé ont permis de centrer chacun des entretiens sur un questionnement de départ spécifique. Cette conduite nous a permis de progresser dans notre compréhension du fonctionnement de l'entreprise et de se focaliser sur le bloc de l'exploitation du big data par l'entreprise.

Enfin, le directeur régional d'ABC Arménie a été retenu comme unique interviewé. La position de notre interlocuteur nous permet de nous situer à un niveau plus stratégique qu'opérationnel du sujet. Une véritable compréhension du fonctionnement d'ABC, en particulier au regard de l'utilisation du big data et des attentes qui y sont attachées, a été explorée.

Au total, 5 entretiens non-directifs ont été réalisés d'une durée d'environ 1h 15 chacun (Tableau 7).

Entretiens exploratoires		N=5
Entreprise	ABC	
Qualité	Directeur régional du groupe (A.A.)	
Date	Mode	Durée
10/05/2019	(F2F) Erévan, Arménie	1h 40m
02/07/2019	(F2F) Erévan, Arménie	1h 15m
25/09/2019	Via Skype	1h 10m
18/10/2019	Via Skype	1h 10m
17/09/2020	Via Zoom	1h 5m
Total		6h 20m

Tableau 7 Détail des entretiens exploratoires

Les deux premiers entretiens se sont déroulés en face à face, et ont fait l'objet de prises de notes.

Les deux autres ont eu lieu *via* Skype, et un autre *via* Zoom ; ils ont été enregistrés. L'ensemble a ensuite été retranscrit de manière à faire l'objet d'une analyse de contenu thématique ; conformément aux étapes énoncées par Bardin (2013) (Figure 37). La langue du travail était l'arménien, les résultats sont ensuite traduits en français, quelques exemples sont illustrés dans le tableau 8.

L'analyse a été guidée par les 7 thèmes du modèle à tester. Autrement dit, les thèmes de l'analyse étaient pré-identifiés avant celle-ci, ils n'ont pas émergé de celle-ci. Au total, les 6h 20m de discours ont permis d'aborder les 7 thèmes du modèle : comme on peut voir dans la figure 38, les facteurs en rouge sont aussi explorés lors de l'analyse.

Figure 37 Les principales étapes de l'analyse de contenu

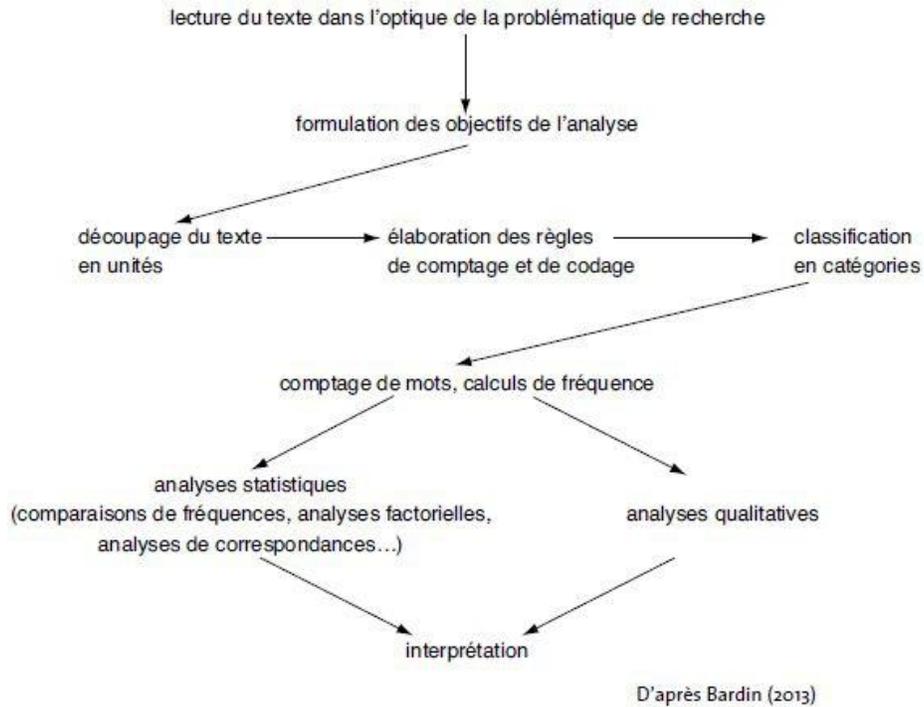


Figure 37 Les principales étapes de l'analyse de contenu

De manière complémentaire, cette recherche s'appuie sur des données secondaires relatives à ce terrain. Il s'agit des données issues du site Internet d'ABC, de sa page Facebook, de ses brochures de marketing, et d'autres documents confidentiels autorisés. Ces données secondaires nous ont essentiellement permis de nous familiariser avec le terrain et d'en rédiger la description détaillée de l'entreprise et sa relation à l'exploitation du big data, présentée dans le chapitre suivant.

Tableau 8 Exemples de codage des entretiens de l'étude qualitative exploratoire

<i>Devis</i>	<i>Code</i>
<p><i>Dans la langue originelle (Arménien)</i> «...որոտում պրոֆեսիոնալներ գտնելու լուրջ խնդիր կա, հատկապես տվյալագետների առումով, լուրջ տեխնիկների, ծրագրի մենեջերների, ովքեր լրջորեն տիրապետում են այս ոլորտի ծրագրերի կառավարման ասպեկտներին...մենք այս խնդրին կարող ենք առնչվել՝ Հայաստանում էլ, թե Ռուսաստանում, թե արևմուտքում...ճիշտա կարողա ամխաներով մասնագետ ման գանք, դժվարությամբ, բայց վերցնում ենք իրենց գործում իսկական</p>	<p>(f5.1) գիտելիքներ և հմտություններ (f5.1) savoirs et compétences des individus</p>

սրտֆեսինալներին...»

Traduite en Français

«...c'est vraiment très difficile de trouver des spécialistes du domaine, des datascientistes, des techniciens, des chefs de projet avec une expertise assez forte en domaine... et l'on rencontre ce problème partout, en Arménie, en Russie, en Europe... Parfois on peut rechercher un spécialiste pendant des mois...mais on recrute vraiment des professionnels de leur métier...»

Dans la langue originelle (Arménien)

«...ֆիրմայի տեխնոլոգիական ծագումն իհարկե նպաստում է սոցիալների կառավարումն ավելի արդյունավետ դարձնելուն: Կոնկրետ էս բրնի համար նպատակն էղերա ի սկզբանե հեշտացնելու մարդկանց կյանքն օնլայն ու օֆֆլայն տիրույթներում, հա սա սսենք որպես միասիս հա կարեղիս անակերպել, որի հիմքում ընկածա էղել տեխնոլոգիական որոտի նվաճումները...»

Traduite en Français

«...née technologique la firme dispose d'une expertise très forte dans le domaine qui assure une culture organisationnelle favorable à l'exploitation des données. Sa mission, c'est aider ses consommateurs à mieux naviguer dans les mondes en ligne ou hors-ligne, en suivant et en utilisant les avancements technologiques dans les services et produits proposés. Sa stratégie, c'est de proposer un meilleur service grâce à la technologie...»

(f6.1) կազմակերպչական հմտություններ

(f6.1) compétence organisationnelle

4.2.2.2 Deuxième étude : Étude quantitative principale

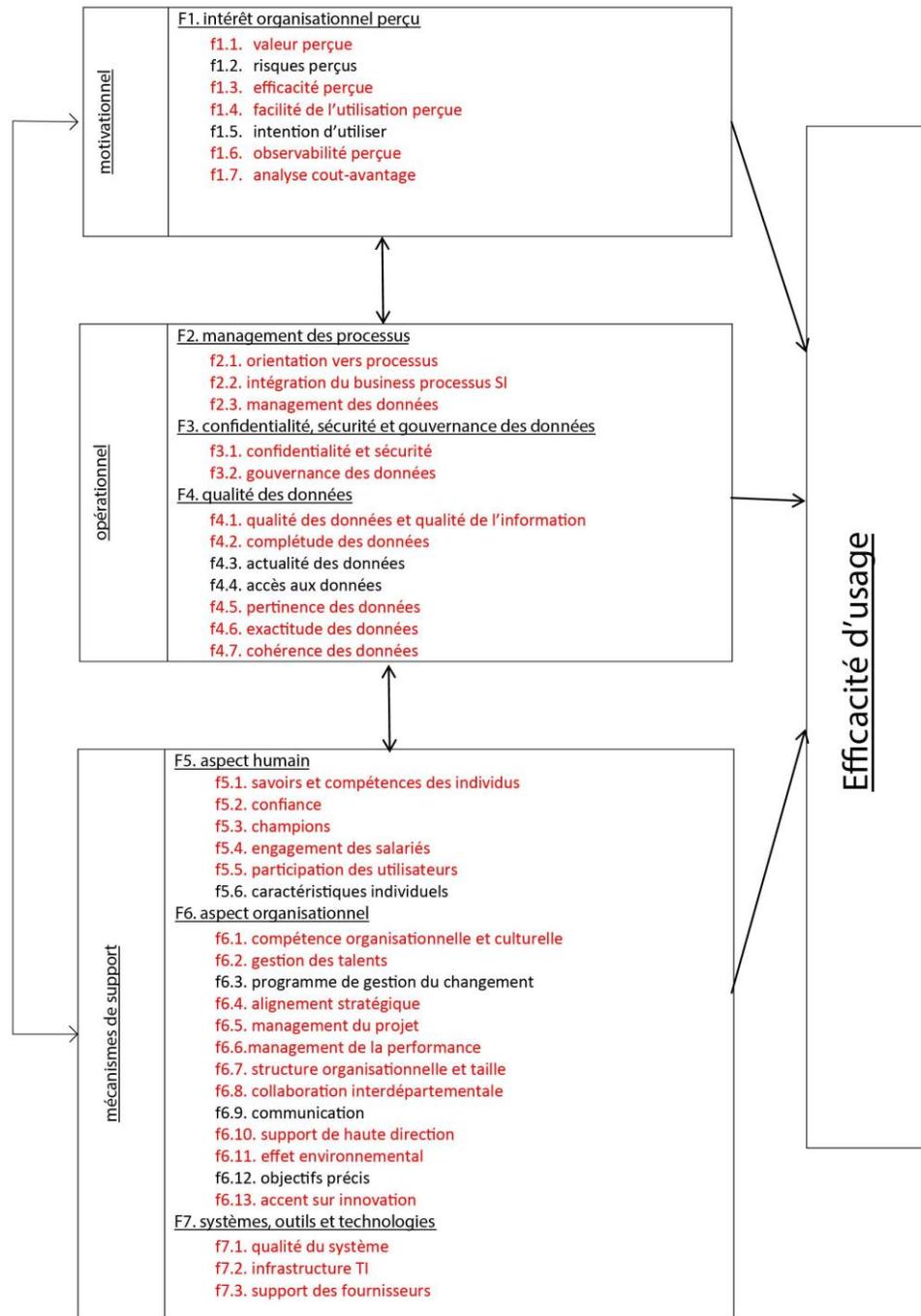
Pour la suite, nous continuons avec une étude quantitative, qui permet d'opérationnaliser les variables du modèle et de le tester. Le test se rapporte à la mise à l'épreuve de la réalité d'un objet théorique. Le test du modèle conceptuel a été réalisé en utilisant la méthode d'enquête auprès des professionnels en big data, représentant 16 entreprises arméniennes.

Nous proposons de résumer la conduite de l'étude quantitative par les 9 étapes.

4.2.2.2.1 Première étape : Opérationnalisation des variables composantes du modèle conceptuel

Pour la collecte des données quantitatives, nous avons développé un instrument de mesure de l'efficacité d'usage.

Figure 38 Premiers résultats de l'analyse thématique



Nous avons utilisé une méthode bibliométrique à partir des 40 articles analysés, où les échelles de mesure de 41 facteurs du modèle sont recensées et reconnues comme étant scientifiquement solides.

Les construits et les éléments de l'instrument d'enquête sont détaillés dans le tableau 9, une description plus complète des éléments par les thèmes et les facteurs du modèle est présenté en annexe 2. Cette opérationnalisation est aussi une des suggestions de recherche future de Surbakti et al. (Surbakti et al., 2019) (pour le construit d'efficacité d'usage en particulier).

4.2.2.2 Deuxième étape : Construction de l'instrument d'enquête

Nous avons développé l'instrument d'enquête (détaillé en Annexe 3) en utilisant une échelle de réponse de Likert à sept échelons (allant de 1- pas du tout d'accord à 7- tout à fait d'accord). Ainsi, cet instrument nous permet de mesurer les perceptions relatives à l'efficacité d'usage du big data auprès des professionnels du domaine. Le choix d'une échelle de Likert est conditionné d'abord par le fait, qu'elle s'utilise souvent dans les recherches en SI et donc connue par les interviewés. Ainsi, elle présente une facilité d'administration pour la recherche.

Tous les éléments de l'enquête sont basés sur des variables latentes publiées précédemment avec des propriétés psychométriques qui soutiennent leur validité (Tableau 9).

Tableau 9 Construits avec les éléments de l'instrument d'enquête

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Qualité du système (QS) System Quality	Big data system reliability	<ul style="list-style-type: none"> The system operates reliably to manage big data. The system performs reliably to manage big data. The operation of the system is technically available to manage big data. 	(Ji-fan Ren et al., 2017)
Qualité du système (QS) System Quality	Big data system adaptability	<ul style="list-style-type: none"> The system can be adapted to meet a variety of business needs. The system can flexibly adjust to new demands or conditions during the treatment of big data. The system is flexible in addressing needs as they arise during the treatment of big data. 	(Ji-fan Ren et al., 2017)
Qualité du système (QS) System Quality	Big data system integration	<ul style="list-style-type: none"> The system effectively integrates data from different areas of the company. The system pulls together data that used to come from different places in the company. The system effectively combines different types of data from all areas of the company. 	(Ji-fan Ren et al., 2017)

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Qualité du système (QS) System Quality	Big data system accessibility	<ul style="list-style-type: none"> The system allows information to be readily accessible as requested. The system makes information very accessible. The system makes information easy to access. 	(Ji-fan Ren et al., 2017)
Qualité du système (QS) System Quality	Big data system response time	<ul style="list-style-type: none"> The system does not take long time to process the requests. The system provides information in a timely fashion. The system processes the requests quickly. 	(Ji-fan Ren et al., 2017)
Qualité du système (QS) System Quality	Big data system privacy	<ul style="list-style-type: none"> The system protects information about personal data. The system protects personal identifiable data. The system offers a meaningful guarantee that it will not share private information. 	(Ji-fan Ren et al., 2017)
Infrastructure TI (ITI) IT infrastructure	Infrastructure TI	<ul style="list-style-type: none"> My company uses the latest technology to manage big data. Our technological competency helps us to enhance big data management. We use a variety of technological tools to manage big data. Our big data technological tools are more effective than those used by others in the industry. 	(S. Shamim et al., 2019)
Support Fournisseurs (SF) Vendor support	Support Fournisseurs	<ul style="list-style-type: none"> The developers of the big data system have provided adequate training to users. 	(Robey, 1979)
Qualité des données (QD) Data quality	Data quality and information quality	<ul style="list-style-type: none"> Our big data sources provide accurate data. Our big data providers have a very good reputation in the industry. We are satisfied with the quality of the data provided by our big data sources. 	(S. Shamim et al., 2019)
Qualité des données (QD) Data quality	Complétude des données Data completeness	<p>The Big data system used:</p> <ul style="list-style-type: none"> ___ provides a complete set of information. ___ produces comprehensive information. ___ provides all the information needed. 	(Ji-fan Ren et al., 2017)
Qualité des données (QD) Data quality	Actualité des données Data currency	<p>The Big data system used:</p> <ul style="list-style-type: none"> ___ provides the most recent information. ___ produces the most current information. ___ always provides up-to-date information. 	(Ji-fan Ren et al., 2017)
Qualité des données (QD) Data quality	Accès aux données Data access	<ul style="list-style-type: none"> Needed data would be easily retrievable. Needed data would be easily accessible. Needed data would be easily obtainable. The data would be quickly accessible when needed. 	(Demoulin & Coussement, 2020)
Qualité des données (QD) Data quality	Pertinence des données Data relevance	<ul style="list-style-type: none"> Information is informative for my needs. Information is valuable for my needs. 	(Zheng et al., 2013)
Qualité des données (QD) Data quality	Exactitude des données Data accuracy	<p>The Big data system used:</p> <ul style="list-style-type: none"> ___ produces correct information. ___ provides few errors in the information. ___ provides accurate information. 	(Ji-fan Ren et al., 2017)

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Qualité des données (QD) Data quality	Cohérence des données Data consistency	<ul style="list-style-type: none"> The Big data system assures the consistency of data to users (or applications). 	(Wixom & Watson, 2001)
Aspect Humain (AH) Human Aspect	Savoirs et compétences des individus People's knowledge and skills	<ul style="list-style-type: none"> Our 'big data system' related staff has the right skills to accomplish their jobs successfully. Our 'big data system' related staff is well trained. The company provides big data system operation's related training to its own employees. Our 'big data system' related staff has suitable education to fulfill their jobs. 	(Mikalef, Boura, Lekakos, & Krogstie, 2019 a)
Aspect Humain (AH) Human Aspect	Confiance Trust	<ul style="list-style-type: none"> I believe in the benefits of the big data system. My peers believe in the benefits of the big data system. My management team believes in the project benefits related the big data system. 	(Amoako-Gyampah & Salam, 2004)
Aspect Humain (AH) Human Aspect	Champions Champions	<ul style="list-style-type: none"> A high-level champion(s) for the big data system came from IS. A high-level champion(s) for the big data system came from a functional area(s). 	(Wixom & Watson, 2001)
Aspect Humain (AH) Human Aspect	Engagement des salariés Employee engagement	<ul style="list-style-type: none"> The big data system had support from people throughout the organization. 	(Wixom & Watson, 2001)
Aspect Humain (AH) Human Aspect	Participation des utilisateurs User participation	<ul style="list-style-type: none"> IS and users worked together as a team on the big data system- project. Users were assigned full-time to parts of the big data system- project. Users performed hands-on activities during the big data system- project. 	(Wixom & Watson, 2001)
Confidentialité, sécurité et gouvernance des données (CSG)	Confidentialité et sécurité Data privacy and security	<ul style="list-style-type: none"> The Big data system protects information against unauthorized access. The Big data protects the personal information. We feel safe to deal with the big data system. 	(Zheng et al., 2013)
Confidentialité, sécurité et gouvernance des données (CSG)	Gouvernance des données Data governance	<ul style="list-style-type: none"> In our organization, we have controlled practices regarding data management /e.g., in terms of setting retention policies (e.g. time to live) of data/ backup routines/ establishing/monitoring access (e.g. user access) to data/ classifying data according to value/ monitoring costs versus value of data/. 	(Mikalef et al., 2019 a)
Intérêt organisationnel perçu (IOP)	Valeur perçue Perceived value	<p>The use of big data :</p> <ul style="list-style-type: none"> ___ creates competitive advantage. ___ aligns analytics with business strategy. ___ establishes useful links with other organisations. ___ enables quicker response to change. ___ improves customer relations. ___ provides better products or services to customers. 	(Ji-fan Ren et al., 2017)

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Intérêt organisationnel perçu (IOP)	Risques perçus Perceived risk	The risks related the uses of big data are: <ul style="list-style-type: none"> • ___reluctance of employees to adapt to changes. • ___lack of Information System infrastructure support. • ___technical uncertainty. • ___minimal IT expertise. • ___uncertainty about how to measure potential benefits. • ___uncertainty about how to measure the involved costs • ___capital outlay with no guarantee of likely returns. • ___security issues. • ___reassignment of personnel trained on big data analytics solutions. • ___privacy issues. 	(Elisabetta Raguseo, 2018)
Intérêt organisationnel perçu (IOP)	Efficiency perçue Perceived usefulness	<ul style="list-style-type: none"> • Using the big data system would improve the performance in the job's outcome. • Using the big data system would increase the productivity. • Using the big data system would enhance the effectiveness in my job. • I would find the big data system to be useful in my job. 	(Demoulin & Coussement, 2020)
Intérêt organisationnel perçu (IOP)	Facilité perçue de l'utilisation Perceived ease of use	<ul style="list-style-type: none"> • My interaction with the big data system would be clear and understandable. • I would find the big data system easy to use. • I would find it easy to get the big data system to do what I want it to do. 	(Demoulin & Coussement, 2020)
Intérêt organisationnel perçu (IOP)	Intention d'utilisation Intention to use	<ul style="list-style-type: none"> • I want to use the big data system in the future. • I expect that I will use the big data system later. • I plan to use the big data system in the future. 	(Venkatesh et al., 2003)
Intérêt organisationnel perçu (IOP)	Observabilité perçue Perceived observability	<ul style="list-style-type: none"> • Many competitors or business partners in the market have started using Big Data Systems. • Using Big Data System helps my company to connect with both domestic and international business partners. • There are many computers that people in the company can access to Big Data System. • There are many computers that people in the company can access to use Big Data System. • Big Data Systems shows improved results over doing business the traditional way. 	(Maroufkhani, Tseng, et al., 2020)
Intérêt organisationnel perçu (IOP)	Analyse cout-avantage Cost-benefit analysis	<ul style="list-style-type: none"> • The costs involved in the adoption of big data system would be far greater than the expected benefits. • The cost of maintaining big data system would be very high for our enterprises. • The cost involved in providing support systems for big data system would be too high. • The amount of money invested in staff training would be very high. 	(Maduku et al., 2016)

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Management des processus (MP)	Orientation vers processus Process orientation	<ul style="list-style-type: none"> • Our company is better than competitors in connecting (e.g., communication and information sharing) parties within a business process. • Our company is better than competitors in reducing cost within a business process. • Our company is better than competitors in bringing complex analytical methods to bear on a business process. • Our company is better than competitors in bringing detailed information into a business process. 	(Wamba et al., 2017)
Management des processus (MP)	Intégration du business processus TI IT business process integration	<ul style="list-style-type: none"> • We have the ability to integrate the processes involved in the big data chain (i.e. data collection, preparation, analysis and decision making). • The integration of the processes involved in the big data chain reduces the cost of big data use. • The integration of the processes involved in the big data chain reduces the efforts necessary to analyze big data. 	(S. Shamim et al., 2019)
Management des processus (MP)	Management des données Data management	<ul style="list-style-type: none"> • In my organization there is a good fit between the analytical tools we have access to and the data we process. • The present analytical tools my organization has access to fulfill our data analysis needs. • The analytical tools that my organization currently has access to provide pretty much everything that we need to analyze our data properly. 	(Ghasemaghaei et al., 2017)
Aspect organisationnel (AO) Organizational aspects	Compétence organisationnelle et culturelle	<ul style="list-style-type: none"> • Our decisions are based on data. • A dependency on hunches for decision making is strongly discouraged in our organization. • Depending on data is part of our organizational routine. • We have a culture of data driven work. 	(S. Shamim et al., 2019)
Aspect organisationnel (AO) Organizational aspects	Gestion du talent Talent management	<ul style="list-style-type: none"> • We prefer to hire employees who understand big data. • We have the ability to recruit expert users of big data. • We plan to enhance the big data management skills of our staff. • We take special care in the retention of big data experts in our organization. 	(S. Shamim et al., 2019)
Aspect organisationnel (AO) Organizational aspects	Programme de gestion du changement Change manage program	<ul style="list-style-type: none"> • Change in the organization created by Big Data System was managed effectively. 	(Wixom & Watson, 2001)
Aspect organisationnel (AO) Organizational aspects	Alignement stratégique Strategic alignment	<ul style="list-style-type: none"> • The use of big data aligns with the company's mission, goals, objectives, and strategies. • The use of big data contains quantified goals and objectives. • The use of big data contains detailed action plans/strategies that support company direction. • We prioritize major big data usage related investments by the expected impact on business performance. 	(Aker et al., 2016)

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Aspect organisationnel (AO) Organizational aspects	Management du projet Project management	<ul style="list-style-type: none"> • We continuously examine the innovative opportunities for the strategic use of big data analytics. • We enforce adequate plans for the introduction and utilization of big data analytics. • We perform big data project planning processes in systematic and formalized ways. • We frequently adjust big data project plans to better adapt to changing conditions. 	(Akter et al., 2016)
Aspect organisationnel (AO) Organizational aspects	Management de la performance Performance management	<p>We perform much better than our main competitors in terms of:</p> <ul style="list-style-type: none"> • Profitability • Profits as percentage of sales • Decreasing product or service delivery cycle time • In reducing operating costs • In profit growth rates • Rapid response to market demand • Rapid confirmation of customer orders • Increasing customer satisfaction • Providing better product and service quality • In reducing operating costs 	(Mikalef et al., 2019 a)
Aspect organisationnel (AO) Organizational aspects	Structure organisationnelle et taille Organizational structure and size	<ul style="list-style-type: none"> • We can cost effectively change our organizational structure to respond to changing business conditions. • We can change our organizational structure without negatively impacting service/product quality. • Our current organization structure enables us to adapt to changing business conditions. • Our organization is more flexible than our competitors in changing our organizational structure. 	(Srinivasan & Swink, 2018)
Aspect organisationnel (AO) Organizational aspects	Collaboration interdépartementale Interdepartmental collaboration	<ul style="list-style-type: none"> • There is collaboration among big data analysts and big data providers. • There is collaboration among big data analysts and decision makers. • There is collaboration among big data providers and decision makers. 	(S. Shamim et al., 2019)
Aspect organisationnel (AO) Organizational aspects	Communication Communication	<ul style="list-style-type: none"> • Our employees transfer their knowledge about data. • Knowledge about how data are collected is exchanged within the organization. • Knowledge about how data are collected is exchanged within the organization. • Knowledge about how data are processed is exchanged within our firm. • The exchange of knowledge makes it easy for us to analyze data. 	(S. Shamim et al., 2019)

Thème/Facteur	Facteur/Élément	Echelle de mesure	Source
Aspect organisationnel (AO) Organizational aspects	Support de haute direction Top management support	<ul style="list-style-type: none"> • The top management is greatly interested in using big data systems. • The top management is aware of the benefits of big data systems for future success. • The top management will allocate adequate financial and other resources to the development and operation of big data systems. • The top management has a vision to project our company as the market leader in the use of big data systems. 	(Demoulin & Coussement, 2020)
Aspect organisationnel (AO) Organizational aspects	Effet environnemental Environmental effect	<p>The use of big data in our organization is affected by:</p> <ul style="list-style-type: none"> • The extent of technological turbulence in the environment. • The leadership in product/process innovation. • The impact of new technology on operations. • The extent of market turbulence. • The frequent changes in customer preferences. • The need to reduce market uncertainty • The need to respond to market opportunities 	(Côte-Real et al., 2017)
Aspect organisationnel (AO) Organizational aspects	Effet environnemental Environmental effect	<p>The use of big data in our organization is affected by:</p> <ul style="list-style-type: none"> • The extent of technological turbulence in the environment. • The leadership in product/process innovation. • The impact of new technology on operations. • The extent of market turbulence. • The frequent changes in customer preferences. • The need to reduce market uncertainty • The need to respond to market opportunities 	(S. Shamim et al., 2019)
Aspect organisationnel (AO) Organizational aspects	Effet environnemental Environmental effect	<p>The use of big data in our organization is affected by:</p> <ul style="list-style-type: none"> • The extent of technological turbulence in the environment. • The leadership in product/process innovation. • The impact of new technology on operations. • The extent of market turbulence. • The frequent changes in customer preferences. • The need to reduce market uncertainty • The need to respond to market opportunities 	(G. Kim et al., 2011)

4.2.2.2.3 Étape 3 : Validation technique globale de l'instrument

Avant de passer à l'adaptation du questionnaire au contexte de recherche, nous avons besoin de nous rassurer sur la validité de l'instrument et sa capacité de traitement statistique. Cette étape a été approuvée par un expert d'analyses quantitatives. Après une mise en place de

quelques amendements mineurs sur l'instrument, conseillés par l'expert, nous sommes passés à son adaptation au contexte et à une étude-pilote.

4.2.2.2.4 Étape 4 : Adaptation au contexte et étude-pilote

Au total, les 8 professionnels du domaine et un académicien ont été mobilisés : les 2 professionnels sont des experts qui travaillent à QRS Co. Ces deux professionnels, ayant une expérience du domaine de plus de 10 ans nous ont aidés à nous rassurer sur les trois éléments importants suivants :

- Intelligibilité et compréhension des questions composantes de l'instrument ;
- Adaptation au contexte ;
- Fiabilité des construits.

Pour la fiabilité des construits, nous avons retenu un coefficient de cohérence interne de Cronbach's Alpha ≥ 0.9 (0.93) (Hair, Black, Babin, & Anderson, 2010) (Annexe 4).

La procédure de cette étude-pilote dans son ensemble nous a permis d'évaluer la validité apparente et de contenu des éléments de l'instrument développé. Ainsi, cela assure que, d'une part, nos répondants seraient en place pour comprendre l'enquête comme prévu ; d'autre part, les propriétés statistiques sont valides tout en examinant leur exhaustivité.

4.2.2.2.5 Étape 5 : Construction du questionnaire final

Suite aux remarques et commentaires des participants en étude-pilote et des experts du domaine, le questionnaire a été revu, traduit en anglais (une langue de travail des professionnels concernés en Arménie) et préparé pour l'étude principale. L'instrument final est détaillé en annexe 3.

4.2.2.2.6 Étape 6 : Choix d'échantillon, distribution du questionnaire et recueil des données

L'auditoire ciblé (défini comme une population dans les manuels de statistique) regroupe des professionnels relatifs au traitement et usage du big data (les Data Scientists, Data Analystes, Data Engineer, Chef du projet du traitement big data) au sein des entreprises en Arménie, exploitant le big data.

Pour construire notre échantillon, deux directives principales nous ont guidés :

- Directives de Bryant et Yarnold (1995), qui notent que l'échantillon doit être 5 fois plus grand que le nombre des variables (dans notre cas, $41 \text{ facteurs} * 5 = 205$) ;
- Directives de Heckler et Hatcher (1996) et par « la loi de 100 », qui proposent que la taille minimale d'un échantillon ne doive pas être moins de 100 et encore non moins que 5 fois le nombre des variables (Hair et al., 2010).

La nature de notre recherche nous impose d'adapter un échantillonnage non probabiliste. M. Bidan (2020) décrit la méthode de l'échantillonnage non probabiliste ou empirique comme la construction d'un échantillon préalablement ciblé dit « de convenance » (*auprès d'étudiants, de collègues, d'experts, etc.*) ou encore de retenir des individus qui auront des caractéristiques précises (*géographie, CSP, revenu, taille, âge, etc.*) qui constitueront autant de quotas pensés comme incontournables pour faire partie de l'échantillon ciblé.

Ainsi, notre échantillon est composé de 219 professionnels en big data qui travaillent dans les 16 entreprises arméniennes qui exploitent potentiellement le big data dans leur activité. Parmi ces entreprises, 19% sont de petites entreprises (moins de 50 salariés), 33% sont de taille moyenne (50-249 salariés) et 48% sont de grandes entreprises (plus de 250 salariés), selon la définition européenne (Elisabetta Raguseo, 2018). La majorité des répondants (41%) possède une expérience dans le domaine de 3-5 ans, 18% ont une expérience de 6-10 ans. La forte majorité des répondants ont un âge compris entre 26 et 35 ans (76%) et sont des hommes (61%) (Annexe 5).

La phase de cette étape est réalisée par une mobilisation des contacts personnels et professionnels. Une pré-phase introductive a été organisée pour chaque contact-entreprise afin de rassurer de son éligibilité et son intérêt pour notre étude. Une brève présentation introductive de 5 minutes environ sur le contexte de recherche et les concepts-clés a été réalisée pour chaque contact-entreprise (N=16). Une description complète sur chacune de ces 16 entreprises est détaillée dans le chapitre suivant.

Ainsi, nous avons offert 2 possibilités pour remplir le questionnaire, pour une meilleure accessibilité des répondants : Un lien sur l'outil « Google Survey » ou un questionnaire papier, imprimé et distribué par le contact-entreprise.

La période de recueil des données pris environ 9 mois (Jan-2021-Sépt-2021) ; sa phase préparatoire à peu près de 3 mois, donc au total presque 12 mois.

4.2.2.2.7 Étape 7 : Administration, préparation des données et leur traitement

4.2.2.2.7.1 Administration et préparation des données

Au total, nous avons obtenu 219 questionnaires remplis, dont 7 étaient incomplets et l'un ne répondait pas aux critères d'inclusion, laissant 211 questionnaires exploitables après une phase de nettoyage des données.

Pour administrer et préparer les données (Annexe 14) à leur traitement principal, nous avons suivi une procédure en 2 étapes :

- ✓ Structuration des données de manière à ce qu'elles correspondent à notre question de recherche ($EUBD^{13} = w_1X_1 + w_2X_2 + \dots + w_{41}X_{41}$) (Hair et al., 2010) ; qu'elles soient cohérentes avec le plan d'analyse et mises en forme pour le traitement (une ligne pour chaque cas unique et une colonne pour chaque variable) (Annexe 6);
- ✓ Nettoyage des données, où nous avons éliminé les cas non-valides pour l'analyse (N=7), et rejeté des cas avec les données peu fiables (N=1) (Annexe 7).

4.2.2.2.7.2 Traitement des données

Une fois les données administrées, nettoyées et préparées à l'analyse, nous avons procédé à l'étape principale du traitement statistique des données (Annexe 10). Pour ce faire, nous avons utilisé les logiciels Excel avec sa fonction de « Data Analysis » et XLSTAT. Pour pouvoir exploiter nos données par l'application des fonctionnalités de ces logiciels, nous nous sommes basés sur les directives méthodologiques de Yesayan (2016).

Pour le traitement des données, nous avons adopté les principales directives d'analyse des données quantitatives de Raguseo (2018). Ces directives s'appuient principalement sur la logique de classement et de hiérarchisation, basée sur un ordre des variables selon leurs valeurs moyennes. Ce choix de méthode d'analyse nous semble le plus pertinent à appliquer à nos données récoltées, pour 2 raisons principales :

- La nature de notre question de recherche impose de proposer une certaine liste des facteurs de 41 composants du modèle par leur impact le plus fort à contribuer à l'efficacité d'usage du big data ;
- La simplicité qui rend les résultats plus intelligibles.

13 EUBD-efficacité d'usage du big data

Dans les lignes qui suivent, nous détaillerons les directives de traitement, adapté à la nature de notre question de recherche.

4.2.2.2.7.2.1 Première étape de traitement

La première étape consiste à nous assurer de la fiabilité d'une comparaison par les moyennes des variables. Pour ce faire, le coefficient de variation (CV) (le rapport de l'écart-type à la moyenne) a été calculé pour chaque facteur ($CV = \sigma/\mu$ ou $CV = s/x$) ; l'annexe 8 montre que pour les données étudiées, ce coefficient est dans une zone fiable ($CV < 15\%$). Ainsi, ces résultats nous confirment que la comparaison des moyennes peut être un choix fiable, et que nous pouvons considérer les moyennes comme représentatives.

4.2.2.2.7.2.2 Deuxième étape de traitement

Par cette deuxième étape, nous allons recenser un classement hiérarchisé d'une liste des variables qui permettront de bien répondre à notre question de recherche. Au préalable, nous nous sommes intéressés à l'existence de différences des résultats entre certains groupes des répondants homogènes. Ce sont les réponses des répondants aux questions démographiques du questionnaire qui nous permettent de le faire (Annexe 3).

Les classements ont été calculés selon la taille de l'entreprise, le secteur industriel d'entreprises et leur relation avec les big data. Un test d'analyse de variance (ANOVA) nous a permis de constater une différence statistique significative pour les 2 groupes des répondants (Annexe 9, où $SSBW < SSWG$, avec une valeur $p < 0.05$). Cette différence correspond à la qualité de la relation de l'entreprise du répondant avec l'exploitation du big data ; la question concernée dans la partie de l'information démographique de l'instrument propose 2 choix de réponses : « l'entreprise du répondant a adopté une infrastructure nécessaire pour l'exploitation du big data » ou « l'entreprise du répondant est basée initialement sur l'idée de l'exploitation du big data ». Ce résultat nous renvoie plus tard à proposer des noms qualifiant ces 2 groupes : « utilisateur novice » et « utilisateur avancé ».

Ainsi, nous avons constaté cette différence dans les 2 groupes, que nous proposons d'appeler « les utilisateurs novices » et « les utilisateurs avancés »¹⁴ dans le reste de ce document. Ces 2 titres concernent bien les entreprises.

Sous le titre « utilisateur avancé », nous désignons toute entreprise qui s'est fondée sur l'idée de l'exploitation du big data, qui en possède une forte expertise et son activité quotidienne y est liée particulièrement.

Et sous le titre « utilisateur novice », nous désignons toute entreprise qui a commencé à exploiter le big data au cours de son activité principale « traditionnelle ».

4.2.2.2.7.2.3 Troisième étape de traitement

A ce stade de développement, nous avons établi 2 listes correspondant à chaque type d'utilisateur. Cette approche permet une appréhension plus facile de la réalité.

Dans les facteurs du modèle conceptuel, ont été classés et hiérarchisés par un ordre selon leurs valeurs moyennes, où le critère important d'inclusion dans la liste est considéré la valeur critique de moyenne égale ou plus (\geq) 5.0 (point critique dans les choix de réponse).

Les résultats de cette étude quantitative seront détaillés et interprétés dans les deux derniers chapitres du travail.

4.2.2.3 Troisième étude : Étude qualitative confirmatoire

Cette dernière approche méthodologique a trois objectifs principaux :

- Fournir une illustration des résultats obtenus par l'enquête quantitative (qui sert également comme une sorte de confirmation des résultats) ;
- Mettre en évidence les 2 types différents d'entreprises à partir de 2 études de cas bien précis (un « utilisateur avancé » ; société CDE et un « utilisateur novice », société BCD), pour comprendre comment les facteurs qui influencent le plus l'efficacité d'usage du big data émergent concrètement dans chacune de 2 types d'entreprises ;

¹⁴ Dans le dernier chapitre, concerné à la discussion, nous détaillons bien les caractéristiques de ces 2 groupes d'utilisateurs en y proposant les 2 définitions.

- Comprendre la place du « contexte arménien » dans les résultats d'enquête quantitative (Société basée aux Etats-Unis, QRS Co.).

Pour ce faire, nous avons conduit au total 5 entretiens dans les 3 sociétés. Le tableau 10 détaille les profils des répondants.

Entretiens confirmatoires		N=5
Entreprise	<i>QRS Co.</i>	
Qualité	<i>Senior Data Scientiste, expérience 10 ans+, R.R.</i>	
Date	Mode	Durée
<i>18/11/2021</i>	<i>Via Zoom</i>	<i>1h</i>
Entreprise	<i>BCD / « Utilisateur novice », C.C.</i>	
Qualité	<i>Responsable du département : Data Management</i>	
Date	Mode	Durée
<i>15/12/2021</i>	<i>Via Zoom</i>	<i>1h 5m</i>
<i>10/01/2022</i>	<i>Via Zoom</i>	<i>1h</i>
Entreprise	<i>CDE / « Utilisateur avancé », D.D.</i>	
Qualité	<i>Senior Data Scientiste, Membre des Fondateurs</i>	
Date	Mode	Durée
<i>03/01/2022</i>	<i>Via Zoom</i>	<i>1h</i>
<i>31/01/2022</i>	<i>Via Zoom</i>	<i>1h 10m</i>
Total	5h 15m	

Tableau 10 Détail des entretiens confirmatoires

Après avoir retenu des résultats de l'étude quantitative, nous nous sommes fixés un objectif de nous rassurer que les résultats différents que nous avons pour les 2 groupes distingués ont une logique abstraite et ne dépend pas du contexte étudié.

Pour cela, nous avons conduit un entretien semi-directif (le guide d'entretien est détaillé dans l'annexe 18) avec un des 2 experts du groupe QRS Co., qui a participé à l'étude-pilote.

La langue de l'entretien était l'anglais.

Après, une fois assurés dans l'indépendance du rôle de contexte aux groupes des résultats, nous avons conduit 2 entretiens semi-directifs (guide d'entretien adapté de (Mikalef, Boura, Lekakos, & Krogstie, 2019 b), détaillées en annexe 19 avec un représentant de chaque groupe d'entreprises. La langue des entretiens était l'arménien.

Nous avons appliqué la même tactique d'analyse du contenu que dans la première étude qualitative exploratoire (Bardin, 2013), (Philippe Wanlin, 2007) détaillée au préalable. L'analyse a été guidée par les résultats obtenus de l'étude quantitative.

Au total, les 5h 15m de discours ont permis de décrire un cas bien précis d'efficacité d'usage de chaque groupe d'entreprises en illustrant les rôles des facteurs qui influencent le plus dans les activités détaillées.

L'ensemble des résultats issus de ces trois études présentées nous permet de répondre complètement à la question de recherche du travail. Une présentation détaillée et une discussion sur les résultats seront présentées dans le dernier chapitre de la thèse.

Chapitre 5 Terrain de recherche

Avant de passer à une présentation détaillée du terrain de recherche, nous allons d'abord traiter les quatre éléments importants de contextualisation de la recherche. Il s'agit donc de créer d'abord une image claire sur le contexte d'examen avant de passer à la présentation du terrain.

Ainsi, ce chapitre se construit autour de 4 axes principaux suivants :

- Intérêt de la contextualisation de la recherche en science de gestion ;
- Intérêt du profil de l'Arménie au regard d'une contextualisation à 3 pôles : culturel, institutionnel et post-soviétique et de l'aspect technologique du contexte d'examen ;
- Présentation détaillée du terrain de la recherche.

5.1 Pourquoi faut-il contextualiser une recherche en science de gestion ?

Encore aujourd'hui, les spécificités historiques, culturelles, éthiques, politiques, sociales, humaines sont souvent considérées par les chercheurs *comme des contraintes à dépasser par la recherche. Une fois ces spécificités dépassées, la recherche peut être « normalisée » et se soumettre au canevas méthodologique* (Buchanan & Bryman, 2007).

Pourtant, de plus en plus la recherche en management peine à se réfugier dans sa confortable tour d'ivoire épistémologique et méthodologique, dans son « *business as usual* », qui privilégient les choses bien définies et stables, les relations causales élémentaires et simples, univoques et déterministes, l'élaboration de lois générales réputées valides hors du temps, de l'espace et des contextes (Martinet, 2015).

Nous sommes d'accord dans la vision de Louitri et Sahraoui (2014), qui notent qu'une idée comme « toute chose est égale par ailleurs » peut biaiser les résultats des recherches ; pourtant son apport managérial majeur consiste dans son actionnabilité pour le contexte étudié.

Ainsi, la contextualisation d'abord permet aux recherches en management d'intégrer le contexte et ses spécificités, de la définition de la problématique jusqu'à la formulation des conclusions et recommandations managériales (Doha et Abdenbi, 2020).

Donc, de notre point de vue, la prise en compte des spécificités contextuelles est une tâche nécessaire dans toute étude en science de gestion comme dans les sciences sociales où la production des connaissances contextualisées est souvent liée aux spécificités du contexte d'étude.

Et comme noté par Charreire Petit et Durieux (2014), le chercheur ne traite pas avec les lois ou théories universelles, mais il élabore ou teste des théories qui sont généralement qualifiées de substantives. Ici, une distinction proposée par Glaser et Strauss (1967) admet que la théorie substantive est un développement théorique en relation directe avec un domaine empirique. Pourtant une théorie formelle (universelle) concerne un domaine conceptuel. Et donc l'élaboration d'une théorie universelle se fait par l'intégration d'une multitude de théories substantives issues de différents contextes empiriques.

Dans ce travail de recherche, nous proposons une co-construction des connaissances à partir d'un dialogue horizontal. Cette co-construction exige de reconnaître l'importance de la contextualisation /dans une logique de l'épistémologie frontalière. Ce dernier porte l'idée que les *différentes formes de savoirs sont incomplètes et la considération de cette imperfection réciproque est une condition à l'obtention de la justice cognitive*. Ce courant regroupe des auteurs provenant de diverses disciplines, de l'anthropologie à l'histoire, parmi lesquels Enrique Dussel, Walter Mignolo, Santiago Castro-Gómez et Ramón Grosfoguel (Frimousse, 2020).

Nous suivons la proposition de D.S. Bentaleb et C. Bentaleb (2020) et présentons la démarche de contextualisation qui se subdivise en 2 parties : la contextualisation théorique et la contextualisation empirique.

5.1.1. Contextualisation théorique

La contextualisation au niveau théorique, prend en compte 2 éléments principaux : le choix du thème et la recherche bibliographique.

Le choix du thème est composé de 2 aspects importants : la réalité socio-économique et la contrainte légale du contexte de recherche que nous détaillons dans la section suivante.

Pour la recherche bibliographique, il faut qu'on détaille bien l'intérêt. Le pays dans lequel l'étude est réalisée (Arménie) est doté d'une recherche scientifique dont le niveau de

structuration et de maturité sont insuffisantes ; tandis que le niveau pratique (TI développée), ainsi que le potentiel de développement reste prometteur en affichant des avancées majeures et donc présente l'intérêt principal.

D'une part, ce qui semble représenter une difficulté est le manque de recherches scientifiques dans certains domaines et contextes, restreignant ainsi le pouvoir de comparaison et de généralisation. Cependant, cette difficulté est considérée également comme une des contributions de ce travail de recherche qui vise combler le manque de connaissances académiques et pratiques liées.

Ainsi, nous tentons de mobiliser les productions scientifiques de qualité sur le domaine, ainsi que toute documentation fiable qui porte sur le thème et le contexte étudié afin de positionner le sujet et le rapprocher au mieux des réalités contextuelles.

5.1.2. Contextualisation empirique

Le deuxième composant le plus important de la démarche de contextualisation est la contextualisation empirique.

Dans notre cas, l'objectif est de soumettre le modèle conceptuel de (Surbakti et al., 2019) aux réalités contextuelles liées au domaine de la recherche. Ici, il faut bien noter que le but de la contextualisation empirique n'est pas le rejet de l'universalisme, mais l'intérêt de mener une recherche actionnable, plus particulièrement de présenter les résultats qui intègrent le contexte arménien. Ces résultats tendent à apporter un éclairage nouveau aux spécificités organisationnelles dans un pays autre que celui à l'origine des théories managériales liées.

Les différentes phases d'étude contextuelle, ainsi que les instruments développés et adaptés sont détaillées dans le chapitre 4.

Ainsi, une recherche contextualisée nous permet de valider/rejeter ou supprimer la présence des variables étudiées et de préciser ou développer le sens de ces variables au regard du contexte étudié, ce qui est dans l'intérêt particulier de notre travail de recherche.

5.2. Profil de l'Arménie : Contextualisation à 3 pôles de Livian (2020)

Dans les paragraphes suivants, nous allons traiter la question suivante : « Quel est le profil de l'Arménie au regard d'une contextualisation à 3 pôles : culturel, institutionnel et post-soviétique pour un travail de recherche dans un domaine de la gestion de technologie de l'information ? »

Nous allons démarrer par une brève présentation de l'Arménie. Par la suite, nous examinons les 3 pôles de contextualisation proposé et adapté de Livian (2020) : «pôle culturaliste », « pôle institutionnaliste » et « pôle post-soviétique ». A la fin de cette partie, nous présentons un résumé sur l'environnement technologique arménien.

5.2.1. Arménie : une brève présentation

Le nom officiel est « République d'Arménie », la nature du système politique est une république parlementaire. Depuis le 8 mai 2018, réélu le 20 juin 2021, le chef du gouvernement est le Premier ministre Nikol Pachinian.

Données géographiques	
Superficie	29 800 km ²
Capitale	Erevan
Villes principales	Gumri, Vanadzor, Kapan
Langue officielle	arménien, (russe très répandu)
Monnaie	Drame
Fête nationale	21 septembre (Fête de l'indépendance)
Données démographiques	
Population (2019) :	2,958 millions d'habitants
Croissance démographique	0%
Espérance de vie	74,9 ans
Taux d'alphabétisation	99,4%
Religion(s)	Arméniens orthodoxes (94%)
Indice de développement humain (Classement ONU 2017)	0,74 (84ème rang)
Données économiques	
PIB (2018)	12,43 Mds USD
PIB par habitant (2018)	4,19 USD
Espérance de vie	74,9 ans
Taux de croissance (2018)	5,2%
Taux de chômage (2019)	17,7%
Exportations de la France vers l'Arménie	0, 41,5 M EUR en 2017
Importations françaises de l'Arménie	8,5 M EUR en 2017

Tableau 11 Profil du pays en quelques chiffres clés

Le tableau 11 présente quelques chiffres clés sur le profil du pays¹⁵.

5.2.2. Trois pôles de contextualisation de Livian (2020)

Dans les lignes qui suivent, nous allons traiter la question abordée pour cette section à partir des trois pôles de contextualisation conseillés par Livian (2020). Selon lui, le chercheur peut utiliser ces trois pôles pour situer sa recherche dans un ancrage contextuel. Dans son article l'auteur présente le cas africain et propose de discuter du pôle « culturaliste », du pôle « institutionnaliste » et du pôle « postcolonial » au regard de la contextualisation africaine. Nous suivons principalement les recommandations de cet auteur en appliquant quelques aménagements nécessaires pour les adapter au contexte arménien.

5.2.2.1. Le pôle « culturaliste »

Ici, la finalité consiste à reconnaître et prendre en compte les spécificités culturelles des réalités arméniennes (rôle de la famille, relation aux ancêtres, notion de hiérarchie, importance du prestige, générosité) qui peuvent potentiellement impacter le management arménien. Cela est une démarche indispensable à la compréhension du degré de compatibilité des méthodes de management occidentales avec les traditions locales.

5.2.2.1.1. Culture arménienne

La culture arménienne est séculaire, elle a été aussi bien influencée par sa diaspora de par le monde, que par la culture occidentale comme par son héritage post-soviétique.

La vie familiale, la religion et les traditions sont des éléments importants qui ont forgé et contribué à la conservation de l'identité et de l'héritage arménien.

L'Arménie est le premier État à avoir officiellement adopté le christianisme, en 301.

Au fur et à mesure que l'alphabet arménien est créé en 405 par Mésrop Mashtots, une floraison de littérature religieuse chrétienne apparaît. Les érudits religieux arméniens ont fait un apport très important au monde chrétien.

¹⁵ <https://armstat.am/en/>; <https://am.ambafrance.org/>

Tout au long de l'histoire de l'Arménie, les émigrés (actuellement la diaspora arménienne qui compte plus de 10 millions de personnes et c'est presque trois fois plus que la population arménienne en Arménie) ont emporté leur foi avec eux, ce qui a servi à maintenir leur identité. Même quand l'Arménie a été envahie et a perdu son indépendance plusieurs fois lors de son histoire, l'Église Arménienne a survécu en assumant un rôle de guide et a contribué à sauver l'identité arménienne.

5.2.2.1.2. Rôle de la famille

L'autre élément clé de la culture, c'est le rôle de la famille. Un aspect important de la culture à noter ici, c'est le paternalisme dominant dans les familles arméniennes. Et là aussi, la foi et l'Église ont un impact important. Pour fonder une famille on n'accepte que le mariage officiel qui se fait en trois étapes : un acte officiel à la mairie, un acte religieux à l'Église (où le prêtre couronne les mariés, qui symbolise la création d'un nouveau « royaume-foyer » en tant que famille) et un acte non-officiel, mais obligatoire – une fête avec un grand nombre d'amis et de membres de la famille. Ces rituels religieux et traditionnels se mettent en place avec un espoir que le couple marié soit béni d'avoir des enfants. La naissance d'enfants dans la famille est célébrée aussi avec une grande cérémonie, cet amour et la grande attention aux enfants à l'enfance est la conséquence d'une structure familiale élargie aux beaux-parents, vivant souvent ensemble. Le respect des aïeux est à ce point important que le prénom de l'enfant aîné de famille identifie souvent un ancêtre.

Un autre aspect culturel, est la communauté : l'importance de l'avis, des valeurs de la communauté et son impact sur les vies des gens. Il faut noter que les arméniens, surtout la génération des aînés donnent une importance majeure à l'avis de la communauté sur différents aspects de la vie. Cela peut donc influencer leur choix sur un large spectre d'évènements.

Ainsi, nous voyons que ce premier pôle, l'approche culturaliste peut donner une vision assez large sur le contexte, mais les conclusions tirées de ces analyses sont trop générales ; pour pouvoir y recenser des spécificités organisationnelles et managériales, il faut mettre en place une confrontation des modèles. Par exemple un modèle arménien vs un modèle occidental ; *afin de pouvoir survivre à l'ère de la mondialisation de la concurrence et la globalisation des économies, une expérience multiculturelle est indispensable* (Mutabazi, 2006). Ainsi, la finalité

est de trouver un modèle organisateur-circulatoire commun et répandu en Arménie ; qui peut être étudié dans des cas spécifiques liés au contexte. Cette comparaison n'entre pas dans l'objectif de notre travail de thèse, mais elle peut être un sujet bien intéressant pour les futures recherches.

L'aspect culturel contribue également à situer la position du contexte étudié sur un vecteur des traditions à la modernité. Ce mouvement, *apparu à la fin des années 1990*, dit un *courant de la « crossvergence »* : *ce néologisme sert à distancier à la fois de ceux qui pensent à une « convergence » des modes de management dans le contexte de la mondialisation, et de ceux qui, au contraire, insistent essentiellement sur les « divergences » culturelles et sociétales* (Livian, 2020). La théorie de la « crossvergence » a été introduite pour la première fois en 1993 dans l'un des articles de David A Ralston et ses coauteurs en JIBS (Journal of International Business Studies)¹⁶. Ce courant propose plutôt une combinaison d'une part des spécificités et des forces socio-culturelles et d'autre part la logique du business international. Ainsi, comprendre les prédictors des valeurs et des comportements au niveau individuel qui sont de leur côté influencés par les éléments culturels est une des premières étapes pertinentes à pouvoir agir dans l'arène du business international (Ralston, 2008).

Ainsi une contextualisation culturelle englobe les caractéristiques des éléments-composants détaillés qui peuvent être prises en compte afin de comprendre les situations managériales.

5.2.2.2. Pôle « institutionnaliste »

Pour présenter ce deuxième pôle-composant de la contextualisation proposée par Livian, l'on peut toutefois s'appuyer sur les perspectives institutionnelles de la théorie de l'organisation : *les différents aspects institutionnels (la légitimité, les logiques institutionnelles, l'agence institutionnelle...) présentant une ressource importante qui potentiellement impacte l'action organisationnelle* (Greenwood, Oliver, Lawrence, & Meyer, 2018) ; (Whitley, 1999). Selon cette théorie, chaque pays a un certain « profil institutionnel », influencé par la structuration locale gouvernementale, administrative, législative et bien d'autres structures y fonctionnant.

¹⁶ «Differences in Managerial values: A study of US, Hong Kong and PRC managers'' (Ralston, Gustafson, Cheung, & Terpstra, 1993)

Ainsi pour détailler le contexte institutionnel arménien, nous allons nous appuyer en particulier sur les éléments suivants :

5.2.2.2.1. Structuration des marchés, profil économique

Après l'indépendance en 1991, l'Arménie a connu une croissance économique solide qui l'a propulsé parmi le classement des pays à revenu intermédiaire (PIB par habitant de 4 595 USD en 2021). L'effondrement de l'Union soviétique et la planification centralisée de l'économie ont entraîné une contraction initiale du PIB au début des années 1990. Cependant, le PIB par habitant (mesuré en dollars en 2001 à parité de pouvoir d'achat \$) a plus que quadruplé, passant de 2000 \$ en 1994 à 8800 \$ en 2017. Cette croissance s'est produite pendant les années de boom économique (début des années 2000). Après la crise financière mondiale, l'Arménie a connu quelques années de stagnation, bien que la croissance ait repris en 2017 et 2018. A l'issue de cette croissance, l'Arménie a subi une transformation structurelle importante d'une économie agricole à une productivité faible vers une économie dotée d'un secteur de services en expansion. (Une composition sectorielle est détaillée en Figure 39).¹⁷

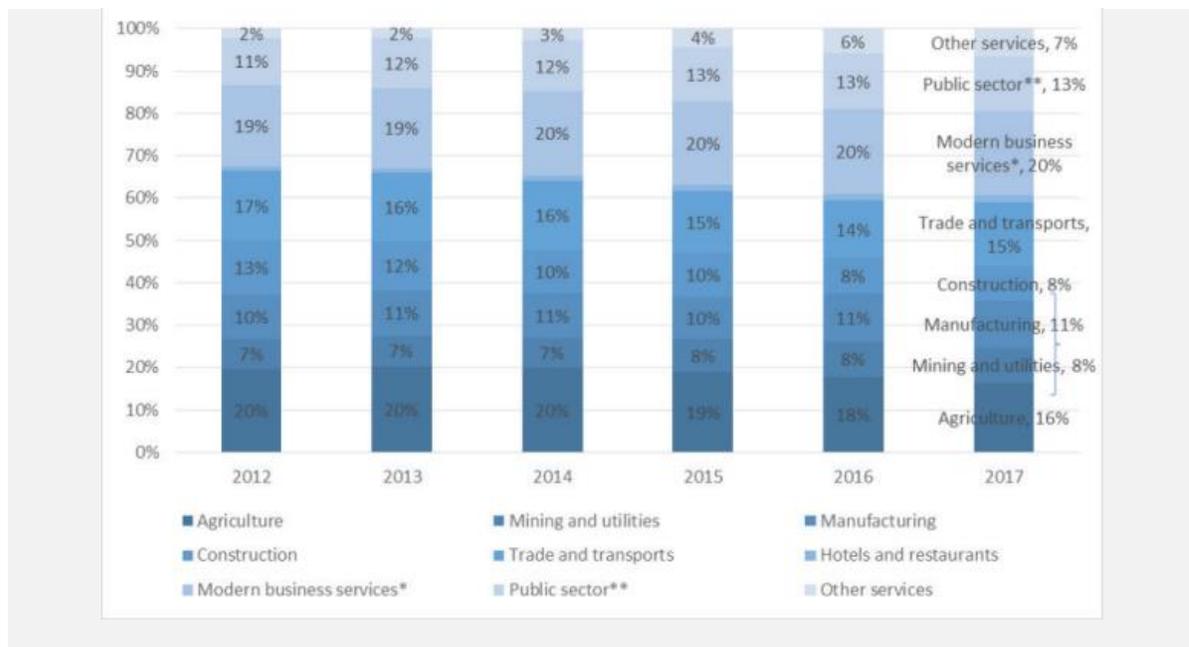
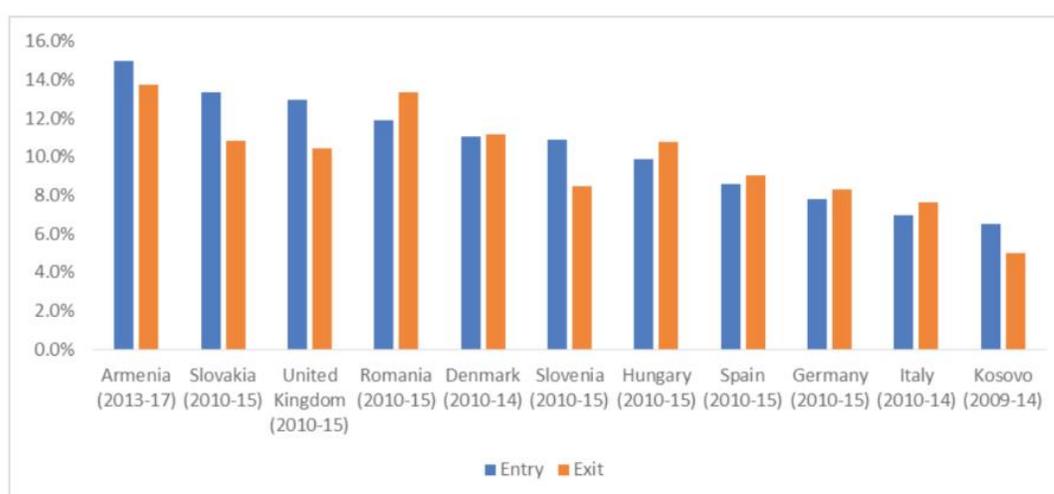


Figure 39 Composition sectorielle détaillée, 2012-17

¹⁷ <https://armstat.am/en/>; <https://www.mineconomy.am/en>

La pandémie de Covid-19 et la guerre du Haut-Karabakh ont durement touché l'économie du pays, qui a enregistré une lourde récession en 2020-2021. L'Arménie tente de mobiliser les ressources financières et de suivre un plan de rattrapage économique.

Un projet de recherche sur l'environnement du business en Arménie, réalisée par la Banque Mondiale¹⁸, montre qu'au cours de la période de 2013-2017, en moyenne 14 % des entreprises arméniennes en activité ont quitté le marché, ce qui correspond à peu près à la sortie de 14 600 entreprises chaque année. Le taux de sortie des entreprises arméniennes reste élevé au cours de cette période (Figure 40). Les barrières à l'entrée ne sont peut-être pas particulièrement élevées, mais la probabilité de survie des micro-entreprises est faible. A cet égard, une analyse plus approfondie est nécessaire pour déterminer dans quelle mesure les pressions concurrentielles impactent les taux de sortie. L'on n'écartera pas non plus comme cause le climat dysfonctionnel des affaires dans le pays.



Source: For Armenia: SRC firm-level data for 2013-17; for other countries: OECD Business Demography Indicators.

Figure 40 Taux d'entrée et de sortie des entreprises, en comparaison avec d'autres pays

Sous l'ancien système soviétique de planification centrale, l'Arménie avait développé un secteur industriel moderne pour l'époque, fournissant des machines-outils, textiles, et autres produits manufacturés aux républiques-sœurs en échange de matières premières et d'énergie. Depuis l'effondrement de l'URSS en décembre 1991, l'Arménie s'est convertie à l'agriculture de

¹⁸ <https://documents1.worldbank.org/curated/en/387401564380250230/pdf/Work-for-a-Better-Future-in-Armenia-An-Analysis-of-Jobs-Dynamics.pdf>

petite taille, loin de grands complexes agro-industriels de l'époque soviétique. Le secteur agricole a des besoins à long terme en investissements et modernisation de sa technologie. La privatisation de l'industrie s'est réalisée à un rythme plus lent, mais a reçu un plus grand soutien de la part de l'Etat. L'Arménie est importatrice en produits alimentaires, et ses dépôts de minerai (or, bauxite) sont de peu d'importance. Le conflit permanent avec l'Azerbaïdjan à propos de la région sous contrôle arménien du Haut-Karabakh et la dissolution du système économique, d'économie dirigiste de l'ancienne Union soviétique ont contribué au grave déclin économique au début des années 90. Cependant, le gouvernement a lancé en 1994 un programme économique ambitieux soutenu par le Fonds Monétaire International, qui a eu comme conséquence des taux de croissance positifs en 1995-1998. L'Arménie est également parvenue à réduire l'inflation et à privatiser la plupart des petites et moyennes entreprises. Le déficit chronique en énergie de ces dernières années a été en grande partie compensé par l'apport de la centrale nucléaire de Metsamor. La crise financière russe a renouvelé les inquiétudes sur la performance économique de l'Arménie en 1998. L'inflation est tombée à 10% et le PIB a crû d'environ 6%, mais le secteur industriel est demeuré moribond. Une grande partie de la population arménienne reste fortement dépendante de l'aide et des transferts de l'étranger, et les aides venant de Russie ont chuté brutalement en 1998.¹⁹

5.2.2.2.2. Modes de gouvernance

La République d'Arménie est un État souverain, démocratique, social et de droit. La souveraineté nationale appartient au peuple. Son pouvoir est exercé par la voie d'élections libres et des référendums, ainsi que par la voix des autorités nationales et locales, prévus par la Constitution de la République de l'Arménie. Conformément à la Constitution et aux lois locales, le pouvoir de l'État s'exerce sur la base de la séparation et de l'équilibre des pouvoirs législatif, exécutif et judiciaire. Le chef de l'Etat est le Président de la République qui veille aussi au à la conformité de la Constitution et ces pouvoirs sont définis par celle-ci²⁰. Le Président est impartial, il est guidé exclusivement par les intérêts d'Etat et de la nation. Le Président de la République est élu par l'Assemblée nationale pour un mandat de sept ans.

¹⁹ <https://www.gov.am/en/>

²⁰ <https://www.president.am/en/>

Le gouvernement est l'organe suprême du pouvoir exécutif qui élabore et met en œuvre la politique extérieure et intérieure de l'Etat en accordance avec son programme, effectue la gestion générale des institutions d'administration publique.

La Constitution et les lois définissent les pouvoirs du Gouvernement. Il est composé du Premier ministre, des Vice-premiers ministres et des Ministres²¹. Le premier ministre est élu par la majorité parlementaire, il est nommé par le Président de la République.

Les principales lignes de la politique du Gouvernement sont désignées par le Premier-ministre, accordées avec le Programme général du Gouvernement.

La République d'Arménie est composée de 10 régions (appelé marz) ; ces derniers sont composés des communautés municipaux et ruraux où s'effectue la gouvernance locale.

5.2.2.2.3. Systèmes juridiques et financiers

Dans la République d'Arménie, la justice est administrée par les tribunaux: la Cour constitutionnelle qui garantit la suprématie de la Constitution, la Cour de cassation qui garantit l'application uniforme des lois et des actes juridiques normatifs en révisant les actes judiciaires dans le cadre de ses pouvoirs constitutionnels, ainsi qu'elle élimine les violations des droits et les libertés fondamentales de l'homme ; des Cours d'appel, des tribunaux de compétence générale et le tribunal administratif ; il est possible aussi la création des cours spéciales, si nécessaires. Le Conseil supérieur de la magistrature garantit l'indépendance des tribunaux et des juges. C'est un institut d'Etat indépendant composé de 10 membres : 5 membres du Conseil judiciaire suprême, élus par l'Assemblée générale des juges, et les autres 5 membres sont élus par l'Assemblée nationale.²²

Depuis 1993, la Banque Centrale privilégie une approche très libérale avec un haut niveau de désintermédiation : le système bancaire arménien est peu concentré ; ses indicateurs de stabilité financière sont plutôt satisfaisants. Le marché financier est composé de 17 banques (entre elles il y a aussi les représentations de banques étrangères) représentant 83,5% des actifs totaux du secteur financier : au début 2020, l'indice de HerfindalHirschmann qui mesure

²¹ <https://www.gov.am/en/structure/>

²² https://www.moj.am/en/staff/structure_of_the_ministry

concentration d'un marché donné (0-1, 0 correspondant à un niveau de concentration minimal) s'est calculé de 0,089.²³ Le système bancaire est bien capitalisé (l'exigence pour les fonds propres est de 12% ; en 2021 s'est consolidé à 17%) ; le taux des crédits non performants dans le portefeuille s'est approché à 3.4% en 2021 ; la dollarisation des banques reste élevée : les crédits (46%) et les dépôts (52%), respectivement en 2021. Dans son ensemble, le système bancaire constate une rentabilité faible : l'accord des fonds au secteur privé augmente très lentement ; peu de banques arméniennes sont en mesure d'accorder des crédits de plus de 20 000 dollars à moyen et long terme. Les banques privées proposent des crédits essentiellement à court terme (3-5 an au moyen), et à des taux excessivement élevés de 18%-24% annuels.²⁴

Plusieurs organisations financières internationales disposent de représentations en Arménie : la Banque Européenne de la Reconstruction et le Développement (BERD), la Banque Mondiale, les Fonds Monétaire International, la Fondation Eurasia. Ces organismes soutiennent le développement du secteur privé ; les projets sont principalement destinés aux PME locales.

5.2.2.2.4. Organisation des partenaires sociaux

Les organismes principaux de régulation du travail en Arménie sont gérés principalement par le Ministère du Travail et des Affaires sociales qui définit les orientations des branches du système de sécurité sociale. La protection sociale est gérée aussi par l'Agence pour l'emploi attaché au Ministère du Travail et des Affaires sociales, qui est responsable des paiements des indemnités de chômage et de la mise en œuvre de la politique de l'emploi. Le ministère de la Santé, qui définit les orientations du système de santé et veille à leur mise en œuvre et l'Agence attaché au Ministère est responsable du financement du système des soins de santé. L'Arménie est membre de l'Organisation Internationale du Travail dont l'action consiste en la promotion des programmes de développement du dialogue social, de la protection sociale, de la sécurité, la santé, l'égalité des genres, le refus de la discrimination en raison d'un état de santé et/ou d'un handicap au travail ainsi que l'élimination du travail des enfants notamment. La Confédération des Syndicats d'Arménie est un centre important de syndicats dans le pays ; elle est la

²³ <https://www.cba.am/am/SitePages/Default.aspx>

²⁴ <https://armstat.am/en/>

transformation de ce qui reste de la structure syndicale de l'ex Union soviétique. Les estimations de cet institut montrent que 290.000 personnes sont membres de syndicats en Arménie.²⁵.

5.2.2.2.5. Système de formation

L'Arménie a un système développé d'établissements d'enseignement et d'instituts de recherches et de développement. Ce système est dispensé par différents types d'établissements d'enseignement supérieurs : Universités, Instituts, Académies et Conservatoires. Actuellement en Arménie fonctionnent 26 établissements d'enseignement supérieur soutenu par l'Etat et 33 établissements privés. Ce réseau a permis le développement de nombreuses industries de technologie de pointe, expliquant la croissance substantielle de l'économie de l'Arménie ces dernières décennies. L'Académie Nationale des Sciences, fondée en 1943, a développé un réseau de 30 centres de recherche scientifique. La qualité des résultats en recherche fondamentale dans différents domaines scientifiques bénéficiait d'une notoriété internationale dans les années 1950-1990. Pourtant, la crise économique a sérieusement affecté le système de formation, réduisant sensiblement le financement des établissements de recherche scientifique. En 1990, 130 établissements employant 20.000 chercheurs existaient dans le pays. En 1995, ce nombre s'était réduit à 119 établissements avec un peu plus de 5000 employés. Le profond déclin de l'économie a diminué les besoins des industries en développement technologique. Ceci a brisé les relations entre les instituts de recherche et l'industrie qui avaient existé ces dernières années. La stabilisation de l'économie et en particulier le programme de privatisation à grande échelle des entreprises a donné aux centres de recherche scientifique l'occasion de reprendre une place importante grâce à la présence d'un marché demandeur d'innovation²⁶.

En 2000 a été créé l'Université française en Arménie (UFAR) en application de l'accord de coopération culturelle, scientifique et technique signé en 1995 entre les gouvernements français et arménien. Ses fondateurs, les Ministères de l'Éducation et de la Science et des Affaires étrangères de la République d'Arménie, le Ministère français des Affaires étrangères et du Développement international, et l'Ambassade de France en Arménie, en ont fait un fleuron de la coopération entre l'Arménie et la France, en manifestant l'importance qu'ils attachent à l'éducation et à la formation de jeunes Arméniens, au service de leur pays et des relations franco-

²⁵ <https://www.mlsa.am/>

²⁶ <https://escs.am/am>

arméniennes. L'UFAR aspire à devenir un leader de l'enseignement supérieur dans le Caucase et une université française de référence à l'étranger. En fournissant les outils nécessaires à la formation de dirigeants intègres, responsables et ouverts sur le monde, elle donne aux jeunes l'opportunité de devenir des spécialistes hautement qualifiés, aptes à mobiliser les connaissances et savoir-faire acquis, au profit du développement économique de l'Arménie et des relations entre l'Arménie, la France et l'Europe²⁷.

5.2.2.3. Troisième pôle

Le troisième pôle de contextualisation proposé par Livian (2020) est celui qui concerne le statut postcolonial pour le cas du management africain.

Afin de pouvoir présenter l'Arménie sous l'angle d'une logique proche, nous proposons ainsi la réappropriation de ce pôle post-colonial par un pôle « post-soviétique ».

Ainsi, dans les lignes qui suivent nous soulignerons les points importants du statut « post-soviétique » qui peuvent potentiellement constituer une base pour la contextualisation.

5.2.2.3.1. Annonce historique

La République socialiste soviétique d'Arménie (RSS d'Arménie) est créée après la soviétisation de l'Arménie le 29 novembre 1920. De 1828 jusqu'à la Révolution Russe en 1917, l'Arménie Orientale était intégrée à l'Empire Russe.

Après la Révolution russe d'Octobre, le gouvernement de Lénine reconnaît le droit à l'autodétermination des minorités de l'empire. Par la suite l'Arménie se déclare indépendante (1918). L'Arménie n'était pas capable de résister à sa soviétisation en 1920 car était affaiblie des conséquences du Génocide (1915) et des dévastations de l'Arménie Occidentale comme des pertes humaines subies lors de la guerre turque-arménienne (septembre-décembre 1920). Bien que l'Arménie fût dépourvue d'une réelle souveraineté, l'on conviendra que la période de l'adhésion à l'Union Soviétique était très importante pour la conservation de son identité. Et ici, nous sommes bien d'accord avec Yves Ternon (1983), qui écrit «*Quel que soit le regard que l'on porte sur l'Arménie soviétique, sans les bolcheviks, l'Arménie orientale aurait été anéantie ; bolchevisée, elle demeure un sanctuaire, la dépositaire du patrimoine et la gardienne de*

²⁷ <https://www.ufar.am/>

l'identité... qu'on les considère comme des libérateurs ou des occupants, les bolcheviks, eux, ne cherchaient pas à détruire l'Arménie, mais à la remodeler... ».

En 1920, les représentants de l'Arménie (République d'Arménie) et la Russie soviétique signe un traité de paix qui oblige l'Arménie à reconnaître les territoires de Nakhitchevan, de Zanguezour et de Haut-Karabakh « litigieux », étant donné que le sort de ces territoires pouvait être décidé au moyen d'un référendum populaire.²⁸

Après, le 29 novembre 1920 l'Armée Rouge avec les bolcheviks arméniens entre à Karvansara du côté de l'Azerbaïdjan qui était déjà soviétisé (avril 1920) et déclare l'Arménie comme « soviétique ».

5.2.2.3.2. Profil du pays en tant que membre de l'US (Union Soviétique)

Il est à noter qu'au cours des années soviétiques l'Arménie devient un centre industriel important et agricole. De grands progrès sont réalisés dans le domaine de la science. Les instituts scientifiques arméniens deviennent reconnus, ainsi que le système d'enseignement constate un développement de haut niveau. Au cours ces années ont été fondées l'Académie Nationale des sciences (1943), qui dirigeait à l'époque 30 centres de recherches, seule période où la qualité des recherches menées jouissait d'une reconnaissance et d'une notoriété internationales. Aujourd'hui, en raison des restrictions budgétaires, l'importance et la qualité des projets de recherches sont loin de ce qu'ils étaient auparavant. A l'époque essor dans plusieurs domaines a permis à l'Arménie de structurer un profil institutionnel assez fort parmi les autres républiques de l'URSS (Union des Républiques socialistes soviétiques). Notons également que l'Arménie ainsi que les autres républiques membres de l'Union étaient indépendants « théoriquement » : tous les domaines régaliens étaient gérés et « dictés » par le pouvoir et la volonté centralisatrice des bolcheviks. La culture était isolée de la langue et de l'église ; la russification dominait dans tous les domaines. Cette période fut marquée par de nombreuses exactions d'opposants et d'intellectuels arméniens. C'est à cette époque que le conflit du Haut Karabakh apparaît et qui n'est toujours pas résolu à ce jour. (Khatisyan, 1968), (Karapetyan, 2003).

²⁸ En avril 1920, après la soviétisation de l'Azerbaïdjan, l'Armée Rouge déclare ces 3 territoires arméniens dans l'ensemble de l'Azerbaïdjan.

5.2.2.3.3. Effondrement de l'US (Union Soviétique)

Après l'effondrement de l'Union Soviétique, l'Arménie retrouve son indépendance et l'autonomie (le 21 septembre, 1991). Les années qui suivent deviennent très dures surtout sur le plan économique. Cependant le pays retrouve sa liberté à se réapproprier des concepts, des modes de pensée, des valeurs culturelles remises en cause pendant presque 70 ans. Ce processus de revalorisation continue encore, plusieurs domaines sont encore en cours de reconstruction. Après l'indépendance relativement récente, depuis plus de 30 ans, l'Arménie continue de dépendre des anciennes puissances, surtout sur le plan politique et économique. De même, il y a une grande dépendance financière et politique des institutions internationales (la Banque Mondiale, l'ONU, la Banque Européenne pour la Reconstruction et le Développement (BERD), les Fonds Monétaire International, la Fondation Eurasia.)²⁹ .

Donc cette brève description du pôle post-soviétique que l'on a proposé nous semble pertinente pour tout membre de l'Union Soviétique ; qui porte encore les caractéristiques communes héritées de l'URSS.

Cette présentation des 3 pôles fait office d'un résumé approfondi du contexte de recherche. Leurs interconnexions et les transformations entre des pratiques issues du terroir local et d'autres dominés de l'extérieur renvoient à un processus d'hybridation et une idée de management hybride qui prend d'abord en compte l'importance des contextes historiques et culturels d'où émerge l'hybridité et, d'autre part, permet d'identifier les éléments qui changent et se transforment ainsi que ceux qui persistent lorsque les pratiques de gestion externes sont imposés (Yousfi, 2014).

5.2.3. Environnement technologique

L'aspect technologique du contexte arménien, malheureusement insuffisamment traité dans la littérature, porte un intérêt particulier pour notre sujet de recherche.

Il faut bien noter que le potentiel technologique et scientifique compose une des ressources importantes du développement économique de l'Arménie.

²⁹ <https://www.mfa.am/fr/history>

Dans les lignes qui suivent, nous allons détailler quelques aspects de ce domaine qui nous paraît important pour une vision plus claire sur le terrain de notre travail.

Actuellement, l'industrie technologique en Arménie est en plein essor : au cours de la dernière décennie, le secteur a enregistré une croissance significative et dynamique malgré des débuts laborieux.

Le nombre d'entreprises opérant activement dans l'espace de TI dépasse 1 000, employant plus de 20000 travailleurs et rapportant plus d'un milliard de dollars par an. Cette industrie connaît une croissance de 20 % par an, y compris en 2020³⁰ où la croissance fut régulière malgré les effets du COVID-19 et des hostilités régionales.

5.2.3.1. Histoire du secteur

L'Arménie a eu une longue histoire en tant qu'un centre technologique et scientifique au sein de l'Union Soviétique. Elle comptait plus d'ingénieurs et de scientifiques par habitant que toute autre république soviétique et, bien qu'elle soit la plus petite république de l'Union, elle était responsable de la production de 30 à 40 % des composants électroniques et informatiques des industries de la défense et de l'aérospatiale soviétiques. Plus de 7 000 personnes travaillaient dans le très secret Institut de recherche et de développement informatique d'Erevan (capitale)³¹. Une brève analyse des brevets d'inventions déposés par des scientifiques et des ingénieurs d'Arménie dans la période de 1980 à 1990 met en évidence l'importance de l'impact des centres de recherches sur des technologies industrielles : dans cette période 4263 brevets ont été déposés, soit 22 brevets par an pour 1000 chercheurs. Cependant, une grande partie de cette culture et de ces connaissances ont été perdue au cours des années chaotiques suivant l'indépendance de l'Arménie (1991), surtout en raison de l'émigration massive et du manque de biens de première nécessité dans le pays.³²

Ce n'est qu'il y a une dizaine d'années que l'industrie a recommencé à se remettre sur pieds. En 2008, le gouvernement a annoncé un plan de route de 10 ans pour créer un écosystème technologique doté des infrastructures, des capitaux et des établissements d'enseignement

³⁰ <https://armstat.am/en/>

³¹ <http://www.encyclopedia.am/pages.php?hId=1341>

³² Encyclopédie « Arménie Soviétique », 1987

nécessaires qui permettraient au domaine de devenir le fondement et la force motrice de l'économie³³.

5.2.3.2. Les chiffres clés et les réussites du secteur

L'an 2008 a marqué le début du taux de croissance annuel d'environ 23% du secteur qui se poursuit aujourd'hui. Il est passé de 1 % de part dans l'économie nationale à 6 %. Même s'il ne représente encore qu'une petite partie de l'économie, il est responsable d'une croissance économique hors du commun. Actuellement, le secteur des technologies de l'information dans son ensemble emploie 15 350 personnes et vaut 765,1 millions USD, et exporte pour 338,6 millions USD. Le secteur technologique à lui seul, qui exclut les fournisseurs de services Internet, emploie 9 057 personnes et vaut 612,7 millions USD. Dans un pays où les salaires moyens sont d'environ 400 USD par mois et le PIB par habitant de 3 880 USD par an, le secteur technologique permet aux jeunes adultes de gagner des milliers de dollars par mois et a une production industrielle par habitant de 40 109 USD par an. En grande partie ce sont des filiales de recherche et développement de grandes entreprises occidentales comme Synopsys, National Instruments et Oracle, mais elles font rapidement de la place aux entreprises locales qui deviennent tout aussi influentes. Il s'agit notamment de EFG, mieux connue sous le nom de sa subdivision FGH, officiellement la deuxième plus grande entreprise technologique d'Arménie après Synopsys, et PQR, l'entreprise technologique locale la plus connue publiquement avec plus de cinq cents millions de téléchargements de l'application ; ce sont des entreprises qui possèdent une infrastructure du big data. Parmi les autres réussites nationales dans ce secteur, citons celles du domaine des services telles que GG, la société de covoiturage et menu.am, le service de livraison de nourriture, qui ont tous deux réussi non seulement à dominer le marché intérieur face à la concurrence étrangère (Yandex.Taxi particulièrement), mais aussi à entrer sur les marchés étrangers. Une autre tendance tout aussi importante dans l'industrie, tant pour les entreprises nationales que pour les entreprises étrangères ayant des succursales locales, a été la transition d'un travail externalisé bas de gamme vers un travail beaucoup plus complexe et innovant. Ceci est essentiel dans la mesure où il représente un début de transformation du secteur depuis le bas de l'échelle à un centre d'innovation et de développement d'aspiration mondiale.³⁴

³³ <https://www.gov.am/en/>

³⁴ <https://hti.am/>

Un autre projet d'Etat à noter qui contribue énormément à ce développement technologique, est la création de « Armath Engineering Laboratories » : depuis 2011, l'Union des entreprises de technologie avancée (UATE) s'engage à mettre en place ce programme éducatif. L'UATE a lancé aussi le processus d'exportation du modèle arménien d'enseignement technologique dans le monde entier. En 2022, 623 laboratoires d'ingénierie opèrent sur le territoire de l'Arménie, de l'Artsakh, de la Géorgie et de l'Inde. Environ 17 000 étudiants bénéficient d'une formation gratuite en ingénierie. Aux laboratoires d'ingénierie d'Armath, les enfants âgés de 10 à 18 ans sont initiés à l'enseignement des sciences, de la technologie, de l'ingénierie et des mathématiques grâce à des cours interactifs après l'école. Les jeunes ingénieurs ont la possibilité de concevoir, construire, tester et améliorer leurs propres créations dans un environnement sûr, en créant des startups. Tenant compte du fait que la sphère de la technologie de pointe est l'un des moteurs de l'économie mondiale et du fait que le besoin d'une main-d'œuvre qualifiée dans les domaines de la science, de la technologie, de l'ingénierie et des mathématiques augmente de façon exponentielle, les laboratoires d'ingénierie Armath s'efforcent de jouer un rôle important dans le soutien des tendances de croissance mondiale.³⁵

5.2.3.3. *Perspectives et défis de développement du secteur*

Souvent, on pose la question suivante : comment se fait-il qu'en Arménie dont l'économie est faible et les secteurs peu développés, l'industrie technologique prospère à ce point ?

Nous proposons une combinaison des facteurs qui à notre avis offrent le potentiel principal à cet essor de croissance rapide et la possibilité de salaires élevés qui ne sont offerts dans aucun autre secteur en Arménie :

- tout d'abord, c'est *la nature digitale* de l'industrie qui signifie que le secteur reste moins touché des conséquences de la pléthore des complications géopolitiques, pandémiques et surtout celles de la guerre (2020) qui ont entravé énormément presque tous les secteurs économiques en Arménie. Ainsi, nous constatons, que les entreprises technologiques en Arménie sont en mesure de fonctionner et travailler avec des marchés riches sans

³⁵ <https://armath.am/hy>

ressentir les inconvénients de la fermeture des frontières ou les limitations des infrastructures de transport.

- De plus, comme dit précédemment, cette industrie est dotée *d'une culture nationale et d'une prédisposition* à exceller dans le domaine.
- Enfin, le secteur n'est pas à forte *dépendance de capital ou de ressources*, ce qui est essentiel pour un pays qui en a peu.

Le fait que dans ce secteur l'Arménie soit confrontée à des problématiques similaires à ce que l'on peut retrouver dans d'autres pays, ne l'empêche pas d'avoir ses problèmes spécifiques.

Les défis majeurs qui empêchent le développement de ce secteur, spécifiques en Arménie, sont :

- *L'émigration massive* : les individus émigrent pour obtenir des salaires plus élevés et pour se perfectionner professionnellement dans de grandes entreprises comme Google ou Yandex. Ce problème touche surtout les développeurs de logiciels seniors, qui sont les plus critiques pour la croissance des entreprises, dans le mentorat des jeunes développeurs et dans l'incitation des entreprises étrangères ouvrent des succursales en Arménie ;
- *L'émigration* au niveau des startups qui choisissent d'émigrer pour avoir la chance d'un accès direct aux marchés occidentaux qu'elles ciblent. Les experts technologiques occidentaux qui visitent l'Arménie - tout en complétant le niveau d'expertise technique - énumèrent souvent les lacunes et des problèmes dans le marketing et les ventes des entreprises arméniennes. Par conséquent, les startups recherchent des pays occidentaux où elles peuvent surmonter les barrières linguistiques et culturelles auxquelles elles sont confrontées en Arménie³⁶. De plus, les startups émigrent à la recherche d'investisseurs et d'entrepreneurs et de mentorat technique, ce qui peut être difficile à atteindre en Arménie.
- *La croissance disparate du secteur* dans la capitale (Erevan) par rapport aux régions de province. Bien qu'il existe une multitude d'opportunités pour l'enseignement dans la capitale, elles sont limitées ou le plus souvent n'existent pas dans la plupart des régions. Il en résulte que les deux tiers du potentiel du pays sont inutilisés. De plus, les

³⁶ <http://csi.am/sci.am/en>

opportunités d'emploi pour les cadres supérieurs sont limitées dans les régions où existent des entreprises technologiques.

- *La déficience du système d'enseignement* : le secteur technologique en Arménie est capable de recruter 5 000 nouveaux employés par an, et devra en fait le faire s'il veut atteindre certains objectifs ambitieux qui ont été fixés par le gouvernement dans leur plan de projection de 5 ans. Cependant, les universités et les institutions de formation en Arménie ne sont capables de former que 1 300 diplômés en technologie par an. Pire encore, seulement la moitié de ces diplômés sont qualifiés pour entrer sur le marché du travail sans formation complémentaire. Cela signifie que le système d'enseignement arménien ne répond qu'à un peu plus de dix pour cent de la demande du marché de la technologie³⁷. C'est vrai, que partout dans le monde, des Etats-Unis au Japon, les entités d'enseignement ne parviennent pas à suivre le rythme des exigences de l'industrie technologique, mais ces problèmes sont aggravés en Arménie en raison surtout d'un budget gouvernemental très limité et un système de formation inefficace, besoin d'une restructuration profonde après l'effondrement de l'Union Soviétique. De ce point de vue-là, la création du projet des laboratoires Armath a une importance stratégique, il est certes nécessaire de créer une base de connaissances mais ce n'est pas suffisant.

Donc, il est raisonnable de supposer que le potentiel de croissance dans l'industrie technologique peut être soutenu en Arménie, voire accéléré, et donc il est primordial d'entreprendre des actions afin de résoudre correctement et le plus rapidement possible ces problèmes.

5.2.3.4. Technologie et sécurité de l'Arménie

Enfin, l'aspect assez critique à notre avis et le moins discuté par le secteur technologique arménien sont ses potentielles implications militaires. Développer correctement le secteur technologique en disposant des institutions et des experts nécessaires est une exigence essentielle si l'Arménie espère développer une industrie de défense nationale viable qui sera capable de relever les défis de la sécurité de l'Etat.

Compte tenu de l'évolution rapide de la nature de la guerre récente du Haut-Karabagh en 2020 et le rôle croissant de l'intelligence artificielle dans l'armement militaire, il est essentiel

³⁷ <https://armstat.am/en/>

pour l'Arménie de pousser son développement technologique au maximum pour maintenir un degré plus élevé d'indépendance et être en capacité d'assurer sa propre sécurité.

Et si l'Arménie est capable de faire des progrès dans ce domaine, cela aura des impacts politiques et diplomatiques importants. En renforçant ainsi sa défense, le pays peut retrouver sa crédibilité dans les relations internationales aussi bien avec ses partenaires qu'avec ses ennemis.

Tout aussi important, cela fera de l'Arménie un partenaire intéressant et attractif aux yeux de la communauté internationale, source de prospérité dans un monde interconnecté.

5.3. Présentation détaillée des acteurs du terrain

Dans ce travail de recherche, la construction des connaissances s'appuie sur 2 grands processus : l'exploration et le test.

Dans les sections qui suivent, nous détaillons les trois approches successives mobilisées au niveau de terrain de recherche.

Pour notre première approche exploratoire, il s'agit d'une étude qualitative dont l'objectif est de mettre en évidence la présence de 7 thèmes du modèle de Surbakti et al. (2019) dans une activité d'entreprise qui exploite le big data. Pour ce faire, nous étudions un cas bien précis – un service de taxi. Ce terrain nous permet de vérifier que les composants du modèle sont bien présents parmi les préoccupations des acteurs.

Après avoir eu les premiers résultats de cette étude qualitative, nous testons quantitativement le modèle sur le terrain, représenté par 211 professionnels du big data qui travaillent dans 16 entreprises arméniennes.

Ainsi, pour la suite nous allons proposer de brèves présentations de ces 16 entreprises.

Pour la troisième approche, nous avons étudiés les 2 cas bien précis d'efficacité d'usage chez les 2 acteurs-participants dans notre étude quantitative. Ainsi, pour ces deux entreprises, nous présenterons de façon plus détaillée, les deux cas bien précis.

5.3.1. Première étude exploratoire : Étude d'un expert en utilisation du big data

ABC³⁸ est une organisation qui s'est installée en Arménie en 2018. C'est un moteur de recherche d'envergure mondiale. La mission de cette entreprise est d'aider les consommateurs (particuliers et entreprises) à mieux naviguer dans les mondes en ligne ou hors-ligne. Cette entreprise technologique crée des produits et services intelligents utilisant le Machine Learning (ML). C'est une technologie d'intelligence artificielle permettant aux ordinateurs d'apprendre sans avoir été programmés explicitement à cet effet. Cet apprentissage des ordinateurs repose sur les gros volumes de données disponibles à analyser. Ainsi le big data et ML entretiennent des liens étroits, puisque le big data est la « matière première » de Machine Learning, qui constitue la technologie permettant d'exploiter pleinement le big data.

ABC dispose donc d'un vrai savoir-faire et d'une forte expertise dans le domaine de l'exploitation du big data. C'est en particulier cette expertise qui fait d'ABC un cas particulièrement intéressant pour nous assurer de la présence des composants du modèle conceptuel de l'efficacité d'usage du big data de Surbakti et al. (2019).

Comme le premier terrain de notre recherche, nous avons choisi l'une des entités d'ABC, celle du service de taxi qui fonctionne sur le marché arménien. C'est un service de covoiturage lancé en 2011 par ABC, entreprise-mère. La nature du service ABC est un service de taxi en ligne qui fonctionne sur un principe de maraude et une application de géolocalisation. Lors de son apparition, ce type de service a été considéré comme « une innovation de rupture » dans la littérature (Utterback & Acee, 2005), (Christensen, 1997). A l'époque les chercheurs s'intéressaient surtout à son impact sur l'économie, l'environnement et la société (Majchrzak, Lynne Markus, & Wareham, 2016). Donc, il se trouve que ce service novateur de taxi a totalement révolutionné l'industrie du service traditionnel des taxis (Castellano, 2016).

Notons, que ce type de service de taxi a été initialement lancé en 2010 à San Francisco par Uber (Berger, Chen, & Frey, 2018). Le fonctionnement du service et le big data à l'instar d'Uber, ABC ne possède pas de voiture, ni de chauffeurs salariés.

³⁸ <https://ABC.com/company/>

Pour servir ses clients, l'entreprise met en relation des chauffeurs-partenaires et des passagers via une application mobile. A l'aide de cette application, le passager peut commander une course dans les villes où le service est disponible. La commande passée est reçue par les chauffeurs à proximité. Une fois acceptée par l'un d'entre eux, l'application indique au client le délai d'arrivée du chauffeur sur le lieu de prise en charge, puis elle lui signale l'arrivée du chauffeur sur le lieu de rendez-vous. L'application transmet également au client la marque et le modèle du véhicule, le numéro de taxi ainsi que le moyen de contacter le chauffeur. Le client peut quant à lui indiquer au chauffeur s'il le souhaite, l'itinéraire à choisir. L'application lui permet également d'obtenir une estimation du prix de la course qui se calcule automatiquement. Une fois arrivé à destination, le prix de la course peut, soit être débité sur la carte du client, soit être payé en espèces. Enfin pour clore la transaction, l'application sollicite le client pour noter son chauffeur et laisser un commentaire. Le chauffeur peut également laisser son commentaire concernant son client. Cette brève description du processus de fonctionnement d'ABC permet d'illustrer les empreintes numériques générées par le service. Ces empreintes sont aussi l'une des sources du big data utilisées chez ABC. Les données sont stockées et traitées en temps réel. Par exemple, le suivi du trajet de la course en temps réel permet de détecter le comportement potentiellement irrégulier d'un chauffeur, qui sans raison ne suivrait pas l'itinéraire recommandé par l'application pour une course donnée, itinéraire qui est généré par l'algorithme « On A Trade ». Autre exemple, l'algorithme « hausse de prix » permet de connaître l'évolution de la demande du service en temps réel, d'estimer l'élasticité de la demande par rapport au prix, et de gérer ainsi le surplus de clients sur une zone donnée, à un moment donné.

Les algorithmes sophistiqués de traitement du big data d'ABC permettent ainsi à l'entreprise d'accomplir trois activités majeures :

- Déterminer la demande ;
- Octroyer les ressources nécessaires pour satisfaire la demande ;
- Fixer les tarifs.

Les capteurs GPS qui équipent les véhicules du service de taxi laissent de nombreuses empreintes numériques des activités et des mouvements quotidiens de leurs clients. Elles constituent ainsi d'importantes et riches sources de données, qui sont combinées à d'autres données et permettent de mieux comprendre et d'analyser en

profondeur les dynamiques de comportement d'un individu ou d'un groupe social. Pour conclure cette présentation, soulignons qu'ABC utilise ses propres technologies de cartographie (mapping), de routage (routing) et de navigation.

Une partie du big data crée dans le cadre de l'activité de taxi est composé de données structurées (celles issues des capteurs GPS par exemple), de données semi-structurées et données non-structurées (commentaires des clients et des chauffeurs-partenaires, enregistrements audio), ce qui suppose notamment des capacités de traitement adaptées pour une utilisation efficace.

5.3.2. Étude quantitative : profils des participants

5.3.2.1. PQR³⁹

Cette société a été fondée en 2011. Elle emploie de plus de 400 salariés. C'est un éditeur photo et vidéo spécialisé au niveau international. Plus précisément, il s'agit d'une application de montage de photo et vidéo avec une plate-forme sociale pour les artistes. Ces derniers créent une communauté autour d'images éditées de manière créative.

PQR maintient sa position parmi les 50 meilleures applications sur Google Play. De plus, elle occupait la deuxième place du classement international iTunes des applications gratuites. L'application est traduite en 30 langues et attire le public du monde entier.

À l'heure actuelle, PQR a atteint 130 millions d'utilisateurs actifs mensuels et de 600 millions d'installations.

Elle exploite des pipelines de données par lots et en temps réel à grande échelle avec des infrastructures de traitement de données Spark.

Pour son fonctionnement quotidien, l'entreprise utilise les données massives hétérogènes et les systèmes Hadoop, HDFS et Parquet ; les bases de données MongoDB, Redis.

³⁹ <https://pqr.com/>

5.3.2.2. DEF⁴⁰

Fondée en 2010, cette société emploie plus de 150 salariés. Son siège social est situé à Sunnyvale, en Californie, un bureau est aux Philippines. Environ 50 personnes travaillent dans le bureau arménien de la société.

DEF est une plate-forme de création de publicités qui transforme le marketing des applications mobiles à l'aide du big data, de l'apprentissage automatique et d'une large clientèle. Les centres de données mondiaux de la société atteignent plus de 300 000 requêtes par seconde. Ces données offrent des informations approfondies sur les intentions et les habitudes d'utilisation des clients, puis sont utilisées pour construire des modèles d'utilisateurs prédictifs et mettre en œuvre des enchères ciblées. Une telle approche aide les entreprises travaillant avec DEF à se développer et à réengager leurs clients.

En 2018, DEF a été classé sur Technology Fast 500 de Deloitte, un classement des 500 entreprises technologiques, de médias, de télécommunications, de sciences de la vie et de technologie énergétique comme ayant la croissance la plus rapide en Amérique du Nord et numéro 4 dans la Bay Area.

⁴⁰ <https://www.def.com/>

5.3.2.3. *EFG et FGH*⁴¹

Fondé par les frères Aramyan en 2010, EFG est un développeur de technologies innovantes pour plusieurs domaines TI en Arménie. La société a réussi à créer des labels leaders sur le marché tels que FGH, RST, STU, TUV, UVX et plus encore. Marques indépendantes, toutes ces marques forment une grande famille EFG.

Dans leurs courses, les technologies du groupe d'experts d'EFG a initié de nouvelles approches pour travailler avec les consommateurs en proposant des produits et des solutions spécifiquement conçus pour l'industrie.

À ce stade, EFG compte plus de 5 000 employés, plus de 10 marques et un certain nombre de succursales opérant dans le monde entier.

5.3.2.4. *GHI*⁴²

Cette société, fondée en 2019, propose des services de développement de logiciels, avec une expertise dans la fourniture de solutions d'apprentissage automatique et de science des données. Elle dispose de développeurs backend et frontend hautement qualifiés travaillant avec les piles JS - React, Vue.JS et .NET. Ils emploient les meilleurs et les plus brillants talents d'Arménie dans la création de technologies et de produits de pointe.

Cette entreprise aide à utiliser des informations commerciales précieuses à partir des données disponibles pour mieux comprendre le public, prévoir les demandes, réduire les risques, prévenir les dépassements de coûts, augmenter la productivité, etc.

Les services du bloc de big data permettent de résoudre des problèmes dans différents domaines de l'IA, notamment les algorithmes supervisés et non supervisés, le traitement du langage naturel, la vision par ordinateur et bien d'autres. Notamment, ils aident à prendre des décisions commerciales en fonction des données, via des analyses statistiques, des modèles prédictifs et des visualisations avancées.

⁴¹ <https://www.efg.com/>

⁴² <https://ghi.com/>

5.3.2.5. *HIJ*⁴³

Cette société se positionne comme une machine d'intelligence artificielle conçue pour créer des décisions en temps réel. Elle propose des services d'aide à la décision en temps réel adaptés à la plateforme ou au système du client.

A l'aide d'une approche personnalisée appliquée à l'analyse des données du client, elle aide à prendre des décisions en temps réel à l'aide d'outils d'intelligence artificielle adaptés à ses objectifs et projections spécifiques.

Elle propose 2 options de coopération à partir de leur propre plateforme :

- Modèles créés par les utilisateurs ;
- Modèles créés par la société elle-même.

Dans le premier cas, le client reçoit l'accès à leur plate-forme avec toute la formation et la documentation nécessaires sur la façon d'utiliser ses fonctionnalités.

Dans le deuxième cas, une équipe travaille en étroite collaboration avec le client pour identifier les problèmes solubles grâce à la modélisation. Ils aident donc à explorer et à définir les données utilisables dont le client dispose.

5.3.2.6. *KLM*⁴⁴

Fondée en 2014, cette société propose des services d'exploration de données et d'analyses statistiques.

Ces services sont regroupés en 3 domaines :

- Apprentissage machine et analyse statistique ;
- Recherche et analyse marketing ;
- Visualisation des données et formation.

Ils planifient, mettent en œuvre et finalisent des projets liés aux données depuis des problématique métier jusqu'au déploiement du modèle. Ils offrent également une formation en

⁴³ <https://hij.ai/>

⁴⁴ <https://klm.com/>

entreprise dans le cadre de leurs services pour aider à maîtriser les données de manière plus efficace et proactive.

Leur mission consiste à aider les spécialistes du marketing, les décideurs, les entrepreneurs et les entreprises, en général, à prendre des décisions basées sur les données et ainsi à stimuler leurs ventes.

5.3.2.7. LMN⁴⁵

Cette société a été fondée en 2018 ; créée par Aram Aramyan et Armen Armenyan, elle emploie environ de 100 salariés. Elle a été implantée avec une mission d'utiliser l'intelligence arménienne pour créer des technologies d'IA de classe mondiale.

Elle propose des solutions d'intelligence artificielle et d'apprentissage automatique pour les entreprises. La technologie développée par la société est applicable dans les grandes et petites entreprises, les sociétés, les institutions et les gouvernements. Récemment, elle a lancé une plate-forme d'IA effectuant des prédictions intelligentes en temps réel sur le comportement des clients en fonction des données massives du client. La plate-forme donne un aperçu de l'interaction du public avec l'entreprise en ligne, augmente l'efficacité des campagnes marketing avec la segmentation et le ciblage appropriés du marché.

Parmi les produits développés par la société figure aussi le système de reconnaissance faciale entièrement développé et mis en œuvre avec succès en Arménie. Le modèle d'apprentissage automatique du système divise un visage en neuf parties différentes et peut identifier des personnes même dans une seule image. Le système a été entraîné sur une base de données d'environ 1 million de personnes à différents moments de leur vie.

5.3.2.8. JKL⁴⁶

Cette société est positionnée comme un fournisseur des informations innovantes et fiables aux acteurs de l'industrie du divertissement ; spécialisée dans les jeux vidéo et le merchandising, elle capte des données massives sur les ventes et le streaming.

⁴⁵ <https://lmn.ai/>

⁴⁶ <https://www.jkl.com/>

La stratégie de la société consiste à fournir le service d'exploitation de grand volume des données de façon fiable et professionnelle, au travers de leur plateforme EMBER.

La société propose principalement 2 types de services :

- *Services d'analyse du marché* (présenter les points clés sur la performance du marché, sur les ventes digitales et au détail, sur l'engagement des communautés à l'aide des API publiques et privées à partir de plate-formes numériques)
- *Consulting* (tenir au courant des dernières tendances du secteur avec des informations personnalisées et analyses ciblées).

5.3.2.9. *MNO*⁴⁷

La société a été fondée en 2016 à Erevan. La mission de cette société est de simplifier les opérations commerciales axées sur les données, en rendant les données plus accessibles pour en extraire des informations utiles. Elle emploie près de 200 salariés ; une équipe dédiée de DataScientists et d'experts en sciences de la vie, passionnés par l'innovation technologique en science des données et en apprentissage automatique favorise la croissance d'une culture organisationnelle axée sur les données.

La vision de cette entreprise est d'appliquer l'IA et l'apprentissage en profondeur pour l'automatisation des flux de travail et d'obtenir des informations significatives sur les données. Ils aident leurs clients à prendre des décisions basées sur des données de plusieurs milliards de dollars.

Cette société pratique une combinaison d'Intelligence Artificielle, d'Apprentissage Machine, de Cloud Computing, de Big Data et de Blockchain.

5.3.2.10. *NOP*⁴⁸

Ce service à la demande, a été créé en 2014 par Arsen Arsenyan et Avet Avetyan. La société a obtenu des investissements de la société de capital-risque arménienne VXY⁴⁹. Elle a aujourd'hui plus de 300 partenaires.

⁴⁷ <https://mno.ai/>

⁴⁸ <https://www.nop.com/>

Cette société propose une plate-forme de transport à la demande reliant les conducteurs et les passagers dans toutes les grandes villes d'Arménie, fournissant des transferts interurbains, des services de camions et d'évacuateurs partant d'Arménie. NOP propose ses services désormais en Géorgie et en Russie avec plus de 1000000 utilisateurs actifs par mois. Comme dans le cas d'ABC, l'usage du big data leur permet de régler d'abord à des problématiques opérationnelles, comme la régulation de l'offre et de la demande.

5.3.2.11. OPQ⁵⁰

Cette société, fondée en 2019, emploie environ 100 salariés. Elle propose le conseil et développement en science des données/apprentissage automatique. Elle se présente comme l'un des principaux fournisseurs de services d'IA avec un portefeuille diversifié de projets d'apprentissage automatique en science des données.

Leur mission est de doter les entreprises d'intelligence artificielle et d'informations basées sur les données ; en convertissant leurs données en informations exploitables. Elle propose 4 types de services :

- *Intelligence*, qui comprend une chaîne de processus suivants « visualisation et reporting », « stratégie de développement », « optimisation du flux de travail », « feuille de route d'adoption de la technologie », « sélection des KPI » ;
- *Analytics* qui comprend « segmentation de la clientèle », « systèmes de recommandation », « action client », « test A/B », « analyse des préférences », « optimisation de la valeur du cycle de vie », « analyse des sentiments » ;
- *Ressources des données* avec « web scraping », « survey design », « text analytics », « computer vision », « architecture de stockage ».

5.3.2.12. UTS⁵¹

La société a été fondée par Ara et Avet Aramyan. Créée en 1993, elle fonctionnait sous de différentes marques.

⁴⁹ Une société de capital-risque à fournir une expertise et des réseaux d'investissement aux startups du monde entier qui exploitent le potentiel de l'Arménie en tant que pôle technologique émergent.

⁵⁰ <https://opq.am/>

⁵¹ <https://www.uts.am/>

Elle propose aujourd'hui des services de communication téléphonique GSM 900/1800, 3G (UBCD) et de communication mobile standard 4G (LTE), de communication téléphonique fixe, et fournit également un accès Internet utilisant les technologies FTTH, CDMA, ADSL, FTTB et VDSL.

5.3.2.13. VUT⁵²

La société a été fondée en 2009. C'est un leader sur le marché de l'IPTV et de l'internet fixe en Arménie et occupe la position clé sur le marché arménien de l'internet mobile.

Agissant dans un marché hautement concurrentiel et en évolution rapide, où les services voix et SMS traditionnels connaissent un net déclin, la société a revu ses priorités et ses approches commerciales : elle accorde plus d'importance à l'implication active de ses clients dans son activité et essaie de coopérer étroitement avec eux, afin de connaître les exigences de la prochaine génération pour les produits/services ou les modèles de consommation.

En tant que leader de la technologie et de l'innovation, elle essaie de renforcer continuellement les principales orientations de ses activités actuelles, tout en créant de nouvelles opportunités de revenus en développant des solutions numériques innovantes et différenciées : VUT est l'une des rares entreprises traditionnelles arméniennes à implanter une infrastructure appropriée au traitement du big data.

5.3.3. Deux cas bien précis : « utilisateur novice » et « utilisateur avancé »

5.3.3.1. BCD Arménie, « utilisateur novice » du big data

BCD Arménie a été lancé en 2005. Son arrivée sur le marché arménien a mis fin au monopole dans le domaine où la situation a rapidement changé de manière importante.

BCD possède son propre réseau de câbles à fibre optique dans tout le pays, s'étendant du nord de l'Arménie au sud. La Société importe Internet par ses propres canaux pour les besoins du marché domestique.

La société compte actuellement environ 1 200 employés et plus de 2 millions d'abonnés.

⁵² <https://www.vut.am/>

5.3.3.1.1. Profil du répondant

C.C.

Data Science Team Lead chez BCD Arménie dès l'implantation du bloc (2018).

Titulaire du Master en BA « Finance et Management » et Master en « Ingénierie nucléaire »; enseignant du cours de Marketing Analytics.

Plus de 9 ans d'expérience comme Scientifique des données.

Ses domaines d'expertise comprennent l'analyse client, la gestion de la valeur client, la géo-analyse et l'apprentissage automatique.

5.3.3.1.2. Cas détaillé

Depuis 2019, l'entreprise exploite pratiquement le big data. Dès le début de lancement du projet de l'exploitation du big data, l'entreprise a considéré un lancement dans une philosophie complète « du big data ». Le premier problème qui a été rencontré, était le manque de ressources intellectuelles et humaines spécialisées dans le domaine.

Comme dans un cas général, ils ont considéré les 2 cas traditionnels du développement de technologies du big data : « in-house » et « sur les solutions Cloud ». Ces 2 cas diffèrent complètement dans leurs scénarios de développement pour l'entreprise. Pour le premier cas, l'entreprise a considéré le développement « in-house » comme une dépense d'investissement et à partir de quelques scénarios, ils ont discuté avec les experts et ont calculé son apport sur 10 ans. Les calculs ont montré que pour ce premier cas les investissements dans un développement « in-house » infrastructure, ne sont pas justifiés, compte tenu sa valeur attendue.

Un des projets en cours du département DataManagement a été détaillé par C.C. pour lequel l'entreprise gagne de la valeur et a été présenté comme un cas d'efficacité d'usage du big data par M C. C.

Grace à un outil DPI (Deep Packed Inspection), qui est un outil payant, ils peuvent traquer tous les applications et les pages Web que les abonnés de l'entreprise utilisent (Annexe 23).

La question qui a été posée, c'est de comprendre les profils différents qu'ils ont parmi leurs utilisateurs.

Ainsi un besoin secondaire a été défini. Il s'agissait de regrouper ces profils par leurs usages spécifiques de certaines applications pour une période donnée. Donc, une des tâches consistait à collecter des données historiques concernées (plus, le trafic en MB utilisé en fonction du temps). Donc c'est un processus en continue : ces données sont enregistrées sous différents protocoles.

Une fois, les données collectées et stockées, l'on passe à l'étape suivante pour laquelle il s'agit de comprendre quelle part de données stockées est utile : (1) quelle période est utile à considérer (2) et quelles sont les applications /les sources à considérer comme utiles (3), compte tenu la question principale posée.

Une fois que ces 3 questions ont trouvé leurs réponses, une fois la partie utile des données mise en évidence ; elles passent à un tri journalier et horaire.

L'étape suivante, débute avec la préparation du rapport de prise de décision. Ce processus prend du temps. Une fois que l'entreprise a fini de construire ces profils de clients, elle les utilise dans ces différentes campagnes.

Pour l'analyse et le traitement de toutes ces données, l'entreprise utilisent des méthodes de « clustering » (regroupement) ; ce qui permet d'abord de comprendre « qui utilise quoi », par les flèches-transcriptions que leur transmet l'entreprise (d'où l'on peut comprendre quel « ip » appartient à quelle application de marques bien connues) et des méthodes de modélisation prédictive (predictive modeling).

Par la suite, les informations reçues des clusters sont utilisées pour construire des insights, les perspicacités qui sont après utilisées largement dans les développements des produits, dans les stratégies de rétention, dans les stratégies d'acquisition, dans les ventes incitatives et croisées.

5.3.3.2. CDE⁵³, « utilisateur avancé » du big data

Cette société a été fondée en 2016. Elle emploie de plus de 150 salariés. Elle propose une plate-forme (PaaS) fournissant des données sur les consommateurs. CDE permet de comprendre ce que les gens pensent et font tout au long de l'expérience de la marque.

CDE collecte, analyse et fournit des données de qualité sur les consommateurs aux plus grandes agences d'études de marché, sociétés d'analyse et marques du monde.

La mission de l'entreprise est d'aider les clients à découvrir le « pourquoi » derrière les opinions et le comportement des consommateurs dans le but d'améliorer le processus de prise de décision grâce au big data.

En 2018, la société a levé 13,5 millions USD auprès de la société XYZ⁵⁴.

5.3.3.2.1. Profil du répondant

D.D.

Sénior Scientifique des données chez CDE depuis 2019, responsable du bloc de développement des systèmes de recommandation et de modèles de prédiction.

Docteur en « management du commerce électronique », titulaire du Master en « Innovation et Management », Professeur en « Programmation pour la science des données ».

Plus de 10 ans d'expérience comme Scientifique des données.

Ses principales compétences et son domaine d'expertise comprennent l'apprentissage automatique, les réseaux de neurones, l'analyse de données massives (big data).

5.3.3.2.2. Cas détaillé

Initialement fondée en 2015 comme une start-up sur un investissement de 100 millions de dollars, cette entreprise constate un fort potentiel de croissance avec son modèle économique innovant pour le marché arménien. CDE est basé initialement sur l'idée des données et toute son

⁵³ <https://www.CDE.com/>

⁵⁴ XYZ est une société de capital-risque axée sur la connaissance qui investit dans des entreprises en démarrage utilisant l'IA, les données massives et la science comportementale pour fournir des informations exploitables et ouvrir de nouveaux marchés.

activité s'appuie sur les 2 types de données : les données comportementales et les données d'opinions.

Comme les données comportementales (behavioural data), ils stockent des données des utilisateurs qui se sont inscrit préalablement et ont donné leur accord pour que leurs données soient traitées. Quand l'utilisateur est en ligne, CDE peut stocker les données sur les sites visités, les données sur les recherches en ligne, les données sur les achats en lignes – tout type des données peuvent être conservées. Actuellement, plus de 1 milliard de points de contact sont stockés par mois (touchpoint), donc toutes les actions des utilisateurs sur le réseau sont considérées séparément.

La société puis développe de nombreux produits-offres à partir du traitement de ces données collectées. Dans ce segment, ils traitent 2 types de données : les données comportementales et les données d'opinion. Sur les données comportementales, ils ont développé le produit « mesure du média numérique », par exemple, l'entreprise X qui prépare un budget de 1 million de dollars pour sa nouvelle publicité, et avant de la distribuer veut comprendre son impact potentiel, et un test d'influence.

La 2^{ème} orientation d'activité représente le secteur des études de marché basées sur les données d'opinion (opinion data). Ce type d'activité est nouvelle sur le marché arménien ; mais largement répandu en occident et courant aux Etats-Uni, il est largement utilisé comme un outil de prise de décision basée sur les données (DDD-Data Driven Decision). Ici l'on prend compte l'opinion des consommateurs. Cela se fait par les enquêtes électroniques en ligne. Ce marché est assez intéressant pour CDE, il intègre un large public pour ce type d'activité ; ce sont des personnes qui s'inscrivent sur leur plateforme, où ils prennent des enquêtes et gagnent des points qu'ils peuvent transformer en carte cadeaux. Ce mécanisme de collecte des données d'opinion permet les utilisateurs de gagner près de 1000 dollars par mois ; comme cela ils collectent près de 3 millions d'enquêtes par mois.

On exploite chez CDE une infrastructure AWS (Amazon Web Services)⁵⁵ qui propose beaucoup d'outils nécessaires pour traiter et analyser des données. Les scientifiques des données

⁵⁵ Une plateforme cloud assez complète et largement adoptée au monde qui propose plus de 200 services issus de centres de données du monde entier.

de CDE utilisent largement SageMaker⁵⁶ dans leurs tâches quotidiennes pour rendre le ML plus accessible, et pour préparer les données à grande échelle, accélérer le développement et rationaliser le cycle de vie du ML.

L'outil technique, principalement utilisé par l'entreprise qui permet de travailler avec big data, c'est Spark. Il y en a 2 types, Spark qui est basé sur le langage de programmation Python – PySpark et la programmation de Scala qui sont exploités fréquemment.

Notons qu'un autre outil existe, c'est Databricks, qui est une plate-forme de traitement et d'analyse de données développée par les mêmes personnes responsables de la création de Spark. Son cœur est une instance Spark modifiée appelée Databricks Runtime, qui est hautement optimisée même au-delà d'un cluster Spark normal. Il est également occupé avec plusieurs technologies de gestion de données et outils de visualisation. En tant que plate-forme cloud, elle peut être déployée sur les différents fournisseurs de cloud et il est facile de créer et de mettre à l'échelle des clusters à l'aide de l'interface de Databricks. Donc tous ces plate-formes et outils de traitement du big data sont utilisés dans le Cloud au niveau d'AWS.

Les coûts opérationnels liés à l'usage du big data par CDE sont des frais, payés mensuellement à AWS (le montant n'est pas fixe, en fonction du volume des données qui augmente, les frais augmentent aussi). Chez CDE, existe un indicateur à partir duquel les frais mensuels d'AWS qui ne doivent pas dépasser un certain pourcentage du CA du mois en cours, qui rend cet usage efficient, donc il est appliqué scrupuleusement, avec un suivi du stockage des données au meilleur format et sans coûts excessifs. Ainsi, le développement de nouveaux offres-produits à partir des données disposées est une tâche régulière de l'entreprise.

⁵⁶ Proposé par Amazon, ce service permet aux scientifiques des données et aux développeurs de créer, former et déployer rapidement et facilement des modèles de machine learning à n'importe quelle échelle.

Chapitre 6 Résultats

Dans le cadre de ce travail de thèse, nous avons réalisé trois études.

Notre première étude qualitative exploratoire a été réalisée chez une société experte d'exploitation du big data, ABC.

Notre objectif pour cette étude était de vérifier que les composants du modèle conceptuel de Surbakti et al. (2019) sont bien présents dans l'activité d'une entreprise ayant une forte expertise dans le domaine de l'exploitation du big data.

Le fait de mener les entretiens avec un membre représentant de la haute direction nous a permis de toucher les thèmes plutôt du point de vue stratégique qu'opérationnel.

La deuxième étude quantitative a été conduite auprès de 211 professionnels du big data, représentant 16 entreprises arméniennes, en utilisant la méthode d'enquête. Elle avait pour objectif principal de tester le modèle conceptuel de Surbakti et ses coauteurs (2019) et répondre à la question de recherche.

Ainsi, l'analyse statistique des données nous a permis de présenter les deux grands résultats de ce travail. Ces résultats nous permettent de répondre complètement à la question de recherche proposée.

Et la troisième étude qualitative confirmatoire a porté principalement sur une étude des cas bien précis d'efficacité d'usage du big data chez deux types d'utilisateurs (« un utilisateur novice » et « un utilisateur avancé »). En plus, nous avons réalisé un entretien avec l'un des experts du groupe QRS Co., qui avait participé aussi à l'étude-pilote de notre recherche quantitative.

Cette troisième approche nous a permis de :

- Montrer, comment émergent les facteurs les plus importants issus de l'étude quantitative dans chacun de ses cas ;
- Proposer un prolongement des résultats ;

- Comprendre les particularités du contexte arménien.

Ainsi, les paragraphes qui suivent, présentent bien les trois principaux résultats de ce travail de recherche.

6.1 Résultat 1: « Utilisateur novice » et « utilisateur avancé »

Déjà détaillé dans la partie méthodologique de la thèse, le traitement des données de l'étude quantitative a permis de constater une différence statistique significative pour les deux groupes des répondants (Annexe 9). Cette différence correspond à la qualité de la relation de l'entreprise du répondant à l'exploitation du big data. Ainsi, la question concernée dans la partie de l'information démographique de l'instrument proposait deux choix de réponse possible : « l'entreprise du répondant a adopté une infrastructure nécessaire pour l'exploitation du big data » ou « l'entreprise du répondant est basée initialement sur l'idée de l'exploitation du big data ».

Ainsi, ce résultat nous permet de proposer une distinction en deux groupes d'utilisateurs du big data quand l'on traite l'efficacité d'usage en entreprise en contexte arménien : « utilisateur novice » et « utilisateur avancé ».

Dans le cadre de ce travail de thèse, nous avons proposé les définitions à ces deux groupes d'utilisateurs du big data :

- Une société de type « utilisateur avancé » du big data s'est fondée sur l'idée de l'exploitation du big data, elle en possède une forte expertise et son fonctionnement y est lié principalement.
- Une société de type « utilisateur novice », exerce une activité principale et a commencé à exploiter le big data à un moment donné de son fonctionnement.

Il faut bien noter, que cette étape de définition est primordiale pour ce travail de recherche. L'importance de ces deux définitions est attribuée à deux pistes.

Premièrement, ces définitions permettent de bien décrire le profil d'une société en tant qu'un utilisateur du big data.

Et deuxièmement, c'est une introduction importante dans ce domaine de recherche (Barki, Paré, & Sicotte, 2008). Ainsi, nous essayons de proposer des définitions consensuelles qui doivent être prises en compte et développées dans les futures recherches du domaine (Ronda-Pupo & Guerras-Martin, 2012).

Nous avons donc mis en évidence que le niveau de l'utilisateur détermine les caractéristiques théoriques qui doivent être employées pour l'analyse. Il y aura donc deux variantes du modèle selon le niveau d'utilisateur, comme nous allons le voir dans le second résultat.

6.2 Résultat 1 bis : Société ABC, « utilisateur avancé » du big data

L'analyse thématique nous a permis d'identifier la présence des 7 thèmes du modèle de Surbakti et al. (2019) dans l'exploitation du big data par ABC, une entreprise de type « utilisateur avancé ». Nous avons pu ainsi identifier clairement les 33 sur les 41 facteurs inclus dans les 7 thèmes. La figure 41 illustre les facteurs explorés qui sont marqués en rouge.

Les paragraphes suivants présentent brièvement les résultats sur chaque thème du modèle, en insérant les verbatims les plus illustratifs.

6.2.1 F1 : Intérêt organisationnel perçu

Ce thème est considéré comme la priorité initiale d'ABC pour construire un avantage comparatif. « *ABC est né technologique, sa mission, c'est aider ses consommateurs à mieux naviguer dans les mondes en ligne ou hors-ligne, en suivant et en utilisant les avancements technologiques dans les services et produits proposés. Sa stratégie, c'est de proposer un meilleur service grâce à la technologie* »

Nous avons repéré les facteurs suivants :

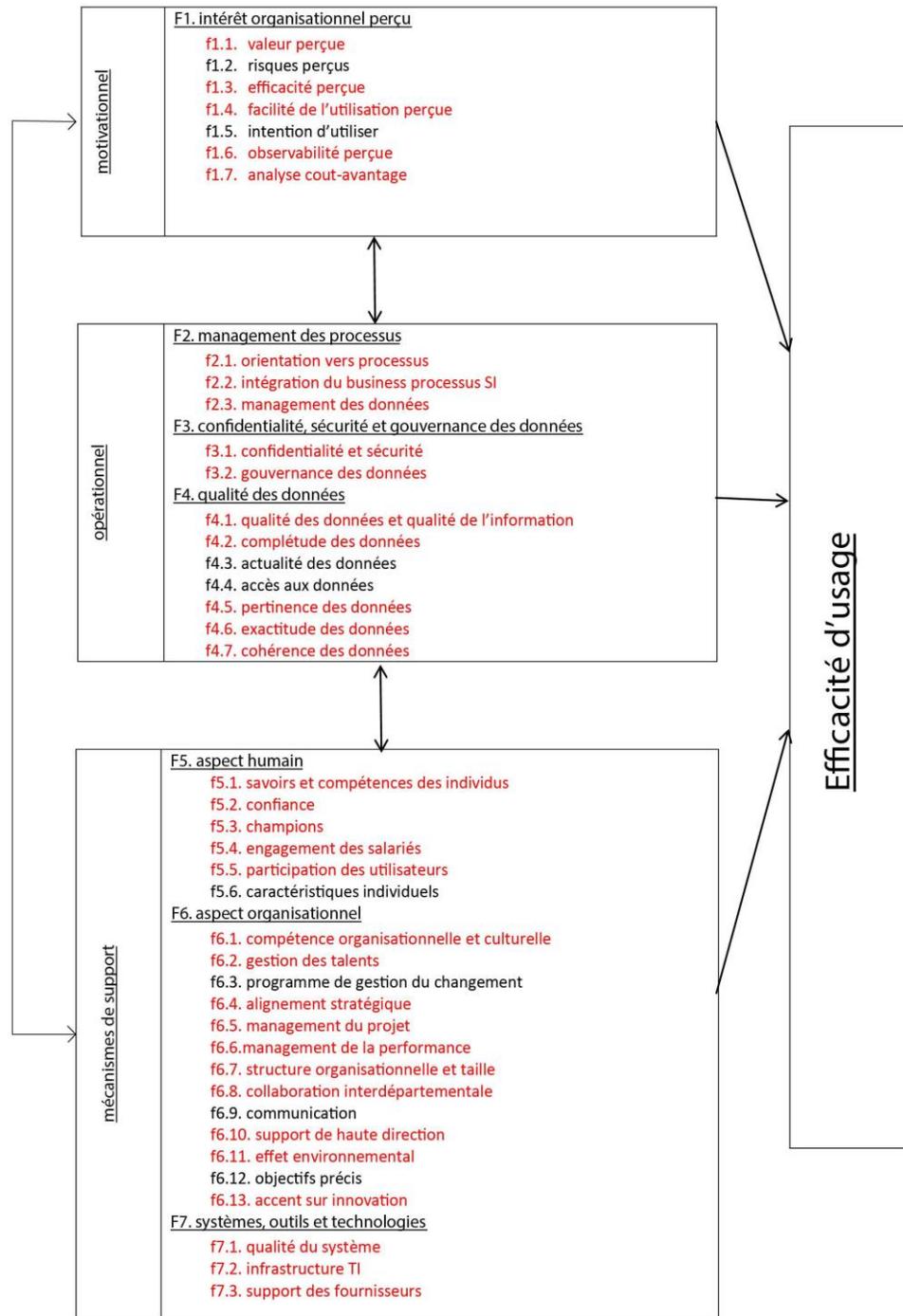
- Valeur perçue du big data (f1.1) et son efficacité perçue (f1.3) (Figure 41) : « *...chez nous, la valeur des services et celle de toute l'organisation repose sur nos capacités technologiques d'exploitation du big data ; ce type de service de taxi n'a de sens sans*

l'exploitation du big data ... » ; «...pour un projet de service de taxi, on ne voit pas de grande valeur dans un modèle traditionnel aujourd'hui ».

- Facilité perçue d'utilisation (f1.4) : *«...l'exploitation du big data dans ce service innovateur de taxi rend le service plus rapide, plus sécuritaire, plus confortable et donc plus agréable pour nos consommateurs... ».*
- Observabilité perçue (f1.6) et analyse coût-avantage (f1.7) : *« ...par rapport au service traditionnel du taxi, qui va, à notre avis, bientôt disparaître, ce service innovateur renvoie aux résultats plus visibles et plus avantageux dans les divers contextes ... ».*

On observe ainsi, que cette catégorie « motivationnelle » est éminemment stratégique dans le cas d'une entreprise technologique puisque le big data et son usage sont intimement liés aux objectifs stratégiques de l'entreprise. Autrement dit, l'intérêt organisationnel perçu du big data dans une organisation née technologique apparaît nécessairement fort et par ricochet son impact sur l'efficacité d'usage l'est aussi. Pour autant, les perceptions des professionnels devront compléter ces premiers résultats, particulièrement, celles des scientifiques des données et des ingénieurs des données, pour lesquels le big data est au cœur de leur l'activité quotidienne.

Figure 41 Modèle de l'efficacité d'usage, traduit de Surbakti et al. (2019)⁵⁷



⁵⁷ Les facteurs, marqués en rouge, ont été identifiés lors de l'analyse thématique réalisée dans le cadre de la première étude exploratoire.

6.2.2 F2 : Management des processus

Dans la catégorie « opérationnelle », le premier thème est celui du management des processus. Il s'agit de la capacité d'une organisation à structurer ou restructurer ses processus pour tirer parti des opportunités du big data. Le succès d'ABC souligne un rapport qualité-prix compétitif du service proposé par l'entreprise. Plus précisément, ABC a su prendre des parts de marché au service du taxi traditionnel, en proposant un service de qualité supérieure, à un prix supérieur au service traditionnel sans toutefois être proportionnel au niveau d'augmentation de la qualité.

Nous avons pu repérer tous les facteurs de ce thème opérationnel :

- Orientation vers processus (f2.1) : *« ...c'est sûrement la puissance opérationnelle qui est la plus critique pour l'exploitation du big data, la réussite pour un projet axé sur big data... »* ; *« ...notre entreprise bénéficie de la forte compétence et expérience développée par l'entreprise-mère dans ses diverses activités, notamment dans la gestion de ses données et dans l'alignement de ses processus métiers avec ceux de big data ».*
- Intégration du business processus SI (f2.2) et management des données (f2.3) : *«...sûrement, pour exploiter le big data, l'entreprise doit avoir une maturité en management de ses processus..., le management de ses données : les opérations avec les bases de données pour qu'elles soient propres, la sécurité des données, leur stockage, etc... ».*

6.2.3 F3 : Confidentialité, sécurité et gouvernance des données

Le deuxième thème de la catégorie « opérationnel » concerne la confidentialité, sécurité et gouvernance des données. La gouvernance des données s'applique à l'échelle d'une organisation. Elle se définit comme l'ensemble des procédures mises en place au sein de cette organisation afin d'encadrer la collecte de données et leur utilisation. Il s'agit à la fois de respecter les obligations légales imposées, et d'instaurer un cadre permettant d'optimiser l'utilisation des données. Ce thème joue un rôle majeur dans la confiance que les clients accordent à une entreprise, et donc dans leur propension à lui confier des données personnelles. Par ailleurs, sans données, pas d'exploitation du big data et pour une entreprise comme ABC, pas de service de taxi !

Dans les propos de son directeur régional, nous avons pu repérer les deux facteurs contenus dans ce 2^e thème opérationnel :

- Confidentialité et sécurité (f3.1) et gouvernance des données (f3.2) : « ...nous stockons d'énormes masses de données personnelles, avant de les traiter, nous sécurisons leur confidentialité au niveau opérationnel... Une fois l'application téléchargée, le client doit accepter les modalités et les conditions d'emploi qui définissent particulièrement que l'utilisateur donne son accord à ABC de traiter son information personnelle, y compris, ses données personnelles. En plus, il donne son consentement aussi pour que cette information soit transmise aux partenaires de l'entreprise pour de futurs traitements... En plus, le passager accepte que toutes ses conversations avec des personnes de l'entreprise puissent être enregistrées... ».

Il apparaît clairement que chez ABC, la réglementation concernant la captation et l'utilisation des données personnelles est respectée. Par ailleurs, des mécanismes techniques et organisationnels sont déployés afin de sécuriser et protéger les données contre la destruction non autorisée, la perte, l'abus et des formes illégales de traitement.

6.2.4 F4 : Qualité des données

Le dernier thème de la catégorie « opérationnel » concerne la qualité des données. Pour une entreprise comme ABC, dont les données sont la principale ressource clé du modèle d'affaire, ce thème est fondamental. Dans ce thème-là, nous avons repéré les facteurs suivants :

- Qualité des données et qualité de l'information (f4.1) : « ...le point le plus délicat avec l'exploitation du big data chez nous, c'est qu'on n'a pas de contrôle sur la qualité des données dans leur ensemble... ».
- Complétude des données (f4.2) et cohérence des données (f4.7) : « ...une autre partie vient des sources comme les commentaires et les enregistrements des conversations, ou les déclarations... leur qualité est moins facile à vérifier, elles sont de la nature semi-structurée et non structurée et la cohérence n'est pas toujours possible à constater... » ; « ...l'idéal, c'est quand, par exemple, la note du client peut se compléter par un commentaire qui est prouvé par les empreintes structurées du GPS, mais ce n'est pas toujours le cas... ».

- Pertinence des données (f4.5) : « ...une autre partie des données qui vient des empreintes numériques des capteurs GPS, constitue la majeure partie pour l'exécution de l'activité principale de la firme, ce sont des données faciles, car elles sont pour la plupart de nature structurée... ».
- Exactitude des données (f4.6) : « ...une grosse partie des données est saisie dans la base par les clients et les partenaires eux-mêmes, mais ils disent ce qu'ils veulent... ».

En réalité, tous les facteurs de ce thème sont présents dans les propos du directeur régional d'ABC. Il précise ensuite que c'est la capacité de l'entreprise, en matière de mécanismes de supports technologiques, qui va permettre d'appliquer des processus de nettoyage des données, pour assurer la qualité nécessaire à une utilisation efficace du big data.

6.2.5 F5 : Aspect humain

L'aspect humain est le premier thème de la catégorie « mécanismes de support ». Il représente l'un des moteurs essentiels de l'efficacité d'usage du big data. C'est à travers le facteur humain que tous les autres thèmes sont maîtrisés. Dans ce thème, nous avons repéré les facteurs suivants :

- Savoirs et compétences des individus (f5.1) et champions (f5.3) : « ...c'est vraiment très difficile de trouver des spécialistes du domaine, des data scientists, des techniciens, des chefs de projet avec une expertise assez forte en big data... et l'on rencontre ce problème partout, en Arménie, en Russie, en Europe... Parfois on peut rechercher un spécialiste pendant des mois..., mais on recrute vraiment des professionnels de leur métier... ».
- Confiance (f5.2) et engagement des salariés (f5.4) : «...évidemment, il y a une relation symbiotique entre l'entreprise et ces spécialistes : ils sont persuadés que dans ce domaine, ABC possède une forte expertise et la compétence, qui assure que tout souci lié aux tâches recevra l'attention nécessaire et sera réglé, ce qui confirme aussi leur fort engagement... ».
- Participation des utilisateurs (f5.5) : « ...la plupart des individus recrutés sont chez nous des hommes dans la tranche d'âge 26-35. Même pour les chauffeurs partenaires ... On constate que les personnes plus âgées ont plus de mal avec les technologies novatrices, et ont de l'anxiété

à utiliser les nouveautés. Il est plus difficile de les inclure dans les processus... Dans un pays africain, par exemple, nous avons des cours préparatoires pour les chauffeurs pendant plus d'un mois afin de les former à utiliser l'application...et c'est vraiment compliqué... ».

Le facteur relatif aux savoirs et compétences des individus est considéré comme crucial dans les processus d'exploitation du big data. Il apparaît également problématique car il y a pénurie de spécialistes. Dans ce contexte, l'attractivité de l'entreprise, puis l'implication et l'engagement de ces spécialistes, lorsqu'ils ont pu être captés, reposent en partie sur leur confiance dans les compétences de l'entreprise en big data.

6.2.6 F6 : Aspect organisationnel

Dans ce 2^e thème de la catégorie de « mécanismes de support », nous avons repéré les facteurs suivants :

- Compétence organisationnelle (f6.1) et accent sur l'innovation (f6.13) : *«... ABC, née technologique a une expertise très forte dans le domaine qui assure une culture organisationnelle favorable à l'exploitation du big data. Sa mission, c'est d'aider ses consommateurs à mieux naviguer dans les mondes en ligne ou hors-ligne, en suivant et en utilisant les avancements technologiques dans les services et produits proposés. Sa stratégie, c'est de proposer un meilleur service grâce à la technologie... ».*
- Gestion des talents (f6.2), collaboration interdépartementale (f6.8), support de haute direction (f6.10), taille de l'entreprise (f6.7) : *« ...l'entreprise-mère joue un rôle majeur sur l'activité ABC en matière de partage des connaissances, d'expérience, des compétences, des bonnes pratiques, des perspectives de développement... Il s'agit bien d'un partage de ressources en particulier humaines et technologiques... ».*
- Alignement stratégique (f6.4), management du projet (f6.5), management de la performance (f6.6) : *« ...le big data est dans le cœur de l'activité d'ABC...on développe et poursuit le "best case practice" dans le Groupe et évidemment, on applique souvent les méthodes prédéfinies et valides qui assurent le succès... ».*

- Effet environnemental (f6.11) : « ... *soulignons, que c'est au niveau du Groupe que l'accent est mis sur l'innovation et que la veille environnementale est assurée...* ».

6.2.7 F7 : Systèmes, outils et technologies

Le dernier thème de la catégorie « mécanismes de support » dans l'efficacité d'usage du big data regroupe les facteurs « qualité du système », « infrastructure TI » et « support des fournisseurs ». Ces facteurs assurent, le soutien technologique nécessaire aux initiatives big data. Conformément au modèle de Surbakti et al. (2019), pour permettre une utilisation efficace du big data, les systèmes doivent être fiables, flexibles face aux nouvelles exigences ou conditions, permettre l'intégration des données provenant de différents ensembles, être facilement accessibles aux utilisateurs, et doivent fournir des données en temps opportun tout en protégeant les informations. Par ailleurs, les initiatives big data ont besoin d'outils et technologies appropriés pour aider les organisations à exploiter à leur profit les données volumineuses. Elles ont besoin d'une infrastructure informatique flexible pour s'adapter facilement aux exigences d'utilisations diverses des projets big data. Elles ont besoin de plates-formes robustes pour traiter de multiples sources de données.

Ainsi, nous avons repéré tous les 3 facteurs composant du thème :

- Actualité du système (f7.1) et infrastructure TI (f7.2) : « ...*notre entreprise a développé ses propres technologies de cartographie, de routage, de navigation qui assurent la qualité du système pour le traitement du BD et l'adéquation de son infrastructure SI aux besoins du service...* ».
- Support des fournisseurs (f7.3) : « ...*en utilisant nos propres technologies de routage, cartographie, navigation, nous enlevons presque le risque de dépendance de bon fonctionnement des fournisseurs clés...* ».

Ces résultats, ainsi que le cas de la société ABC sont bien détaillés dans notre article présenté lors du 25^e Conférence de l'AIM (Grigoryan & Lobre-Lebraty, 2020) (Annexe 21).

6.3 Résultat 2 : Deux modèles différents de l'efficacité d'usage du big data en entreprise en contexte arménien

Selon le type d'utilisateur, en contexte arménien, le modèle conceptuel principal de Surbakti et al. (2019) a été aménagé de deux manières suivantes : un modèle d'efficacité d'usage du big data pour un « utilisateur novice » et un modèle d'efficacité d'usage du big data pour un « utilisateur avancé ».

Ainsi, nous proposons dans les paragraphes suivants de présenter ces deux modèles.

6.3.1 Modèle d'efficacité d'usage du big data pour un « utilisateur novice »

Dans un contexte arménien, une entreprise de type « utilisateur novice » distingue les 29 facteurs qui contribuent le plus à l'efficacité d'usage du big data sur les 41 du modèle. La figure 42 illustre bien ces facteurs (1-29 par l'importance hiérarchique, où « 1 » est le plus important) regroupés en thématiques proposés de Surbakti et al. (2019).

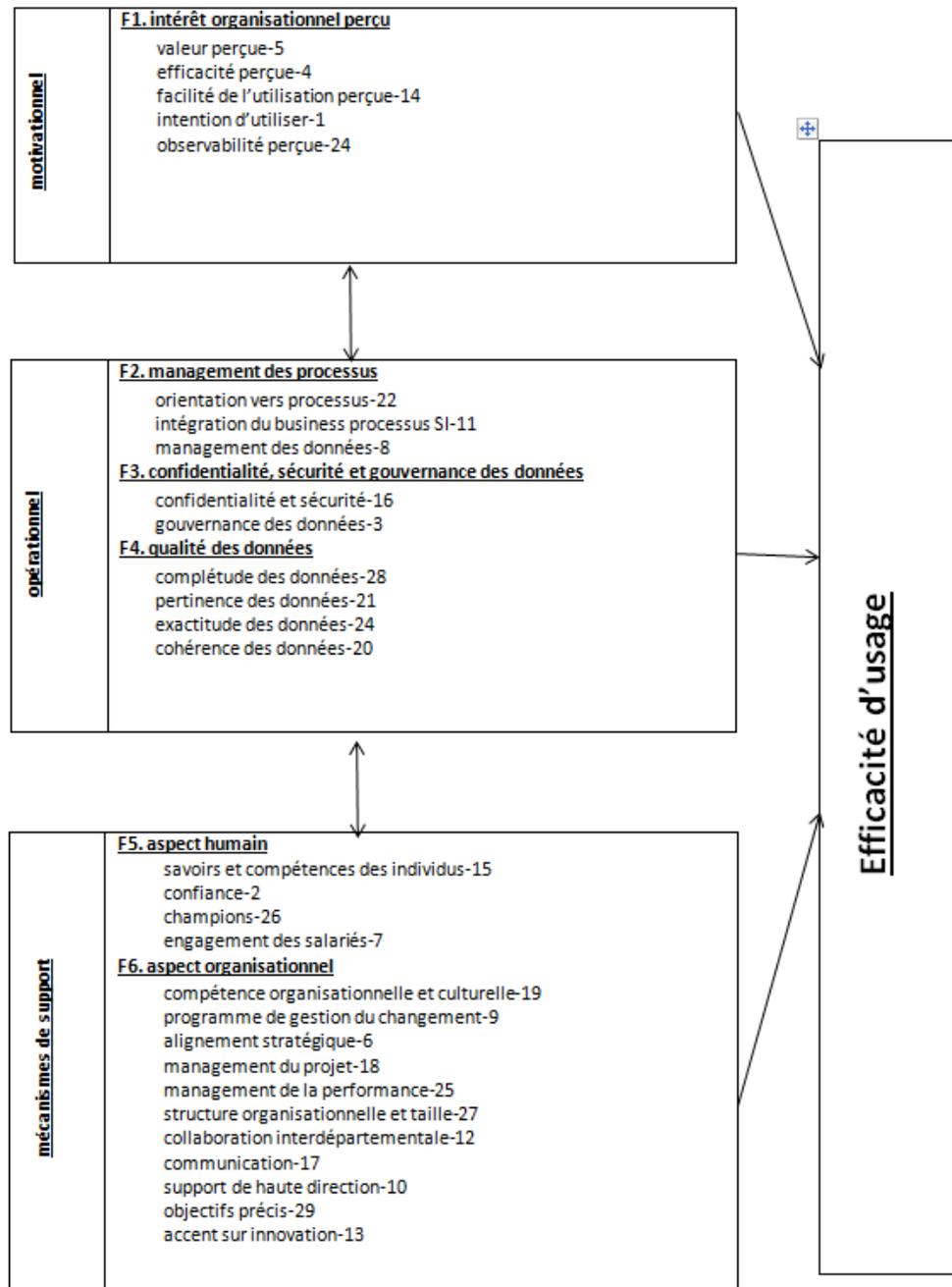
Nous pouvons bien noter que ces facteurs peuvent être présentés au niveau de 3 catégories suivantes :

- Motivationnel ;
- Opérationnel ;
- Mécanismes de support.

Ces résultats nous permettent de constater que pour une entreprise de type « utilisateur novice », l'efficacité d'usage du big data dépend d'abord des facteurs qui consistent l'intérêt perçu de l'organisation à exploiter le big data. Deuxièmement, c'est la confiance à l'information issue du traitement de big data. La gouvernance des données est traitée comme le troisième aspect important.

Ainsi, nous pouvons constater que pour un « utilisateur novice » l'efficacité d'usage du big data, se cache plutôt dans une chaîne des facteurs de type « motivationnel-mécanismes de support (aspect humain)-opérationnel ».

Figure 42 Modèle de l'efficacité d'usage du big data pour un utilisateur de type « utilisateur novice »



Pour ce groupe d'utilisateurs, nous avons identifié aussi des résultats surprenants. Lorsqu'on compare le modèle de l'efficacité d'usage du big data pour un utilisateur de type « utilisateur novice » avec le modèle conceptuel de Surbakti et al. (2019), nous pouvons toutefois remarquer que les 3 facteurs (« qualité du système », « infrastructure TI », « support des

fournisseurs ») regroupés dans le thématique « systèmes, outils et technologie » n'émergent pas parmi les 29 facteurs importants. Nous revenons à ce résultat surprenant et l'interprétons dans la partie de discussion de ce chapitre.

6.3.2 Modèle d'efficacité d'usage du big data pour un « utilisateur avancé »

Pour le deuxième type d'utilisateur, « l'utilisateur avancé » les 16 facteurs sur 41 composants du modèle conceptuel (Surbakti et al., 2019) ont été recensés à contribuer le plus à l'efficacité d'usage du big data en entreprise en contexte arménien. La figure 43 illustre bien le modèle qui valorise celui de Surbakti et al. (2019).

Ainsi, pour ce deuxième type d'utilisateur, nous constatons qu'au niveau des catégories des facteurs qui contribuent le plus à l'efficacité d'usage du big data il y en a deux :

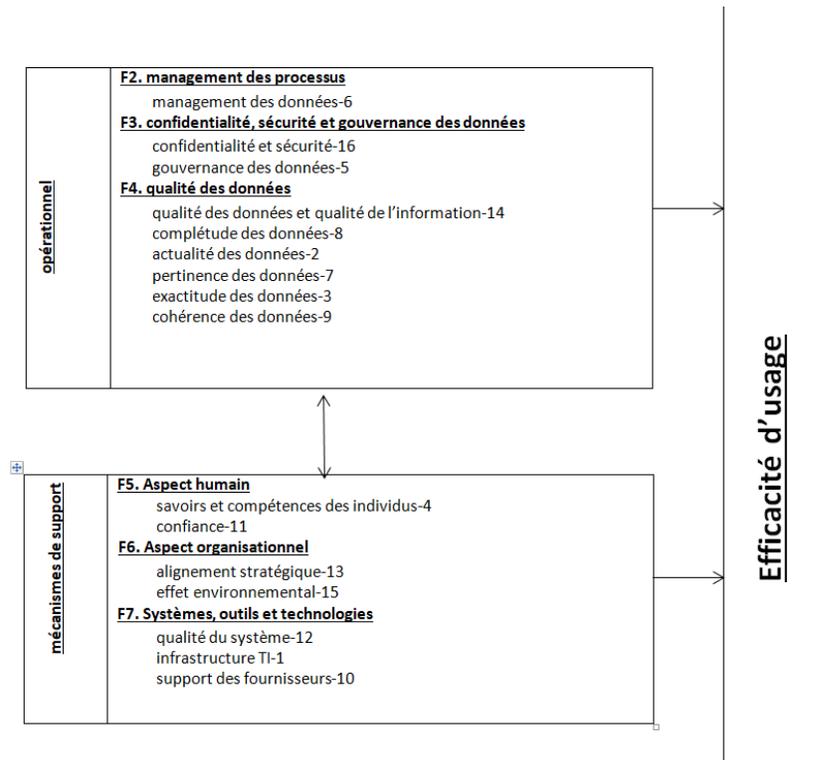
- Niveau opérationnel ;
- Mécanismes de support.

Nous constatons ainsi que pour un « utilisateur avancé » l'efficacité d'usage du big data dépend premièrement de sa capacité en infrastructure technologique.

Deuxièmement, au niveau opérationnel, il faut noter la qualité des données. Les compétences et les savoirs du personnel représentent le troisième aspect qui contribue à l'efficacité d'usage du big data chez ce type d'utilisateurs.

Pour ce deuxième groupe d'utilisateur aussi, nous avons identifié des résultats surprenants. Nous constatons que tous les 7 facteurs motivationnels, représentant le thématique « intérêt organisationnel perçu » (« valeur perçu », « risques perçus », « efficacité perçue », « facilité d'utilisation perçue », « intention d'utiliser », « observabilité perçue » et « analyse coût-avantage » n'émergent pas parmi les plus importants à contribuer à l'efficacité d'usage du big data. Ce résultat sera également interprété dans la section concernée de la partie de discussion.

Figure 43 Modèle de l'efficacité d'usage du big data pour un utilisateur de type « utilisateur avancé »



6.4 Résultat 2 bis : Efficacité d'usage du big data chez une entreprise « utilisateur avancé » et « utilisateur novice »

6.4.1 CDE, « utilisateur avancé »

Dans les propos de M D.D., le sénior scientifique des données chez CDE, nous avons pu repérer tous les 16 facteurs identifiés comme les plus importants selon les résultats de l'enquête quantitative.

Ainsi, dans les paragraphes suivants, nous allons présenter les verbatims qui illustrent le mieux ces résultats.

Systèmes, outils et technologie (F7) : (f7.1 « qualité du système » ; f7.2 « infrastructure TI » ; f7.3 « support des fournisseurs »)

« ...Dans le cas où une entreprise se déciderait d'initier une activité dont le cœur est l'utilisation du big data, la question de l'infrastructure technologique est un point de départ critique. ...D'abord, il faut comprendre que chez les grandes entreprises qui ont déjà une activité principale, à un moment donné elles peuvent discuter la question du traitement du big data

comme une opportunité, qui peut se faire ou pas ; une fois décidées à l'utiliser, ces grandes entreprises peuvent intégrer les gens et les solutions qui peuvent s'en occuper...

...Ce n'est pas facile à imaginer ce que représente le volume des données à stocker et traiter. Pour illustration de toutes les actions des utilisateurs en ligne à partir de leur chaque cliquet, la capacité infrastructurelle doit avant tout le permettre : naturellement ce volume des données demande des technologies spécifiques de traitement, une infrastructure technologique vraiment très puissante... ; ... Chez nous, les données sont considérées comme « big » à partir des outils de traitement et des outils de développement des modèles. Comme un scientifique des données, ayant une expérience du domaine de plus de 10 ans, j'ai beaucoup travaillé avec la base de données RedShift, qui permet d'y sortir les données aux différents paramètres et formats souhaités ; ainsi cette plate-forme permet de réaliser ce type d'actions plus de 10 fois plus rapidement qu'à partir des technologies traditionnelles... ».

Nous pouvons ainsi bien noter que lors de présentations de leur cas de l'efficacité d'usage du big data, M D.D. se focalise toujours sur les aspects de l'infrastructure technologique et de qualité du système technologique :

« ...Chez CDE, nous exploitons une infrastructure AWS (Amazon Web Services)⁵⁸ qui propose beaucoup d'outils nécessaires à traiter et à analyser des données (f7.3). Nos scientifiques des données largement utilisent aussi SageMaker⁵⁹ dans leurs tâches quotidiennes pour rendre le ML (Machine Learning) plus accessible à préparer les données à grande échelle, à accélérer le développement et rationaliser le cycle de vie de ML. Aussi, l'outil technique principalement utilisé par l'entreprise qui permet de travailler avec big data, c'est Spark. Il y en a 2 types, Spark à la langue de programmation Python – PySpark et par la programmation de Scala. Le dernier est exploité fréquemment... ».

Management des processus (F2) et Confidentialité, sécurité et gouvernance des données (F3)

Dans ces deux thèmes du modèle, nous avons les facteurs suivants :

⁵⁸ Une plate-forme cloud assez complète et largement adoptée au monde qui propose plus de 200 services issus de centres de données du monde entier.

⁵⁹ Proposé par Amazon, ce service permet aux scientifiques des données et aux développeurs de créer, former et déployer rapidement et facilement des modèles de machine learning à n'importe quelle échelle.

- Management des données (f2.3) ; Confidentialité et sécurité des données (f3.1) et Gouvernance des données (f3.2)

« Au niveau des régulations en matière de management et gouvernance des données, nous sommes conformes à 2 régulations courantes, le Règlement Général sur la Protection des Données (GDPR⁶⁰) et le CCPA (California Consumer Privacy Act). Ces deux régulations assurent le traitement sécurisé des données au niveau des lois et les réglementations. Le RGPD est un règlement adopté par le Parlement européen qui consiste des lois qui protègent les personnes concernées par un traitement de leurs données à caractère personnel et la responsabilisation des acteurs de ce traitement. En plus, nous sommes conformes à la loi arménienne sur les régulations de l'usage des données personnelles (Annexe 20) : il s'agit de la loi de régulation sur la sécurité et l'utilisation des données personnelles, que l'Arménie a adopté en 2015. Cette loi régleme principalement la procédure de traitement des données personnelles et l'exercice du contrôle de l'État sur celles-ci. Cette utilisation et le traitement des données personnelles doivent se faire en accord avec la personne concernée et pour des raisons légitimes. La loi comprend 29 articles sur les différents aspects sur l'utilisation, le traitement, le transfert des données personnelles. Mais toutefois, je peux noter et non seulement moi, plusieurs représentants du domaine TI notent, que cette loi a plusieurs défauts où plusieurs aspects de l'utilisation des données et leur traitement ne sont pas définis ; ... en même temps, l'absence de strictes restrictions peut se considérer comme une opportunité. Ça peut faciliter le processus de l'utilisation et du traitement des données si l'on se limite au marché local arménien... ».

- Qualité des données

Dans ce thème, nous avons exploré les 6 facteurs retenus comme importants : « qualité des données et qualité de l'information » (f4.1) ; « complétude des données » (f4.2) ; « actualité des données » (f4.3) ; « pertinence des données » (f4.5) ; « exactitude des données » (f4.6) ; « cohérence des données » (f4.7).

« ...A l'aide du code écrit par nos programmeurs, toutes les actions des utilisateurs (qui préalablement installent chez eux une extension ou une application) sont suivies. Ainsi, tous ces flux des données deviennent accessibles à l'entreprise (f4.3 ; f4.5). Ces données sont stockées dans les appareils ou l'ordinateur de l'utilisateur. Une fois stockées pour une périodicité fixée, les données sont transmises à l'entreprise au format "raw" qui sont après stockées dans une

⁶⁰ General Data Protection Regulation

base selon une certaine logique (f4.2 ; f4.6). Cette base de données peut être une base de données ordinaire ou une base de données destinée au big data, comme par exemple RedShift. Une fois les données matières sont collectées, elles passent par certains processus, ce que nous appelons les ‘‘jobs’’, dont s’occupent principalement les data ingénieurs qui créent un ‘‘pipeline’’⁶¹(f4.7). Nous utilisons également des bases de données qui sont mises à jour au temps réel et selon les cas d’utilisation, les données subissent certaines agrégations. Il est possible de les garder dans d’autres bases de données ou sur d’autres tableaux ; par exemple, toutes les données sont au niveau des cliques, et l’on veut transformer au niveau des utilisateurs, on veut comprendre les actions de certains utilisateurs ; ou bien, on peut faire certaines agrégations afin de convertir les données au format préféré (f4.1). Ainsi, les mêmes données peuvent être utilisées pour différentes raisons : moi, en tant que scientifique des données, je peux l’utiliser pour développer un ML modèle, dans un autre cas les données matières peuvent être vendues, dans un autre cas l’on peut le recenser ‘‘un insight’’ selon une requête client. Pour un autre cas, on peut réaliser un ‘‘ad measurement’’ (f4.3 ; f4.7). Ça peut se faire par l’utilisation d’un ‘‘pipeline’’ pour les utilisateurs afin de générer les données concrètes pour procéder à ces processus... ».

- Aspect humain (« savoirs et compétences » f5.1)

« ...mais le cas est très différent si toute l’activité cœur de l’entreprise consiste dans l’utilisation du big data. Donc, comme j’ai déjà noté l’autre fois, la première chose dont on s’occupe c’est l’infrastructure qui est très chère ; les questions qui poursuivent concernent si cette infrastructure « est hébergée » sur-site (‘‘on premises’’) ou hors site avec un Cloud. D’un côté, la solution peut paraître même plus chère, mais elle permet de se servir des solutions techniques plus facilement...Si pour un moment l’on imagine qu’une entreprise X désire créer une infrastructure locale, il est très compliqué en Arménie (c’est une question actuelle partout dans le monde, mais surtout en Arménie) de trouver des gens qui peuvent construire cette infrastructure sur-site, il y a beaucoup de manque des connaissances techniques ; souvent l’on peut chercher des mois quelqu’un qui connaît bien Spark, par exemple... ».

- Aspect organisationnel (« alignement stratégique » (f6.4) et « effet environnemental » (f6.11)

⁶¹ Également appelé pipeline de données, consiste d’un ensemble d’éléments de traitement de données connectés en série, où la sortie d’un élément est l’entrée du suivant.

« ... Le développement de nouvelles offres-produits à partir des données disposées est dans le cœur de la stratégie de notre firme (f6.4)...Ce choix est souvent conditionné par l'environnement concurrentiel qui nous dicte de différents lancements... (f6.11) ».

6.4.2 BCD Arménie, « utilisateur novice »

Dans les propos de M C.C., le chef de département DATASCIENCE chez BCD Arménie, nous avons pu repérer tous les 29 facteurs identifiés comme les plus importants selon les résultats de l'enquête quantitative.

Ainsi, dans les paragraphes suivants, nous allons présenter les verbatims qui illustrent le mieux ces résultats. Pour une illustration plus structurée, nous proposons de regrouper ces résultats en leurs 3 catégories : « motivationnel », « opérationnel » et « mécanismes de support ».

- Facteurs de caractère « motivationnel »

Dans notre discours avec M C.C., nous avons pu repérer les 4 facteurs motivationnels du modèle :

«...Il y a trois ans que nous exploitons pratiquement le big data. Je peux dire que toute cette histoire a commencé quand l'entreprise a appris que le concurrent fait un projet de lancement... (f1.5) oui, c'est vrai, la motivation majeure de cette initiative était avant tout la peur de ne pas retarder dans la concurrence ... » (f1.6). « ...il faut toujours se poser la question de quelle sera la valeur du big data pour tel ou tel projet de l'entreprise; car en matière de volume, chaque entreprise génère bien le « big data », et si à un instant l'on aperçoit que la question n'est pas encore bien définie ou conçue, pour laquelle il faut adopter l'infrastructure nécessaire du traitement du big data, donc, il n'a pas de sens pour l'entreprise...Donc, le point de départ est toujours la conscience du besoin justifié ; car même tels géants comme Facebook, Amazon, Google ne conserve et ne traite pas toutes leurs données générées ; ce n'est pas qu'ils ne peuvent pas les stocker, c'est une question de sens du point de vue du business qui manque... » (f1.1 ; f1.3).

- Facteurs de caractère « opérationnel »

Dans la présentation d'un des projets courant de l'entreprise comme un cas d'efficacité d'usage du big data, nous repérons les 4 facteurs concernant la qualité des données : « complétude des données » ; « pertinence des données » ; « exactitude des données » et « cohérence des données ».

« ... L'une des premières tâches consiste à collecter des données historiques concernées, particulièrement les données concernant le trafic en MB (Mégabyte) utilisé en fonction du temps. Donc c'est un processus en continu : ces données sont enregistrées sous les protocoles différents ; on comprend bien qu'une conversation zoom ou une conversation vidéo ont des protocoles différents. Une fois, les données sont collectées et stockées, l'on passe à l'étape suivante pour laquelle la question est définie de comprendre 3 questions : quelle part de données stockées est utile (1) ; quelle période est utile à considérer (2) et quelles sont les applications /les sources à considérer comme utiles (3), compte tenu de la question principale posée... » (f4.2 ; f4.5 ; f4.6 ; f4.7).

En matière de management des processus et de la gouvernance des données, nous avons repéré les interprétations de tous les 5 facteurs : « orientation vers processus » ; « intégration du business processus SI », « management des données » ; « confidentialité et sécurité » ; « gouvernance des données » :

« Au niveau de l'équipe du département de Data Management, nous recevons une requête, nous la complétons avec les exigences minimales pour passer un « visibility check » afin de comprendre si l'on peut recenser ces données...toutes ces données sont anonymes. Nos processus sont suivis par les 2 normes ; le GDPR (General Data Protection Regulation) et la loi locale arménienne...Après nos data ingénieurs procèdent au stockage de ces données, après un travail commun avec eux, nous recevons les données purifiées utiles pour alimenter notre modèle... (f2.1 ; f2.2 ; f2.3 ; f3.1 ; f3.2).

- Facteurs de caractère de « mécanismes de support »

Dans cette catégorie, nous repérons les interprétations sur les facteurs suivants :

- F5.1 « savoirs et compétences des individus » : *« le problème qui est toujours actuel est le manque des ressources intellectuelles et humaines qui sont spécialisées dans ce domaine. Ce problème était surtout grave au début de l'implantation ; où l'on était en recherche du cœur de l'équipe. Cette phase a duré presque 2 ans... ».*

- F6.4 « alignement stratégique » ; « collaboration interdépartementale »

« ...Nous sommes particulièrement intéressés par la possibilité de faire un regroupement des abonnés ; par exemple, si tel ou tel abonné est un cinéphile ou quelqu'un qui aime faire du shopping ; cette information nous aide à faire une offre bien ciblée avec une mise en

collaboration du département marketing. Ainsi, nous sommes particulièrement intéressés par la construction des macro-profils de nos clients abonnés : au niveau de notre métier, il s'agit de pouvoir les connaître et les décrire dans leur nature au maximum de dimensions possibles... Pour ce cas, l'entreprise collecte, le « streaming data qui est stocké dans un serveur intermédiaire, d'où l'on transmet dans une base de données et après c'est transmis dans HDFS où les données sont classées dans les fichiers Parquet⁶² ce qui réduit le volume en le compressant quelques fois...pour la suite déjà, les informations reçues des clusters sont utilisées pour avoir des insights, les perspicacités qui sont après utilisées largement dans les développements des produits, dans les stratégies de rétention, dans les stratégies d'acquisition, dans les ventes incitatives et croisées... ».

6.5 Résultats 3 : L'existence de deux types d'utilisateurs ne dépend pas du contexte d'examen

Notre troisième résultat est que l'existence de deux types d'entreprises utilisateurs, « utilisateur novice » et « utilisateur avancé » n'est pas une spécificité liée au contexte d'examen. Ainsi, dans les propos d'un expert à QRS Co., nous avons pu repérer la réponse suivante à ce questionnement : « ...cela a du sens que les facteurs qui influencent l'efficacité d'usage du big data peuvent ou doivent se différencier en fonction de la relation de la firme avec les données; /entreprises qui sont créés sur les données ou elles ont adopté une infrastructure appropriée à exploiter le big data à un moment donnée de leur activité principale/, comme le premier groupe des sociétés que vous mentionnez sont celles qui se procurent des données et les utilisent comme produit pour servir les autres, tandis que le second groupe de sociétés s'abonne à ces données pour atteindre leurs objectifs. Ainsi, les deux sociétés ont des mesures de réussite différentes en matière d'utilisation des données. Par exemple, une entreprise qui achète des données LinkedIn en tant que service (les récupère de LinkedIn et les vend à d'autres), ses mesures de réussite peuvent bien concerner ses efforts de placement de données. Alors que les entreprises, comme les Cabinets de conseil en gestion qui ont besoin de ces données pour leurs engagements, définissent l'efficacité de l'usage des données autour de mesures qui démontrent le succès de

⁶² Un fichier Parquet se compose d'un ou plusieurs groupes de lignes, un partitionnement logique des données en lignes.

l'engagement.... J'oserais supposer que cela a à voir avec le fait qu'ils ne sont pas des producteurs de données et que les données ne sont pas leur produit principal, mais plutôt un moyen d'atteindre un objectif. Par conséquent, ils disposent de plus de mesures commerciales concernant l'efficacité d'usage des données...), traduit de l'anglais'' « ... It does make sense since the first set of corporations you mentioned, are those who procure data and use it as product to serve others, while the second set of corporations subscribe to this data to achieve their goals. So both corporations have different success metrics of data usage. For example a company that procures LinkedIn data as their service (scrape it from LinkedIn and sell it to others), their success metrics would be around their data placement efforts. While companies, like Management Consultancies that need this data for their engagements, define their data usage effectiveness around metrics that demonstrate engagement success.... I would venture to assume that it has to do with the fact that they are not data producers and data is not their main product, rather it is a means to an end. Therefore, they have more business metrics around data usage... ''.

Ainsi, ce résultat vient d'affirmer que notre premier résultat (Résultat 1) n'est pas une spécificité liée au contexte d'examen.

6.6 Résultat 3 bis : Particularités contextuelles arméniennes

Lors de notre étude, nous avons pu repérer ainsi les 4 particularités contextuelles. Dans les propos des représentants, elles s'interprètent comme des éléments contextuels qui peuvent potentiellement impacter l'efficacité d'usage du big data en contexte arménien.

L'ensemble de 4 entretiens réalisés avec les représentants de 2 types d'entreprises utilisateurs du big data, nous ont permis de repérer les éléments contextuels suivants :

- **Compétences techniques et ressources humaines** : cet élément a été repéré dans les propos de tous les deux représentants des entreprises « utilisateur novice » et « utilisateur avancé ». C'était mentionné comme un défi sérieux du contexte arménien qui freine l'efficacité d'usage du big data.

« ... Il est très compliqué en Arménie (c'est une question actuelle partout dans le monde, mais surtout en Arménie) de trouver des gens de qualité...donc, beaucoup de manque des

connaissances techniques ; très souvent, on cherche des mois quelqu'un qui connaît bien Spark... » D.D., Société CDE.

« ...Nous sommes en recherche permanente des mecs compétents dans ce métier ; c'est une question toujours actuelle pour nous, pas juste pour nous. Mais c'est lié au marché limité de ressources humaines en Arménie... » C.C. ; Société BCD.

- **Réglementation législative arménienne :** cet élément a été repéré dans les propos de D.D. (Société CDE) comme un facilitateur de traitement du big data au niveau de processus de gouvernance des données.

« ...La réglementation législative locale ne définit pas de strictes restrictions sur l'usage et le traitement des données personnelles ; et si l'activité d'une firme se limite au marché local, cette condition facilite beaucoup l'utilisation... ».

Les deux entreprises utilisatrices du big data ont quand même adopté le RGPD afin de s'assurer contre des problèmes graves. Ainsi, cet élément peut être considéré comme une opportunité de la part d'une société, mais l'on comprend bien qu'il s'agit de la conséquence d'une loi incomplète arménienne qui a été adoptée en 2015.

- **Inertie de dupliquer :** cet élément intéressant a été repéré dans les propos de représentant de l'utilisateur novice, société BCD. Il s'agit de dupliquer des précurseurs. Comme c'était présenté par M C.C., l'intention d'adoption et de l'usage de l'infrastructure et du big data a été motivée par la peur de ne pas retarder les concurrents.

« ...ça peut vous paraître trop simpliste, mais la société BCD a adopté cette infrastructure quand son concurrent principal a commencé à utiliser pratiquement le big data... et c'est comme par inertie que presque 80% des acteurs dans notre secteur ont passée par cette « inertie de dupliquer » ...autre question combien d'entre eux l'utilise efficacement, sûrement moins que la moitié... ».

- **Marché limité :** cet élément a été repéré dans les propos du représentant de la société « utilisateur avancé ». Le marché limité est considéré comme une particularité contextuelle arménienne qui freine le développement de ce type utilisateur du big data.

« ...les entreprises comme la nôtre ne peuvent pas se limiter au marché local, comme j'ai déjà mentionné, nous avons des clients potentiels des Etats-Unis, de Californie particulièrement : nous avons adapté CCPA, (California Consumer Privacy Act) , qui est un règlement issu assez récemment sur la confidentialité des données qui régleme la

manière dont les entreprises du monde entier peuvent traiter les informations personnelles des résidents de la Californie... ».

6.7 Conclusion des résultats

Dans ce chapitre, nous avons présenté concrètement les résultats de 3 études réalisées. En effet, les deux résultats de l'étude quantitative permettent de répondre de façon complète à la question de recherche du travail. Les résultats des recherches qualitatives, de leur part, nous permettent d'interpréter l'ensemble des résultats et de comprendre l'effet du contexte arménien. Nous évoquons ci-après les six conclusions des résultats de ces trois études.

La première conclusion est que les sept thèmes contribuant à l'efficacité d'usage du big data (Surbakti et al., 2019) sont bien présents chez ABC. Autrement dit, dans ce cas bien précis, chaque thème semble bien contribuer à ce que l'usage du big data contribue aux objectifs d'ABC. Donc, le modèle conceptuel de Surbakti et al. (2019) peut être utilisé pour évaluer l'efficacité d'usage du big data en entreprise, utilisateur des données massives.

La deuxième conclusion est que l'efficacité d'usage du big data dans une entreprise telle qu'ABC (utilisateur avancé) est certainement valorisée autrement que chez une entreprise qui en possède une vision et des connaissances limitées (utilisateur novice).

La troisième conclusion est que parmi les entreprises qui exploitent du big data en Arménie, il y a 2 types différents d'utilisateurs : « utilisateur novice » et « utilisateur avancé ».

La quatrième conclusion est qu'à chaque type d'utilisateur, il y a une adaptation du modèle. Un « utilisateur novice » perçoit l'efficacité d'usage du big data dans les facteurs motivationnels comme l'intérêt organisationnel, les facteurs concernant des aspects humains et organisationnels. Ce type d'entreprises adopte une infrastructure à exploiter le big data afin de réaliser une valeur, conforme à MITMO (Moteur-Information-Technologie-Méthodes-Objectif) (Surbakti, 2020), présenté dans le chapitre de discussion de la thèse.

Ainsi, l'efficacité d'usage du big data chez « un utilisateur novice » en contexte arménien, dépend de 29 facteurs sur les 41 du modèle.

Par contre, pour un « utilisateur avancé », ce sont les facteurs liés aux mécanismes de support et les facteurs opérationnels qui contribuent le plus à l'efficacité d'usage du big data. Ce type d'entreprises, créées sur les données, les utilisent principalement pour assurer leur fonctionnement (niveau opérationnel).

Ainsi, l'efficacité d'usage du big data chez « un utilisateur avancé » en contexte arménien, dépend de 16 facteurs sur les 41 du modèle.

La cinquième conclusion est qu'en fonction de la relation de la firme avec les données (« utilisateur novice » ou « utilisateur avancé »), les facteurs qui influencent l'efficacité d'usage du big data peuvent et doivent être différenciés indépendamment du contexte d'examen.

La sixième conclusion est que l'efficacité d'usage du big data en contexte arménien met en avant les quatre particularités contextuelles. Elles sont bien liées à trois pôles de contextualisation proposés par Livian (2020). De nature culturelle, institutionnelle et d'un statut « post-soviétique », ces quatre particularités sont les compétences techniques et ressources humaines, la réglementation législative, le marché limité et l'imitation des précurseurs.

Ainsi, dans le chapitre suivant, le dernier de ce travail de thèse, nous allons répondre à la question de recherche et développer une discussion globale sur les résultats et le travail de recherche.

Chapitre 7 Conclusion et discussion

L'objectif principal de ce dernier chapitre est de répondre à la question principale de la thèse : « **Quels sont les facteurs qui contribuent le plus à l'efficacité d'usage du big data en entreprise en contexte arménien ?** ». Il s'agit de déterminer quels sont parmi les 41 facteurs composant du modèle conceptuel de Surbakti et al. (Surbakti et al., 2019) ceux qui contribuent le plus à l'efficacité d'usage du big data en entreprise en contexte arménien.

Ainsi, dans la première partie du chapitre, nous allons répondre à la question de recherche du travail. Dans une deuxième partie, nous développons déjà une discussion portée sur les trois axes suivants :

- Prolongement des résultats ;
- Place du contexte arménien dans cette étude ;
- Mise en perspective de notre recherche

Pour la suite, les contributions du travail seront détaillées. Nous finaliserons ce chapitre par la présentation des limites et des perspectives de recherche.

7.1 Réponse à la question de recherche

La question de recherche que nous avons formulée suite à la revue de littérature, était la suivante :

« **Quels sont les facteurs qui contribuent le plus à l'efficacité d'usage du big data en entreprise en contexte arménien ?** ».

Comme nous avons déjà vu, le premier grand résultat de l'étude quantitative suggère qu'il y a deux cas d'entreprises distincts : novice et avancé. Discutons ci-après chacun de ces cas.

7.1.1 Le cas de l'entreprise « utilisateur novice »

Dans un contexte arménien, pour une entreprise qui a commencé à exploiter le big data à un moment donné de son activité initiale (traditionnelle)⁶³, l'efficacité d'usage de ces données massives dépend de 3 catégories des facteurs. Par leur importance à contribuer à l'efficacité d'usage du big data en entreprise en contexte arménien, elles sont classées comme suit :

- Motivationnel ;
- Mécanismes de support ;
- Opérationnel.

Dans la catégorie motivationnelle, les 3 premiers facteurs les plus importants sont « l'intention d'utiliser », « l'efficacité perçue » et « la valeur perçue ».

Il faut noter que chez une entreprise « utilisateur novice » cette catégorie des facteurs motivationnels impacte l'ensemble de cycle de vie des étapes d'exploitation du big data (« l'adoption » et « l'utilisation potentielle »).

Donc pour une entreprise novice le point de démarrage pour l'efficacité d'usage du big data consiste d'abord dans l'intention d'entreprise d'utilisation du big data : dans le cas précis étudié chez une entreprise de télécommunications cette intention était liée à la peur de ne pas rater son concurrent. Dans leur cas d'efficacité d'usage du big data « segmentation des clients », l'entreprise perçoit l'efficacité comme une valeur qui contribue à l'augmentation de son CA de X%.

Dans la catégorie des facteurs liés aux mécanismes de support, les 3 facteurs les plus importants sont « la confiance » ; « l'alignement stratégique » et « l'engagement des salariés ».

Pour le cas étudié, tous les deux facteurs, « la confiance » et « l'engagement des salariés » étaient mentionnés comme critiques. Ici, il faut noter que la confiance concerne « l'output » du système et donc à l'information pour un décideur. Dans le cas de BCD Arménie,

⁶³ Dans ce contexte, une activité traditionnelle est celle qui n'est pas initialement basée sur l'exploitation du big data.

l'engagement du personnel dépend fortement du soutien de la part de la haute direction de la firme.

Utilisateur novice			
Mean	(f)Code	Rang	Factor
5.9578	F23	1	Intention to use
5.9276	F12	2	Trust
5.9121	F18	3	Data governance
5.826	F21	4	Perceived usefulness
5.7519	F19	5	Perceived value
5.7051	F32	6	Strategic alignment
5.6801	F14	7	Employee engagement
5.6353	F28	8	Data management
5.63	F31	9	Change management program
5.6206	F38	10	Top management support
5.6166	F27	11	IT BP integration
5.6125	F36	12	Interdepartmental collaboration
5.6117	F41	13	Focus on innovation
5.5914	F22	14	Perceived ease of use
5.5827	F11	15	Peoples knowledge and skills
5.5442	F17	16	Data privacy and security
5.535	F37	17	Communication
5.5324	F33	18	Project management
5.431	F29	19	Organizational cultural competence
5.4139	F10	20	Data consistency
5.3339	F8	21	Data relevance
5.3303	F26	22	Process orientation
5.3275	F24	23	Perceived observability
5.2572	F9	24	Data accuracy
5.2013	F34	25	Performance management
5.0818	F13	26	Champions
5.0797	F35	27	Organizational structure and size
5.0749	F5	28	Data Completeness
5.0085	F40	29	Clear goals

Tableau 12 Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur novice »

Le facteur « alignement stratégique » montre que l'efficacité d'usage du big data dépend du niveau de la maturité de l'usage du big data dans la stratégie de l'entreprise.

Dans la catégorie opérationnelle, les 3 facteurs les plus importants sont « la gouvernance des données », « le management des données » et « l'intégration du business processus en SI ».

Il faut noter que la gouvernance des données concerne les aspects juridiques et éthiques liés au traitement et l'usage des données. En Arménie, 3 régulations sont pratiquées : le RGPD (Règlement Général sur la Protection des Données), CCPA (*California Consumer Privacy Act*) et la loi locale (Annexe 20).

Le management des données est traité au niveau des processus. Il s'agit donc de l'ensemble des processus liés aux différentes tâches (jobs) pour traiter le big data (collecte, intégration, exécution.).

Et donc au niveau du management des processus, l'autre facteur qui est retenu comme important montre la capacité d'adaptation du business et du système informatique existant à la TI émergente.

Nous avons présenté les 3 facteurs les plus importants à partir d'un classement de 3 catégories.

Notre étude quantitative a retenu les 29 facteurs comme important à contribuer à l'efficacité d'usage du big data en entreprise en contexte arménien. Ainsi, le tableau 12 présente la liste complète de ces 29 facteurs.

7.1.2 Le cas de l'entreprise « utilisateur avancé »

Dans un contexte arménien, pour une entreprise qui est basée initialement sur l'idée d'exploitation du big data, l'efficacité de ces données massives dépend de 2 catégories des facteurs : premièrement, ce sont les facteurs liés de mécanismes de support, en matière de systèmes, outils et technologies ; et deuxièmement ce sont les facteurs de catégorie opérationnelle.

Dans la catégorie des facteurs de mécanismes de support, les 3 premiers facteurs les plus importants sont « l'infrastructure TI », « les savoirs et compétences des individus » et « le support des fournisseurs ».

Utilisateur avancé			
(f) code	Mean	Rang	Factor
F2	6.3442	1	Infrastructure IT
F6	6.0167	2	Data Currency
F9	6.0127	3	Data accuracy
F11	6.0001	4	Peoples knowledge and skills
F18	5.8663	5	Data governance
F28	5.8597	6	Data management
F8	5.81	7	Data relevance
F5	5.6859	8	Data Completeness
F10	5.6809	9	Data consistency
F3	5.5243	10	Vendor Support
F12	5.4691	11	Trust
F1	5.3647	12	System Quality
F32	5.1702	13	Strategic alignment
F4	5.1033	14	Data Quality
F39	5.0779	15	Environmental effect
F17	5.0051	16	Data privacy and security

Tableau 13 Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur avancé »

Dans le cas bien précis de CDE, les 3 facteurs ont été notés dans les propos de M D.D. comme les facteurs critiques qui impactent l'efficacité d'usage du big data dans leur entreprise. Pour une entreprise dont le cœur du métier est l'exploitation du big data, les capacités infrastructurelles et le support technique nécessaire assurent le fonctionnement quotidien de l'entreprise. Le manque de professionnels de ce métier est noté comme un défi sur le marché arménien. Ce facteur est ainsi considéré comme un des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un utilisateur avancé dans le contexte arménien.

Dans la catégorie opérationnelle, les 3 premiers facteurs les plus importants sont « la qualité des données », « la gouvernance des données » et « le management des données ».

Nous pouvons toutefois noter que dans cette catégorie opérationnelle « la gouvernance des données » et « le management des données » sont aussi les 2 facteurs importants comme dans le cas des entreprises de type « utilisateur novice ».

Ainsi, quel que soit le type d'utilisateur, le management des données au niveau des processus et la gouvernance des données au niveau de réglementation juridique et éthique sont les deux facteurs importants qui contribuent à l'efficacité d'usage du big data.

Comme les données sont dans le cœur de l'activité de l'entreprise d'utilisateur avancé, et son bon fonctionnement y est lié pratiquement, la qualité des données émerge comme le troisième facteur important qui contribue à l'efficacité d'usage du big data.

Nous avons présenté les 3 facteurs les plus importants à partir d'un classement de 2 catégories.

Notre étude quantitative a retenu les 16 facteurs comme important pour contribuer à l'efficacité d'usage du big data en entreprise en contexte arménien pour une entreprise « utilisateur avancé ». Ainsi, dans le tableau 13, nous présentons l'ensemble des 16 facteurs, classés selon leur importance.

7.2 Trois axes de discussion

Après avoir répondu à la question de recherche, nous proposons de développer ici une discussion qui portera sur les 3 éléments suivants :

- Prolongement des résultats ;
- Place du contexte arménien dans cette étude ;
- Mise en perspective de notre recherche

A la fin de la discussion de ces 3 éléments, nous présenterons les contributions de la thèse, ainsi que les limites et les perspectives de recherche.

7.2.1 Prolongement des résultats

Dans cette partie, nous allons nous focaliser sur l'interprétation des résultats surprenants. Pour construire une compréhension claire, nous essayons de les présenter dans un contexte de comparaison pour les deux types d'utilisateurs : « utilisateur novice » et « utilisateur avancé » (Annexe 17).

D'après l'étude quantitative, nous avons deux résultats nous paraissant surprenants :

- Les 3 facteurs de la catégorie de « mécanismes de support », liés aux « systèmes, outils et technologies » n'émergent pas parmi ceux qui contribuent le plus à l'efficacité d'usage du big data pour une entreprise de type « utilisateur novice » en contexte arménien (Annexe 15). Tandis que, ces 3 facteurs (« qualité du système », « infrastructure TI », « support des fournisseurs ») émergent comme les plus importants pour les entreprises de type « utilisateur avancé » en contexte arménien (Annexe 13).
- Pour une entreprise de type « utilisateur avancé » une grande majorité (11 facteurs sur 13) des facteurs liés aux aspects organisationnels (Annexe 16) n'émerge pas parmi les facteurs qui contribuent le plus à l'efficacité d'usage du big data en contexte arménien (Annexe 12).

L'analyse thématique de l'étude confirmatoire nous permet d'apporter 2 interprétations logiques à ces résultats surprenants :

7.2.1.1 Interprétation 1

Les entreprises créées sur les données (« utilisateur avancé ») les utilisent pour assurer leur fonctionnement au niveau opérationnel.

Comme cela a été interprété par un expert à QRS Co. (ayant participé dans notre étude qualitative confirmatoire), pour le cas des entreprises dont le produit principal sont les données, l'usage efficace est considéré comme un moyen opérationnel à atteindre leurs objectifs.

Donc, pour un utilisateur avancé, à toute étape de son activité, l'infrastructure et les ressources technologiques restent les éléments les plus importants pour l'efficacité d'usage du big data.

Cette nature justifie aussi l'absence de 11 facteurs sur 13 de nature organisationnelle parmi les facteurs qui contribuent le plus à l'efficacité d'usage du big data en contexte arménien.

Un utilisateur avancé, autrement dit, une entreprise née technologique possède dès sa naissance une infrastructure organisationnelle permettant l'usage efficace du big data (« compétence organisationnelle et culturelle », « collaboration interdépartementale », « accent sur innovation », etc.)

Si l'on reprend le cas bien précis de CDE, « utilisateur avancé », pour leur service « *ad measurement* », l'efficacité d'usage des données est considérée juste comme un processus opérationnel dans la chaîne de réalisation de valeur de ce service-là.

7.2.1.2 *Interprétation 2*

Les entreprises qui adoptent une infrastructure TI permettant d'exploiter leur big data à un moment donné de leur activité principale, visent *un objectif de réaliser une valeur conforme à MITMO* (Moteur-Information-Technologie-Méthodes-Objectif) (Surbakti, 2020).

Comme l'a interprété un expert de QRS Co. (ayant participé dans notre étude qualitative confirmatoire) dans le cas des utilisateurs novices, les entreprises cherchent à créer des opportunités et une valeur commerciale dans l'exploitation du big data. Ce qui encore justifie les différentes valorisations du modèle conceptuel principal.

Mais avant de passer à détailler cette interprétation, nous allons, décrire la chaîne MITMO (Moteur-Information-Technologie-Méthodes-Objectif), proposé de Surbakti (2020).

La réalisation de valeur conforme à MITMO (Moteur-Information-Technologie-Méthodes-Objectif) est un ensemble de cinq thèmes à partir desquels Surbakti (2020) propose de définir l'efficacité d'usage du big data :

- M (Moteur) – il s'agit d'abord de l'exploitation de la valeur du big data par la définition du problème et la mise en évidence de sa pertinence pour l'organisation ;
- I (Information) – il s'agit d'apporter une solution à l'aide des informations fournies par les données;
- T (Technologie) – il s'agit de bâtir une infrastructure technologique pérenne, pouvant être mise à l'échelle ;
- M (Méthodes) – la définition du problème doit prendre en compte le niveau de flexibilité et de viabilité infrastructurelle de l'entreprise ;
- O (Objectif) - apporter une solution efficace, capable de résoudre la question posée.

A partir des propos du représentant de l'entreprise « utilisateur novice » nous avons repéré les 3 points suivants :

- Les tâches décrites dans leur cas bien précis de l'efficacité d'usage, « regroupement des clients », suivent une chaîne logique de MITMO (Moteur-Information-Technologie-Méthodes-Objectif) ;
- La valeur réalisée de cette activité est traduite en terme d'augmentation du CA de x% ;
- En présentant leur cas d'efficacité d'usage, le représentant présente les 2 phases : adoption et utilisation pratique.

7.2.2 Place du contexte arménien dans cette étude

Un des principaux enjeux du problème, qui fait que les entreprises ne réalisent pas une valeur commerciale de l'implantation du big data, *même en imitant les précurseurs qui ont réussi* (Kiron, 2017), est qu'elles ne fonctionnent pas dans les mêmes environnements. Il est nécessaire d'avoir une approche différente concernant la façon dont les entreprises investissent et déploient leurs ressources pour la réalisation de la valeur de big data (Abbasi et al., 2016).

Ainsi, nous sommes d'accord avec Mikalef et ses coauteurs (Mikalef et al., 2019a), dans l'idée que pour réussir un projet du big data, il faut prendre en compte les particularités contextuelles dans lesquelles les entreprises opèrent.

D'après les résultats de notre dernière étude qualitative, nous avons pu repérer 4 éléments particuliers du contexte arménien lié à l'efficacité d'usage du big data en entreprises :

- Compétences techniques et ressources humaines ;
- Réglementation législative arménienne ;
- Marché limité ;
- Imitation des précurseurs.

Nous allons discuter ces particularités contextuelles de point de vue de 3 pôles de contextualisation proposés et adaptés de Livian (2020) : pôle culturaliste, pôle institutionnaliste, pôle statut « post soviétique ».

- **Compétences techniques et ressources humaines**

A ce stade de connaissances, nous pouvons interpréter le manque de ressources humaines dans les métiers liés au big data au niveau du pôle institutionnaliste.

Evoquons ici un bref résumé sur le profil institutionnel arménien lié à la formation des métiers en TI.

Le système arménien d'établissements d'enseignement et d'instituts de recherche et de développement est dispensé par 4 types de structures d'enseignement supérieur : Universités, Instituts, Académies et Conservatoires. Actuellement en Arménie fonctionnent 26 établissements d'enseignement supérieur soutenus par l'Etat et 33 établissements privés. Ce réseau a permis le développement de nombreuses industries de technologie de pointe, expliquant la croissance substantielle de l'économie de l'Arménie pendant ces dernières décennies⁶⁴.

Pourtant, le système de formation reste incapable de satisfaire la demande du marché de ressources humaines pour les métiers technologiques.

Nous avons repéré une statistique issue de l'Agence Nationale de la Statistique, qui montre bien la déficience du système d'enseignement (2019-2021). Alors, le secteur technologique en Arménie est capable de recruter 5 000 nouveaux employés par an. Cependant, les universités et les institutions de formation en Arménie ne sont capables de former que 1 300 diplômés en TI par an. Pire encore, seulement la moitié de ces diplômés sont qualifiés pour entrer sur le marché du travail sans formation complémentaire. Cela signifie que le système d'enseignement arménien ne répond qu'à un peu plus de dix pour cent de la demande du marché de TI⁶⁵.

Notons, que le secteur des technologies de l'information dans son ensemble emploie 15 350 personnes et vaut 765,1 millions USD, et exporte pour 338,6 millions USD. Dans un pays où les salaires moyens sont d'environ 400 USD par mois et le PIB par habitant est de 3 880 USD par

⁶⁴ <https://armstat.am/en/>

⁶⁵ <https://armstat.am/en/>

an, le secteur technologique permet aux jeunes adultes de gagner des milliers de dollars par mois et a une production industrielle par habitant de 40109 USD par an⁶⁶.

Donc, nous constatons bien que le manque de ressources humaines dans les métiers technologiques est le vrai défi qui freine le développement du secteur. Cette incapacité au niveau du pôle institutionnaliste doit attirer l'attention des structures gouvernementales pour pouvoir atteindre leurs objectifs ambitieux fixés pour ce secteur.

- **Réglementation législative arménienne et marché limité**

Nous proposons de discuter ces 2 aspects ensemble pour 2 raisons : premièrement, ils sont interconnectés dans notre recherche qualitative et deuxièmement, ils sont liés tous les deux au pôle « institutionnaliste ».

L'analyse thématique de notre dernière étude qualitative nous permet de faire 2 conclusions :

Conclusion 1 : Même si le RGPD s'applique, les représentants des 2 types d'entreprises, « utilisateur novice » et « utilisateur avancé » constatent que l'absence de strictes restrictions dans la réglementation législative arménienne peut être considérée comme une opportunité pour l'entreprise qui traite le big data et limite en même temps son activité au marché local.

Conclusion 2 : Une entreprise de type « utilisateur avancé » qui est basée sur les données, ne peut pas se limiter au marché local.

Ces deux conclusions (opportunité et développement international) peuvent servir de base à des recherches en management international, mais aussi constituent des lignes directrices pour les managers de ces entreprises.

Sur le plan de la structuration du marché (aspect du pôle institutionnaliste), l'Arménie avait développé un secteur industriel moderne sous l'ancien système soviétique de planification centrale, fournissant des machines-outils, textiles et autres produits manufacturés aux républiques sœurs en échange des matières premières et d'énergie. Depuis l'effondrement de l'URSS en décembre 1991, l'Arménie s'est convertie à l'agriculture de petite taille, loin de grands

⁶⁶ <https://hti.am/>

complexes agro-industriels de l'époque soviétique. L'Arménie a eu une longue histoire en tant qu'un centre technologique et scientifique au sein de l'Union Soviétique. C'est pourquoi, en 2008, le gouvernement arménien a annoncé un plan de route de 10 ans pour créer un écosystème technologique doté d'infrastructures, de capitaux et d'établissements d'enseignement nécessaires qui permettraient au domaine de devenir le fondement et la force motrice de l'économie. Cela fait une dizaine d'années que l'industrie a recommencé à se remettre sur pied. Aujourd'hui, le potentiel technologique et scientifique compose une des ressources importantes du développement économique de l'Arménie.

- **Imitation des précurseurs**

L'imitation des précurseurs est une pratique dans un monde d'affaires : dans la littérature sur l'adoption du big data, nous constatons que les échecs des projets du big data sont liés à *l'imitation des pratiques des entreprises qui ont réussi dans l'exploitation du big data* (Kiron, 2017), *l'imitation sans tenir compte des particularités contextuelles* (Mikalef, Pappas, Krogstie, & Giannakos, 2018).

L'entreprise « utilisateur novice » ayant participé à notre étude qualitative constate que l'intention d'adoption et d'usage du big data a été liée d'abord à la peur de la haute direction de ne pas prendre du retard sur les concurrents donc, une imitation de son concurrent précurseur. Selon ses mots, c'est le cas d'autres acteurs du secteur. Pourtant, entreprise X et Y ont constaté de grands échecs dans cette implantation.

Il est difficile à ce stade de développement de définir si cette imitation des précurseurs prend ses racines dans les aspects du pôle culturaliste ou c'est une inertie organisationnelle.

En tout cas, c'est un résultat intéressant qui reste à explorer.

7.2.3 Mise en perspective de notre recherche

La littérature sur le sujet du big data met l'accent sur les caractéristiques des entreprises qui l'adoptent à un moment donné dans leur activité principale (Batko & Ślęzak, 2022), (J. Wang et al., 2022), (Hamdam et al., 2022), (Ghasemaghaei, 2021), (Elisabetta Raguseo et al., 2020), (Mikalef, Krogstie, Pappas, & Pavlou, 2020), (Wamba et al., 2020), (Dubey et al., 2020), (Dubey, Gunasekaran, Childe, Fosso Wamba, et al., 2019), (Raguseo, 2018), (P. S. Fosso

Wamba, 2017). La problématique globale de ces travaux consiste dans l'étude de la réalisation de valeur commerciale des opportunités du big data. L'efficacité d'usage du big data représente le cœur de la réalisation de valeur du big data, pourtant elle reste rarement traitée dans la littérature et très peu explorée empiriquement. De plus, beaucoup de travaux au sujet de l'impact du big data sur la performance de la firme essaient de mettre en évidence les résultats positifs ou négatifs liés à l'usage du big data, mais ce qui représente l'efficacité d'usage du big data, reste intouchable (Behl, 2022).

Dans leur théorie de l'efficacité d'usage, Burton-Jones et Grange (2013) distinguent 2 types d'objectifs : l'efficacité d'usage, qui est considérée comme un objectif moyen, et la performance, comme un objectif global. Notre travail de recherche s'est focalisé donc sur les aspects les plus fins qui représentent l'efficacité d'usage, autrement dit « l'objectif moyen ».

Pour pouvoir aussi interpréter le contexte arménien dans cette étude sur l'efficacité d'usage du big data et situer ce travail sur les recherches du domaine, nous proposons de :

- Présenter dans cette section une synthèse des résultats des études de 4 travaux qui nous intéressent particulièrement dans le domaine : (Surbakti, 2020), (F. Sejahtera, Wang, Indulska, & Sadiq, 2018), (Mikalef et al., 2019a), (Maroufkhani, Wan Ismail, et al., 2020) et (Maroufkhani, Tseng, et al., 2020) ;
- Comparer les résultats de nos études avec ceux de ces 4 travaux.

Dans le tableau récapitulatif (Tableau 14), nous avons synthétisé les résultats des études empiriques. L'une de ces études porte sur la définition de l'efficacité d'usage du big data (1), l'autre sur les facilitateurs et les inhibiteurs de l'efficacité d'usage du big data (2) et 2 études examinent la réalisation de valeur du big data. Parmi elles, le travail de Mikalef et ses coauteurs (Mikalef et al., 2019a) insiste sur le contexte d'examen, qui nous intéresse en particulier.

Tableau 14 Présentation synthétisée des résultats des études sur l'efficacité d'usage du big data et la réalisation de valeur

<p>Référence (Surbakti, 2020)</p> <p>Problématique Etude sur la proposition d'une définition au concept de l'efficacité d'usage du big data</p> <p>Fondements théoriques Théorie d'efficacité d'usage du SI (TEU) (Burton-Jones & Grange, 2013)</p> <p>Méthodes utilisées Etude qualitative exploratoire à partir des entretiens semi-directifs</p>	<p>Résultats Etude exploratoire avec 12 participants représentant les 8 organisations Définition de l'efficacité d'usage du big data à partir de 5 thèmes :</p> <ul style="list-style-type: none"> - Driver (Moteur), définir un problème qui apportera une valeur à l'entreprise ; - Information, transformation des données nécessaires aux informations actionnables à résoudre un problème posé; - Technology (Technologie), mise en place de l'infrastructure nécessaire au support de processus efficace de transformation des données aux informations actionnables ; - Methods (Méthodes), définir le problème, prise en compte le niveau de flexibilité et viabilité de son infrastructure ; - Goal (Objectif), apporter une solution efficace, capable à résoudre la question posée.
<p>Référence (F. Sejahtera et al., 2018)</p> <p>Problématique Etude sur les facilitateurs et les inhibiteurs de l'efficacité d'usage du big data</p> <p>Fondements théoriques Théorie d'efficacité d'usage du SI (TEU) (Burton-Jones & Grange, 2013)</p> <p>Méthodes utilisées Etude de cas exploratoire</p>	<p>Résultats Etude de cas exploratoire, 5 participants d'un projet qui collecte des données au travers d'une université australienne 4 principaux facilitateurs retenus :</p> <ul style="list-style-type: none"> - Capacités adéquates du système - Culture de collaboration établie - Bonne attitude de travail - Champions <p>11 principaux inhibiteurs retenus :</p> <ul style="list-style-type: none"> - Mauvaise qualité de données - Manque de compréhension des données - Silos des données (données isolées) - Manque du temps - Absence d'analyse coûts-avantages - Manque de soutien par la haute direction - Manque des compétences techniques - Manque des processus établis du management des données - Manque du personnel compétent - Manque des objectifs clairs - La sécurité et la confidentialité de mauvaises données
<p>Référence (Mikalef et al., 2019a)</p> <p>Problématique La relation des éléments contextuels avec les facteurs influençant la réalisation de la valeur du big data en entreprise</p> <p>Fondements théoriques Théorie de la complexité (Cabrera, Cabrera, Powers, Solin, & Kushner, 2018)</p> <p>Méthodes utilisées Approche mixte : quantitative (méthode</p>	<p>Résultats Etude quantitative, réalisée dans les entreprises grecques <i>Différents scénarios contextuels à la réalisation de valeur de l'usage du big data</i></p> <ul style="list-style-type: none"> - <i>Grandes entreprises et PME fonctionnant dans des contextes à fort dynamisme et d'absence d'hétérogénéité (secteur de biens de consommation, de médias, transport)</i> <p><i>Aspects essentiels de réalisation de valeur du big data :</i></p> <ul style="list-style-type: none"> • « les données » • « les ressources technologiques » <p><i>Aspects nécessaires :</i></p> <ul style="list-style-type: none"> • « les compétences techniques » <p>Elément non-important dans les PME :</p> <ul style="list-style-type: none"> • « Etablissement des pratiques procédurales » - Entreprises fonctionnant dans un environnement de « haute incertitude », dans des conditions de dynamisme et d'hostilité élevés (industries de pétrole et gaz,

d'enquête) et qualitative (entretiens semi-directif)	<p>ainsi que ceux du secteur bancaire et secteur financier)</p> <p><i>Aspects essentiels de réalisation de valeur du big data :</i></p> <ul style="list-style-type: none"> • « les données » • « les compétences techniques » <p><i>Aspects nécessaires :</i></p> <ul style="list-style-type: none"> • « les compétences managériales » <p><i>Éléments critiques :</i></p> <ul style="list-style-type: none"> • « les pratiques structurelles et procédurales » • « la culture de l'entreprise axée sur les données » <p><i>Élément moins important :</i></p> <ul style="list-style-type: none"> • « les ressources technologiques » <p><u>Résultats</u></p> <p><i>Etude qualitative</i></p> <p>3 éléments essentiels à la réalisation de valeur, recensés des entretiens :</p> <ul style="list-style-type: none"> • Stratégie big data • Inertie organisationnelle • Etique et réglementation législative
<p>Référence (Maroufkhani, Wan Ismail, et al., 2020) ; (Maroufkhani, Tseng, et al., 2020)</p> <p>Problématique Etude de l'adoption et l'usage de BDA (BIG DATA Analytics) par les PME en Iran au travers d'un modèle axé sur technologie-organisation-contexte environnementale (TOE ; (Tornatzky & Klein, 1982)</p> <p>Fondements théoriques Théorie du management par les ressources (Barney, 2001)</p> <p>Méthodes utilisées Etude quantitative, PLS-SEM appliqué</p>	<p><u>Résultats</u></p> <p>Etude dans les PME iraniens du secteur manufacturier</p> <p>L'influence sur l'adoption et l'usage du BD / facteurs par 3 thèmes</p> <p>Contexte Technologique</p> <ul style="list-style-type: none"> - avantage relatif (no impact) - compatibilité (no impact) - complexité (effet négatif) - risque et insécurité (effet négatif) - possibilité d'essai (effet positif) - observabilité (effet positif) <p>Contexte Organisationnel</p> <ul style="list-style-type: none"> - support de haute direction (facteur critique) - préparation organisationnelle (lien important) <p>Contexte Environnemental</p> <ul style="list-style-type: none"> - pression concurrentielle (no impact) - soutien extérieur (lien important) - règlements gouvernementaux (no impact) <p>Ces résultats ont des contradictions avec des résultats d'études sur les sujets similaires (Lai, Sun, & Ren, 2018), (Verma & Bhattacharyya, 2017), (D. Q. Chen et al., 2015), (Ghobakhloo, Sadegh Sabouri, Sai Hong, & Zulkifli, 2011) ce qui se justifie par les particularités contextuelles iraniennes (spécificité du marché iranien, conséquence des sanctions du pays, effets mineurs de globalisations, manque d'investissements internationaux..).</p>

7.2.3.1 Étude sur la proposition d'une définition du concept de l'efficacité d'usage du big data

Comme déjà noté, l'efficacité d'usage du big data est très peu explorée dans la littérature. Les résultats de 2 premières études empiriques présentées dans le tableau 14, ont été réalisées par des auteurs du modèle conceptuel, Surbakti et ses coauteurs (Surbakti, 2020), (F. Sejahtera et al., 2018). Ces 2 études empiriques sont parmi les rares qui développent le fondement théorique de l'efficacité d'usage du big data. La théorie de l'efficacité d'usage (TEU) développée au niveau du SI et appliquée au MIS (Management du Système d'Information) par les auteurs Burton-Jones

et Grange (Burton-Jones & Grange, 2013) consiste bien en une base théorique essentielle pour développer ce concept dans le domaine du big data. Mais, comme elle l'étudie au niveau du système et ne traite pas les caractéristiques spécifiques du big data, il est nécessaire qu'elle soit adaptée à ce contexte.

A cette tâche importante ont été consacrés les travaux de Feliks Prasepta Sejahtera Surbakti (2020) après le développement du modèle conceptuel de l'efficacité d'usage du big data (Surbakti et al., 2019).

Dans ce travail, F. P. S. Surbakti propose donc une première définition consensuelle de l'efficacité d'usage du big data au niveau organisationnel : sa contribution théorique consiste en la proposition d'une définition qui est fondée sur un raisonnement empirique et éclairée par la théorie (Burton-Jones & Grange, 2013) et (De Mauro, Greco, & Grimaldi, 2015), (De Mauro et al., 2016).

Le processus de l'efficacité d'usage du big data est présenté par F. Surbakti (2020) comme un ensemble de 5 éléments importants que nous avons schématisés dans la figure 44.

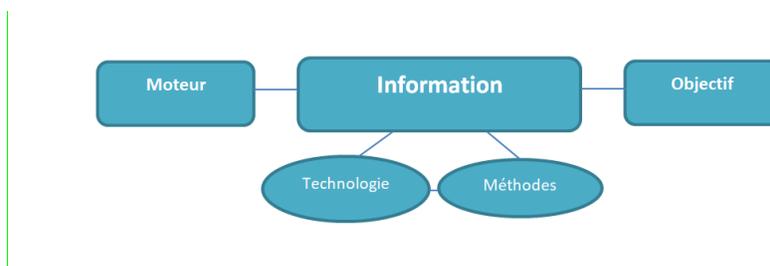


Figure 44 Éléments-composants de l'efficacité d'usage du big data en entreprise

Nous constatons que ce schéma peut être appliqué à chacun des exemples bien précis de l'efficacité d'usage du big data des 2 types d'utilisateurs entreprises, décrits dans le chapitre 5. Nous constatons ainsi que les interprétations fournies par les répondants, à partir des résultats de l'enquête quantitative, viennent principalement soutenir dans sa globalité la définition de l'efficacité d'usage du big data proposée par F. Surbakti (2020).

Nous suggérons ci-après une traduction de l'anglais de la définition consensuelle du concept de l'efficacité d'usage du big data, proposée par F. Surbakti : « *L'efficacité d'usage du big data se réfère à l'exploitation de la valeur du big data par la définition du problème et la*

mise en évidence de sa pertinence pour l'organisation, afin d'apporter une solution à l'aide des informations fournies par les données tout en bâtissant une infrastructure technologique pérenne et pouvant être mise à l'échelle. »⁶⁷.

Pour une société de type « utilisateur novice » comme c'est le cas d'entreprise BCD, la proposition de définition de F. Surbakti semble pouvoir s'appliquer. Donc, il s'agit bien de la réalisation de valeur pour l'entreprise grâce à l'usage de l'information issue de l'exploitation du big data.

Pour l'autre cas de type « utilisateur avancé », l'efficacité d'usage du big data peut bien constituer un moyen important, un objectif moyen (Burton-Jones & Grange, 2013), mais pas suffisant à la réalisation de sa propre valeur. Par exemple, chez ABC, l'efficacité d'usage du big data est considérée juste comme un moyen opérationnel de prestation de leur service, chez CDE, l'efficacité d'usage du big data « *...est un bloc nécessaire et important pour le fonctionnement de notre firme, et nous jouons un rôle intermédiaire dans ce processus, à partir des données que nous 'vendons' l'information dont la partie tierce peut profiter pour la réalisation de valeur...D.D.* ». Ainsi, dans ce deuxième cas, nous voyons que ça peut être aussi bien lié à la réalisation de valeur pour la partie tierce, où l'utilisateur avancé occupe juste un rôle « d'intermédiaire modérateur ».

Ainsi, en l'état, nous constatons que la définition de l'efficacité d'usage du big data peut être appliquée pour les sociétés de type « utilisateurs-novices ». En revanche, pour les « utilisateurs avancés », cette définition n'est adaptée que partiellement, nécessitant de futurs ajustements et raffinements.

7.2.3.2 Étude sur les facilitateurs et les inhibiteurs de l'efficacité d'usage du big data

La deuxième étude sur le sujet réalisée par Surbakti et ses coauteurs (F. Sejahtera et al., 2018) essaie de mettre en évidence les éléments facilitateurs et les éléments inhibiteurs de l'efficacité d'usage du big data à partir d'une étude de cas exploratoire.

⁶⁷ Définition proposée de F. Surbakti : *“Effective use of Big Data refers to harnessing value from Big Data by defining the problem, showing the value to the organization, and making solution by using insights from data and build Big Data infrastructure and data analytics by considering scalability and sustainability.”* (Surbakti, 2020)

Les 4 principaux éléments facilitateurs retenus par cette étude sont « les capacités adéquates du système », « la culture de collaboration établie », « la bonne attitude de travail » et « les champions ».

Les résultats issus de notre étude quantitative sont consistants partiellement avec les conclusions de l'étude de Surbakti et ses coauteurs (F. Sejahtera et al., 2018). Ainsi, pour le groupe « utilisateur novice » trois de ces éléments facilitateurs (« la culture de collaboration établie » (F29-19), « la bonne attitude de travail » (F37-17) et « les champions » (F13-26)) émergent comme les facteurs qui contribuent le plus à l'efficacité d'usage du big data au contexte arménien, tandis que les aspects du système et technologie ne sont pas recensés comme les plus importants à contribuer l'efficacité d'usage du big data (Annexe 12). Ils sont en effet considérés comme les aspects critiques pour la phase de l'adoption. Ce résultat met en évidence à nouveau que la mauvaise pratique répandue dans ce domaine, due à une compréhension limitée, pousse la forte concentration sur les aspects technologiques. Elle résulte d'une négligence de l'importance des facteurs organisationnels, considérée comme la raison principale des projets « fiasco » du big data (Kiron, 2017), (Brynjolfsson & McAfee, 2012).

Tandis que les entreprises, qui génèrent une vraie valeur commerciale de l'exploitation du big data, notent bien d'autres facteurs que nous avons détaillés dans la section précédente comme importants « *...le problème est posé, notre tâche, c'est produire l'information basée sur l'évidence pratique, nous sommes conformes avec les contraintes, mais le reste qui dépend bien d'autres participants...et si nous manquons d'une coopération forte entre les départements et si le chef du département préfère revenir à ces pratiques traditionnelles, ce qui arrive très souvent et c'est déjà la tâche d'une direction de plus hauts niveaux de mettre en place une stratégie permettant de rendre ces « informations actionnables... » C.C., BCD.*

Les auteurs ont identifié les 11 éléments inhibiteurs de l'efficacité d'usage du big data :

- « Mauvaise qualité des données » ;
- « Manque de compréhension des données » ;
- « Silos des données ou données isolées » ;
- « Manque du temps » ;
- « Absence d'analyse coûts-avantages » ;

- « Manque de soutien par la haute direction » ;
- « Manque de compétences techniques » ;
- « Manque des processus établis du management des données » ;
- « Manque du personnel compétent » ;
- « Manque des objectifs clairs » ;
- « Sécurité et confidentialité de mauvaises données ».

Pour pouvoir comparer les résultats issus de nos études avec ces conclusions, nous considérons aussi les groupes des éléments qui contribuent à l'efficacité d'usage du big data, dans une logique que leur absence est un élément inhibiteur. Cela nous permet de distinguer les groupes suivants :

- Les données (qualité, compréhension),
- Le management des données (processus établis, management des données isolées silos des données, la sécurité et la confidentialité des données),
- L'aspect humain (compétences techniques et personnelles) ;
- L'analyse coût-avantage ;
- Le soutien par la haute direction ;
- Les objectifs clairs et le temps.

Ce regroupement nous permet de mettre en évidence les similitudes et les différences de ces résultats avec les nôtres :

- La qualité des données avec ses 4 caractéristiques (complétude /F5-28/, pertinence /F8-21/, exactitude /F9-24/ et cohérence /F10-20/ (Annexe 12)) émerge comme un facteur des plus importants qui contribue à l'efficacité d'usage du big data dans nos résultats obtenus ; « *...la qualité des données, c'est d'abord son potentiel à apporter une solution au problème posé...dans ce sens-là, si l'on ne parle pas des caractéristiques plus techniques et je vous n'embêterai pas avec cela ...bref, une donnée peut être considérée « riche » dans un contexte et « pauvre » dans l'autre, ... , et donc ça dépend de la question posée et donc il faut bien être prudent dans ses conclusions... ».*
- En termes de management des données /F28-8/ (ce qui considère aussi les processus établis pour éviter les situations avec « silos des données » que les participants du projet

d'étude de Surbakti et ses coauteurs ont rencontré (F. Sejahtera et al., 2018), la confidentialité et la sécurité des données /F17-16/, le soutien de haute direction /F38-10/ et les objectifs précis /F40-29/ nous avons des conclusions similaires (Annexe 12).

- Le facteur « analyse coût-avantage » que nous avons étudié dans le modèle conceptuel, n'était pas retenu comme un des 29 facteurs les plus importants qui contribue à l'efficacité d'usage du big data au contexte arménien. Ce facteur a émergé lors des entretiens, comme un facteur ayant un impact critique dans la phase initiale préparatoire de l'adoption pour les utilisateurs-novices ; « ...nous avons bien fait une analyse coût-avantage dès le début du lancement du projet de l'exploitation du big data par notre entreprise...les 2 cas traditionnels du développement des technologies du big data ont été considérés ; « in house » et « sur les solutions Cloud » et c'est là, que cette analyse coût-avantage a pris assez de temps et avait une grande importance sur le choix, nous avons discuté quelques scénarios de développement « in house » en considération de ces dépenses liées, les investissements et son apport sur 10 ans pour l'entreprise...cette phase préparatoire a pris plus d'un an...les calculs ont montré que les investissements dans le développement de l'infrastructure « in house » ne sont pas justifiés pour la période considérée, compte tenu ses apports attendus...mais pour tel ou tel projet du big data...». C.C., BCD.

Tandis que, le manque de l'analyse coût-avantage apparaît dans l'étude du projet de Surbakti et ses coauteurs (F. Sejahtera et al., 2018) comme un inhibiteur où il s'agit d'un seul projet, l'autre inhibiteur, « le manque du temps » est aussi spécifique du projet étudié. Nous ne pouvons apporter une interprétation comparative à cet élément-là au niveau de nos résultats.

7.2.3.3 Étude sur la relation des éléments contextuels avec les facteurs influençant la réalisation de la valeur du big data en entreprise

La troisième étude (Mikalef et al., 2019b) nous intéresse particulièrement, car elle est de rares travaux de recherche dans le domaine de gestion du big data qui met en évidence l'importance du contexte. Ce travail de recherche vise à « casser » les préjugés que toutes les entreprises sont confrontées aux mêmes défis et devraient donc concentrer leurs investissements sur un ensemble uniforme d'aspects. Le rôle du contexte est bien traité dans la littérature du SI (Bechor, Neumann, Zviran, & Glezer, 2010), (Petter, Delone, & McLean, 2013). Pourtant, pour

un domaine émergent du big data, où les études empiriques sont peu nombreuses, il apparaît que ces études empiriques sur les sujets manquent bien d'attribuer une attention au contexte traité.

Très souvent, les échecs des projets du big data sont liés à l'imitation des pratiques des entreprises réussies dans le domaine de l'utilisation du big data sans avoir pris en compte des particularités du contexte (Kiron, 2017). Ces particularités contextuelles sont importantes, car elles définissent la manière dont les facteurs – composants de réalisation de valeur du big data sont structurées et dans quelle mesure elles émergent (Mikalef et al., 2018).

Donc, cette étude examine les entreprises qui adoptent les initiatives du big data, un groupe (utilisateurs novices) qui possède des connaissances et des compétences assez limitées dans le domaine et essaie de créer une valeur commerciale à partir du potentiel de big data. Sur la base d'une méthode mixte, cette étude empirique démarre par une enquête quantitative auprès des entreprises grecques. Ensuite, elle continue examiner 3 études de cas et en s'appuyant sur une base théorique proposée de (Bechor et al., 2010) et (Petter et al., 2013) essaie de mettre en avant l'idée que le contexte joue un rôle important et que les mêmes ressources dans les contextes différents peuvent avoir une grande ou une moindre importance dans la réalisation de valeur du big data. Surtout, ils supposent que la signification de ces éléments-là sous différentes conditions contextuelles n'est pas la même.

Dans leur modèle et l'étude empirique, Bechor et ses coauteurs (Bechor et al., 2010) examinent le contexte de planification (SISP-Planification Stratégique du Système D'information) comme l'ensemble des éléments représentant l'organisation et son environnement; notamment les conditions qui peuvent avoir un impact sur le processus de SISP et son succès (niveau d'incertitude environnementale, structure organisationnelle, rôle de TI à l'organisation, etc.). Ils considèrent le processus de SISP comme une variable qui dépend de 3 dimensions : des facteurs clés de succès, de l'approche SISP (un ensemble des éléments de décision – orientation de planification, le temps qu'elle recouvre, etc.) et le contexte du SISP. L'idée principale du modèle, validé empiriquement par ces auteurs, est que la relation entre le succès de SISP et les facteurs clés de succès dépend de la cohérence entre l'approche et le contexte. L'importance du rôle de contexte est aussi prise en compte par Petter et ses coauteurs (Petter et al., 2013) dans leur étude de succès du SI.

Il convient néanmoins de noter que ces études ont été réalisées dans un domaine du SI traditionnel, tandis que Mikalef et ses coauteurs (Mikalef et al., 2019a) essaient de remplir le manque dans le sujet du big data afin de démontrer empiriquement l'importance de l'impact de différents éléments sur la réalisation de la valeur du big data, en prenant en compte le contexte.

La mauvaise pratique répandue dans les milieux économiques est de supposer que toute entreprise doit se concentrer dans la même mesure des éléments de succès dans la réalisation de valeur du big data, pourtant Mikalef et ses coauteurs (Mikalef et al., 2019a) essaient de démontrer empiriquement, qu'il existe de différences.

Ces auteurs partent de l'idée principale que les entreprises diffèrent dans leur mode de fonctionnement, dans leur environnement et le contexte, nécessitant une attention plus agile dans les différents ensembles de ces aspects.

Les résultats principaux de leur étude quantitative suggèrent que selon le contexte de l'examen, les mêmes facteurs de « données », des « ressources technologiques », des « compétences techniques », d'« établissement des pratiques procédurales », des « compétences managériales », de « culture de l'entreprise axée sur les données » reçoivent de différentes valorisations pour la réalisation de valeur du big data : « éléments critiques », « éléments moins importants », « aspects essentiels », « aspects nécessaires ».

Les résultats issus de nos études complètent les conclusions de Mikalef et al (Mikalef et al., 2019a), de la manière suivante :

L'intérêt organisationnel perçu émerge comme le moteur d'adoption du big data pour une entreprise « utilisateur novice », le facteur le plus important, retenu par l'étude quantitative, l'« intention d'utiliser » /F23-1/ (Annexe 12) s'interprète comme impacté par l'inertie de dupliquer le précurseur en contexte arménien « ...j'oserai dire que c'est plutôt la peur de retarder les précurseurs du domaine qui a conditionné cette implantation que la vraie conscience de la part de la « haute direction » quelle valeur peut créer l'usage des informations issues de l'exploitation du big data pour l'entreprise..., la ressource humaine compétente est toujours un point critique (C.C.)... il est très compliqué en Arménie (c'est une question actuelle partout dans le monde, mais surtout en Arménie) de trouver des gens ... il y a beaucoup de manque des connaissances techniques ; souvent l'on peut chercher des mois quelqu'un qui

connaît bien Spark...(D.D.)...oui, big data c'est beaucoup trop des données, pour mieux imaginer l'espace sur mon portable, c'est 500GB, si vous voulez physiquement, ça représente un volume de 1 million de livres au format de texte...mais ce que l'on peut garder sur ce portable je n'oserai pas appeler « le big data », vous imaginez maintenant au niveau organisationnel de quoi l'on parle...(A.D., Big Data Engineer chez ABC)...

Ces illustrations issues des entretiens s'ajoutent aux résultats obtenus de l'étude quantitative qui montrent dans un ordre hiérarchique les facteurs comme contributeurs importants à l'efficacité d'usage du big data au niveau d'entreprise fonctionnant dans des conditions de haut dynamisme au contexte arménien : les données, les pratiques structurelles et procédurales, l'alignement stratégique, les compétences managériales, détaillés dans la section précédente.

Nos résultats soutiennent également l'idée que le niveau de maturité de l'entreprise affecte potentiellement l'adoption de nouvelles technologies sur les aspects différents (Trieu, 2013), comme, mise en évidence par Mikalef et ses coauteurs (Mikalef et al., 2019a) sur les aspects de ressources technologiques et les pratiques structurelles / procédurales. Ceci est conforme avec l'idée que les entreprises (grandes entreprises) qui disposent de plus de ressources et une structure déjà établie des procédures de management des données, sont davantage prêtes à réaliser une valeur du big data.

7.2.3.4 Étude de l'adoption et l'usage de BDA (Big Data Analytics) par les PME's en Iran

La quatrième étude présente un grand intérêt dans le sens où elle a été réalisée en Iran, un pays ayant des spécificités de marché, un marché qui manque des investissements internationaux, avec un faible impact de globalisation, assorti de sanctions.

Ainsi, les facteurs influençant la réalisation de valeur du big data dans les entreprises iraniennes au contexte organisationnel et aussi certains facteurs technologiques sont recensés aussi par nos résultats comme les plus importants qui contribuent à l'efficacité d'usage du big data :

- Au niveau du contexte organisationnel ; la « compétence organisationnelle » /F29-19/, « support de haute direction » /F38-10/ ; contexte technologique : « facilité perçue de

l'utilisation » /F22-14/, « confidentialité et la sécurité » /F17-16/, « observabilité perçue » /F24-23/) (Annexe 12).

- Au niveau des éléments contextuels de la pression concurrentielle et les règlements gouvernementaux qui n'ont pas un effet sur la réalisation de valeur du big data par les entreprises iraniennes, nous constatons des contradictions avec nos résultats (F18-3, F28-8) (Annexe 12). Certains de ces résultats ont des contradictions aussi avec des résultats d'études sur les sujets similaires (Lai et al., 2018), (Verma & Bhattacharyya, 2017), (D. Q. Chen et al., 2015), (Ghobakhloo et al., 2011).

Tout d'abord, ces contradictions peuvent être justifiées par les particularités contextuelles iraniennes et surtout par les spécificités du marché iranien, la conséquence des sanctions sur le pays, les effets mineurs de la mondialisation et le manque d'investissements internationaux. Ce fait renforce encore une fois la constatation que *les particularités contextuelles définissent la manière dont les facteurs – composants de réalisation de valeur du big data sont structurés et dans quelles mesure ils sont importants (facteurs contextuels)* (Mikalef et al., 2018). De même, *la signification de ces éléments sous différentes conditions contextuelles n'est pas la même* (Mikalef et al., 2019a), donc sans prise en compte les spécificités contextuelles et une « aveugle imitation » des pratiques des entreprises réussies dans le domaine de l'utilisation du big data peuvent bien les amener aux échecs (Kiron, 2017).

7.3 Contributions théoriques et managériales de la thèse

7.3.1 Contributions théoriques

Au niveau des contributions théoriques, ce travail de thèse répond aux appels des auteurs de 3 travaux principaux sur le thème (Burton-Jones & Grange, 2013), (Surbakti et al., 2019), (Surbakti, 2020) et vise compléter la littérature existante de plusieurs façons :

- Ce travail de recherche démontre comment l'efficacité d'usage peut être empiriquement explorée et contextualisée dans un domaine particulier celui du big data. Plus particulièrement, nous avons proposé :
 - Un guide méthodologique détaillé pour étudier le sujet au niveau organisationnel ;

- 2 instruments développés et adaptés ; un questionnaire, l’outil principal de l’enquête quantitative et un guide d’entretien, l’outil principal de l’enquête qualitative.

Pour développer l’instrument principal de l’enquête quantitative, une opérationnalisation de 41 variables par une méthode bibliométrique a été réalisée. Ainsi, cet instrument, adapté au sujet de recherche par des experts du domaine, recense les échelles de mesure de 41 variables du modèle conceptuel de Surbakti et al. (2019). Nous l’avons appliqué pour la première fois à un contexte arménien, mais nous l’avons mis en épreuve aussi par les professionnels qui travaillent dans un contexte américain et ils affirment son applicabilité à n’importe quel contexte.

Le guide d’entretien a été adapté de (Mikalef et al., 2019a) et ajusté au contexte de l’examen. Il peut être utilisé pour conduire une étude qualitative exploratoire sur l’efficacité d’usage du big data en entreprise.

- En appliquant les 3 pôles théoriques de contextualisation de Livian (Livian, 2020), nous avons exploré le sujet avec un ancrage du contexte arménien, un contexte d’examen inexploré dans les recherches sur le sujet. Il faut noter, que même dans les recherches du management, le contexte arménien reste très peu étudié (Sargsyan, 2020).
- A notre connaissance, nous avons réalisé le premier test empirique du modèle de Surbakti et al. (2019).
- Ce travail complète la théorie de l’efficacité d’usage, explorée dans le domaine du MIS (Burton-Jones & Grange, 2013) et propose un développement sur le contexte spécifique du big data.
- Ce travail de thèse soutient le développement de définition consensuelle de l’efficacité d’usage du big data et propose une définition ajustée et appropriée pour les 2 groupes différents d’utilisateur du big data *en français* :
 - Pour un « utilisateur novice », l’efficacité d’usage du big data consiste en *‘l’exploitation de la valeur du big data par la définition du problème et la mise en évidence de sa pertinence pour l’organisation, afin d’apporter une solution à l’aide des informations fournies par les données tout en bâtissant une infrastructure technologique pérenne et pouvant être mise à l’échelle’*.

- Pour le cas « d'utilisateur avancé », l'efficacité d'usage du big data peut bien consister en *‘fonctionnement opérationnel par l'usage des informations fournies par les données ce qui peut bien constituer un moyen important, mais pas suffisant à la réalisation de sa propre valeur ou bien contribuer à la réalisation de valeur pour le tiers ’*.
- Enfin, ce travail a un apport pour la littérature francophone sur le sujet qui est assez limité et les grandes thématiques restent encore discutées par une forte domination de travaux anglophones.

7.3.2 Contributions managériales

Ce travail de thèse propose aussi plusieurs implications intéressantes pour les praticiens non seulement de l'environnement d'affaires arménien.

Nous voulons présenter ces contributions managériales par les 3 axes selon leurs applications :

- « compréhension » ;
- « action » ;
- « évaluation ».

Au niveau des compréhensions plus fines du sujet, pour les professionnels du domaine et les managers de différents niveaux :

- Ce travail de thèse montre que l'efficacité d'usage du big data est valorisé différemment, compte tenu du type d'entreprise ; « utilisateur novice » ou « utilisateur avancé » ; et donc doit avoir une approche différente des actions à mettre en œuvre pour son adaptation et utilisation pratique.
- Ce travail propose une consolidation de la littérature du domaine sur les 41 facteurs composants du modèle conceptuel qui peuvent potentiellement impacter l'efficacité d'usage du big data en entreprise. Ce résumé tente de combler le manque de compréhension profonde du sujet chez les praticiens. Il présente un grand intérêt non seulement pour les managers dont les entreprises exploitent déjà le big data, mais également pour ceux qui en projettent l'adoption.

- Les praticiens du domaine de chaque type « utilisateur entreprise » peuvent bien se servir de nos résultats de l'étude quantitative comme un guide pour développer une meilleure compréhension d'usage efficace du big data dans leur entreprise.

Ici, donc, l'apport pratique principal est de mettre en évidence l'importance de la prise en compte de tous les facteurs qui peuvent potentiellement contribuer à l'efficacité d'usage du big data dans chaque groupe d'entreprises.

La deuxième sous-catégorie de cet axe consiste en un projet qui est déjà en cours de préparation ; il s'agit de l'insertion d'un cursus basique dans les divers champs disciplinaires comme un programme obligatoire dans les institutions d'enseignement supérieur.

Nous croyons, que dans cette ère des données, une compréhension structurée sur l'usage du big data par les spécialistes de divers domaines manque beaucoup, en particulier dans les institutions supérieures d'enseignement en Arménie, « ...souvent il arrive que même une vision générale sur les choses manque chez les candidatures dans nos entretiens pour les postes dans les projets du big data ... » (Mikalef et al., 2019a).

Donc, ce projet est dans la continuité de ce travail de recherche, qui vise une contribution managériale de nature stratégique et de long terme.

Dans le deuxième axe des contributions managériales « action », nous voulons mettre en évidence les deux principales orientations des actions que permet notre travail. Ces deux orientations consistent en une conduite importante dans la phase d'adoption du big data par les entreprises arméniennes qui exercent déjà une activité :

- La première orientation envisage une meilleure vision sur les facteurs importants à considérer dès le début et même avant l'adoption. Son apport consiste dans la minimisation de l'échec et dans la contribution au bon fonctionnement de l'implantation.
- La deuxième orientation envisage un autre scénario possible, qui est aussi un résultat positif, est, après prise en compte de tous les facteurs potentiels, d'en refuser l'adoption.

Au niveau de l'évaluation, l'instrument de mesure que nous avons développé peut effectivement être utilisé. Grâce à cet instrument il est possible de mettre en évidence les forces et les faiblesses de 2 types d'utilisateurs en matière de l'efficacité d'usage du big data.

Dans cette perspective, nos résultats peuvent servir à l'entreprise comme un outil de contrôle pour comprendre son niveau de maturité et agir en cohérence. Ayant à leur disposition un tel outil de contrôle, les entreprises, qui exploitent déjà le big data dans les deux profils d'utilisateurs, peuvent mettre en place une évaluation de leur état d'efficacité d'usage. Si dans le premier groupe d'utilisateur (utilisateur novice), il peut y avoir une implication au niveau stratégique, pour le deuxième groupe d'utilisateurs (utilisateur avancé), elle peut aussi bien servir comme un outil de contrôle opérationnel.

7.4 Limites et pistes de recherche

Dans le cadre de notre travail de thèse, nous avons certainement démontré que les deux types d'utilisateurs du big data peuvent disposer de différentes mesures commerciales concernant l'efficacité d'usage des données et ça ne dépend pas du contexte d'examen. Ainsi, compte tenu des impératifs temporels que requiert la réalisation d'une thèse, nous n'avons pas pu explorer l'effet des caractéristiques contextuelles sur les résultats. Et donc cela pourrait bien être une autre voie de recherche envisagée sur le sujet.

Par ailleurs, une autre perspective de recherche que nous proposons réside dans l'étude plus profonde de chacun des deux groupes d'utilisateurs, « utilisateur avancé » et « utilisateur novice », en appliquant des méthodes de recherche qui prennent appui principalement sur des données qualitatives afin de laisser un maximum de latitude de compréhension des choses.

Une dernière voie de recherche que les impératifs du temps n'ont pas permis de réaliser dans le cadre de ce travail de recherche et que nous envisageons d'entreprendre après la thèse consiste dans l'étude des corrélations des facteurs pour chacun des groupes d'utilisateurs. L'exploitation des données déjà collectées, nous a permis de dresser les matrices de corrélation pour chacun des deux groupes d'utilisateurs que nous proposons en annexe 11. Pour une telle perspective d'étude, nous proposons d'adopter une démarche de *recherche à visée transformative* (Pastorelli, 2000). Dans ce cas, une insertion dans les pratiques organisationnelles est nécessaire pour la mise en évidence d'une étude plus profonde de ces relations corrélées.

Au terme de notre étude réalisée, nous espérons que ce travail de thèse va contribuer à une meilleure compréhension de l'efficacité d'usage du phénomène émergent du big data tant au niveau académique que pratique.

Nous envisageons de continuer nos recherches sur le sujet et entreprendre un travail important par la suite avec une proposition d'un cursus académique sur les aspects managériaux du big data dans les formations universitaires, tout particulièrement dans les champs disciplinaires de gestion qui nous paraissent aujourd'hui nécessaires de donner aux futurs professionnels.

Table des illustrations

Figure 1 Volume des données accessibles aux entreprises par rapport aux volumes qu'elles sont en capacité de traiter (Zikopoulos & Chris Eaton, 2011)	28
Figure 2 Fréquence d'apparition du terme "big data" (Gandomi & Haider, 2015).....	29
Figure 3 Évolution des données générées (Yaqoob et al., 2016c)	31
Figure 4 Modèle conceptuel du big data pour les entreprises du domaine de « High-Frequency Trading » (Seddon & Currie, 2017).....	33
Figure 5 Nuage de mots dans le corpus d'analyse du terme « big data » (De Mauro et al., 2016).....	34
Figure 6 Les paradigmes de la science (Hey, Tansley, & Tolle, 2009)	35
Figure 7 Résultats de l'enquête: Définitions du big data proposées par les cadres exécutifs (Gandomi & Haider, 2015)	36
Figure 8 Un exemple de « smart city » (Yaqoob et al., 2016a)	38
Figure 9 Plateformes du big data (Hu et al., 2014).....	41
Figure 10 Genèse des applications du big data (Abolfazli et al., 2014).....	42
Figure 11 Système du big data à 3 catégories (Hu et al., 2014).....	43
Figure 12 Cycle de vie des données (N. Khan et al., 2014)	45
Figure 13 2 Scénarios de traitement du big data (Hashem et al., 2015).....	49
Figure 14 Processus du traitement des données non-structurées (Siddiqi et al., 2016).....	50
Figure 15 Processus de production d'un insight par l'exploitation du big data	53
Figure 16 Structures d'un SI (Wand & Weber, 1995).....	59
Figure 17 Niveaux de structures du SI (Burton-Jones & Grange, 2013)	61
Figure 18 Cadre général pour étudier l'efficacité d'usage (Burton-Jones & Grange, 2013)	62
Figure 19 Zoom sur le lien 1 de figure 18 (Burton-Jones & Grange, 2013).....	65
Figure 20 Efficacité d'usage et ses effets sur la performance (Burton-Jones & Grange, 2013).....	65
Figure 21 Facteurs déterminants la performance individuelle (Campbell, 1990).....	67
Figure 22 Facteurs influençant l'efficacité d'usage (Burton-Jones & Grange, 2013)	69
Figure 23 TAM original (Legris et al., 2003)	70
Figure 24 Modèle UTAUT (Venkatesh et al., 2003)	72
Figure 25 Modèle initial du succès d'un SI (DeLone & McLean, 1992)	72
Figure 26 Modèle proposé de Goodhue et Thompson –TTF (Goodhue & Thompson, 1995).....	73
Figure 27 Construits Majeurs et les contributions de la théorie AST (DeSanctis & Poole, 1994)	74
Figure 28 Modèle de TTF (Task-Technology Fit) avec les évaluations des utilisateurs (Goodhue, 1995) 75	
Figure 29 Modèle de l'efficacité d'usage, traduit de (Surbakti et al., 2019)	80
Figure 30 Compétences nécessaires pour la gestion du big data (Fogelman-Soulié & Lu, 2016).....	98
Figure 31 Statut accordé aux données (Mbengue & Vandangeon-Derumez, 1999).....	122
Figure 32 Conception du réel et les paradigmes épistémologiques (Allard-Poesi & Perret, 2014).....	123
Figure 33 Dépendance du chercheur et de l'objet de l'étude (Mbengue & Vandangeon-Derumez, 1999)124	
Figure 34 Modes de raisonnement et connaissance scientifique (Charreire-Petit & Durieux, 2014).....	125
Figure 35 Construction des connaissances et les observations empiriques (Mbengue & Vandangeon-Derumez, 1999).....	126
Figure 36 Modèle conceptuel de l'efficacité d'usage proposé de Surbakti et al. (2019).....	132
Figure 37 Les principales étapes de l'analyse de contenu	135
Figure 38 Premiers résultats de l'analyse thématique.....	137

Figure 39 Composition sectorielle détaillée, 2012-17	160
Figure 40 Taux d'entrée et de sortie des entreprises, en comparaison avec d'autres pays	161
Figure 41 Modèle de l'efficacité d'usage, traduit de Surbakti et al. (2019)	195
Figure 42 Modèle de l'efficacité d'usage du big data pour un utilisateur de type « utilisateur novice » ..	202
Figure 43 Modèle de l'efficacité d'usage du big data pour un utilisateur de type « utilisateur avancé »	204
Figure 44 Éléments-composants de l'efficacité d'usage du big data en entreprise	230
Tableau 1 Comparaison entre les données traditionnelles et le big data (Hu, Wen, Chua, & Li, 2014).....	33
Tableau 2 Rapidité de génération des données non-structurées (N. Khan et al., 2014).....	37
Tableau 3 Comparaison entre les 2 plateformes de traitement du big data (Philip Chen & Zhang, 2014).	41
Tableau 4 Comparaison entre les données structurées et non structurées (N. Khan et al., 2014).....	49
Tableau 5 Exemples des composants du système suite les types du SI (Burton-Jones & Grange, 2013)..	59
Tableau 6 Paradigme épistémologique réaliste critique (transcendantal) (Gavard-Perret et al., 2012)....	119
Tableau 7 Détail des entretiens exploratoires	134
Tableau 8 Exemples de codage des entretiens de l'étude qualitative exploratoire	135
Tableau 9 Construits avec les éléments de l'instrument d'enquête.....	138
Tableau 10 Détail des entretiens confirmatoires	150
Tableau 11 Profil du pays en quelques chiffres clés	156
Tableau 12 Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur novice ».....	217
Tableau 13 Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur avancé»	219
Tableau 14 Présentation synthétisée des résultats des études sur l'efficacité d'usage du big data et la réalisation de valeur	228

Bibliographie

- Abbasi, A., Adjeroh, D., Dredze, M., Paul, M. J., Zahedi, F. M., Zhao, H., ... Ross, A. (2014). Social media analytics for smart health. *IEEE Intelligent Systems*, 29(2), 60–80. <https://doi.org/10.1109/MIS.2014.29>
- Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, 17(2), 1–32. <https://doi.org/10.17705/1jais.00423>
- Abd Wahab, M., Mohd, M., Hanafi, H., & Mohamad Mohsin, M. (2008). Data pre-processing on web server logs for generalized association rules mining algorithm. *World Academy of Science, Engineering and Technology*, 48.
- Aberg, J., & Shahmehri, N. (2001). An empirical study of human web assistants: Implications for user support in web information systems. In *Conference on Human Factors in Computing Systems - Proceedings*.
- Abolfazli, S., Sanaei, Z., Gani, A., Xia, F., & Yang, L. T. (2014). Rich mobile applications: Genesis, taxonomy, and open issues. *Journal of Network and Computer Applications*. <https://doi.org/10.1016/j.jnca.2013.09.009>
- Adrian, C., Abdullah, R., Atan, R., & Jusoh, Y. Y. (2016). Towards developing strategic assessment model for big data implementation: A systematic literature review. *International Journal of Advances in Soft Computing and Its Applications*.
- Adrian, C., Abdullah, R., Atan, R., & Jusoh, Y. Y. (2017). Factors influencing to the implementation success of big data analytics: A systematic literature review. In *International Conference on Research and Innovation in Information Systems, ICRIIS*. <https://doi.org/10.1109/ICRIIS.2017.8002536>
- Agarwal R. (2000). Individual Acceptance of Information Technologies. *Framing the Domains of IT Management Projecting the Future through the Past*.
- Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/3250951>
- Ahmad, A., Ahmad, R., & Hashim, K. F. (2016). Innovation traits for business intelligence successful deployment. *Journal of Theoretical and Applied Information Technology*, 89(1).
- Ahmed Oussous a, Fatima-Zahra Benjelloun a, Ayoub Ait Lahcen a, b,†, S. B. (2018). Big Data technologies: A survey. *Journal of King Saud University – Computer and Information Sciences*, 431–448.
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32(4). <https://doi.org/10.1111/j.1559-1816.2002.tb00236.x>

- Akhtar, P., Frynas, J. G., Mellahi, K., & Ullah, S. (2019). Big Data-Savvy Teams' Skills, Big Data-Driven Actions and Business Performance. *British Journal of Management*. <https://doi.org/10.1111/1467-8551.12333>
- Akhtar, P., Khan, Z., Frynas, J. G., Tse, Y. K., & Rao-Nicholson, R. (2018). Essential Micro-foundations for Contemporary Business Operations: Top Management Tangible Competencies, Relationship-based Business Networks and Environmental Sustainability. *British Journal of Management*, 29(1). <https://doi.org/10.1111/1467-8551.12233>
- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2016.08.018>
- Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6(1). <https://doi.org/10.1186/s13174-015-0041-5>
- Alam, S. S., Alam, S. S., Khatibi, A., Ahmad, M. I. S., & Ismail, H. Bin. (2008). Factors affecting e-commerce adoption in the electronic manufacturing companies in Malaysia. *International Journal of Commerce and Management*, 17(1–2). <https://doi.org/10.1108/10569210710776503>
- Aljumah, A. I., Nuseir, M. T., & Alam, M. M. (2021). Organizational performance and capabilities to analyze big data: do the ambidexterity and business value of big data analytics matter? *Business Process Management Journal*, 27(4). <https://doi.org/10.1108/BPMJ-07-2020-0335>
- Allard-Poesi, F., & Perret, V. (2014). Chapitre 1. Fondements épistémologiques de la recherche. In *Méthodes de recherche en management*. <https://doi.org/10.3917/dunod.thiet.2014.01.0014>
- Altmann, N., Köhler, C., & Meil, P. (2017). *Technology and work in German industry*. *Technology and Work in German Industry*. <https://doi.org/10.4324/9781315212487>
- Amaratunga, D., & Baldry, D. (2002). Moving from performance measurement to performance management. *Facilities*, 20. <https://doi.org/10.1108/02632770210426701>
- Amoako-Gyampah, K., & Salam, A. F. (2004). An extension of the technology acceptance model in an ERP implementation environment. *Information and Management*. <https://doi.org/10.1016/j.im.2003.08.010>
- Avenier, M. J., & Thomas, C. (2011). Mixer quali et quanti pour quoi faire ? Méthodologie sans épistémologie n'est que ruine de réflexion. <halshs-00644303>.
- Avenier, M., & Thomas, C. (2015). Finding one's way around various methodological guidelines for doing rigorous case studies: A comparison of four epistemological frameworks. *Systèmes d'information & Management*, 20(1). <https://doi.org/10.3917/sim.151.0061>
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016a). Transformational issues of big data and analytics in networked business. *MIS Quarterly: Management*

- Information Systems*, 40(4), 807–818. <https://doi.org/10.25300/MISQ/2016/40:4.03>
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016b). Transformational Issues of Big Data and Analytics in Networked Business. *MIS Quarterly*. <https://doi.org/10.25300/MISQ/2016/40:4.03>
- Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2019). Big data adoption: State of the art and research challenges. *Information Processing and Management*. <https://doi.org/10.1016/j.ipm.2019.102095>
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37(2). <https://doi.org/10.1037/0003-066X.37.2.122>
- Bardin, L. (2013). *L'analyse de contenu*. Paris cedex 14, France: Presses Universitaires de France. <https://doi.org/10.3917/puf.bard.2013.01>
- Barki, H., Paré, G., & Sicotte, C. (2008). Linking IT implementation and acceptance via the construct of psychological ownership of information technology. *Journal of Information Technology*, 23(4). <https://doi.org/10.1057/jit.2008.12>
- Barki, H., Titah, R., & Boffo, C. (2007). Information system use-related activity: An expanded behavioral conceptualization of individual-level information system use. *Information Systems Research*, 18(2). <https://doi.org/10.1287/isre.1070.0122>
- Barkin, S. R., & Dickson, G. W. (1977). An investigation of information system utilization. *Information and Management*. [https://doi.org/10.1016/0378-7206\(77\)90007-6](https://doi.org/10.1016/0378-7206(77)90007-6)
- Barnes, S. J. (2011). Understanding use continuance in virtual worlds: Empirical test of a research model. *Information and Management*, 48(8). <https://doi.org/10.1016/j.im.2011.08.004>
- Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of Management*. [https://doi.org/10.1016/S0149-2063\(01\)00115-5](https://doi.org/10.1016/S0149-2063(01)00115-5)
- Batistič, S., & van der Laken, P. (2019). History, Evolution and Future of Big Data and Analytics: A Bibliometric Analysis of Its Relationship to Performance in Organizations. *British Journal of Management*. <https://doi.org/10.1111/1467-8551.12340>
- Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. *Journal of Big Data*, 9(1). <https://doi.org/10.1186/s40537-021-00553-4>
- Bauer, R. A. (1960). Consumer Behaviour as Risk Taking. In *Dynamic marketing for a changing world. Proceedings of the 43rd National Conference of the American Marketing Association, Chicago*.
- Baumard, P., & Ibert, J. (2014). Chapitre 4. Quelles approches avec quelles données ? In *Méthodes de recherche en management*. <https://doi.org/10.3917/dunod.thiet.2014.01.0105>
- Bechor, T., Neumann, S., Zviran, M., & Glezer, C. (2010). A contingency model for estimating

- success of strategic information systems planning. *Information and Management*, 47(1).
<https://doi.org/10.1016/j.im.2009.09.004>
- Behl, A. (2022). Antecedents to firm performance and competitiveness using the lens of big data analytics: a cross-cultural study. *Management Decision*, 60(2). <https://doi.org/10.1108/MD-01-2020-0121>
- Benoit-Smullyan, E., & Bachelard, G. (1940). La formation de l'esprit scientifique. *American Sociological Review*. <https://doi.org/10.2307/2084470>
- Berger, T., Chen, C., & Frey, C. B. (2018). Drivers of disruption? Estimating the Uber effect. *European Economic Review*, 110, 197–210.
<https://doi.org/10.1016/j.eurocorev.2018.05.006>
- Bertereau, C., Marbot, E., & Chaudat, P. (2019). Positionnement épistémologique et orientation de la recherche : un focus sur l'étude des stéréotypes. *RIMHE : Revue Interdisciplinaire Management, Homme & Entreprise*, n 34; 8(1), 51–66.
<https://doi.org/10.3917/rimhe.034.0051>
- Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: Toward a next generation of insights. *MIS Quarterly: Management Information Systems*, 37(2), 471–482. <https://doi.org/10.25300/MISQ/2013/37:2.3>
- Bharadwaj, A., Sambamurthy, V., & Zmud, R. W. (1999). IT capabilities: theoretical perspectives and empirical operationalization. *Management Science*.
<https://doi.org/10.1145/352925.352962>
- Bhaskar, R. (1998). Philosophy and Scientific Realism. *Critical Realism: essential readings*.
Critical Realism: Essential Readings.
- Bhaskar, R. (2013). *A realist theory of science. A Realist Theory of Science*.
<https://doi.org/10.4324/9780203090732>
- Bidan, M. (2020). Chapitre 16. Une introduction aux méthodologies quantitatives. In *Produire du savoir et de l'action*. <https://doi.org/10.3917/ems.peret.2020.01.0183>
- Bischoff, S., Aier, S., Haki, K., & Winter, R. (2015). Understanding Continuous Use of Business Intelligence Systems: A Mixed Methods Investigation. *Journal of Information Technology Theory and Application (JITTA)*, 16(2).
- Bishop, C. M. (2014). Bishop - Pattern Recognition And Machine Learning - Springer 2006.
Antimicrobial Agents and Chemotherapy, 58(12).
- Bjeladinovic, S. (2018). A fresh approach for hybrid SQL/NoSQL database design based on data structuredness. *Enterprise Information Systems*, 12(8–9).
<https://doi.org/10.1080/17517575.2018.1446102>
- Blanc, A., Drucker-Godard, C., & Ehlinger, S. (2014). Chapitre 17. Exploitation des données textuelles. In *Méthodes de recherche en management*.
<https://doi.org/10.3917/dunod.thiet.2014.01.0551>

- Boisot, M., & Mckelvey, B. (2011). Integrating Modernist And Postmodernist Perspectives On Organizations: A Complexity Science Bridge. *Academy of Management Review*, 35(3). <https://doi.org/10.5465/amr.2010.51142028>
- Boudreau, M.-C., & Seligman, L. (2011). Quality of Use of a Complex Technology. *Journal of Organizational and End User Computing*, 17(4). <https://doi.org/10.4018/joeuc.2005100101>
- Boudreau, M. C., & Seligman, L. (2005). Quality of use of a complex technology: A learning-based model. *Journal of Organizational and End User Computing*. <https://doi.org/10.4018/joeuc.2005100101>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society*, 15(5). <https://doi.org/10.1080/1369118X.2012.678878>
- Brinch, M., Gunasekaran, A., & Fosso Wamba, S. (2021). Firm-level capabilities towards big data value creation. *Journal of Business Research*, 131. <https://doi.org/10.1016/j.jbusres.2020.07.036>
- Brous, P., Janssen, M., Schraven, D., Spiegelner, J., & Duzgun, B. C. (2017). Factors influencing adoption of IoT for data-driven decision making in asset management organizations. In *IoTBDS 2017 - Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security*. <https://doi.org/10.5220/0006296300700079>
- Bryant, F. B., & Yarnold, P. R. (1995). Principal-Components Analysis and Exploratory and Confirmatory Factor Analysis. *Reading and Understanding Multivariate Statistics*.
- Brynjolfsson, E., & McAfee, A. (2012). Big Data: The Management Revolution. *Harvard Business Review*.
- Buchanan, D. A., & Bryman, A. (2007). Contextualizing methods choice in organizational research. *Organizational Research Methods*, 10(3). <https://doi.org/10.1177/1094428106295046>
- Buhl, H. U., Röglinger, M., Moser, F., & Heidemann, J. (2013). Big data: A fashionable topic with(out) sustainable relevance for research and practice? *Business and Information Systems Engineering*. <https://doi.org/10.1007/s12599-013-0249-5>
- Bunge, M. (1993). Realism and antirealism in social science. *Theory and Decision*, 35(3). <https://doi.org/10.1007/BF01075199>
- Burton-Jones, A., & Grange, C. (2013). From use to effective use: A representation theory perspective. *Information Systems Research*, 24(3), 632–658. <https://doi.org/10.1287/isre.1120.0444>
- Burton-Jones, A., & Straub, D. W. (2006). Reconceptualizing system usage: An approach and empirical test. *Information Systems Research*, 17(3), 228–246. <https://doi.org/10.1287/isre.1060.0096>
- Burton-Jones, A., Wand, Y., & Weber, R. (2009). Guidelines for empirical evaluations of

- conceptual modeling grammars. *Journal of the Association for Information Systems*, 10(6). <https://doi.org/10.17705/1jais.00201>
- Byrd, T. A., & Turner, D. E. (2000). Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. *Journal of Management Information Systems*. <https://doi.org/10.1080/07421222.2000.11045632>
- Cabrera, D., Cabrera, L., Powers, E., Solin, J., & Kushner, J. (2018). Applying systems thinking models of organizational design and change in community operational research. *European Journal of Operational Research*, 268(3). <https://doi.org/10.1016/j.ejor.2017.11.006>
- Calic, G., & Ghasemaghaei, M. (2021). Big data for social benefits: Innovation as a mediator of the relationship between big data and corporate social performance. *Journal of Business Research*, 131, 391–401. <https://doi.org/10.1016/j.jbusres.2020.11.003>
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In *Handbook of Industrial and Organizational Psychology*.
- Campbell, J. P., & Wiernik, B. M. (2015). The Modeling and Assessment of Work Performance. *Annual Review of Organizational Psychology and Organizational Behavior*. <https://doi.org/10.1146/annurev-orgpsych-032414-111427>
- Carlaw, K. I., & Oxley, L. (2008). Resolving the productivity paradox. *Mathematics and Computers in Simulation*, 78(2–3). <https://doi.org/10.1016/j.matcom.2008.01.029>
- Carver, C. S., & Scheier, M. F. (1998). *On the Self-Regulation of Behavior*. *On the Self-Regulation of Behavior*. <https://doi.org/10.1017/cbo9781139174794>
- Castellano, S. (2016). Be Afraid, Be Uber-Afraid. *TD Magazine*, 70(2).
- Cetindamar, D., Shdifat, B., & Erfani, E. (2022). Understanding Big Data Analytics Capability and Sustainable Supply Chains. *Information Systems Management*, 39(1). <https://doi.org/10.1080/10580530.2021.1900464>
- Chalus-Sauvannet, M.-C. (2021). *Veille stratégique à l'usage des décideurs*. (E. Editions, Ed.).
- Charreire-Petit, S., & Durieux, F. (2014). Chapitre 3. Explorer et tester : les deux voies de la recherche. In *Méthodes de recherche en management*. <https://doi.org/10.3917/dunod.thiet.2014.01.0076>
- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*. <https://doi.org/10.1145/1978542.1978562>
- Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*. <https://doi.org/10.1080/07421222.2015.1138364>
- Chen, J., Wang, C., & Wang, R. (2009). Using stacked generalization to combine SVMs in magnitude and shape feature spaces for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7).

<https://doi.org/10.1109/TGRS.2008.2010491>

- Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. In *Mobile Networks and Applications* (Vol. 19). <https://doi.org/10.1007/s11036-013-0489-0>
- Cho, J., & Garcia-Molina, H. (2000). The evolution of the web and implications for an incremental crawler. In *Proceedings of the 26th International Conference on Very Large Data Bases, VLDB'00*.
- Christensen, C. M. (1997). The Innovator ' s Dilemma : When New Technologies Cause Great Firms to Fail (Management of Innovation and Change) Hardcover –. *Harvard Business School Press*, 1–14.
- Clarke, R. (2016). Big data, big risks. *Information Systems Journal*. <https://doi.org/10.1111/isj.12088>
- Comuzzi, M., & Patel, A. (2016). How organisations leverage: Big Data: A maturity model. *Industrial Management and Data Systems*. <https://doi.org/10.1108/IMDS-12-2015-0495>
- Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*. <https://doi.org/10.1057/jit.2014.17>
- Córdoba, J. R., Pilkington, A., & Bernroider, E. W. N. (2012). Information systems as a discipline in the making: Comparing EJIS and MISQ between 1995 and 2008. *European Journal of Information Systems*, 21(5). <https://doi.org/10.1057/ejis.2011.58>
- Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70, 379–390. <https://doi.org/10.1016/j.jbusres.2016.08.011>
- Cortina, J. M., & Landis, R. S. (2013). *Modern research methods for the study of behavior in organizations. Modern Research Methods for the Study of Behavior in Organizations*. <https://doi.org/10.4324/9780203585146>
- Court, D., & Barton, D. (2012). Making Advanced Analytics Work For You. *Harvard Business Review*.
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and Conducting Mixed Methods Research* / SAGE Publications Ltd. SAGE Publications, Inc.
- Davenport, T. H., Barth, P., & Bean, R. (2012). How “big data” is different. *MIT Sloan Management Review*.
- Davenport, T., & Harris, J. (2007). The architecture of business intelligence. *Competing on Analytics*.
- David, A. (2012). Logique, épistémologie et méthodologie en sciences de gestion. In *Les nouvelles fondations des sciences de gestion*. [https://doi.org/10.1016/S0166-2236\(02\)02264-6](https://doi.org/10.1016/S0166-2236(02)02264-6)
- Davis, F. (1989). Perceived Usefulness, Perceived Ease Of Use, And User Accep. *MIS*

Quarterly.

- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*.
<https://doi.org/10.2307/249008>
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982–1003. <https://doi.org/10.1287/mnsc.35.8.982>
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology*, 22(14).
<https://doi.org/10.1111/j.1559-1816.1992.tb00945.x>
- De Luca, L. M., Herhausen, D., Troilo, G., & Rossi, A. (2021). How and when do big data investments pay off? The role of marketing affordances and service innovation. *Journal of the Academy of Marketing Science*, 49(4). <https://doi.org/10.1007/s11747-020-00739-x>
- De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. In *AIP Conference Proceedings*.
<https://doi.org/10.1063/1.4907823>
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*. <https://doi.org/10.1108/LR-06-2015-0061>
- Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1). <https://doi.org/10.1145/1327452.1327492>
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*. <https://doi.org/10.1287/isre.3.1.60>
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. In *Journal of Management Information Systems* (Vol. 19, pp. 9–30). <https://doi.org/10.1080/07421222.2003.11045748>
- Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. *Information and Management*.
<https://doi.org/10.1016/j.im.2018.10.006>
- DeSanctis, G., & Poole, M. S. (1994). Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. *Organization Science*, 5(2).
<https://doi.org/10.1287/orsc.5.2.121>
- Doha .S. et Abdenbi. L. (2020). Contexte et contextualisation Mosaïque d’approches en management. *EMI 2020 Edition Management et Innovation*.
- Donate, M. J., & Sánchez de Pablo, J. D. (2015). The role of knowledge-oriented leadership in knowledge management practices and innovation. *Journal of Business Research*, 68(2).
<https://doi.org/10.1016/j.jbusres.2014.06.022>

- Dorr, B. J., Greenberg, C. S., Fontana, P., Przybocki, M., Le Bras, M., Ploehn, C., ... Chang, W. (2015). The NIST data science evaluation series: Part of the NIST information access division data science initiative. In *Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015*. <https://doi.org/10.1109/BigData.2015.7364096>
- Drucker-Godard, C., Ehlinger, S., & Grenier, C. (2014). Chapitre 10. Validité et fiabilité de la recherche. In *Méthodes de recherche en management*. <https://doi.org/10.3917/dunod.thiet.2014.01.0297>
- Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture. *British Journal of Management*. <https://doi.org/10.1111/1467-8551.12355>
- Dubey, R., Gunasekaran, A., Childe, S. J., Bryde, D. J., Giannakis, M., Foropon, C., ... Hazen, B. T. (2020). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2019.107599>
- Dubey, R., Gunasekaran, A., Childe, S. J., Fosso Wamba, S., Roubaud, D., & Foropon, C. (2019). Empirical investigation of data analytics capability and organizational flexibility as complements to supply chain resilience. *International Journal of Production Research*. <https://doi.org/10.1080/00207543.2019.1582820>
- Dutta, D., & Bose, I. (2015). Managing a big data project: The case of Ramco cements limited. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2014.12.032>
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897–904. <https://doi.org/10.1016/j.jbusres.2015.07.001>
- Esteves, J., & Curto, J. (2013). A risk and benefits behavioral model to assess intentions to adopt big data. *Journal of Intelligence Studies in Business*, 3(3). <https://doi.org/10.37380/jisib.v3i3.74>
- Fishbein, M., & Ajzen, I. (1975). Strategies of Change: Active Participation. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*.
- Fogelman-Soulié, F., & Lu, W. (2016). Implementing Big Data Analytics Projects in Business. https://doi.org/10.1007/978-3-319-26989-4_6
- Fosso Wamba, P. S. (2017). Big data analytics and business process innovation. *Business Process Management Journal*. <https://doi.org/10.1108/bpmj-02-2017-0046>
- Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How “big data” can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2014.12.031>
- Fotaki, M., Voudouris, I., Lioukas, S., & Zyglidopoulos, S. (2021). More Accountable, More

- Ethical, Yet Less Trusted: Misplaced Corporate Governance in the Era of Big Data. *British Journal of Management*, 32(4). <https://doi.org/10.1111/1467-8551.12447>
- Frimousse, S. (2020). Chapitre 8. Pour une épistémologie frontalière en sciences de gestion. In *Produire du savoir et de l'action*. <https://doi.org/10.3917/ems.peret.2020.01.0089>
- Frisk, J. E., & Bannister, F. (2017). Improving the use of analytics and big data by changing the decision-making culture: A design approach. *Management Decision*. <https://doi.org/10.1108/MD-07-2016-0460>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Gangwar, H. (2018). Understanding the Determinants of Big Data Adoption in India. *Information Resources Management Journal*, 31(4). <https://doi.org/10.4018/irmj.2018100101>
- Gantz, B. J., Reinsel, D., & Shadows, B. D. (2012). The Digital Universe in 2020. *IDC IView "Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East."*
- Gao, Jerry, Xie, C., & Tao, C. (2016). Big data validation and quality assurance - Issues, challenges, and needs. In *Proceedings - 2016 IEEE Symposium on Service-Oriented System Engineering, SOSE 2016*. <https://doi.org/10.1109/SOSE.2016.63>
- Gao, Jing, Koronios, A., & Selle, S. (2015). Towards a process view on critical success factors in Big Data analytics projects. In *2015 Americas Conference on Information Systems, AMCIS 2015*.
- Gasser, L. (1986). The Integration of Computing and Routine Work. *ACM Transactions on Information Systems (TOIS)*. <https://doi.org/10.1145/214427.214429>
- Gavard-Perret, M.-L., Gotteland, D., Haon, C., & Jolibert, A. (2008). *Méthodologie de la recherche, Réussir son mémoire ou sa thèse en sciences de gestion*. Pearson France.
- Gavard-Perret, M.-L., Gotteland, D., Haon, C., & Jolibert, A. (2012). *Méthodologie de la recherche en sciences de gestion*. Pearson France.
- Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and tam in online shopping: AN integrated model. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/30036519>
- Gefen, D., & Straub, D. (2000). The Relative Importance of PEOU in IS Adoption: A Study of E-Commerce Adoption. *JAIS*, 1(8).
- Gephart, R. P. (2013). Doing research with words: Qualitative methodologies and industrial/organizational psychology. In *Modern Research Methods for The Study of Behavior in Organizations*. <https://doi.org/10.4324/9780203585146>
- Ghasemaghahi, M. (2019a). Are firms ready to use big data analytics to create value? The role of

- structural and psychological readiness. *Enterprise Information Systems*, 13(5), 650–674. <https://doi.org/10.1080/17517575.2019.1576228>
- Ghasemaghaei, M. (2019b). Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency. *Decision Support Systems*. <https://doi.org/10.1016/j.dss.2019.03.004>
- Ghasemaghaei, M. (2020). The role of positive and negative valence factors on the impact of bigness of data on big data analytics usage. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2018.12.011>
- Ghasemaghaei, M. (2021). Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2019.102055>
- Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of big data on firm innovation performance: Big data is not always better data. *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2019.09.062>
- Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *Journal of Strategic Information Systems*. <https://doi.org/10.1016/j.jsis.2017.10.001>
- Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*. <https://doi.org/10.1016/j.dss.2017.06.004>
- Ghobakhloo, M., Hong, T. S., Sabouri, M. S., & Zulkifli, N. (2012). Strategies for successful information technology adoption in small and medium-sized enterprises. *Information (Switzerland)*. <https://doi.org/10.3390/info3010036>
- Ghobakhloo, M., Sadegh Sabouri, M., Sai Hong, T., & Zulkifli, N. (2011). Information Technology Adoption in Small and Medium-sized Enterprises; An Appraisal of Two Decades Literature. *Interdisciplinary Journal of Research in Business*.
- Giannelloni, J.-L., & Le Nagard, E. (2016). Big Data et marketing prédictif. Vers un « consommateur nu »? *Décisions Marketing*, 82. <https://doi.org/10.7193/dm.082.05.15>
- Glaser, B., & Strauss, A. (1967). Applying Grounded Theory. The discovery of grounded theory: strategies of qualitative research. *The Grounded Theory Review*, 13(1).
- Goes, P. B. (2014). Big Data and IS Research [Editor's Comments]. *MIS Quarterly*.
- Gohar Sargsyan. (2020). *La veille stratégique comme dispositif de soutien à l'innovation dans un contexte arménien*No Title.
- Golightly, D., Kefalidou, G., & Sharples, S. (2018). A cross-sector analysis of human and organisational factors in the deployment of data-driven predictive maintenance. *Information Systems and E-Business Management*, 16(3). <https://doi.org/10.1007/s10257-017-0343-1>

- Goodhue, D. L. (1995). Understanding User Evaluations of Information Systems. *Management Science*. <https://doi.org/10.1287/mnsc.41.12.1827>
- Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly: Management Information Systems*, 19(2). <https://doi.org/10.2307/249689>
- Gopalkrishnan, V., Steier, D., Lewis, H., & Guszczka, J. (2012). Big data, big business: Bridging the gap. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. <https://doi.org/10.1145/2351316.2351318>
- Gorla, N., Somers, T. M., & Wong, B. (2010). Organizational impact of system quality, information quality, and service quality. *Journal of Strategic Information Systems*, 19(3). <https://doi.org/10.1016/j.jsis.2010.05.001>
- Graupner, E., Berner, M., Maedche, A., & Jegadeesan, H. (2014). Business intelligence & analytics for processes - A visibility requirements evaluation. In *Tagungsband Multikonferenz Wirtschaftsinformatik 2014, MKWI 2014*.
- Greenwood, R., Oliver, C., Lawrence, T., & Meyer, R. (2018). *The SAGE Handbook of Organizational Institutionalism. The SAGE Handbook of Organizational Institutionalism*. <https://doi.org/10.4135/9781526415066>
- Gregor, S. (2006). The nature of theory in Information Systems. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/25148742>
- Gregor, S., Martin, M., Fernandez, W., Stern, S., & Vitale, M. (2006). The transformational dimension in the realization of business value from information technology. *Journal of Strategic Information Systems*. <https://doi.org/10.1016/j.jsis.2006.04.001>
- Grigoryan, E., & Lobre-Lebraty, K. (2020). VIANDOX Taxi et Big Data : un cas d'efficacité d'usage. In *AIM 2020 : Travailler et manager à l'ère des technologies émergentes*.
- Grublješič, T., & Jaklič, J. (2015). Conceptualization of the business intelligence extended use model. *Journal of Computer Information Systems*. <https://doi.org/10.1080/08874417.2015.11645774>
- Guimaraes, T., Igarria, M., & Lu, M. -t. (1992). The Determinants of DSS Success: An Integrated Model. *Decision Sciences*. <https://doi.org/10.1111/j.1540-5915.1992.tb00397.x>
- Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26(3), 191–209. <https://doi.org/10.1016/j.jsis.2017.07.003>
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, 53(8), 1049–1064. <https://doi.org/10.1016/j.im.2016.07.004>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate Data Analysis. *Vectors*. <https://doi.org/10.1016/j.ijpharm.2011.02.019>
- Halaweh, M., & Ahmed El Massry. (2015). Conceptual Model for Successful Implementation of

Big Data in Organizations. *Journal of International Technology and Information Management*, 24(2).

- Halaweh, M., & El Massry, A. (2017). A synergetic model for implementing big data in organizations: An empirical study. *Information Resources Management Journal*, 30(1). <https://doi.org/10.4018/IRMJ.2017010104>
- Hambrick, D. C., & Mason, P. A. (1984). Upper Echelons: The Organization as a Reflection of Its Top Managers. *Academy of Management Review*, 9(2). <https://doi.org/10.5465/amr.1984.4277628>
- Hamdam, A., Jusoh, R., Yahya, Y., Abdul Jalil, A., & Zainal Abidin, N. H. (2022). Auditor judgment and decision-making in big data environment: a proposed research framework. *Accounting Research Journal*, 35(1). <https://doi.org/10.1108/ARJ-04-2020-0078>
- Hashem, I. A. T., Anuar, N. B., Gani, A., Yaqoob, I., Xia, F., & Khan, S. U. (2016). MapReduce: Review and open challenges. *Scientometrics*, 109(1). <https://doi.org/10.1007/s11192-016-1945-y>
- Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., ... Chiroma, H. (2016). The role of big data in smart city. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2016.05.002>
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*. <https://doi.org/10.1016/j.is.2014.07.006>
- Hawley, D. (2016). Implementing Business Analytics within the Supply Chain: Success and Fault Factors. *Electronic Journal of Information Systems Evaluation*.
- Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2014.04.018>
- He, W., Tian, X., Hung, A., Akula, V., & Zhang, W. (2018). Measuring and comparing service quality metrics through social media analytics: a case study. *Information Systems and E-Business Management*, 16(3). <https://doi.org/10.1007/s10257-017-0360-0>
- Heckler, C. E., & Hatcher, L. (1996). A Step-by-Step Approach to Using the SAS® System for Factor Analysis and Structural Equation Modeling. *Technometrics*, 38(3). <https://doi.org/10.2307/1270628>
- Hey, T., Tansley, S., & Tolle, K. (2009). The Fourth Paradigm. *Data-Intensive Scientific Discovery*. Microsoft Research.
- Hong, K. K., & Kim, Y. G. (2002). The critical success factors for ERP implementation: An organizational fit perspective. *Information and Management*. [https://doi.org/10.1016/S0378-7206\(01\)00134-3](https://doi.org/10.1016/S0378-7206(01)00134-3)

- Hou, C. K. (2012). Examining the effect of user satisfaction on system usage and individual performance with business intelligence systems: An empirical study of Taiwan's electronics industry. *International Journal of Information Management*, 32(6).
<https://doi.org/10.1016/j.ijinfomgt.2012.03.001>
- Howell, J. M., & Higgins, C. A. (1990). Champions of Technological Innovation. *Administrative Science Quarterly*. <https://doi.org/10.2307/2393393>
- Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access*, 2. <https://doi.org/10.1109/ACCESS.2014.2332453>
- Huang, T. C. K., Liu, C. C., & Chang, D. C. (2012). An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2011.11.006>
- Hunt, S. D. (1990). Truth in Marketing Theory and Research. *Journal of Marketing*, 54(3).
<https://doi.org/10.1177/002224299005400301>
- Hunt, S. D. (1993). Objectivity in Marketing Theory and Research. *Journal of Marketing*, 57(2).
<https://doi.org/10.1177/002224299305700206>
- Hunt, S. D. (2005). For truth and realism in management research. *Journal of Management Inquiry*, 14(2). <https://doi.org/10.1177/1056492605275242>
- Igbaria, M., & Tan, M. (1997). The consequences of information technology acceptance on subsequent individual performance. *Information and Management*.
[https://doi.org/10.1016/S0378-7206\(97\)00006-2](https://doi.org/10.1016/S0378-7206(97)00006-2)
- Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2016.08.007>
- Jha, A. K., Agi, M. A. N., & Ngai, E. W. T. (2020). A note on big data analytics capability development in supply chain. *Decision Support Systems*, 138.
<https://doi.org/10.1016/j.dss.2020.113382>
- Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*.
<https://doi.org/10.1080/00207543.2016.1154209>
- Johnson, H. G., & Wiener, N. (1949). Cybernetics: Control and Communication in the Animal and the Machine. *The Economic Journal*, 59(236). <https://doi.org/10.2307/2226579>
- K. S. Cunningham. (1966). Standards. *Australian Journal of Education*, 10, 1–11.
- Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/249751>
- Karapetyan, M. S. (2003). ՀԱՅԱՍՏԱՆԸ 1912-1920 ԹՎԱԿԱՆՆԵՐԻՆ. Yerevan.

- Karoui, M., Davauchelle, G., & Duzert, A. (2014). Perspective and Challenges for Business. *Ingenierie Des Systemes d'Information*, 19(3), 73–92. <https://doi.org/10.3166/ISI.19.3.73-92>
- Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., ... Gani, A. (2014). Big data: Survey, technologies, opportunities, and challenges. *Scientific World Journal*. <https://doi.org/10.1155/2014/712826>
- Khan, Z., & Vorley, T. (2017). Big data text analytics: an enabler of knowledge management. *Journal of Knowledge Management*, 21(1). <https://doi.org/10.1108/JKM-06-2015-0238>
- Khatisyan, A. (1968). Հայաստանի Հանրապետության ծագումն ու զարգացումը. Beirut.
- Kim, B. (2012). The diffusion of mobile data services and applications: Exploring the role of habit and its antecedents. *Telecommunications Policy*, 36(1). <https://doi.org/10.1016/j.telpol.2011.11.011>
- Kim, G., Shin, B., Kim, K. K., & Lee, H. G. (2011). IT capabilities, process-oriented dynamic capabilities, and firm financial performance. *Journal of the Association for Information Systems*. <https://doi.org/10.17705/1jais.00270>
- Kim, M. K., & Park, J. H. (2017). Identifying and prioritizing critical factors for promoting the implementation and usage of big data in healthcare. *Information Development*, 33(3). <https://doi.org/10.1177/0266666916652671>
- Kiron, D. (2017). Lessons from Becoming a Data-Driven Organization. *MIT Sloan Management Review*, 58(2).
- Kiron, D., Prentice, P. K., & Ferguson, R. B. (2014). The Analytics Mandate. *MIT Sloan Management Review*.
- Kiron, D., & Shockley, R. (2011). Creating business value with analytics. *MIT Sloan Management Review*.
- Klein, B. D., Goodhue, D. L., & Davis, G. B. (1997). Can humans detect errors in data? impact of base rates, incentives, and goals. *MIS Quarterly: Management Information Systems*, 21(2). <https://doi.org/10.2307/249418>
- Kocabyk, S., & Erduran, I. (2000). The effect of valine substitution for glycine in the dimer interface of citrate synthase from thermoplasma acidophilum on stability and activity. In *Biochemical and Biophysical Research Communications* (Vol. 275, pp. 460–465). <https://doi.org/10.1006/bbrc.2000.3310>
- Kreps, J., Narkhede, N., & Rao, J. (2011). Kafka: a Distributed Messaging System for Log Processing. *ACM SIGMOD Workshop on Networking Meets Databases*.
- Krishnan, T. N., & Scullion, H. (2017). Talent management and dynamic view of talent in small and medium enterprises. *Human Resource Management Review*, 27(3). <https://doi.org/10.1016/j.hrmr.2016.10.003>

- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Kung, L., Kung, H. J., Jones-Farmer, A., & Wang, Y. (2015). Managing Big data for firm performance: A configurational approach. In *2015 Americas Conference on Information Systems, AMCIS 2015*.
- Kwahk, K. Y., & Lee, J. N. (2008). The role of readiness for change in ERP implementation: Theoretical bases and empirical validation. *Information and Management*.
<https://doi.org/10.1016/j.im.2008.07.002>
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2014.02.002>
- Lai, Y., Sun, H., & Ren, J. (2018). Understanding the determinants of big data analytics (BDA) adoption in logistics and supply chain management: An empirical investigation. *International Journal of Logistics Management*. <https://doi.org/10.1108/IJLM-06-2017-0153>
- Lamb, R., & Kling, R. (2003). Reconceptualizing users as social actors in information systems research. *MIS Quarterly: Management Information Systems*, 27(2).
<https://doi.org/10.2307/30036529>
- Landset, S., Khoshgoftaar, T. M., Richter, A. N., & Hasanin, T. (2015). A survey of open source tools for machine learning with big data in the Hadoop ecosystem. *Journal of Big Data*, 2(1). <https://doi.org/10.1186/s40537-015-0032-1>
- Lankton, N. K., Harrison Mcknight, D., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*.
<https://doi.org/10.17705/1jais.00411>
- Lave, J., & Wenger, E. (1991). *Situated Learning Legitimate Peripheral Participation | Developmental psychology |*. Cambridge University Press.
- Le Moigne, J.-L. (1994). Sur la capacité de la raison à discerner rationalité substantive et rationalité procédurale. *Cahiers d'économie Politique*, 24(1).
<https://doi.org/10.3406/cep.1994.1168>
- Leavitt, N. (2013). Bringing big analytics to the masses. *Computer*, 46(1).
<https://doi.org/10.1109/mc.2013.9>
- Lebraty, J. F. (2011). Du concept d'information à celui de systèmes d'information. In *Master : systèmes d'Information* (Thévenot, pp. 15–34). Paris: ESKA.
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management*.
[https://doi.org/10.1016/S0378-7206\(02\)00043-5](https://doi.org/10.1016/S0378-7206(02)00043-5)
- Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information and Management*, 40(3),

- 191–204. [https://doi.org/10.1016/S0378-7206\(01\)00143-4](https://doi.org/10.1016/S0378-7206(01)00143-4)
- Lépinard, P., Lebraty, J.-F., & Lobre, K. (2013). 18ème Colloque de l'Association Information et Management. In *Big Data ou Big Bazar ?* Lyon.
- Li, C., Aggarwal, C. C., & Wang, J. (2011). On anonymization of multi-graphs. In *Proceedings of the 11th SIAM International Conference on Data Mining, SDM 2011*.
<https://doi.org/10.1137/1.9781611972818.61>
- Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly: Management Information Systems*, 31(1). <https://doi.org/10.2307/25148781>
- Liang, H., & Xue, Y. (2009). Avoidance of information technology threats: A theoretical perspective. *MIS Quarterly: Management Information Systems*, 33(1).
<https://doi.org/10.2307/20650279>
- Liu, Y., Wang, W., & Zhang, Z. (Justin). (2022). The dual drivetrain model of digital transformation: role of industrial big-data-based affordance. *Management Decision*, 60(2).
<https://doi.org/10.1108/MD-12-2019-1664>
- Livian, Y. F. (2020). Contextualiser le management africain ? Oui, mais comment ? *Revue Francaise de Gestion*. <https://doi.org/10.3166/rfg.2020.00445>
- Lobre, K., & Lebraty, J.-F. (2012). L'open Data : nouvelle pratique managériale risquée ? *Gestion 2000*, (4).
- Loia, V., Pedrycz, W., Senatore, S., & Sessa, M. I. (2006). Web navigation support by means of proximity-driven assistant agents. *Journal of the American Society for Information Science and Technology*, 57(4). <https://doi.org/10.1002/asi.20306>
- Longwell, H. C., & Rousselot, P. (1911). L'intellectualisme de Saint Thomas. *The Philosophical Review*, 20(3). <https://doi.org/10.2307/2177857>
- Louitri, A., & Sahraoui, D. (2014). Contexte et contextualisation dans les recherches en Management au Maroc. *Revue Marocaine Des Sciences de Management*.
- Love, P. E. D., Irani, Z., Standing, C., Lin, C., & Burn, J. M. (2005). The enigma of evaluation: benefits, costs and risks of IT in Australian small–medium-sized enterprises. *Information & Management*, 42(7), 947–964.
- Lynn Segarra, L., Almalki, H., Elabd, J., Gonzalez, J., Marczewski, M., Alrasheed, M., & Rabelo, L. (2016). A Framework for Boosting Revenue Incorporating Big Data. *Journal of Innovation Management*, 4(1). https://doi.org/10.24840/2183-0606_004.001_0005
- Madlberger, M. (2009). A model of antecedents of RFID adoption intention in the supply chain. In *Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS*. <https://doi.org/10.1109/HICSS.2009.22>
- Maduku, D. K., Mpinganjira, M., & Duh, H. (2016). Understanding mobile marketing adoption

- intention by South African SMEs: A multi-perspective framework. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2016.04.018>
- Majchrzak, A., Lynne Markus, M., & Wareham, J. (2016). Designing for digital transformation: Lessons for information systems research from the study of ICT and societal challenges. *MIS Quarterly: Management Information Systems*, *40*(2), 267–277. <https://doi.org/10.25300/MISQ/2016/40>
- Marcolin, B. L., Compeau, D. R., Munro, M. C., & Huff, S. L. (2000). Assessing User Competence: Conceptualization and Measurement. *Information Systems Research*, *11*(1), 37–60. <https://doi.org/10.1287/isre.11.1.37.11782>
- Markus, M. L., & Topi, H. (2015). Big Data, Big Decisions for Government, Business, and Society. Report on a Research Agenda Setting Workshop. *Technical Report*. National Science Foundation, USA.
- Maroufkhani, P., Tseng, M. L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2020.102190>
- Maroufkhani, P., Wan Ismail, W. K., & Ghobakhloo, M. (2020). Big data analytics adoption model for small and medium enterprises. *Journal of Science and Technology Policy Management*. <https://doi.org/10.1108/JSTPM-02-2020-0018>
- Marr, B. (2016). *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics*. John Wiley & Sons. Retrieved from https://books.google.ch/books/about/Big_Data_in_Practice.html?id=UIU7jgEACAAJ&redir_esc=y%0Ahttps://ebookcentral-proquest-com.ezproxy.hamk.fi/lib/hamk-ebooks/reader.action?docID=4455265
- Marshall, A., Mueck, S., & Shockley, R. (2015). How leading organizations use big data and analytics to innovate. *Strategy and Leadership*. <https://doi.org/10.1108/SL-06-2015-0054>
- Martinet, A. C. (2015). Chapitre 1. La recherche en management : une épistémologie pragmatiste. In *La création de connaissance par les managers* (pp. 23–52). <https://doi.org/10.3917/ems.beaul.2015.01.0023>
- Mason, R. O. (1978). Measuring information output: A communication systems approach. *Information and Management*. [https://doi.org/10.1016/0378-7206\(78\)90028-9](https://doi.org/10.1016/0378-7206(78)90028-9)
- Mawed, M., & Aal-Hajj, A. (2017). Using big data to improve the performance management: a case study from the UAE FM industry. *Facilities*, *35*(13–14). <https://doi.org/10.1108/F-01-2016-0006>
- Maximilien Dossa. (2019). *Aide à la modélisation et au traitement de données massives : proposition d'un guide méthodologique*.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Boston, MA: Houghton Mifflin Harcourt.

- Mbengue, A., & Vandangeon-Derumez, I. (1999). Positions épistémologiques et outils de recherche en management stratégique. In *Conférence de l'AIMS, Ecole Centrale de Paris, 26-28 Mai*.
- McAfee, A., & Brynjolfsson, E. (2012a). Big data: the management revolution. *Harvard Business Review*. <https://doi.org/10.1007/s12599-013-0249-5>
- McAfee, A., & Brynjolfsson, E. (2012b). Big data: The management revolution. *Harvard Business Review*.
- Mcfarlan, F. W. (1981). Portfolio approach to information systems. *Harvard Business Review*, 59(5).
- Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of it business value. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/25148636>
- Merino, J., Caballero, I., Rivas, B., Serrano, M., & Piattini, M. (2016). A Data Quality in Use model for Big Data. *Future Generation Computer Systems*, 63, 123–130. <https://doi.org/10.1016/j.future.2015.11.024>
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019a). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. <https://doi.org/10.1016/j.jbusres.2019.01.044>
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019b). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management*, 30(2), 272–298. <https://doi.org/10.1111/1467-8551.12343>
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information and Management*. <https://doi.org/10.1016/j.im.2019.05.004>
- Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and E-Business Management*, 16(3), 547–578. <https://doi.org/10.1007/s10257-017-0362-y>
- Mikalef, P., Pappas, I. O., Krogstie, J., & Pavlou, P. A. (2020). Big data and business analytics: A research agenda for realizing business value. *Information and Management*. <https://doi.org/10.1016/j.im.2019.103237>
- Mikalef, P., van de Wetering, R., & Krogstie, J. (2021). Building dynamic capabilities by leveraging big data analytics: The role of organizational inertia. *Information and Management*, 58(6). <https://doi.org/10.1016/j.im.2020.103412>
- Miles, M. B., & Huberman, M. A. (2003). *Analyse des données qualitatives. Méthodes en sciences humaines*.

- Mithas, S., Lee, M. R., Earley, S., Murugesan, S., & Djavanshir, R. (2013). Leveraging big data and business analytics. *IT Professional*. <https://doi.org/10.1109/MITP.2013.95>
- Molla, A., & Licker, P. (2001). E-commerce systems success: An attempt to partially extend and respecify the Delone and Mclean model of IS success. *Journal of Electronic Commerce Research*.
- Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*. <https://doi.org/10.1287/isre.2.3.192>
- Munawar, H. S., Ullah, F., Qayyum, S., & Shahzad, D. (2022). Big Data in Construction: Current Applications and Future Opportunities. *Big Data and Cognitive Computing*. <https://doi.org/10.3390/bdcc6010018>
- Muniz Félix, B., Tavares, E., & Freitas Cavalcante, N. W. (2018). Critical success factors for Big Data adoption in the virtual retail: Magazine Luiza case study. *Review of Business Management*. <https://doi.org/10.7819/rbgn.v20i1.3627>
- Mutabazi, E. (2006). Face à la diversité des cultures et des modes de gestion : le modèle circulaire de management en Afrique. *Management & Avenir, n° 10(4)*. <https://doi.org/10.3917/mav.010.0179>
- Nanopoulos, A., Zakrzewicz, M., Morzy, T., & Manolopoulos, Y. (2002). Indexing Web Access-Logs for Pattern Queries. In *Proceedings of the International Workshop on Web Information and Data Management*. <https://doi.org/10.1145/584943.584945>
- Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of Management Information Systems*. <https://doi.org/10.1080/07421222.2005.11045823>
- Neumeyer, L., Robbins, B., Nair, A., & Kesari, A. (2010). S4: Distributed stream computing platform. In *Proceedings - IEEE International Conference on Data Mining, ICDM*. <https://doi.org/10.1109/ICDMW.2010.172>
- Noël, A. (2012). *Conduite d'une recherche : mémoires d'un directeur*.
- Oh, W., & Pinsonneault, A. (2007). On the assessment of the strategic value of information technologies: Conceptual and analytical approaches. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/25148790>
- Okwechime, E., Duncan, P., & Edgar, D. (2018). Big data and smart cities: a public sector organizational learning perspective. *Information Systems and E-Business Management, 16(3)*. <https://doi.org/10.1007/s10257-017-0344-0>
- Oliveira, S. F., Furlinger, K., & Kranzlmüller, D. (2012). Trends in computation, communication and storage and the consequences for data-intensive science. In *Proceedings of the 14th IEEE International Conference on High Performance Computing and Communications, HPCC-2012 - 9th IEEE International Conference on Embedded Software and Systems, ICESS-2012*. <https://doi.org/10.1109/HPCC.2012.83>

- Orlikowski, W. J. (2000). Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. *Organization Science*, 11(4), 404–428.
<https://doi.org/10.1287/orsc.11.4.404.14600>
- Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, 2(1).
<https://doi.org/10.1287/isre.2.1.1>
- Ostrom, E. (2009). *Understanding institutional diversity. Understanding Institutional Diversity*.
<https://doi.org/10.1057/palgrave.ces.8100208>
- Papadopoulos, T., & Balta, M. E. (2022). Climate Change and big data analytics: Challenges and opportunities. *International Journal of Information Management*, 63.
<https://doi.org/10.1016/j.ijinfomgt.2021.102448>
- Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL a multiple-item scale for assessing electronic service quality. *Journal of Service Research*.
<https://doi.org/10.1177/1094670504271156>
- Parsons, J., & Cole, L. (2005). What do the pictures mean? Guidelines for experimental evaluation of representation fidelity in diagrammatical conceptual modeling techniques. In *Data and Knowledge Engineering* (Vol. 55). <https://doi.org/10.1016/j.datak.2004.12.008>
- Pastorelli, I. (2000). Quelles pratiques pour une connaissance fondée sur l'action ? Le cas d'une recherche-action en contrôle. *Comptabilité - Contrôle - Audit*, 6(3).
<https://doi.org/10.3917/cca.063.0095>
- Pauleen, D. J., & Wang, W. Y. C. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. *Journal of Knowledge Management*.
<https://doi.org/10.1108/JKM-08-2016-0339>
- Perret V. et Seville M. (2007). Fondements épistémologiques de la recherche. *Recherche En Management*, (Dunod), 13–33.
- Petter, S., Delone, W., & McLean, E. R. (2013). Information systems success: The quest for the independent variables. *Journal of Management Information Systems*.
<https://doi.org/10.2753/MIS0742-1222290401>
- Philip Chen, C. L., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314–347.
<https://doi.org/10.1016/j.ins.2014.01.015>
- Philippe Wanlin. (2007). L'analyse de contenu comme méthode d'analyse qualitative d'entretiens : une comparaison entre les traitements manuels et l'utilisation de logiciels. In *RECHERCHES QUALITATIVES – Hors Série – numéro 3*. Association pour la recherche qualitative.
- Phillips-Wren, G., & Hoskisson, A. (2015). An analytical journey towards big data. *Journal of Decision Systems*. <https://doi.org/10.1080/12460125.2015.994333>

- Piaget J. (1967). *Logique et Connaissance scientifique*. Paris: Gallimard.
- Pickering, M. (2011). Le positivisme philosophique : Auguste Comte. *Revue Interdisciplinaire d'études Juridiques*, 67(2). <https://doi.org/10.3917/riej.067.0049>
- Pierre Romelaer, & Michel Kalika. (2016). *Comment réussir sa thèse*.
- Pitoura, E., & Bhargava, B. (1999). Data consistency in intermittently connected distributed systems. *IEEE Transactions on Knowledge and Data Engineering*, 11(6). <https://doi.org/10.1109/69.824602>
- Polites, & Karahanna. (2012). Shackled to the Status Quo: The Inhibiting Effects of Incumbent System Habit, Switching Costs, and Inertia on New System Acceptance. *MIS Quarterly*. <https://doi.org/10.2307/41410404>
- Popovič, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, 20(2), 209–222. <https://doi.org/10.1007/s10796-016-9720-4>
- Popper, E. R. (1959). The Logic of Scientific Discovery. *Physics Today*, 12(11). <https://doi.org/10.1063/1.3060577>
- Premkumar, G., & Ramamurthy, K. (1995). The Role of Interorganizational and Organizational Factors on the Decision Mode for Adoption of Interorganizational Systems. *Decision Sciences*. <https://doi.org/10.1111/j.1540-5915.1995.tb01431.x>
- Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *Eurasip Journal on Advances in Signal Processing*. <https://doi.org/10.1186/s13634-016-0355-x>
- Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187–195.
- Raguseo, Elisabetta. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1). <https://doi.org/10.1016/j.ijinfomgt.2017.07.008>
- Raguseo, Elisabetta, Vitari, C., & Pigni, F. (2020). Profiting from big data analytics: The moderating roles of industry concentration and firm size. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2020.107758>
- Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/25148729>
- Ralston, D. A. (2008). The crossvergence perspective: Reflections and projections. *Journal of International Business Studies*, 39(1). <https://doi.org/10.1057/palgrave.jibs.8400333>
- Ramamurthy, K. (Ram), Sen, A., & Sinha, A. P. (2008). An empirical investigation of the key

- determinants of data warehouse adoption. *Decision Support Systems*.
<https://doi.org/10.1016/j.dss.2007.10.006>
- Ranjan, J., & Foropon, C. (2021). Big Data Analytics in Building the Competitive Intelligence of Organizations. *International Journal of Information Management*.
<https://doi.org/10.1016/j.ijinfomgt.2020.102231>
- Rao, T. R., Mitra, P., Bhatt, R., & Goswami, A. (2019). The big data system, components, tools, and technologies: a survey. *Knowledge and Information Systems*, 60(3).
<https://doi.org/10.1007/s10115-018-1248-0>
- Reeves, C. A., & Bednar, D. A. (1994). DEFINING QUALITY: ALTERNATIVES AND IMPLICATIONS. *Academy of Management Review*.
<https://doi.org/10.5465/amr.1994.9412271805>
- Riccucci, N. M. (2010). Envisioning public administration as a scholarly field in 2020: Rethinking epistemic traditions. *Public Administration Review*, 70(SUPPL. 1).
<https://doi.org/10.1111/j.1540-6210.2010.02294.x>
- Robey, D. (1979). User Attitudes and Management Information System Use. *Academy of Management Journal*. <https://doi.org/10.5465/255742>
- Rogers, E. M. (1983). *Diffusion of innovations Third Edition. An Integrated Approach to Communication Theory and Research, Third Edition*.
- Rogers, E. M. (1995). *Diffusion of Innovations, Fourth Edition. Elements of Diffusion*.
- Rogers EM. (1983). *Diffusion of Innovations (4th ed.): The Free Press, New York, NY. New York NY Free Press*.
- Ronda-Pupo, G. A., & Guerras-Martin, L. A. (2012). Dynamics of the evolution of the strategy concept 1962-2008: A co-word analysis. *Strategic Management Journal*, 33(2).
<https://doi.org/10.1002/smj.948>
- Rondeau, P. J., Ragu-Nathan, T. S., & Vonderembse, M. A. (2006). How involvement, IS management effectiveness, and end-user computing impact IS performance in manufacturing firms. *Information and Management*.
<https://doi.org/10.1016/j.im.2005.02.001>
- Russom, P. (2013). Manageing Big Data. *Tdwi.Org*.
- Sahraoui Bentaleb, D., & Bentaleb, C. (2020). Chapitre 15. La contextualisation de la recherche : une démarche processuelle. In *Produire du savoir et de l'action*.
<https://doi.org/10.3917/ems.peret.2020.01.0175>
- Saltz, J. S., & Shamshurin, I. (2016). Big data team process methodologies: A literature review and the identification of key factors for a project's success. In *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016*.
<https://doi.org/10.1109/BigData.2016.7840936>

- Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review*, 58(3). <https://doi.org/10.1525/cmr.2016.58.3.26>
- Seddon, J. J. J. M., & Currie, W. L. (2017). A model for unpacking big data analytics in high-frequency trading. *Journal of Business Research*, 70, 300–307. <https://doi.org/10.1016/j.jbusres.2016.08.003>
- Seddon, P. B. (1997). A Respecification and Extension of the DeLone and McLean Model of IS Success. *Information Systems Research*, 8(3), 240–253. <https://doi.org/10.1287/isre.8.3.240>
- Sejahtera, F. P., Wang, W., Indulska, M., & Sadiq, S. (2018). Enablers and inhibitors of effective use of big data: Insights from a case study. In *Proceedings of the 22nd Pacific Asia Conference on Information Systems - Opportunities and Challenges for the Digitized Society: Are We Ready?, PACIS 2018*.
- Sejahtera, F., Wang, W., Indulska, M., & Sadiq, S. (2018). Enablers and Inhibitors of Effective Use of Big Data: Insights from a Case Study Harmonized Messaging View project Effective Use of Big Data View project. *PACIS 2018 Proceedings*. 27, (July).
- Sena, V., Bhaumik, S., Sengupta, A., & Demirbag, M. (2019). Big Data and Performance: What Can Management Research Tell us? *British Journal of Management*, Vol. 30, 2.
- Setia, P., & Patel, P. C. (2013). How information systems help create OM capabilities: Consequents and antecedents of operational absorptive capacity. *Journal of Operations Management*. <https://doi.org/10.1016/j.jom.2013.07.013>
- Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management*. <https://doi.org/10.1016/j.im.2018.12.003>
- Shamim, Saqib, Cang, S., & Yu, H. (2019). Impact of knowledge oriented leadership on knowledge management behaviour through employee work attitudes. *International Journal of Human Resource Management*. <https://doi.org/10.1080/09585192.2017.1323772>
- Shannon, C. E., & warren weaver. (1949). The mathematical theory of communication. *The University Of Illinois Press*. <https://doi.org/10.1145/584091.584093>
- Sharda, R., Delen, D., Turban, E., & King, D. (2014). Business intelligence: a managerial perspective on analytics. *Pearson*.
- Shaul, L., & Tauber, D. (2013). Critical success factors in enterprise resource planning systems: Review of the last decade. *ACM Computing Surveys*. <https://doi.org/10.1145/2501654.2501669>
- Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2017.06.006>
- Shim, J. P., French, A. M., Guo, C., & Jablonski, J. (2015). Big data and analytics: Issues,

- solutions, and ROI. *Communications of the Association for Information Systems*.
- Shin, D. H. (2016). Demystifying big data: Anatomy of big data developmental process. *Telecommunications Policy*, 40(9). <https://doi.org/10.1016/j.telpol.2015.03.007>
- Shuradze, G., & Wagner, H. T. (2016). Towards a conceptualization of data analytics capabilities. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2016.626>
- Siddiqa, A., Hashem, I. A. T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., & Nasaruddin, F. (2016). A survey of big data management: Taxonomy and state-of-the-art. *Journal of Network and Computer Applications*. <https://doi.org/10.1016/j.jnca.2016.04.008>
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286. <https://doi.org/10.1016/j.jbusres.2016.08.001>
- Sonntag, S. (2002). *Psychological Management of Individual Performance* (John Wiley). Chichester, UK: John Wiley & Sons, Chichester, UK.
- Soon, K. W. K., Lee, C. A., & Boursier, P. (2016). A study of the determinants affecting adoption of big data using integrated Technology Acceptance Model (TAM) and diffusion of innovation (DOI) in Malaysia. *International Journal of Applied Business and Economic Research*, 14(1).
- Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. *Production and Operations Management*. <https://doi.org/10.1111/poms.12746>
- Sumbal, M. S., Tsui, E., & See-to, E. W. K. (2017). Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector. *Journal of Knowledge Management*, 21(1), 180–196. <https://doi.org/10.1108/JKM-07-2016-0262>
- Surbakti, F. P. S. (2020). What is Effective Use of Big Data? the Consensual Definition of Effective Use of Big Data. In *IOP Conference Series: Materials Science and Engineering*. <https://doi.org/10.1088/1757-899X/847/1/012003>
- Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2019). Factors influencing effective use of big data: A research framework. *Information and Management*, 57(1). <https://doi.org/10.1016/j.im.2019.02.001>
- Tai, C. H., Yu, P. S., Yang, D. N., & Chen, M. S. (2011). Structural diversity for privacy in publishing social networks. In *Proceedings of the 11th SIAM International Conference on Data Mining, SDM 2011*. <https://doi.org/10.1137/1.9781611972818.4>
- Tambe, P. (2014). Big Data Investment, Skills, and Firm Value. *Management Science*. <https://doi.org/10.1287/mnsc.2014.1899>
- Taylor, S., & Todd, P. (1995). Assessing IT usage: The role of prior experience. *MIS Quarterly*:

- Management Information Systems*. <https://doi.org/10.2307/249633>
- Ternon, Y. (1983). *La Cause arménienne. La Cause arménienne*. <https://doi.org/10.3917/lsrel.terno.1983.01>
- Thiétart, R.-A. (2014). *Méthodes de recherche en management. Méthodes de recherche en management*. <https://doi.org/10.3917/dunod.thiet.2014.01>
- Thiétart, R. (2014). *Méthodes de recherche en management. Paris: Dunod*. <https://doi.org/10.3917/dunod.thiet.2014.01>
- Thompson, N., Ravindran, R., & Nicosia, S. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly*, 32(3). <https://doi.org/10.1016/j.giq.2015.05.001>
- Thusoo, A., Sarma, J. Sen, Jain, N., Shao, Z., Chakka, P., Anthony, S., ... Murthy, R. (2009). Hive - A warehousing solution over a map-reduce framework. *Proceedings of the VLDB Endowment*, 2(2). <https://doi.org/10.14778/1687553.1687609>
- Tolle, K. M., Tansley, D. S. W., & Hey, A. J. G. (2011). The fourth Paradigm: Data-intensive scientific discovery. In *Proceedings of the IEEE* (Vol. 99). <https://doi.org/10.1109/JPROC.2011.2155130>
- Tornatzky, L. G., & Klein, K. J. (1982). Innovation Characteristics And Innovation Adoption-Implementation: A Meta-Analysis Of Findings. *Ieee Transactions On Engineering Management*. <https://doi.org/10.1109/tem.1982.6447463>
- Triandis. (1977). Triandis' Theory of Interpersonal Behaviour. *Brooks/Cole Pub. Co.*
- Trieu, T. V. H. (2013). Extending the theory of effective use: The impact of enterprise architecture maturity stages on the effective use of business intelligence systems. In *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information Systems Design*.
- Utterback, J. M., & Acee, H. J. (2005). Disruptive Technologies: an Expanded View. *International Journal of Innovation Management*, 09(01), 1–17. <https://doi.org/10.1142/s1363919605001162>
- Vassakis, K., Petrakis, E., & Kopanakis, I. (2018). Big Data Analytics: Applications, Prospects and Challenges. *Mobile Big Data. Lecture Notes on Data Engineering and Communications Technologies, Springer, Cham, 10*, 3–20. https://doi.org/10.1007/978-3-319-67925-9_1
- Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>
- Venkatesh, V., & Davis, F. D. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of

- information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.
<https://doi.org/10.2307/30036540>
- Verma, S. (2017). The Adoption Of Big Data Services By Manufacturing Firms: An Empirical Investigation In India. *Journal of Information Systems and Technology Management*, (1).
<https://doi.org/10.5465/ambpp.2016.12647abstract>
- Verma, S., & Bhattacharyya, S. S. (2017). Perceived strategic value-based adoption of Big Data Analytics in emerging economy. *Journal of Enterprise Information Management*.
<https://doi.org/10.1108/jeim-10-2015-0099>
- W. Form, R.L. Kaufman, T.L. Parcel, M. W. (1988). The Impact of Technology on Work Organization and Work Outcomes. Industries, Firms, and Jobs (pp. 303–328). Springer New York LLC.
- Wamba, S. F., & Akter, S. (2015). Big data analytics for supply chain management: A literature review and research agenda. In *Lecture Notes in Business Information Processing*.
https://doi.org/10.1007/978-3-319-24626-0_5
- Wamba, S. F., Dubey, R., Gunasekaran, A., & Akter, S. (2020). The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism. *International Journal of Production Economics*.
<https://doi.org/10.1016/j.ijpe.2019.09.019>
- Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11). <https://doi.org/10.1145/240455.240479>
- Wand, Y., & Weber, R. (1988). An Ontological Analysis Of Some Fundamental Information Systems Concepts. In *ICIS 1988 Proceedings* (Vol. 35).
- Wand, Y., & Weber, R. (1990). An Ontological Model of an Information System. *IEEE Transactions on Software Engineering*, 16(11), 1282–1292.
<https://doi.org/10.1109/32.60316>
- Wand, Y., & Weber, R. (1995). On the deep structure of information systems. *Information Systems Journal*. <https://doi.org/10.1111/j.1365-2575.1995.tb00108.x>
- Wang, R. Y., and Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*. *Journal of Management Information Systems*, 12(4).
- Wang, B. L., Fang, B. X., & Yun, X. C. (2005). Study and implementation of zero-copy packet capture platform. *Jisuanji Xuebao/Chinese Journal of Computers*, 28(1).
- Wang, F., & Liu, J. (2011). Networked wireless sensor data collection: Issues, challenges, and approaches. *IEEE Communications Surveys and Tutorials*.
<https://doi.org/10.1109/SURV.2011.060710.00066>
- Wang, J., Xu, C., Zhang, J., & Zhong, R. (2022). Big data analytics for intelligent manufacturing systems: A review. *Journal of Manufacturing Systems*, 62.

<https://doi.org/10.1016/j.jmsy.2021.03.005>

- Wang, Y., Kung, L. A., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126. <https://doi.org/10.1016/j.techfore.2015.12.019>
- Weber, R. (2004). The rhetoric of positivism versus interpretivism: A personal view. *Material Religion*, 2(3).
- Weber, Ron. (2003). Editor's Comments: Still Desperately Seeking the IT Artifact. *Management Information Systems*. <https://doi.org/10.1177/0899764098272001>
- Weirich, P. (2005). Annie Petit (Editor). Auguste Comte: Trajectoires positivistes 1798–1998. (Épistémologie et Philosophie des Sciences.) 438 pp., bibl., index. Paris: L'Harmattan, 2003. . *Isis*, 96(3). <https://doi.org/10.1086/498813>
- White, K. B., & Leifer, R. (1986). Information systems development success: Perspectives from project team participants. *MIS Quarterly: Management Information Systems*, 10(3). <https://doi.org/10.2307/249253>
- Whitley, R. (1999). *Divergent Capitalisms. Higher Education* (Vol. 6).
- Wielki, J. (2013). Implementation of the Big Data concept in organizations - Possibilities, impediments and challenges. In *2013 Federated Conference on Computer Science and Information Systems, FedCSIS 2013*.
- Willaert, P., Van Den Bergh, J., Willems, J., & Deschoolmeester, D. (2007). The process-oriented organisation: A holistic view developing a framework for business process orientation maturity. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 4714 LNCS). https://doi.org/10.1007/978-3-540-75183-0_1
- Willcocks, L., & Margetts, H. (1994). Risk assessment and information systems. *European Journal of Information Systems*, 3(2). <https://doi.org/10.1057/ejis.1994.13>
- Winograd, T., & Flores, F. (1987). On understanding computers and cognition: A new foundation for design. A response to the reviews. *Artificial Intelligence*. [https://doi.org/10.1016/0004-3702\(87\)90026-9](https://doi.org/10.1016/0004-3702(87)90026-9)
- Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*. <https://doi.org/10.1287/isre.1050.0042>
- Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/3250957>
- Wu, X., Liang, L., & Chen, S. (2022). How big data alters value creation: through the lens of big data competency. *Management Decision*, 60(3). <https://doi.org/10.1108/MD-09-2021-1199>

- Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, 69(5). <https://doi.org/10.1016/j.jbusres.2015.10.017>
- Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016a). Big data: From beginning to future. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2016.07.009>
- Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016b). Big data: From beginning to future. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2016.07.009>
- Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016c, December 1). Big data: From beginning to future. *International Journal of Information Management*. Elsevier Ltd. <https://doi.org/10.1016/j.ijinfomgt.2016.07.009>
- Yesayan, A. (2016). “Հսկանականության տեսության և մարտնաստիկական վիճակագրության դասականի մասին արևմտահայաստանյայցիները” *"Les particularités d'enseignement de la théorie de probabilité et de statistique mathématique"*. Yerevan.
- Ylijoki, O., & Porras, J. (2016). Conceptualizing Big Data: Analysis of Case Studies. *Intelligent Systems in Accounting, Finance and Management*. <https://doi.org/10.1002/isaf.1393>
- Yoon, Y., Guimaraes, T., & O'Neal, Q. (1995). Exploring the factors associated with expert systems success. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.1590/s0104-530x1997000100001>
- Yousfi, H. (2014). Rethinking Hybridity in Postcolonial Contexts: What Changes and What Persists? The Tunisian case of Poulina's managers. *Organization Studies*, 35(3). <https://doi.org/10.1177/0170840613499751>
- Zhan, Y., Tan, K. H., Ji, G., Chung, L., & Tseng, M. (2017). A big data framework for facilitating product innovation processes. *Business Process Management Journal*, 23(3). <https://doi.org/10.1108/BPMJ-11-2015-0157>
- Zhang, L., Stoffel, A., Behrisch, M., Mittelstädt, S., Schreck, T., Pompl, R., ... Keim, D. (2012). Visual analytics for the big data era - A comparative review of state-of-the-art commercial systems. In *IEEE Conference on Visual Analytics Science and Technology 2012, VAST 2012 - Proceedings*. <https://doi.org/10.1109/VAST.2012.6400554>
- Zheng, Y., Zhao, K., & Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision Support Systems*. <https://doi.org/10.1016/j.dss.2012.11.008>
- Zhou, Z., Dou, W., Jia, G., Hu, C., Xu, X., Wu, X., & Pan, J. (2016). A method for real-time trajectory monitoring to improve taxi service using GPS big data. *Information and Management*, 53(8), 964–977. <https://doi.org/10.1016/j.im.2016.04.004>
- Zhu, K., & Kraemer, K. L. (2005). Post-adoption variations in usage and value of e-business by

organizations: Cross-country evidence from the retail industry. *Information Systems Research*. <https://doi.org/10.1287/isre.1050.0045>

Zikopoulos, P. C., & Chris Eaton. (2011). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill Osborne Media.

Table des matières

Introduction générale	17
Première partie : Le choix d'un cadre théorique au big data : L'efficacité d'usage	25
Chapitre 1 Utilisation du big data par les entreprises	27
1.1 La genèse et les définitions du big data.....	28
1.1.1 Définition du big data à partir de ses caractéristiques spécifiques	30
1.1.2 Définition du big data par sa dimension technologique.....	34
1.1.3 Autres dimensions pour définir le big data	35
1.2 Les principales sources du big data.....	37
1.3 L'usage et le traitement du big data	39
1.3.1 Le traitement du big data	40
1.3.2 Les applications du big data.....	42
1.3.3 Les systèmes de stockage du big data.....	50
1.3.4 Traitement du big data par la société JKL Armenia	52
Chapitre 2 Efficacité d'usage: Fondements théoriques	55
2.1 Travaux fondateurs.....	55
2.2 Efforts de définition d'usage du système	56
2.3 Théorie de l'efficacité d'usage (TEU)	57
2.3.1 Théorie générale du SI.....	57
2.3.1.1 Deux vues de représentation du SI.....	58
2.3.1.2 Trois structures du SI.....	59
2.3.2 Visions de développement de TEU.....	60
2.3.3 Efficacité d'usage d'un SI : Développement de théorie à 3 étapes.....	61
2.3.3.1 Cadre théorique.....	61
2.3.3.2 Application du cadre théorique	64
2.3.3.3 Modèles proposés à tester	64
2.3.4 Développement de la discussion sur la théorie d'efficacité d'usage	66
2.3.5 La place de la théorie d'efficacité d'usage dans les théories existantes	70
2.3.5.1 TEU et TAM.....	70
2.3.5.2 TEU et UTAUT	71

2.3.5.3	TEU et le modèle de DeLone et McLean.....	72
2.3.5.4	TEU et TTF.....	73
2.3.5.5	TEU et AST	74
2.3.5.6	TEU et ISURA.....	75
2.3.6	Contribution de Burton-Jones et Grange (Burton-Jones & Grange, 2013)	76
Chapitre 3 Efficacité d’usage du big data : Modèle conceptuel mobilisé.....		79
3.1	Thème 1 : Intérêt organisationnel perçu.....	81
3.1.1	Valeur perçue	81
3.1.2	Risques perçus	82
3.1.3	Efficacité perçue	84
3.1.4	Facilité perçue de l’utilisation.....	85
3.1.5	Intention à utiliser	85
3.1.6	Observabilité perçue	86
3.1.7	Analyse coût-avantage	87
3.2	Thème 2: Management des processus.....	88
3.2.1	Orientation processus.....	88
3.2.2	Intégration du business processus en TI	89
3.2.3	Management des données	90
3.3	Thème 3: Confidentialité, sécurité et gouvernance des données	91
3.3.1	Confidentialité et sécurité	91
3.3.2	Gouvernance des données.....	92
3.4	Thème 4: Qualité des données	93
3.4.1	Qualité des données et qualité de l’information	93
3.4.2	Complétude des données.....	94
3.4.3	Actualité des données	94
3.4.4	Accès aux données.....	95
3.4.5	Pertinence des données	96
3.4.6	Exactitude des données.....	96
3.4.7	Cohérence des données	96
3.5	Thème 5 : Aspect Humain.....	97
3.5.1	Savoirs et compétences des individus.....	97

3.5.2	Confiance	99
3.5.3	Champions	99
3.5.4	Engagement du personnel	100
3.5.5	Participation des utilisateurs	100
3.5.6	Caractéristiques individuelles	101
3.6	Thème 6 : Aspect organisationnel	101
3.6.1	Compétence organisationnelle et culturelle	102
3.6.2	Gestion du talent	103
3.6.3	Programme de gestion du changement	103
3.6.4	Alignement stratégique	104
3.6.5	Management du projet	105
3.6.6	Management de la performance	105
3.6.7	Structure organisationnelle et taille	106
3.6.8	Collaboration interdépartementale	106
3.6.9	Communication	107
3.6.10	Soutien de la haute direction	108
3.6.11	Effet environnemental	108
3.6.12	Objectifs précis	109
3.6.13	Accent sur l'innovation	109
3.7	Thème 7 : Systèmes, outils et technologies	110
3.7.1	Qualité du système	110
3.7.2	Infrastructure TI	112
3.7.3	Support Fournisseurs	113
Seconde partie : Les enseignements de l'efficacité d'usage du big data dans les entreprises arméniennes		115
Chapitre 4 Positionnement épistémologique et méthodologie de la recherche		117
4.1	Positionnement épistémologique et paradigme	117
4.1.1	Positionnements épistémologiques et trois principaux paradigmes en SI	117
4.1.2	Paradigme de notre travail de recherche	119
4.1.2.1	Justification du choix du positionnement épistémologique et mode de raisonnement adopté à partir de quatre axes de Mbengue et Vandangeon-Derumez (1999)	121

4.1.2.2	Statut accordé aux données	122
4.1.2.3	Mode de collecte des données.....	124
4.1.2.4	Relation entre la théorie et les observations empiriques	125
4.1.2.5	Critères de scientificité de la recherche	126
4.2	Méthodologie de la recherche	127
4.2.1	Présentation de la question de recherche à partir de six paramètres de Noël (2012) 128	
4.2.1.1	Élément 1 : Clarté	128
4.2.1.2	Élément 2 : Faisabilité.....	128
4.2.1.3	Élément 3 : Pertinence	129
4.2.1.4	Élément quatre : Interrogation	130
4.2.1.5	Élément 5 : Auditoire pour les résultats.....	130
4.2.1.6	Élément 6 : Intelligence	130
4.2.2	Deux volets de la méthodologie de recherche	131
4.2.2.1	Première étude : Étude qualitative exploratoire	133
4.2.2.2	Deuxième étude : Étude quantitative principale	136
4.2.2.2.1	Première étape : Opérationnalisation des variables composantes du modèle conceptuel 136	
4.2.2.2.2	Deuxième étape : Construction de l'instrument d'enquête	138
4.2.2.2.3	Étape 3 : Validation technique globale de l'instrument	144
4.2.2.2.4	Étape 4 : Adaptation au contexte et étude-pilote.....	145
4.2.2.2.5	Étape 5 : Construction du questionnaire final	145
4.2.2.2.6	Étape 6 : Choix d'échantillon, distribution du questionnaire et recueil des données 145	
4.2.2.2.7	Étape 7 : Administration, préparation des données et leur traitement.....	147
4.2.2.2.7.1	Administration et préparation des données	147
4.2.2.2.7.2	Traitement des données.....	147
4.2.2.2.7.2.1	Première étape de traitement	148
4.2.2.2.7.2.2	Deuxième étape de traitement.....	148
4.2.2.2.7.2.3	Troisième étape de traitement	149
4.2.2.3	Troisième étude : Étude qualitative confirmatoire	149
Chapitre 5 Terrain de recherche.....		153
5.1	Pourquoi faut-il contextualiser une recherche en science de gestion ?	153

5.1.1.	Contextualisation théorique	154
5.1.2.	Contextualisation empirique	155
5.2.	Profil de l'Arménie : Contextualisation à 3 pôles de Livian (2020)	156
5.2.1.	Arménie : une brève présentation	156
5.2.2.	Trois pôles de contextualisation de Livian (2020).....	157
5.2.2.1.	Le pôle « culturaliste ».....	157
5.2.2.1.1.	Culture arménienne	157
5.2.2.1.2.	Rôle de la famille	158
5.2.2.2.	Pôle « institutionnaliste »	159
5.2.2.2.1.	Structuration des marchés, profil économique	160
5.2.2.2.2.	Modes de gouvernance.....	162
5.2.2.2.3.	Systèmes juridiques et financiers	163
5.2.2.2.4.	Organisation des partenaires sociaux	164
5.2.2.2.5.	Système de formation.....	165
5.2.2.3.	Troisième pôle.....	166
5.2.2.3.1.	Annonce historique	166
5.2.2.3.2.	Profil du pays en tant que membre de l'US (Union Soviétique).....	167
5.2.2.3.3.	Effondrement de l'US (Union Soviétique)	168
5.2.3.	Environnement technologique	168
5.2.3.1.	Histoire du secteur	169
5.2.3.2.	Les chiffres clés et les réussites du secteur	170
5.2.3.3.	Perspectives et défis de développement du secteur.....	171
5.2.3.4.	Technologie et sécurité de l'Arménie	173
5.3.	Présentation détaillée des acteurs du terrain	174
5.3.1.	Première étude exploratoire : Étude d'un expert en utilisation du big data.....	175
5.3.2.	Étude quantitative : profils des participants.....	177
5.3.2.1.	PQR.....	177
5.3.2.2.	DEF	178
5.3.2.3.	EFG et FGH	179
5.3.2.4.	GHI	179
5.3.2.5.	HIJ	180
5.3.2.6.	KLM.....	180

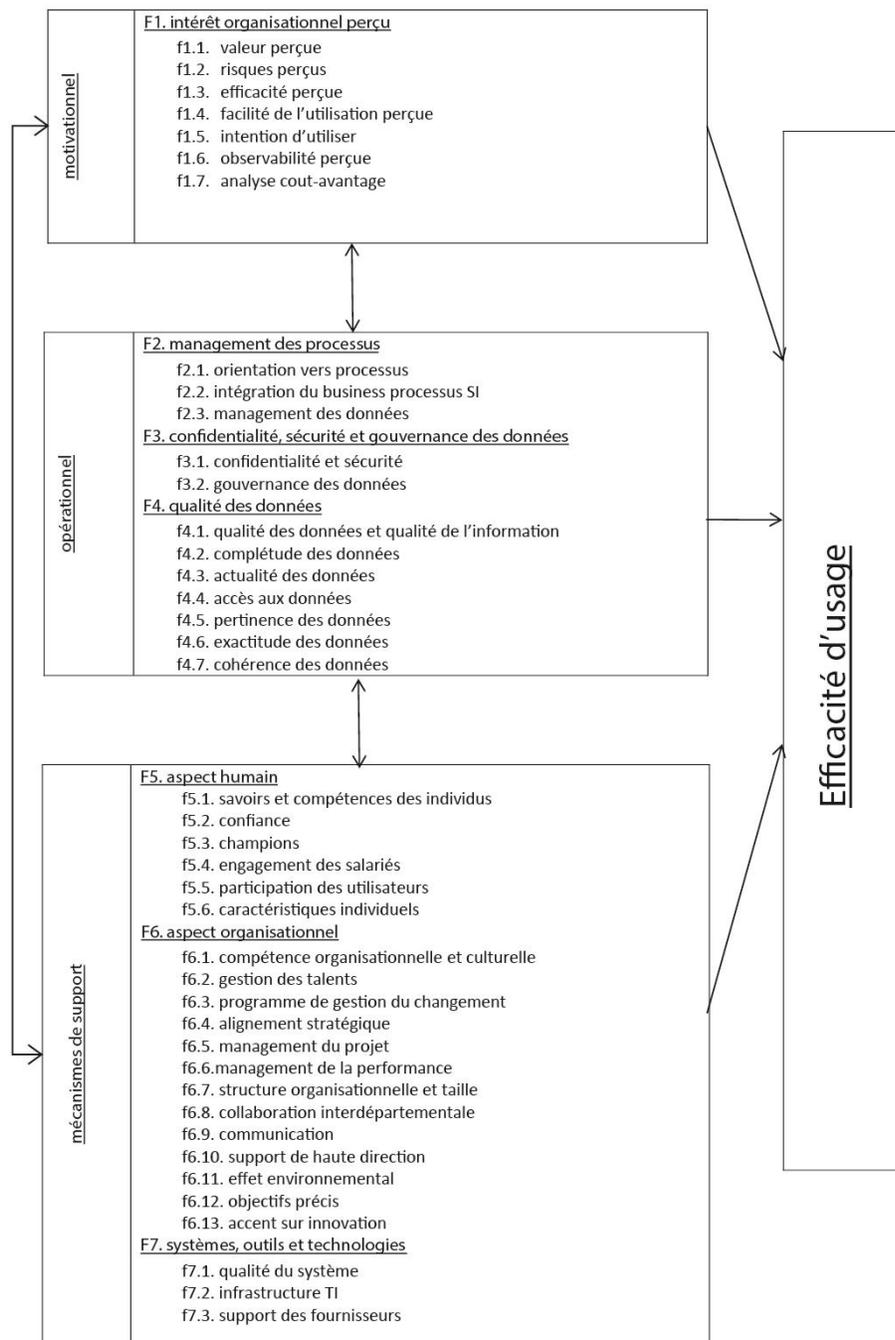
5.3.2.7.	LMN.....	181
5.3.2.8.	JKL.....	181
5.3.2.9.	MNO	182
5.3.2.10.	NOP.....	182
5.3.2.11.	OPQ.....	183
5.3.2.12.	UTS	183
5.3.2.13.	VUT	184
5.3.3.	Deux cas bien précis : « utilisateur novice » et « utilisateur avancé ».....	184
5.3.3.1.	BCD Arménie, « utilisateur novice » du big data	184
5.3.3.1.1.	Profil du répondant.....	185
5.3.3.1.2.	Cas détaillé.....	185
5.3.3.2.	CDE, « utilisateur avancé » du big data.....	187
5.3.3.2.1.	Profil du répondant.....	187
5.3.3.2.2.	Cas détaillé.....	187
Chapitre 6 Résultats		191
6.1	Résultat 1: « Utilisateur novice » et « utilisateur avancé »	192
6.2	Résultat 1 bis : Société ABC, « utilisateur avancé » du big data	193
6.2.1	F1 : Intérêt organisationnel perçu	193
6.2.2	F2 : Management des processus	196
6.2.3	F3 : Confidentialité, sécurité et gouvernance des données	196
6.2.4	F4 : Qualité des données	197
6.2.5	F5 : Aspect humain	198
6.2.6	F6 : Aspect organisationnel	199
6.2.7	F7 : Systèmes, outils et technologies	200
6.3	Résultat 2 : Deux modèles différents de l'efficacité d'usage du big data en entreprise en contexte arménien	201
6.3.1	Modèle d'efficacité d'usage du big data pour un « utilisateur novice ».....	201
6.3.2	Modèle d'efficacité d'usage du big data pour un « utilisateur avancé »	203
6.4	Résultat 2 bis : Efficacité d'usage du big data chez une entreprise « utilisateur avancé » et « utilisateur novice »	204
6.4.1	CDE, « utilisateur avancé »	204
6.4.2	BCD Arménie, « utilisateur novice ».....	208

6.5	Résultats 3 : L'existence de deux types d'utilisateurs ne dépend pas du contexte d'examen.....	210
6.6	Résultat 3 bis : Particularités contextuelles arméniennes.....	211
6.7	Conclusion des résultats.....	213
Chapitre 7 Conclusion et discussion.....		215
7.1	Réponse à la question de recherche.....	215
7.1.1	Le cas de l'entreprise « utilisateur novice ».....	216
7.1.2	Le cas de l'entreprise « utilisateur avancé ».....	218
7.2	Trois axes de discussion.....	220
7.2.1	Prolongement des résultats.....	220
7.2.1.1	Interprétation 1.....	221
7.2.1.2	Interprétation 2.....	222
7.2.2	Place du contexte arménien dans cette étude.....	223
7.2.3	Mise en perspective de notre recherche.....	226
7.2.3.1	Étude sur la proposition d'une définition du concept de l'efficacité d'usage du big data 229	
7.2.3.2	Étude sur les facilitateurs et les inhibiteurs de l'efficacité d'usage du big data.....	231
7.2.3.3	Étude sur la relation des éléments contextuels avec les facteurs influençant la réalisation de la valeur du big data en entreprise.....	234
7.2.3.4	Étude de l'adoption et l'usage de BDA (Big Data Analytics) par les PME en Iran....	237
7.3	Contributions théoriques et managériales de la thèse.....	238
7.3.1	Contributions théoriques.....	238
7.3.2	Contributions managériales.....	240
7.4	Limites et pistes de recherche.....	242
Table des illustrations.....		245
Bibliographie.....		247
Annexes.....		285
Annexe 1: Modèle conceptuel de l'efficacité d'usage du big data traduit de Surbakti et al. (2019)285		
Annexe 2: Description complète des construits et des éléments de l'instrument d'enquête... 286		
Annexe 3 : Instrument final de l'étude quantitative (questionnaire).....		328

Annexe 4: Détail “Cronbach's Alpha”	339
Annexe 5 : Caractéristiques de l'échantillon	340
Annexe 6 : Administration du questionnaire, préparation des données (capture d'écran)	341
Annexe 7 : Préparation des données ; une ligne pour chaque cas unique et une colonne pour chaque variable (capture d'écran)	342
Annexe 8 : Le rapport de l'écart-type à la moyenne en pourcentage pour 41 facteurs	343
Annexe 9 : Test d'analyse de variance pour les 2 groupes de répondants « utilisateurs novices » et « utilisateurs avancés ».....	345
Annexe 10 : Données brutes par facteur (capture d'écran).....	346
Annexe 11 : Matrices de corrélation (Pearson) par groupe d'utilisateurs (captures d'écrans)	347
Annexe 12 : Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur novice » (capture d'écran).....	352
Annexe 13 : Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur avancé» (capture d'écran)	353
Annexe 14 : Codage des facteurs par questions (capture d'écran)	354
Annexe 15 : Le reste des facteurs qui contribuent à l'efficacité d'usage du big data chez un « utilisateur novice » (capture d'écran).....	355
Annexe 16 : Le reste des facteurs qui contribuent à l'efficacité d'usage du big data chez un « utilisateur avancé » (capture d'écran)	356
Annexe 17 : Colonne comparatives de 41 facteurs pour les 2 groupes « utilisateurs avancés » et « utilisateurs novices »	357
Annexe 18: Guide d'entretien	358
Annexe 19 : Guide d'entretien adapté de (Mikalef et al., 2019a).....	359
Annexe 20 : Régulation arménienne de la sécurité et l'utilisation des données personnelles	363
Annexe 21 : Article présenté lors de 25 ^{ème} conférence de l'AIM 2020	384
Annexe 22: Facteurs marqués en rouge sont repérés lors de l'analyse de l'étude exploratoire	400
Annexe 23: Cas « utilisateur novice ».....	401

Annexes

Annexe 1: Modèle conceptuel de l'efficacité d'usage du big data traduit de Surbakti et al. (2019)



Annexe 2: Description complète des construits et des éléments de l'instrument d'enquête

Thème : Systèmes, Outils et Technologies

1 : Qualité du système 1. Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*. LOG, Rang 2 (FNEGE_2019)

/ constructs were measured using the 7-point Likert scale (e.g. strongly disagree – strongly agree)/

BD System quality is defined as:

systems reliability

(The system operates reliably for the analytics.

The system performs reliably for the analytics.

The operation of the system is dependable for the analytics.)

system adaptability

(The system can be adapted to meet a variety of analytics needs.

The system can flexibly adjust to new demands or conditions during analytics.

The system is flexible in addressing needs as they arise during the analytics.)

system integration

(The system effectively integrates data from different areas of the company.

The system pulls together data that used to come from different places in the company.

The system effectively combines different types of data from all areas of the company.)

system accessibility

(The system allows information to be readily accessible to me.

The system makes information very accessible.

The system makes information easy to access.)

system response time

(The system does not take long time to process my requests.

The system provides information in a timely fashion.

The system processes my requests quickly.)

system privacy

(The system protects information about personal issues.

This system protects information about personal identity.

The system offers a meaningful guarantee that it will not share private

information.)

2. Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

3. Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL a multiple-item scale for assessing electronic service quality. *Journal of Service Research*. MKG, Rang 1 (FNEGE_2019)

2 : Infrastructure TI

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

(The DW technology that the project team used worked well with technology already in place in the organization.

Appropriate technology was available to implement DW.)

2. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/ items were measured using a 7-point Likert scale/

(We use the latest technology to manage big data

Our technological competency helps us to enhance big data management

We use a variety of technological tools to manage big data

Our big data technological tools are more effective than those used by others in the industry)

3 : Support Fournisseurs

1. Maroufkhani, P., Tseng, M. L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management. MIS, Rang 2 (FNEGE_2019)*

/ a five-point Likert scale with anchors ranging from 1 (strongly disagree) to 5 (strongly agree) utilized/

External Support

(Community agencies/vendors can provide required training for Big Data Analytics adoption

Community agencies/vendors can provide effective technical support for Big Data Analytics adoption

Vendors actively market Big Data Analytics adoption)

2. Robey, D. (1979). User Attitudes and Management Information System Use. *Academy of Management Journal. GEN MAN, Rang 1 (FNEGE_2019)*

/ All items are measured on a 5-point scale to indicate whether the respondent strongly agrees, agrees, is uncertain, disagrees, or

strongly disagrees with the statement./

(The developers of the system have provided adequate training to users.)

Thème : Qualité des données

1 : Qualité des données et qualité de l'information

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/ items were measured using a 7-point Likert scale/

(Our big data sources provide accurate data

Our big data providers have a very good reputation in the industry

We are satisfied with the quality of the data provided by our big data sources)

2. Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of Management Information Systems. MIS, Rang 1 (FNEGE_2019)*

/ constructs were measured using the 7-point Likert scale (e.g. strongly disagree – strongly agree)/

(Overall, I would give the information from ____ high marks.

Overall, I would give the information provided by ____ a high rating in terms of quality.

In general, ____ provides me with high-quality information.)

2 : Complétude des données

1. Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research. LOG, Rang 2 (FNEGE_2019)*

2. Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of*

Management Information Systems. MIS, Rang 1 (FNEGE_2019)

/ constructs were measured using the 7-point Likert scale (e.g. strongly disagree – strongly agree)/

(The business analytics used:

_____ provides a complete set of information.

_____ produces comprehensive information.

_____ provides all the information needed.)

3. Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. Information and Management. MIS, Rang 1 (FNEGE_2019)

/ All items are measured on a 0 to 10 scale where 0 is not at all and 10 is completely; the mid-point of the scale, 5, labeled as “average.”/

(This information includes all necessary values.

This information is incomplete.

This information is complete.

This information is sufficiently complete for our needs.

This information covers the needs of our tasks.

This information has sufficient breadth and depth for our task.)

3 : Actualité des données

1. Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. International Journal of Production Research. LOG, Rang 2 (FNEGE_2019)

2. Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. Journal of Management Information Systems. MIS, Rang 1 (FNEGE_2019)

/ constructs were measured using the 7-point Likert scale (e.g. strongly disagree – strongly agree)/

(The business analytics used:

_____ *provides the most recent information.*

_____ *produces the most current information.*

_____ *always provides up-to-date information.)*

4 : Accès aux données

1. Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/7-point Likert scales used/

(Text-based data would be easily retrievable.

Text-based data would be easily accessible.

Text-based data would be easily obtainable.

Text-based data would be quickly accessible when needed.)

2. Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/ All items are measured on a 0 to 10 scale where 0 is not at all and 10 is completely; the mid-point of the scale, 5, labeled as “average.”/

(This information is easily retrievable.

This information is easily accessible.

This information is easily obtainable.

This information is quickly accessible when needed.)

5 : Pertinence des données

1. Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/ All items are measured on a 0 to 10 scale where 0 is not at all and 10 is completely; the mid-point of the scale, 5, labeled as “average.”/

(This information is useful to our work.

This information is relevant to our work.

This information is appropriate for our work.

This information is applicable to our work.)

2. Zheng, Y., Zhao, K., & Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision Support Systems*. MIS, Rang 1 (FNEGE_2019)

(5-point Likert scale: 1 — strongly disagree and 5 — strongly agree)

/ Information is informative for my needs.

Information is valuable for my needs. /

6 : Exactitude des données

1. Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*. LOG, Rang 2 (FNEGE_2019)

2. Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: An empirical examination within the context of data warehousing. *Journal of Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/ constructs were measured using the 7-point Likert scale (e.g. strongly disagree – strongly agree)/

(The business analytics used:

_____ produces correct information.

_____ provides few errors in the information.

_____ provides accurate information.)

7 : Cohérence des données

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success.

MIS Quarterly: Management Information Systems. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

(DW has improved the consistency of data to users (or applications) over that of source systems.)

2. Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management. MIS, Rang 2 (FNEGE_2019)

/A seven-point Likertscale ranging from "strongly disagree" (1) to "strongly agree" (7)was used to measure responses/

(1: An automatic method of maintaining data consistency is being used.

2: A common definition of the main data source (order forms, componentinformation, etc.) is being used.

3: All data (ordering details, material inventory, etc.) are managed the sameway throughout the organization.

4: There is no input error in all data.

Thème : Aspect Humain

1 : Savoirs et compétences des individus

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. MIS Quarterly: Management Information Systems. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

(Members of the DW team (including consultants) had the right technical skills for DW.

Members of the DW team had good interpersonal skills.)

2. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-

method approach. *Journal of Business Research*, 98, 261–276.
MKG, Rang 2 (FNEGE_2019)

/ a 7-point Likert scale utilized/

(TS1. Our 'big data analytics' staff has the right skills to accomplish their jobs successfully

TS2. Our 'big data analytics' staff is well trained

TS3. We provide big data analytics training to our own employees

TS4. Our 'big data analytics' staff has suitable education to fulfil their jobs)

2 : Confiance

1. Amoako-Gyampah, K., & Salam, A. F. (2004). An extension of the technology acceptance model in an ERP implementation environment. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/ Likert-type scale ranging from 1 to 7 with 1 being strongly agree and 7 being strongly disagree/

/beliefs in the benefits of project/

(I believe in the benefits of the NEW system

My peers believe in the benefits of the new system

My management team believes in the project benefits)

2. Lankton, N. K., Harrison Mcknight, D., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems. MIS*, Rang 2 (FNEGE_2019)

/7-point Likert scale from (1) not true at all to (7) very true/

Trusting intention

(1. When I [do a class assignment/ network socially online], I feel I can depend on [Microsoft Access/MySNW.com].

2. I can always rely on [Microsoft Access/MySNW.com] for [a tough class assignment/ online social networking].

3. I feel I can count on [Microsoft Access/MySNW.com] when

[doing my assignments/ networking online].)

3 : Champions

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

(A high-level champion(s) for DW came from IS.

A high-level champion(s) for DW came from a functional area(s).)

2. Premkumar, G., & Ramamurthy, K. (1995). The Role of Interorganizational and Organizational Factors on the Decision Mode for Adoption of Interorganizational Systems. *Decision Sciences*. MIS/LOG, Rang 2 (FNEGE_2019)

/ 7-point Likert type scale ranging from strongly disagree to strongly agree/

(Champion

1. Existence of a single person committed to introducing EDI in the organization)

** electronic data interchange (EDI)*

4 : Engagement des salariés

1. Rondeau, P. J., Ragu-Nathan, T. S., & Vonderembse, M. A. (2006). How involvement, IS management effectiveness, and end-user computing impact IS performance in manufacturing firms. *Information and Management*. MIS, Rang 1 (FNEGE_2019)

/ Please circle the appropriate number which best indicates your existing level of end-user involvement in software application development.

1 = None, 2 = Low, 3 = Moderate, 4 = High, 5 = Very

High, NA = Not Applicable, or Do Not Know/

(EI1 Design of manufacturing software applications.

EI2 Development of manufacturing software applications.

EI3 Analysis of manufacturing software application problems and opportunities.

EI4 Testing of manufacturing software applications.

EI5 Specification of manufacturing software application requirements.

EI6 Management of manufacturing software application development projects.

EI7 Implementation of manufacturing software applications.)

2. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

* widespread support from personnel

(The DW had support from people throughout the organization.)

5 : Participation des
utilisateurs

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

(IS and users worked together as a team on the DW project.

Users were assigned full-time to parts of the DW project.

Users performed hands-on activities (e.g., data modeling) during the DW project.)

2. Yoon, Y., Guimaraes, T., & O'Neal, Q. (1995). Exploring the factors associated with expert systems success. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/ Measuring scale for all variables: 1 = completely disagree, 2 = moderately disagree, 3 = somewhat disagree, 4= neither

agree/disagree, 5 = somewhat agree, 6= moderately agree, 7= completely agree./

(Initiating the project

Establishing project objective

Determining user requirements

Determining ways to meet requirements

Identifying sources of data/information

Outlining information flow

Developing input forms/screens

Developing output forms/screens

Determining systems availability/access)

6 : Caractéristiques
individuels

1. Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/otes: All constructs except Self Efficacy are seven-point scales with the anchors 1 = Strongly disagree, 4 = Neutral, 7=Strongly agree).

Personal Innovativeness

(PIIT1. If I heard about a new information technology, I would look for ways to experiment with it.

PIIT2. In general, I am hesitant to try out new information technologies.

PIIT3. Among my peers, I am usually the first to try out new information technologies.

PIIT4. I like to experiment with new information technologies.)

2. Yoon, Y., Guimaraes, T., & O'Neal, Q. (1995). Exploring the factors associated with expert systems success. *MIS Quarterly:*

Management Information Systems. MIS, Rang 1 (FNEGE_2019)

/ Measuring scale for all variables: 1 = completely disagree, 2 = moderately disagree, 3 = somewhat disagree, 4= neither agree/disagree, 5 = somewhat agree, 6= moderately agree, 7= completely agree./

(Positive attitude on ES

Expectation

Computer/AI knowledge)

**ES-Expert Systems*

3. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. MIS, Rang 1 (FNEGE_2019)

/ Sevenpoint scales were used for all of the constructs' measurement, with 1 being the negative end of the scale and 7 being the positive end of the scale./

(Anxiety

ANX1: I feel apprehensive about using the system.

ANX2: It scares me to think that I could lose a lot of information using the system by hitting the wrong key.

ANX3: I hesitate to use the system for fear of making mistakes I cannot correct.

ANX4: The system is somewhat intimidating to me.)

Thème : Confidentialité, sécurité et gouvernance des données

1 : Confidentialité et sécurité

1. Ghasemaghaei, M. (2020). The role of positive and negative valence factors on the impact of bigness of data on big data analytics usage. *International Journal of Information Management*. MIS, Rang 2 (FNEGE_2019)

/ 7-point Likert scales ranging from “strongly disagree” to

“strongly agree”/

(In my organization, the data used by data analytics tools cannot be manipulated by inappropriate parties (reverse coded).

In my organization, the data used by data analytics tools is secured (reverse coded).

In my organization, inappropriate parties cannot deliberately view the data used by data analytics tools (reverse coded).

2. Zheng, Y., Zhao, K., & Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision Support Systems*. MIS, Rang 1 (FNEGE_2019)

/5-point Likert scale: 1 — strongly disagree and 5 — strongly agree/

(XXX protects information against unauthorized access.

XXX protects my personal information.

We feel safe to participate in XXX.)

2 : Gouvernance des données

1. Byrd, T. A., & Turner, D. E. (2000). Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. *Journal of Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/ a 5-point Likert scale utilized, from strongly disagree to strongly agree/

(Our organization utilizes online analytical processing OLAP

Our corporate database is able to communicate through many different protocols (e.g.,SQL, ODB)

Mobile users have ready access to the same data used at desktops

A common view of our organization's customer is available to everyone in the Organization

Our organization easily adapts to various vendors' data- base management systems protocols and standards

Data captured in one part of our organization are immediately available to everyone in the organization

Our IT organization handles variances in corporate data formats and standards

Data rules and relations (e.g., tax regulations) are hard coded into applications)/

2. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. MKG, Rang 2 (FNEGE_2019)

/ a 7-point Likert scale utilized/

Procedural practices

(In our organization, we have controlled practices regarding data management in terms of _____

PCR1. setting retention policies (e.g. time to live) of data

PCR2. backup routines

PCR3. establishing/monitoring access (e.g. user access) to data

PCR4. classifying data according to value

PCR5. monitoring costs versus value of data)

Thème : Intérêt organisationnel perçu

1 : Valeur perçue (stratégiques)

1. Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2016). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, 55(17), 5011–5026. LOG, Rang 2 (FNEGE_2019)

/constructs were measured using the 7-point Likert scale (e.g. strongly disagree – strongly agree/.

(Creating competitive advantage

Aligning analytics with business strategy

Establishing useful links with other organisations

Enabling quicker response to change

Improving customer relations

Providing better products or services to customers)

2. Gregor, S., Martin, M., Fernandez, W., Stern, S., & Vitale, M. (2006). The transformational dimension in the realization of business value from information technology. *The Journal of Strategic Information Systems*, 15(3), 249–270. MIS, Rang 2 (FNEGE_2019)

/Responses were on a scale of 1–10, where 1 was strongly disagree and 10 was strongly agree./

(The creation of competitive advantage

Aligning ICT strategy with business strategy

Establishing useful links with other organizations such as suppliers

Enabling the organization to respond more quickly to change

Improving customer relations

Providing better products or services to customers)

3. Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187–195. MIS, Rang 2 (FNEGE_2019)

/ measurement items were based on a seven-point Likert scale, with responses ranging from “completely disagree” (-3) to “completely

agree” (+3)/

(Creating a competitive advantage

Aligning IT with a business strategy

Establishing useful links with other organizations

Enabling a quicker response to change

Improving customer relations

Providing better products or services)

2 : Risques perçus

1. Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187–195. MIS, Rang 2 (FNEGE_2019)

/ measurement items were based on a seven-point Likert scale, with responses ranging from “completely disagree” (-3) to “completely agree” (+3)/

(Reluctance of employees to adapt to changes

Lack of Information System infrastructure support

Technical uncertainty

Minimal IT expertise

Uncertainty about how to measure potential benefits

Uncertainty about how to measure the involved costs

Capital outlay with no guarantee of likely returns

Security issues

Reassignment of personnel trained on big data analytics solutions

Privacy issues)

2. Love, P. E. D., Irani, Z., Standing, C., Lin, C., & Burn, J. M. (2005). The enigma of evaluation: benefits, costs and risks of IT in Australian small–medium-sized enterprises. *Information &*

Management, 42(7), 947–964. MIS, Rang 1 (FNEGE_2019)

(Reluctance of employees to adapt to change

Lack of IS infrastructure support

Technical uncertainty and lack of knowledge

Minimal IT expertise

Maintenance costs

Uncertainty about how to measure potential benefits

Uncertainty about how to measure the costs involved

Capital outlay with no guarantee of likely returns

Security issues

Training expenses on staff that leave the organization

Theft of software and hardware)

3 : Efficiency perçue

1. Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. Information and Management. MIS, Rang 1 (FNEGE_2019)

/7-point Likert scales used/

(Using text-mining software would improve my performance in my job.

Using text-mining software in my job would increase my productivity.

Using text-mining software would enhance my effectiveness in my job.

I would find text-mining software to be useful in my job.)

2. Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. Information Systems Research. MIS, Rang 1 (FNEGE_2019)

--

PWS-Personal Work Station

Notes a. The indicated items were all deleted as the result of the first factor analysis and hence were not in the final scales.

*—indicates items suggested for inclusion in any "short" scales.

/seven-point Likert scale ranging from "extremely disagree" to "extremely agree" chosen as the response format. /

- (1. Using a PWS enables me to accomplish tasks more quickly.
2. Using a PWS improves the quality of work I do.
3. Using a PWS makes it easier to do my job.
4. The disadvantages of my using a PWS far outweigh the advantages. (See Note a.)
5. Using a PWS improves my job performance.
6. Overall, I find using a PWS to be advantageous in my job.
- 7.* Using a PWS enhances my effectiveness on the job.
- 8.* Using a PWS gives me greater control over my work.
9. Using a PWS increases my productivity.)

4 : Facilité perçue de l'utilisation

1. Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/7-point Likert scales used/

(My interaction with text-mining software would be clear and understandable.

I would find text-mining software easy to use.

I would find it easy to get text-mining software to do what I want it to do.)

2. Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research. MIS*, Rang

1 (FNEGE_2019)

--

PWS-Personal Work Station

Notes a. The indicated items were all deleted as the result of the first factor analysis and hence were not in the final scales.

**—indicates items suggested for inclusion in any "short" scales.*

/seven-point Likert scale ranging from "extremely disagree" to "extremely agree" chosen as the response format. /

(1. I believe that a PWS is cumbersome to use.

2. It is easy for me to remember how to perform tasks using a PWS. (See Note a.)

3. My using a PWS requires a lot of mental effort.

4. Using a PWS is often frustrating.

5. My interaction with a PWS is clear and understandable. (See Note a.)*

6. I believe that it is easy to get a PWS to do what I want it to do.*

7. Overall, I believe that a PWS is easy to use.*

8. Learning to operate a PWS is easy for me.)*

5 : Intention d'utilisation

1. Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/7-point Likert scales used/

(Assuming I have access to text-mining software, I intend to use it in the next few months.

Assuming I have access to text-mining software, I predict that I will use it in the next few months.

I plan to use text-mining software in the next few months.)

2. Kwahk, K. Y., & Lee, J. N. (2008). The role of readiness for change in ERP implementation: Theoretical bases and empirical validation. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/ All question items were measured using a seven-point Likert-type scale with anchors ranging from “strongly disagree” to “strongly agree.”/

(I intend to use the ERP system for performing my job as often as needed

To the extent possible, I would frequently use the ERP system in my job)

3. Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology. *Journal MIS Quarterly. MIS*, Rang 1 (FNEGE_2019)

/ items were represented as a

seven-point Likert type scale (strongly disagree = 1 to strongly agree = 7)/

(I want to use the mobile office in the future

I expect that I will use the mobile office later

I plan to use the mobile office in the future)

6 : Observabilité perçue

1. Maroufkhani, P., Tseng, M. L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management. MIS*, Rang 2 (FNEGE_2019)

/ a five-point Likert scale with anchors ranging from 1 (strongly disagree) to 5 (strongly agree) utilized/

(Many competitors or business partners in the market have started using Big Data Analytics

Using Big Data Analytics helps my company to connect with both domestic and international business partners

There are many computers that people in the company can access to Big Data Analytics

There are many computers that people in the company can access to use Big Data Analytics

Big Data Analytics shows improved results over doing business the traditional way)

2. Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*. MIS, Rang 1 (FNEGE_2019)

--

PWS-Personal Work Station

Notes a. The indicated items were all deleted as the result of the first factor analysis and hence were not in the final scales.

**—indicates items suggested for inclusion in any "short" scales.*

/seven-point Likert scale ranging from "extremely disagree" to "extremely agree" chosen as the response format. /

(1. I would have no difficulty telling others about the results of using a PWS. 2* I believe I could communicate to others the consequences of using a PWS.*

3. The results of using a PWS are apparent to me.*

4. I would have difficulty explaining why using a PWS may or may not be beneficial.)*

3. Huang, T. C. K., Liu, C. C., & Chang, D. C. (2012). An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management*. MIS, Rang 2 (FNEGE_2019)

/ variables were measured on a seven-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (7)/

(1. I have no difficulty telling others about the results of using the data mining tool.

2. I believe I could communicate to others the consequences of using the data

mining tool.

3. The results of using the data mining tool are apparent to me.

4. I would have difficulty explaining why using the data mining tool may or

may not be beneficial.)

7 : analyse cout-avantage

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*, MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

** Net Benefits*

(DW has changed my job significantly.

DW has reduced the time it takes to support decision making to the end-user community.

DW has reduced the effort it takes to support decision making to the end-user community.)

2. Maduku, D. K., Mpinganjira, M., & Duh, H. (2016). Understanding mobile marketing adoption intention by South African SMEs: A multi-perspective framework. *International Journal of Information Management*, MIS, Rang 2 (FNEGE_2019)

/ seven-point Likert-type scale with anchors ranging from 1 ('strongly disagree') to 7 ('strongly agree')/

(The costs involved in the adoption of mobile marketing would be

far greater than the expected benefits

The cost of maintaining mobile marketing would be very high for our enterprises

The cost involved in providing support systems for mobile marketing would be too high

The amount of money invested in training employees to use mobile marketing would be very high)

3. Polites, & Karahanna. (2012). Shackled to the Status Quo: MIS, Rang 1 (FNEGE_2019) The Inhibiting Effects of Incumbent System Habit, Switching Costs, and Inertia on New System Acceptance. MIS Quarterly. MIS, Rang 1 (FNEGE_2019)

/ All constructs were measured on a 1–7 scale/

Relative Advantage

(Using Google Docs to collaborate / share files with my teammates, rather than our current method of collaborating / sharing files, would enhance my group's effectiveness.

Using Google Docs to collaborate / share files with my teammates, rather than our current method of collaborating / sharing files, would increase my group's productivity.

Using Google Docs to collaborate / share files with my teammates, rather than our current method of collaborating / sharing files, would improve my group's performance.)

Thème : Management des processus

1 : Orientation vers processus

1. Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. fan, Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research. MKG, Rang 2 (FNEGE_2019)

/ constructs were measured using 7-point Likert scales (strongly disagree – strongly agree)/

Process-oriented dynamic capabilities

(Our company is better than competitors in connecting (e.g., communication and information sharing) parties within a business process.

Our company is better than competitors in reducing cost within a business process.

Our company is better than competitors in bringing complex analytical methods to bear on a business process.

Our company is better than competitors in bringing detailed information into a business process.)

2. Hong, K. K., & Kim, Y. G. (2002). The critical success factors for ERP implementation: An organizational fit perspective. Information and Management. MIS, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

(The processes built in ERP meet all needs required from organizational processes.

The processes flow built in ERP correspond to flow of organizational processes.

The processes built in ERP accommodate the change required from organizational processes

The processes built in ERP correspond to business practices of our firm)

2 : Intégration du business processus SI

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. Information and Management. MIS, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

(We have the ability to integrate the processes involved in the big data chain (i.e. data collection, preparation, analysis and decision making)

The integration of the processes involved in the big data chain

reduces the cost of big data use

The integration of the processes involved in the big data chain reduces the efforts necessary to analyse big data)

2. Bharadwaj, A., Sambamurthy, V., & Zmud, R. W. (1999). IT capabilities: theoretical perspectives and empirical operationalization. *Management Science*. GEN MAN/LOG, Rang 1 (FNEGE_2019)

/ Respondents were asked to rate their firm's performance on each item, relative to other firms in their industry, using a five-point Likert scale ranging from "exceptionally well" to "poorer than most."

(Consistency of IT application portfolios with business processes

Restructuring of Business work processes to leverage opportunities

Restructuring of IT work processes to leverage opportunities)

3 : Management des données

1. Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, 53(8), 1049–1064. *MIS*, Rang 1 (FNEGE_2019)

/ We have access to very large, unstructured, or fast-moving data for analysis

We integrate data from multiple internal sources into a data warehouse or mart for easy access

We integrate external data with internal to facilitate high-value analysis of our business environment/

2. Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*. *MIS*, Rang 1 (FNEGE_2019)

/ 7-point Likert scales ranging from "strongly disagree" to "strongly agree":/

(• In my organization there is a good fit between the analytical tools we have access to and the data we process.

• The present analytical tools my organization has access to fulfill

our data analysis needs.

• *The analytical tools that my organization currently has access to provide pretty much everything that we need to analyze our data properly.)*

Thème : Aspect organisationnel

1 : compétence
organisationnelle et
culturelle

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Organizational culture of big data

(Our decisions are based on data

A dependency on hunches for decision making is strongly discouraged in our organization

Depending on data is part of our organizational routine

We have a culture of data driven work)

2. Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

(We consider data a tangible asset

We base our decisions on data rather than on instinct

We are willing to override our own intuition when data contradict our viewpoints

We continuously assess and improve the business rules in response to insights extracted from data

We continuously coach our employees to make decisions based on data)

2 : gestion du talent

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Talent management for big data

(We prefer to hire employees who understand big data

We have the ability to recruit expert users of big data

We plan to enhance the big data management skills of our staf

We take special care in the retention of big data experts in our organization)

2. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics. LOG 1* (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Big data analytics talent capability/technical level

(Our analytics personnel are very capable in terms of programming

Skills

Our analytics personnel are very capable in terms of managing project life cycles.

Our analytics personnel are very capable in the areas of data and network management and maintenance.

Our analytics personnel create very capable decision support systems

driven by analytics.

3 : Programme de gestion

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making

du changement

capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Leadership focus on big data

(Our leadership provides a clear vision

Our leadership sets clear goals

Our leadership encourages big data decision making

Our leadership shows great interest in the big data chain

Our leadership shows concern for the use of big data

Our leadership is very active in managing big data)

2. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems. MIS*, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

* change management

(Change in the organization created by DW was managed effectively.)

4 : Alignement stratégique

1. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics. LOG* 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Analytics Capability–Business Strategy Alignment

(The big data analytics plan aligns with the company's mission, goals,

objectives, and strategies.

The big data analytics plan contains quantified goals and objectives.

The big data analytics plan contains detailed action plans/strategies that

support company direction.

We prioritize major big data analytics investments by the expected impact on business performance.)

2. Setia, P., & Patel, P. C. (2013). How information systems help create OM capabilities: Consequents and antecedents of operational absorptive capacity. *Journal of Operations Management*. LOG 1 (FNEGE_2019)

[1= strongly disagree to 7= strongly agree]

(The IT plan aligns with the company's mission, goals, objectives, and

strategies.

The IT plan contains quantified goals and objectives

The IT plan contains detailed action plans/strategies that support company

Direction

We prioritize major IT investments by the expected impact on business

performance.

)

5 : Management du projet

1. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*. LOG 1 (FNEGE_2019)

/ items were measured using a 7-point Likert /

Big data analytics (project) management capabilities

(We continuously examine the innovative opportunities for the strategic use of big data analytics.

We enforce adequate plans for the introduction and utilization of big data analytics.

We perform big data analytics planning processes in systematic and formalized ways.

We frequently adjust big data analytics plans to better adapt to changing conditions.)

2. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*, MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

**Project Implementation Success*

(The DW project met its critical project deadlines (eg., rollout deadline, initial development deadline).

The cost of the DW did not exceed its budgeted amount.

The DW project provided all of the DW functionality that it was supposed to provide.)

6 : management de la performance

1. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. MKG, Rang 2 (FNEGE_2019)

/ a 7-point Likert scale utilized/

(We perform much better than our main competitors in terms of:

PER1. Profitability

PER2. Profits as percentage of sales

PER3. Decreasing product or service delivery cycle time

PER4. In reducing operating costs

PER5. In profit growth rates

PER6. Rapid response to market demand

PER7. Rapid confirmation of customer orders

PER8. Increasing customer satisfaction

PER9. Providing better product and service quality

PER10. In reducing operating costs)

2. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*. LOG 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Using big data analytics improved_____during the last 3 years relative to competitors:

FPER1 _____Customer retention

FPER2 _____Sales growth

FPER3 _____Profitability

FPER4 _____Return on investment)

7 : Structure
organisationnelle et taille

1. Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. *Production and Operations Management*. LOG, Rang 1 (FNEGE_2019)

(Scale - “Strongly disagree” (1) to “Strongly agree” (5))

Organizational flexibility

(We can quickly change our organizational structure to respond to changing business conditions.

We can cost effectively change our organizational structure to

respond to changing business conditions.

We can change our organizational structure without negatively impacting service quality.

Our current organization structure enables us to adapt to changing business conditions.

Our organization is more flexible than our competitors in changing our organizational structure)

2. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. MKG, Rang 2 (FNEGE_2019)

/ a 7-point Likert scale utilized/

(In our organization, we _____

a/ have identified key IT and non-IT decision makers to have the responsibility regarding data ownership, value analysis and cost management.

b/ use steering committees to oversee and assess data values and costs)

8 : Collaboration
interdépartementale

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS*, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

(There is collaboration among big data analysts and big data providers

There is collaboration among big data analysts and decision makers

There is collaboration among big data providers and decision makers)

2. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*. LOG 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

Big data analytics Coordination

(In our organization, business analysts and line people meet frequently to

discuss important issues both formally and informally.

In our organization, business analysts and line people from various

departments frequently attend cross-functional meetings.

In our organization, business analysts and line people coordinate their efforts harmoniously.

In our organization, information is widely shared between business analysts and line people so that those who make decisions or perform jobs have access to all available know-how.

9 : Communication

1. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management*. MIS, Rang 1 (FNEGE_2019)

/ items were measured using a 7-point Likert scale/

(Our employees transfer their knowledge about data

Knowledge about how data are collected is exchanged within the organization

Knowledge about how data are used is exchanged within our firm

Knowledge about how data are processed is exchanged within our firm

The exchange of knowledge makes it easy for us to analyse data)

2. Kim, G., Shin, B., & Kwon, O. (2012). Investigating the value of sociomaterialism in conceptualizing its capability of a firm. *Journal of Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

IT coordination

(In our organization, IS and line people meet frequently to discuss important issues both formally and informally.

In our organization, IS people and line people from various departments frequently attend cross-functional meetings.

In our organization, IS and line people coordinate their efforts harmoniously.

In our organization, information is widely shared between IS and line people so that those who make decisions or perform jobs have access to all available know-how.)

10 : Support de haute direction

1. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*. MIS, Rang 1 (FNEGE_2019)

/Items were measured based on a sevenpoint Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree."/

(Overall, management has encouraged the use of DW.

User satisfaction has been a major concern of management.)

2. Maroufkhani, P., Tseng, M. L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management*. MIS, Rang 2 (FNEGE_2019)

/ a five-point Likert scale with anchors ranging from 1 (strongly disagree) to 5 (strongly agree) utilized/

(Our top management promotes the use of Big Data Analytics in

the organization

Our top management creates support for Big Data Analytics initiatives within the organization

Our top management promotes Big Data Analytics as a strategic priority within the organization

Our top Management is interested in the news about using Big Data Analytics adoption)

3. Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. Information and Management. MIS, Rang 1 (FNEGE_2019)

/7-point Likert scales used/

(The top management is greatly interested in using text-mining software.

The top management is aware of the benefits of text-mining software for future success.

The top management will allocate adequate financial and other resources to the development and operation of text-mining software.

The top management has a vision to project our company as the market leader in the use of text-mining software.)

11 : Effet environmental

1. Zhu, K., & Kraemer, K. L. (2005). Post-adoption variations in usage and value of e-business by organizations: Cross-country evidence from the retail industry. *Information Systems Research*. MIS, Rang 1 (FNEGE_2019)

/ 1~5,5-point Likert scale/

(Degree affected by competitors in the local market

Degree affected by competitors in the national market)

2. Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70, 379–390. MKG, Rang 2 (FNEGE_2019)

/ Items were measured using a 7-point numerical scale (1 is strongly disagree and 7 is strongly agree/

(Please indicate the degree to which you agree with the following statements.

Technological turbulence

Extent of technological turbulence in the environment

Leadership in product/process innovation

Impact of new technology on operations

Market turbulence

Extent of market turbulence in the market

Frequent changes in customer preferences

Ability to reduce market uncertainty

Ability to respond to market opportunities)

12 : Objectifs précis

1. Robey, D. (1979). User Attitudes and Management Information System Use. *Academy of Management Journal*. GEN MAN, Rang 1 (FNEGE_2019)

/ All items are measured on a 5-point scale to indicate whether the respondent strongly agrees, agrees, is uncertain, disagrees, or strongly disagrees with the statement./

(Company goals have become more clear.)

2. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management. MIS, Rang 1 (FNEGE_2019)*

/ items were measured using a 7-point Likert scale/

(Our leadership sets clear goals)

13 : Accent sur
l'innovation

1. Kim, G., Shin, B., Kim, K. K., & Lee, H. G. (2011). IT capabilities, process-oriented dynamic capabilities, and firm financial performance. *Journal of the Association for Information Systems. MIS 2 (FNEGE_2019)*

/ 7-point Likert scale/

IT planning

(We continuously examine the innovative opportunities for the strategic use of IT.

We enforce adequate plans for the introduction and utilization of IT.

We perform IT planning processes in systematic and formalized ways.

We frequently adjust IT plans to better adapt to changing conditions)

2. Ghasemaghaei, M., & Calic, G. (2019). Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research. MKG 2 (FNEGE_2019)*

/ 7-point Likert scale from “Not at all” to “To a very great extent”./

Innovation competency

Exploitation competency : Please indicate to what extent has your firm

Upgraded current knowledge and skill for familiar products and

technologies. • Invested in exploiting mature markets and technologies. • Enhanced competencies in searching for solutions to customer problems that are near to existing solutions rather than completely new solutions. • Upgraded skills in which your firm already possesses significant experience. • Strengthened knowledge and skills for projects that improve efficiency of existing activities.

Exploration competency:

Acquired new-to-the-firm technologies and skills. • Learned skills and processes entirely new to the industry. • Acquired entirely new managerial and organizational skills. • Invested in exploring new markets and technologies. • Strengthened innovation skills in new areas.

DW – Data Warehousing /One of the foundational technologies of Big Data Analytics.

References

1. Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/3250951>
2. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2016.08.018>
3. Amoako-Gyampah, K., & Salam, A. F. (2004). An extension of the technology acceptance model in an ERP implementation environment. *Information and Management*. <https://doi.org/10.1016/j.im.2003.08.010>
4. Bharadwaj, A., Sambamurthy, V., & Zmud, R. W. (1999). IT capabilities: theoretical perspectives and empirical operationalization. *Management Science*. <https://doi.org/10.1145/352925.352962>
5. Byrd, T. A., & Turner, D. E. (2000). Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. *Journal of Management Information Systems*. <https://doi.org/10.1080/07421222.2000.11045632>

6. Côte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70, 379–390.
<https://doi.org/10.1016/j.jbusres.2016.08.011>
7. Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. *Information and Management*.
<https://doi.org/10.1016/j.im.2018.10.006>
8. Dubey, R., Gunasekaran, A., Childe, S. J., Fosso Wamba, S., Roubaud, D., & Foropon, C. (2019). Empirical investigation of data analytics capability and organizational flexibility as complements to supply chain resilience. *International Journal of Production Research*. <https://doi.org/10.1080/00207543.2019.1582820>
9. Ghasemaghaei, M. (2020). The role of positive and negative valence factors on the impact of bigness of data on big data analytics usage. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2018.12.011>
10. Ghasemaghaei, M., & Calic, G. (2019). Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2019.07.006>
11. Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*.
<https://doi.org/10.1016/j.dss.2017.06.004>
12. Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, 53(8), 1049–1064.
<https://doi.org/10.1016/j.im.2016.07.004>
13. Hong, K. K., & Kim, Y. G. (2002). The critical success factors for ERP implementation: An organizational fit perspective. *Information and Management*.
[https://doi.org/10.1016/S0378-7206\(01\)00134-3](https://doi.org/10.1016/S0378-7206(01)00134-3)
14. Huang, T. C. K., Liu, C. C., & Chang, D. C. (2012). An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2011.11.006>
15. Kim, G., Shin, B., Kim, K. K., & Lee, H. G. (2011). IT capabilities, process-oriented dynamic capabilities, and firm financial performance. *Journal of the Association for Information Systems*. <https://doi.org/10.17705/1jais.00270>
16. Kim, G., Shin, B., & Kwon, O. (2012). Investigating the value of sociomaterialism in conceptualizing its capability of a firm. *Journal of Management Information Systems*.
<https://doi.org/10.2753/MIS0742-1222290310>
17. Kwahk, K. Y., & Lee, J. N. (2008). The role of readiness for change in ERP implementation: Theoretical bases and empirical validation. *Information and Management*. <https://doi.org/10.1016/j.im.2008.07.002>
18. Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2014.02.002>

19. Lankton, N. K., Harrison Mcknight, D., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*. <https://doi.org/10.17705/1jais.00411>
20. Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management*. [https://doi.org/10.1016/S0378-7206\(02\)00043-5](https://doi.org/10.1016/S0378-7206(02)00043-5)
21. Lu, Y., & Ramamurthy, K. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS Quarterly: Management Information Systems*. <https://doi.org/10.2307/41409967>
22. Maduku, D. K., Mpinganjira, M., & Duh, H. (2016). Understanding mobile marketing adoption intention by South African SMEs: A multi-perspective framework. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2016.04.018>
23. Maroufkhani, P., Tseng, M. L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2020.102190>
24. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. <https://doi.org/10.1016/j.jbusres.2019.01.044>
25. Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*. <https://doi.org/10.1287/isre.2.3.192>
26. Polites, & Karahanna. (2012). Shackled to the Status Quo: The Inhibiting Effects of Incumbent System Habit, Switching Costs, and Inertia on New System Acceptance. *MIS Quarterly*. <https://doi.org/10.2307/41410404>
27. Premkumar, G., & Ramamurthy, K. (1995). The Role of Interorganizational and Organizational Factors on the Decision Mode for Adoption of Interorganizational Systems. *Decision Sciences*. <https://doi.org/10.1111/j.1540-5915.1995.tb01431.x>
28. Robey, D. (1979). User Attitudes and Management Information System Use. *Academy of Management Journal*. <https://doi.org/10.5465/255742>
29. Rondeau, P. J., Ragu-Nathan, T. S., & Vonderembse, M. A. (2006). How involvement, IS management effectiveness, and end-user computing impact IS performance in manufacturing firms. *Information and Management*. <https://doi.org/10.1016/j.im.2005.02.001>
30. Setia, P., & Patel, P. C. (2013). How information systems help create OM capabilities: Consequents and antecedents of operational absorptive capacity. *Journal of Operations Management*. <https://doi.org/10.1016/j.jom.2013.07.013>
31. Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A

dynamic capabilities view. *Information and Management*.
<https://doi.org/10.1016/j.im.2018.12.003>

32. Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. *Production and Operations Management*.
<https://doi.org/10.1111/poms.12746>
33. Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology. *Journal MIS Quarterly*.
34. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.
<https://doi.org/10.2307/30036540>
35. Vitari, C., & Raguseo, E. (2016). Digital data, dynamic capability and financial performance: an empirical investigation in the era of Big Data. *Systèmes d'information & Management*. <https://doi.org/10.3917/sim.163.0063>
36. Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. fan, Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2016.08.009>
37. Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*.
<https://doi.org/10.2307/3250957>
38. Yoon, Y., Guimaraes, T., & O'Neal, Q. (1995). Exploring the factors associated with expert systems success. *MIS Quarterly: Management Information Systems*.
<https://doi.org/10.1590/s0104-530x1997000100001>
39. Zheng, Y., Zhao, K., & Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision Support Systems*.
<https://doi.org/10.1016/j.dss.2012.11.008>
40. Zhu, K., & Kraemer, K. L. (2005). Post-adoption variations in usage and value of e-business by organizations: Cross-country evidence from the retail industry. *Information Systems Research*. <https://doi.org/10.1287/isre.1050.0045>

Annexe 3 : Instrument final de l'étude quantitative (questionnaire)

Hello,

I am Elen Grigoryan, PhD student in the field of big data management at University of Lyon, France.

I am very interested to request your participation in a professional survey.

Your responses as a field professional will help us evaluate the factors influencing the effective use of big data in organizations located in Armenia.

This survey is estimated about 20 minutes to complete.

Your participation in the survey is completely voluntary and all of your responses will be kept confidential. No personally identifiable information will be associated with your responses to any reports of these data.

Thank you very much for your time and cooperation.

Please note, that your response to this survey is very important contribution to the development of Information Systems Science in the domain of big data management.

Please return the completed document to your colleague (Name, Surname).

Regards,

Elen Grigoryan

- Other

—

YOUR CHARACTERISTICS

Q1: Your age?

- 18-25 years
- 26-35 years
- 36-45 years
- 46-55 years
- 56 years and older

Q2: You are?

- Man
- Woman

Q3: Your position?

- General Manager /CEO
- CIO
- Person with appropriate qualification to work with Big Data/BD technologies (Data Scientist, Data-Engineer, other)

Q4: Your experience with Big Data/BD?

- Less than a year
- 1-2 years
- 3-5 years
- 6-10 years
- More than 10 years

Q5: Company size?

- Small (-50 employees)
- Medium (50-249 employees)
- Large (250 employees and more)

Q6: Industrial sector of your company?

- Manufacturing
- ICT
- Services
- Construction

Q7: Your Company?

- _____ is based initially on big data / start-up
- _____ has adopted technologies to manage big data

Effective use - using a system in a way that helps attain the goals for using the system.

1	2	3	4	5	6	7
Strongly disagree	Disagree	Disagree somewhat	Neither agree nor disagree	Agree somewhat	Agree	Strongly agree

Q: Please indicate (circle) how strongly you agree or disagree with the following items:

Big data system reliability

1. The system operates reliably to manage big data.

1 2 3 4 5 6 7

2. The system performs reliably to manage big data.

1 2 3 4 5 6 7

3. The operation of the system is technically available to manage big data.

1 2 3 4 5 6 7

Big data system adaptability

4. The system can be adapted to meet a variety of business needs.

1 2 3 4 5 6 7

5. The system can flexibly adjust to new demands or conditions during the treatment of big data.

1 2 3 4 5 6 7

6. The system is flexible in addressing needs as they arise during the treatment of big data.

1 2 3 4 5 6 7

Big data system integration

7. The system effectively integrates data from different areas of the company.

1 2 3 4 5 6 7

8. The system pulls together data that used to come from different places in the company.

1 2 3 4 5 6 7

9. The system effectively combines different types of data from all areas of the company.

1 2 3 4 5 6 7

Big data system accessibility

10. The system allows information to be readily accessible as requested.

1 2 3 4 5 6 7

11. The system makes information very accessible.

1 2 3 4 5 6 7

12. The system makes information easy to access.

1 2 3 4 5 6 7

Big data system response time

13. The system does not take long time to process the requests.

1 2 3 4 5 6 7

14. The system provides information in a timely fashion.

1 2 3 4 5 6 7

15. The system processes the requests quickly.

1 2 3 4 5 6 7

Big data system privacy

16. The system protects information about personal data.

1 2 3 4 5 6 7

17. The system protects personal identifiable data.

1 2 3 4 5 6 7

18. The system offers a meaningful guarantee that it will not share private information.

1 2 3 4 5 6 7

The system quality enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F1.2: IT infrastructure

19. My company uses the latest technology to manage big data.

1 2 3 4 5 6 7

20. Our technological competency helps us to enhance big data management.

1 2 3 4 5 6 7

21. We use a variety of technological tools to manage big data.

1 2 3 4 5 6 7

22. Our big data technological tools are more effective than those used by others in the industry.

1 2 3 4 5 6 7

The IT infrastructure enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F1.3: Vendor support

23. The developers of the big data system have provided adequate training to users.

1 2 3 4 5 6 7

The Vendor support enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.1 : Data quality and information quality

24. Our big data sources provide accurate data.

1 2 3 4 5 6 7

25. Our big data providers have a very good reputation in the industry.

1 2 3 4 5 6 7

26. We are satisfied with the quality of the data provided by our big data sources.

1 2 3 4 5 6 7

The Data quality and information quality enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.2 : Data completeness

The Big data system used:

27. ____ provides a complete set of information.

1 2 3 4 5 6 7

28. ____ produces comprehensive information.

1 2 3 4 5 6 7

29. ____ provides all the information needed.

1 2 3 4 5 6 7

The Data completeness enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.3 : Data currency

The Big data system used:

30. ____ provides the most recent information.

1 2 3 4 5 6 7

31. ____ produces the most current information.

1 2 3 4 5 6 7

32. ____ always provides up-to-date information.

1 2 3 4 5 6 7

The Data currency enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.4 : Data access

33. Needed data would be easily retrievable.

1 2 3 4 5 6 7

34. Needed data would be easily accessible.

1 2 3 4 5 6 7

35. Needed data would be easily obtainable.

1 2 3 4 5 6 7

36. The data would be quickly accessible when needed.

1 2 3 4 5 6 7

The Data access enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.5 : Data relevance

37. Information is informative for my needs.

1 2 3 4 5 6 7

38. Information is valuable for my needs.

1 2 3 4 5 6 7

The Data relevance enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.6 : Data accuracy

The Big data system used:

39. ____ produces correct information.

1 2 3 4 5 6 7

40. ____ provides few errors in the information.

1 2 3 4 5 6 7

41. ____ provides accurate information.

1 2 3 4 5 6 7

The Data accuracy enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F2.4 : Data consistency

42. The Big data system assures the consistency of data to users (or applications).

1 2 3 4 5 6 7

The Data consistency enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F3.1 : People's knowledge and skills

43. Our 'big data system' related staff has the right skills to accomplish their jobs successfully.

1 2 3 4 5 6 7

44. Our 'big data system' related staff is well trained.

1 2 3 4 5 6 7

45. The company provides big data system operation's related training to its own employees.

1 2 3 4 5 6 7

46. Our 'big data system' related staff has suitable education to fulfill their jobs.

1 2 3 4 5 6 7

The People's knowledge and skills enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F3.2 : Trust

47. I believe in the benefits of the big data system.

1 2 3 4 5 6 7

48. My peers believe in the benefits of the big data system.

1 2 3 4 5 6 7

49. My management team believes in the project benefits related the big data system.

1 2 3 4 5 6 7

The Trust enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F3.3 : Champions

50. A high-level champion(s) for the big data system came from IS.

1 2 3 4 5 6 7

51. A high-level champion(s) for the big data system came from a functional area(s).

1 2 3 4 5 6 7

The Champions enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F3.4 : Employee engagement

52. The big data system had support from people throughout the organization.

1 2 3 4 5 6 7

The Employee engagement enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F3.5 : User participation

53. IS and users worked together as a team on the big data system- project.

1 2 3 4 5 6 7

54. Users were assigned full-time to parts of the big data system- project.

1 2 3 4 5 6 7

55. Users performed hands-on activities during the big data system- project.

1 2 3 4 5 6 7

The User participation enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F3.6 : Individual Characteristics

56. I feel apprehensive about using the big data system.

1 2 3 4 5 6 7

57. It scares me to think that I could lose a lot of information using the big data system by hitting the wrong key.

1 2 3 4 5 6 7

58. I hesitate to use the big data system for fear of making mistakes I cannot correct.

1 2 3 4 5 6 7

59. The big data system is somewhat intimidating to me.

1 2 3 4 5 6 7

The Individual Characteristics enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F4.1 : Data privacy and security

60. The Big data system protects information against unauthorized access.

1 2 3 4 5 6 7

61. The Big data protects the personal information.

1 2 3 4 5 6 7

62. We feel safe to deal with the big data system.

1 2 3 4 5 6 7

The Data privacy and security enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F4.2 : Data governance

63. In our organization, we have controlled practices regarding data management /e.g. in terms of setting retention policies (e.g. time to live) of data/ backup routines/ establishing/monitoring access (e.g. user access) to data/ classifying data according to value/ monitoring costs versus value of data/.

1 2 3 4 5 6 7

The Data governance enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.1 : Perceived value

The use of big data :

64. ___ creates competitive advantage.

1 2 3 4 5 6 7

65. ___ aligns analytics with business strategy.

1 2 3 4 5 6 7

66. ___ establishes useful links with other organisations.

1 2 3 4 5 6 7

67. ___ enables quicker response to change.

1 2 3 4 5 6 7

68. ___ improves customer relations.

1 2 3 4 5 6 7

69. ___ provides better products or services to customers.

1 2 3 4 5 6 7

The Perceived value enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.2 : Perceived risk

The risks related the uses of big data are:

70. ___reluctance of employees to adapt to changes.

1 2 3 4 5 6 7

71. ___lack of Information System infrastructure support.

1 2 3 4 5 6 7

72. ___technical uncertainty.

1 2 3 4 5 6 7

73. ___minimal IT expertise.

1 2 3 4 5 6 7

74. ___uncertainty about how to measure potential benefits.

1 2 3 4 5 6 7

75. ___uncertainty about how to measure the involved costs.

1 2 3 4 5 6 7

76. ___capital outlay with no guarantee of likely returns.

1 2 3 4 5 6 7

77. ___security issues.

1 2 3 4 5 6 7

78. ___reassignment of personnel trained on big data analytics solutions.

1 2 3 4 5 6 7

79. ___privacy issues.

1 2 3 4 5 6 7

The Perceived risks enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.3 : Perceived usefulness

80. Using the big data system would improve the performance in the job's outcome.

1 2 3 4 5 6 7

81. Using the big data system would increase the productivity.

1 2 3 4 5 6 7

82. Using the big data system would enhance the effectiveness in my job.

1 2 3 4 5 6 7

83. I would find the big data system to be useful in my job.

1 2 3 4 5 6 7

The Perceived usefulness enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.4 : Perceived ease of use

84. My interaction with the big data system would be clear and understandable.

1 2 3 4 5 6 7

85. I would find the big data system easy to use.

1 2 3 4 5 6 7

86. I would find it easy to get the big data system to do what I want it to do.

1 2 3 4 5 6 7

The Perceived ease of use enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.5 : Intention to use

87. I want to use the big data system in the future.

1 2 3 4 5 6 7

88. I expect that I will use the big data system later.

1 2 3 4 5 6 7

89. I plan to use the big data system in the future.

1 2 3 4 5 6 7

The Intention to use enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.6 : Perceived observability

90. Many competitors or business partners in the market have started using Big Data Systems.

1 2 3 4 5 6 7

91. Using Big Data System helps my company to connect with both domestic and international business partners.

1 2 3 4 5 6 7

92. There are many computers that people in the company can access to Big Data System.

1 2 3 4 5 6 7

93. There are many computers that people in the company can access to use Big Data System.

1 2 3 4 5 6 7

94. Big Data Systems shows improved results over doing business the traditional way.

1 2 3 4 5 6 7

The Perceived observability enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F5.7 : Cost-benefit analysis

95. The costs involved in the adoption of big data system would be far greater than the expected benefits.

1 2 3 4 5 6 7

96. The cost of maintaining big data system would be very high for our enterprises.

1 2 3 4 5 6 7

97. The cost involved in providing support systems for big data system would be too high.

1 2 3 4 5 6 7

98. The amount of money invested in training employees to use big data system would be very high.

1 2 3 4 5 6 7

The Cost-benefit analysis enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F6.1 : Process orientation

99. Our company is better than competitors in connecting (e.g., communication and information sharing) parties within a business process.

1 2 3 4 5 6 7

100. Our company is better than competitors in reducing cost within a business process.

1 2 3 4 5 6 7
101. Our company is better than competitors in bringing complex analytical methods to bear on a business process.

1 2 3 4 5 6 7
102. Our company is better than competitors in bringing detailed information into a business process.

1 2 3 4 5 6 7
The Process orientation enables effective use of Big Data in our organization.

1 2 3 4 5 6 7
F6.2 : IT business process integration
103. We have the ability to integrate the processes involved in the big data chain (i.e. data collection, preparation, analysis and decision making).

1 2 3 4 5 6 7
104. The integration of the processes involved in the big data chain reduces the cost of big data use.

1 2 3 4 5 6 7
105. The integration of the processes involved in the big data chain reduces the efforts necessary to analyze big data.

1 2 3 4 5 6 7
The IT business process integration enables effective use of Big Data in our organization.

1 2 3 4 5 6 7
F6.3 : Data management
106. In my organization there is a good fit between the analytical tools we have access to and the data we process.

1 2 3 4 5 6 7
107. The present analytical tools my organization has access to fulfill our data analysis needs.

1 2 3 4 5 6 7
108. The analytical tools that my organization currently has access to provide pretty much everything that we need to analyze our data properly.

1 2 3 4 5 6 7

The Data management enables effective use of Big Data in our organization.

1 2 3 4 5 6 7
Theme 7: Organizational aspects
109. Our decisions are based on data.

1 2 3 4 5 6 7
110. A dependency on hunches for decision making is strongly discouraged in our organization.

1 2 3 4 5 6 7
111. Depending on data is part of our organizational routine.

1 2 3 4 5 6 7
112. We have a culture of data driven work.

1 2 3 4 5 6 7
The Organizational cultural competence enables effective use of Big Data in our organization.

1 2 3 4 5 6 7
F7.2 : Talent management
113. We prefer to hire employees who understand big data.

1 2 3 4 5 6 7
114. We have the ability to recruit expert users of big data.

1 2 3 4 5 6 7
115. We plan to enhance the big data management skills of our staff.

1 2 3 4 5 6 7
116. We take special care in the retention of big data experts in our organization.

1 2 3 4 5 6 7
The Talent management enables effective use of Big Data in our organization.

1 2 3 4 5 6 7
F7.3 : Change management program
117. Change in the organization created by Big Data System was managed effectively.

1 2 3 4 5 6 7
The Change management program enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.4 : Strategic alignment

118. The use of big data aligns with the company’s mission, goals, objectives, and strategies.

1 2 3 4 5 6 7

119. The use of big data contains quantified goals and objectives.

1 2 3 4 5 6 7

120. The use of big data contains detailed action plans/strategies that support company direction.

1 2 3 4 5 6 7

121. We prioritize major big data usage related investments by the expected impact on business performance.

1 2 3 4 5 6 7

The Strategic alignment enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.5 : Project management

122. We continuously examine the innovative opportunities for the strategic use of big data analytics.

1 2 3 4 5 6 7

123. We enforce adequate plans for the introduction and utilization of big data analytics.

1 2 3 4 5 6 7

124. We perform big data project planning processes in systematic and formalized ways.

1 2 3 4 5 6 7

125. We frequently adjust big data project plans to better adapt to changing conditions.

1 2 3 4 5 6 7

The Project management enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.6 : Performance management

We perform much better than our main competitors in terms of:

126. Profitability

1 2 3 4 5 6 7

127. Profits as percentage of sales

1 2 3 4 5 6 7

128. Decreasing product or service delivery cycle time

1 2 3 4 5 6 7

129. In reducing operating costs

1 2 3 4 5 6 7

130. In profit growth rates

1 2 3 4 5 6 7

131. Rapid response to market demand

1 2 3 4 5 6 7

132. Rapid confirmation of customer orders

1 2 3 4 5 6 7

133. Increasing customer satisfaction

1 2 3 4 5 6 7

134. Providing better product and service quality

1 2 3 4 5 6 7

135. In reducing operating costs

1 2 3 4 5 6 7

The Performance management enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.7 : Organizational structure and size

136. We can cost effectively change our organizational structure to respond to changing business conditions.

1 2 3 4 5 6 7

137. We can change our organizational structure without negatively impacting service/product quality.

1 2 3 4 5 6 7

138. Our current organization structure enables us to adapt to changing business conditions.

1 2 3 4 5 6 7

139. Our organization is more flexible than our competitors in changing our organizational structure.

1 2 3 4 5 6 7

The Organizational structure and size enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.8: Interdepartmental collaboration

140. There is collaboration among big data analysts and big data providers.

1 2 3 4 5 6 7

141. There is collaboration among big data analysts and decision makers.

1 2 3 4 5 6 7

142. There is collaboration among big data providers and decision makers.

1 2 3 4 5 6 7

The Interdepartmental collaboration enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.9: Communication

143. Our employees transfer their knowledge about data.

1 2 3 4 5 6 7

144. Knowledge about how data are collected is exchanged within the organization.

1 2 3 4 5 6 7

145. Knowledge about how data are used is exchanged within the organization.

1 2 3 4 5 6 7

146. Knowledge about how data are processed is exchanged within our firm.

1 2 3 4 5 6 7

147. The exchange of knowledge makes it easy for us to analyze data.

1 2 3 4 5 6 7

The Communication enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.10: Top management support

148. The top management is greatly interested in using big data systems.

1 2 3 4 5 6 7

149. The top management is aware of the benefits of big data systems for future success.

1 2 3 4 5 6 7

150. The top management will allocate adequate financial and other resources to

the development and operation of big data systems.

1 2 3 4 5 6 7

151. The top management has a vision to project our company as the market leader in the use of big data systems.

1 2 3 4 5 6 7

The Top management support enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.11: Environmental effect

The use of big data in our organization is affected by:

152. The extent of technological turbulence in the environment.

1 2 3 4 5 6 7

153. The leadership in product/process innovation.

1 2 3 4 5 6 7

154. The impact of new technology on operations.

1 2 3 4 5 6 7

155. The extent of market turbulence.

1 2 3 4 5 6 7

156. The frequent changes in customer preferences.

1 2 3 4 5 6 7

157. The need to reduce market uncertainty

1 2 3 4 5 6 7

158. The need to respond to market opportunities

1 2 3 4 5 6 7

The Environmental effect enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.12: Clear goals

159. By using big data company goals have become clearer.

1 2 3 4 5 6 7

The Clear goals enable effective use of Big Data in our organization.

1 2 3 4 5 6 7

F7.13: Focus on innovation

160. We continuously examine the innovative opportunities for the strategic use of IT.

1 2 3 4 5 6 7

161. We enforce adequate plans for the introduction and utilization of IT.

1 2 3 4 5 6 7

162. We perform IT planning processes in systematic and formalized ways.

1 2 3 4 5 6 7

163. We frequently adjust IT plans to better adapt to changing conditions.

1 2 3 4 5 6 7

The Focus on innovation enables effective use of Big Data in our organization.

1 2 3 4 5 6 7

Mandatory comment

In our organization the value realization from big data depends on the above listed 41 factors:

1 2 3 4 5 6 7

Annexe 4: Détail “Cronbach's Alpha”

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Rows	162.876	7	23.2680	43.8513	1.63E-	2.01579
	2		3	5	56	7
Columns	784.825	210	3.73726	7.04331	6.6E-	1.17999
	8		6	9	119	5
Error	779.998	1470	0.53061			
	8		1			
Total	1727.70	1687				
	1					
Cronbach's Alpha	0.97719					
	6					

Annexe 5 : Caractéristiques de l'échantillon

Characteristic	Frequency	Percentage
Big Data Experience		
< 1 year	15	7.11
1-2 years	72	34.12
3-5 years	86	40.76
6-10 years	38	18.01
10 + years	-	-
Age		
18-25 years	46	21.80
26-35 years	161	76.30
36-45 years	4	1.90
46-55 years	-	-
56 + years	-	-
Gender		
Male	130	61.61
Female	81	38.39
Company size		
Small (<50 employees)	40	18.96
Medium (50-249 employees)	69	32.70
Large (250 employees and more)	102	48.34
Sector type		
ICT	117	55.45
Services	94	44.55
Relation with Big Data		
Based initially on big data	94	44.55
Adopted big data technologies	117	55.45

Annexe 6 : Administration du questionnaire, préparation des données (capture d'écran)

	A	B	C	D	E	F
4	Quest_3	System Quality	2	F3	Vendor Support	Quest_23
5	Quest_4	System Quality	1	F4	Data Quality	Quest_24-Quest_26
6	Quest_5	System Quality	1	F5	Data Completeness	Quest_27-Quest_29
7	Quest_6	System Quality	6	F6	Data Currency	Quest_30-Quest_32
8	Quest_7	System Quality	5	F7	Data access	Quest_33-Quest_36
9	Quest_8	System Quality	2	F8	Data relevance	Quest_37-Quest_38
10	Quest_9	System Quality	8	F9	Data accuracy	Quest_39-Quest_41
11	Quest_10	System Quality	2	F10	Data consistency	Quest_42
12	Quest_11	System Quality	11	F11	Peoples knowledge and skills	Quest_43-Quest_46
13	Quest_12	System Quality	13	F12	Trust	Quest_47-Quest_49
14	Quest_13	System Quality	2	F13	Champions	Quest_50-Quest_51
15	Quest_14	System Quality	1	F14	Employee engagement	Quest_52
16	Quest_15	System Quality	2	F15	User participation	Quest_53-Quest_55
17	Quest_16	System Quality	8	F16	Individual characteristics	Quest_56-Quest_59
18	Quest_17	System Quality	8	F17	Data privacy and security	Quest_60-Quest_62
19	Quest_18	System Quality	8	F18	Data governance	Quest_63
20	Quest_19	IT infrastructure	2	F19	Perceived value	Quest_64-Quest_69
21	Quest_20	IT infrastructure	2	F20	Perceived risk	Quest_70-Quest_79
22	Quest_21	IT infrastructure	2	F21	Perceived usefulness	Quest_80-Quest_83
23	Quest_22	IT infrastructure	2	F22	Perceived ease of use	Quest_84-Quest_86
24	Quest_23	Vendor support	3	F23	Intention to use	Quest_87-Quest_89
25	Quest_24	Data quality and information quality	2	F24	Perceived observability	Quest_90-Quest_94
26	Quest_25	Data quality and information quality	2	F25	Cost-Benefit analysis	Quest_95-Quest_98
27	Quest_26	Data quality and information quality	1	F26	Process orientation	Quest_99-Quest_102
28	Quest_27	Data completeness	5	F27	IT BP integration	Quest_103-Quest_105
29	Quest_28	Data completeness	5	F28	Data management	Quest_106-Quest_108
30	Quest_29	Data completeness	5	F29	Organizational cultural competence	Quest_109-Quest_112
31	Quest_30	Data currency	17	F30	Talent management	Quest_113-Quest_116
32	Quest_31	Data currency	17	F31	Change management program	Quest_117
33	Quest_32	Data currency	4	F32	Strategic alignment	Quest_118-Quest_121
34	Quest_33	Data access	8	F33	Project management	Quest_122-Quest_125
35	Quest_34	Data access	8	F34	Performance management	Quest_126-Quest_135

Annexe 7 : Préparation des données ; une ligne pour chaque cas unique et une colonne pour chaque variable (capture d'écran)

	Quest_1	Quest_2	Quest_3	Quest_4	Quest_5	Quest_6	Quest_7	Quest_8	Quest_9	Quest_10	Quest_11	Quest_12	Quest_13	Quest_14	Quest_15	Quest_16	Quest_17	Quest_18	moyenne F1	Sys Cntrl	Quest_19	Quest_20
ref02	6	6	5	5	5	5	6	6	6	6	6	6	4	4	6	4	4	5	5.2778	6	7	
ref05	6	6	5	6	5	5	6	6	6	6	6	6	4	4	6	4	4	4	5.2778	7	7	
ref08	7	6	7	7	7	7	7	7	5	6	6	6	5	6	6	6	6	7	6.3333	6	6	
ref09	5	4	5	7	6	5	7	7	6	7	6	6	7	6	6	7	7	7	6.1667	6	7	
ref10	5	5	5	7	6	5	7	7	6	7	6	6	7	6	6	7	7	7	6.2222	6	6	
ref15	5	6	5	5	5	5	6	6	5	6	6	6	4	4	6	4	4	5	5.1667	6	7	
ref16	5	5	5	7	6	5	7	7	6	7	6	6	7	7	6	7	7	7	6.2778	6	6	
ref20	6	6	6	6	6	6	6	6	6	6	6	6	4	4	5	4	4	4	5.3889	7	7	
ref24	5	5	5	5	5	5	6	6	5	6	6	5	4	4	6	4	4	5	5.0556	6	7	
ref30	6	6	5	6	5	5	6	6	5	6	6	6	4	4	6	4	4	5	5.2778	6	7	
ref34	6	6	6	6	5	5	5	6	6	6	6	6	4	5	5	4	4	4	5.2778	7	7	
ref36	6	6	6	5	5	5	6	6	6	6	6	6	4	4	5	4	4	4	5.2222	7	6	
ref38	5	5	6	6	5	5	6	6	6	6	6	6	4	4	5	4	4	4	5.1667	7	6	
ref40	6	6	6	6	5	5	6	6	6	6	6	6	5	5	5	4	4	4	5.3889	7	7	
ref44	6	6	6	5	5	6	6	6	6	6	6	6	5	5	5	5	5	5	5.5556	7	7	
ref46	6	6	6	5	5	5	6	6	6	6	6	6	4	4	4	4	4	4	5.1667	7	6	
ref48	7	6	7	7	7	7	7	7	5	6	6	6	5	6	6	6	6	7	6.3333	7	6	
ref51	5	4	5	7	6	5	7	7	6	7	6	6	7	6	6	7	7	7	6.1667	6	7	
ref52	5	5	5	5	6	5	7	6	6	7	6	6	7	7	6	7	7	7	6.1111	6	6	
ref53	6	6	6	6	5	5	6	6	6	6	6	6	4	4	5	4	4	4	5.2778	7	7	
ref56	6	6	5	5	5	5	6	6	6	6	6	6	4	4	4	4	4	4	5.1111	7	7	
ref58	5	5	5	7	6	5	7	7	6	7	6	6	7	6	6	7	7	7	6.2222	5	7	
ref60	6	6	7	7	7	7	7	7	5	6	6	6	5	6	6	6	6	7	6.2778	7	7	
ref62	6	6	7	7	7	7	7	7	5	6	6	6	5	6	6	6	6	7	6.2778	6	7	
ref64	5	5	6	6	5	5	6	6	6	6	6	6	5	4	5	4	4	4	5.2222	7	7	
ref66	6	6	5	7	5	5	7	7	6	7	6	6	7	6	6	7	7	7	6.2778	5	7	
ref68	7	6	7	7	7	7	7	7	5	6	6	6	5	6	6	6	6	7	6.3333	7	6	
ref70	5	4	5	7	6	5	7	7	6	7	6	6	7	6	6	7	7	7	6.1667	7	7	
ref72	6	6	6	6	5	5	6	6	6	6	6	6	5	4	5	6	4	4	5.4444	7	7	

Annexe 8 : Le rapport de l'écart-type à la moyenne en pourcentage pour 41 facteurs

<i>Factor</i>	<i>x</i>	<i>S</i>	<i>CV (%)</i>
Syst_Qual	5.821513	0.51834	8.903865
Infra_IT	6.617021	0.734794	11.1046
Vend_Supp	5.851064	0.82901	14.16853
Data_Qual	5.939716	0.832383	14.01386
Data_Complet	6.049645	0.493585	8.158904
Data_Curr	6.400709	0.386696	6.04145
Data_acc	6.085106	0.61626	10.12734
Data_relev	6.058511	0.642811	10.61005
Data_accur	6.049645	0.738039	12.1997
Data_consistency	5.734043	0.467791	8.15814
Peop_knowl_skls	6.231383	0.38493	6.177282
Trust	6.794326	0.293727	4.323126
Champions	5.202128	0.453891	8.725101
Empl_engage	5.074468	0.42094	8.295256
User_particip	4.670213	0.593664	12.7117
Individ_charac	4.978723	0.528311	10.61137
Data_priv_sec	5.131206	0.469419	9.14831
Data_gov	6.138298	0.824374	13.43001
Pvd_value	5.39539	0.454622	8.426116
Pvd_risk	4.893617	0.426743	8.720395
Pvd_usefls	5.06383	0.321572	6.35038
Pvd_ease_use	5.719858	0.310329	5.425462

Intent_use	5.58156	0.804642	14.41608
Pvd_observ	5.085106	0.37844	7.442121
Cost-Ben_analys	4.765957	0.425669	8.931451
Process_orient	5.585106	0.740126	13.25178
IT_BP_integr	5.312057	0.262369	4.939124
Data_manag	6.829787	0.377835	5.53216
Organl_cult_compet	4.893617	0.451237	9.220928
Tal_mangmnt	4.691489	0.529284	11.2818
Chnge_managmnt_prog	4.553191	0.649516	14.26508
Str_alignmnt	4.542553	0.542097	11.93374
Projct_managmnt	5.968085	0.594819	9.966658
Perf_managmnt	5.893617	0.30998	5.259595
Organl_struct_size	4.5	0.523635	11.63633
Interdep_collab	6.723404	0.449713	6.688774
Communication	6.7234	0.4497	6.6888
Top_managmnt_supprt	6.861702	0.347063	5.057978
Envir_effect	5.606383	0.55296	9.863052
Clr_goals	6.159574	0.368175	5.977273
Focus_innov	6.130319	0.575869	9.393778

Annexe 9 : Test d'analyse de variance pour les 2 groupes de répondants « utilisateurs novices » et « utilisateurs avancés »

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Column 1	41	214.0698	5.221215	0.343809
Column 2	41	182.4538	4.450093	1.207111

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	12.1899	1	12.1899	15.71956	0.000159	3.960352
Within Groups	62.03683	80	0.77546			
Total	74.22672	81				

Annexe 10 : Données brutes par facteur (capture d'écran)

D1	DJ	DK	DL	DM	DN	DO	DP	DQ	DR	DS	DT	DU	DV	DW	DX	DY	DZ	EA	EB	EC	ED	
F16 Indiv characteristics					F17 Data privacy and security			F18 Data govern			govern			Quest_64		Quest_65						
	Quest_56	Quest_57	Quest_58	Quest_59	moyenne F16	charac cntrl	charac cntrl coef	Quest_60	Quest_61	Quest_62	moyenne F17	pr sec cntrl	pr sec cntrl coef	Quest_63	govern cntrl	govern cntrl coef						
3	5.428571	4	1	1	4	2.5	7	1	2.5	4	4	6	4.666667	7	1	4.666667	7	7	1	7	5	6
4	2.857143	5	1	1	1	2	4	0.571429	1.142857	4	4	6	4.666667	7	1	4.666667	7	7	1	7	5	6
5	5.428571	2	1	1	1	1.25	7	1	1.25	6	6	6	6	7	1	6	7	6	0.857143	6	6	6
6	5.428571	5	1	1	5	3	7	1	3	5	5	6	5.333333	7	1	5.333333	7	7	1	7	5	6
7	2.857143	5	1	1	1	2	4	0.571429	1.142857	4	4	6	4.666667	7	1	4.666667	7	7	1	7	5	6
8	6	6	6	5	5	5.5	7	1	5.5	6	6	6	6	7	1	6	6	7	1	6	7	7
9	2	4	4	4	4	4	7	1	4	4	6	4	4.666667	7	1	4.666667	4	6	0.857143	3.428571	5	6
10	3.238095	6	6	5	4	5.25	4	0.571429	3	6	6	5	5.666667	6	0.857143	4.857143	6	6	0.857143	5.142857	6	6
11	3.571429	1	4	4	5	3.5	5	0.714286	2.5	5	5	6	5.333333	6	0.857143	4.571429	5	6	0.857143	4.285714	5	5
12	3.571429	1	4	4	5	3.5	5	0.714286	2.5	5	5	5	5	6	0.857143	4.285714	5	6	0.857143	4.285714	5	5
13	7	7	5	5	6	5.75	7	1	5.75	5	6	6	5.666667	6	0.857143	4.857143	6	6	0.857143	5.142857	7	7
14	4.285714	1	1	1	1	1	7	1	1	7	7	7	7	6	0.857143	6	7	6	0.857143	6	6	7
15	1.714286	6	2	2	1	2.75	7	1	2.75	7	7	7	7	7	1	7	6	7	1	6	7	7
16	5	3	4	4	2	3.25	7	1	3.25	4	3	5	4	7	1	4	5	6	0.857143	4.285714	6	4
17	2.857143	5	1	1	1	2	4	0.571429	1.142857	4	4	6	4.666667	7	1	4.666667	7	7	1	7	5	5
18	3.571429	1	4	4	5	3.5	4	0.571429	2	5	5	5	5	6	0.857143	4.285714	5	6	0.857143	4.285714	5	5
19	6.333333	2	1	1	1	1.25	7	1	1.25	6	6	6	6	7	1	6	7	6	0.857143	6	6	6
20	4.571429	6	5	6	5	5.5	7	1	5.5	6	6	5	5.666667	7	1	5.666667	5	7	1	5	5	6
21	5.428571	2	1	1	1	1.25	6	0.857143	1.071429	6	6	6	6	6	0.857143	5.142857	7	6	0.857143	6	6	6
22	2.85714	5	1	1	1	2	5	0.714286	1.428571	4	5	6	5	7	1	5	7	6	0.857143	6	6	6
23	7	2	2	2	2	2	7	1	2	7	7	7	7	7	1	7	7	7	1	7	7	7
24	5.142857	6	5	1	1	3.25	6	0.857143	2.785714	5	5	5	5	6	0.857143	4.285714	7	6	0.857143	6	6	7
25	5.333333	3	3	2	1	2.25	7	1	2.25	4	4	4	4	7	1	4	7	7	1	7	7	7
26	2.857143	5	1	1	1	2	4	0.571429	1.142857	4	4	6	4.666667	7	1	4.666667	7	7	1	7	5	6
27	4.571429	6	5	6	5	5.5	7	1	5.5	6	6	5	5.666667	7	1	5.666667	5	7	1	5	5	6
28	4.285714	3	4	4	2	3.25	7	1	3.25	4	3	5	4	7	1	4	5	7	1	5	6	4
29	5.333333	2	2	2	2	2	7	1	2	6	6	6	6	7	1	6	7	1	6	6	5	6
30	1.714286	6	2	2	1	2.75	7	1	2.75	7	7	7	7	7	1	7	6	6	0.857143	5.142857	7	7
31	4.285714	3	4	4	2	3.25	6	0.857143	2.785714	4	3	5	4	6	0.857143	3.428571	5	6	0.857143	4.285714	6	4

Annexe 11 : Matrices de corrélation (Pearson) par groupe d'utilisateurs (captures d'écrans) Utilisateurs avancés

File Home Insert Page Layout Formulas Data Review View XLSTAT

Start XLSTAT

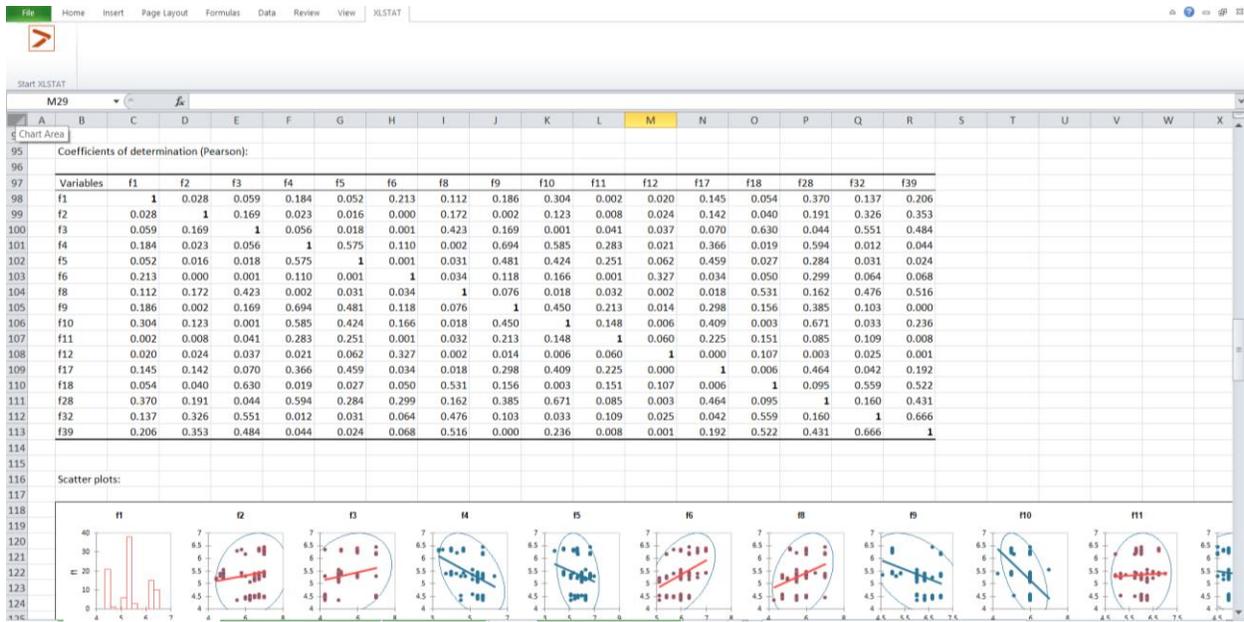
A10

	A	B	C	D	E	F	G	H	I	J
10										
11		Summary statistics:								
12										
13		Variable	Observations	Obs. with missing	Obs. without	Minimum	Maximum	Mean	Std. deviation	
14		f1	94	0	94	4.333	6.444	5.365	0.649	
15		f2	94	0	94	4.464	7.000	6.344	0.768	
16		f3	94	0	94	4.286	7.000	5.524	0.919	
17		f4	94	0	94	3.333	6.333	5.103	0.697	
18		f5	94	0	94	3.810	7.000	5.686	0.676	
19		f6	94	0	94	5.143	7.000	6.017	0.531	
20		f8	94	0	94	4.286	7.000	5.810	0.649	
21		f9	94	0	94	4.571	7.000	6.013	0.754	
22		f10	94	0	94	4.286	7.000	5.681	0.498	
23		f11	94	0	94	4.929	7.000	6.001	0.497	
24		f12	94	0	94	4.524	7.000	5.469	0.702	
25		f17	94	0	94	4.000	6.000	5.005	0.526	
26		f18	94	0	94	4.286	7.000	5.866	0.868	
27		f28	94	0	94	3.714	7.000	5.860	1.113	
28		f32	94	0	94	2.286	4.821	3.357	0.860	
29		f39	94	0	94	2.857	6.571	4.188	1.056	
30										
31										

Start XLSTAT

M29

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
31																			
32		Correlation matrix (Pearson):																	
33																			
34		Variables	f1	f2	f3	f4	f5	f6	f8	f9	f10	f11	f12	f17	f18	f28	f32	f39	
35		f1	1	0.166	0.242	-0.429	-0.229	0.462	0.335	-0.431	-0.551	0.046	-0.140	0.381	-0.232	-0.609	0.371	0.454	
36		f2	0.166	1	0.411	-0.151	-0.127	0.006	0.414	-0.041	-0.350	0.088	0.155	0.377	-0.201	-0.438	0.571	0.594	
37		f3	0.242	0.411	1	0.236	0.136	0.031	0.650	0.411	-0.028	-0.204	-0.191	0.265	-0.794	-0.209	0.742	0.696	
38		f4	-0.429	-0.151	0.236	1	0.758	-0.332	0.040	0.833	0.765	-0.532	-0.144	-0.605	-0.138	0.771	0.110	-0.209	
39		f5	-0.229	-0.127	0.136	0.758	1	0.027	0.176	0.693	0.652	-0.501	-0.249	-0.678	-0.165	0.533	0.176	-0.156	
40		f6	0.462	0.006	0.031	-0.332	0.027	1	0.186	-0.343	-0.407	-0.031	-0.572	0.186	-0.225	-0.547	0.253	0.262	
41		f8	0.335	0.414	0.650	0.040	0.176	0.186	1	0.276	-0.135	-0.178	-0.048	0.133	-0.729	-0.403	0.690	0.719	
42		f9	-0.431	-0.041	0.411	0.833	0.693	-0.343	0.276	1	0.671	-0.461	-0.120	-0.546	-0.395	0.621	0.321	-0.013	
43		f10	-0.551	-0.350	-0.028	0.765	0.652	-0.407	-0.135	0.671	1	-0.385	-0.075	-0.640	0.059	0.819	-0.181	-0.486	
44		f11	0.046	0.088	-0.204	-0.532	-0.501	-0.031	-0.178	-0.461	-0.385	1	0.244	0.474	0.389	-0.292	-0.330	-0.090	
45		f12	-0.140	0.155	-0.191	-0.144	-0.249	-0.572	-0.048	-0.120	-0.075	0.244	1	-0.008	0.327	0.051	-0.157	-0.023	
46		f17	0.381	0.377	0.265	-0.605	-0.678	0.186	0.133	-0.546	-0.640	0.474	-0.008	1	-0.075	-0.681	0.204	0.438	
47		f18	-0.232	-0.201	-0.794	-0.138	-0.165	-0.225	-0.729	-0.395	0.059	0.389	0.327	-0.075	1	0.308	-0.748	-0.723	
48		f28	-0.609	-0.438	-0.209	0.771	0.533	-0.547	-0.403	0.621	0.819	-0.292	0.051	-0.681	0.308	1	-0.400	-0.656	
49		f32	0.371	0.571	0.742	0.110	0.176	0.253	0.690	0.321	-0.181	-0.330	-0.157	0.204	-0.748	-0.400	1	0.816	
50		f39	0.454	0.594	0.696	-0.209	-0.156	0.262	0.719	-0.013	-0.486	-0.090	-0.023	0.438	-0.723	-0.656	0.816	1	
51																			
52																			



Utilisateurs novices

File Home Insert Page Layout Formulas Data Review View XLSTAT

Start XLSTAT

A11

Variable	Observations	Obs. with missing	Obs. without	Minimum	Maximum	Mean	Std. deviation
f5	117	0	117	3.143	7.000	5.075	0.812
f8	117	0	117	2.571	7.000	5.334	1.113
f9	117	0	117	3.429	7.000	5.257	1.050
f10	117	0	117	3.429	7.000	5.414	0.836
f11	117	0	117	2.357	7.000	5.583	0.841
f12	117	0	117	4.286	7.000	5.928	0.827
f13	117	0	117	2.857	7.000	5.082	0.870
f14	117	0	117	2.286	7.000	5.680	1.050
f17	117	0	117	3.429	7.000	5.544	0.990
f18	117	0	117	3.429	7.000	5.912	0.946
f19	117	0	117	2.762	7.000	5.752	0.963
f21	117	0	117	3.571	7.000	5.826	0.794
f22	117	0	117	2.286	7.000	5.591	1.064
f23	117	0	117	3.429	7.000	5.958	0.865
f24	117	0	117	2.400	7.000	5.327	0.818
f26	117	0	117	2.571	7.000	5.330	1.001
f27	117	0	117	3.238	7.000	5.617	0.840
f28	117	0	117	2.857	7.000	5.635	1.176
f29	117	0	117	2.250	7.000	5.431	0.848
f31	117	0	117	1.000	7.000	5.630	1.279
f32	117	0	117	1.000	7.000	5.705	1.004
f33	117	0	117	0.857	14.500	5.955	1.896
f34	117	0	117	3.429	7.000	5.201	0.832
f35	117	0	117	2.857	7.000	5.080	1.078
f36	117	0	117	2.000	7.000	5.613	1.058
f37	117	0	117	2.000	7.000	5.535	1.275
f38	117	0	117	1.250	7.000	5.621	1.048
f40	117	0	117	2.571	7.000	5.009	1.372
f41	117	0	117	0.857	7.000	5.612	1.091

File Home Insert Page Layout Formulas Data Review View XLSTAT

Start XLSTAT

B45 Correlation matrix (Pearson):

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
46	Variables	f5	f8	f9	f10	f11	f12	f13	f14	f17	f18	f19	f21	f22	f23	f24	f26	f27	f28	f29	f31	f32	f33	f34
47	f5	1	0.318	0.681	0.528	0.316	0.267	0.117	-0.165	0.237	0.336	0.176	0.249	0.361	-0.115	-0.325	0.438	0.269	0.297	0.364	0.131	0.100	0.191	0.264
48	f8	0.318	1	0.230	0.629	0.513	0.118	0.234	0.213	0.719	0.316	0.265	0.117	0.343	0.082	-0.281	0.603	0.592	0.786	-0.014	0.223	0.240	0.310	0.342
49	f9	0.681	0.230	1	0.439	0.051	-0.036	-0.312	-0.488	0.123	0.162	0.052	0.067	0.403	-0.047	-0.364	0.255	0.162	0.063	0.207	-0.137	-0.251	0.195	0.086
50	f10	0.528	0.629	0.439	1	0.465	0.243	-0.025	-0.019	0.337	0.367	0.100	0.166	0.231	-0.147	-0.264	0.290	0.243	0.421	0.127	0.046	0.175	0.340	0.092
51	f11	0.316	0.513	0.051	0.465	1	0.506	0.410	0.467	0.522	0.589	0.210	0.275	0.314	0.112	-0.165	0.347	0.339	0.620	0.277	0.370	0.447	0.309	0.322
53	f12	0.267	0.118	-0.036	0.243	0.506	1	0.468	0.432	0.294	0.194	0.132	0.219	-0.076	-0.044	-0.201	0.099	0.181	0.322	0.567	0.423	0.465	0.309	0.322
54	f13	0.117	0.234	-0.312	-0.025	0.410	0.468	1	0.623	0.457	0.308	0.296	0.266	0.091	0.102	0.099	0.461	0.456	0.425	0.354	0.338	0.464	-0.088	0.404
55	f14	-0.165	0.213	-0.488	-0.019	0.467	0.432	0.623	1	0.403	0.171	0.269	0.221	-0.001	0.086	0.271	0.171	0.349	0.362	0.339	0.680	0.634	0.188	0.165
56	f17	0.237	0.719	0.123	0.337	0.522	0.294	0.457	0.403	1	0.367	0.371	0.279	0.390	-0.025	-0.299	0.652	0.660	0.762	0.185	0.459	0.329	0.452	0.274
57	f18	0.336	0.316	0.162	0.367	0.589	0.194	0.308	0.171	0.367	1	0.234	0.119	0.260	0.062	-0.161	0.466	0.313	0.494	-0.026	0.061	0.135	0.034	0.180
58	f19	0.176	0.265	0.052	0.100	0.210	0.132	0.296	0.269	0.371	0.234	1	0.642	0.644	0.309	0.342	0.631	0.407	0.515	0.115	0.351	0.142	0.226	0.161
59	f21	0.249	0.137	0.067	0.166	0.275	0.219	0.266	0.221	0.279	0.119	0.642	1	0.670	0.489	0.338	0.402	0.335	0.245	0.390	0.290	0.312	0.131	-0.084
60	f22	0.361	0.343	0.403	0.231	0.314	-0.076	0.091	-0.001	0.390	0.260	0.644	0.670	1	0.532	0.097	0.600	0.430	0.444	0.106	0.181	0.036	0.173	-0.065
61	f23	-0.115	0.082	-0.047	-0.147	0.112	-0.044	0.102	0.086	-0.025	0.062	0.309	0.489	0.532	1	0.320	0.255	0.215	0.141	0.059	0.004	0.115	-0.241	-0.023
62	f24	-0.325	-0.281	-0.364	-0.264	-0.165	-0.201	0.099	0.271	-0.299	-0.161	0.342	0.338	0.097	0.320	1	-0.074	0.007	-0.251	0.175	0.108	0.188	-0.227	-0.180
63	f26	0.438	0.603	0.255	0.290	0.347	0.099	0.461	0.171	0.652	0.466	0.631	0.402	0.600	0.255	-0.074	1	0.761	0.676	0.076	0.268	0.205	0.083	0.404
64	f27	0.269	0.592	0.162	0.243	0.339	0.181	0.456	0.349	0.660	0.313	0.407	0.335	0.430	0.215	0.007	0.761	1	0.574	0.297	0.461	0.476	0.208	0.461
65	f28	0.297	0.786	0.053	0.421	0.620	0.322	0.425	0.362	0.762	0.494	0.515	0.245	0.444	0.141	-0.251	0.676	0.574	1	0.043	0.387	0.318	0.356	0.410
66	f29	0.364	-0.014	0.207	0.127	0.277	0.567	0.354	0.339	0.185	-0.026	0.115	0.390	0.106	0.059	0.175	0.076	0.297	0.043	1	0.512	0.576	0.208	0.243
67	f31	0.131	0.223	-0.137	0.046	0.370	0.423	0.338	0.680	0.459	0.061	0.351	0.290	0.181	0.004	0.108	0.268	0.461	0.387	0.512	1	0.733	0.543	0.209
68	f32	0.100	0.240	-0.251	0.175	0.447	0.465	0.476	0.634	0.329	0.135	0.142	0.312	0.036	0.115	0.188	0.205	0.476	0.318	0.576	0.733	1	0.327	0.329
69	f33	0.191	0.310	0.195	0.340	0.285	0.309	-0.088	0.188	0.452	0.034	0.226	0.131	0.173	-0.241	-0.227	0.083	0.208	0.356	0.208	0.543	0.327	1	0.067
70	f34	0.264	0.342	0.086	0.092	0.137	0.322	0.404	0.165	0.274	0.180	0.161	-0.084	-0.065	-0.023	-0.180	0.404	0.461	0.410	0.243	0.209	0.329	0.067	1
71	f35	0.354	0.246	0.105	0.228	0.241	0.634	0.452	0.331	0.404	0.208	0.480	0.192	0.132	-0.133	-0.126	0.476	0.411	0.467	0.416	0.540	0.406	0.417	0.481
72	f36	0.256	0.516	-0.088	0.391	0.468	0.510	0.507	0.387	0.466	0.343	0.238	0.084	-0.009	-0.119	-0.174	0.431	0.479	0.655	0.251	0.347	0.553	0.325	0.682
73	f37	0.198	0.529	-0.158	0.252	0.421	0.392	0.472	0.486	0.483	0.277	0.419	0.042	0.102	-0.090	-0.085	0.513	0.405	0.706	0.074	0.528	0.490	0.324	0.517
74	f38	0.410	0.502	0.178	0.404	0.381	0.448	0.417	0.288	0.516	0.255	0.424	0.184	0.133	-0.159	-0.158	0.559	0.491	0.610	0.354	0.347	0.392	0.309	0.638
75	f40	0.309	0.206	0.205	0.102	0.149	-0.010	0.313	0.103	0.493	0.208	0.641	0.317	0.497	-0.054	0.086	0.618	0.336	0.409	0.116	0.142	-0.074	0.111	0.205
76	f41	0.195	0.402	0.455	0.209	0.136	0.002	-0.102	-0.102	0.274	0.122	0.192	-0.079	0.287	-0.010	-0.258	0.269	0.370	0.372	0.095	0.153	0.073	0.359	0.541

File Home Insert Page Layout Formulas Data Review View XLSTAT

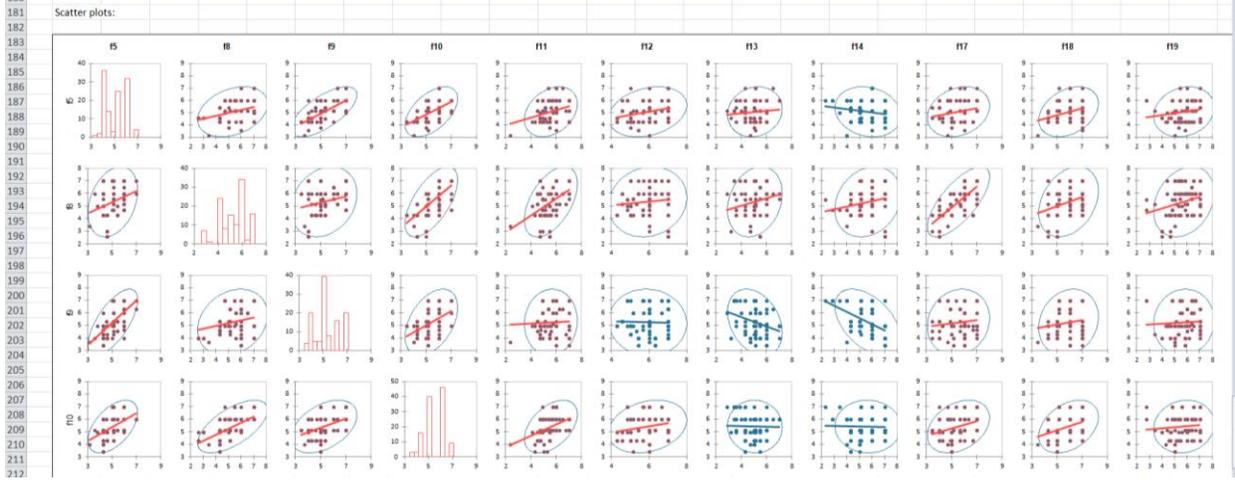
Start XLSTAT

Q47 f24

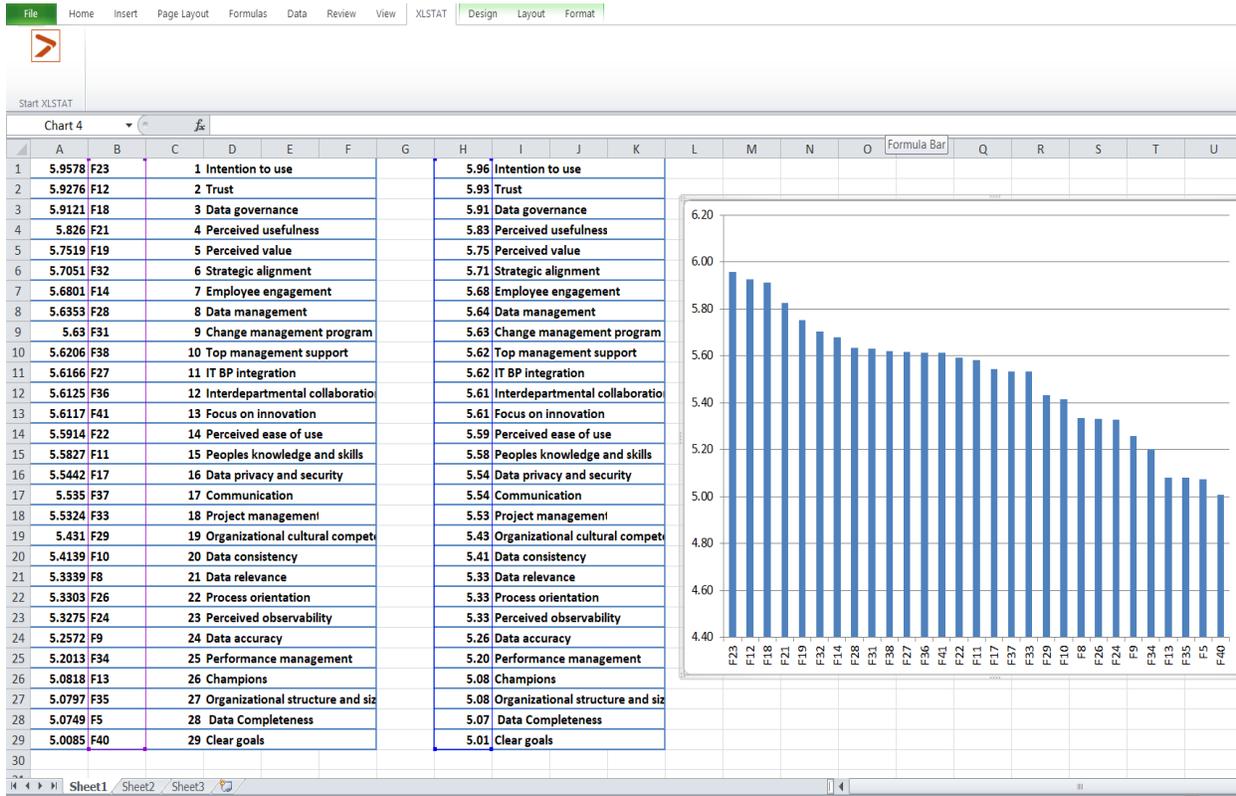
M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	
46	f19	f21	f22	f23	f24	f26	f27	f28	f29	f31	f32	f33	f34	f35	f36	f37	f38	f40	f41
48	0.176	0.249	0.361	-0.115	-0.325	0.438	0.269	0.297	0.364	0.131	0.100	0.191	0.264	0.354	0.256	0.198	0.410	0.309	0.195
49	0.265	0.117	0.343	0.082	-0.281	0.603	0.592	0.786	-0.014	0.223	0.240	0.310	0.342	0.246	0.516	0.529	0.502	0.206	0.402
50	0.052	0.067	0.403	-0.047	-0.364	0.255	0.162	0.063	0.207	-0.137	-0.251	0.195	0.342	0.105	-0.088	-0.158	0.178	0.205	0.455
51	0.100	0.166	0.231	-0.147	-0.264	0.290	0.243	0.421	0.127	0.046	0.175	0.340	0.092	0.228	0.391	0.252	0.404	0.102	0.209
52	0.210	0.275	0.314	0.112	-0.165	0.347	0.339	0.620	0.277	0.370	0.447	0.285	0.137	0.241	0.468	0.421	0.381	0.149	0.136
53	0.132	0.219	-0.076	-0.044	-0.201	0.099	0.181	0.322	0.567	0.423	0.465	0.309	0.322	0.634	0.510	0.392	0.448	-0.010	0.002
54	0.296	0.266	0.091	0.102	0.099	0.461	0.456	0.425	0.354	0.338	0.476	-0.088	0.404	0.452	0.507	0.472	0.417	0.313	-0.102
55	0.269	0.221	-0.001	0.086	0.271	0.171	0.349	0.362	0.339	0.680	0.634	0.188	0.165	0.331	0.387	0.486	0.288	0.103	-0.102
56	0.371	0.279	0.390	-0.025	-0.299	0.652	0.660	0.762	0.185	0.459	0.329	0.452	0.274	0.404	0.466	0.483	0.516	0.439	0.274
57	0.234	0.119	0.260	0.062	-0.161	0.466	0.313	0.494	-0.026	0.061	0.135	0.034	-0.180	0.208	0.343	0.277	0.255	0.208	0.122
58	1	0.642	0.644	0.309	0.342	0.631	0.407	0.515	0.115	0.351	0.142	0.226	0.161	0.480	0.238	0.419	0.424	0.641	0.192
59	0.642	1	0.670	0.489	0.338	0.402	0.335	0.245	0.390	0.290	0.312	0.131	-0.084	0.192	0.084	0.042	0.184	0.317	-0.079
60	0.644	0.670	1	0.532	0.097	0.600	0.430	0.444	0.106	0.181	0.036	0.173	-0.065	0.132	-0.009	0.102	0.133	0.497	0.287
61	0.309	0.489	0.532	1	0.320	0.255	0.215	0.141	0.059	0.004	0.115	-0.241	-0.023	-0.133	-0.119	-0.090	-0.159	-0.054	-0.010
62	0.342	0.338	0.097	0.320	1	-0.074	0.007	-0.251	0.175	0.108	0.188	-0.227	-0.180	-0.126	-0.174	-0.085	-0.158	0.086	-0.258
63	0.631	0.402	0.600	0.255	-0.074	1	0.761	0.676	0.076	0.268	0.205	0.083	0.404	0.476	0.431	0.513	0.559	0.618	0.269
64	0.407	0.335	0.430	0.215	0.007	0.761	1	0.574	0.297	0.461	0.476	0.208	0.461	0.411	0.479	0.405	0.491	0.336	0.370
65	0.515	0.245	0.444	0.141	-0.251	0.676	0.574	1	0.043	0.387	0.318	0.356	0.410	0.467	0.655	0.706	0.610	0.409	0.372
66	0.115	0.390	0.106	0.059	0.175	0.076	0.297	0.043	1	0.512	0.576	0.208	0.243	0.416	0.251	0.074	0.354	0.116	0.095
67	0.351	0.290	0.181	0.004	0.108	0.268	0.461	0.387	0.512	1	0.733	0.543	0.209	0.540	0.347	0.528	0.347	0.142	0.153
68	0.142	0.312	0.036	0.115	0.188	0.205	0.476	0.318	0.576	0.733	1	0.327	0.329	0.406	0.553	0.490	0.392	-0.074	0.073
69	0.226	0.131	0.173	-0.241	-0.227	0.083	0.208	0.356	0.208	0.543	0.327	1	0.067	0.417	0.32				

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	
147		Coefficients of determination (Pearson):																									
148																											
149		Variables	f5	f8	f9	f10	f11	f12	f13	f14	f17	f18	f19	f21	f22	f23	f24	f26	f27	f28	f29	f31	f32	f33	f34		
150	f5		1	0.101	0.464	0.279	0.100	0.071	0.014	0.027	0.056	0.113	0.031	0.062	0.131	0.013	0.106	0.192	0.072	0.089	0.132	0.017	0.010	0.037	0.070		
151	f8			1	0.053	0.396	0.263	0.014	0.055	0.045	0.517	0.100	0.070	0.014	0.117	0.007	0.079	0.363	0.351	0.617	0.000	0.050	0.058	0.096	0.117		
152	f9				1	0.193	0.003	0.001	0.097	0.238	0.015	0.026	0.003	0.004	0.163	0.002	0.132	0.065	0.026	0.004	0.043	0.019	0.063	0.038	0.007		
153	f10					1	0.216	0.059	0.001	0.000	0.114	0.135	0.010	0.028	0.053	0.022	0.069	0.084	0.059	0.177	0.016	0.002	0.031	0.115	0.009		
154	f11						1	0.256	0.168	0.218	0.272	0.346	0.044	0.076	0.099	0.013	0.027	0.120	0.115	0.385	0.077	0.137	0.200	0.081	0.019		
155	f12							1	0.219	0.187	0.086	0.038	0.017	0.048	0.006	0.002	0.040	0.010	0.033	0.103	0.322	0.179	0.216	0.095	0.104		
156	f13								1	0.388	0.209	0.095	0.088	0.071	0.008	0.010	0.010	0.212	0.208	0.180	0.125	0.114	0.227	0.008	0.163		
157	f14									1	0.162	0.029	0.073	0.049	0.000	0.007	0.074	0.029	0.122	0.131	0.115	0.463	0.403	0.205	0.027		
158	f17										1	0.134	0.137	0.078	0.152	0.001	0.089	0.425	0.435	0.580	0.034	0.211	0.108	0.204	0.075		
159	f18											1	0.055	0.014	0.068	0.004	0.026	0.218	0.098	0.244	0.001	0.004	0.018	0.001	0.032		
160	f19												1	0.413	0.415	0.095	0.117	0.398	0.166	0.265	0.013	0.123	0.020	0.051	0.026		
161	f21													1	0.450	0.239	0.114	0.162	0.112	0.060	0.152	0.084	0.097	0.017	0.007		
162	f22														1	0.283	0.010	0.359	0.185	0.197	0.011	0.033	0.001	0.030	0.004		
163	f23															1	0.102	0.065	0.046	0.020	0.003	0.000	0.013	0.058	0.001		
164	f24																1	0.006	0.000	0.063	0.030	0.012	0.025	0.052	0.032		
165	f26																	1	0.580	0.456	0.006	0.072	0.042	0.007	0.164		
166	f27																		1	0.329	0.088	0.212	0.226	0.043	0.212		
167	f28																			1	0.002	0.150	0.101	0.126	0.168		
168	f29																				1	0.262	0.331	0.043	0.059		
169	f31																					1	0.538	0.295	0.043		
170	f32																						1	0.107	0.108		
171	f33																							1	0.004		
172	f34																								1		
173	f35																									1	
174	f36																									1	
175	f37																									1	
176	f38																									1	
177	f40																									1	
178	f41																									1	

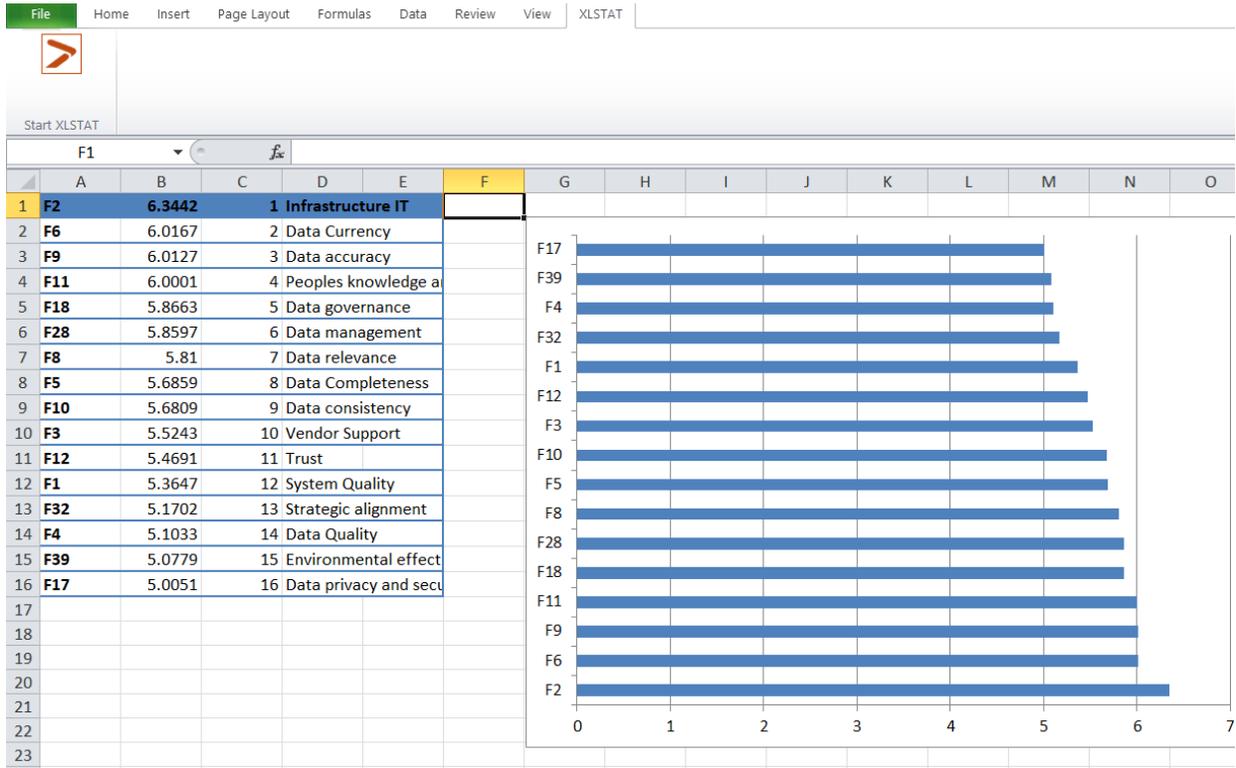
		I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF							
147																																
148																																
149		f13	f14	f17	f18	f19	f21	f22	f23	f24	f26	f27	f28	f29	f31	f32	f33	f34	f35	f36	f37	f38	f40	f41								
150	f13		0.014	0.027	0.056	0.113	0.031	0.062	0.131	0.013	0.106	0.192	0.072	0.089	0.132	0.017	0.010	0.037	0.070	0.125	0.066	0.039	0.168	0.095	0.038							
151	f14			0.055	0.045	0.517	0.100	0.070	0.014	0.117	0.007	0.079	0.363	0.351	0.617	0.000	0.050	0.058	0.096	0.117	0.061	0.266	0.280	0.252	0.043	0.162						
152	f17				0.097	0.238	0.015	0.026	0.003	0.004	0.163	0.002	0.132	0.065	0.026	0.004	0.043	0.019	0.063	0.038	0.007	0.011	0.008	0.025	0.032	0.042	0.207					
153	f18					0.001	0.000	0.114	0.135	0.010	0.028	0.053	0.022	0.069	0.084	0.059	0.177	0.016	0.002	0.031	0.115	0.009	0.052	0.153	0.064	0.163	0.010	0.044				
154	f19						0.216	0.059	0.001	0.000	0.114	0.135	0.010	0.028	0.053	0.022	0.069	0.084	0.059	0.177	0.016	0.002	0.031	0.115	0.009	0.052	0.153	0.064	0.163	0.010	0.044	
155	f21							1	0.256	0.168	0.218	0.272	0.346	0.044	0.076	0.099	0.013	0.027	0.120	0.115	0.385	0.077	0.137	0.200	0.081	0.019						
156	f22								1	0.219	0.187	0.086	0.038	0.017	0.048	0.006	0.002	0.040	0.010	0.033	0.103	0.322	0.179	0.216	0.095	0.104						
157	f23									1	0.388	0.209	0.095	0.088	0.071	0.008	0.010	0.010	0.212	0.208	0.180	0.125	0.114	0.227	0.008	0.163						
158	f24										1	0.162	0.029	0.073	0.049	0.000	0.007	0.074	0.029	0.122	0.131	0.115	0.463	0.403	0.205	0.027						
159	f26											1	0.134	0.137	0.078	0.152	0.001	0.089	0.425	0.435	0.580	0.034	0.211	0.108	0.204	0.075						
160	f27												1	0.055	0.014	0.068	0.004	0.026	0.218	0.098	0.244	0.001	0.004	0.018	0.001	0.032						
161	f28													1	0.413	0.415	0.095	0.117	0.398	0.166	0.265	0.013	0.123	0.020	0.051	0.026						
162	f29														1	0.450	0.239	0.114	0.162	0.112	0.060	0.152	0.084	0.097	0.017	0.007						
163	f31															1	0.283	0.010	0.359	0.185	0.197	0.011	0.033	0.001	0.030	0.004	0.017	0.000	0.010	0.018	0.247	0.082
164	f32																1	0.102	0.065	0.046	0.020	0.003	0.000	0.013	0.058	0.001	0.018	0.014	0.008	0.025	0.003	0.000
165	f33																	1	0.006	0.000	0.063	0.030	0.012	0.035	0.052	0.032	0.016	0.030	0.007	0.025	0.007	0.067
166	f34																		1	0.580	0.456	0.006	0.072	0.042	0.007	0.164	0.227	0.186	0.263	0.313	0.382	0.072
167	f35																			1	0.329	0.088	0.212	0.226	0.043	0.212	0.169	0.229	0.164	0.241	0.113	0.137
168	f36																				1	0.002	0.150	0.101	0.126	0.168	0.219	0.429	0.499	0.373	0.167	0.138
169	f37																					1	0.262	0.331	0.043	0.059	0.173	0.063	0.006	0.125	0.013	0.009
170	f38																						1	0.538	0.295	0.043	0.291	0.120	0.278	0.120	0.020	0.023
171	f40																							1	0.107	0.108	0.165	0.306	0.240	0.153	0.005	0.005
172	f41																								1	0.004	0.174	0.106	0.105	0.095	0.012	0.129
173																										1	0.231	0.465	0.267	0.407	0.022	0.293
174																											1	0.398	0.472	0.509	0.148	



Annexe 12 : Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur novice » (capture d'écran)



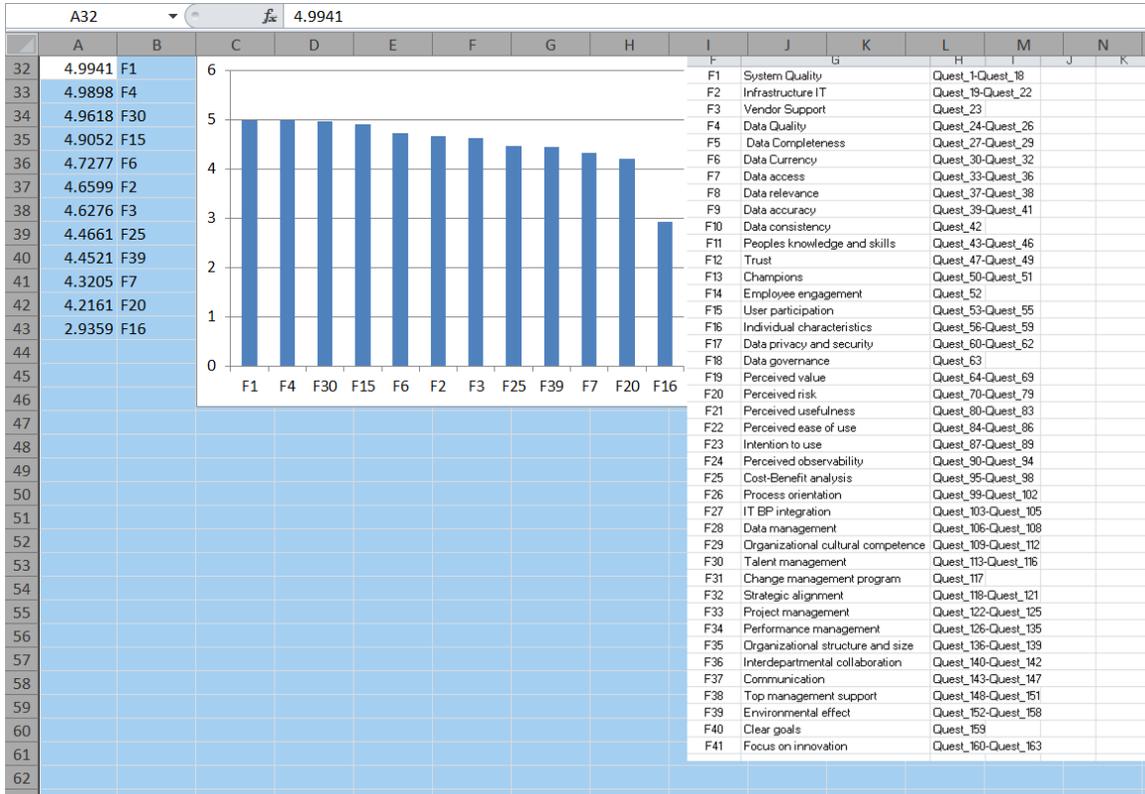
Annexe 13 : Liste des facteurs les plus importants qui contribuent à l'efficacité d'usage du big data chez un « utilisateur avancé » (capture d'écran)



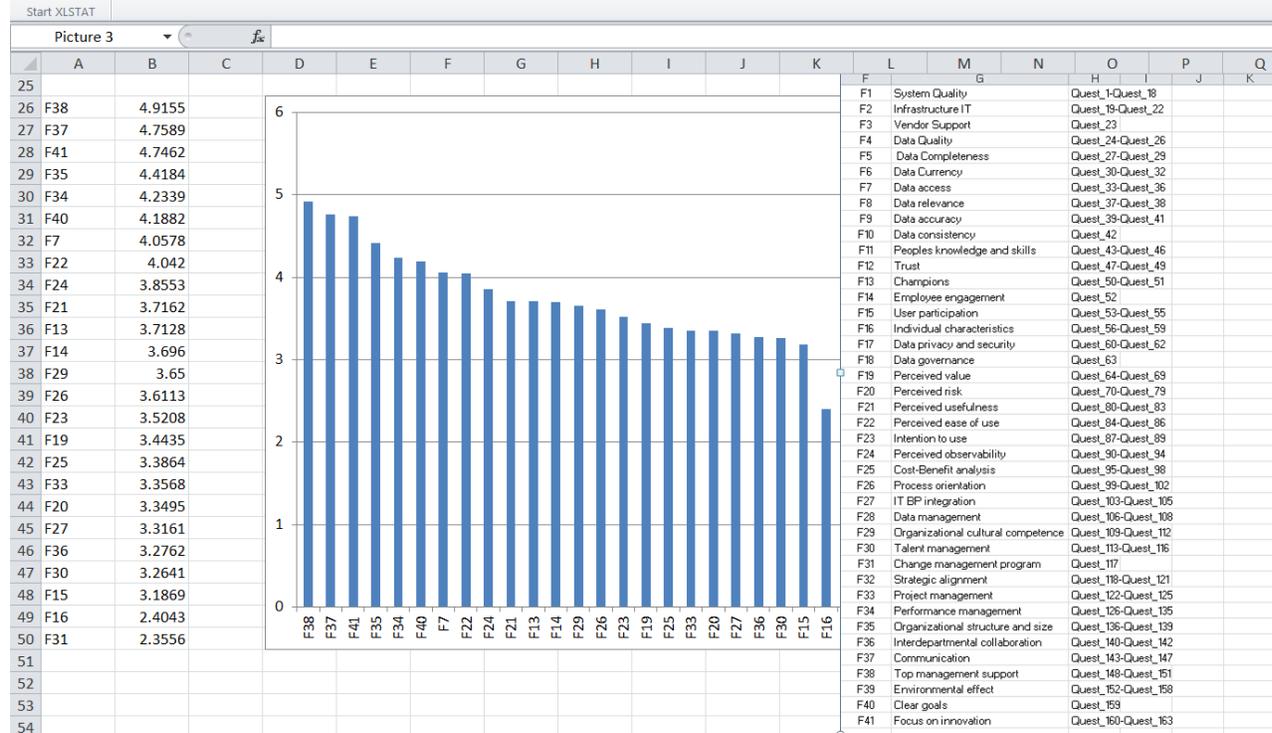
Annexe 14 : Codage des facteurs par questions (capture d'écran)

		F1			
C	D	E	F	G	H
0.832395			F1	System Quality	Quest_1-Quest_18
0.654646			F2	Infrastructure IT	Quest_19-Quest_22
0.500076			F3	Vendor Support	Quest_23
0.507315			F4	Data Quality	Quest_24-Quest_26
0.552696			F5	Data Completeness	Quest_27-Quest_29
0.499362			F6	Data Currency	Quest_30-Quest_32
0.535098			F7	Data access	Quest_33-Quest_36
0.554898			F8	Data relevance	Quest_37-Quest_38
0.675638			F9	Data accuracy	Quest_39-Quest_41
0.420453	0.607365	F1	F10	Data consistency	Quest_42
0.543562			F11	Peoples knowledge and skills	Quest_43-Quest_46
0.393669			F12	Trust	Quest_47-Quest_49
0.468223			F13	Champions	Quest_50-Quest_51
0.711766			F14	Employee engagement	Quest_52
0.478074			F15	User participation	Quest_53-Quest_55
0.887494			F16	Individual characteristics	Quest_56-Quest_59
0.888966			F17	Data privacy and security	Quest_60-Quest_62
0.828229			F18	Data governance	Quest_63
0.879991			F19	Perceived value	Quest_64-Quest_69
0.848252	0.85403	F2	F20	Perceived risk	Quest_70-Quest_79
0.891536			F21	Perceived usefulness	Quest_80-Quest_83
0.796343			F22	Perceived ease of use	Quest_84-Quest_86
0.539348	0.539348	F3	F23	Intention to use	Quest_87-Quest_89
0.636582			F24	Perceived observability	Quest_90-Quest_94
0.580458	0.554113	F4	F25	Cost-Benefit analysis	Quest_95-Quest_98
0.445299			F26	Process orientation	Quest_99-Quest_102
0.872904			F27	IT BP integration	Quest_103-Quest_105
0.764639	0.775517	F5	F28	Data management	Quest_106-Quest_108
0.689009			F29	Organizational cultural competence	Quest_109-Quest_112
0.643132			F30	Talent management	Quest_113-Quest_116
0.809254	0.668645	F6	F31	Change management program	Quest_117
0.553551			F32	Strategic alignment	Quest_118-Quest_121
0.453605			F33	Project management	Quest_122-Quest_125
0.618331	0.563204	F7	F34	Performance management	Quest_126-Quest_135
0.649276			F35	Organizational structure and size	Quest_136-Quest_139
0.531604			F36	Interdepartmental collaboration	Quest_140-Quest_143
0.57245			F37	Communication	Quest_143-Quest_147
0.624796	0.598623	F8	F38	Top management support	Quest_148-Quest_151
0.858075			F39	Environmental effect	Quest_152-Quest_158
0.777865	0.837484	F9	F40	Clear goals	Quest_159
0.876513			F41	Focus on innovation	Quest_160-Quest_163
0.521132	0.521132	F10			

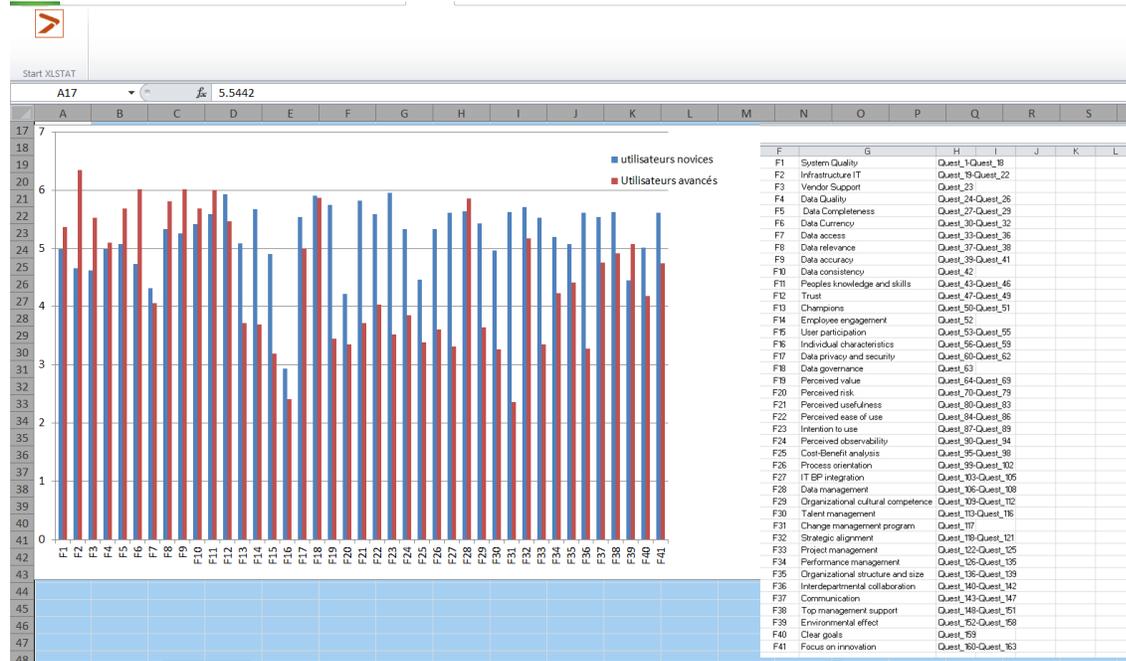
Annexe 15 : Le reste des facteurs qui contribuent à l'efficacité d'usage du big data chez un « utilisateur novice » (capture d'écran)



Annexe 16 : Le reste des facteurs qui contribuent à l'efficacité d'usage du big data chez un « utilisateur avancé » (capture d'écran)



Annexe 17 : Colonne comparatives de 41 facteurs pour les 2 groupes « utilisateurs avancés » et « utilisateurs novices »



Annexe 18: Guide d'entretien

A: Introduction

Would you tell me about your background and experience with your company?

B: Big data context –Results interpretation

Introduction to results/2 groups

Would you share your professional opinion on our main result?

If you find that this is a context specific result?

C: Differences interpretation

Would you share your professional opinion on the fact that the factors in these two groups have different classifications depending on their influence?

Would you please give an example for illustration?

D: Value creation in two groups

How the value creation process by using big data may differ in these two groups?

Annexe 19 : Guide d'entretien adapté de (Mikalef et al., 2019a)

Guide d'entretien

Զրույցի հարցերի շրջանակ

1. Votre formation et expérience en relation du BD

L'activité générale de l'organisation : Le rôle de l'usage du big data au sein de l'organisation

Ձեր մասնագիտական փորձն ու առնչությունը մեծածավալ տվյալների /big data/ հետ

Կազմակերպության գործունեությունը: Մեծածավալ տվյալների /big data/ օգտագործումը կազմակերպության կողմից

2. A : Si l'entreprise préalablement est fondée sur l'exploitation du big data / comment l'idée a été engendrée?

B : Si l'entreprise s'est fait l'idée d'exploiter du big data à un moment donné ; comment s'est fait cette décision ?

A : Եթե կազմակերպությունը հիմնված է նախապես big data-ի օգտագործման գաղափարի վրա, գաղափարի ստեղծման նախադրյալները/խնդիրները որոնք են?

B: Եթե կազմակերպությունը իր գործունեության որոշակի փուլում է ներդրել big data-ի օգտագործման տեխնոլոգիաներ և պայմաններ, ինչն է դրան դրդել և ինչ խնդիրների է առնչվել?

3. Un exemple bien précis de l'efficacité d'usage du big data : Utilisateur-avancé VS utilisateur-novice

/Quel objectif : Quelles données : Quelles big data technologies sont utilisées ; quelles techniques analytiques ; par exemple/ML, Recommendation system, sentiment analysis, time series analysis, SNA, AI simulations, .../ quelles technologies analytiques et de visualisation, par exemple/Excel, SAS, Stata, SPSS, R, Python, Mahout, Tableau : Si les outils sont utilisés par une voie locale/cloud : Comment cet usage crée une valeur pour l'entreprise : Quel type de prise de décision est affecté par cet usage ? Est-ce que les résultats de ces outils alimentent les rapports de gestion ? Si, oui qui est l'utilisateur final des rapports de ces sorties ? /

1 օրինակ մեծածավալ տվյալների արդյունավետ օգտագործման /effective use of big data/

« փորձառու օգտագործող» թե «սկսնակ օգտագործող», lifecycle

/Ինչ խնդիր կարող է դրված լինել: Ինչ տվյալներ պետք է օգտագործվեն: Ինչ big data տեխնոլոգիաներ /analytics techniques: e.g./ML, Recommendation system, sentiment analysis, time series analysis, SNA, AI simulations, .../ինչ desktop analytics and visualization տեխնոլոգիաներ e.g. /Excel, SAS, Stata, SPSS, R, Python, Mahout, Tableau/ technologies գործիքներ/մեթոդներ պետք է օգտագործվեն; օգտագործվում են լոկալ թե քլաուդ-ում: Ինչպես այն կարող է ստեղծել արժեք կազմակերպության համար: Ինչ տեսակի որոշումների վրա կարող է ազդել /օգտագործվել:/Արդյոք այդ վերջնարդյունքը ներառվում է որոշ հաշվետվություններում?Եթե այո, ով է դրա վերջնական սպառողը?

4. Quelles sont les contraintes réglementaires ou légales ? Est-ce que cela inhibe ce qui peut être fait avec le big data? Quelles sont les questions éthiques qui vous préoccupent (elles peuvent être légales mais pas quelque chose que vous voudriez faire). À l'avenir, quels autres obstacles à l'utilisation et à la création de valeur à partir du Big Data voyez-vous ? Par exemple, organisationnel, managérial, technique. Si elles sont spécifiques au contexte arménien ?

Ինչպիսի սահմանափակումներ կան /իրավական/էթիկ/այլ, որոնք խոչընդոտում են big data –ից առավել շատ արդյունքի ստացումը: Ինչ ուրիշ սահմանափակումներ կան ապագայի համար: Որոնք են առանձնահատուկ հայկական կոնտեքստի համար:

5. Résultats de l'étude quantitative / Քանակական հետազոտության նախնական արդյունքները⁶⁸

« Utilisateurs novices » ; 29 facteurs/41 les plus importants

Սկսնակ օգտագործողներ, 29 կարևոր ազդեցության գործոններ/41-ից

F23	1	Intention to use
F12	2	Trust
F18	3	Data governance
F21	4	Perceived usefulness
F19	5	Perceived value
F32	6	Strategic alignment
F14	7	Employee engagement
F28	8	Data management
F31	9	Change management program

⁶⁸ Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2019). Factors influencing effective use of big data: A research framework. Information and Management, 57(1). <https://doi.org/10.1016/j.im.2019.02.001>

F38	10	Top management support
F27	11	IT BP integration
F36	12	Interdepartmental collaboration
F41	13	Focus on innovation
F22	14	Perceived ease of use
F11	15	Peoples knowledge and skills
F17	16	Data privacy and security
F37	17	Communication
F33	18	Project management
F29	19	Organizational cultural competence
F10	20	Data consistency
F8	21	Data relevance
F26	22	Process orientation
F24	23	Perceived observability
F9	24	Data accuracy
F34	25	Performance management
F13	26	Champions
F35	27	Organizational structure and size
F5	28	Data Completeness
F40	29	Clear goals

Votre avis sur quelques relations corrélatives entre les facteurs :

Կան արդյոք տրամաբանական մեկնաբանություններ հետևյալ կորելացված կոմբինացիաներում

F8-F28; F14-F31; F28-F17...

« Utilisateurs avancés » ; 16 facteurs/41 les plus importants

«Փորձառու օգտագործողներ», 16 կարևոր ազդեցության գործոններ/41-ից

F2	1	Infrastructure IT
F6	2	Data Currency
F9	3	Data accuracy
		Peoples knowledge and
F11	4	skills
F18	5	Data governance
F28	6	Data management
F8	7	Data relevance
F5	8	Data Completeness
F10	9	Data consistency
F3	10	Vendor Support
F12	11	Trust
F1	12	System Quality
F32	13	Strategic alignment

F4	14	Data Quality
F39	15	Environmental effect
F17	16	Data privacy and security

Votre avis sur quelques relations corrélatives entre les facteurs :

Կան արդյոք տրամաբանական մեկնաբանություններ հետևյալ կորելացված կոմբինացիաներում

F9-F4 ; F10-F28 ; F32-F39 ; F18-F3 /- / ...

Annexe 20 : Régulation arménienne de la sécurité et l'utilisation des données personnelles

ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ

Օ Ր Ե Ն Ք Ը

Ընդունված է 2015 թվականի մայիսի 18-ին

ԱՆՁՆԱԿԱՆ ՏՎՅԱԼՆԵՐԻ ՊԱՇՏՊԱՆՈՒԹՅԱՆ ՄԱՍԻՆ

Գ Լ ՈՒ Խ 1

ԸՆԴՀԱՆՈՒՐ ԴՐՈՒՅԹՆԵՐ

Հոդված 1. Օրենքի կարգավորման առարկան

1. Սույն օրենքը կարգավորում է պետական կառավարման կամ տեղական ինքնակառավարման մարմինների, պետական կամ համայնքային հիմնարկների կամ կազմակերպությունների, իրավաբանական կամ ֆիզիկական անձանց կողմից անձնական տվյալները մշակելու, դրանց նկատմամբ պետական հսկողություն իրականացնելու կարգն ու պայմանները:

2. Պետական և ծառայողական, բանկային, նոտարական, փաստաբանական, ապահովագրական գաղտնիք համարվող, ազգային անվտանգությանը կամ պաշտպանությանն առնչվող գործողությունների ժամանակ օգտագործվող, ինչպես նաև փողերի լվացման և ահաբեկչության դեմ պայքարի, օպերատիվ-հետախուզական գործունեության ընթացքում կամ դատավարություններում օգտագործվող անձնական տվյալների հետ կապված առանձնահատկությունները կարգավորվում են այլ օրենքներով:

3. Սույն օրենքով սահմանված անձնական տվյալներ մշակելու սահմանափակումները չեն տարածվում բացառապես լրագրության, գրական և գեղարվեստական նպատակներով մշակվող անձնական տվյալների վրա:

4. Այլ օրենքներով կարող են սահմանվել անձնական տվյալներ մշակելու, հսկողություն իրականացնելու առանձնահատկություններ: Այլ օրենքներով հսկողություն իրականացնող մարմին սահմանված լինելու դեպքում լիազոր մարմինն իր լիազորություններն իրականացնում է սույն օրենքով սահմանված կարգով:

Հոդված 2. Անձնական տվյալների մասին Հայաստանի Հանրապետության օրենսդրությունը

1. Անձնական տվյալները մշակելու հետ կապված հարաբերությունները կարգավորվում են Հայաստանի Հանրապետության Սահմանադրությամբ, Հայաստանի Հանրապետության միջազգային պայմանագրերով, սույն օրենքով, այլ օրենքներով:

Հոդված 3. Օրենքի հիմնական հասկացությունները

1. Սույն օրենքում օգտագործվում են հետևյալ հիմնական հասկացությունները.

1) անձնական տվյալ` ֆիզիկական անձին վերաբերող ցանկացած տեղեկություն, որը թույլ է տալիս կամ կարող է թույլ տալ ուղղակի կամ անուղղակի կերպով նույնականացնել անձի ինքնությունը.

2) անձնական տվյալների մշակում` անկախ իրականացման ձևից և եղանակից (այդ թվում` ավտոմատացված, տեխնիկական ցանկացած միջոցներ կիրառելու կամ առանց դրանց) ցանկացած գործողություն կամ գործողությունների խումբ, որը կապված է անձնական տվյալները հավաքելու կամ ամրագրելու կամ մուտքագրելու կամ համակարգելու կամ կազմակերպելու կամ պահպանելու կամ օգտագործելու կամ վերափոխելու կամ վերականգնելու կամ փոխանցելու կամ ուղղելու կամ ուղեփակելու կամ ոչնչացնելու կամ այլ գործողություններ կատարելու հետ.

3) անձնական տվյալների փոխանցում երրորդ անձանց` անձնական տվյալները որոշակի կամ անորոշ շրջանակի այլ անձանց փոխանցելուն կամ դրանց հետ ծանոթացնելուն ուղղված գործողություն, այդ թվում` զանգվածային լրատվության միջոցներով անձնական տվյալները հրապարակելը, տեղեկատվական հաղորդակցման ցանցերում տեղադրելը կամ այլ եղանակով անձնական տվյալներն այլ անձի մատչելի դարձնելը.

4) անձնական տվյալների օգտագործում` անձնական տվյալների հետ կատարվող գործողություն, որի ուղղակի կամ անուղղակի նպատակը կարող է լինել որոշումներ ընդունելը կամ կարծիք ձևավորելը կամ իրավունքներ ձեռք բերելը կամ իրավունքներ կամ արտոնություններ տրամադրելը կամ իրավունքները սահմանափակելը կամ զրկելը կամ այլ նպատակի իրագործումը, որոնք տվյալների սուբյեկտի կամ երրորդ անձանց համար առաջացնում կամ կարող են առաջացնել իրավական հետևանքներ կամ այլ կերպ առնչվել նրանց իրավունքներին ու ազատություններին.

5) անձնական տվյալներ մշակող` պետական կառավարման կամ տեղական ինքնակառավարման մարմին, պետական կամ համայնքային հիմնարկ կամ կազմակերպություն, իրավաբանական կամ ֆիզիկական անձ, որը կազմակերպում և (կամ) իրականացնում է անձնական տվյալների մշակում.

6) տվյալների սուբյեկտ` ֆիզիկական անձ, որին վերաբերում են անձնական տվյալները.

7) տվյալների բազա` որոշակի հատկանիշներով համակարգված անձնական տվյալների ամբողջություն.

8) տեղեկատվական համակարգ` տվյալների բազայում ներառված անձնական տվյալների, դրանց էլեկտրոնային կամ ոչ էլեկտրոնային եղանակով մշակելու համար կիրառվող տեղեկատվական տեխնոլոգիաների կամ տեխնիկական միջոցների ամբողջություն.

9) անձնական տվյալների ապանձնավորում գործողություններ, որոնց արդյունքում հնարավոր չէ որոշել տվյալների պատկանելությունը կոնկրետ տվյալների սուբյեկտին.

10) անձնական տվյալների ուղեփակում անձնական տվյալները հավաքելու կամ ամրագրելու կամ համակարգելու կամ փոխանցելու կամ օգտագործելու հնարավորության ժամանակավոր կասեցում.

11) անձնական տվյալների ոչնչացում գործողություն, որի արդյունքում հնարավոր չէ վերականգնել տեղեկատվական համակարգում առկա անձնական տվյալների բովանդակությունը.

12) անձնական կյանքի տվյալներ՝ անձի անձնական կյանքի, ընտանեկան կյանքի, ֆիզիկական, ֆիզիոլոգիական, մտավոր, սոցիալական վիճակի վերաբերյալ կամ նման այլ տեղեկություններ.

13) կենսաչափական անձնական տվյալներ՝ անձի ֆիզիկական, ֆիզիոլոգիական և կենսաբանական առանձնահատկությունները բնութագրող տեղեկություններ.

14) հատուկ կատեգորիայի անձնական տվյալներ՝ անձի ռասայական, ազգային պատկանելությանը կամ էթնիկ ծագմանը, քաղաքական հայացքներին, կրոնական կամ փիլիսոփայական համոզմունքներին, արհեստակցական միությանն անդամակցությանը, առողջական վիճակին ու սեռական կյանքին վերաբերող տեղեկություններ.

15) հանրամատչելի անձնական տվյալներ՝ տեղեկություններ, որոնք տվյալների սուբյեկտի համաձայնությամբ կամ իր անձնական տվյալները հանրամատչելի դարձնելուն ուղղված գիտակցված գործողությունների կատարմամբ մատչելի են դառնում որոշակի կամ անորոշ շրջանակի անձանց համար, ինչպես նաև այն տեղեկությունները, որոնք օրենքով նախատեսված են որպես հանրամատչելի տեղեկություններ.

16) լիազորված անձ՝ իրավաբանական կամ ֆիզիկական անձ, պետական կառավարման կամ տեղական ինքնակառավարման մարմին, պետական կամ համայնքային հիմնարկ կամ կազմակերպություն, որին տվյալների մշակողի կողմից օրենքով սահմանված դեպքերում կամ պայմանագրի հիման վրա պատվիրվել է հավաքել, մուտքագրել, համակարգել կամ այլ կերպ մշակել անձնական տվյալներ.

17) երրորդ անձ՝ ցանկացած անձ, մարմին, հիմնարկ կամ կազմակերպություն, որը չի հանդիսանում տվյալների սուբյեկտ, անձնական տվյալների մշակող կամ լիազորված անձ, և որի իրավունքները կամ օրինական շահերը շոշափվում կամ կարող են շոշափվել անձնական տվյալները մշակելու արդյունքում:

Գ Լ ՈՒ Խ 2

ԱՆՁՆԱԿԱՆ ՏՎՅԱԼՆԵՐ ՄՇԱԿԵԼՈՒ ՀԻՄՆԱԿԱՆ ՍԿՁԲՈՒՆՔՆԵՐԸ

Հոդված 4. Օրինականության սկզբունքը

1. Անձնական տվյալներ մշակողը պարտավոր է հետևել և ապահովել, որ տվյալները մշակվեն օրենքի պահանջների պահպանմամբ:
2. Անձնական տվյալները մշակվում են օրինական և որոշակի նպատակներով և առանց տվյալների սուբյեկտի համաձայնության չեն կարող օգտագործվել այլ նպատակներով:

Հոդված 5. Համաչափության սկզբունքը

1. Տվյալների մշակումը պետք է հետապնդի օրինական նպատակ, դրան հասնելու միջոցները պետք է լինեն պիտանի, անհրաժեշտ և չափավոր:
2. Անձնական տվյալներ մշակողը պարտավոր է անձնական տվյալները մշակել այն նվազագույն քանակով, որն անհրաժեշտ է օրինական նպատակներին հասնելու համար:
3. Արգելվում է այնպիսի անձնական տվյալների մշակումը, որոնք անհրաժեշտ չեն տվյալները մշակելու նպատակի համար կամ անհամատեղելի են դրա հետ:
4. Արգելվում է անձնական տվյալների մշակումը, եթե տվյալները մշակելու նպատակին հնարավոր է հասնել ապանձնավորված կերպով:
5. Անձնական տվյալները պետք է պահպանվեն այնպես, որ բացառվի տվյալների սուբյեկտի հետ դրանց նույնականացումն ավելի երկար ժամկետով, քան անհրաժեշտ է դրանց նախօրոք որոշված նպատակներին հասնելու համար:

Հոդված 6. Հավաստիության սկզբունքը

1. Մշակվող անձնական տվյալը պետք է լինի ամբողջական, ճշգրիտ, պարզ և հնարավորինս թարմացված:

Հոդված 7. Սուբյեկտների նվազագույն ներգրավման սկզբունքը

1. Անձնական տվյալների մշակումն իրականացվում է սուբյեկտների նվազագույն ներգրավման սկզբունքով:
2. Այն դեպքում, երբ պետական կառավարման կամ տեղական ինքնակառավարման մարմինը, նոտարը միասնական էլեկտրոնային տեղեկատվական համակարգի միջոցով կարող են անձնական տվյալը ձեռք բերել այլ մարմնից, ապա անձնական տվյալների սուբյեկտից չի պահանջվում ներկայացնել որոշակի գործողությունների համար անհրաժեշտ անձնական տվյալը:

3. Անձնական տվյալների սուբյեկտի գրավոր համաձայնության դեպքում անձնական տվյալներ մշակող համարվող ֆիզիկական կամ իրավաբանական անձինք կարող են պետական կամ տեղական ինքնակառավարման մարմնից ստանալ որոշակի գործողության համար անհրաժեշտ և անձնական տվյալների սուբյեկտի գրավոր համաձայնության մեջ ուղղակիորեն մատնանշված անձնական տվյալը:

4. Էլեկտրոնային տեղեկատվական համակարգի միջոցով անձնական տվյալների փոխանցման կարգը սահմանում է Հայաստանի Հանրապետության կառավարությունը:

Գ Լ ՈՒ Խ 3

ԱՆՁՆԱԿԱՆ ՏՎՅԱԼՆԵՐ ՄՇԱԿԵԼՈՒ

Հոդված 8. Անձնական տվյալները մշակելու օրինականությունը

1. Անձնական տվյալներ մշակելը օրինական է, եթե՝

1) տվյալները մշակվել են օրենքի պահանջների պահպանմամբ, և տվյալների սուբյեկտը տվել է իր համաձայնությունը, բացառությամբ սույն օրենքով կամ այլ օրենքներով ուղղակիորեն նախատեսված դեպքերի, կամ

2) մշակվող տվյալները ձեռք են բերվել անձնական տվյալների հանրամատչելի աղբյուրներից:

Հոդված 9. Տվյալների սուբյեկտի համաձայնությունը

1. Տվյալների սուբյեկտը կարող է տալ իր համաձայնությունը անձամբ կամ ներկայացուցչի միջոցով, եթե լիազորագրով հատուկ նախատեսված է նման լիազորություն:

2. Համաձայնության հիման վրա մշակվող տվյալները պահպանվում են այն ժամկետով, որն օբյեկտիվորեն անհրաժեշտ է տվյալներ մշակելու նպատակներն իրականացնելու համար, կամ համաձայնությամբ սահմանված ժամկետով:

3. Տվյալների սուբյեկտն իրավունք ունի հետ կանչելու իր համաձայնությունը սույն օրենքով, ինչպես նաև այլ օրենքներով նախատեսված դեպքերում և կարգով:

4. Տվյալների սուբյեկտի համաձայնությունը համարվում է տրված, և մշակողն այն մշակելու իրավունք ունի, երբ՝

1) մշակողին հասցեագրած և տվյալների սուբյեկտի ստորագրած փաստաթղթում նշված են անձնական տվյալները, բացառությամբ այն դեպքերի, երբ փաստաթուղթն իր բովանդակությամբ հանդիսանում է անձնական տվյալները մշակելու դեմ առարկություն.

2) մշակողը տվյալները ստացել է տվյալների սուբյեկտի հետ կնքված պայմանագրի հիման վրա և օգտագործում է այդ պայմանագրով սահմանված գործողությունների նպատակով.

3) տվյալների սուբյեկտն իր կամքով, օգտագործելու նպատակով բանավոր փոխանցում է մշակողին իր անձնական տվյալների մասին տեղեկությունները:

5. Առանց տվյալների սուբյեկտի համաձայնության անձնական տվյալներ կարող են մշակվել, եթե տվյալներ մշակելն ուղղակիորեն նախատեսված է օրենքով:

6. Անձնական տվյալներ մշակողը կամ լիազորված անձը տվյալների սուբյեկտի գրավոր համաձայնությունն ստանալու նպատակով տվյալների սուբյեկտին ծանուցում է տվյալներ մշակելու մտադրության մասին:

7. Տվյալների սուբյեկտի համաձայնությունը տրվում է գրավոր կամ էլեկտրոնային եղանակով՝ հաստատված էլեկտրոնային թվային ստորագրությամբ, բանավոր համաձայնության դեպքում՝ այնպիսի հավաստի գործողությունների միջոցով, որոնք ակնհայտորեն կվկայեն տվյալների սուբյեկտի՝ անձնական տվյալները օգտագործելու համաձայնության մասին:

8. Տվյալների սուբյեկտի համաձայնությունը ստանալու փաստն ապացուցելու, իսկ հանրամատչելի անձնական տվյալներ մշակելու դեպքում դրանց հանրամատչելի լինելու փաստն ապացուցելու պարտականությունը կրում է մշակողը:

9. Տվյալների սուբյեկտի անգործունակության կամ սահմանափակ գործունակության կամ մինչև 16 տարեկան անչափահաս լինելու դեպքում նրա անձնական տվյալները մշակելու համար համաձայնություն է տալիս տվյալների սուբյեկտի օրինական ներկայացուցիչը:

10. Տվյալների սուբյեկտի մահվան կամ դատարանի վճռով մահացած ճանաչվելու դեպքում նրա անձնական տվյալները մշակելու համար համաձայնություն են տալիս տվյալների սուբյեկտի բոլոր ըստ օրենքի ժառանգները, ժառանգներ չունենալու դեպքում՝ ժառանգության բացման վայրի համայնքի ղեկավարը, իսկ անհայտ բացակայող ճանաչվելու դեպքում՝ անհայտ բացակայող ճանաչված անձի գույքի հավատարմագրային կառավարիչը, եթե մինչ այդ տվյալների սուբյեկտը նման համաձայնություն չի տվել:

11. Տվյալների սուբյեկտի մահվան դեպքում նրա անձնական տվյալները կարող են մշակվել առանց համաձայնության, եթե մշակվող տվյալները պետք է լինեն մահացած անձի անունը, սեռը, ծննդյան և մահվան տարին, ամիսը, ամսաթիվը: Մշակույթի, արվեստի, գիտության, կրթության, սպորտի, կրոնի և հանրային այլ բնագավառների գործչի մահվան դեպքում նրա անձնական կյանքի տվյալները կարող են մշակվել առանց համաձայնության, եթե մահվան օրվանից անցել է հիսուն տարի:

Հոդված 10. Անձնական տվյալներ մշակելու համաձայնությունն ստանալու նպատակով տվյալների սուբյեկտին ներկայացվող ծանուցումը

1. Անձնական տվյալներ մշակողը կամ սույն օրենքի 14-րդ հոդվածով նախատեսված լիազորված անձը տվյալների սուբյեկտի համաձայնությունն ստանալու նպատակով ծանուցում է տվյալներ մշակելու մտադրության մասին:

2. Ծանուցման մեջ նշվում են՝

1) տվյալների սուբյեկտի ազգանունը, անունը, հայրանունը.

2) անձնական տվյալների մշակման իրավական հիմքերը և նպատակը.

3) մշակման ենթակա անձնական տվյալների ցանկը.

4) անձնական տվյալների հետ կատարման ենթակա գործողությունների ցանկը, որոնց համար հայցվում է տվյալների սուբյեկտի համաձայնությունը.

5) այն անձանց շրջանակը, որոնց կարող են փոխանցվել անձնական տվյալները.

6) անձնական տվյալների սուբյեկտի համաձայնությունը հայցող մշակողի կամ նրա ներկայացուցչի անվանումը (ազգանունը, անունը, հայրանունը, պաշտոնը) և գտնվելու կամ հաշվառման (փաստացի բնակության) վայրը.

7) տվյալների սուբյեկտի կողմից անձնական տվյալների ուղղում, ոչնչացում, տվյալների մշակման դադարեցում պահանջելու կամ մշակման հետ կապված այլ գործողություն կատարելու վերաբերյալ տեղեկություններ.

8) հայցվող համաձայնության գործողության ժամկետը, ինչպես նաև համաձայնությունը հետ կանչելու կարգը և դրա հետևանքները:

Հոդված 11. Հանրամատչելի անձնական տվյալները

1. Տվյալների սուբյեկտի համաձայնությամբ կամ օրենքով նախատեսված դեպքերում կարող է սահմանվել անձնական տվյալների հանրամատչելի տեղեկատվության ռեժիմ (հեռախոսային տեղեկատուներ, հասցեների գրքեր, կենսագրական տեղեկատուներ, մասնավոր հայտարարություններ, եկամուտների հայտարարագրում և այլն): Հանրամատչելի են համարվում անձի անունը, ազգանունը, ծննդյան տարին, ամիսը և ամսաթիվը, վայրը, մահվան վայրը, տարին, ամիսը և ամսաթիվը, ինչպես նաև այն անձնական տվյալները, որոնք տվյալների սուբյեկտի կողմից հանրամատչելի դարձնելուն ուղղված գիտակցված գործողությունների կատարմամբ մատչելի են դառնում որոշակի կամ անորոշ շրջանակի անձանց համար:

2. Տվյալների սուբյեկտի վերաբերյալ տեղեկությունները, բացառությամբ սույն հոդվածի 1-ին մասով նախատեսված տեղեկությունների, կարող են հանվել անձնական տվյալների հանրամատչելի աղբյուրներից տվյալների սուբյեկտի պահանջով կամ դատական կարգով:

3. Պայմանագրի հիման վրա մշակվող տվյալները կարող են հանվել անձնական տվյալների հանրամատչելի աղբյուրներից փոխադարձ համաձայնությամբ կամ դատական կարգով:

Հոդված 12. Հատուկ կատեգորիայի անձնական տվյալներ մշակելու առանձնահատկությունները

1. Առանց անձի համաձայնության արգելվում է հատուկ կատեգորիայի անձնական տվյալներ մշակելը, բացառությամբ, երբ տվյալի մշակումն ուղղակիորեն նախատեսված է օրենքով:
2. Սույն հոդվածի 1-ին մասով նախատեսված անձնական տվյալների մշակումն անհապաղ դադարեցվում է, եթե վերացել են տվյալները մշակելու հիմքերը և նպատակը:

Հոդված 13. Կենսաչափական անձնական տվյալներ մշակելու առանձնահատկությունները

1. Կենսաչափական անձնական տվյալներ մշակվում են միայն տվյալների սուբյեկտի համաձայնությամբ, բացառությամբ օրենքով նախատեսված դեպքերի, և եթե օրենքով հետապնդվող նպատակը հնարավոր է իրականացնել միայն այդ կենսաչափական տվյալները մշակելու միջոցով:

Հոդված 14. Տվյալներ մշակողի առաջադրանքով անձնական տվյալների մշակումը լիազորված անձի կողմից

1. Անձնական տվյալները կարող է մշակել նաև տվյալները մշակողի առաջադրանքով լիազորված անձը:
2. Առաջադրանքը տրվում է գրավոր, որի մեջ պետք է շարադրված լինեն անձնական տվյալներ մշակելու իրավական հիմքերը և պայմանները, նպատակը, մշակման ենթակա անձնական տվյալների ցանկը, տվյալների սուբյեկտների շրջանակը, այն անձանց շրջանակը, որոնց կարող են փոխանցվել անձնական տվյալները, անձնական տվյալների պաշտպանության տեխնիկական և կազմակերպչական միջոցառումները և անհրաժեշտ այլ տեղեկություններ:
3. Անձնական տվյալները մշակվում են միայն առաջադրանքի շրջանակներում: Առաջադրանքի շրջանակներում անձնական տվյալների մշակման համար պատասխանատու է տվյալներ մշակողը: Եթե առաջադրանքը չի համապատասխանում օրենքի պահանջներին, ապա լիազորված անձը պետք է այդ մասին գրավոր տեղեկացնի տվյալներ մշակողին և հրաժարվի մշակումից:
4. Անձնական տվյալները պետական կառավարման կամ տեղական ինքնակառավարման մարմինների, պետական կամ համայնքային հիմնարկների կամ կազմակերպությունների առաջադրանքով մշակվում են սույն օրենքի պահանջների պահպանմամբ:

5. Այլ օրենքներով կամ տվյալներ մշակողի և լիազորված անձի միջև կնքված պայմանագրերով կարող են սահմանվել լիազորված անձի կողմից անձնական տվյալներ մշակելու առանձնահատկություններ, որոնք չեն կարող վերաբերել այլ անձանց իրավունքներին և պարտականություններին:

Գ Լ ՈՒ Խ 4

ՏՎՅԱԼՆԵՐԻ ՍՈՒԲՅԵԿՏԻ ԻՐԱՎՈՒՆՔՆԵՐԸ

Հոդված 15. Տվյալների սուբյեկտի՝ իր անձնական տվյալների վերաբերյալ տեղեկություններ ստանալու իրավունքը

1. Տվյալների սուբյեկտն իրավունք ունի ստանալու տեղեկություններ իր անձնական տվյալների, տվյալները մշակելու, մշակելու հիմքերի և նպատակների, տվյալները մշակողի, նրա գտնվելու վայրի մասին, ինչպես նաև այն անձանց շրջանակի մասին, որոնց կարող են փոխանցվել անձնական տվյալները:

2. Տվյալների սուբյեկտն իրավունք ունի ծանոթանալու իր անձնական տվյալներին, մշակողից պահանջելու ուղղել, ուղեփակել կամ ոչնչացնել իր անձնական տվյալները, եթե անձնական տվյալներն ամբողջական կամ ճշգրիտ չեն կամ հնացած են կամ ձեռք են բերվել անօրինական ճանապարհով կամ անհրաժեշտ չեն մշակելու նպատակներին հասնելու համար:

3. Անձնական տվյալները մշակողի կողմից ուղղված, ուղեփակված կամ ոչնչացված լինելու վերաբերյալ կասկածների դեպքում տվյալների սուբյեկտն իրավունք ունի դիմելու անձնական տվյալների պաշտպանության լիազոր մարմին՝ իր անձնական տվյալներն ուղղված, ուղեփակված կամ ոչնչացված լինելու փաստը պարզելու և իրեն տեղեկություններ տրամադրելու մասին պահանջով:

4. Մշակողի կողմից տվյալների սուբյեկտին անձնական տվյալների վերաբերյալ տեղեկությունները պետք է տրամադրվեն հասանելի ձևով և չպետք է պարունակեն տվյալների այլ սուբյեկտի վերաբերող անձնական տեղեկություններ:

5. Տվյալների սուբյեկտին անձնական տվյալները տրամադրվում են տվյալների սուբյեկտի կամ լիազորագրով հանդես եկող ներկայացուցչի կամ օրինական ներկայացուցչի գրավոր հարցման հիման վրա: Հարցումը կարող է ներկայացվել էլեկտրոնային եղանակով՝ հաստատված էլեկտրոնային թվային ստորագրությամբ:

6. Տվյալների սուբյեկտն իրավունք ունի տեղեկատվություն ստանալու իր անձնական տվյալների մշակման վերաբերյալ, այդ թվում՝

1) անձնական տվյալները մշակելու փաստը հաստատելու և մշակելու նպատակի վերաբերյալ:

- 2) անձնական տվյալները մշակելու եղանակների վերաբերյալ.
- 3) այն սուբյեկտների վերաբերյալ, որոնց տրամադրվել կամ կարող են տրամադրվել անձնական տվյալները.
- 4) մշակվող անձնական տվյալների ցանկի և դրանք ձեռք բերելու աղբյուրի վերաբերյալ.
- 5) անձնական տվյալները մշակելու ժամկետների վերաբերյալ.
- 6) անձնական տվյալները մշակելու հետևանքով տվյալների սուբյեկտի համար առաջացող հնարավոր իրավական հետևանքների վերաբերյալ:
7. Տեղեկությունները տվյալների սուբյեկտին տրամադրվում են անվճար, եթե օրենքով այլ բան նախատեսված չէ:

Հոդված 16. Տվյալների սուբյեկտի իրավունքները անձնական տվյալները մշակելու հիման վրա որոշումներ ընդունելիս

1. Արգելվում է անձնական տվյալներ մշակելու նպատակներից չբխող այնպիսի որոշումներ ընդունել, որոնք տվյալների սուբյեկտի համար առաջացնում են իրավական հետևանքներ կամ այլ կերպ առնչվում են նրա իրավունքներին ու օրինական շահերին, բացառությամբ սույն հոդվածի 2-րդ մասով նախատեսված դեպքերի:
2. Անձնական տվյալները մշակելու հիման վրա տվյալների սուբյեկտի համար իրավական հետևանքներ առաջացնող կամ այլ կերպ նրա իրավունքներին ու օրինական շահերին առնչվող որոշումներն ընդունվում են տվյալների սուբյեկտի համաձայնությամբ կամ օրենքով նախատեսված դեպքերում:

Հոդված 17. Մշակողի գործողությունները կամ անգործությունը բողոքարկելու իրավունքը

1. Եթե տվյալների սուբյեկտը համարում է, որ իր անձնական տվյալների մշակումն իրականացվում է սույն օրենքի պահանջների խախտմամբ կամ այլ կերպ խախտում է իր իրավունքներն ու ազատությունները, ապա նա իրավունք ունի մշակողի գործողությունները կամ անգործությունը կամ որոշումները բողոքարկելու անձնական տվյալների պաշտպանության լիազոր մարմին կամ դատական կարգով:
2. Տվյալների սուբյեկտն ունի օրենքով սահմանված կարգով վնասի հատուցման իրավունք:

Գ Լ ՈՒ Խ 5

ԱՆՁՆԱԿԱՆ ՏՎՅԱԼՆԵՐ ՄՇԱԿՈՂԻ ՊԱՐՏԱԿԱՆՈՒԹՅՈՒՆՆԵՐԸ

Հոդված 18. Անձնական տվյալներ մշակողի պարտականություններն անձնական տվյալները հավաքելու ընթացքում

1. Անձնական տվյալները մշակելու ընթացքում մշակողը պարտավոր է տվյալների սուբյեկտի պահանջով նրան տրամադրել սույն օրենքի 15-րդ հոդվածով նախատեսված տեղեկատվությունը:

2. Ոչ ամբողջական, ոչ ճշգրիտ, հնացած, անօրինական ճանապարհով ձեռք բերված կամ մշակելու նպատակներին հասելու համար ոչ անհրաժեշտ անձնական տվյալների դեպքում անձնական տվյալներ մշակողը պարտավոր է անհրաժեշտ գործողություններ իրականացնել դրանք ամբողջացնելու, թարմացնելու, ուղղելու կամ ոչնչացնելու ուղղությամբ:

3. Մշակողը պարտավոր է տվյալների սուբյեկտին գրավոր պարզաբանել անձնական տվյալները չտրամադրելու հետևանքները, այդ թվում՝ անձնական տվյալների սուբյեկտի իրավունքները:

4. Եթե անձնական տվյալներն ստացվել են ոչ տվյալների սուբյեկտից, բացառությամբ օրենքով նախատեսված դեպքերի, ինչպես նաև հանրամատչելի անձնական տվյալների, ապա մշակողը մինչև այդպիսի անձնական տվյալները մշակելը պարտավոր է տվյալների սուբյեկտին տրամադրել հետևյալ տեղեկատվությունը.

1) մշակողի կամ նրա լիազորած անձի (առկայության դեպքում) անվանումը (ազգանունը, անունը, հայրանունը) և գտնվելու կամ հաշվառման (փաստացի բնակության) վայրը.

2) անձնական տվյալները մշակելու նպատակը և իրավական հիմքը, մշակվող տվյալների ցանկը.

3) անձնական տվյալների հավանական օգտագործողների շրջանակը.

4) տվյալների սուբյեկտի՝ սույն օրենքով սահմանված իրավունքները:

Հոդված 19. Անձնական տվյալներ մշակելու անվտանգության ապահովման միջոցները և մշակողի պարտականությունները

1. Մշակողը պարտավոր է ոչնչացնել կամ ուղեփակել անձնական այն տվյալները, որոնք անհրաժեշտ չեն օրինական նպատակին հասնելու համար:

2. Անձնական տվյալները մշակելու ընթացքում մշակողը պարտավոր է օգտագործել գաղտնագրման միջոցներ՝ անձնական տվյալներ պարունակող տեղեկատվական

համակարգերի պաշտպանվածությունը պատահական կորստից, տեղեկատվական համակարգեր անօրինական մուտք գործելուց, անձնական տվյալների անօրինական օգտագործումից, ձայնագրումից, ոչնչացումից, վերափոխումից, ուղեփակումից, կրկնօրինակումից, տարածումից և այլ միջամտությունից ապահովելու համար:

3. Մշակողը պարտավոր է կանխել անձնական տվյալների մշակման համապատասխան տեխնոլոգիաների մատչելիությունը դրա իրավունքը չունեցող անձանց համար և ապահովել, որ այդ համակարգերի օրինական օգտագործողի համար հասանելի լինեն միայն իր կողմից մշակման ենթակա տվյալները և այն տվյալները, որոնցից թույլատրված է օգտվել:

4. Տեղեկատվական համակարգերում անձնական տվյալները մշակելու անվտանգությունն ապահովելուն ներկայացվող պահանջները, կենսաչափական անձնական տվյալների նյութական կրիչներին և տեղեկատվական համակարգերից դուրս այդ անձնական տվյալները պահպանելու տեխնոլոգիաներին ներկայացվող պահանջները սահմանվում են Հայաստանի Հանրապետության կառավարության որոշմամբ:

5. Օրենքով հսկողություն իրականացնող այլ մարմին սահմանված լինելու դեպքում այդ մարմինը օրենքով իրեն վերապահված լիազորությունների շրջանակներում կարող է սահմանել սույն հոդվածի 4-րդ մասով նախատեսված` Հայաստանի Հանրապետության կառավարության որոշմամբ սահմանված պահանջներից ավելի բարձր պահանջներ:

6. Տեղեկատվական համակարգերից դուրս կենսաչափական անձնական տվյալների օգտագործումն ու պահպանումը կարող են իրականացվել միայն այնպիսի նյութական կրիչների միջոցով, տեխնոլոգիաների կիրառմամբ կամ ձևերով, որոնք ապահովում են այդ տվյալների պաշտպանվածությունը դրանց անօրինական մուտք գործելուց, անձնական տվյալների անօրինական օգտագործումից, ոչնչացումից, վերափոխումից, ուղեփակումից, կրկնօրինակումից, տարածումից և այլն:

7. Անձնական տվյալներ մշակողները կամ սույն օրենքով նախատեսված այլ անձինք պարտավոր են պահպանել անձնական տվյալների գաղտնիությունը ինչպես անձնական տվյալները մշակելու հետ առնչվող ծառայողական կամ աշխատանքային պարտականությունները կատարելու ընթացքում, այնպես էլ դրա ավարտից հետո:

8. Սույն հոդվածի պահանջների կատարման նկատմամբ հսկողությունն իրականացնում է անձնական տվյալների պաշտպանության լիազոր մարմինը` առանց տեղեկատվական համակարգերում մշակվող անձնական տվյալները մշակելու իրավունքի:

9. Անձնական տվյալներ մշակող իրավաբանական անձինք իրենց տիրապետման տակ գտնվող անձնական տվյալներ մշակող էլեկտրոնային համակարգերը բավարար պաշտպանության մակարդակ ունեցող ճանաչելու և ռեեստրում ընդգրկելու նպատակով կարող են դիմել անձնական տվյալների պաշտպանության լիազոր մարմին:

Հոդված 20. Անձնական տվյալներ մշակողի պարտականությունները տվյալների սուբյեկտի կամ լիազոր մարմնի գրավոր հարցման, անձնական տվյալների ծանոթացման, մշակողի կամ լիազորված անձի կողմից խախտումներ հայտնաբերելու դեպքերում

1. Մշակողը պարտավոր է սույն օրենքի 15-րդ հոդվածով սահմանված կարգով տվյալների սուբյեկտին և լիազոր մարմնին տեղեկատվություն տրամադրել տվյալների սուբյեկտի վերաբերյալ անձնական տվյալների առկայության մասին կամ հնարավորություն ընձեռել ծանոթանալու դրանց գրավոր հարցումը ստանալուց հետո՝ հինգ օրվա ընթացքում:

2. Մշակողը պարտավոր է տվյալների սուբյեկտին հնարավորություն ընձեռել անվճար ծանոթանալու տվյալների սուբյեկտին վերաբերող անձնական տվյալներին: Եթե տվյալների սուբյեկտի անձնական տվյալներն ամբողջական կամ ճշգրիտ չեն կամ հնացած են կամ ձեռք են բերվել անօրինական ճանապարհով կամ անհրաժեշտ չեն մշակելու նպատակներին հասնելու համար, ապա մշակողի կամ լիազորված անձի կողմից դրանք հայտնաբերվելու կամ տվյալների սուբյեկտի կամ օրինական ներկայացուցչի (կամ լիազորված անձի) կողմից դիմում ստանալուց հետո մշակողը պարտավոր է անհապաղ կամ նման հնարավորության բացակայության դեպքում երեք աշխատանքային օրվա ընթացքում անհրաժեշտ գործողություններ իրականացնել դրանք ամբողջացնելու, թարմացնելու, ուղղելու, ուղեփակելու կամ ոչնչացնելու ուղղությամբ:

3. Մշակողը պարտավոր է անձնական տվյալների պաշտպանության լիազոր մարմնի գրավոր հարցման հիման վրա նրա գործունեության իրականացման համար անհրաժեշտ տեղեկատվությունը տրամադրել հարցումն ստանալու օրվանից հինգ օրվա ընթացքում:

4. Տվյալների սուբյեկտի գրավոր պահանջի հիման վրա տվյալների սուբյեկտի անձնական տվյալները տրամադրելը, ուղղելը, ուղեփակելը կամ ոչնչացնելը մերժելու դեպքում մշակողը պարտավոր է տվյալների սուբյեկտին և լիազոր մարմնին հարցումն ստանալու օրվանից հինգ օրվա ընթացքում տրամադրել պատճառաբանված գրավոր որոշում՝ հղում կատարելով այն օրենքի դրույթներին, որոնք հիմք են հանդիսացել որոշում ընդունելու համար:

5. Անձնական տվյալները տրամադրելը, ուղղելը, ուղեփակելը կամ ոչնչացնելը մերժելու հիմքերը լիազոր մարմնի կողմից հիմնավորված չհամարվելու դեպքում մշակողը պարտավոր է անհապաղ տրամադրել, ուղղել, ուղեփակել կամ ոչնչացնել տվյալների սուբյեկտի անձնական տվյալները կամ լիազոր մարմնի որոշումը բողոքարկել դատական կարգով:

Հոդված 21. Մշակողի պարտականություններն անձնական տվյալներ մշակելու ընթացքում թույլ տված օրենսդրության խախտումները վերացնելիս, անձնական տվյալներն ուղղելիս, ուղեփակելիս կամ ոչնչացնելիս

1. Տվյալների սուբյեկտի կամ անձնական տվյալների պաշտպանության լիազոր մարմնի հարցման հիման վրա անձնական տվյալների հավաստիությունը կամ մշակելու օրինականությունը վիճարկվելու դեպքում մշակողը պարտավոր է ուղեփակել տվյալների սուբյեկտին վերաբերող անձնական տվյալները՝ սկսած հարցումն ստանալու պահից մինչև ստուգման գործողությունների ավարտը:

2. Անձնական տվյալների ոչ ճշգրիտ լինելը հաստատվելու դեպքում մշակողը պարտավոր է տվյալների սուբյեկտի կամ անձնական տվյալների պաշտպանության լիազոր մարմնի

ներկայացրած փաստաթղթերի կամ այլ անհրաժեշտ փաստաթղթերի հիման վրա ուղղել անձնական տվյալները և վերացնել դրանց ուղեփակումը:

3. Անձնական տվյալների հետ իրականացվող անօրինական գործողություններ հայտնաբերելու դեպքում մշակողը պարտավոր է անհապաղ, բայց ոչ ուշ, քան երեք աշխատանքային օրվա ընթացքում վերացնել թույլ տված խախտումները: Խախտումները վերացնելու անհնարինության դեպքում մշակողը պարտավոր է անհապաղ ոչնչացնել անձնական տվյալները: Խախտումները վերացնելու կամ անձնական տվյալները ոչնչացնելու մասին մշակողը պարտավոր է երեք աշխատանքային օրվա ընթացքում տեղեկացնել տվյալների սուբյեկտին կամ նրա ներկայացուցչին, իսկ այն դեպքում, երբ հարցումն ստացվել է անձնական տվյալների պաշտպանության լիազոր մարմնից՝ նաև այդ մարմնին:

4. Էլեկտրոնային համակարգերից անձնական տվյալների արտահոսքի դեպքում մշակողը պարտավոր է այդ մասին անհապաղ հրապարակել հայտարարություն՝ միաժամանակ արտահոսքի վերաբերյալ հայտնելով Հայաստանի Հանրապետության ոստիկանությանը և անձնական տվյալների պաշտպանության լիազոր մարմնին:

5. Անձնական տվյալները մշակելու նպատակին հասնելու դեպքում մշակողը պարտավոր է անհապաղ դադարեցնել տվյալներ մշակելը, եթե այլ բան նախատեսված չէ օրենքով:

6. Տվյալների սուբյեկտի համաձայնությունը գրավոր՝ հաստատված ստորագրությամբ կամ Էլեկտրոնային եղանակով՝ հաստատված Էլեկտրոնային թվային ստորագրությամբ, հետ կանչվելու դեպքում մշակողը պարտավոր է դադարեցնել անձնական տվյալներ մշակելը և ոչնչացնել տվյալները հետկանչն ստանալու օրվան հաջորդող տասն աշխատանքային օրվա ընթացքում, եթե այլ բան նախատեսված չէ տվյալների սուբյեկտի և մշակողի փոխադարձ համաձայնությամբ կամ օրենքով: Մշակողը պարտավոր է ոչնչացնելու պահից երեք աշխատանքային օրվա ընթացքում տվյալների սուբյեկտին տեղեկացնել անձնական տվյալները ոչնչացնելու մասին:

Հոդված 22. Լիազոր մարմնի լիազորությունների իրականացման կարգը հսկողություն և իրականացնող այլ մարմնի միջոցով

1. Եթե օրենքով սահմանված է հսկողություն իրականացնող այլ մարմին, ապա հսկողություն իրականացնող այդ մարմինը լիազոր մարմնին տրամադրում է իրավաբանական անձանց անձնական տվյալներ մշակող Էլեկտրոնային համակարգերը բավարար պաշտպանության մակարդակ ունեցող ճանաչելու վերաբերյալ եզրակացություն:

2. Եթե օրենքով սահմանված է հսկողություն իրականացնող այլ մարմին, ապա լիազոր մարմնին անձնական տվյալների պաշտպանության մասին դիմումները, ինչպես նաև անձնական տվյալների պաշտպանության վերաբերյալ տեղեկությունները փոխանցում է հսկողություն իրականացնող այդ մարմնին:

3. Հսկողություն իրականացնող այլ մարմինն անձնական տվյալների պաշտպանության ոլորտում իր կայացրած որոշումները կամ անձնական տվյալների պաշտպանության ուղղությամբ կատարված գործողությունների մասին տեղեկությունները սեղմ ժամկետում ուղարկում է լիազոր մարմին:

4. Հսկողություն իրականացնող այլ մարմնի որոշումները, գործողություններն ու անգործությունը լիազոր մարմինը կարող է բողոքարկել դատական կարգով:

Հոդված 23. Անձնական տվյալները մշակելու վերաբերյալ լիազոր մարմինն ծանուցելը

1. Մշակողը մինչև անձնական տվյալները մշակելը իրավունք ունի անձնական տվյալների պաշտպանության լիազոր մարմինն ծանուցելու տվյալներ մշակելու մտադրության մասին:

2. Լիազոր մարմնի պահանջով մշակողը պարտավոր է ծանուցում ուղարկել լիազոր մարմին:

3. Մշակողը մինչև կենսաչափական կամ հատուկ կատեգորիայի անձնական տվյալներ մշակելը պարտավոր է անձնական տվյալների պաշտպանության լիազոր մարմինն ծանուցել տվյալներ մշակելու մտադրության մասին:

4. Ծանուցման մեջ նշվում են հետևյալ տեղեկությունները.

1) մշակողի կամ լիազորված անձի (առկայության դեպքում) անվանումը (ազգանունը, անունը, հայրանունը), գտնվելու կամ հաշվառման վայրը.

2) անձնական տվյալներ մշակելու նպատակը և իրավական հիմքերը.

3) անձնական տվյալների շրջանակը.

4) տվյալների սուբյեկտների շրջանակը.

5) անձնական տվյալների հետ կատարվող գործողությունների ցանկը, մշակողի կողմից անձնական տվյալների մշակման եղանակների ընդհանուր նկարագիրը.

6) այն միջոցների նկարագիրը, որոնք մշակողը պարտավորվում է իրականացնել անձնական տվյալները մշակելու անվտանգության ապահովման ուղղությամբ.

7) անձնական տվյալների մշակումն սկսելու ամսաթիվը.

8) անձնական տվյալների մշակումն ավարտելու ժամկետները և պայմանները:

5. Անձնական տվյալների պաշտպանության լիազոր մարմինը ծանուցումն ստանալուց հետո՝ 30 օրվա ընթացքում, սույն հոդվածի 2-րդ մասով նախատեսված տեղեկությունները, ինչպես նաև տվյալ ծանուցումն ուղարկելու ամսաթվի վերաբերյալ տեղեկությունները մուտքագրում է մշակողների ռեեստր:

6. Մշակողի վրա չեն կարող դրվել անձնական տվյալների պաշտպանության լիազոր մարմնի կողմից անձնական տվյալներ մշակելու վերաբերյալ ծանուցման քննարկման, ինչպես նաև տեղեկությունները մշակողների ռեեստր մուտքագրելու հետ կապված ծախսերը:

7. Այն դեպքում, երբ մշակողի կողմից ներկայացված՝ սույն հոդվածի 2-րդ մասով նախատեսված տեղեկությունները լրիվ կամ ճշգրիտ չեն, ապա անձնական տվյալների պաշտպանության լիազոր մարմինն իրավունք ունի մշակողից պահանջելու հստակեցնել ներկայացված տեղեկությունները մինչև դրանք մշակողների ռեեստր մուտքագրելը:

8. Սույն հոդվածի 2-րդ մասով նախատեսված տեղեկությունների փոփոխման դեպքում մշակողը պարտավոր է փոփոխությունների մասին ծանուցել անձնական տվյալների պաշտպանության լիազոր մարմին՝ փոփոխություններն ի հայտ գալու օրվանից տասն աշխատանքային օրվա ընթացքում:

Գ Լ ՈՒ Խ 6

ԱՆՁՆԱԿԱՆ ՏՎՅԱԼՆԵՐ ՄՇԱԿԵԼՈՒ ՕՐԻՆԱԿԱՆՈՒԹՅԱՆ ՆԿԱՏՄԱՄԲ ՀՍԿՈՂՈՒԹՅՈՒՆԸ

Հոդված 24. Անձնական տվյալների պաշտպանության լիազոր մարմինը

1. Անձնական տվյալների պաշտպանությունն իրականացնում է լիազոր մարմինը, որը գործում է Հայաստանի Հանրապետության կառավարության որոշմամբ սահմանված կառուցվածքով:

2. Անձնական տվյալների պաշտպանության լիազոր մարմինը գործում է անկախ՝ օրենքի և այլ իրավական ակտերի հիման վրա:

3. Անձնական տվյալների պաշտպանության լիազոր մարմինը՝

1) ստուգում է իր նախաձեռնությամբ կամ համապատասխան դիմումի հիման վրա անձնական տվյալների մշակման համապատասխանությունը սույն օրենքի պահանջներին.

2) սույն օրենքի պահանջների խախտման դեպքում կիրառում է օրենքով սահմանված վարչական պատասխանատվության միջոցներ.

3) պահանջում է արգելափակել, կասեցնել կամ դադարեցնել սույն օրենքի պահանջները խախտող անձնական տվյալների մշակումը.

4) օրենքով նախատեսված հիմքերի առկայության դեպքում մշակողից պահանջում է անձնական տվյալների ուղղում, փոփոխում, ուղեփակում կամ ոչնչացում.

5) անձնական տվյալներ մշակելու վերաբերյալ մշակողի ծանուցման ուսումնասիրության արդյունքում ամբողջությամբ կամ մասամբ արգելում է անձնական տվյալների մշակումը:

6) վարում է անձնական տվյալներ մշակողների ռեեստր:

7) իրավաբանական անձանց անձնական տվյալներ մշակող էլեկտրոնային համակարգերը ճանաչում է բավարար պաշտպանության մակարդակ ունեցող և դրանք ներառում է ռեեստրում:

8) ստուգում է տվյալներ մշակելու համար օգտագործվող սարքերը և փաստաթղթերը, այդ թվում առկա տվյալները և համակարգչային ծրագրերը:

9) օրենքով նախատեսված դեպքերում դիմում է դատարան:

10) իրականացնում է օրենքով սահմանված այլ լիազորություններ:

11) պահպանում է իր գործունեության ընթացքում իրեն վստահված կամ հայտնի դարձած անձնական տվյալների գաղտնիությունը:

12) ապահովում է տվյալների սուբյեկտի իրավունքների պաշտպանությունը:

13) քննում է անձնական տվյալների մշակմանը վերաբերող հարցերով ֆիզիկական անձանց դիմումները և իր լիազորությունների սահմաններում ընդունում որոշումներ:

14) տարեկան մեկ անգամ ներկայացնում է հրապարակային հաշվետվություն՝ անձնական տվյալների պաշտպանության բնագավառում առկա իրավիճակի և նախորդ տարվա գործունեության վերաբերյալ:

15) կատարում է հետազոտություններ և մշակողների դիմումների կամ լուսաբանումների հիման վրա տալիս տվյալներ մշակելու վերաբերյալ խորհրդատվություն կամ տեղեկացնում է անձնական տվյալներ մշակելու վերաբերյալ լավագույն փորձի մասին:

16) իրավապահ մարմիններին հաղորդում է ներկայացնում իր գործունեության ընթացքում քրեաիրավական բնույթի խախտումների վերաբերյալ կասկածներ ի հայտ գալու դեպքում:

4. Անձնական տվյալների պաշտպանության լիազոր մարմնի որոշումները կարող են բողոքարկվել դատական կարգով:

5. Անձնական տվյալների պաշտպանության լիազոր մարմնի գործունեությունը ֆինանսավորվում է պետական բյուջեի միջոցների հաշվին՝ առանձին տողով:

6. Անձնական տվյալների պաշտպանության լիազոր մարմնին կից կարող է հասարակական հիմունքներով գործել խորհրդատվական մարմին, որի ձևավորման ու գործունեության կարգը սահմանվում է անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարի հրամանով:

Հոդված 25. Անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարի նշանակումը, լիազորությունների դադարումը և նրան ներկայացվող պահանջները

1. Անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարը նշանակվում է հինգ տարի ժամկետով Հայաստանի Հանրապետության վարչապետի կողմից՝ Հայաստանի Հանրապետության արդարադատության նախարարի ներկայացմամբ՝ իրավապաշտպան գործունեություն իրականացնող առնվազն հինգ հասարակական կազմակերպությունների համատեղ առաջարկությունների հիման վրա: Հայաստանի Հանրապետության արդարադատության նախարարի կողմից Հայաստանի Հանրապետության վարչապետին ներկայացվող լիազոր մարմնի ղեկավարի թեկնածուն պետք է լինի հասարակական կազմակերպությունների առաջարկած թեկնածությունների ցանկից:

2. Հասարակական կազմակերպությունների կողմից թեկնածություններ առաջարկելու կարգը սահմանում է Հայաստանի Հանրապետության կառավարությունը:

3. Մինևույն անձը չի կարող երկու անգամ անընդմեջ նշանակվել անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարի պաշտոնում:

4. Անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարը ղեկավարում է անձնական տվյալների պաշտպանության լիազոր մարմնի գործունեությունը և պատասխանատու է անձնական տվյալների պաշտպանության լիազոր մարմնի լիազորությունների իրականացման համար:

5. Անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարն ունի օրենքով և այլ իրավական ակտերով սահմանված իրավունքներ և պարտականություններ:

6. Անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարը՝

1) պետք է ունենա բարձրագույն կրթություն, բարձր հեղինակություն և առնվազն հինգ տարվա կառավարման ոլորտի աշխատանքային փորձ.

2) պետք է զերծ մնա իր անկախությունը և անաչառությունը կասկածի տակ դնող գործողություններից:

7. Անձնական տվյալների պաշտպանության լիազորված մարմնի ղեկավարը պաշտոնից ազատվում է հետևյալ հիմքերի առկայության դեպքում՝

1) գրավոր դիմումի հիման վրա.

2) լրացել է նրա 65 տարին (պաշտոնավարման տարիքը), կամ լրացել է պաշտոնավարման ժամկետը.

3) ընտրվել կամ նշանակվել է այլ պաշտոնում կամ անցել է անձնական տվյալների պաշտպանության լիազոր մարմնի ղեկավարի պաշտոնի հետ անհամատեղելի այլ աշխատանքի.

4) ժամանակավոր անաշխատունակության հետևանքով ավելի քան 120 օր անընդմեջ կամ վերջին 12 ամսվա ընթացքում ավելի քան 140 օր ծառայության չներկայանալու դեպքում՝ չհաշված հղիության և ծննդաբերության կամ երեխային ինամելու համար արձակուրդը.

5) ավելի քան հինգ օր անընդմեջ առանց հարգելի պատճառի չի ներկայացել աշխատանքի.

6) դատարանի՝ օրինական ուժի մեջ մտած վճռի հիման վրա ճանաչվել է անգործունակ, սահմանափակ գործունակ, անհայտ բացակայող կամ մահացած.

7) նրա նկատմամբ դատարանի մեղադրական դատավճիռն օրինական ուժի մեջ մտնելու դեպքում:

Գ Լ ՈՒ Խ 7

ԱՆՁՆԱԿԱՆ ՏԿՅԱԼՆԵՐԸ ԵՐՐՈՐԴ ԱՆՁԱՆՑ ԵՎ ԱՅԼ ՊԵՏՈՒԹՅՈՒՆՆԵՐ ՓՈԽԱՆՑԵԸ

Հոդված 26. Անձնական տվյալները երրորդ անձանց փոխանցելը

1. Առանց անձնական տվյալների սուբյեկտի համաձայնության մշակողը կարող է անձնական տվյալները փոխանցել երրորդ անձանց կամ տվյալներից օգտվելու հնարավորություն տրամադրել, եթե դա նախատեսված է օրենքով և ունի բավարար պաշտպանության մակարդակ:

2. Առանց անձնական տվյալների սուբյեկտի համաձայնության մշակողը կարող է հատուկ կատեգորիայի անձնական տվյալներ փոխանցել երրորդ անձանց կամ տվյալներից օգտվելու հնարավորություն տրամադրել, եթե՝

1) տվյալներ մշակողը հանդիսանում է օրենքով կամ միջպետական պայմանագրով սահմանված հատուկ կատեգորիայի անձնական տվյալներ մշակող, այդ տեղեկության փոխանցումը ուղղակիորեն նախատեսված է օրենքով և ունի բավարար պաշտպանության մակարդակ.

2) օրենքով նախատեսված բացառիկ դեպքերում հատուկ կատեգորիայի անձնական տվյալները կարող են փոխանցվել տվյալների սուբյեկտի կյանքի, առողջության կամ ազատության պաշտպանության համար:

Հոդված 27. Անձնական տվյալներն այլ պետություններ փոխանցելը

1. Անձնական տվյալները կարող են այլ երկիր փոխանցվել տվյալների սուբյեկտի համաձայնությամբ, կամ եթե տվյալների փոխանցումը բխում է անձնական տվյալների մշակման նպատակներից և (կամ) անհրաժեշտ է այդ նպատակների իրագործման համար:

2. Առանց լիազոր մարմնի թույլտվության անձնական տվյալները կարող են փոխանցվել այլ պետություն, եթե այդ պետությունում ապահովված է անձնական տվյալների պաշտպանության բավարար մակարդակ: Անձնական տվյալների պաշտպանության բավարար մակարդակը համարվում է ապահովված, եթե՝

1) անձնական տվյալները փոխանցվում են միջազգային պայմանագրերին համապատասխան.

2) անձնական տվյալները փոխանցվում են լիազոր մարմնի կողմից պաշտոնական հրապարակված ցուցակում ընդգրկված որևէ երկիր:

3. Անձնական տվյալները կարող են փոխանցվել բավարար պաշտպանության մակարդակ չապահովող պետության տարածք միայն լիազոր մարմնի թույլտվությամբ, եթե անձնական տվյալները փոխանցվում են պայմանագրի հիման վրա, և պայմանագրով նախատեսված են անձնական տվյալների պաշտպանության այնպիսի երաշխիքներ, որոնք լիազոր մարմնի կողմից հաստատվել են որպես բավարար պաշտպանություն ապահովող:

4. Սույն հոդվածի 3-րդ մասում նշված դեպքերում անձնական տվյալներ մշակողը պարտավոր է նախքան այլ երկիր տվյալներ փոխանցելը գրավոր դիմել լիազոր մարմին՝ թույլտվություն ստանալու հայտով: Հայտում անձնական տվյալներ մշակողը պարտավոր է նշել այն երկիրը, որտեղ փոխանցվում են անձնական տվյալները, անձնական տվյալներն ստացող սուբյեկտի նկարագրությունը (անվանումը, իրավակազմակերպչական ձևը), անհատական տվյալների նկարագրությունը (պարունակությունը), անձնական տվյալների մշակման, անձնական տվյալները փոխանցելու նպատակը և պայմանագիրը կամ դրա նախագիծը: Լիազոր մարմինը 30 օրվա ընթացքում պարտավոր է թույլատրել կամ մերժել հայտը: Լիազոր մարմինը կարող է անձնական տվյալներ մշակողից պահանջել լրացուցիչ տեղեկություններ՝ պահպանելով հայտը դիտարկելու ժամկետը: Այն դեպքում, երբ լիազոր մարմինը կգտնի, որ պայմանագրային երաշխիքները բավարար չեն, պարտավոր է նշել այն անհրաժեշտ փոփոխությունները, որոնք կապահովեն անձնական տվյալների պաշտպանության երաշխիքներ:

5. Անձնական տվյալների պաշտպանության լիազոր մարմինը պարբերաբար, սակայն ոչ պակաս, քան տարին մեկ անգամ, պարտավոր է վերանայել անձնական տվյալների պաշտպանության բավարար մակարդակն ապահովող երկրների ցուցակը և հրապարակել փոփոխությունները պաշտոնական տեղեկագրում և իր պաշտոնական ինտերնետային կայքում:

6. Պետական մարմինների տնօրինության տակ գտնվող անձնական տվյալները կարող են փոխանցվել օտարերկրյա պետական մարմիններին միայն միջպետական պայմանագրերի շրջանակներում, իսկ ոչ պետական մարմիններին՝ սույն հոդվածի նորմերին համապատասխան:

Գ Լ ՈՒ Խ 8

ԵԶՐԱՓՈՎԿԻՉ ՄԱՍ ԵՎ ԱՆՑՈՒՄԱՅԻՆ ԴՐՈՒՅԹՆԵՐ

Հողված 28. Եզրափակիչ մաս

1. Սույն օրենքն ուժի մեջ է մտնում 2015 թվականի հուլիսի 1-ից:
2. Սույն օրենքն ուժի մեջ մտնելու պահից ուժը կորցրած ճանաչել «Անհատական տվյալների մասին» Հայաստանի Հանրապետության 2002 թվականի հոկտեմբերի 8-ի ՀՕ-422-Ն օրենքը:
3. Սույն օրենքի 7-րդ հոդվածն ուժի մեջ է մտնում 2019 թվականի հունվարի 1-ից:

Հողված 29. Անցումային դրույթներ

1. Սույն օրենքն ուժի մեջ մտնելուց հետո մինչև սույն օրենքն ուժի մեջ մտնելը մշակվող անձնական տվյալների մշակումը շարունակվում է իրականացվել սույն օրենքով սահմանված կարգով:
2. Այն մշակողները, որոնք անձնական տվյալներ մշակել են մինչև սույն օրենքն ուժի մեջ մտնելը և շարունակում են անձնական տվյալներ մշակել սույն օրենքն ուժի մեջ մտնելուց հետո, պարտավոր են սույն օրենքով նախատեսված պարտադիր ծանուցումն ուղարկել անձնական տվյալների պաշտպանության լիազոր մարմին մինչև 2015 թվականի սեպտեմբերի 1-ը:

Հայաստանի Հանրապետության

Նախագահ Ս. Սարգսյան

2015 թ. հունիսի 13

Երևան

ՀՕ-49-Ն

Annexe 21 : Article présenté lors de 25^{ème} conférence de l'AIM 2020



VIANDOX Taxi et Big Data : un cas d'efficacité d'usage

*Elen Grigoryan **

*Katia Lobre-Lebraty ***

* IAE, Université Jean Moulin Lyon 3, France

** IAE, Université Jean Moulin Lyon 3, France

Résumé : Cette communication s'intéresse aux aspects managériaux du phénomène de Big Data. Nous proposons de traiter la question suivante : comment évaluer l'impact du Big Data en termes d'efficacité d'usage ? Nous appliquons le modèle de l'efficacité d'usage du Big Data, proposé par Surbakti et al. (2019), au cas de VIANDOX Taxi qui utilise les technologies d'analyse des données massives dans son activité quotidienne. A partir de 4 entretiens avec le directeur régional de cette entreprise, nous avons pu valider une première partie du modèle. L'analyse thématique a en effet permis d'identifier la majorité des facteurs contenus dans les 7 thèmes du modèle.

Mots clés : Big Data, Analyse des données, Efficacité d'usage, Etude de cas, Evaluation.

Introduction

Les entreprises disposent désormais d'un volume de données qu'elles n'arrivent pas à traiter (Zhou et al., 2016). S'intéresser à l'ère de cette explosion des données, où leur masse dépasse souvent la capacité technologique de l'entreprise à les stocker, traiter, analyser et visualiser, consiste à s'intéresser au phénomène du « Big Data » ((Mikalef et al., 2019b); (Vassakis et al., 2018); (Sumbal, Tsui, & See-to, 2017); (Gupta & George, 2016); (Karoui et al., 2014)). Après le client au cœur de l'activité, depuis quelques années, ce sont l'ensemble des données qui tendent à être positionner à cette place centrale au sein des organisations. Ainsi, les principaux acteurs des différentes industrie ont les yeux rivés sur les potentialités que la technologie peut leur offrir pour acquérir, stocker et analyser les données (Marr, 2016).

Le Big Data (noté BD dans la suite de ce document) est un concept ancien mais dont l'usage demeure un phénomène émergent et non normé dans les entreprises. Il existe d'ailleurs plusieurs approches du BD (Markus & Topi, 2015). Pour comprendre les fondements du BD ou des données massives, on peut toutefois s'appuyer sur les relations qui s'établissent entre « Données – Informations – Connaissances » : les données sont des faits représentant une certaine réalité, elles sont indépendantes de celui que les utilise (Kocabyk & Erduran, 2000); les informations sont des données auxquelles un individu ou un groupe a ajouté du sens ; et le passage de la notion d'information à celle de connaissance s'opère au travers d'un mécanisme de cognition sur les informations (Lebraty, 2011). Dès lors, on comprend bien que les données seules ne représentent aucune valeur pour l'entreprise si elles ne sont pas analysées. On comprend aussi la nécessité des outils et procédures analytiques et technologiques afin de les traiter (Bharadwaj, El Sawy, Pavlou, & Venkatraman, 2013). Il y a donc un processus complet qui consiste à capturer, stocker, transférer, analyser et visualiser une énorme masse de données structurées et non-structurées (Erevelles et al., 2016). Du point de vue de l'utilisateur en entreprise, seules les deux dernières étapes sont visibles (analyse et visualisation). Aussi, le concept de Big Data Analytique se confond avec le terme plus générique de Big Data.

L'intérêt porté au BD est actuellement très grand, tant au niveau académique que managérial. Dans la littérature l'accent est surtout mis sur les aspects technologiques du phénomène BD, et les cas empiriques d'analyse de l'usage et des effets du BD restent limités. Bien entendu, les entreprises essaient de tirer de la valeur des gros volumes de données dont elles disposent, mais comment procèdent-elles concrètement ? Et avec quels résultats ? Les connaissances des entreprises sont assez vagues sur le sujet, aussi la question d'une exploitation efficace du potentiel des avancées technologiques et des gros volumes de données reste à explorer.

Dans ce papier, nous aborderons la problématique de l'efficacité d'usage du BD, ce qui nous conduira à répondre à la question de recherche suivante : « **comment évaluer l'impact du Big Data en termes d'efficacité d'usage ?** ». En effet, le terme « d'efficacité d'usage du BD » est une traduction de l'expression anglo-saxonne « effective use of BD » qui renvoie à l'efficacité d'utilisation du BD, autrement dit la capacité d'une organisation à améliorer l'atteinte de ses objectifs grâce au BD.

Nous travaillerons à partir de l'étude du cas de l'entreprise VIANDOX qui propose un service de taxi fondé sur la maraude et une application de géolocalisation. Plus précisément, nous testerons empiriquement le modèle de l'efficacité d'usage du BD développé par Surbakti et al. (2019) dans le cas du service de taxi de VIANDOX. Ce modèle est focalisé sur les facteurs influençant l'efficacité d'usage du BD. Ces facteurs sont regroupés selon leur nature : motivationnelle, opérationnelle et liée à différents mécanismes de support.

Du point de vue académique, cette recherche a pour l'objectif de comprendre et de développer la notion « d'efficacité d'usage » dans un contexte particulier, celui du BD. Du point de vue managérial, il s'agit pour nous de contribuer à l'établissement d'un guide pour les entreprises, leur permettant de bien comprendre les facteurs potentiels qui favorisent une utilisation efficace de leur BD. Autrement dit une utilisation du BD qui contribue à améliorer l'atteinte de leurs objectifs. Pour cela nous testons empiriquement le modèle proposé par Surbakti et al. (2019), ce que ses auteurs n'ont pu faire et qui constitue la principale limite de leur travail. Le modèle comporte 41 facteurs influençant l'efficacité d'usage du BD dans l'entreprise. Ces facteurs sont consolidés dans les 7 thèmes suivants : « qualité des données » ; « confidentialité, sécurité et gouvernance des données » ; « l'intérêt organisationnel perçu » ; « management des processus » ; « aspect humain » ; « aspect organisationnel » et « systèmes, outils et technologies ». Selon ses auteurs et à notre connaissance, il s'agit de la première tentative d'élaboration d'un modèle complet des facteurs qui influencent l'efficacité d'usage du BD, et du premier test empirique de ce modèle.

Ce papier est construit comme suit : dans la première partie, un point sera fait sur les fondements théoriques de l'efficacité d'usage en SI et dans le contexte de BD, puis la méthodologie de recherche utilisée sera exposée, le terrain mobilisé sera présenté ainsi que le modèle à tester. Enfin les premiers résultats seront proposés ainsi que les conclusions préliminaires qui peuvent en être tirées.

Théories et modèles mobilisés sur l'efficacité d'usage

L'efficacité d'usage d'un système d'information

Pour Burton-Jones et Grange (2013), l'efficacité d'usage d'un SI (Système d'Information) est un usage du SI qui contribue à l'amélioration de l'atteinte des objectifs

de l'organisation. Ces auteurs mettent alors en avant une théorie de l'efficacité d'usage (TEU) dans le SI. Le concept présenté est fondé sur les facteurs issus de la théorie de la représentation (RT) (Wand & Weber, 1990, 1995). La RT est une théorie générale appliquée à un SI. Elle constate qu'un SI est composé de plusieurs structures qui servent à représenter une partie du monde de manière compréhensible pour les usagers et pour les autres parties prenantes. Comme le souligne Orlikowski (Orlikowski, 2000; p. 425), « *la technologie en soi ne peut pas augmenter ou diminuer la productivité des salariés, seul l'usage peut le faire* ». De plus, « *simplement dire qu'un plus grand usage procurera plus d'avantages sans tenir compte de la nature de cet usage est bien insuffisant* », (DeLone & McLean, 2003; p. 16). Enfin l'usage d'un système n'est pas suffisant pour en obtenir des avantages (P. B. Seddon, 1997), c'est pourquoi l'usage doit être efficace (Burton-Jones & Grange, 2013). A première vue, cela peut sembler évident, mais en réalité ce concept est plus complexe à aborder qu'il n'y paraît, et encore insuffisamment étudié. Cette lacune est soulignée depuis longtemps par Marcolin (2000; p. 53) qui préconisent le passage « de l'étude de l'usage, à l'étude de l'efficacité d'usage ».

Les 3 éléments vecteurs de l'efficacité d'usage sont : les compétences et les motivations des usagers (1) ; l'intentionnalité du système (2) ; les caractéristiques des tâches (3). Les usagers ont la possibilité d'agir dans le sens de l'amélioration de l'efficacité d'usage par le biais de l'adaptation et de l'apprentissage (Burton-Jones & Straub, 2006). Remarquons que ces vecteurs sont intimement liés aux facteurs catégorisés dans les thèmes « organisationnel » et « mécanismes de support » du modèle de l'efficacité d'usage du BD de Surbakti et al. (2019).

Le TAM (Technology Acceptance Model), initialement proposé par Fred Davis en 1985 (Legris et al., 2003) est un modèle bien connu qui porte sur l'usage des TI (F. D. Davis et al., 1989). Adapté de la théorie socio-psychologique/comportemental, ce modèle a été développé par de nombreux chercheurs. Etant donné le gap entre, d'un côté les énormes investissements réalisés pour l'implantation du SI au sein des organisations, et d'un autre côté des taux de réussite relativement faibles, les chercheurs ont mobilisé leurs efforts pour identifier les facteurs qui facilitent l'acceptation d'une technologie par ses utilisateurs. Le TAM met en avant deux facteurs : facilité perçue d'utilisation et utilité perçue, comme les antécédents de l'efficacité d'un SI. Ces facteurs influencent le comportement des utilisateurs et déterminent leur acceptation de la technologie.

Un modèle dérivé, l'UTAUT (Unified Theory of Acceptance and Use of Technology), se propose d'unifier les huit modèles déclinant le TAM. Selon ce modèle (Venkatesh et al., 2003), les quatre principaux déterminants de l'acceptation par l'utilisateur et le comportement dans l'usage, qui sont considérés comme les déterminants directs sont : l'espérance de performance, l'espérance d'effort, l'influence sociale et les conditions facilitatrices. En dehors des déterminants directs, le modèle inclut également des éléments modérateurs clés tels que : le sexe, l'âge, le volontarisme de l'utilisateur et

l'expérience. L'espérance de la performance traduit le degré de croyance d'un individu dans le système, c'est-à-dire la capacité du système à soutenir son travail, à améliorer ses performances dans son travail. L'espérance d'effort souligne le degré de facilité d'usage du système. Ces deux déterminants rejoignent bien les facteurs motivationnels de l'efficacité d'usage du BD du modèle de Surbakti et al. (2019) ; (1) l'efficacité perçue et (2) la facilité de l'usage perçue. Les deux autres déterminants directs sont : l'influence sociale qui représente le degré d'importance accordé par l'utilisateur à la perception des tiers concernant le système, et les conditions facilitatrices qui traduisent la perception de l'utilisateur concernant l'infrastructure organisationnelle et technique destinée à soutenir l'usage du système. Ces déterminants sont présents dans la catégorie de mécanismes de support du modèle de Surbakti et al. (2019), thèmes : « Aspects humains » (F5) et « Systèmes, outils et technologies » (F7). Les modérateurs de l'UTAUT sont inclus dans le thème « Aspects humain » facteur « caractéristiques humaines » (F5.6) du modèle à tester.

On peut constater ainsi que la Théorie de l'Efficacité d'usage (TEU) est une évolution naturelle des modèles TAM et UTAUT liée au passage nécessaire « de l'étude de l'usage à l'étude de l'efficacité d'usage » (Marcolin et al., 2000).

L'efficacité d'usage dans le contexte du big data

Les modèles et théories présentés ci-dessus ont été développés sur la base d'un SI traditionnel qui n'intègre pas les caractéristiques spécifiques aux données massives dites Vs : 3Vs (Volume, Vitesse, Variété) - (Philip Chen & Zhang, 2014) ; 4Vs (Volume, Vitesse, Variété, Vérité) - (Sivarajah, Kamal, Irani, & Weerakkody, 2017; De Mauro, Greco, & Grimaldi, 2016; Abbasi, Sarker, & Chiang, 2016); 5Vs (Volume, Vitesse, Variété, Vérité, Valeur) - (Sharda et al., 2014) et 7Vs (Volume, Vitesse, Variété, Vérité, Valeur, Variabilité et Visualisation) - (J. J. J. M. Seddon & Currie, 2017).

Dans le contexte spécifique du BD, Mérimo et al, (2016) ont tenté de développer la Théorie de l'efficacité d'usage en proposant un modèle permettant d'évaluer les niveaux de qualité du BD dans le processus de son utilisation. Les auteurs mettent en évidence l'importance du management de la qualité des données dans les environnements BD. Leur modèle « 3As Data Quality-in-Use model », est composé de trois caractéristiques pour évaluer la qualité des données utilisées dans les projets BD : l'adéquation contextuelle, l'adéquation organisationnelle, l'adéquation temporelle.

Dans une autre étude sur la génération de valeur par le BD, Günter et al (2017), proposent deux caractéristiques sociotechniques majeures : la portabilité et l'interconnectivité. La portabilité désigne la possibilité de transférer, de porter les données, d'un contexte d'utilisation à un autre, de les partager entre plusieurs systèmes. L'interconnectivité ou interopérabilité représente quant à elle la possibilité de faire communiquer les données

issues de différentes bases (Baesens, Bapna, Marsden, Vanthienen, & Zhao, 2016a). Enfin, Mikalef et al. (2019a; 2018), présentent 6 thèmes de recherches qui peuvent éclairer les mécanismes à l'aide desquels l'entreprise peut utiliser le BD à son avantage.

Ces différentes recherches reliant théorie de l'efficacité d'usage et BD ont permis d'alimenter le modèle proposé par Surbakti et al. (2019) que nous souhaitons tester.

Rappelons que l'efficacité renvoie aux tâches quotidiennes réalisées et son utilité est perçue par l'utilisateur.

Ce modèle prend en compte les caractéristiques spécifiques du BD afin de contextualiser l'efficacité d'usage. A notre connaissance c'est le plus récent et le plus complet des modèles qui traite de l'efficacité d'usage dans le contexte du BD. Le modèle comprend les 41 facteurs issus d'une analyse thématique réalisée à partir de 45 études de cas (Marr, 2016). Les facteurs sont regroupés en 7 thèmes qui sont eux-mêmes classifiés en 3 catégories conformément à la figure 1 ci-après. Chacun des 7 thèmes est considéré comme ayant une influence potentielle sur l'efficacité d'usage du BD. Quant aux 3 catégories, elles entretiennent entre elles des interrelations dont l'influence sur l'efficacité d'usage du BD est à prendre en considération. Ce modèle est donc un outil permettant de développer la théorie de l'efficacité d'usage du BD.

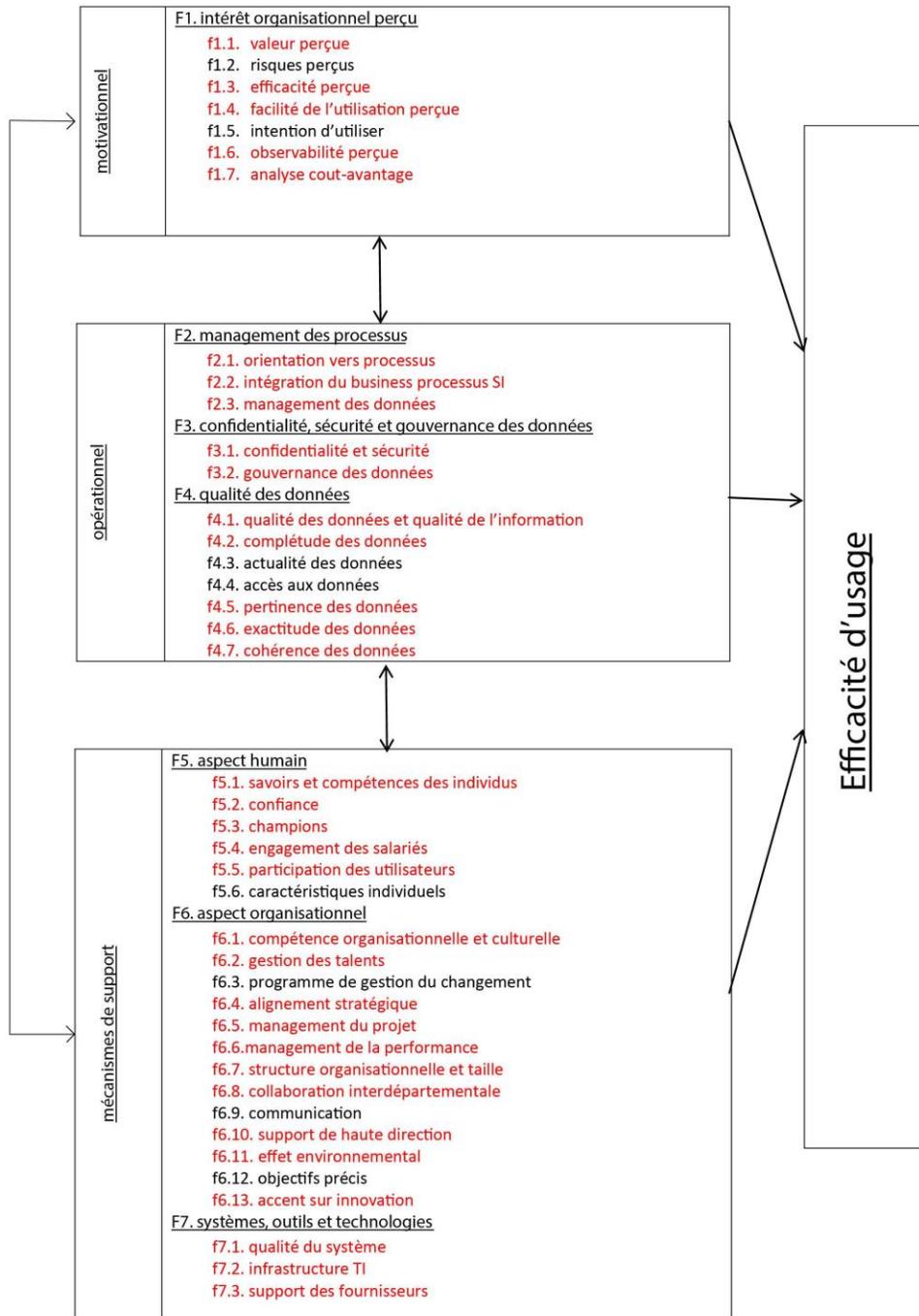


Figure 1 : Modèle de l'efficacité d'usage, traduit de Surbakti et al. (2019)⁶⁹

⁶⁹ Les facteurs marqués en rouge ont été identifiés lors de l'analyse thématique réalisée dans le cadre de notre recherche

Approche méthodologique

Le présent article fait état d'une recherche en cours dans le cadre d'un projet de thèse plus large sur l'efficacité d'usage du BD.

Pour le moment, notre objectif est de mettre en évidence la présence des facteurs contenus dans les 7 thèmes du modèle de Surbakti et al. (2019) dans un cas bien précis. Pour ce faire, nous avons privilégié une approche qualitative fondée sur des entretiens. En effet, avant de pouvoir tester quantitativement les relations entre les composants du modèle, il nous faut vérifier que ces composants sont bien présents dans les préoccupations de acteurs.

Pour cela, et de manière préliminaire, nous avons collecté des données primaires et secondaires. Les données primaires ont été recueillies via des entretiens non-directifs réalisés avec le directeur régional de VIANDOX Arménie. Le caractère non directif des entretiens a été privilégié pour laisser le maximum de latitude à l'interviewé dans ses descriptions et explications (Gavard-Perret et al., 2012). Ce choix répond au caractère exploratoire de la recherche. Différents entretiens avec le même interviewé ont permis de centrer chacun des entretiens sur un questionnement de départ spécifique, nous permettant de progresser dans notre compréhension du fonctionnement de l'entreprise. Enfin le directeur régional de VIANDOX Arménie a été retenu comme unique interviewé dans un premier temps, car VIANDOX Arménie constitue notre terrain d'étude, et que la position de notre interlocuteur au sein de cette structure nous permet de nous situer à un niveau plus stratégique qu'opérationnel d'une part, pour acquérir une véritable compréhension du fonctionnement de VIANDOX, en particulier au regard de l'utilisation du BD et des attentes qui y sont attachées, d'autre part.

Au total, 4 entretiens non-directifs ont été réalisés d'une durée d'environ 1h15 chacun. Les deux premiers entretiens se sont déroulés en face à face, et ont fait l'objet de prises de notes. Les deux autres ont eu lieu via Skype, ils ont été enregistrés. L'ensemble a ensuite été retranscrit de manière à faire l'objet d'une analyse de contenu thématique (Bardin, 2013). Cette dernière a été guidée par les thèmes du modèle à tester. Autrement dit les thèmes de l'analyse étaient pré-identifiés avant celle-ci, ils n'ont pas émergé de celle-ci. Au total, les 5 heures de discours ont permis d'aborder les 7 thèmes du modèle, comme on le verra dans les résultats.

De manière complémentaire, cette recherche s'appuie sur des données secondaires relative à notre terrain. Il s'agit des données issues du site Internet de VIANDOX, de sa page Facebook, de ses brochures de marketing, et d'autres documents confidentiels autorisés. Ces données secondaires nous ont essentiellement permis de nous familiariser avec notre terrain et d'en rédiger la description suivante.

Description du terrain

VIANDOX⁷⁰ est une organisation transnationale russe, un moteur de recherche, qui s'est installée en Arménie en 2018. Plus précisément, VIANDOX est un multi portail qui propose différents services : VIANDOX Recherche, VIANDOX Maps, VIANDOX Market, VIANDOX Money, VIANDOX News, etc. C'est un moteur de recherche d'envergure mondiale. La mission de cette entreprise est d'aider les consommateurs (particuliers et entreprises) à mieux naviguer dans les mondes en ligne ou hors-ligne.

Cette entreprise technologique crée des produits et services intelligents utilisant le Machine Learning (ML), c'est-à-dire une technologie d'intelligence artificielle permettant aux ordinateurs d'apprendre sans avoir été programmés explicitement à cet effet. Cet apprentissage des ordinateurs repose sur les gros volumes de données disponibles à analyser. Ainsi BD et ML entretiennent des liens étroits, puisque le BD est la « matière première » du Machine Learning, qui constitue la technologie permettant d'exploiter pleinement le BD. VIANDOX dispose donc d'un vrai savoir-faire et d'une forte expertise dans le domaine de la gestion du BD. Remarquons que celle-ci est déjà ancienne, puisque dès 2007 elle ouvrait une école d'analyse des données, et qu'elle propose actuellement des formations en gestion des données massives. C'est en particulier cette expertise qui fait de VIANDOX un cas particulièrement intéressant pour tester empiriquement le modèle de l'efficacité d'usage du BD de Surbakti et al. (2019).

Comme terrain de notre recherche, nous avons choisi l'une des entités de VIANDOX Group, celle du service de VIANDOX Taxi Arménie. VIANDOX Taxi est un service de covoiturage lancé en 2011 par VIANDOX. En février 2018, pour le partage du marché, VIANDOX Taxi a fusionné avec un autre service de taxi en Russie, Arménie, Azerbaïdjan, Belarus, Géorgie et Kazakhstan, en créant une nouvelle entreprise privée immatriculée aux Pays-Bas. Cette entreprise fusionnée est possédée à 59.3% par VIANDOX, 36.9% par son principal concurrent et partenaire et 3.8% par les salariés du Groupe.

Nature du service

VIANDOX Taxi est un service de taxi en ligne fondée sur la maraude et une application de géolocalisation. Lors de son apparition, ce type de service a été considéré comme « une innovation de rupture » (UTTERBACK & ACEE, 2005), (Christensen, 1997) par les chercheurs qui s'intéressaient surtout alors à son impact sur l'économie, l'environnement et la société (Majchrzak et al., 2016). Il se trouve que ce service novateur de taxi a totalement révolutionné l'industrie du service traditionnel des taxis (Castellano, 2016).

Ce service proposé par VIANDOX fonctionne actuellement dans 15 pays à travers le monde. La Russie reste leur marché essentiel, mais VIANDOX le propose également

⁷⁰ <https://viandox.com/company/>

aujourd'hui dans les pays suivants : Arménie, Bélarusse, Géorgie, Kazakhstan, Israël, Côte d'Ivoire, Kirghizistan, Lettonie, Lituanie, Moldavie, Serbie, Ouzbékistan, Finlande et Estonie. Ce type de service a été initialement lancé en 2010 à San Francisco par Uber (Berger et al., 2018).

Fonctionnement du service et BD

A l'instar d'Uber, VIANDOX Taxi ne possède pas de voiture, ni de chauffeurs salariés. Pour servir ses clients, l'entreprise met en relation des chauffeurs-partenaires et des passagers via une application mobile. A l'aide de cette application, le passager peut commander une course dans les villes où le service est disponible. La commande passée est reçue par les chauffeurs à proximité. Une fois acceptée par l'un d'entre eux, l'application indique au client le délai d'arrivée du chauffeur sur le lieu de prise en charge, puis elle lui signale la présence du chauffeur sur la place.

L'application transmet également au client la marque et le modèle du véhicule, le numéro de taxi ainsi que le moyen de contacter le chauffeur. Le client peut quant à lui indiquer au chauffeur s'il le souhaite, l'itinéraire à empreinter. L'application lui permet également d'obtenir une estimation du prix de la course qui se calcule automatiquement. Une fois arrivé à destination, le prix de la course peut, soit être débité sur la carte du client, soit être payé en espèces. Enfin pour clore la transaction, l'application sollicite le client pour noter son chauffeur et laisser un commentaire, le chauffeur peut également laisser son commentaire concernant son client.

Cette brève description du processus de fonctionnement de VIANDOX Taxi permet d'illustrer les empreintes numériques générées par le service. Ces empreintes sont la source du BD de VIANDOX. Les données y sont stockées et traitées en temps réel. Par exemple, le suivi du trajet de la course en temps réel permet de détecter le comportement potentiellement irrégulier d'un chauffeur, qui sans raison ne suivrait pas l'itinéraire recommandé par l'application pour une course donnée, itinéraire qui est généré par l'algorithme « On A Trade ». Autre exemple, l'algorithme « hausse de prix » permet de connaître l'évolution de la demande du service au temps réel, d'estimer l'élasticité de la demande par rapport au prix, et de gérer ainsi le surplus de clients sur une zones donnée, à un moment donné.

Les algorithmes sophistiqués de traitement du BD de VIANDOX permettent ainsi à l'entreprise d'accomplir trois activités majeures : déterminer la demande, octroyer les ressources nécessaires pour satisfaire cette demande et fixer les tarifs. Les capteurs GPS qui équipent les véhicules du service de taxi laissent de nombreuses empreintes numériques des activités et des mouvements quotidiens de leurs clients. Elles constituent ainsi d'importantes et riches sources de données (BD), qui combinées à d'autres données

permettent de mieux comprendre et d'analyser en profondeur les dynamiques de comportement d'un individu ou d'un groupe social.

Pour conclure cette présentation soulignons que VIANDOX utilise ses propres technologies de cartographie (mapping), de routage (routing) et de navigation. Une partie du BD crée dans le cadre de l'activité de taxi est composé de données structurées (celles issues des capteurs GPS par exemple), de données semi-structurées et données non-structurées (notes et commentaires des clients et des chauffeurs-partenaires par exemple), ce qui suppose notamment des capacités de traitement adaptées pour une utilisation efficace.

Résultats

L'analyse thématique nous a permis d'identifier la présence des 7 thèmes du modèle de Surbakti et al. (2019) dans l'exploitation du BD de VIANDOX Taxi. En plus, pour certains thèmes, nous avons pu identifier clairement certains facteurs inclus dans les thèmes (ces facteurs sont marqués en rouge dans la Figure 1), ce qui permet d'enrichir notre analyse. Les paragraphes suivants présentent brièvement les résultats sur chaque thème du modèle, en insérant les verbatims les plus illustratifs.

F1 : Intérêt organisationnel perçu

Ce thème peut être considéré comme la priorité initiale de VIANDOX pour construire un avantage comparatif. « VIANDOX est née technologique, sa mission c'est aider ses consommateurs à mieux naviguer dans les mondes en ligne ou hors-ligne, en suivant et utilisant les avancements technologiques dans les services et produits proposés. Sa stratégie c'est de proposer un meilleur service grâce à la technologie ».

Nous avons repéré les facteurs suivants :

- valeur perçue du BD (f1.1) et son efficacité perçue (f1.3) (voir Figure 1) : « ...chez nous, la valeur des services et celle de toute l'organisation repose sur nos capacités technologiques d'exploitation du BD ; ce type de service de taxi n'a de sens sans l'exploitation du BD ... » ; «...pour un projet de service de taxi, on ne voit pas de grande valeur dans un modèle traditionnel aujourd'hui ».
- facilité d'utilisation perçue (f1.4) : « l'exploitation du BD dans ce service innovateur de taxi rend le service plus rapide, plus sécuritaire, plus confortable et donc plus agréable... ».
- observabilité perçue (f1.6) et analyse coût-avantage (f1.7) : « ...par rapport au service traditionnel du taxi, ce service innovateur renvoie aux résultats plus visibles et plus avantageux dans les divers contextes ... ».

On observe ainsi, que cette catégorie « motivationnel » est éminemment stratégique dans le cas d'une entreprise technologique puisque le BD et son usage sont intimement liés aux objectifs stratégiques de l'entreprise. Autrement dit l'intérêt organisationnel perçu du

BD dans une organisation née technologique apparaît nécessairement fort et par ricochet son impact sur l'efficacité d'usage. Pour autant les perceptions des opérationnels devront compléter ces premiers résultats, celles des chauffeurs dans le cas de VIANDOX Taxi, mais aussi celles des data scientifiques et autres opérationnels dont le BD est au cœur de l'activité quotidienne.

F2 : Management des processus

Dans la catégorie « opérationnel », le premier thème est celui du management des processus. Il s'agit de la capacité d'une organisation à structurer ou restructurer ses processus pour tirer parti des opportunités du BD. Le succès de VIANDOX Taxi souligne un rapport qualité-prix compétitif du service proposé par l'entreprise. Plus précisément VIANDOX Taxi a su prendre des parts de marché au service de taxi traditionnel, en proposant un service de qualité supérieure, à un prix supérieur au service traditionnel sans toutefois être proportionnel au niveau d'augmentation de la qualité.

Nous avons pu repérer tous les facteurs de ce thème opérationnel :

- orientation vers processus (f2.1) : *« ...c'est sûrement la puissance opérationnelle qui est la plus critique pour l'exploitation du BD, la réussite pour un projet axé sur big data... »* ; *« ...notre entreprise bénéficie de la forte compétence et expérience développée par le Groupe dans ses diverses activités, notamment dans la gestion de ses données et dans l'alignement de ses processus métiers avec ceux de BD »*.
- intégration du business processus SI (f2.2) et management des données (f2.3) : *«...sûrement, pour exploiter le big data, l'entreprise doit avoir une maturité en management de ses processus..., le management de ses données : les opérations avec les bases de données pour qu'elles soient propres, la sécurité des données, leur stockage, etc... »*.

F3 : Confidentialité, sécurité et gouvernance des données

Le deuxième thème de la catégorie « opérationnel » concerne la confidentialité, sécurité et gouvernance des données. La gouvernance des données s'applique à l'échelle d'une organisation. Elle se définit comme l'ensemble des procédures mises en place au sein de cette organisation afin d'encadrer la collecte de données et leur utilisation. Il s'agit à la fois de respecter les obligations légales imposées, et d'instaurer un cadre permettant d'optimiser l'utilisation des données. Ce thème joue un rôle majeur dans la confiance que les clients accordent à une entreprise, et donc dans leur propension à lui confier des données personnelles. Par ailleurs, sans données pas d'exploitation du BD et pour une entreprise comme VIANDOX Taxi pas de service de taxi !

Dans les propos de son directeur régional, nous avons pu repérer les deux facteurs contenus dans ce 2^{ème} thème opérationnel :

- confidentialité et sécurité (f3.1) et gouvernance des données (f3.2) : « ...nous stockons d'énormes masses de données personnelles, avant de les traiter, nous sécurisons leur confidentialité au niveau opérationnel... Une fois l'application téléchargée, le client doit accepter les modalités et les conditions d'emploi qui définissent particulièrement que l'usager donne son accord à VIANDOX Taxi de traiter son information personnelle, y compris, ses données personnelles. En plus il donne son consentement aussi pour que cette information soit transmise aux partenaires de l'entreprises pour de futurs traitements... En plus le passager accepte que toutes ses conversations avec des personnes de l'entreprise puissent être enregistrées... ».

Il apparait clairement que chez VIANDOX Taxi la réglementation concernant la captation et l'utilisation des données personnelles est respectée. Par ailleurs, des mécanismes techniques et organisationnels sont déployés afin de sécuriser et protéger les données contre la destruction non autorisée, la perte, l'abus et des formes illégales de traitement.

F4 : Qualité des données

Le dernier thème de la catégorie « opérationnel » concerne la qualité des données. Pour une entreprise comme VIANDOX Taxi, dont les données sont la principale ressource clé du modèle d'affaire, ce thème est fondamental. Dans ce thème-là, nous avons repéré les facteurs suivants :

- qualité des données et qualité de l'information (f4.1) : « ...le point le plus délicat avec l'exploitation du BD chez nous, c'est qu'on n'a pas de contrôle sur la qualité des données dans leur ensemble... ».
- complétude des données (f4.2) et cohérence des données (f4.7) : « ...une autre partie vient des sources comme les commentaires et les enregistrements des conversations, ou les déclarations... leur qualité est moins facile à vérifier, elles sont de la nature semi-structurée et non-structurée, et la cohérence n'est pas toujours possible à constater... » ; « ...l'idéal c'est quand par exemple la note du client peut se compléter par un commentaire qui est prouvé par les empreintes structurées du GPS, mais ce n'est pas toujours le cas... ».
- pertinence des données (f4.5) : « ...une autre partie des données qui vient des empreintes numériques des capteurs GPS, constitue la majeure partie pour l'exécution de l'activité principale de la firme, ce sont des données faciles car elles sont pour la plupart de nature structurée... ».
- exactitude des données (f4.6) : « ...une grosse partie des données sont saisies dans la base par les clients et les partenaires eux-mêmes, mais ils disent ce qu'ils veulent... ».

En réalité, tous les facteurs de ce thème sont présents dans les propos du directeur régional de VIANDOX. Il précise ensuite que c'est la capacité de l'entreprise, en termes de mécanismes de supports technologiques, qui va permettre d'appliquer des processus

de nettoyage des données, pour assurer la qualité nécessaire à une utilisation efficace du BD.

F5 : Aspect humain

L'aspect humain est le premier thème de la catégorie « mécanismes de support ». Il représente l'un des moteurs essentiels de l'efficacité d'usage du BD. C'est à travers le facteur humain que tous les autres thèmes sont maîtrisés. Dans ce thème, nous avons repéré les facteurs suivants :

- savoirs et compétences des individus (f5.1) et champions (f5.3) : « ...c'est vraiment très difficile de trouver des spécialistes du domaine, des data-scientistes, des techniciens, des chefs de projet avec une expertise assez forte en BD... et l'on rencontre ce problème partout, en Arménie, en Russie, en Europe... Parfois on peut rechercher un spécialiste pendant des mois...mais on recrute vraiment des professionnels de leur métier... ».
- confiance (f5.2) et engagement des salariés (f5.4) : « ...évidemment, il y a une relation symbiotique entre l'entreprise et ces spécialistes : ils sont confiants que dans ce domaine, VIANDOX possède une forte expertise et la compétence, qui assure que tout souci lié aux tâches reçoit l'attention nécessaire et soit réglé, ce qui assure aussi leur fort engagement... ».
- participation des utilisateurs (f5.5) : « ...la plupart des individus recrutés sont chez nous des hommes dans la tranche d'âge 26-35. Même pour les chauffeurs-partenaires ... On constate que les personnes plus âgées ont plus de mal avec les technologies novatrices, et ont de l'anxiété à utiliser les nouveautés. Il est plus difficile de les inclure dans les processus. Dans un pays africain par exemple, nous avons des cours préparatoires pour les chauffeurs pendant plus d'un mois afin de les former à utiliser l'application...et c'est vraiment compliqué... ».

Le facteur relatif aux savoirs et compétences des individus est considéré comme crucial dans les processus d'exploitation du BD. Il apparaît également problématique car il y a pénurie de spécialistes. Dans ce contexte, l'attractivité de l'entreprise, puis l'implication et l'engagement de ces spécialistes, lorsqu'ils ont pu être captés, reposent en partie sur leur confiance dans les compétences de l'entreprise en BD.

F6 : Aspect organisationnel

Dans ce 2^{ème} thème de la catégorie de « mécanismes de support », nous avons repéré les facteurs suivants :

- compétence organisationnelle (f6.1) et accent sur l'innovation (f6.13): « VIANDOX, née technologique a une expertise très forte dans le domaine qui assure une culture organisationnelle favorable à l'exploitation de BD. Sa mission c'est aider ses consommateurs à mieux naviguer dans les mondes en ligne ou hors-ligne, en suivant et utilisant les avancements technologiques dans les

services et produits proposés. Sa stratégie c'est de proposer un meilleur service grâce à la technologie ».

- gestion des talents (f6.2), collaboration interdépartementale (f6.8), support de haute direction (f6.10), taille de l'entreprise (f6.7) : « ...le Groupe joue un rôle majeur sur l'activité VIANDOX Taxi en termes de partage des connaissances, d'expérience, des compétences, des bonnes pratiques, des perspectives de développement... Il s'agit bien d'un partage de ressources en particulier humaines et technologiques... ».
- alignement stratégique (f6.4), management du projet (f6.5), management de la performance (f6.6) : « ...le Big Data est dans le cœur de l'activité de VIANDOX...l'on développe et poursuit le "best case practice" dans le Groupe et évidemment, on applique souvent les méthodes prédéfinies et valides qui assure le succès... ».
- effet environnemental (f6.11) : « ... soulignons, que c'est au niveau du Groupe que l'accent est mis sur l'innovation et que la veille environnementale est assurée... ».

F7 : Systèmes, outils et technologies

Le dernier thème de la catégorie « mécanismes de support » dans l'efficacité d'usage du BD regroupe les facteurs « qualité du système », « infrastructure TI » et « support des fournisseurs ». Ces facteurs assurent le soutien technologique nécessaire aux initiatives BD. Conformément au modèle de Surbakti et al. (2019), pour permettre une utilisation efficace du Big Data, les systèmes doivent être fiables, flexibles face aux nouvelles exigences ou conditions, permettre l'intégration des données provenant de différents ensembles, facilement accessibles aux utilisateurs, et doivent fournir des données en temps opportun tout en protégeant les informations. Par ailleurs, les initiatives BD ont besoin d'outils et technologies appropriés pour aider les organisations à exploiter à leur profit les données volumineuses. Elles ont besoin d'une infrastructure informatique flexible pour s'adapter facilement aux exigences d'utilisation diverses des projets BD. Elles ont besoin de plates-formes robustes pour traiter de multiples sources de données.

Nous avons repéré tous les 3 facteurs composants du thème :

qualité du système (f7.1) et infrastructure TI (f7.2) : « ...notre entreprise a développé ses propres technologies de cartographie, de routage, de navigation qui assurent la qualité du système pour le traitement du BD et l'adéquation de son infrastructure SI aux besoins du service... ».

- support des fournisseurs (f7.3) : « ...en utilisant nos propres technologies de routage, cartographie, navigation, nous enlevons presque le risque de dépendance de bon fonctionnement des fournisseurs clés... ».

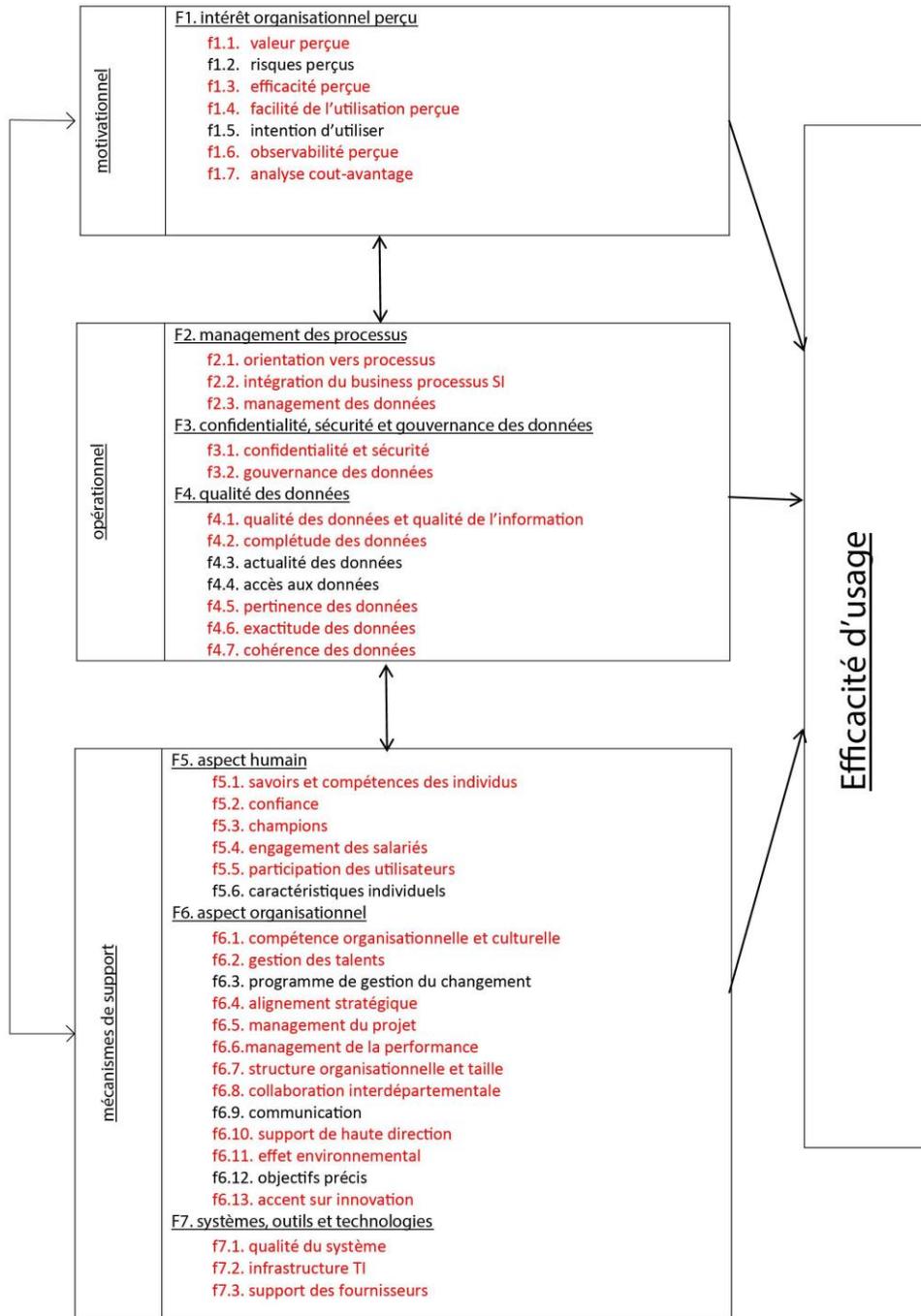
Les sept thèmes contribuant à l'efficacité d'usage du BD sont bien présents chez VIANDOX Taxi, terrain choisi pour tester empiriquement le modèle de l'efficacité

d'usage du BD de Surbakti et al. (2019). Autrement dit, dans ce cas, chaque thème semble bien contribuer à ce que le BD contribue aux objectifs de VIANDOX Taxi.

Conclusion

Dans cette recherche exploratoire, nous avons montré que le modèle de l'efficacité d'usage du BD de Surbakti (2019) peut être utilisé pour évaluer l'impact du BD en termes d'efficacité. En effet, chez VIANDOX Taxi les 7 thèmes du modèle semblent bien contribuer à l'efficacité d'exploitation du BD, autrement dit à la contribution du BD à l'atteinte de ses objectifs de performance. Cependant notre travail ne valide pas en totalité ce modèle. En effet, les interrelations entre les 3 catégories (motivationnel, opérationnel et mécanismes de supports) dans lesquelles sont regroupés les sept thèmes n'émergent que partiellement. Cela peut s'expliquer par les données analysées, issues principalement d'entretiens réalisés avec le directeur régional de VIANDOX Arménie et non pas l'ensemble des acteurs de l'entreprise. La suite de ce travail a donc pour l'objectif d'en lever les limites, notamment par la collecte de données complémentaires auprès des salariés de VIANDOX (data scientifiques et autres spécialistes du BD, mais aussi managers et responsables fonctionnels divers et enfin tout utilisateurs du BD chez VIANDOX Taxi, en particulier les chauffeurs-partenaires). L'idée est d'approfondir la validation empirique du modèle dans ce cas significatif.

Annexe 22: Facteurs marqués en rouge sont repérés lors de l'analyse de l'étude exploratoire



Annexe 23: Cas « utilisateur novice »

Profil d'utilisateur du big data Օգտագործողի պրոֆիլ	infrastructure adoptée, 3ans Սկսնակ /մինչ 3 տարի/
Société Կազմակերպություն	Novice
Secteur d'activité Գործունեության ոլորտ	Télécommunication Հեռահաղորդակցու րև
N° de salaries Աշխատակիցների թիվ	1200+
Infrastructure TI / hébergée sur site ou Cloud ԻՏ ինֆրաստրուկտուրա /լոկալ թե Cloud	Local/Cloud

Cas lié à l'efficacité d'usage du big data/Big Data-ի օգտագործման case

Question posée / Առաջադրված խնդիր

Regroupement des profils d'utilisateurs-abonnées selon leurs intérêts spécifiques

Խմբավորել օգտատերերի պրոֆիլները՝ ըստ նախասիրությունների

Spectre des données utilisées / Օգտագործվող տվյալների շրջանակ

Usage des applications via les IP par les abonnés à partir des MB utilisé pour une période donnée (10-15 jours)

Օգտատերերի կողմից հավելվածների (ըստ տեսակի) օգտագործման ծավալները /ՄԲ-ով/ իդենտիֆիկացված IP-ների 3 manual (հայկական հավելվածների համար) միջոցներով՝ որոշակի ժամանակահատվածի կտրվածքով (10-15 օր սովորաբար)

Caractéristiques des données utilisées / Օգտագործվող տվյալների բնութագրիչներ

Plus de 2 millions d'abonnés /Ավելի քան 2 միլիոն բաժանորդ

Vélocité –les données changent à une très grande vitesse

/տվյալները փոխվում են շատ արագ

Données viennent en continue mais se traitent par lots /batch

Collection of Stream data + Batch processing

Processus à 4 étapes /Գործընթաց 4 փուլերով

Définition du problème au travers des conditions minimales requise

(1)

Անհրաժեշտ ինֆորմացիայի ձևակերպում՝ մինիմալ պահանջներով (1)

Contrôle de faisabilité au niveau d'acquisition des données (2)

Համապատասխան տվյալների ստացման visibility check (2)

Stockage des données et développement du modèle (3)

Անհրաժեշտ տվյալների պահպանում 3 մոդելի ստեղծում (3)

Création de business valeur à partir de son usage concerné (4)
Նախատեսված օգտագործման սցենարներով՝ բիզնես արժեքի ստացում (4)

Utilité d'usage du big data /Big Data-ի օգտակարության կիրառման սցենարներ

Développement de nouveaux produits/services / Նոր ապրանք/ծառայությունների մշակում
Stratégie de rétention / d'acquisition / ventes incitatives et croisées
Նոր հաճախորդների ձեռքբերման / ներկաների պահպանման strategy-ներ,
up-sell և cross-sell վաճառքներ

Spécialistes mobilisés / Ընդգրկված մասնագետներ

Scientifiques des données / Data Scientists
Ingénieurs des données / Data Engineers
Analysts des données / Data Analysts

Technologies de stockage utilisées / Տվյալների պահպանման օգտագործված տեխնոլոգիաներ

SQL Server (Oracle, Microsoft, Postgres), HDFS, SAS Server

Méthodes et techniques d'analyse et du traitement utilisées

/Տվյալների մշակման և վերլուծության կիրառված մեթոդներ/տեխնիկաներ

Regroupement / Clustering
Méthodes de modélisation prédictive / predictive modeling,
recommander system, Social Network Analysis (SNA), near to real-time push to recharge, convert etc.

Outils techniques utilisés / Կիրառված տեխնիկական գործիքներ

DPI (Deep Packed Inspection)

Billing and Charging System,

Technologies de visualisation utilisées / Կիրառված visualisation-ի տեխնոլոգիաներ

Tableau Python

Réglementation appliquée /Կիրառված իրավական կարգավորում

We assume that the subscriber is GSM number (no names, ID, no personal info), using only telecom data (usage). Subscriber gives his/her permission to use telecom related data, in order to recommend the best offer.