

HEC MONTRÉAL

**Formes et antécédents de l'utilisation des
mégadonnées (Big Data) par les villes**

Par Hamza ALI

Sous la direction de Ryad TITAH

**Sciences de la gestion
(Option Technologies de l'information)**

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)*

Juillet 2018
© Hamza ALI, 2018

Table des matières

Sommaire	7
Remerciements	8
Chapitre 1 : Introduction	10
1.1 Mise en contexte	10
1.2 Problématique	11
1.3 Contributions potentielles	15
1.3.1 Contributions théoriques	15
1.3.2 Contributions pratiques	16
1.3.3 Limites de l'étude	16
1.3.4 Structure de l'étude	17
Chapitre 2 : Revue de la littérature	19
2.1 Approche de recensement des écrits	19
2.1.1 Développement d'un protocole de revue	19
2.1.2 Recherche de la littérature	20
2.1.3 Sélection des études	21
2.1.4 Extraction des données et des concepts	21
2.2 Clarification de la notion de mégadonnées	22
2.2.1 Définition des mégadonnées	22
2.2.2 Caractéristiques des mégadonnées	26
2.3 Conception de l'utilisation des TI par les villes	36
2.3.1 L'acceptation, l'utilisation continue et l'utilisation efficace	36
2.3.2 L'utilisation des TI par les villes	38
2.4 Antécédents organisationnels de l'utilisation des mégadonnées	40
2.4.1 Perspectives théoriques	40
2.4.2 Catégorisation des antécédents de l'utilisation	46
2.5 Modèle de recherche	54
2.5.1 Variables dépendantes	54
2.5.2 Variables indépendantes	54
2.5.3 Variables de contrôle	64

Chapitre 3 : Méthodologie	68
3.1 Sélection de l'approche méthodologique	68
3.2 Sélection de l'échantillon	69
3.2.1 Organisations visées	69
3.2.2 Répondants	70
3.2.3 Considérations éthiques	70
3.3 Opérationnalisation des variables	71
3.3.1 Démarche d'élaboration du questionnaire	71
3.3.2 Mesure des variables indépendantes	75
3.3.2 Mesure de la variable dépendante	76
3.4 Démarche d'évaluation des mesures et d'analyse des données	76
3.4.1 Approche d'évaluation des qualités psychométriques des mesures	77
3.4.2 Approche statistique d'analyse des données	80
3.4.3 Approche de codification des items de mesure	80
Chapitre 4 : Analyse des résultats	83
4.1 Analyse statistique des données collectées	83
4.1.1 Collecte et préparation des données	83
4.1.2 Données démographiques	83
4.1.3 Évaluation des qualités psychométriques des items réfléchifs	84
4.1.4 Évaluation des qualités psychométriques des items formatifs	90
4.1.5 Analyse de régression	92
4.2 Discussion des résultats	108
Chapitre 5 : Conclusion	113
5.1 Rappel des objectifs de l'étude et de l'approche méthodologique	113
5.2 Synthèse des principaux résultats	113
5.3 Apports de l'étude	114
5.4 Limites de l'étude et perspectives de recherche	116
Annexes	119

Liste des figures

Figure 1 : Processus d'implantation des SI, adapté de Kwon et Zmud (1987 : 233).....	14
Figure 2 : Schéma du processus de sélection des écrits.....	21
Figure 3 : Modèle d'utilisation du GE au niveau municipal (Titah, 2010 : 21).....	38
Figure 4 : Déterminants de l'adoption des innovations (Rogers, 2003 : 222).....	40
Figure 5 : Cadre TOE (Tornatzky et Fleischer, 1990 : 153).....	42
Figure 6: Cadre conceptuel de l'utilisation des mégadonnées par les villes	45
Figure 7 : Modèle de recherche de l'utilisation des mégadonnées par les villes	66
Figure 8 : Design du protocole méthodologique	68
Figure 9 : Illustration du tableau d'extraction des items	71
Figure 10 : Illustration de la nature des échelles de mesure	74
Figure 11 : Modèle MIMIC de l'avantage relatif	93
Figure 12 : Modèle MIMIC de la compatibilité.....	94
Figure 13 : Modèle MIMIC de la complexité	94
Figure 14 : Modèle MIMIC des préoccupations de confidentialité.....	95
Figure 15 : Modèle MIMIC des compétences technologiques	96
Figure 16 : Modèle MIMIC du dynamisme de l'écosystème.....	97
Figure 17 : Modèle MIMIC du support de la haute direction	97
Figure 18 : Modèle MIMIC de l'utilisation informationnelle	98
Figure 19 : Modèle MIMIC de l'utilisation intégrationnelle.....	99
Figure 20 : Modèle MIMIC de l'utilisation émergente	100
Figure 21 : Les antécédents de l'utilisation informationnelle.....	103
Figure 22 : Les antécédents de l'utilisation intégrationnelle	104
Figure 23 : Les antécédents de l'utilisation intégrationnelle	104

Liste des tableaux

Tableau 1 : Plan de concepts.....	20
Tableau 2 : Définition des éléments de la pyramide DIKW	23
Tableau 3 : Principales définitions des mégadonnées dans la littérature.....	24
Tableau 4 : Principales caractéristiques des mégadonnées dans la littérature	26
Tableau 5 : Classification des articles retenus selon les caractéristiques des mégadonnées..	27
Tableau 6 : Unités de mesure des volumes de données (The Economist, 2010).....	28
Tableau 7 : Définition des formes d'utilisation du GE [adapté de Titah (2010 : 143-144)]	39
Tableau 8 : Caractéristiques de l'innovation selon Rogers (2003 : 15-16).....	41
Tableau 9 : Composantes du cadre TOE selon Tornatzky et Fleischer (1990 : 152-154).....	42
Tableau 10 : Revue de la littérature sur les antécédents de l'utilisation des TI	47
Tableau 11 : Antécédents significatifs de l'utilisation des TI.....	49
Tableau 12 : Catégorisation des antécédents de l'utilisation.....	52
Tableau 13 : Table de correspondance DOI-Hypothèses-Caractéristiques.....	55
Tableau 14 : Liste des hypothèses	65
Tableau 15 : Approche d'évaluation des mesures et d'analyse des données.....	77
Tableau 16 : Mesures de fiabilité des items réfléchifs.....	85
Tableau 17 : Configuration des calculs dans SmartPLS.....	86
Tableau 18 : Validité convergente des items réfléchifs.....	87
Tableau 19 : Validité discriminante des items réfléchifs (Fornell-Larcker Criterion).....	89
Tableau 20 : Validité discriminante des items réfléchifs (ratio Heterotrait-Monotrait).....	90
Tableau 21 : Coefficients d'inflation de variance (VIF).....	91
Tableau 22 : Synthèse des résultats des tests des modèles structurels	102
Tableau 23 : Résultats des tests sur les variables de contrôle	105
Tableau 24 : Synthèse de l'analyse des modèles structurels	107
Tableau 25 : Comparatif des antécédents de l'utilisation des mégadonnées par les villes....	108
Tableau 26 : Contributions factorielles croisées (utilisation informationnelle)	153
Tableau 27 : Contributions factorielles croisées (utilisation intégrationnelle).....	154
Tableau 28 : Contributions factorielles croisées (utilisation émergente).....	155

Liste des annexes

Annexe 1 : Listes des requêtes effectuées pour la revue de la littérature	120
Annexe 2 : Liste des articles retenus pour la revue systématique de la littérature	123
Annexe 3 : Certificat d'approbation éthique.....	126
Annexe 4 : Mise en application des principales exigences éthiques.....	127
Annexe 5 : Littérature sur les antécédents organisationnels de l'utilisation des TI.....	128
Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche	131
Annexe 7 : Items de mesure des variables du modèle de recherche	142
Annexe 8 : Structure de codification des items de mesure.....	149
Annexe 9 : Données démographiques.....	152
Annexe 10 : Contributions factorielles croisées	153
Annexe 11 : Significativité des items des modèles MIMIC	156

Liste des abréviations, sigles et acronymes utilisés

ACIS	Americas Conference on Information Systems
ACM	Association for Computing Machinery
ANOVA	Analysis of Variance
ASIMC	Association des systèmes d'information municipale du Canada
BDA	Big Data Analytics / Analytique des mégadonnées
BOLD	Big and Open Linked Data
CDO	Chief Data Officer
CEO	Chief Executive Officer
CER	Comité d'éthique de la recherche
CFA	Component factor analysis
CIO	Chief Information Officer
DIKW	Data, information, knowledge and wisdom
DOI	Diffusion of Innovation theory
EBSCO	Business Source Complete
ECIS	European Conference on Information Systems
EDI	Échange électronique des données
GE	Gouvernement électronique
GIGO	Garbage In, Garbage Out
HICSS	Hawaii International Conference on System Sciences
HiPPO	Highest Paid Person's Opinion
IaaS	Infrastructure as a Service
ICIS	International Conference on Information Systems
IDC	International Data Corporation
IMap	Motivation intrinsèque vers l'accomplissement
IMkw	Motivation intrinsèque à connaître
Imst	Motivation intrinsèque pour expérimenter la stimulation
INV	Utilisation innovante
IoT	Internet of Things / internet des objets ou objets connectés
ISURA	IS use-related activity
LISREL	Linear Structural Relations
MAC	Media Access Control
MANOVA	Multivariate Analysis of Variance
MRP	Planification des besoins matériels
PaaS	Platform as a Service
PEOU	Perception de la facilité de l'utilisation
PIIT	Innovation personnelle avec les TI
PLS	Partial Least Squares regression
PU	Perception d'utilité
RBV	Resource-based View
RFID	Radio Frequency Identification
RIM	Rich Intrinsic Motivation
RSL/SLR	Revue systématique de la littérature / Systematic Literature Revue
RTN	Utilisation de routine
SaaS	Software as a Service
SEM	Structural Equation Modeling
SPSS	Statistical Package for the Social Sciences
TAM	Modèle d'acceptation des technologies
TIC	Technologies de l'information et de communication
TMT	Top Management
TOE	Technology-Organization-Environment framework
TPB	Théorie du comportement planifiée
TRA	Théories de l'action raisonnée
UTAUT	Théorie unifiée de l'acceptation et de l'utilisation des technologies

Sommaire

En 2008, la population urbaine a, pour la première fois, dépassé la population rurale et d'ici 2050, 70% de la population sera citadine (Nations Unies, 2014). Cette urbanisation rapide met beaucoup de pression sur les villes et pose des défis importants en ce qui a trait à plusieurs aspects de la vie urbaine. L'urgence entourant ces défis a amené de nombreuses villes à chercher des solutions technologiques pour les relever. La firme Gartner prévoit que, d'ici 2020, plus de 6.8 milliards d'objets connectés seront installés dans les villes.

Cette interconnexion de milliards d'objets générera de très grands volumes de données (Big Data). Si ce grand volume de données est géré et analysé correctement, il constitue un avantage important et peut créer un impact réel sur le fonctionnement des villes. En dépit de ces avantages, seules quelques organisations l'exploitent et certaines études estiment que moins de 0.5% de l'ensemble des données disponibles a été exploré. Ainsi, il s'avère important de comprendre les mécanismes sous-jacents à la décision des villes à utiliser ou non les mégadonnées. La présente étude cherche donc à répondre à deux questions de recherche : quelles sont les formes de l'utilisation des mégadonnées par les villes ? Et quels sont les antécédents de l'utilisation des mégadonnées par les villes ?

Pour répondre à ces deux questions, nous avons développé un modèle de recherche afin d'étudier les antécédents organisationnels de l'utilisation des mégadonnées aux étapes de post-adoption du cadre de Kwon et Zmud (1987). Ainsi, une revue systématique de la littérature (RSL) portant sur les mégadonnées dans un contexte de ville est effectuée. Les résultats de cette revue ont alimenté les éléments constitutifs du modèle de recherche. Pour les besoins de l'étude, nous avons intégré trois perspectives théoriques: le cadre technologie-organisation-environnement (TOE) de Tornatzky et Fleischer (1990), la théorie de la diffusion de l'innovation (DOI) de Rogers (2003) et le modèle d'utilisation du gouvernement électronique de Titah (2010).

Un questionnaire a été élaboré afin de mesurer les différents construits du modèle de recherche. Les échelles de mesure sont constituées d'items réfléchitifs et d'items formatifs, et ce afin d'obtenir un modèle de type Multiple indicators and multiple causes (MIMIC). Ceci a permis d'offrir une riche conceptualisation des différents types d'utilisation des mégadonnées par les villes, de même que des antécédents retenus dans le cadre de la présente étude.

Les hypothèses du modèle ont été testées à partir de données d'enquête auprès de 106 villes canadiennes. Les résultats des analyses effectuées par l'approche des moindres carrés partiels (PLS) suggèrent que les antécédents sont différents en fonction du type d'utilisation des mégadonnées. Pour l'utilisation informationnelle, les antécédents significatifs sont l'avantage relatif, la compatibilité, le dynamisme de l'écosystème. L'effet modérateur des compétences technologiques sur la complexité des mégadonnées est également significatif, ainsi que l'effet des pressions compétitives sur le support de la haute direction. Pour l'utilisation intégrationnelle, les antécédents significatifs sont l'avantage relatif, ainsi que le support de la haute direction qui lui-même est influencé par les pressions compétitives. Enfin, pour l'utilisation émergente, les antécédents significatifs sont l'avantage relatif, la compatibilité et le support de la haute direction.

Mots clés : mégadonnées, Big Data, villes, municipalités, smart city, post-adoption, TOE, DOI, eGovernment, SEM-PLS, questionnaire, utilisation organisationnelle.

Remerciements

Cette étude marque l'accomplissement de 18 mois de travail et de trois années d'étude, riches en enseignements professionnels et personnels. J'aimerais donc sincèrement remercier les personnes qui m'ont apporté leur soutien tout au long de ce projet et sans qui ce mémoire n'aurait probablement pas vu le jour.

Je voudrais tout d'abord adresser toute ma gratitude au directeur de ce mémoire, Ryad T'itah, pour sa disponibilité, son soutien, sa rigueur, ses judicieux conseils et dont l'étendue des connaissances a grandement stimulé ma réflexion.

Je remercie également l'ensemble des professeurs du Service de l'enseignement des TI de HEC Montréal qui m'ont transmis leur passion et qui m'ont fourni les outils nécessaires à la réussite de mes travaux académiques et professionnels. Je tiens à remercier spécialement Gregory Vial et Ana Ortiz de Guinea qui ont eu une contribution directe dans la réalisation de ce mémoire.

Je suis aussi extrêmement reconnaissant envers les répondants qui ont accepté de participer à cette étude, de même qu'au pré-test du questionnaire.

Je voudrais exprimer ma reconnaissance envers les amis et collègues qui m'ont apporté leur support moral et intellectuel tout au long de ma démarche. Je souhaite particulièrement remercier Jorge Passalacqua pour son soutien indéfectible, ce qui m'a grandement facilité la conciliation travail-études. Je tiens également à témoigner toute ma gratitude à Radia Benheddi, pour son support inestimable et pour m'avoir fourni de précieux conseils quant à la mise au point de l'outil de mesure.

Enfin, ce travail est dédié à tous les membres de ma famille, notamment à mes parents, à ma grand-mère, à mes frères et sœurs, à ma conjointe et à mes deux garçons. Un grand merci du fond du cœur pour être ce qui fait de moi un homme comblé.

“ In God we trust. All others must bring data ”
W. Edwards Deming

Chapitre 1 : Introduction

L'objectif de ce chapitre est de présenter l'objet du présent mémoire qui porte sur la conceptualisation de l'utilisation des mégadonnées par les villes. Dans un premier temps, une mise en contexte des défis auxquels les villes font face sera présentée. Elle sera suivie de la définition de la problématique de recherche, de l'identification des limites et des contributions potentielles et, enfin, de la manière dont l'étude est structurée.

1.1 Mise en contexte

La proportion de la population mondiale habitant dans les villes augmente rapidement. Selon un rapport de l'Organisation des Nations Unies (2014), cette proportion ne représentait que 8% au début du 19^e siècle et a atteint 20% un siècle plus tard. Depuis, le monde est entré dans un « millénaire urbain » et près de 150 000 personnes s'installent en ville chaque jour (International Electrotechnical Commission, 2014). En 2008, la population urbaine a, pour la première fois, dépassé la population rurale et les projections indiquent que, d'ici 2050, 70% de la population sera citadine (représentant ainsi près de 6.5 milliards de personnes). Cette urbanisation accélérée est notamment observée en Afrique avec un taux de croissance urbaine de 3.4 % et en Asie avec un taux d'urbanisation de 53% (Nations Unies, 2014).

Cette urbanisation rapide est notamment motivée par la volonté de profiter des opportunités d'emploi offertes par les villes, de même que par l'aspiration à une meilleure qualité de vie. Elle a, cependant, mis beaucoup de pression sur les villes (Belanche-Gracia, Casaló-Ariño et Pérez-Rueda, 2015; Koppenjan et Enserink, 2009) et pose actuellement des défis importants en ce qui a trait au logement, à la prestation de services publics, à la gestion des infrastructures, à la mobilité, à l'énergie, à l'environnement et à la gestion de l'eau et de l'assainissement (Klopp et Petretta, 2017; van Zoonen, 2016).

L'urgence entourant ces défis a amené de nombreuses villes à travers le monde à chercher des solutions technologiques pour les gérer (Mehmood *et al.*, 2017; Ransbotham *et al.*, 2016; Tachizawa, Alvarez-Gil et Montes-Sancho, 2015; van Zoonen, 2016). Ceci a notamment été facilité par le fait que les technologies de l'information et de communication (TIC) sont devenues de plus en plus répandues dans les zones urbaines (Fietkiewicz, Mainka et Stock, 2017; Gagliardi *et al.* ; Kourtit, Nijkamp et Steenbruggen) et peuvent permettre aux citoyens de façonner leur environnement urbain en participant à identifier et à mettre en œuvre des solutions à des problèmes urbains (van der Graaf et Veeckman, 2014).

Les initiatives de numérisation des villes visent à rendre l'utilisation des ressources plus efficace, à améliorer la qualité de la prestation de service, à fournir de l'information en temps réel aux citoyens et à favoriser un développement économique et social durable (Aguilera *et al.* ; Jara, Genoud et Bocchi, 2015). Des technologies, de plus en plus accessibles, telles que l'internet des objets (IoT) et la radio-identification (RFID) sont désormais intégrées aux diverses infrastructures urbaines ou disponibles grâce à des applications intégrées dans les points d'accès et les téléphones mobiles (Jara, Genoud et Bocchi, 2015; Rathore *et al.*, 2016; Re Calegari, Celino et Peroni, 2016). D'ailleurs, la firme de conseil Gartner (2016) prévoit que, d'ici 2020, plus de 25 milliards d'objets connectés seront

installés dans le monde. De l'ensemble de ces objets, plus d'un quart (6.8 milliards) sera introduit dans les villes.

Cette interconnexion de milliards d'objets et l'utilisation croissante de technologies généreront de très grands volumes de données à analyser et à exploiter (Popescul et Radu, 2016; Xiong, Zheng et Li, 2014). Selon Éric Schmidt, ancien CEO de Google, environ 2,5 quintillions d'octets (Exabytes) sont produits chaque jour, ce qui représente la moitié des données créées entre l'aube de la civilisation et l'année 2003 (Galizia, Gee et Landis, 2011 : 16; van Loenen, Kulk et Ploeger, 2016). De plus, un rapport de l'International Data Corporation, publié en juin 2011, indiquait que le volume mondial de données devrait augmenter de près de 50% par année.

Si ce grand volume de données, appelé Big Data, est géré et analysé correctement, il constitue un avantage important (LaValle *et al.*, 2011; Lycett, 2013) et peut créer un réel impact sur le fonctionnement des villes. D'ailleurs, dans leur article introductif à l'édition spéciale de la revue MIS Quarterly qui a porté sur l'intelligence d'affaires, H. C. Chen, Chiang et Storey (2012) indiquent que le gouvernement électronique (GE, eGovernment) figure parmi les cinq (5) domaines où le Big Data (ci-dessous appelé « mégadonnées ») va avoir le plus d'impact. Notons que le GE constitue l'une des bases de l'innovation urbaine, et est considéré comme l'un des aspects les plus importants à prendre en considération dans la recherche sur les villes (Fietkiewicz, Mainka et Stock, 2017).

Les mégadonnées représentent « une collection de jeux de données qui est si vaste et complexe qu'il devient difficile de les traiter à l'aide des outils et applications traditionnels de gestion et de traitement de bases de données » (Clarke, 2016 : 77). Même si elles réfèrent généralement à de grands ensembles de données (Kallinikos et Constantiou, 2015), le volume seul n'aurait jamais suffi à encapsuler la nouveauté du phénomène des mégadonnées (Boyd et Crawford, 2012). De ce fait, elles sont également définies dans la littérature par leurs attributs dont le nombre peut dépasser la dizaine (Markus, 2015). Toutefois, le nouveau paradigme entourant les mégadonnées vient plutôt de la combinaison de l'ensemble de ces attributs (Goes, 2014).

Pour les besoins de la présente étude, la notion de mégadonnées s'étend au-delà des ensembles générés par les activités des individus sur internet (Constantiou et Kallinikos, 2015), et comprend l'ensemble des données auxquelles une ville peut avoir accès, soit les données présentes dans les registres des villes, les données des enquêtes gouvernementales ou d'entreprises, les données provenant des médias sociaux (Mergel, Rethemeyer et Isett, 2016) de même que les données qui peuvent être acquises gratuitement ou en contrepartie d'une contribution.

1.2 Problématique

Il est estimé que les villes produisent près de 4.1 téraoctets de données par jour par kilomètre carré (Dobre et Xhafa, 2014). Jamais auparavant autant d'information sur les personnes, les événements, les lieux et les objets n'a été disponible (Gschwender, Munizaga et Simonetti, 2016). Cette masse de données résulte principalement de la numérisation (conversion de l'information analogique) des activités (Loebbecke et Picot, 2015) et de la forte volonté de tout mesurer. Elle a été facilitée par la baisse des coûts des composantes électroniques et des supports de stockage des données (Müller *et al.*, 2016).

La disponibilité de ces quantités massives de données offre des opportunités intéressantes afin, notamment, de comprendre la ville et les activités qui s'y produisent (Picone *et al.*, 2015 2836). Elle devrait également permettre de représenter, de modéliser et de simuler les résultats d'éventuelles interventions urbaines (Kitchin, 2014).

Ainsi, on retrouve, dans la littérature, plusieurs avantages que les mégadonnées devraient procurer aux organisations. Les mégadonnées sont utiles pour la conception de nouveaux produits et services, et permettent de transformer les modèles d'affaires (Loebbecke et Picot, 2015). Elles constituent une base essentielle à la productivité et à l'innovation (Zhou *et al.*, 2016) et à l'optimisation des processus d'affaires (D. Chen, Preston et Swink, 2015). En outre, les mégadonnées sont de plus en plus considérées comme un avantage concurrentiel majeur (LaValle *et al.*, 2011; Lycett, 2013) et un élément d'actif, à part entière (Abbasi, Sarker et Chiang, 2016; Janssen et Kuk, 2016a). Des preuves du pouvoir des mégadonnées peuvent être trouvées dans les secteurs privé et public.

Plusieurs exemples d'application des mégadonnées dans les entreprises sont cités dans la littérature. Les cas les plus édifiants concernent leur utilisation comme intrant aux fonctions prédictives (construction de modèles) et prescriptives (actions prises suite à une prédiction) des systèmes de prise de décision des firmes. Ainsi, dans le domaine du commerce du détail, par exemple, les algorithmes de la firme Target ont permis, en 2012, de découvrir la grossesse d'une adolescente, dont les parents n'étaient pas au courant, et ce en se basant sur ses habitudes d'achat (voir notamment Duhigg (2012) et Hill (2012)). Aussi, l'analyse des données effectuée par Wal-Mart a permis de découvrir que l'arrivée d'un ouragan dans une zone fait augmenter la demande sur les lampes de poche et les batteries, en même temps que celle sur les biscuits à la fraise (Pop-Tarts) (Hayashi, 2014; Loebbecke et Picot, 2015).

Dans le domaine de la politique, la campagne de 2012 de Barack Obama doit beaucoup de son succès à l'analyse quantitative qui a permis l'identification des personnes susceptibles d'être influencées par un dépliant, un appel téléphonique ou une visite à domicile (Hayashi, 2014). De plus, lors de ces élections de 2012, le statisticien Nate Silver a correctement prédit le résultat de vote dans les 50 États américains (PricewaterhouseCoopers, 2013).

En ce qui concerne le secteur public en particulier, plusieurs rapports anticipent que les mégadonnées vont favoriser la croissance économique (Foulonneau *et al.*, 2014). La firme McKinsey indique que la valeur des mégadonnées pour le secteur public en Europe représente un potentiel de 250 milliards USD par année (Manyika *et al.*, 2011 : 8). Malomo et Sena (2017) citent quelques exemples d'économies de coûts conséquentes réalisées par des villes britanniques grâce à la mise en œuvre d'applications basées sur les mégadonnées. À titre d'illustration, l'arrondissement londonien de Hammersmith & Fulham a économisé 1,15 M GBP par année grâce à un portail en libre-service et le district de Spelthorne réalise des économies annuelles d'environ 43 800 GBP et une baisse de 10% des appels aux services à la clientèle grâce à une application améliorant la communication entre le conseil et les citoyens (Malomo et Sena (2017 : 10).

En dépit de ces avantages, et en dépit du fait que l'utilisation de l'information est profondément ancrée dans les fonctions essentielles des gouvernements (Janssen et van den Hoven, 2015), et même si les organisations ont de plus en plus accès à de grandes quantités de données, quelques-unes d'entre elles seulement les exploitent (Martens *et al.*, 2016; Tachizawa, Alvarez-Gil et Montes-Sancho,

2015). D'ailleurs, un rapport de l'IDC (2012) indique que seule une petite fraction, estimée à 0.5%, de l'ensemble des données générées a été explorée (Gantz et Reinsel, 2012 : 3). De plus, un sondage de PricewaterhouseCoopers (2013 : 7) indique que la majorité des entreprises sondées n'ont pas encore adopté de technologies de mégadonnées. Enfin, une enquête menée par la firme Gartner indique que 30% des organisations sondées ont déclaré avoir investi dans les technologies de mégadonnées et que seul un quart d'entre elles (8%) a pu les mettre en production (Gartner, 2013).

Il existe plusieurs raisons pour lesquelles l'utilisation des mégadonnées pourrait ne pas être bénéfique (D. Chen, Preston et Swink, 2015). Tout d'abord, il est probable que d'importants changements structurels soient nécessaires à la réalisation des bénéfices, et les organisations peuvent être réticentes à effectuer ces mises à niveau. En outre, le dynamisme (vélocité et variabilité) des mégadonnées peut provoquer de l'incertitude quant à la concrétisation des bénéfices. Une telle incertitude rend difficile l'identification précise des ajustements organisationnels nécessaires. Aussi, on évalue que 80% de ces données sont constituées de données non structurées (Tachizawa, Alvarez-Gil et Montes-Sancho, 2015) que les systèmes traditionnels de gestion de données ne peuvent gérer (Davenport, Barth et Bean, 2012). En outre, Gartner (2013) indique qu'une grande confusion entoure les mégadonnées et une incertitude existe en ce qui trait à leur capacité à générer des bénéfices tangibles pour les organisations.

Ainsi, les organisations rencontrent aujourd'hui d'énormes défis liés à la manière dont les mégadonnées peuvent être utilisées (Goes, 2014). Ceci est d'autant plus frustrant qu'elles consacrent des ressources importantes à la collecte de ces données (H. C. Chen, Chiang et Storey, 2012; Saboo, Kumar et Park, 2016) ou à leur acquisition. Selon le même rapport de l'IDC, les investissements en mégadonnées ne cessent d'augmenter alors que les coûts des composantes baissent (Gantz et Reinsel, 2012).

En lien avec la valorisation des données, un des nouveaux paradigmes introduits par les mégadonnées concerne l'absence de buts préalables aux collectes des données. Ainsi, et contrairement aux pratiques traditionnelles, les mégadonnées sont souvent recueillies sans but préalablement défini, tout en espérant qu'un jour elles seraient utiles (Constantiou et Kallinikos, 2015; Müller *et al.*, 2016). Le schéma global qui est mis en évidence peut être résumé comme suit « first data then search for any possible uses of what is already available as data » (Constantiou et Kallinikos, 2015 : 51). Dans la même perspective, Chris Anderson (2008) indique que le monde actuel nous oblige, d'abord, à voir les données sous une perspective mathématique, et à établir le contexte plus tard.

De ce qui précède, nous comprenons, d'une part, que le simple accès aux données ne conduit pas nécessairement à la création de la valeur (Malomo et Sena, 2017; Mergel, Rethemeyer et Isett, 2016) et que ces grands ensembles de données ne sont utiles que lorsqu'ils ont été rigoureusement examinés, évalués et analysés (Constantiou et Kallinikos, 2015). D'autre part, même s'il n'existe aucune donnée officielle concernant l'utilisation des données par les villes (Malomo et Sena, 2017 : 10), il s'avère que la plupart des municipalités rencontrent d'importants défis quant à l'exploitation des données collectées. Malomo et Sena (2017 : 22) soulignent d'ailleurs « notre analyse suggère que la plupart des gouvernements locaux n'utilisent pas leurs données à leur plein potentiel ... et ce à cause de plusieurs facteurs ».

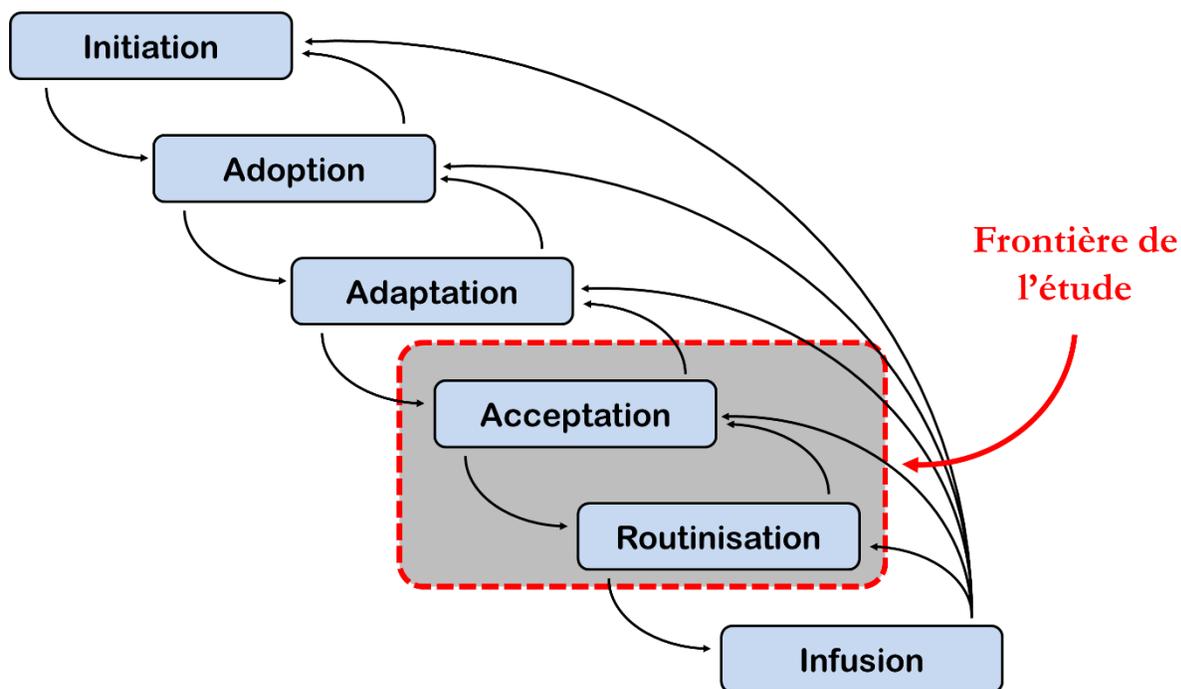
Ainsi, les mécanismes sous-jacents à l'utilisation des mégadonnées par les organisations méritent une enquête approfondie (D. Chen, Preston et Swink, 2015). Par conséquent, et compte tenu du contexte dans lequel évoluent les villes, il s'avère pertinent d'étudier les facteurs influençant l'utilisation des mégadonnées par les villes, de même que la nature de cette utilisation. Ainsi, la présente étude cherche à répondre aux deux questions de recherche suivantes :

QR1 | Quelles sont les formes de l'utilisation des mégadonnées par les villes?

QR2 | Quels sont les antécédents de l'utilisation des mégadonnées par les villes?

Pour répondre à ces questions de recherche, nous développerons un modèle de recherche afin d'étudier les antécédents de l'utilisation des mégadonnées aux étapes de post-adoption du cadre de Kwon et Zmud (1987), illustrées dans la [figure 1](#). Il s'agit en particulier de l'étape d'acceptation qui se traduit justement par l'utilisation de la technologie à l'étude, ainsi que de l'étape de routinisation durant laquelle l'utilisation de la technologie est encouragée et devient, graduellement, une activité normale et routinière (Cooper et Zmud, 1990 : 124).

Figure 1 : Processus d'implantation des SI, adapté de Kwon et Zmud (1987 : 233)



Ainsi, une revue systématique de la littérature (RSI) portant sur les mégadonnées dans le contexte des villes sera effectuée. Les résultats de cette revue alimenteront les éléments constitutifs d'un modèle conceptuel, qui sera traduit en modèle de recherche. Une enquête de terrain par questionnaire auto-administré sera conduite afin de tester le modèle de recherche.

1.3 Contributions potentielles

Les mégadonnées bouleversent la chaîne de valeur de l'information avec des répercussions généralisées (Abbasi, Sarker et Chiang, 2016; Chris Anderson, 2008; Constantiou et Kallinikos, 2015). Une telle perturbation présente des opportunités intéressantes (Abbasi, Sarker et Chiang, 2016). Ainsi, la présente étude examine l'utilisation des mégadonnées par les villes. À caractère empirique, elle vise à fournir des contributions intéressantes pour les praticiens ainsi que pour la recherche.

1.3.1 Contributions théoriques

Sur le plan théorique, les mégadonnées sont devenues, depuis une dizaine d'années, un centre d'intérêt pour la recherche, pour les industries de même que pour les gouvernements (De Gennaro, Paffumi et Martini, 2016; Zhou *et al.*, 2016). Toutefois, la littérature s'est souvent plainte de la faiblesse des bases théoriques concernant l'analyse des mégadonnées (Aaltonen et Tempini, 2014 ; D. Chen, Preston et Swink, 2015; Goes, 2014). La présente étude s'inscrit donc en réponse aux appels de la littérature à constituer des assises conceptuelles solides au phénomène des mégadonnées (Goes, 2014) et constitue une contribution théorique au domaine de l'étude des systèmes d'information.

En outre, plusieurs auteurs soulignent la nécessité à ce que la littérature sur les systèmes d'information prenne le leadership en étudiant les implications des mégadonnées dans les organisations et en dirigeant la manière dont les transformations organisationnelles doivent être menées (Goes, 2014 ; Mergel, Rethemeyer et Isett, 2016). La présente étude constitue une contribution qui s'inscrit dans cette optique.

Par ailleurs, contrairement à la riche littérature qui se concentre sur la prédiction de l'adoption des TI par les organisations à partir de l'analyse des intentions au niveau individuel, les études sur le comportement d'adoption au niveau organisationnel restent relativement rares, malgré l'importance de ce phénomène (D. Chen, Preston et Swink, 2015 ; Thong, 1999). Traitant de l'utilisation des mégadonnées au niveau organisationnel, la présente étude constitue une contribution dans cette perspective. Notre examen, à la fois, des antécédents, et de l'utilisation des mégadonnées par les municipalités devrait fournir une profonde compréhension des enjeux et des implications de la diffusion de cette innovation au sein des villes.

En outre, le nombre limité d'études portant sur l'adoption des TI au niveau organisationnel se concentre généralement sur les intentions d'adoption plutôt que sur les comportements post-adoption (D. Chen, Preston et Swink, 2015; Zhu *et al.*, 2006). Ainsi, la présente étude met l'accent sur l'utilisation post-adoption des innovations et devrait nous permettre d'obtenir une meilleure compréhension de la transformation des villes par les mégadonnées, et ce à travers la compréhension des facteurs d'influence et des formes de leur utilisation.

Pour les besoins de l'étude, nous avons intégré trois perspectives théoriques afin d'analyser et expliquer les forces qui favorisent les initiatives organisationnelles d'utilisation des mégadonnées : le cadre technologie-organisation-environnement (TOE) de Tornatzky et Fleischer (1990), la théorie de la diffusion de l'innovation (DOI) de Rogers (2003) et le modèle d'utilisation du gouvernement électronique de Titah (2010). Cette étude constitue donc une combinaison utile de ces cadres et

devrait permettre l'obtention d'une explication originale, ainsi qu'un modèle cohérent et complet de l'utilisation organisationnelle des mégadonnées par les villes.

Enfin, cette étude devrait offrir une contribution empirique aux questions liées à la valorisation des mégadonnées au sein des villes, un domaine dans lequel la majorité de la littérature est de nature conceptuelle (Chauhan, Agarwal et Kar, 2016).

1.3.2 Contributions pratiques

Sur le plan pratique, il s'avère important de rappeler que le gouvernement électronique a permis l'ajout de valeur de nombreuses façons (Ransbotham *et al.*, 2016). Les changements introduits par les mégadonnées sont tout aussi importants du fait qu'elles peuvent provoquer des restructurations organisationnelles au niveau des villes et au niveau de leurs relations avec les citoyens (Constantiou et Kallinikos, 2015). Toutefois, la majorité de la littérature sur les mégadonnées dans le domaine des systèmes d'information se concentre sur la description des défis que les mégadonnées posent aux organisations. Les mégadonnées sont là pour rester (Woerner et Wixom, 2015) et les gouvernements locaux devront tenir compte de leur nature problématique. La présente étude constitue une tentative de proposition de solutions pratiques aux gestionnaires publics (Mergel, Rethemeyer et Isett, 2016) en contribuant à offrir un aperçu de la façon dont certains de ces défis devraient être relevés.

L'application de l'intelligence d'affaires aux mégadonnées dans le contexte des municipalités reste à ses débuts (Hashem *et al.*, 2016) et cette étude devrait contribuer à appréhender les défis liés au processus de valorisation des mégadonnées collectées au sein des espaces urbains. Elle devrait également aider les praticiens, et notamment les gestionnaires des TI au sein des villes, à définir les facteurs à mettre en œuvre en vue de l'utilisation de cette innovation. Ainsi, un contexte adéquat est indispensable à la réussite des initiatives d'adoption du concept de la ville intelligente et des conditions clés de succès sont nécessaires. Enfin, cette étude devrait contribuer à la compréhension de la manière dont les mégadonnées sont utilisées dans un contexte urbain.

1.3.3 Limites de l'étude

La présente étude présente plusieurs limites en lien avec la méthodologie et avec le contexte de la recherche.

Premièrement, l'étude porte principalement sur les étapes post-adoption du processus d'implantation des mégadonnées au sein des villes, et plus particulièrement sur les phases d'acceptation et de routinisation du cadre de Kwon et Zmud (1987). Toutefois, nous ne pourrions pas évaluer avec précision le degré de maturité des technologies utilisées par les villes participantes à l'étude. La sophistication de l'utilisation est le seul indicateur qui permet d'évaluer cette maturité. Par conséquent, il se peut que certaines villes soient à divers stades du processus d'implantation.

Deuxièmement, même si beaucoup d'attention a été consacrée à l'élaboration du questionnaire afin qu'il soit le plus clair et le plus précis possible (notamment à travers l'incorporation d'une section comportant des définitions et des exemples d'utilisation), les répondants pourraient avoir des perceptions différentes de ce que représentent les concepts de « mégadonnées », « utilisation des mégadonnées » et « entités utilisatrices des mégadonnées ».

Enfin, l'approche de validation du modèle de recherche qui a été retenue est le questionnaire. Ceci limite la profondeur de l'analyse vu qu'il ne permet pas de recueillir toute la richesse de l'expérience qu'on peut obtenir de tels projets.

1.3.4 Structure de l'étude

Ce mémoire est structuré en cinq chapitres, y compris la présente introduction. Le prochain chapitre porte sur la revue de la littérature, et a pour but de recenser l'ensemble des écrits pertinents en lien avec le sujet des mégadonnées au niveau municipal. Pour cela, on s'appuie d'abord sur la littérature pour clarifier le concept de mégadonnées et ses caractéristiques dans un contexte urbain. Par la suite, nous aborderons les concepts liés à l'utilisation des TI au niveau municipal. Enfin, nous analyserons les antécédents découlant de la littérature sur l'utilisation organisationnelle des TI. Cette revue de la littérature nous permettra de concevoir le modèle de recherche de la présente étude.

Le troisième chapitre du mémoire précise la méthodologie et offre une description détaillée du devis de recherche. Le choix de l'enquête de terrain par questionnaire auto-administré comme approche de validation du modèle de recherche et comme méthode de collecte de données y sera justifié. La justification de l'unité d'analyse, des caractéristiques de l'échantillon, de la démarche d'analyse des données, de même que les considérations éthiques de l'étude sont également précisées dans ce chapitre.

Le quatrième chapitre a trait aux à l'analyse des données et à la discussion des résultats. Les données collectées grâce au questionnaire permettront de tester le modèle de recherche dans l'optique de mieux comprendre les antécédents de l'utilisation des mégadonnées au sein des villes ainsi que la manière dont ces mégadonnées sont utilisées par les villes.

Finalement, le dernier chapitre conclut ce mémoire. Il présente les conclusions et synthétise les contributions pour la recherche et pour la pratique. Les limites et les pistes de recherches futures y sont également mentionnées.

**"If we have data, let's look at data.
If all we have are opinions, let's go with mine"**

JIM BARKSDALE

Chapitre 2 : Revue de la littérature

L'étude des formes et des antécédents de l'utilisation des mégadonnées par les villes est à l'intersection de plusieurs sous-domaines de la discipline des systèmes d'information : l'implantation des technologies de l'information dans les organisations, l'intelligence d'affaires, la gestion des savoirs et des connaissances, l'administration publique et le gouvernement électronique (eGovernment).

Pour couvrir la littérature portant sur ces dimensions, une revue systématique de la littérature a été effectuée, dont la démarche est présentée à la première section du présent chapitre. Les autres sections portent sur la clarification de la notion des mégadonnées, la conceptualisation de l'utilisation des mégadonnées, de même que la conception des antécédents de l'utilisation des mégadonnées. La dernière section de ce chapitre propose un modèle de recherche basé sur le modèle d'utilisation des TI par les villes (Titah, 2010) et la catégorisation des antécédents de l'utilisation des TI au niveau organisationnel.

2.1 Approche de recensement des écrits

En se basant sur les lignes directrices suggérées par Paré *et al.* (2016), une revue systématique de la littérature a été réalisée dans cette étude. Ces lignes directrices comprennent plusieurs activités telles que le développement d'un plan de revue de littérature, la recherche des documents et la sélection des études.

Les paragraphes qui suivent (Section 2.1.1) visent à décrire ces activités, tandis que les autres sections de ce chapitre aborderont les activités d'extraction des données et des concepts clés des études retenues, de même que l'analyse et la synthèse des données et informations recueillies.

2.1.1 Développement d'un protocole de revue

Dans la présente étude, l'examen de la littérature vise principalement le développement d'un modèle de recherche sur les formes et antécédents de l'utilisation des mégadonnées par les villes. Elle a également pour but d'identifier les différents items permettant de mesurer les construits de ce modèle de recherche. Enfin, elle permet de contextualiser l'étude en identifiant les recherches menées sur le sujet et les domaines qui ne sont toujours pas explorés (Webster et Watson, 2002).

Afin d'atteindre ces objectifs, le protocole de revue a porté sur deux volets. Premièrement, nous avons réuni les principaux articles de référence portant sur l'utilisation et sur les antécédents de l'utilisation des systèmes d'information. Nous avons également utilisé les techniques de recherche arrière (Backward Search) et avant (Forward search) afin d'identifier l'ensemble des documents pertinents à notre sujet. Pour les aspects liés aux mégadonnées dans un contexte de ville, un plan de concepts a été développé d'une manière itérative. Ainsi, un plan initial contenant deux principales catégories de termes (Big Data et Digital City) a été élaboré et testé sur quelques bases de données. Les résultats ont permis d'enrichir la liste des termes de recherche. Le [tableau 1](#) ci-dessous présente le plan de concepts final.

Tableau 1 : Plan de concepts

Big Data	Digital City	
"Big Data"	"Smart City"	"Smart Cities"
"Huge Data"	"Digital City"	"Digital Cities"
"Massive Data"	"Ubiquitous City"	"Ubiquitous Cities"
"Large Data"	"Connected City"	"Connected Cities"
"Urban Data"	"Intelligent City"	"Intelligent Cities"
"City Data"	"Real-time City"	"Real-time Cities"
"Municipality Data"	"e-Government"	

2.1.2 Recherche de la littérature

Pour la recherche sur les antécédents et les formes de l'utilisation des mégadonnées par les villes, nous nous sommes principalement basés sur les huit revues du Senior Scholars' Basket of Journals. Ce choix est motivé par la haute qualité des articles qui y sont publiés. Nous avons également inclus les revues spécialisées dans la publication d'articles portant sur l'administration électronique : Electronic Journal of e-government, Government Information Quarterly et Public Administration Review. Vu le peu de résultats obtenus dans ces revues, nous avons décidé, dans un second temps, de ne pas combiner les deux concepts et de formuler plutôt une requête par concept de recherche.

Comme l'étude des mégadonnées dans les villes constitue un sujet en émergence, nous avons utilisé la base de données IEEEExplore pour la récupération des comptes-rendus des meilleures conférences dans le domaine des systèmes d'information, à savoir : ICIS, ACIS, ECIS et HICSS. Enfin, afin d'élargir le bassin des revues scientifiques potentielles, d'autres bases de données ont été utilisées. La liste finale des bases de données utilisées pour la revue de la littérature est la suivante : Business Source Complete (EBSCO), ABI/INFORM (ProQuest), ScienceDirect, Wiley Online Library, IEEEExplore et Emerald.

Pour la recherche d'articles, nous avons utilisé la fonction de « Recherche avancée » des bases de données citées ci-dessus. En fonction du plan de concepts présenté dans le [tableau 1](#) ci-dessus, voici la requête générique effectuée :

```
("Big Data" OR "Huge Data" OR "Massive Data" OR "Large Data" OR "Urban data" OR "City Data" OR "Municipality Data") AND ("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Real-time Cities" OR "e-Government")
```

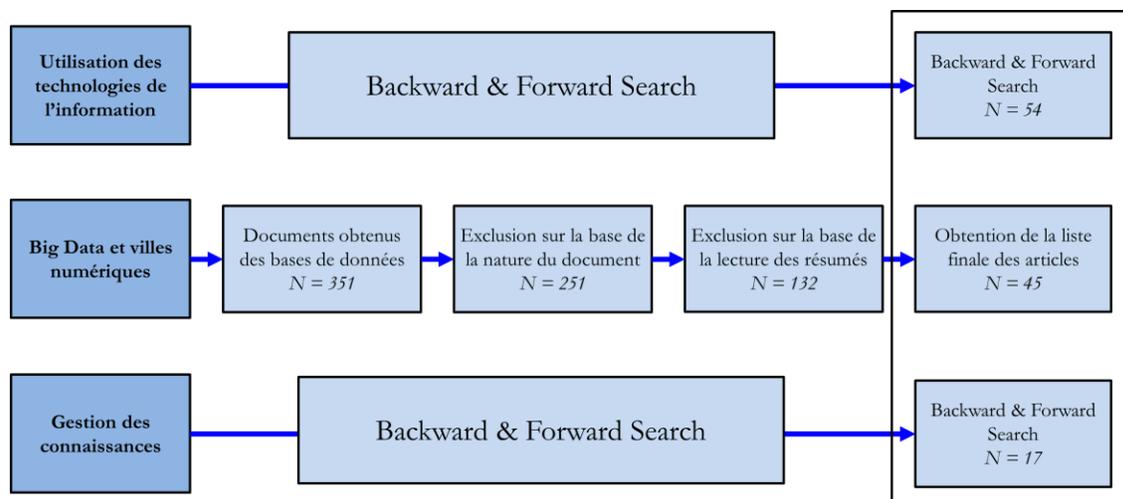
La requête indiquée ci-dessus a dû être ajustée dans certains cas, vu que, par exemple, le moteur de recherche de la base de données n'accepte que quinze (15) termes par requête. La liste complète des requêtes effectuées est indiquée à l'Annexe 1. Enfin, la version Microsoft Windows du logiciel de gestion de références bibliographiques EndNote a été utilisée pour gérer l'ensemble des références mobilisées dans le cadre de la présente étude.

2.1.3 Sélection des études

Le processus de sélection des études obtenues avec le protocole de recherche indiqué ci-dessus s'est effectué en trois étapes. Dans un premier temps, nous avons procédé à l'exclusion des articles en double, des éditoriaux, des commentaires, des revues d'articles, et des préfaces. Ainsi, cette étape a permis d'identifier 251 documents de recherche. Dans un deuxième temps, une sélection s'est effectuée sur la base de la lecture du résumé (Abstract) de chaque document. C'est à cette étape que nous avons donc exclu les articles qui, bien qu'ils contenaient les mots clés du plan de concept, n'étaient pas pertinents au contexte de notre étude. À la suite de cette étape, un total de 132 articles de recherche a été retenu.

Enfin, la sélection finale s'est effectuée sur la base de la lecture de l'intégralité de chaque article de recherche. Deux critères de sélection, correspondant aux objectifs de la présente revue de la littérature, ont été retenus. Les articles devaient satisfaire à une des deux questions suivantes : l'article discute-t-il des formes ou antécédents de l'utilisation des mégadonnées ? ou l'article permet-il d'avoir une compréhension des défis posés par les mégadonnées dans un contexte de villes ? Ainsi, ces critères ont permis l'identification de 45 articles de recherche qui pourraient contribuer à satisfaire aux objectifs de la présente étude. La liste finale des études retenues est indiquée à l'annexe 2. Enfin, le processus de sélection et le nombre des documents de recherche identifiés à chaque étape sont présentés à la [figure 2](#) ci-dessous.

Figure 2 : Schéma du processus de sélection des écrits



2.1.4 Extraction des données et des concepts

L'extraction des concepts s'est effectuée grâce à plusieurs tableaux qu'on retrouve dans les sections 2.2 à 2.5 ci-dessous, et notamment dans les tableaux 3, 4, 5, 6, 11, 12 et 13, de même que dans l'annexe 3. Ces tableaux ont permis une extraction uniforme des concepts et la présentation synthétisée de l'information pertinente. Ainsi, les deux dernières étapes suggérées par Paré *et al.* (2016) (c.-à-d. l'extraction des concepts clés et l'analyse et la synthèse des concepts extraits) seront présentées dans les sections 2.2 à 2.5.

2.2 Clarification de la notion de mégadonnées

Cette section vise à clarifier la notion des mégadonnées dans la littérature, et ce dans un contexte de ville numérique. En effet, si certains auteurs les définissent en fonction des défis qu'elles posent, d'autres les définissent par leurs caractéristiques. Nous allons donc dans un premier temps définir la notion des mégadonnées. Les caractéristiques des mégadonnées seront abordées dans un second temps.

2.2.1 Définition des mégadonnées

Le terme mégadonnées (Big Data) est un mot composé d'un adjectif (méga, ou grand) et d'un nom (donnée). Essayons d'abord de clarifier la signification du concept de « donnée » dans la littérature avant de discuter de la définition de la notion de « mégadonnées ».

On retrouve plusieurs définitions du concept de données dans la littérature sur la gestion des savoirs et des connaissances : Wersig et Neveling (1975), Cleveland (1982) Zeleny (1987), Cooley (1987), Choo (1996), Awad et Ghaziri (2004), Bellinger, Castro et Mills (2004), Jashapara (2004), Pearlson et Saunders (2004) et Chaffey et White (2010). La définition la plus citée est probablement celle proposée par (Ackoff, 1989: 3), selon laquelle les données sont « les symboles qui représentent les propriétés d'un objet, d'un événement ou de leur environnement ». Notons également la définition de Choo (1996) qui élargit la définition du concept afin d'inclure des éléments des systèmes physiques (tels que les livres ou les tableaux de bord) qui donnent des indices sur les données et la façon dont elles doivent être lues.

Les travaux de Rowley (2007) ont permis d'identifier les éléments communs aux principales définitions retrouvées dans la littérature. Premièrement, dépourvues de contexte et d'interprétation, les données n'ont pas de grande signification ou de valeur. Deuxièmement, les données représentent des faits objectifs ou des observations qui ne sont ni organisés ni transformés et ne donnent aucune signification spécifique. Enfin, les éléments de données sont une description élémentaire et enregistrée des objets, des événements, des activités et des transactions.

Selon Ackoff (1989), les données ne sont d'aucune utilité jusqu'à ce qu'elles soient dans une forme utilisable (c.-à-d. pertinente). De ce fait, ce concept est lié à plusieurs autres concepts : information, connaissance, intelligence, savoir, compréhension et croyance. Les liens entre ces différents concepts ont été formalisés dans la « Wisdom Hierarchy » ou la hiérarchie DIKW, en référence à Data (donnée), Information (information), Knowledge (connaissance), Wisdom (compréhension). Le [tableau 2](#) ci-dessous indique la définition, dans la littérature, de chaque élément de la pyramide DIKW.

De ce qui précède, il ressort que le fait de disposer de jeux de données ne suffit pas, et qu'un processus de création de la valeur à partir des données brutes doit avoir lieu. Ce processus de transformation contribue à l'acquisition de la connaissance et permet une prise de décision éclairée (Choo (1996) ; Bellinger, Castro et Mills (2004)).

Tableau 2 : Définition des éléments de la pyramide DIKW

Couche	Définition
Donnée (Data)	Les données sont les symboles qui représentent les propriétés d'un objet, d'un évènement ou de l'environnement. (Ackoff, 1989). Elles ne sont d'aucune utilité jusqu'à ce qu'elles soient dans une forme utilisable (c'est-à-dire pertinente). La différence entre les données et les informations est fonctionnelle, et non structurelle (Rowley, 2007).
Information (Information)	L'information est un ensemble de données organisées ou structurées. Ce traitement confère aux données une certaine pertinence dans un contexte précis, ce qui les rend utiles, significatives, et pertinentes (Rowley, 2007).
Savoir (Knowledge)	Le savoir est une combinaison d'information, de compréhension, de compétences, de capacités, d'expérience et de valeurs (Rowley, 2007).
Compréhension (Wisdom)	La compréhension est un phénomène interpersonnel, nécessitant l'exercice de l'intuition, de la communication et de la confiance. Elle se caractérise par le fait qu'elle est incorporée dans une action de prise de décision, elle implique l'utilisation des connaissances et l'exercice du jugement dans des situations complexes. Enfin, elle exige l'examen de considérations éthiques et sociales (Rowley et Slack, 2009).

Depuis quelques années, la notion de mégadonnées a gagné en importance parmi les praticiens de même que dans les travaux de recherche. Sémantiquement, ce terme désigne le fait de disposer de grands volumes de données. Pourtant, les préoccupations liées aux volumes des données datent de la moitié du 20e siècle (Press, 2013). Ainsi, Fremont Rider indiquait, en 1944, que les bibliothèques universitaires américaines doublent de taille tous les 16 ans. Compte tenu de ce taux de croissance, Rider prévoyait que la Bibliothèque de Yale en 2040 aurait environ 200 millions de volumes, qui occuperaient plus de 6.000 miles d'étagères (Rider, 1944 : 12). Aussi, en 1961, Derek Price publiait une étude sur la croissance du nombre de revues scientifiques et dans laquelle il constatait que le nombre de nouvelles revues augmentait de façon exponentielle, doublant tous les 15 ans (Price, 1961 : 169). De la lecture des travaux de ces deux auteurs, on comprend que leur principale préoccupation portait surtout sur l'augmentation des coûts liés aux espaces « physiques » nécessaires à l'entreposage des documents.

Il fallait cependant attendre 1999 pour que le terme « Big Data » soit utilisé, et ce par Steve Bryson et ses collègues lors d'une communication de l'ACM (Association for Computing Machinery). Les auteurs soulignaient que des « ordinateurs très puissants sont une bénédiction pour de nombreux domaines d'enquête. Ils sont aussi une malédiction. Des calculs rapides émettent des quantités massives de données. Là où les ensembles de données en mégaoctets étaient jadis considérés comme grands, nous trouvons maintenant des ensembles de données à partir de simulations individuelles dans la gamme de 300 Go. Mais la compréhension des données résultant des calculs haut de gamme est un effort important. Comme l'a dit plus d'un scientifique, il est tout simplement difficile d'examiner tous les chiffres... » (Bryson *et al.*, 1999 : 83). Depuis, les travaux sur les mégadonnées affluent parmi les praticiens et dans le milieu universitaire. La plupart de ces travaux commencent par un effort de définition de la notion des « mégadonnées », ce qui indique qu'une définition cohérente n'a pas encore été établie. Le [tableau 3](#) ci-dessous en présente les principales définitions identifiées dans la littérature.

Tableau 3 : Principales définitions des mégadonnées dans la littérature

Auteurs	Définition/Éléments de définition
Baesens <i>et al.</i> (2016 : 808)	Big data to refer to data that includes information acquired externally together with data gathered internally.
Constantiou et Kallinikos (2015 : 45)	Big data often represents miscellaneous records of the whereabouts of large and shifting online crowds. It is frequently agnostic ... based on varying formats and modes of communication ... raising severe problems of semiotic translation and meaning compatibility. Crucially, the usefulness of big data rests on their steady updatability.
Goes (2014 : vii)	Big data is about massive amounts of observational data, of different types, supporting different types of decisions and decision time frames.
Lycett (2013 : 381)	It is the increasing availability of data that provides the impetus for BI, most typically characterized by industry commentators in the '3 Vs' as follows: Volume, Velocity and Variety.
Mergel, Rethemeyer et Isett (2016 : 931)	Big data is high volume data that frequently combines highly structured administrative data actively collected by public sector organizations with continuously and automatically collected structured and unstructured real-time data that are often passively created by public and private entities through their Internet interactions.
Müller <i>et al.</i> (2016 : 1)	We understand Big Data Analysis (BDA) as the statistical modeling of large, diverse, and dynamic data sets of user-generated content and digital traces.
US National Science Foundation (2012)	Big data includes large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources.
Yang <i>et al.</i> (2014 : 1)	Big data as having five key characteristics, known as "the five Vs": volume, velocity [the speed at which data is created], variety, veracity, and value.
Zuboff (2015 : 78)	The informing of the economy... is constituted by a pervasive and continuous recording of the details of each transaction.

La revue de littérature de Mergel, Rethemeyer et Isett (2016) sur les mégadonnées dans le secteur public a permis de catégoriser les définitions retrouvées dans la littérature selon cinq disciplines : gestion, politiques publiques, sciences politiques, gestion des technologies de l'information et ingénierie informatique. Les définitions proposées par les auteurs reflètent donc les opportunités et les défis offerts par les mégadonnées dans la discipline auxquels ils appartiennent, soit les affaires publiques.

Dans les présents travaux, nous nous intéressons aux antécédents et aux formes de l'utilisation des mégadonnées par les villes. Ainsi, nous nous concentrons sur l'analyse des définitions qui portent sur les facteurs de l'utilisation des mégadonnées. Dans cette perspective, nous remarquons, tout d'abord, que les définitions des mégadonnées proposées dans la littérature peuvent être classées en deux principales catégories : en fonction de leurs caractéristiques et en fonctions de leurs sources.

La définition proposée par Constantiou et Kallinikos (2015) s'inscrit dans la série des tentatives de définition des mégadonnées par leurs caractéristiques (voir par exemple : Goes (2014), Yang *et al.* (2014), Lycett (2013) et Saboo, Kumar et Park (2016). Toutefois, celle de ces deux auteurs est axée

sur les données ayant pour source les médias sociaux vu que, d'après eux, c'est surtout les données générées par les réseaux sociaux qui donnent aux mégadonnées les caractéristiques que nous leur attribuons aujourd'hui. En outre, notons également la définition de Müller *et al.* (2016) qui ajoute que les données peuvent provenir de « traces numériques » laissées suite à une activité donnée. Une attention particulière sera accordée à cet aspect dans la prochaine sous-section.

Par ailleurs, Baesens *et al.* (2016) proposent une définition en fonction de la source des données: données transactionnelles, en opposition aux données non transactionnelles, et données internes en opposition aux données de sources externes. Pour ces auteurs, les mégadonnées sont donc des données transactionnelles ou non, acquises à l'interne ou à l'externe de l'organisation. En outre, selon Zuboff (2015), les mégadonnées sont l'agrégation de multiples petits jeux de données, rendue possible grâce à l'enregistrement omniprésent et continu des détails de chaque activité (transaction).

La définition des mégadonnées, appliquée au secteur public, proposée par Mergel, Rethemeyer et Isett (2016) combine les caractéristiques des mégadonnées à leurs sources. Ainsi, les auteurs avancent que dans les administrations publiques des données hautement structurées, générées par les actions des employés et citoyens, peuvent être combinées à des données non structurées, générées en continu, d'une manière automatique (senseurs, objets connectés, activité sur internet) par des entités publiques ou privées.

Considérant les éléments abordés dans la présente section, nous adoptons la définition suivante dans le cadre de la présente étude : « c'est l'ensemble des données auxquelles une municipalité a accès et qui peuvent être hautement structurées, générées par les actions des employés et citoyens, ou non structurées, générées en continu, d'une manière automatique par des entités publiques ou privées (comme les coordonnées GPS, les vidéos/images des caméras et les capteurs embarqués) ».

D'une manière générale, bien que les mégadonnées réfèrent principalement au volume des ensembles de données, nous porterons une attention particulière, dans les présents travaux, à la nature de leur collecte (ponctuelle, continue ou automatique), à leur format (structurées ou non structurées), à leurs sources (publiques ou privées), à leur granularité (détaillées ou agrégées) et à l'horizon temporel de leur utilité (court terme, très court terme ou en temps réel). Ces éléments reflètent la complexité de ces jeux de données qui possèdent des caractéristiques qui leur sont propres. Ces caractéristiques font justement l'objet d'une analyse plus approfondie dans la prochaine section.

2.2.2 Caractéristiques des mégadonnées

Comme il a été vu précédemment, il est généralement admis que le concept des mégadonnées est multidimensionnel, défini par plusieurs caractéristiques qui commencent pour la plupart par la lettre V (Markus, 2015). Ce sont ces caractéristiques qui ont amené certains auteurs à proposer la définition suivante aux mégadonnées « une collection de jeux de données qui est si vaste et complexe qu'il devient difficile de les traiter à l'aide des outils et applications traditionnels de gestion et de traitement de bases de données » (Clarke, 2016 : 77).

Dans une note de recherche, Laney (2001) offrait une définition tridimensionnelle des mégadonnées liées au commerce électronique : volume (grandes quantités de données), vitesse (vitesse élevée de génération des données) et variété (diversification des sources des données). Une quatrième dimension, la véracité, a été rajoutée par Schroeck *et al.* (2012). Depuis, d'autres caractéristiques se sont ajoutées, dont le nombre peut dépasser la dizaine (Markus, 2015). Ces caractéristiques défient les pratiques conventionnelles de gestion et de gouvernance des données (Woerner et Wixom, 2015).

Dans le [tableau 4](#) ci-dessous, nous clarifions les caractéristiques des mégadonnées, retenues par les auteurs identifiés dans la littérature. Les paragraphes qui suivent visent à décrire les caractéristiques les plus pertinentes au sujet de notre étude. Il s'agit notamment du volume, de la variété, de la véracité, de la vitesse, de la valeur, de la confidentialité, des dynamiques humaines et de l'agnostique. Dans le [tableau 5](#) nous classerons la littérature en fonction de ces caractéristiques.

Tableau 4 : Principales caractéristiques des mégadonnées dans la littérature

Caractéristique	Éléments de définition
Volume	La taille des données disponibles a augmenté à un rythme croissant, favorisée par la multiplication des sources génératrices de données.
Variété	Les données collectées peuvent être de nature très diverse, peuvent prendre des formes variées et hétérogènes et peuvent être structurées ou non structurées.
Vélocité/Vitesse	Les données sont traitées en temps réel, et qu'elles sont consommées (utilisées) au fur et à mesure qu'elles sont produites.
Véracité	La véracité (fiabilité) désigne la dimension qualitative des données et qui est nécessaire au bon fonctionnement des outils de traitement et à une prise de décision éclairée.
Valeur	La valeur indique l'extraction des connaissances et de la compréhension à partir des données pour obtenir un avantage concurrentiel ou pour une performance plus efficace.
Agnostique	Les données sont produites à des fins génériques ou à des fins différentes de celles recherchées par le croisement et l'analyse des données.
Confidentialité et vie privée	Les menaces à la confidentialité des informations et les atteintes à la vie privée constituent l'un des principaux inconvénients des mégadonnées.
Dynamiques humaines	La combinaison des mégadonnées aux algorithmes évolutifs transforme les organisations et a un effet sur plusieurs dimensions humaines.

Tableau 5 : Classification des articles retenus selon les caractéristiques des mégadonnées

Sources	Volume	Variété	Vélocité	Véracité	Valeur	Vie privée	Dynamiques humaines	Agnostique
Abbasi, Sarker et Chiang (2016)	•	•	•	•	•			
Abella, Ortiz-de-Urbina-C. et De-Pablos-H. (2017)					•			
Baesens <i>et al.</i> (2016)				•	•			
Belanche-Gracia, Casaló-Ariño et Pérez-Rueda (2015)						•		
Bergamaschi <i>et al.</i> (2016)						•	•	
Bhimani (2015)		•			•			
Brynjolfsson, Geva et Reichman (2016)		•		•				
Clarke (2016)				•				
Constantiou et Kallinikos (2015)	•	•	•				•	•
D. Chen, Preston et Swink (2015)	•	•	•					
Davenport (2014)	•	•	•					
Höchtel, Parycek et Schöllhammer (2016)	•	•	•	•	•			
Janssen et Kuk (2016b)						•	•	
Janssen et van den Hoven (2015)	•	•				•		
Jara, Genoud et Bocchi (2015)							•	
Luftman <i>et al.</i> (2015)	•	•			•			
Lycett (2013)	•	•	•		•		•	
Markus (2015)						•	•	
Martens <i>et al.</i> (2016)				•				
Mehmood <i>et al.</i> (2017)	•	•	•	•	•			
Menon et Sarkar (2016)						•		
Mergel, Rethemeyer et Isett (2016)	•	•	•					•
Müller <i>et al.</i> (2016)								•
Newell (2015)							•	
Pan <i>et al.</i> (2016)	•	•	•	•	•			
Popescul et Radu (2016)						•		
Ransbotham <i>et al.</i> (2016)						•		
Saboo, Kumar et Park (2016)	•	•	•					
Spiekermann et Korunovska (2016)						•		
Šuh <i>et al.</i> (2015)	•	•	•					
Tachizawa, Alvarez-Gil et Montes-Sancho (2015)	•	•	•					•
van Loenen, Kulk et Ploeger (2016)						•		
van Zoonen (2016)						•		
Zhang, Yang et Chen (2016)						•		
Zhou <i>et al.</i> (2016)	•							

A. Volume

Alors que, traditionnellement, la prise de décision au sein des organisations se confrontait au manque de données, à des données incomplètes, insuffisantes et désuètes (Zhou *et al.*, 2016), les organisations disposent actuellement d'énormes données à exploiter. La disponibilité de ces données a pour origine la numérisation et l'utilisation croissante de technologies (Xiong, Zheng et Li, 2014) de même que la volonté de « tout mesurer ».

En août 2010, l'ancien CEO de Google Éric Schmidt déclarait « il y avait seulement 5 exabytes d'informations créées entre l'aube de la civilisation et l'année 2003, mais cette quantité d'information est maintenant créée tous les 2 jours et le rythme ne fait qu'augmenter » (Galizia, Gee et Landis, 2011 : 16). Un rapport de l'International Data Corporation, publié en juin 2011, indiquait que le volume mondial de données devrait augmenter de près de 50% par année et il prévoyait que ce volume dépasserait les 8 000 exabytes en 2015 (Gantz et Reinsel, 2011 : 3). Par ailleurs, selon les estimations de la firme McKinsey, les données stockées en 2010 par les entreprises ont dépassé les 7 exabytes et celles générées par les individus ont dépassé les 6 exabytes (Manyika *et al.*, 2011 : 15).

Pour mettre ces données en perspective, un exabyte de données est équivalent à 4 000 fois les données de la Bibliothèque du Congrès (estimées à 235 terabytes en avril 2011), qui archive à la fois le contenu en ligne et hors ligne (Manyika *et al.*, 2011 : 3). Ceci souligne l'omniprésence des mégadonnées au sein des organisations de même que dans les activités des individus.

Nous constatons également qu'avec l'émergence des mégadonnées, un nouveau vocabulaire est apparu. Le [tableau 6](#) ci-dessous illustre les unités de mesure des volumes de données.

Tableau 6 : Unités de mesure des volumes de données (The Economist, 2010)

Unit	Size	What it means
Bit (b)	1 or 0	Short for "binary digit", after the binary code (1 or 0) computers use to store and process data
Byte (B)	8 bits	Enough information to create an English letter or number in computer code. It is the basic unit of computing
Kilobyte (KB)	1,000, or 2^{10} , bytes	From "thousand" in Greek. One page of typed text is 2KB
Megabyte (MB)	1,000KB; 2^{20} bytes	From "large" in Greek. The complete works of Shakespeare total 5MB. A typical pop song is about 4MB
Gigabyte (GB)	1,000MB; 2^{30} bytes	From "giant" in Greek. A two-hour film can be compressed into 1-2GB
Terabyte (TB)	1,000GB; 2^{40} bytes	From "monster" in Greek. All the catalogued books in America's Library of Congress total 15TB
Petabyte (PB)	1,000TB; 2^{50} bytes	All letters delivered by America's postal service this year will amount to around 5PB. Google processes around 1PB every hour
Exabyte (EB)	1,000PB; 2^{60} bytes	Equivalent to 10 billion copies of <i>The Economist</i>
Zettabyte (ZB)	1,000EB; 2^{70} bytes	The total amount of information in existence this year is forecast to be around 1.2ZB
Yottabyte (YB)	1,000ZB; 2^{80} bytes	Currently too big to imagine

The prefixes are set by an intergovernmental group, the International Bureau of Weights and Measures. Yotta and Zetta were added in 1991; terms for larger amounts have yet to be established.

Source: *The Economist*

En plus de ces volumes de données, Zuboff (2015) souligne quelques faits importants : Google traite environ quatre milliards de requêtes de recherches par jour, traite 20 petabytes de données par jour et prévoyait, en 2009, une capacité de dix millions de serveurs. Le volume de données chargées sur YouTube représente 72h d'enregistrement vidéo chaque minute ». Facebook compte plus d'un milliard d'utilisateurs actifs et déclare, en 2012, détenir un stock de 100 petabytes de photos et de vidéos sur ses serveurs et que sept milliards de photos sont chargées chaque mois sur sa plateforme.

La disponibilité de grandes quantités de données fait en sorte qu'il existe un avantage majeur à les traiter (Lycett, 2013), mais pose d'importants défis de stockage et de gestion et de gouvernance des données (Chauhan, Agarwal et Kar, 2016). À titre d'illustration, le US Defense Advanced Research Project Agency (DARPA) a lancé, en 2012, le programme XDATA qui mobilise des mathématiciens, des informaticiens et des spécialistes en visualisation afin de trouver des solutions en réponse aux défis imposés par les mégadonnées (H. C. Chen, Chiang et Storey, 2012). Aussi, des solutions de stockage en ligne sont de plus en plus accessibles, telles que les serveurs infonuagiques.

B. Variété

Par définition, les mégadonnées sont l'agrégation d'un nombre varié de jeux de données. La variété fait référence ici au fait que les données collectées peuvent être de nature très diverse, peuvent prendre des formes variées et hétérogènes et peuvent être structurées ou non structurées (H. C. Chen, Chiang et Storey, 2012).

Les systèmes traditionnels de collecte de données ont été rendus possibles notamment grâce à des bases de données relationnelles, à l'entrée systématique de données et à leur traitement. De plus, les données qui y sont stockées sont structurées, en opposition aux données non structurées qui réfèrent à des jeux de données autres qu'alphanumériques (par exemple : texte, vidéo, image et audio) (Lycett, 2013). Toutefois, avec l'émergence des mégadonnées, la majorité des données que les organisations collectent, créent et stockent n'est pas structurée (Bhimani, 2015 ; D. Chen, Preston et Swink, 2015) et les techniques et outils traditionnels démontrent leurs limites (Constantiou et Kallinikos, 2015).

Désormais, les organisations traitent des données structurées, semi-structurées et non structurées, provenant de l'interne ou de l'externe (Mergel, Rethemeyer et Isett, 2016). L'utilisation efficace de la variété des données présente à la fois des opportunités et des défis (Abbasi, Sarker et Chiang, 2016). D'un côté, l'analyse des données non structurées peut engendrer des décisions qu'il n'est possible de prendre qu'en utilisant les mégadonnées (Goes, 2014). Pour certaines organisations, « de l'information stratégique commence à arriver à travers des canaux non structurés » (LaValle *et al.*, 2011). De l'autre côté, certaines considérations doivent être prises en compte et comprennent le degré de perte de l'information lors des activités de nettoyage et de l'intégration sémantique des données (Lycett, 2013) et la perte de l'information résultant de l'isolation des données de leur contexte vu qu'elles résultent de fusion de données provenant de sources différentes (Chauhan, Agarwal et Kar, 2016).

C. Vitesse

La vitesse est associée à la vitesse des flux de production, de collecte et d'analyse des données. Elle réfère donc à deux dimensions : au fait, d'une part, que les données sont générées d'une manière

continue et, d'autre part, au fait qu'elles doivent être traitées en temps réel pour qu'elles soient consommées (utilisées) au fur et à mesure qu'elles sont produites. Ainsi, des décisions instantanées peuvent être prises au besoin.

La vitesse croissante de génération des mégadonnées est grandement attribuée à l'émergence de dispositifs connectés (Constantiou et Kallinikos, 2015). Les données ainsi générées doivent être converties en information utile en temps réel ou en laps de temps proche du temps réel (Goes, 2014), car souvent, la donnée a une durée de vie très courte et certaines décisions ne sont pertinentes que si elles sont prises en temps réel. Une grande partie de ce qui relève de la pertinence des mégadonnées implique des jeux de données dont la valeur est intimement liée à leur mise à jour et à leur fréquence d'actualisation (Constantiou et Kallinikos, 2015).

La vélocité des données et la tendance à la prise de décision axée sur les données ont créé un changement de paradigme dans la façon dont les organisations créent et exploitent les connaissances pour la prise de décision (Abbasi, Sarker et Chiang, 2016). Ce changement est attribuable au fait que « les mégadonnées et la complexité de marchés d'aujourd'hui exigent que les dirigeants prennent des décisions plus rapidement que jamais » (Kiron *et al.*, 2012 : 5). On privilégie ainsi le présent au détriment du passé et du futur et ceci se reflète inévitablement dans la planification et la gestion des organisations (Constantiou et Kallinikos, 2015 : 52).

Cette actualisation continue des données a des avantages, mais a aussi des limites. Constantiou et Kallinikos (2015) en soulignent deux : d'abord, les données disponibles connaissent un degré d'obsolescence très accéléré. De plus, l'horizon temporel de la prise de décision est affecté vu que la distinction entre les décisions opérationnelles et les décisions stratégiques réside à la fois dans l'importance et dans la durée des conséquences qui en découlent. Les décisions stratégiques devraient plutôt se baser sur des données stables décrivant des conditions durables de l'environnement de l'organisation (Lycett, 2013).

D. Véracité

La véracité (fiabilité) désigne la dimension qualitative des données et qui est nécessaire au bon fonctionnement des outils de traitement et à une prise de décision éclairée. Valider la véracité des données, analyser la pertinence, filtrer le bruit et les données malveillantes à partir de grands jeux de données constituent donc des activités d'une extrême importance (Goes, 2014) et le principe de GIGO (Garbage In, Garbage Out) reste encore pertinent dans l'ère des mégadonnées (Baesens *et al.*, 2016).

Généralement, l'objectif sous-jacent à une collecte des données détermine le choix des données à collecter, la granularité, et les compromis à réaliser entre la qualité des données et les coûts de collecte (Clarke, 2016 : 82). Comme nous allons le voir ci-dessous, l'utilisation des mégadonnées est souvent étrangère aux raisons pour lesquelles elles sont collectées. De nombreux problèmes de qualité peuvent donc être observés et sont exacerbés par la perte de contexte entourant les données et par la consolidation de plusieurs jeux de données. Ceci augmente considérablement la probabilité d'une mauvaise interprétation des données Brynjolfsson, Geva et Reichman (2016).

Par ailleurs, certains problèmes liés à la qualité des données résultent du fait que les données subissent un nombre important de manipulations afin de les rendre exploitables. Clarke (2016) indique que les mégadonnées résultent souvent d'une consolidation de données provenant de diverses sources et que les méthodes utilisées dans la collecte sont différentes. Aussi, l'acte de collecte implique des activités d'élimination des données par compression, échantillonnage ou filtrage. De plus, d'autres activités, telle que l'interpolation des données manquantes sont souvent menées. Ces actions affectent la précision, l'exhaustivité, l'exactitude et l'intégrité des données et les coûts nécessaires pour l'audit de la qualité des données peuvent s'avérer trop onéreux. Il existe des exemples qui mettent en évidence la question de la véracité et les conséquences de la mauvaise qualité des données : on estime que 20% des comptes sur les médias sociaux sont de faux comptes (Abbasi, Sarker et Chiang, 2016) et dans certains cas, des poursuites administratives en matière de fraude sont compromises par des doutes sur la véracité des données sur lesquelles les preuves sont fondées (Clarke, 2016).

E. Valeur

Les données d'une manière générale ne sont d'aucune utilité jusqu'à ce qu'ils soient dans une forme utilisable (Ackoff, 1989). Cette règle s'applique également aux mégadonnées qui, sans l'application de techniques d'analyse, ce ne sont que du bruit (Boyd et Crawford, 2012 : 667).

La chaîne de valeur de l'information est « l'ensemble des activités qui permettent de convertir des données en information et l'information en connaissance utile pour la prise de décision » (Abbasi, Sarker et Chiang, 2016 : iii). Ultimement, c'est la mise en œuvre d'actions qui découlent de cette prise de décision qui génère de la valeur (Sharma, Mithas et Kankanhalli, 2014). L'apport des mégadonnées réside notamment dans le fait qu'elles permettent de détecter de petits effets, l'étude de relations complexes, la comparaison entre des sous-échantillons, l'incorporation de plusieurs variables et l'étude d'effets rares (Lycett, 2013).

Chaque étape de la chaîne de valeur de l'information mobilise des personnes, des processus et des technologies (Abbasi, Sarker et Chiang, 2016). L'analyse des données concerne les étapes supérieures de la hiérarchie DIKW (Goes, 2014). Pearlson et Saunders (2004) et Awad et Ghaziri (2004) suggèrent que l'apport de l'humain augmente en montant dans les niveaux supérieurs de la hiérarchie DIKW, tandis que le rôle de la technologie descend. Ainsi, des enquêtes de terrains démontrent que le fait de disposer de capacités d'analyse effective constitue un avantage relatif des organisations performantes (LaValle *et al.*, 2011). En ce qui concerne les processus, Baesens *et al.* (2016) souligne que plusieurs termes liés à l'analyse des données sont utilisés : intelligence d'affaires, Data Science et analytique d'affaires. L'auteur indique que cette dernière pratique est la plus appropriée aux mégadonnées vu qu'elle permet de gérer l'ensemble du cycle de vie de la prise de décision. Enfin, les mégadonnées nécessitent des outils analytiques avancés (Müller *et al.*, 2016) qui aident à transformer les données brutes en informations contextuelles ou en connaissances (Lycett, 2013).

Désormais, l'information provenant de l'analyse des mégadonnées est de plus en plus utilisée dans les organisations afin de guider à la fois les stratégies futures et les opérations courantes (Lycett, 2013). Cependant, la création de la valeur à partir des mégadonnées soulève quelques questions. Ainsi, l'application mécanique des méthodes statistiques ou des analyses avancées aux mégadonnées pose

des pièges potentiels (Müller *et al.*, 2016) et la création de valeur découle plutôt d'une étroite collaboration des analystes avec les décideurs (Sharma, Mithas et Kankanhalli, 2014). Ceci s'ajoute aux défis « techniques » liés à l'inefficacité dans la manipulation des jeux de données, aux difficultés à traiter à temps les données collectées et aux coûts onéreux de traitement des données.

Enfin un des aspects le moins étudiés dans la littérature sur les mégadonnées c'est leur visualisation. Lycett (2013) indique que la sophistication de la visualisation des données augmente parallèlement à la nécessité de présenter des données plus complexes, de manière plus esthétique et plus instructive, à la fois rapidement et clairement. L'auteur souligne les outils de visualisation sont de plus en plus appelés à être dynamiques, interactives et utilisés en libre-service (par l'utilisateur final).

F. Agnostique

Une des caractéristiques déterminantes des mégadonnées est le fait qu'elles sont collectées sans but prédéfini. Ainsi, et contrairement aux données traditionnelles, les mégadonnées sont souvent recueillies sans but préalablement défini, ceci en espérant qu'un jour elles seraient utiles (Constantiou et Kallinikos, 2015; Müller *et al.*, 2016).

Avec l'émergence des mégadonnées, un changement de paradigme s'est opéré. Traditionnellement, les pratiques de gestion et de gouvernance reposent sur des données structurées et intentionnellement collectées afin de supporter le processus de prise de décision (Bergamaschi *et al.*, 2016; Constantiou et Kallinikos, 2015). Aujourd'hui, le schéma global qui est mis en évidence peut être résumé comme suit « first data then search for any possible uses of what is already available as data » (Constantiou et Kallinikos, 2015 : 51).

Les mégadonnées sont un produit dérivé de nos progrès technologiques (Müller *et al.*, 2016) . Une grande partie des mégadonnées existantes résulte d'un comportement aléatoire de la foule des utilisateurs, motivée par des objectifs génériques (Constantiou et Kallinikos, 2015) et sont capturés pour des raisons purement commerciales ou techniques, par exemple la liste d'authentification à une plateforme Web (Mergel, Rethemeyer et Isett, 2016). Les mégadonnées reposent donc, en partie, sur les empreintes et traces numériques (Digital Exhaust) laissées par les utilisateurs lors de leurs diverses activités sur internet.

De telles données brutes ne peuvent pas être directement utilisées d'autant plus qu'elles ne reposent souvent pas sur des constructions précises, valides et fiables (Lazer *et al.*, 2014). La pertinence de ces données est donc remise en question (Müller *et al.*, 2016) même si le PDG d'Amazon Jeff Bezos est d'ailleurs connu pour avoir déclaré « nous ne jetons jamais de données » (Davenport et Kim, 2013 : 8).

Dans un tel contexte, les données de ce type ont souvent peu de valeur intrinsèque. Il n'est donc pas surprenant que ce soit avec leur agrégation que ces données produisent de l'information pertinente (Constantiou et Kallinikos, 2015). L'agrégation occasionne l'analyse et le traitement de données en faisant abstraction des spécificités du contexte dans lequel elles ont été produites ou collectées (Lycett, 2013).

G. Confidentialité et vie privée

Au sein des villes, les problèmes de confidentialité deviennent de plus en plus évidents. Ceci résulte du fait qu'un grand nombre de données générées et échangées par les différents acteurs sont de nature privée ou personnelle, telles que les données détenues par les gouvernements, les informations exclusives des entreprises ou les données personnelles des citoyens (Zhang, Yang et Chen, 2016). Ainsi, la majeure partie des sujets abordant les inconvénients des mégadonnées se concentrent aujourd'hui sur les menaces qui pèsent sur la confidentialité des informations et sur la vie privée (Markus, 2015). Ces menaces peuvent entraîner des conséquences dommageables pour les personnes, telles que le vol d'identité, la cyber-intimidation et la discrimination entre les personnes (Bergamaschi *et al.*, 2016).

Les menaces d'atteinte à la vie privée prennent différentes formes : les menaces de confidentialité, organisationnelles, les menaces d'accessibilité et les menaces sociales (Spiekermann et Korunovska, 2016 : 5). Les menaces organisationnelles se rapportent aux risques associés aux utilisations secondaires des données des personnes. Les problèmes d'accessibilité concernent des degrés indésirables d'exposition de ces données. Enfin, les préoccupations sociales concernent les données qui peuvent provoquer des actes de cyber-harcèlement et d'intimidation.

Les analystes, les investisseurs et les entrepreneurs ont reconnu la valeur des données personnelles. D'ailleurs, Meglena Kuneva, ancienne commissaire aux consommateurs de l'Union européenne, indiquait dans un de ses discours : « les données personnelles sont le nouveau pétrole d'internet et la nouvelle monnaie du monde numérique » (Kuneva, 2009 : 2). Il s'agit ici de l'expression d'une réalité économique de plus en plus manifeste à l'échelle mondiale: les données personnelles émergent en tant qu'actif, négocié sur un marché de données en plein essor (Spiekermann et Korunovska, 2016).

Les spécialistes du marketing collectent des données personnelles en échange de services gratuits (Clarke, 2016). Pourtant, les gens ordinaires n'en sont guère conscients, car la plupart ne lisent pas les conditions concernant l'utilisation de ces services. D'ailleurs, un rapport de la Maison-Blanche américaine soulignait que « seul dans un monde de fantaisie, les utilisateurs vont effectivement lire ces conditions et comprendre leurs implications ...» (cité par Sugimoto, Ekbia et Mattioli (2016 : 12)). Dans cette optique, la recherche a identifié deux paradoxes importants (van Zoonen, 2016 : 474). Premièrement, malgré les inquiétudes exprimées par les gens au sujet de leur vie privée, il y a un manque de comportement approprié envers la sécurité. Deuxièmement, le sentiment d'avoir le contrôle sur les données provoque moins d'inquiétudes quant à la façon dont les données sont utilisées plus tard par d'autres parties.

Les éléments indiqués ci-dessus indiquent l'importance d'apporter une attention particulière aux aspects liés à la confidentialité et à la vie privée des citoyens. D'ailleurs, les gouvernements locaux ouvrent de plus en plus l'accès aux données au public, ce qui ne manque pas de soulever plusieurs questions (van Zoonen, 2016). De plus, les services électroniques collectent des données de manière si répandue (ubiquitaire) et sont tellement interconnectés qu'il est difficile de contrôler qui transfère et traite les données à des fins de ce type (Spiekermann et Korunovska, 2016). Enfin, Kitchin (2014 : 11-12) fait valoir que la collecte ubiquitaire de données au sein des villes entraîne la mise en œuvre de « systèmes qui cherchent à permettre des modes de gouvernance plus efficaces, ce qui peut se faire au détriment des droits à la vie privée, de la confidentialité et de la liberté d'expression ».

H. Dynamiques humaines

La disponibilité des quantités de données fait en sorte qu'il existe un avantage majeur à les traiter. La thèse analytique sous-jacente étant qu'avoir « plus de données est plus efficace que les meilleurs modèles statistiques » (Lycett, 2013 : 381). Toutefois, les mégadonnées peuvent avoir des impacts négatifs, car toutes les technologies ont des « effets secondaires » durables (Markus, 2015 : 58).

Ainsi, les mégadonnées, combinées aux algorithmes évolutifs, ont un impact profond sur les personnes et les procédés à un tel point que ces effets s'étendent à la science du comportement, à l'économie et aux sciences de la gestion (Abbasi, Sarker et Chiang, 2016; Janssen et Kuk, 2016b). Comme l'illustre ironiquement Chris Anderson (2008 : pas de no. de page), « avec des modèles mathématiques appliqués aux mégadonnées, tout devient obsolète, des théories du comportement humain à la linguistique en passant par la sociologie... les pétaoctets permettront de dire: la corrélation est suffisante ».

En raison de la disponibilité de micro-données provenant de différentes sources, les algorithmes permettent de modéliser les individus à un niveau très détaillé (Goes, 2014). La capture précise du comportement individuel et des événements environnants permet également de repérer les tendances de la population et l'impact d'événements sur son comportement (Kallinikos et Constantiou, 2015). Ainsi, il est rapporté qu'en moyenne, les internautes effectuent une centaine de requêtes de recherche mensuellement (Loebbecke et Picot, 2015). Ces requêtes traduisent les intentions, les intérêts et les besoins des utilisateurs et permettent d'élaborer des prévisions et d'influencer la prise de décision.

Les mégadonnées ont également un impact sur les connaissances organisationnelles et sur l'apprentissage. Markus (2015) se demande si l'humain peut encore se réappropriier les activités profondément automatisées. Depuis une dizaine d'années, plusieurs écrits ont alimenté la littérature sur les effets de la délégation de la décision à des logiciels (Clarke, 2016). Constantiou et Kallinikos (2015) se questionnent sur l'effet des mégadonnées sur la stratégie des organisations et indiquent que la façon dont l'information est présentée affecte les schémas cognitifs des dirigeants et donc sur les décisions prises. McAfee et Brynjolfsson (2012 : 65) soulignent que la prise de décision basée sur les données remplace de plus en plus la règle de l'opinion de la personne la mieux payée (HiPPO : Highest Paid Person's Opinion).

À l'ère de l'accélération de la numérisation et de l'analytique avancée, les mégadonnées affectent tous les domaines de la vie et créent de nouvelles façons de travailler, de communiquer et de coopérer (Loebbecke et Picot, 2015). Ceci affecte les emplois de la connaissance « Knowledge jobs », tout comme l'automatisation a affecté le travail manuel (Markus, 2015). L'un des exemples qui pourraient être cités est celui des robots de tradings.

D'autres défis résultent d'un manque d'information sur le contexte dans lequel des données ont été générées ou collectées. Chaque unité de donnée est traitée de manière égalitaire indépendamment de l'auteur ou du contexte de sa création (Lycett, 2013). Les algorithmes seuls ne peuvent pas capturer toute la complexité des mégadonnées (Kallinikos et Constantiou, 2015). S'ils peuvent prédire des événements futurs ou inconnus, les algorithmes sont incapables de fournir des explications pour leurs prédictions (Menon et Sarkar, 2016).

Alternativement, certains auteurs considèrent qu'il faut plutôt profiter des opportunités offertes par cette caractéristique. Baensens *et al.* (2016) indiquent que la grande quantité de données sur les interactions humaines offre des possibilités jamais disponibles dans le monde physique et trouvent qu'il y a un grand bénéfice sociétal à en tirer. Clarke (2016) indique que la prise de décision ne devrait pas être automatisée et que les mégadonnées devraient plutôt assister les décideurs dans leur processus décisionnel. L'auteur souligne, toutefois, que les décideurs peuvent rencontrer des difficultés à saisir les détails, la qualité, la signification et la pertinence des données. Menon et Sarkar (2016) indiquent que le développement d'algorithmes évolutifs demeure un ingrédient de succès des mégadonnées. Cependant, les auteurs soulignent que l'évolutivité seule n'est pas une garantie de la pertinence de l'algorithme ni des résultats de traitement. Enfin, Müller *et al.* (2016) soutiennent que, sans analytique automatisée, les mégadonnées se transforment en un fardeau plutôt qu'une opportunité. Ils indiquent, toutefois, que ces algorithmes ne s'appliquent qu'aux données structurées, alors qu'on estime que les données non-structurées représentent près de 90% de l'ensemble des données créées (Gantz et Reinsel, 2011 : 2). L'analyse de ces données nécessiterait donc le recours à des approches qualitatives.

Les éléments indiqués ci-dessus indiquent la complexité des mégadonnées et ses impacts sur les diverses dimensions des personnes et des organisations. Les praticiens, de même que les travaux de recherche se sont surtout intéressés aux facettes techniques des mégadonnées et ne sont pas suffisamment attardés sur l'étude des effets sur les personnes et leurs environnements institutionnels et sociaux (Marchand et Peppard, 2013 : 7). Toutefois, la réussite des initiatives de mégadonnées dépend de la prise en compte de ces aspects.

Conclusion de la section :

Cette section a donc permis d'identifier et de discuter les caractéristiques les plus importantes des mégadonnées. Ces caractéristiques expliquent des aspects importants des mégadonnées et la mise en place de stratégies adaptées pour leur gestion et leur gouvernance au niveau des organisations ne peut être ignorée. Comme leur nom l'indique, les mégadonnées réfèrent généralement à de grands ensembles de données, générés et mis à disposition sur les diverses plateformes numériques (Kallinikos et Constantiou, 2015). Toutefois, le volume seul n'aurait jamais suffi à encapsuler la nouveauté du phénomène des mégadonnées (Boyd et Crawford, 2012). Ainsi, le nouveau paradigme entourant les mégadonnées vient plutôt de la combinaison de l'ensemble d'un ensemble d'attributs (Goes, 2014). Pris ensemble, ces attributs et les mécanismes par lesquels les mégadonnées sont générées et utilisées apportent des changements importants à plusieurs paradigmes enracinés (Constantiou et Kallinikos, 2015).

Les données générées au sein des villes répondent bien aux définitions des mégadonnées et à leurs diverses caractéristiques indiquées dans la présente section (Malomo et Sena, 2017). Si les mégadonnées procurent plusieurs avantages aux villes, plusieurs défis doivent être relevés pour une gestion efficace des mégadonnées : fiabilité des données et des acteurs internes et externes, des menaces permanentes sur la confidentialité, l'intégrité, l'accessibilité, la protection et la confidentialité des données, auxquels s'ajoute la faible maturité des technologies existantes (Popescul et Radu, 2016). La discussion ci-dessus nous aide à identifier les caractéristiques d'innovation qui influencent l'utilisation des mégadonnées par les villes. La prochaine section porte justement sur la conceptualisation de l'utilisation des TI par les villes.

2.3 Conception de l'utilisation des TI par les villes

Il est généralement admis que les bénéfices organisationnels de l'implantation d'une technologie ne se concrétisent qu'avec l'utilisation de celle-ci. D'ailleurs, plusieurs auteurs la considèrent comme un facteur de succès d'un projet TI. Ainsi, Nelson (2005) conceptualise le succès d'un projet TI en fonction de six (6) critères : respect du budget, respect de l'échéancier, respect de l'envergure, réalisation d'un apprentissage organisationnel, adoption de la solution et son utilisation.

Même si nous ne considérons pas les mégadonnées comme une technologie, nous nous baserons sur le concept de l'utilisation des TI dans la littérature afin de conceptualiser l'utilisation des mégadonnées dans un contexte de ville. Ainsi, dans un premier temps, nous aborderons les concepts liés à l'utilisation (les formes d'utilisation, l'utilisation post-adoption et l'utilisation efficace [effective use]). Dans un second temps, nous allons présenter le cadre conceptuel de l'utilisation des TI au niveau municipal sur lequel le modèle de recherche sera développé.

2.3.1 L'acceptation, l'utilisation continue et l'utilisation efficace

Le concept de l'utilisation est un sujet largement étudié dans la littérature sur les systèmes d'information. Il est d'ailleurs considéré comme étant un facteur clé de succès des projets TI (Nelson, 2005) et une condition essentielle à la réalisation des bénéfices escomptés de l'implantation d'une solution (DeLone et McLean, 2003).

Les travaux réalisés dans le domaine de la psychologie sociale, tels que les théories de l'action raisonnée (TRA) (Ajzen et Fishbein, 1975) et la théorie du comportement planifiée (TPB) (Ajzen, 1991) soutiennent que le comportement d'un individu repose sur des actions planifiées et raisonnées qui découlent d'une intention consciente de l'individu à l'égard de ce comportement. Cette intention d'utilisation est la pierre angulaire des modèles d'acceptation des technologies de l'information telles que le modèle d'acceptation des technologies (TAM) de Davis, Bagozzi et Warshaw (1989), la théorie de la diffusion des innovations (DOI) de Rogers (1995) et la théorie unifiée de l'acceptation et de l'utilisation des technologies (UTAUT) de Venkatesh *et al.* (2003).

De nombreuses études empiriques, menées dans différents contextes d'utilisation des technologies de l'information, confirment que les caractères, les croyances et les attributs d'un individu influencent son attitude qui va ainsi dicter ses intentions et ultimement ses comportements (Bhattacharjee et Lin, 2015). Il est donc démontré que l'intention d'utiliser une technologie est le déterminant le plus influent sur l'utilisation de la technologie.

En ce qui a trait aux antécédents individuels de l'utilisation, Venkatesh *et al.* (2003) ont effectué une revue complète. Parmi les principaux antécédents, on retrouve la perception d'utilité [PU] et la perception de la facilité de l'utilisation [PEOU] (Davis, Bagozzi et Warshaw, 1989), l'attente du niveau de performance [PE], l'influence sociale et les conditions de facilitation (Venkatesh *et al.*, 2003), la facilité d'utilisation, la compatibilité et l'image de l'individu (Moore et Benbasat, 1991), l'attitude envers l'utilisation et les habiletés envers l'utilisation (Shareef *et al.*, 2011), les normes subjectives et le contrôle perçu (Ajzen, 1991), la confiance dans le gouvernement, dans internet et dans le gouvernement électronique (Carter et Bélanger, 2005), la participation de l'utilisateur et l'appropriation psychologique (Barki, Paré et Sicotte, 2008).

Les travaux de Barki, Titah et Boffo (2007 : 15), ont permis de mieux comprendre le concept de l'utilisation d'un système d'information. Les auteurs avancent que les activités reliées à l'utilisation

d'un système d'information (ISURA) prennent trois (3) formes : les activités d'interaction directe avec la technologie (en vue d'accomplir une tâche), les activités d'adaptation de la technologie (en vue de configurer ou de personnaliser la technologie par exemple) et les activités d'adaptation des utilisateurs à la technologie (apprentissage de l'utilisation de la technologie).

L'acceptation d'une nouvelle technologie est une condition nécessaire, mais pas suffisante pour le succès d'un projet. Le succès d'une technologie doit permettre la réalisation de bénéfices qui ne se génèrent qu'avec une utilisation continue de la technologie (Seddon, 1997). Malgré le fait que l'acceptation d'une technologie et la continuité de son utilisation sont des comportements conceptuellement et temporellement distincts (Bhattacharjee et Lin, 2015) les travaux de recherche sur l'usage continu des TI se sont largement basés sur les mêmes théories que celles de l'acceptation (Ortiz de Guinea et Markus, 2009). La prémisse de ces études est que l'utilisation continue des TI est un comportement intentionnel, basé sur une décision consciente et planifiée de l'utilisateur (Bhattacharjee et Lin, 2015). Toutefois, les travaux de Limayem, Hirt et Christy (2007) suggèrent que l'habitude peut influencer l'utilisation continue. Ainsi, l'utilisation continue des TI peut être par habitude plutôt que par comportement intentionnel : « Plus un comportement est réalisé par habitude, moins élevée sera la charge cognitive de la personne qui la réalise » (Limayem, Hirt et Christy, 2007 : 720). En outre, les travaux d'Agarwal et Karahanna (2000) indiquent que l'utilisation continue des TI peut être basée sur des réactions affectives ou émotionnelles telles que l'absorption cognitive ou la satisfaction.

L'utilisation continue de la technologie n'est cependant pas suffisante pour générer les bénéfices pour l'organisation (Seddon, 1997). Faut-il encore que cette utilisation soit efficace, c'est-à-dire, « une utilisation qui contribue à atteindre les objectifs pour lesquels le système a été implanté » (Burton-Jones et Grange, 2013 : 633). Malgré l'importance de l'utilisation effective, très peu d'études ont été conduites sur le sujet (voir sur le sujet: DeSanctis et Poole (1994), Saga et Zmud (1994), Covi et Kling (1996), Chin, Gopal et Salisbury (1997), Palen, Salzman et Youngs (2000), Boudreau et Seligman (2006), Burton-Jones et Straub (2006), Paul A. Pavlou et El Sawy (2006), Barki, Titah et Boffo (2007), LeRouge, Hevner et Collins (2007), Paul A Pavlou, Dimoka et Housel (2008) et Agarwal *et al.* (2010)).

L'étude de Li, Hsieh et Rai (2013) est l'une des rares à discuter de l'utilisation individuelle post-adoption dans un contexte d'intelligence d'affaires. Les auteurs ont identifié deux comportements d'utilisation : l'utilisation de routine (RTN), qui désigne une utilisation standardisée et systématique, et l'utilisation innovante (INV) qui décrit la découverte par les utilisateurs de nouvelles façons d'utiliser l'innovation. Les auteurs proposent une conceptualisation du « Rich intrinsic motivation » (RIM) qui inclut la motivation intrinsèque vers l'accomplissement (IMap), la motivation intrinsèque à connaître (IMkw) et la motivation intrinsèque pour expérimenter la stimulation (IMst). Les implications de l'étude se résument ainsi : la perception d'utilité a un impact plus important sur RTN que sur RIM, IMkw et IMst ont un impact plus fort sur INV que PU et IMap, l'impact de l'innovation personnelle avec les TI (PIIT) a modéré positivement l'impact des trois construits du RIM et de INV.

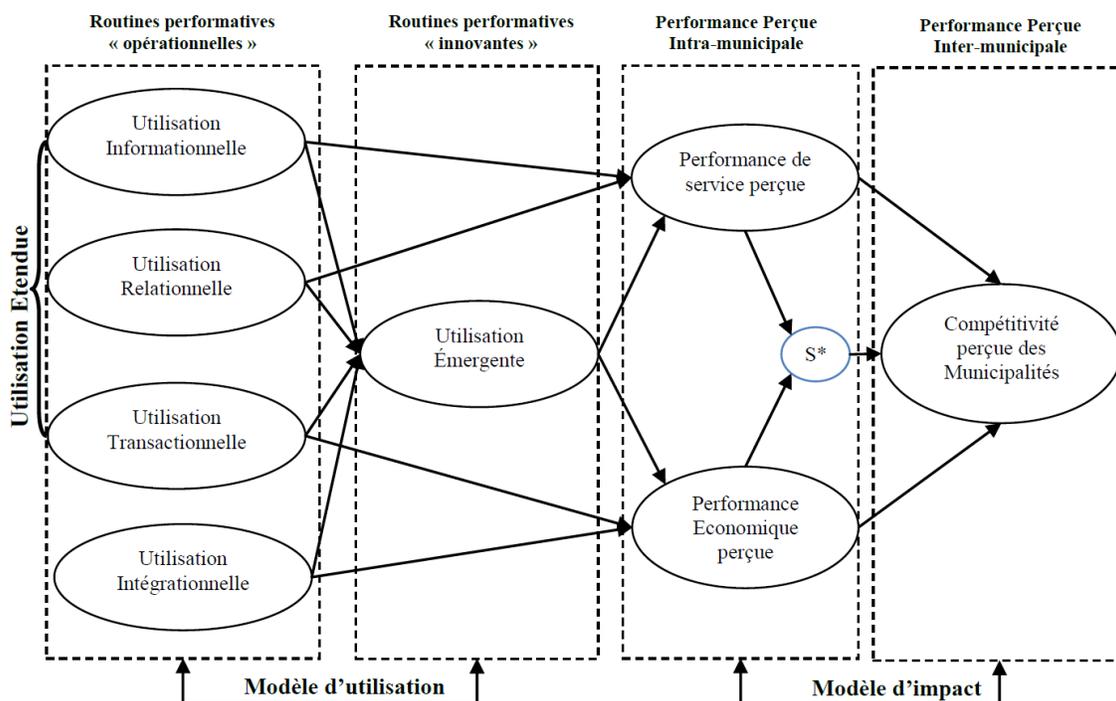
En guise de conclusion, l'utilisation est un construit à la fois important et complexe ; et la richesse des travaux sur le sujet permet à la présente étude de disposer d'un cadre théorique solide. Maintenant que nous avons passé en revue l'utilisation des systèmes d'information au niveau individuel, nous allons aborder l'utilisation au niveau organisationnel, et notamment les diverses formes de l'utilisation du gouvernement électronique au niveau des villes.

2.3.2 L'utilisation des TI par les villes

En dépit de l'importance de la conceptualisation de l'utilisation des technologies de l'information au niveau des villes, la littérature sur le sujet reste peu abondante (Titah et Barki, 2006) (voir sur le sujet : Grönlund (2004), Kim Viborg Anderson et Henriksen (2005), Grönlund et Horan (2005), Heeks et Bailur (2007), Titah (2010) Titah et Barki (2011)). En outre, à notre connaissance, il n'y a aucune étude sur la conceptualisation de l'utilisation des données ou des mégadonnées au niveau des villes, ou même au niveau des autres paliers du gouvernement.

De ce fait, et pour les besoins de la présente étude, nous allons utiliser les cinq (5) formes d'utilisation du modèle conceptuel d'utilisation et d'impact du gouvernement électronique (GE) au niveau municipal proposé par Titah (2010) (voir [figure 3](#) ci-dessous). Ce modèle est notamment basé sur la théorie des routines organisationnelles qui postule que les routines, conceptualisées comme étant les activités ou les processus d'affaires d'une ville, « déterminent la performance organisationnelle » (Titah, 2010 : 21). Ce modèle s'intéresse particulièrement à l'identification des routines exécutées via des systèmes de gouvernement électronique (modèle d'utilisation, partie gauche de la [figure 3](#)) dans le but d'être performant et de réaliser, ultimement, des bénéfices concurrentiels (modèle d'impact, partie droite de la [figure 3](#)).

Figure 3 : Modèle d'utilisation du GE au niveau municipal (Titah, 2010 : 21)



Note :

- L'interaction entre l'utilisation Intégrationnelle et (Utilisation Informationnelle + Utilisation Relationnelle + Utilisation Transactionnelle) n'est pas indiquée pour des considérations de clarté.
- * S = synergie

Pour les besoins de la présente étude, nous nous concentrerons sur l'identification des routines organisationnelles exécutées par les villes, ce qui correspond au modèle d'utilisation [partie gauche] du modèle global. Dans le modèle, l'utilisation des TI au niveau des villes est conceptualisée sous forme de deux types de routines organisationnelles : 1) les routines performatives opérationnelles regroupent l'utilisation informationnelle, l'utilisation relationnelle, l'utilisation transactionnelle et l'utilisation intégrationnelle. 2) les routines performatives innovantes regroupent les utilisations émergentes qui découlent des routines performatives opérationnelles. Le tableau ci-dessous présente les définitions proposées par (Titah, 2010 : 143-144) pour chacune des utilisations.

Tableau 7 : Définition des formes d'utilisation du GE [adapté de Titah (2010 : 143-144)]

Utilisation	Définition
Informationnelle	Processus qui permettent de fournir de l'information sur les différentes activités municipales telles que la diffusion de l'information à l'intention des citoyens en ce qui concerne les réunions du conseil municipal.
Relationnelle	Processus par lesquels la municipalité interagit avec les citoyens, comme par exemple, répondre à des questions ou à des demandes de citoyens par courriel.
Transactionnelle	Processus par lesquels les diverses transactions avec la municipalité peuvent être accomplies via internet, par exemple, le paiement en ligne des taxes et amendes.
Intégrationnelle	Processus par lesquels les activités de la municipalité sont intégrées avec d'autres institutions gouvernementales, comme par exemple, le partage d'informations ou de services avec d'autres paliers de gouvernement.
Émergente	Processus qui n'étaient pas réalisables ou reconnus avant l'exécution des quatre premières catégories de routines, mais qui étant susceptibles d'émerger, par exemple, la génération de sources de revenus alternatives et originales.

L'intérêt de ce modèle réside, notamment, dans le fait qu'il offre une conceptualisation complète et précise de l'utilisation des TI au niveau municipal. De plus, ce modèle a été testé auprès de Maires et hauts responsables municipaux de 221 municipalités au Canada et aux États-Unis et la majorité des hypothèses ont été vérifiées. Plus spécifiquement, la validité et la robustesse des indicateurs formatifs des cinq construits illustrant l'utilisation municipale du gouvernement électronique ont été validées (Titah, 2010 : 143-144). Il offre donc un cadre théorique pertinent pour le sujet de la présente étude.

En plus de ce cadre conceptuel, l'analyse et la catégorisation des antécédents de l'utilisation post-adoption des TI au niveau organisationnel sont nécessaires pour la construction du modèle de recherche de l'étude. C'est l'objet de la prochaine section de ce chapitre.

2.4 Antécédents organisationnels de l'utilisation des mégadonnées

Tel que mentionné plus haut, la littérature sur la conceptualisation de l'utilisation des TI au niveau des divers paliers gouvernementaux est très limitée. Afin d'identifier les antécédents de l'utilisation organisationnelle des mégadonnées par les villes, nous nous sommes basés sur deux perspectives théoriques qui relèvent d'autres champs d'application. Premièrement, la théorie de la diffusion de l'innovation de Rogers (2003) qui met l'accent sur les caractéristiques d'une innovation. Deuxièmement, le cadre technologie-organisation-environnement de Tornatzky et Fleischer (1990) qui met l'accent sur le contexte de l'utilisation d'une innovation.

Dans un premier temps, nous décrivons les principales caractéristiques de la théorie de la diffusion de l'innovation, puis celle du cadre technologie-organisation-environnement. Dans un second temps, nous analyserons les antécédents découlant de la littérature sur l'utilisation organisationnelle des TI.

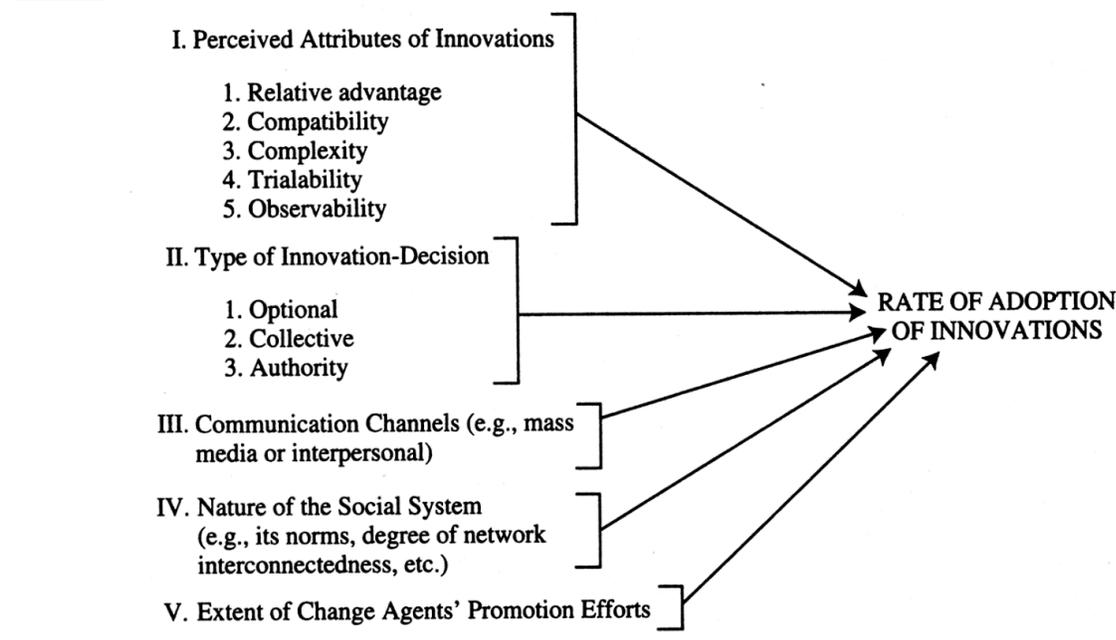
2.4.1 Perspectives théoriques

Deux cadres théoriques ont été mobilisés pour les besoins de la présente étude : la théorie de la diffusion de l'innovation (DOI) et le cadre technologie-organisation-environnement (TOE).

A. Théorie de la diffusion de l'innovation (DOI)

La théorie classique de la diffusion de l'innovation (DOI) attribue l'usage organisationnel d'une innovation à ses caractéristiques (Zhu *et al.*, 2006). Rogers (2003) considère que les innovations ne sont pas toutes équivalentes en termes de diffusion et que la perception de leurs caractéristiques explique leur vitesse de diffusion. Ainsi, il propose cinq (5) principales caractéristiques avec lesquelles une innovation peut être décrite : son avantage relatif, sa compatibilité, sa complexité, sa démontrabilité et son observabilité (figure 4).

Figure 4 : Déterminants de l'adoption des innovations (Rogers, 2003 : 222)



Le [tableau 8](#) indique les définitions de chacune de ces caractéristiques, tel que préconisé par Roger.

Tableau 8 : Caractéristiques de l'innovation selon Rogers (2003 : 15-16)

Caractéristique	Définition
Avantage relatif	La mesure dans laquelle une innovation peut apporter de meilleurs avantages que les précédentes innovations.
Compatibilité	La mesure dans laquelle une innovation est compatible avec les processus, les pratiques et les systèmes de valeurs existants.
Complexité	La mesure dans laquelle une innovation est difficile à comprendre ou à utiliser.
Démonstrabilité	La mesure dans laquelle une innovation peut être expérimentée (Triability).
Observabilité	La mesure dans laquelle les résultats d'une innovation sont visibles par les autres

L'avantage relatif fait référence à la mesure dans laquelle l'innovation est plus efficace, plus efficiente ou moins coûteuse que les pratiques existantes. La compatibilité est le degré auquel l'innovation est perçue comme étant conforme aux valeurs et croyances existantes, aux expériences passées, aux autres innovations et aux besoins. La complexité fait référence au degré auquel une innovation est perçue comme difficile à comprendre ou à utiliser. La démonstrabilité réfère à la réduction du degré de risque lié à l'adoption d'une innovation. L'observabilité réfère au fait que les résultats de certaines innovations peuvent être vus et constatés par les autres. Contrairement à la perception de complexité, la perception de l'avantage relatif, la compatibilité, l'observabilité et la démonstrabilité d'une innovation sont positivement corrélées à son taux d'adoption et de diffusion.

En plus des caractéristiques indiquées ci-dessous, les travaux de Tornatzky et Klein (1982) ont permis d'identifier cinq autres aspects à considérer : le coût, la communicabilité, la divisibilité, la rentabilité et l'approbation sociale. Toutefois, les travaux de ces auteurs indiquent que la compatibilité, l'avantage relatif, et la complexité sont les facteurs les plus significatifs pour la diffusion de l'innovation dans les organisations.

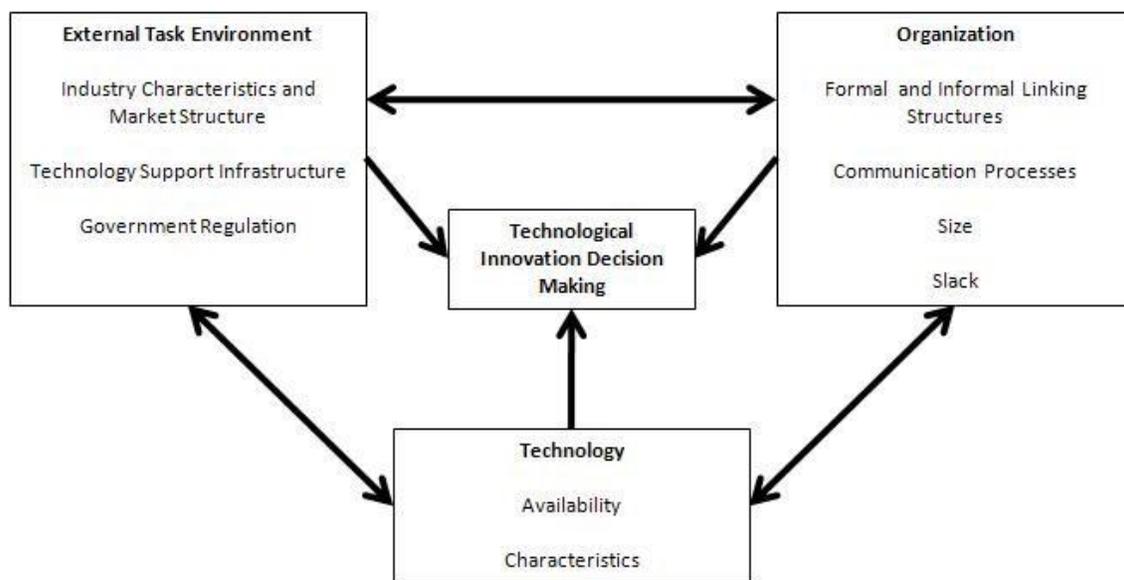
Publiée pour la première fois en 1962, la théorie DOI a été mise à l'épreuve à plusieurs reprises et la littérature démontre qu'elle dispose d'une base théorique solide et d'un soutien empirique cohérent. À titre d'illustration, voici quelques études ayant mobilisé la théorie DOI: Ostlund (1974), Moore et Benbasat (1991), Premkumar, Ramamurthy et Nilakanta (1994), Kendall *et al.* (2001), Beatty, Shim et Jones (2001) et Zhu *et al.* (2006). Enfin, la littérature montre que la théorie DOI a été utile pour l'étude d'une variété d'innovations en systèmes d'information (voir [tableau 10](#) ci-dessous).

Comme nous l'avons vu dans la section 2.2.2, les mégadonnées détiennent des caractéristiques distinctives et qui peuvent avoir des impacts sur l'adoption, la diffusion et l'utilisation de cette innovation. De ce fait, l'utilité de la théorie DOI réside, notamment, dans le fait qu'elle va nous permettre d'analyser avec attention ces caractéristiques et leur impact sur l'utilisation des mégadonnées par les villes.

B. Cadre Technologie-Organisation-Environnement (TOE)

Le contexte dans lequel les mégadonnées sont utilisées mérite d'être étudié, car leur utilisation peut différer d'un environnement à un autre. Dans leurs travaux sur l'étude du processus de l'adoption et de la diffusion des innovations technologiques, Tornatzky et Fleischer (1990) ont développé le cadre TOE qui constitue une perspective théorique importante pour l'étude des facteurs contextuels. Selon les auteurs, trois aspects peuvent influencer l'utilisation organisationnelle d'une innovation technologique : le contexte technologique, le contexte organisationnel et le contexte lié à l'environnement externe. La [figure 5](#) ci-dessous illustre le cadre TOE proposé par les auteurs.

Figure 5 : Cadre TOE (Tornatzky et Fleischer, 1990 : 153)



Le [tableau 9](#) ci-dessous indique la définition de chacune de ces trois composantes du cadre TOE.

Tableau 9 : Composantes du cadre TOE selon Tornatzky et Fleischer (1990 : 152-154)

Contexte	Définition
Technologique	Décrit les technologies en cours d'utilisation ou disponibles sur le marché, ainsi que les compétences techniques pertinentes disponibles au sein de l'organisation.
Organisationnel	Réfère aux caractéristiques internes de l'organisation telles que sa taille, sa localisation et sa culture organisationnelle.
Externe	C'est l'environnement externe dans lequel une organisation mène ses activités, tel que son industrie, ses concurrents et ses partenaires commerciaux.

Le contexte technologique décrit l'ensemble des technologies pertinentes pour l'organisation. Cela inclut les technologies existantes à l'intérieur de l'organisation, ainsi que toutes les technologies offertes sur le marché. Le contexte organisationnel est défini en fonction de plusieurs mesures descriptives : taille et portée de l'organisation, la centralisation, la formalisation et la complexité de sa structure de gestion, la qualité de ses ressources humaines et la quantité de ressources limitées disponibles au sein de l'organisation. Enfin, le contexte (environnement) externe est l'arène dans laquelle une organisation exerce ses activités : son industrie, ses concurrents, l'accès aux ressources et les relations avec le gouvernement (Tornatzky et Fleischer, 1990 : 152-154).

La revue de la littérature effectuée par Zhu, Kraemer et Xu (2003) a démontré l'utilité du cadre TOE pour la compréhension de la diffusion d'une innovation technologique. En outre, Chau et Tam (1997) indiquent, en conclusion de leurs travaux, que « cette étude démontre la valeur d'utiliser le cadre de Tornatzky et Fleischer dans la compréhension de l'adoption d'une innovation [système d'information complexe]... une implication pour de futurs travaux consiste à étendre le cadre proposé [TOE] à d'autres domaines de l'innovation » (Chau et Tam, 1997 : 17).

Le cadre TOE a d'ailleurs été empiriquement testé sur une variété d'innovations. À titre d'illustration, voici quelques exemples qui peuvent être cités : étude des antécédents de l'utilisation de l'échange électronique des données (EDI) par Iacovou, Benbasat et Dexter (1995), étude des logiciels libres par Chau et Tam (1997), étude des antécédents de l'adoption des systèmes de planification des besoins matériels (MRP) par Thong (1999), étude des antécédents de l'utilisation du commerce électronique par Zhu et Kraemer (2005) et étude des antécédents de l'utilisation de l'analytique des mégadonnées (BDA, Big Data Analytics) par D. Chen, Preston et Swink (2015).

Enfin, étant donné que chaque ville a un contexte qui lui est unique, et que ce contexte influence les aspects technologiques, organisationnels et politiques de la municipalité, l'étude de l'utilisation des mégadonnées par les villes peut être considérée comme une interaction contextuelle entre l'innovation technologique et l'innovation managériale et organisationnelle (Gil-Garcia, Zhang et Puron-Cid, 2016).

En s'appuyant sur les éléments indiqués ci-dessus, il s'avère que le cadre TOE est approprié pour étudier les antécédents de l'utilisation des mégadonnées dans un contexte de ville. Ceci est justifié par le fait que l'utilisation des mégadonnées par les villes est influencée par les trois facteurs du cadre.

C. Justification de l'intégration des deux perspectives

Comme nous allons le voir dans la prochaine sous-section (2.4.2), on retrouve dans la littérature des travaux sur l'utilisation d'une variété d'innovations au sein d'organisations opérant dans des contextes différents. Dans cette perspective, plusieurs théories ont été développées et mobilisées, ce qui a donné lieu à un large éventail d'antécédents qui peuvent avoir une incidence sur l'utilisation de l'innovation (Mishra, Konana et Barua, 2007). Toutefois, comme ces cadres théoriques n'étaient pas développés pour une technologie en particulier ou un contexte précis, ces perspectives théoriques se caractérisent par leur haut degré de flexibilité et de généralité. De ce fait, un effort d'adaptation est à prévoir.

Premièrement, la théorie DOI de Rogers a initialement été développée pour aborder la diffusion de l'innovation parmi les individus d'une population (Zhu *et al.*, 2006). Toujours selon les auteurs, elle doit être enrichie si l'accent se concentre sur l'étude d'innovations complexes par les organisations.

Selon Swanson (1994 : 1074), des variables représentant le contexte dans lequel l'innovation est réellement utilisée devraient être prises en compte. Dans ce sens, Chau et Tam (1997 : 3) indiquent que « bon nombre de résultats contradictoires sur l'innovation organisationnelle rapportés dans la littérature pourraient être attribués aux différences contextuelles de ces études. Ainsi, les décisions d'adoption de l'innovation doivent être étudiées dans des contextes appropriés et avec des variables adaptées à la spécificité de l'innovation ».

Deuxièmement, en dépit du fait que le cadre TOE a largement été mobilisé dans la littérature, du fait qu'il a empiriquement été testé et du fait que son utilité a été démontrée, Mishra, Konana et Barua (2007) indiquent que ce cadre ne représente pas une théorie solidement élaborée et ne fournit donc pas la base théorique permettant d'étudier les effets des variables, les unes sur les autres. Toujours selon ces auteurs, les théories individuelles, comme la DOI, n'ont pas la profondeur des construits qu'on retrouve dans le cadre TOE. Ceci justifie la nécessité ressentie de combiner le cadre TOE à d'autres théories dans la littérature (Alkhalil, Sahandi et John, 2017 ; D. Chen, Preston et Swink, 2015; Mishra, Konana et Barua, 2007; Zhu *et al.*, 2006; Zhu et Kraemer, 2005).

En outre, le cadre TOE est cohérent avec la théorie DOI de Rogers, dans laquelle l'emphase est mise sur les caractéristiques technologiques, internes et externes de l'organisation, en tant que moteurs de la diffusion de la technologie (Zhu et Kraemer, 2005 : 63). Toutefois, l'intégration des éléments de contexte de ce cadre contribue à renforcer ce qui a été généralement négligé dans la théorie DOI, c'est-à-dire, les circonstances technologiques et organisationnelles propres à une organisation, de même que les caractéristiques de son contexte externe (Zhu *et al.*, 2006).

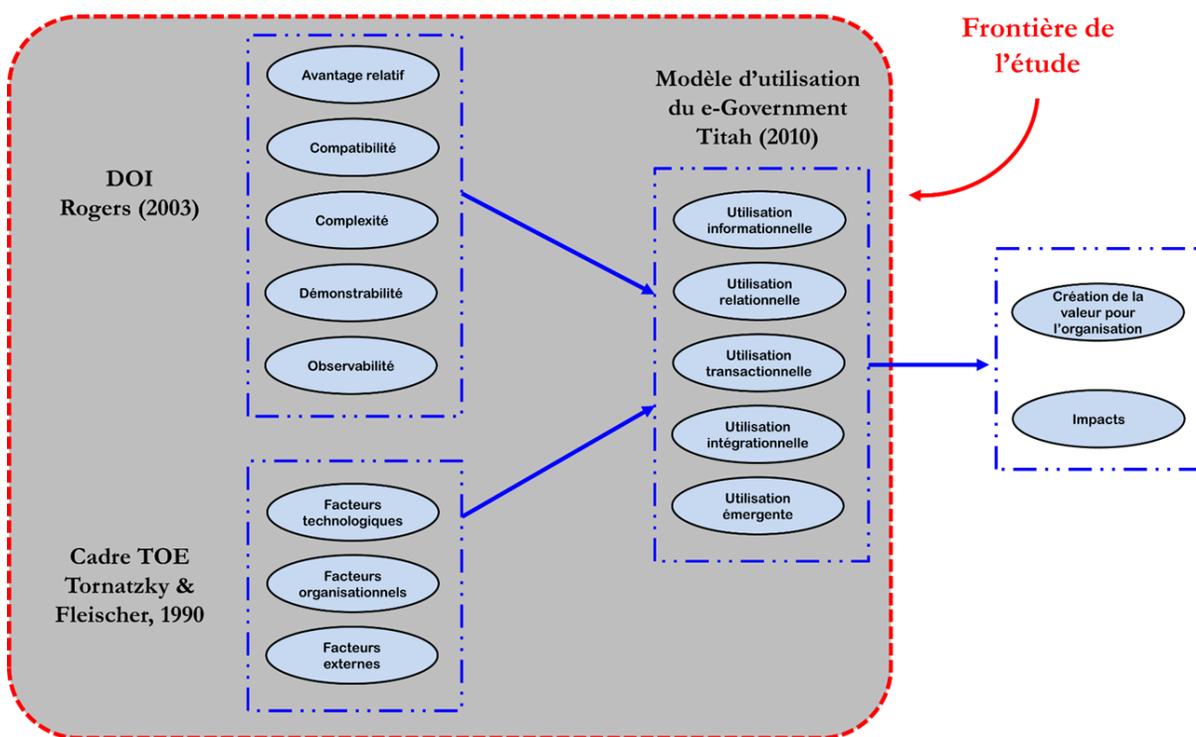
Par ailleurs, nous nous alignons avec Zhu et Kraemer (2005) et Zhu *et al.* (2006) pour mentionner que la nature de la technologie à l'étude nécessite la prise en compte du contexte de son utilisation. Ainsi, dans la littérature professionnelle comme dans la recherche, on retrouve plusieurs catégorisations des innovations. McAfee (2006) indique qu'il existe trois « mondes » de TI sollicitant chacun différentes capacités organisationnelles et exigeant divers types d'intervention et de gestion : « Fonction IT » qui comprend les technologies qui rendent efficaces l'exécution de tâches isolées, « Network IT » qui comprend les technologies qui permettent aux gens de collaborer, et enfin « Enterprise IT » qui est le type de technologies que les organisations adoptent pour structurer les interactions entre l'interne et l'externe. Dans la même optique, Swanson (1994) indique qu'il existe trois types d'innovations: le type 1 concerne les innovations techniques, ayant une envergure restreinte aux tâches fonctionnelles des systèmes d'information. Le type 2 couvre les innovations qui appliquent les systèmes d'information pour soutenir les activités des lignes d'affaires. Enfin, les innovations de type 3 permettent l'intégration des systèmes d'information dans le cœur métier de l'organisation. Selon ces deux typologies, nous considérons les mégadonnées comme une innovation « Enterprise IT », de type 3, en ce sens que leur utilisation transforme les lignes d'affaires en question, que c'est une innovation qui permet d'étendre les produits et services offerts par les villes et que les implications de leur utilisation dépassent les frontières de la municipalité. De ce fait, il est important de prendre en compte les caractéristiques distinctives des mégadonnées, analysées dans les sections ci-dessus, dans un contexte de ville, conceptualisé avec le cadre TOE.

En s'appuyant sur les éléments indiqués ci-dessus, il s'avère que la théorie DOI offre une base conceptuelle solide pour l'étude des antécédents liés aux caractéristiques des mégadonnées, largement analysés dans les sections ci-dessus. De plus, le cadre TOE est approprié pour étudier les antécédents du point de vue contextuel vu que l'environnement dans lequel les villes opèrent influe certainement

sur l'utilisation des mégadonnées. De ce fait, l'intégration des deux perspectives semble pertinente pour constituer une base conceptuelle solide pour la présente étude.

La [figure 6](#) ci-dessous synthétise les éléments discutés dans cette section, de même que ceux portant sur la conception de l'utilisation des TI par les villes (section 2.3). Nous allons maintenant procéder à l'analyse et à la catégorisation des antécédents de l'utilisation organisationnelle des TI (sous-section 2.4.2). La conclusion de ce chapitre vise à présenter le modèle de recherche sous-jacent à cette étude (section 2.5) et qui découle du modèle conceptuel présenté ci-dessous.

Figure 6: Cadre conceptuel de l'utilisation des mégadonnées par les villes



2.4.2 Catégorisation des antécédents de l'utilisation

À notre connaissance, il n'y a pas d'étude portant sur l'utilisation des mégadonnées par les villes dans la littérature. De ce fait, nous nous sommes penchés sur l'analyse des travaux ayant étudié les phases post-adoption du processus d'implantation des TI proposé par Kwon et Zmud (1987). En tout, une douzaine d'études ([tableau 10](#)) a été retenue et analysée pour les besoins des présents travaux.

Certains des articles analysés ont étudié les systèmes d'information d'une manière générale (Armstrong et Sambamurthy, 1999; Boynton, Zmud et Jacobs, 1994; Thong, 1999), tandis que d'autres auteurs ont analysé une technologie particulière telle que le commerce électronique (Lin, 2008; Zhu *et al.*, 2006; Zhu, Kraemer et Dedrick, 2004; Zhu et Kraemer, 2005), l'échange de données informatisées (EDI) (Iacovou, Benbasat et Dexter, 1995; Massetti et Zmud, 1996), l'internet (Mishra, Konana et Barua, 2007), les médias sociaux (Guillamón *et al.*, 2016) et l'analytique des mégadonnées (D. Chen, Preston et Swink, 2015).

En termes de perspectives théoriques, le cadre TOE a été mobilisé dans 5 des 12 études tandis que la théorie DOI a été mobilisée dans 4 études. Ces deux perspectives ont été utilisées conjointement dans une seule étude (Zhu *et al.*, 2006). En plus de la revue de la littérature, d'autres théories ont été mobilisées. Par ailleurs, hormis l'article de Guillamón *et al.* (2016), toutes les études retenues avaient l'entreprise comme contexte organisationnel d'analyse. Enfin, l'étude de Thong (1999) est la seule à analyser les phases pré-adoption et a été incluse dans la présente étude vu la pertinence des antécédents analysés. Le [tableau 10](#) synthétise l'ensemble des éléments de contexte des études indiquées ci-dessus.

Avant de procéder à la catégorisation des antécédents organisationnels de l'utilisation des TI, nous allons synthétiser les principales implications des études retenues. La revue détaillée de chaque article est indiquée dans l'[annexe 5](#).

- Dans l'étude d'Armstrong et Sambamurthy (1999), les connaissances d'affaires et en TI des CIO, le support de la haute direction et la sophistication de l'infrastructure technologique influent de manière significative sur l'assimilation technologique de l'organisation. La taille de l'organisation et les connaissances TI des membres de la haute direction n'ont pas d'influence significative.
- Dans l'étude de Boynton, Zmud et Jacobs (1994), le climat de gestion des TI influence positivement à la fois l'utilisation des TI et l'efficacité du processus de gestion des TI. Cette dernière n'a pas d'effet significatif. Contrairement aux conclusions d'Armstrong et Sambamurthy (1999), l'étude de Boynton, Zmud et Jacobs (1994) indique que les connaissances en TI des gestionnaires ont une influence (positive) sur l'utilisation.
- Dans l'étude de D. Chen, Preston et Swink (2015), l'ensemble des facteurs du cadre TOE ont une influence positive : les facteurs technologiques influent directement sur l'utilisation de l'analytique des mégadonnées, alors que les facteurs organisationnels (prédisposition organisationnelle) et environnementaux (pression concurrentielle) l'influencent indirectement par l'intermédiaire du soutien de la haute direction.
- Dans l'étude de Guillamón *et al.* (2016), le niveau de participation électronique, la taille de la population, le niveau de revenu des citoyens et le niveau d'endettement ont une incidence sur l'utilisation des médias sociaux (Facebook en particulier) par les gouvernements municipaux.

Tableau 10 : Revue de la littérature sur les antécédents de l'utilisation des TI

Source	Objet de l'étude	Contexte	Phase	Base théorique
Armstrong et Sambamurthy (1999)	Assimilation des TI	Entreprises américaines (moyenne et grande taille) de 8 secteurs d'activité différents	Post-adoption (assimilation)	Revue de la littérature sur la gestion des ressources et des connaissances
Boynton, Zmud et Jacobs (1994)	Utilisation des TI	Entreprises commerciales membres du GUIDE international	Post-adoption	Absorptive capacity theory
D. Chen, Preston et Swink (2015)	Big Data Analytics	Gestion de la chaîne d'approvisionnement d'entreprises américaines	Post-adoption	Dynamic capabilities theory TOE framework
Guillamón <i>et al.</i> (2016)	Médias sociaux	Gouvernement municipal utilisant Facebook en Italie et en Espagne	Post-adoption	Agency theory Neoinstitutional framework
Iacovou, Benbasat et Dexter (1995)	Échange des données informatisées (EDI)	PME de la Colombie-Britannique, fournisseurs du gouvernement provincial	Adoption et post-adoption	Diffusion of innovation theory
Lin (2008)	E-Business	Firmes taiwanaises de grande taille utilisant le commerce électronique	Intégration	Diffusion of innovation theory Organizational learning literature
Masseti et Zmud (1996)	Échange des données informatisées (EDI)	Grandes entreprises de fabrication de composants électroniques ou de prestation de services de télécom.	Post-adoption (value and benefits)	Revue de la littérature sur l'EDI
Mishra, Konana et Barua (2007)	Utilisation d'internet	Gestion de la chaîne d'approvisionnement d'entreprises américaines	Post-adoption	Resource-based view (RBV) TOE framework
Thong (1999)	Systèmes d'information	Entreprises de petite ou moyenne taille Singapour	Adoption	Diffusion of innovation theory
Zhu, Kraemer et Dedrick (2004)	E-Business	Entreprises du secteur des services financiers dans 10 pays	Post-adoption (création de la valeur)	TOE framework
Zhu et Kraemer (2005)	E-Business	Entreprises du commerce de détail dans 10 pays (dévelop. et en développement)	Post-adoption (actual use)	Resource-based theory TOE framework
Zhu <i>et al.</i> (2006)	E-Business	Entreprises de services, production, distribution et commerce de détail dans 6 pays européens	Post-adoption (usage)	Diffusion of innovation theory TOE framework

- Dans l'étude de Iacovou, Benbasat et Dexter (1995), les pressions externes (partenaires d'affaires) constituent le facteur le plus important dans l'adoption de l'EDI, alors qu'une grande prédisposition organisationnelle et une volonté de réaliser les bénéfices escomptés constituent des facteurs importants pour l'intégration et la concrétisation des impacts de l'adoption de l'EDI.
- Dans l'étude de Lin (2008), deux caractéristiques de l'innovation (avantage relatif et compatibilité perçus) et quatre capacités d'apprentissage organisationnel (engagement managérial, orientation des systèmes, acquisition de connaissances et diffusion de la connaissance) ont un effet important sur la réussite de la mise en œuvre du commerce électronique. L'effet de la complexité de l'innovation n'a pas été significatif.
- Dans l'étude de Mishra, Konana et Barua (2007), la numérisation des processus d'approvisionnement et la diversité des connaissances organisationnelles sur l'approvisionnement influent sur l'utilisation de l'internet. Aussi, l'infrastructure technologique et la numérisation des processus de vente des fournisseurs ont un impact positif sur l'utilisation de l'internet. Les perceptions sur l'incertitude technologique et sur l'incertitude liée aux volumes sont mitigées.
- Dans l'étude de Thong (1999), les caractéristiques du dirigeant (innovation et niveau de connaissance en TI) et celles de l'innovation (avantage relatif, compatibilité et complexité) sont des déterminants importants de l'adoption. Toutefois, ces caractéristiques n'ont pas d'influence sur l'étendue de l'adoption qui est principalement déterminée par les caractéristiques organisationnelles (taille de l'entreprise et niveau de connaissance des employés). Enfin, la concurrence n'a aucun effet direct sur l'adoption des SI.
- Dans l'étude de Zhu, Kraemer et Dedrick (2004), la prédisposition technologique est le facteur le plus significatif, suivi de la disponibilité de ressources financières, la portée mondiale des opérations et la disponibilité du support réglementaire. Alors que la taille de l'organisation a un effet négatif, l'effet des pressions compétitives n'est pas significatif.
- Dans l'étude de Zhu et Kraemer (2005), la compétence technologique, l'engagement financier, la pression concurrentielle et le soutien réglementaire ont une influence importante. La compétence technologique semble être le facteur le plus important (suivi du soutien réglementaire). La taille de l'organisation a un effet négatif tandis que la portée internationale est positive, mais statistiquement insignifiante.
- Dans l'étude de Zhu *et al.* (2006), l'ensemble des facteurs retenus ont été significatifs. Parmi les caractéristiques d'innovation, la compatibilité est le facteur (positif) le plus significatif, et les préoccupations de sécurité l'emportent sur le coût (négatif). Parmi les variables contextuelles, la compétence technologique, la prédisposition des partenaires et les pressions concurrentielles influencent considérablement l'utilisation du commerce électronique (positif). L'inertie structurelle des grandes entreprises tend à ralentir sa pénétration (négatif).

Le tableau [tableau 11](#) ci-dessous synthétise les antécédents des études analysées dans les présents travaux, de même que l'effet de chaque antécédent sur la variable dépendante. Il est à noter que seuls les antécédents significatifs sont présentés et que l'étude de Massetti et Zmud (1996) n'a pas été considérée vu que les antécédents proposés n'ont pas été empiriquement testés.

Tableau 11 : Antécédents significatifs de l'utilisation des TI

Source	Effet	Antécédents significatifs
Armstrong et Sambamurthy (1999)	Positif Positif	Quality of senior leadership Sophistication of IT infrastructures
Boynton, Zmud et Jacobs (1994)	Positif Positif	IT management Climate Managerial IT Knowledge
D. Chen, Preston et Swink (2015)	Positif Positif Positif Positif Positif	Expected Benefits Technology Compatibility Organizational Readiness (sur le TMT Support) Competitive Pressure (sur le TMT Support) TMT Support
Guillamón <i>et al.</i> (2016)	Positif Positif Positif Positif	Socio-economic determinants: E-participation Index Socio-economic determinants: Citizens' income level Socio-economic determinants: Population size Financial determinants: Debt level of the municipality
Iacovou, Benbasat et Dexter (1995)	Positif Positif Positif	Niveau de prédisposition (sur l'adoption) Pressions externes (sur l'impact) Bénéfices perçus (sur l'impact)
Lin (2008)	Positif Positif Positif Positif Positif Positif	Perceived relative advantage, Compatibility Managerial commitment Systems orientation Knowledge acquisition Knowledge dissemination
Mishra, Konana et Barua (2007)	Positif Positif Positif Mitigé Mitigé	Procurement-process digitalization Diversity of organizational procurement knowledge Suppliers' sales-process digitalization Organizational perceptions of technological uncertainty Organizational perceptions of volume uncertainty
Thong (1999)	Positif Positif Positif Positif Négatif Positif Positif Positif	CEO characteristics: Innovativeness CEO characteristics: Level of IS knowledge Innovation characteristics: Relative advantage Innovation characteristics: Compatibility Innovation characteristics: Complexity of IS Organizational characteristics: business size Organizational characteristics: Level of employees' IS knowledge Organizational characteristics: Informational intensity
Zhu, Kraemer et Dedrick (2004)	Positif Négatif Positif Positif Positif	Technology readiness Firm size Global scope Financial resources Regulatory environment
Zhu et Kraemer (2005)	Positif Négatif Positif Positif Positif	Technology Competence Organization Size Financial commitment Competitive pressure Regulatory support
Zhu <i>et al.</i> (2006)	Positif Positif Négatif Négatif Positif Négatif Positif Positif	Caractéristiques de l'innovation : Avantage relatif Caractéristiques de l'innovation : Compatibilité Caractéristiques de l'innovation : Coûts Caractéristiques de l'innovation : Préoccupations de sécurité Facteurs contextuels : Compétences technologiques Facteurs contextuels : Taille de l'organisation Facteurs contextuels : Pressions compétitives Facteurs contextuels : Prédisposition des partenaires

De l'analyse des conclusions des études indiquées dans le [tableau 10](#) ci-dessus, nous avons pu tirer certains constats. D'abord, la littérature s'est largement intéressée au contexte dans lequel opèrent les organisations. Voici les principaux constats liés à ces aspects :

- Le support de la haute direction est un antécédent qui a largement été pris en compte dans ces études, et ce sous plusieurs formes. Certaines études s'intéressent aux effets des caractéristiques des dirigeants sur leur engagement ou sur les étapes post-adoption. Ainsi, Armstrong et Sambamurthy (1999) introduisent la notion de la qualité du leadership de la haute direction. Boynton, Zmud et Jacobs (1994) se sont intéressés aux connaissances en TI de la haute direction, de même qu'au climat de gestion. (Thong, 1999) s'est également intéressé aux effets des connaissances TI de la haute direction, en plus de l'effet de l'innovation (innovativeness) du CEO sur l'adoption. Les autres études discutent de l'engagement des membres dirigeants. Ainsi, dans l'étude de Lin (2008), on s'intéresse à la notion de l'engagement de la haute direction. Dans l'étude de D. Chen, Preston et Swink (2015), la prédisposition organisationnelle et les pressions compétitives ont un effet positif sur le support de la haute direction qui lui-même impacte positivement l'utilisation de l'analytique des mégadonnées. En somme, l'ensemble de ces formes de support de la haute direction a un effet positif sur les étapes post-adoption.
- On retrouve deux aspects selon lesquels les compétences en TI des organisations sont étudiées. Le premier aspect concerne l'impact de la sophistication de l'infrastructure TI (Armstrong et Sambamurthy, 1999) ou de la numérisation du processus (Mishra, Konana et Barua, 2007 ; Zhu, Kraemer et Dedrick, 2004) alors que le second s'intéresse au niveau de connaissances en TI des employés (Thong, 1999). Dans les études de Zhu et Kraemer (2005) et de Zhu *et al.* (2006), ces deux aspects sont réunis sous le même construit (Technology Competence) qui a un effet positif sur l'utilisation.
- L'effet des pressions compétitives a largement été discuté dans la littérature. Il a un impact positif dans les études de Iacovou, Benbasat et Dexter (1995), de Zhu *et al.* (2006) et de Zhu et Kraemer (2005), alors que l'impact n'a pas été significatif dans les études de Thong (1999) et Zhu, Kraemer et Dedrick (2004).
- Dans les études analysées, les considérations financières se sont traduites sous deux formes. Si certaines études discutent de la disponibilité de ressources financières (D. Chen, Preston et Swink, 2015; Guillamón *et al.*, 2016; Iacovou, Benbasat et Dexter, 1995; Zhu, Kraemer et Dedrick, 2004), l'étude de Zhu *et al.* (2006) analyse l'effet de l'affectation spécifique de ressources à une innovation en particulier. Enfin, Guillamón *et al.* (2016) analysent également l'effet du niveau des dettes de la ville sur l'utilisation des réseaux sociaux alors que Zhu et Kraemer (2005) étudient l'effet des coûts sur l'utilisation du commerce électronique.
- La disponibilité d'un support réglementaire n'a été analysée que dans deux études (Zhu, Kraemer et Dedrick, 2004; Zhu et Kraemer, 2005). Son effet s'est avéré positif sur l'utilisation du commerce électronique par les entreprises.
- La prédisposition des partenaires d'affaires est un antécédent qui a été discuté dans deux études (Mishra, Konana et Barua, 2007; Zhu *et al.*, 2006). Son effet positif y est démontré.

- Enfin, l'effet de la taille de l'organisation sur l'utilisation est mitigé. Dans l'étude d'Armstrong et Sambamurthy (1999) il n'a pas été significatif. Son effet est positif sur l'adoption des systèmes d'information dans l'étude de Thong (1999) alors que son impact sur l'utilisation du commerce électronique s'est avéré négatif dans trois autres études (Zhu *et al.*, 2006; Zhu, Kraemer et Dedrick, 2004; Zhu et Kraemer, 2005). Notons tout de même que dans l'étude de Guillamón *et al.* (2016), l'impact de la variable mesurant la taille de la population d'une ville était positif sur l'utilisation des médias sociaux par la ville. Enfin, l'effet de l'envergure internationale des opérations des entreprises sur l'utilisation du commerce électronique est tantôt significatif (Zhu, Kraemer et Dedrick, 2004) et tantôt sans effet significatif (Zhu et Kraemer, 2005).

La littérature s'est également intéressée aux caractéristiques des innovations pour expliquer l'utilisation post-adoption des innovations. Voici les principaux constats liés à ces aspects :

- L'avantage relatif est certainement la caractéristique de l'innovation la plus étudiée. Elle a un effet positif sur l'adoption (Thong, 1999) et sur les étapes post-adoption (D. Chen, Preston et Swink, 2015; Iacovou, Benbasat et Dexter, 1995; Lin, 2008; Zhu *et al.*, 2006).
- La compatibilité de l'innovation avec les caractéristiques organisationnelles et son système de valeur impacte positivement l'utilisation des TI (D. Chen, Preston et Swink, 2015 ; Lin, 2008; Thong, 1999; Zhu *et al.*, 2006).
- L'effet de la complexité de l'innovation sur l'adoption est négatif dans l'étude de Thong (1999) alors que son effet n'a pas été significatif sur la diffusion du commerce électronique (Lin, 2008).
- Zhu *et al.* (2006) est la seule étude à intégrer les préoccupations de sécurité comme un antécédent des étapes post-adoption. Son effet négatif sur l'utilisation organisationnelle du commerce électronique y est démontré.

Enfin, des dimensions plus spécifiques ont été introduites par les auteurs dans leur quête d'expliquer les étapes post-adoption des innovations. On retrouve ainsi la notion de l'intensité de l'information dans l'étude de Thong (1999 : 196) qui représente la mesure dans laquelle l'information est présente dans le produit ou le service d'une organisation. On retrouve également des effets positifs de la connaissance organisationnelle du processus d'approvisionnement et les effets mitigés des perceptions portant sur le volume des transactions et l'incertitude de la technologie (Mishra, Konana et Barua, 2007). L'étude de Massetti et Zmud (1996) a porté exclusivement sur les caractéristiques de l'EDI. En outre, l'index de la participation électronique des citoyens et leur niveau de revenu se sont avérés avoir un effet positif sur l'utilisation des médias sociaux par les villes. L'étude de Lin (2008 : 65) a démontré que l'alignement (System orientation), l'acquisition et la diffusion des connaissances dans le processus de l'approvisionnement ont un effet positif sur l'utilisation du commerce électronique.

Le [tableau 12](#) offre une synthèse des antécédents identifiés dans la littérature. Le tableau ne reprend que les antécédents qui peuvent être pertinents pour la présente étude. Ces derniers ont été regroupés en fonction des deux perspectives théoriques préalablement expliquées : la théorie de la diffusion des innovations et le cadre TOE.

Tableau 12 : Catégorisation des antécédents de l'utilisation

	Variable	Définition	Littérature
Contexte technologique	Infrastructure technologique	Les technologies installées, qui constituent une base la base pour le développement et l'utilisation de la technologie.	Zhu <i>et al.</i> (2006) Zhu et Kraemer (2005) Armstrong et Sambamurthy (1999) Iacovou, Benbasat et Dexter (1995)
	Compétences technologiques du personnel	C'est les compétences des employés en matière d'utilisation de la technologie.	Zhu <i>et al.</i> (2006) Zhu et Kraemer (2005) Thong (1999) Iacovou, Benbasat et Dexter (1995)
	Prédisposition technologique	Ce construit intègre trois dimensions: les technologies utilisées, la fonctionnalité front end du site Web de l'organisation et l'intégration back-office au sein et au-delà de l'organisation.	Zhu, Kraemer et Dedrick (2004)
Contexte organisationnel	Taille de l'organisation	Représente plusieurs aspects d'une organisation tels que les ressources disponibles, la structure et la flexibilité de la prise de décision.	Zhu <i>et al.</i> (2006) Zhu et Kraemer (2005) Zhu, Kraemer et Dedrick (2004) Armstrong et Sambamurthy (1999) Thong (1999)
	Présence à l'international (global scope)	L'étendue géographique des opérations ou des activités d'une organisation, notamment si elles couvrent plus d'un pays.	Zhu et Kraemer (2005) Zhu, Kraemer et Dedrick (2004)
	Engagement financier	C'est les ressources affectées spécialement au développement et à l'utilisation de la technologie en question (proportion de l'ensemble des ressources de l'organisation).	Zhu et Kraemer (2005) Zhu, Kraemer et Dedrick (2004) Guillamón <i>et al.</i> (2016) Iacovou, Benbasat et Dexter (1995)
	Support de la haute direction	La mesure dans laquelle les gestionnaires comprennent et apprécient les capacités de la nouvelle technologie, de même que le degré de leur implication dans la diffusion des innovations technologiques.	D. Chen, Preston et Swink (2015) Lin (2008)
	Quality of senior leadership	La capacité d'absorption des membres de l'équipe de direction : leur capacité à allier des connaissances technologiques et d'affaires, et à développer et à encourager l'apprentissage et la diffusion des innovations au sein de l'organisation.	Armstrong et Sambamurthy (1999) Thong (1999) Boynton, Zmud et Jacobs (1994)
Contexte externe	Pressions compétitives	La pression exercée par les pairs en adoptant et en utilisant des innovations technologiques.	D. Chen, Preston et Swink (2015) Iacovou, Benbasat et Dexter (1995) Thong (1999) Zhu, Kraemer et Dedrick (2004) Zhu et Kraemer (2005) Zhu <i>et al.</i> (2006)
	Support réglementaire	C'est les actions prises par le gouvernement afin d'encourager l'adoption et l'utilisation d'une technologie.	Zhu et Kraemer (2005) Zhu, Kraemer et Dedrick (2004)
	Prédisposition des partenaires	La mesure dans laquelle les partenaires disposent de technologies similaires ou compatibles avec l'innovation en question.	Zhu <i>et al.</i> (2006) Mishra, Konana et Barua (2007)

Tableau 12 : Catégorisation des antécédents de l'utilisation

	Variable	Définition	Littérature
Caractéristiques de l'innovation	Avantage relatif	Le potentiel d'une innovation à apporter de meilleurs avantages (réduire les coûts ou augmenter les revenus) que les innovations existantes.	Zhu <i>et al.</i> (2006) Thong (1999) Lin (2008) D. Chen, Preston et Swink (2015) Iacovou, Benbasat et Dexter (1995)
	Compatibilité	La mesure dans laquelle une innovation est compatible avec les processus, les pratiques et les systèmes de valeurs existants.	D. Chen, Preston et Swink (2015) Zhu <i>et al.</i> (2006) Lin (2008) Thong (1999)
	Coûts	C'est les dépenses liées à la mise en œuvre de l'innovation, de même que les efforts consacrés à la restructuration organisationnelle et à la réorganisation des processus.	Zhu <i>et al.</i> (2006)
	Préoccupations de sécurité	La mesure dans laquelle l'innovation est jugée peu sécuritaire pour l'utilisation de ses fonctionnalités.	Zhu <i>et al.</i> (2006)
	Complexité	La mesure dans laquelle une innovation est difficile à comprendre ou à utiliser.	Lin (2008) Thong (1999)

La présente section nous a permis de jeter les bases théoriques de notre étude, de même que d'identifier et de catégoriser les principaux antécédents de l'utilisation organisationnelle des innovations. L'objectif de la prochaine section est d'élaborer le modèle de recherche qui découle des conclusions des trois dernières sections (2.2, 2.3 et 2.4).

2.5 Modèle de recherche

La présente section a pour objectif d'identifier les différentes variables découlant du modèle conceptuel présenté précédemment (figure 6). Elle vise également à définir les hypothèses supportant les liens entre ces variables. Le modèle de recherche proposé dans cette étude est illustré dans la figure 7 et la liste des hypothèses du modèle est indiquée dans le tableau 14 ci-dessous.

2.5.1 Variables dépendantes

Les diverses formes d'utilisation des TI constitueront les variables dépendantes du modèle de recherche, conformément au modèle d'utilisation du gouvernement électronique au niveau municipal (Titah, 2010: 21) introduit à la section 2.3.2 (voir figure 3 et tableau 7 ci-dessus). En fonction de ce modèle, les formes d'utilisation des mégadonnées sont de cinq types : utilisation informationnelle, utilisation relationnelle, utilisation intégrationnelle, utilisation transactionnelle et utilisation émergente.

Suite à l'analyse des utilisations des mégadonnées dans la littérature, il s'avère que trois formes d'utilisations sont pertinentes pour la présente étude. Il s'agit de l'utilisation informationnelle, de l'utilisation intégrationnelle et de l'utilisation émergente. En effet, les mégadonnées ne semblent pas être utilisées pour des fins d'interaction avec les citoyens (utilisation relationnelle) ou pour l'accomplissement de transactions (utilisation transactionnelle). De ce fait, nous analyserons les effets des antécédents indiqués ci-dessus sur ces trois formes d'utilisation.

2.5.2 Variables indépendantes

Les mégadonnées au sein des villes présentent des caractéristiques qui leur sont propres et qui font qu'il s'agit d'une innovation différente des pratiques existantes. Nous allons d'abord présenter les hypothèses liées aux caractéristiques des mégadonnées sur la base de la théorie de la diffusion de l'innovation de Rogers (2003). Dans un second temps, nous présenterons les facteurs contextuels des villes à la lumière du cadre TOE de Tornatzky et Klein (1982).

A. Caractéristiques de l'innovation

De par leurs caractéristiques, les mégadonnées posent plusieurs défis à surmonter par les organisations qui désirent en tirer profit (van Zoonen, 2016; Woerner et Wixom, 2015). Les principales caractéristiques des mégadonnées ont été détaillées dans la section 2.2. Pour les besoins de la présente étude, nous en retiendrons quatre : l'avantage relatif, la compatibilité, la complexité et les préoccupations de sécurité. Les trois premières variables sont identifiées en fonction de la théorie DOI (Rogers, 2003). La dernière variable, préoccupation de sécurité, représente une des caractéristiques distinctives des mégadonnées. Le tableau 13 indique la correspondance entre les caractéristiques de la théorie DOI, les hypothèses et les caractéristiques des mégadonnées.

Tableau 13 : Table de correspondance DOI-Hypothèses-Caractéristiques

Théorie DOI	Hypothèse	Caractéristique des mégadonnées
Avantage relatif	Avantage relatif	Valeur
Complexité	Complexité	Volume, variété, vitesse et véridité
Compatibilité	Compatibilité	Dynamiques humaines
-	Confidentialité	Sécurité et vie privée
Démonstrabilité	-	-
Observabilité	-	-

A.1. Avantage relatif (bénéfices attendus)

Comme nous l'avons vu dans la précédente section, l'avantage relatif a été identifié comme un facteur important de l'utilisation organisationnelle des innovations (D. Chen, Preston et Swink, 2015; Iacovou, Benbasat et Dexter, 1995; Lin, 2008; Thong, 1999; Zhu *et al.*, 2006). Il décrit les avantages opérationnels et stratégiques escomptés (attendus) de l'utilisation d'une nouvelle technologie. Ainsi, la perception qu'une nouvelle technologie puisse procurer de meilleurs bénéfices que les technologies en cours d'utilisation au sein de l'organisation devrait favoriser l'adoption et l'utilisation de cette technologie (Rogers, 2003). Comme la présente étude s'intéresse à l'utilisation des mégadonnées au niveau organisationnel, nous nous intéresserons à la perception organisationnelle de réalisation des avantages relatifs qui se traduit par la perception des membres de la haute direction des villes.

La littérature suggère que ces avantages peuvent être de deux natures (D. Chen, Preston et Swink, 2015). Les avantages directs couvrent les économies sur le coût des opérations et les gains d'efficacité qui en découlent (par exemple : élimination ou réduction de l'utilisation du papier, saisie des données et réduction des erreurs). Les avantages indirects sont les opportunités qui émergent de l'utilisation d'une technologie comme l'amélioration du service offert aux citoyens et la possibilité de transformer les processus. Par conséquent, et pour les besoins de l'étude, nous définissons l'avantage relatif des mégadonnées comme son potentiel à contribuer à optimiser les routines internes des villes et à offrir des services innovants à leurs citoyens.

De plus en plus d'organisations traitent les données comme un actif (Abbasi, Sarker et Chiang, 2016) et comme un avantage concurrentiel (LaValle *et al.*, 2011; Lycett, 2013). Plusieurs rapports anticipent que les données vont favoriser la croissance économique (Foulonneau *et al.*, 2014; van Loenen, Kulk et Ploeger, 2016). À titre d'illustration, un rapport de McKinsey indique que la valeur potentielle des mégadonnées pour le secteur public en Europe représente un potentiel de 250 milliards USD par année (Manyika *et al.*, 2011 : 8)

Même si l'utilisation des mégadonnées par les villes est à un stade précoce, un éventail d'avantages opérationnels et stratégiques de leur utilisation est anticipé par la littérature (D. Chen, Preston et Swink, 2015). Citons par exemple la réduction des coûts, l'amélioration des processus, la prestation de nouveaux produits et services, le développement de nouvelles sources de revenus et la stimulation de la croissance économique sur son territoire. Les mégadonnées ont donc le potentiel à apporter des bénéfices importants aux villes, à l'externe comme à l'interne. La perception de réaliser ces bénéfices escomptés devrait donc inciter les villes à utiliser les mégadonnées.

H1

La perception des avantages relatifs des mégadonnées influencera positivement leur utilisation par les villes

A.2. Compatibilité

Dans la littérature sur la diffusion de l'innovation, la compatibilité s'est révélée être un antécédent important pour expliquer l'utilisation des technologies par les organisations (D. Chen, Preston et Swink, 2015; Lin, 2008; Thong, 1999; Zhu *et al.*, 2006). Elle fait référence à la mesure selon laquelle une innovation est compatible avec les processus, les pratiques et les systèmes de valeurs existants au sein de l'organisation (Rogers, 2003). Par conséquent, l'adoption et l'utilisation d'une innovation devraient être facilitées si l'on considère qu'elle est compatible avec le profil de l'organisation. Pour la présente étude, nous considérons qu'il revient à la haute direction d'évaluer le degré de compatibilité des mégadonnées avec les systèmes organisationnels.

Les caractéristiques des mégadonnées étudiées dans ce chapitre indiquent que leur adoption et leur utilisation peuvent ne pas être jugées comme compatibles avec les processus, les pratiques et les systèmes de valeurs existants au sein de certaines organisations. On retrouve dans la littérature des éléments qui supportent cette hypothèse. Bhimani (2015) indique que l'utilisation des mégadonnées est bien plus qu'un effort technique sans conséquences organisationnelles. Les mégadonnées pourraient ainsi affecter la définition des lignes d'autorité et redistribuer le pouvoir d'influence au sein de l'organisation.

Par ailleurs, les travaux de Constantiou et Kallinikos (2015) indiquent que la vélocité des mégadonnées peut remettre en cause la stratégie des organisations. Ceci résulte du fait que les systèmes traditionnels de prise de décision stratégique reposent sur des données stables et des horizons de planification à moyen ou à long terme. De plus, les auteurs avancent que les mégadonnées, jumelées avec la puissance des algorithmes, offrent aux organisations la possibilité d'automatiser la prise de décision opérationnelle et même stratégique, qui est le cœur des emplois des gestionnaires (Markus, 2015). Aussi, selon Mergel, Rethemeyer et Isett (2016), les mégadonnées peuvent constituer une menace pour l'efficacité des gestionnaires publics, vu qu'elles affectent leur panier de compétences. Enfin, Baesens *et al.* (2016) indiquent que, même avec les données de qualité, une organisation ne parviendra pas à améliorer sa prise de décision si les dirigeants ne font pas confiance aux données ou aux techniques analytiques.

Par ailleurs, plusieurs auteurs soulignent la nécessité de remettre le citoyen au centre des débats urbain (voir par exemple : Fung (2015), Kitchin (2014) et March et Ribera-Fumaz (2014)). Ils soulignent, notamment, que le discours entourant la numérisation des villes et l'utilisation intensive des données est porté par une élite technophile, soutenu par des firmes informatiques et répond à des attentes d'une petite partie des citoyens (van Zoonen, 2016). Les implications de ces initiatives sont souvent inconnues de la majorité des habitants qui possèdent des attentes raisonnables quant à la gestion et à la manipulation de ces données (Thompson, Ravindran et Nicosia, 2015).

De ce fait, le manque de compatibilité des mégadonnées avec le système organisationnel et la volonté citoyenne peut entraîner une résistance organisationnelle et/ou populaire ce qui influe négativement

sur leur utilisation par les villes. D'autre part, si leur utilisation est compatible avec les processus, les pratiques et les systèmes de valeurs existants, les villes seront plus susceptibles de les utiliser.

H2

La compatibilité des mégadonnées influencera positivement leur utilisation par les villes

A.3. Complexité

La complexité d'une innovation est un facteur déterminant pour l'adoption d'une innovation (Kendall *et al.*, 2001; Thong, 1999). Cependant, dans la littérature sur les étapes post-adoption des innovations, cet antécédent n'a été considéré que dans les travaux de Lin (2008) sur l'évaluation du succès des initiatives de commerce électronique.

Les mégadonnées détiennent des caractéristiques (notamment le volume, la vélocité et la véracité) qui indiquent qu'il s'agit d'une innovation complexe à comprendre et à utiliser. Ces caractéristiques représentent une barrière à l'utilisation des mégadonnées par les villes dans la mesure où des ressources et des compétences supplémentaires sont nécessaires.

D'abord, le coût de la véracité peut dépasser l'avantage concurrentiel offert par la disponibilité des données (volume) et leur variété (Abbasi, Sarker et Chiang, 2016). En effet, la qualité des données est un facteur important pour la réussite des initiatives publiques basées sur les données. Toutefois, les activités permettant de s'assurer de la qualité des données et de leur véracité posent d'importants défis. Ces implications, liées à la confiance dans les données, peuvent donc influencer l'intention d'utiliser les mégadonnées. Ensuite, la variété et la vélocité des données pourraient avoir des effets potentiellement contrastés sur les perceptions des utilisateurs quant à leur utilité et leur facilité d'utilisation (Abbasi, Sarker et Chiang, 2016). Ceci est d'autant plus vrai que l'exploitation des données nécessite une connaissance du contexte dans lequel les ont été collectées, chose qui est souvent absente dans les mégadonnées (Mergel, Rethemeyer et Isett, 2016).

Par ailleurs, Goes (2014) indique qu'à l'exception de quelques grandes entreprises, notamment celles à forte intensité de l'utilisation de l'information, les dirigeants de la plupart des entreprises sont confus par rapport aux potentiels et aux défis apportés par les mégadonnées. Cette confusion, avance l'auteur, est exacerbée par l'environnement fragmenté des solutions d'analyse des mégadonnées.

Abbasi, Sarker et Chiang (2016 : xviii) illustre la complexité liée aux mégadonnées et à leur exploitation avec l'entreprise Netflix qui, en 2006, a offert un prix de 1 million USD à toute équipe qui pourrait améliorer de 10% son modèle de recommandation de films. À la fin de la compétition, deux ans plus tard, la performance du meilleur modèle prédictif ne s'est améliorée que de moins de 2%, alors que la complexité de l'algorithme a considérablement augmenté. L'auteur cite également le livre de l'ancien analyste de Goldman Sach, Emmanuel Derman, qui met en évidence le rôle joué par les modèles complexes dans la crise financière de 2008 (voir *Models Behaving Badly*, Derman (2011)).

De ce fait, nous pensons que la perception de complexité des mégadonnées influe négativement sur l'utilisation des mégadonnées par les villes. Inversement, la perception de facilité d'exploitation des mégadonnées favorise leur utilisation par les villes.

H3 La perception de complexité des mégadonnées influencera négativement leur utilisation par les villes

A.4. Préoccupations de confidentialité

Les préoccupations de confidentialité ont été négligées dans la littérature sur les étapes post-adoption des innovations. L'étude de Zhu *et al.* (2006) est d'ailleurs la seule à prendre en compte l'aspect sécurité d'une innovation technologique. Dans d'autres études, comme celle de Kendall *et al.* (2001 : 228), cet aspect a été inclus dans la dimension « complexité ». Bien qu'aucune étude l'ait examiné, les problèmes liés à la confidentialité des mégadonnées méritent une attention particulière notamment dans le contexte des villes.

Il existe un grand nombre de données privées recueillies au sein des villes, telles que celles concernant les citoyens, celles exclusives aux organisations et celles consignées dans les registres du gouvernement (Zhang, Yang et Chen, 2016). De ce fait, lors de l'exploitation de ces données, la vie privée est un problème majeur, car ces données pourraient contenir des informations sensibles (Martens *et al.*, 2016).

De plus, le fait de croiser des données, provenant de sources multiples, pourrait permettre de déduire l'identité d'un citoyen ou d'une organisation (Mergel, Rethemeyer et Isett, 2016; van Loenen, Kulk et Ploeger, 2016). Ceci est accentué par les efforts de transparence des gouvernements pour publier des données. À titre d'illustration, les auteurs indiquent que les chercheurs de l'Université Carnegie Mellon ont démontré, en 2009, que les éléments de données communs trouvés sur la plupart des pages Facebook permettent de prédire avec précision les cinq premiers chiffres d'un numéro de sécurité sociale (Mergel, Rethemeyer et Isett, 2016 : 932).

Toutefois, dans la plupart des pays, ces données sont étroitement réglementées en termes de méthodes de collecte et de stockage (Mergel, Rethemeyer et Isett, 2016). Les lois et règlements encadrant l'utilisation de ces données constituent, tout de même, un obstacle pour les gestionnaires publics qui désirent exploiter le potentiel des mégadonnées.

Les gestionnaires publics doivent être conscients des questions juridiques, éthiques et techniques qui affectent les cadres réglementaires existants Menon et Sarkar (2016). Par ailleurs, les auteurs soulignent que la décroissance des activités de partage de données entre partenaires a pour origine des préoccupations concernant la possibilité de relever des informations confidentielles ou sensibles, ou des données exclusives.

Le risque d'une perte massive de données confidentielles est tellement présent qu'il est difficile aux gouvernements locaux de faire des compromis, eux qui doivent suivre une éthique exemplaire et irréprochable (Höchtel, Parycek et Schöllhammer, 2016). Par conséquent, les préoccupations de confidentialité peuvent constituer une barrière à l'utilisation des mégadonnées, notamment en ce qui

trait à la sécurité des données et à la protection de la vie privée des citoyens. Les villes peuvent donc être réticentes envers l'utilisation des mégadonnées si un risque d'exposition à des problèmes de confidentialité ou d'atteinte de vie privée des citoyens est perçu.

H4 Les préoccupations de confidentialité et d'atteinte à la vie privée liées aux mégadonnées influenceront négativement leur utilisation des par les villes

B. Contexte technologique

Le contexte technologique du cadre TOE décrit les technologies en cours d'utilisation ou disponibles sur le marché, de même que les compétences techniques pertinentes disponibles dans l'organisation (Tornatzky et Fleischer, 1990 : 153). Nous suivons ce que suggère Zhu *et al.* (2006) en utilisant une seule variable, « compétences technologiques », afin de capter la capacité des infrastructures technologiques des villes, ainsi que les compétences liées aux mégadonnées des employés. Ceci est cohérent avec notre étude, vu que les mécanismes de génération, de partage et d'exploitation des mégadonnées remettent en cause la pertinence des modèles techniques traditionnels, basés sur l'expertise humaine (Constantiou et Kallinikos, 2015). L'utilisation des mégadonnées nécessite donc une importante mise à niveau des modèles analytiques, des outils de même que des compétences des ressources humaines.

L'infrastructure technologique se réfère aux plateformes technologiques qui constituent une base technologique pour le développement et l'utilisation des systèmes et autres applications (Zhu *et al.*, 2006). Bien que les composantes des infrastructures sont disponibles à la vente sur le marché (Armstrong et Sambamurthy, 1999 : 305), le processus de leur intégration en vue de disposer d'une infrastructure cohérente, évolutive et adaptée au contexte stratégique de l'organisation, est complexe et constitue un avantage concurrentiel pour les organisations (Zhu et Kraemer, 2005).

Les travaux d'Armstrong et Sambamurthy (1999) démontrent que la sophistication de l'infrastructure technologique a un impact important sur le processus d'assimilation des technologies. Selon ces auteurs (1999 : 309-310), la sophistication de l'infrastructure permet une intégration intraorganisationnelle (entre les unités), de même qu'une intégration extraorganisationnelle (partenaires externes). Enfin, les auteurs indiquent que la disponibilité d'une infrastructure technologique sophistiquée permet d'améliorer les capacités des gestionnaires et les amène à encourager la diffusion des innovations au sein de l'organisation (en citant Sambamurthy et Zmud (1996)).

Le fait de disposer d'une infrastructure technologique adaptée n'est pas suffisant pour son utilisation, notamment avec les mouvements d'impartition des fonctions TI et la diffusion des technologies infonuagiques telles que les technologies Infrastructure as a Service (IaaS), Platform as a Service (PaaS) et Software as a Service (SaaS). Les compétences TI des ressources internes demeurent importantes notamment pour fournir les connaissances et les compétences nécessaires à l'exploitation du potentiel de ces technologies (Zhu et Kraemer, 2005). En outre, selon Thong (1999 : 196), il existe des preuves empiriques que les organisations employant des personnes ayant des connaissances technologiques élevées sont susceptibles d'utiliser plus intensivement les technologies en question.

Les villes ont besoin de compétences pour gérer et traiter d'importants flux de données, de les analyser et de les interpréter de manière à appuyer la prise de décision. De nouvelles compétences sont donc nécessaires pour les gestionnaires publics et leurs équipes, notamment celles liées à l'analytique des mégadonnées dans un contexte de temps réel (Mergel, Rethemeyer et Isett, 2016), au cryptage des données et à la sécurisation des réseaux et des objets connectés (Martens *et al.*, 2016).

La compétence technologique est donc un facteur critique qui devrait permettre aux villes de profiter pleinement des avantages relatifs des mégadonnées. La disponibilité de ces compétences devrait favoriser l'utilisation des mégadonnées par les villes. Leur absence devrait retarder ou limiter leur utilisation.

H5

Les compétences technologiques des villes influenceront positivement l'utilisation des mégadonnées

Par ailleurs, la compétence technologique des villes devrait avoir un effet sur la perception de complexité des mégadonnées (Greenfeld et Kuznicki, 1975). Ainsi, plus une ville a des compétences technologiques liés aux mégadonnées, moins sera sa perception de leur complexité. Inversement, moins une ville possède des compétences technologiques, plus sa perception de complexité sera élevée.

H6

Les compétences technologiques des villes modéreront négativement la perception de complexité des mégadonnées

C. Contexte organisationnel

Généralement, le contexte organisationnel couvre les aspects liés à la taille de l'organisation, à la disponibilité des ressources financières, aux mécanismes de communication, au degré de formalisation de la structure hiérarchique et au support de la haute direction (Tornatzky et Fleischer, 1990 : 153). Il s'agit du contexte dans lequel opère l'organisation et qui a certainement un impact sur la chaîne de valeur de l'information (Abbasi, Sarker et Chiang, 2016).

Pour les besoins de la présente étude, nous conceptualiserons les éléments du contexte organisationnel à travers la variable se rapportant au « support la haute direction ». Les autres dimensions liées au contexte, tel que la disponibilité des ressources financières, la taille des villes et leurs populations seront considérées comme des variables de contrôle.

Le support de la haute direction concerne la mesure dans laquelle les membres de la direction comprennent et apprécient les capacités des nouvelles technologies, de même que le degré de leur implication dans la diffusion des innovations technologiques. Il est souvent cité comme un facteur clé de succès des projets TI (Lin, 2008). La littérature suggère que les organisations ayant des membres de direction orientés vers l'utilisation de l'innovation comme levier de transformation organisationnelle prennent plus souvent des mesures pour soutenir l'utilisation des innovations disponibles (D. Chen, Preston et Swink, 2015). Grâce à son soutien financier et ses activités de promotion, de la mise en œuvre et de l'utilisation des innovations au sein de l'organisation, les membres de la haute direction agissent comme des agents de changement. Plus précisément, le support de la haute direction à l'utilisation des mégadonnées devrait se traduire par des actions

concrètes et se cristallise avec deux principaux mécanismes : l'engagement financier envers les mégadonnées et le niveau décisionnel liés aux mégadonnées.

Le premier mécanisme concerne l'affectation de ressources financières spécialement au développement et à l'utilisation des mégadonnées. Ceci est cohérent avec ce qui a été proposé dans la littérature précédemment analysée (Iacovou, Benbasat et Dexter, 1995; Zhu, Kraemer et Dedrick, 2004; Zhu et Kraemer, 2005). Selon Zhu et Kraemer (2005 : 68), ceci est plus approprié, car il reflète les ressources engagées plutôt que simplement celles qui sont disponibles, ce qui indique l'importance accordée par la direction aux mégadonnées. Comme le souligne l'actuel maire du Maryland, Martin O'Malley, qui a implanté les technologies de mégadonnées alors qu'il était le maire de Baltimore, « l'utilisation des mégadonnées dans les affaires municipales nécessite un changement de paradigme ; il faut passer d'une approche axée sur les intrants (montants à débloquent) à une approche axée sur les résultats » (O'Malley, 2014 : 555).

Le second mécanisme concerne le niveau décisionnel en rapport avec les mégadonnées. En effet, les organisations ayant connu une transformation par les mégadonnées sont souvent celles où existe un « mandat de haut niveau pour utiliser l'analyse [des mégadonnées] et soutenir une culture ouverte à de nouvelles idées et à des champions qui guident la méthodologie et les compétences » (Kiron *et al.*, 2012 : 5). Ainsi, les pratiques liées à la gouvernance des données ont d'importantes implications pour la disponibilité, la qualité, la maintenance et la sécurité des données (Abbasi, Sarker et Chiang, 2016). Une des pratiques qui démontre l'engagement des villes envers les mégadonnées consiste à désigner une personne de la haute direction (Chief Data Officer) chargée de s'assurer de la qualité des données au sein de l'ensemble de la ville (Baesens *et al.*, 2016) ou la création d'une structure dédiée aux données, à l'instar de ce qui a été fait par la Ville de New York qui a procédé, en 2013, à la création du bureau du Maire, responsable de l'analytique des données (van Zoonen, 2016).

Le support de la haute direction est donc un antécédent à l'utilisation des mégadonnées au sein des villes et constitue le reflet de sa culture envers les mégadonnées. Un fort engagement de la haute direction des villes devrait favoriser l'utilisation des mégadonnées, alors qu'un faible soutien devrait limiter l'utilisation des mégadonnées.

H7

Le support de la haute direction de la ville influencera positivement l'utilisation des mégadonnées

D. Contexte externe

Le contexte externe concerne l'environnement dans lequel évoluent les organisations. C'est le milieu dans lequel les activités sont menées, il couvre ses concurrents et ses partenaires (Tornatzky et Fleischer, 1990 : 153-154). Ainsi, les principaux acteurs de l'environnement externe peuvent être classés en deux groupes: concurrents horizontaux et partenaires verticaux (H. H. Teo, Wei et Benbasat, 2003).

Pour les besoins de la présente étude, nous retiendrons deux antécédents : l'existence de pressions compétitives et l'existence d'un écosystème offrant des prédispositions d'utilisation des mégadonnées.

D.1. Pressions compétitives

La pression compétitive est une contrainte de l'environnement, exercée par les pairs en adoptant et en utilisant des innovations technologiques. L'effet de cette pression est souvent d'amener l'organisation à embrasser la technologie adoptée par les concurrents et partenaires. Cet antécédent est largement repris dans la littérature sur la post-adoption des innovations. Voir par exemple : D. Chen, Preston et Swink (2015), Iacovou, Benbasat et Dexter (1995), Thong (1999), Zhu, Kraemer et Dedrick (2004), Zhu et Kraemer (2005) et Zhu *et al.* (2006). Il est pertinent de noter que les études de Liang et al. (2007) et de D. Chen, Preston et Swink (2015) sont les seules, à notre connaissance, qui analysent et concluent que les pressions compétitives influencent l'utilisation de l'innovation par l'intermédiaire de l'engagement de la haute direction (effet positif des pressions compétitives sur l'engagement de la haute direction envers l'utilisation de l'innovation).

De plus en plus de villes accentuent leurs efforts de compétitivité en menant des initiatives numérisation (Lee et Lee, 2014). Cette concurrence entre les villes vise à attirer de nouveaux résidents, des entreprises et des visiteurs en mettant en œuvre un haut niveau de qualité de vie élevée et un climat économique dynamique (Chichernea, 2015). D'ailleurs, Titah et Barki (2011 : 2) soulignent que « La concurrence entre États est un phénomène intense et durable. Dans le contexte global actuel, les villes constituent les points centraux par lesquels cette concurrence se manifeste. Cette dernière est déterminée par la capacité des municipalités à attirer et à conserver une proportion optimale de capital, d'organisations et d'individus qualifiés en leur offrant des combinaisons attractives de facilités fiscales et de biens publics (services) ».

De ce fait, les villes subissent de grandes pressions afin d'améliorer l'ensemble des aspects relevant de leur compétence (par exemple : la mobilité, la gestion de l'énergie, du logement, du développement durable et de l'innovation). Les villes sont donc de plus en plus appelées à adopter et à mettre en œuvre des solutions de plus en plus innovantes pour relever les divers défis qui se présentent à elles. Dans cette perspective, nous considérons que plus une ville subit des pressions compétitives, plus le support de la haute direction sera plus important quant à l'utilisation des mégadonnées.

H8

Les pressions compétitives influenceront positivement l'engagement de la haute direction des villes vis-à-vis de l'utilisation des mégadonnées

D.2. Dynamisme de l'écosystème

La prédisposition des partenaires est la mesure dans laquelle les partenaires disposent de technologies similaires ou compatibles avec l'innovation en question. Dans le contexte des villes, il s'agit de la disponibilité d'un écosystème d'entreprises, d'organismes, d'incubateurs et de communautés de développeurs dans le domaine des mégadonnées. En effet, l'utilisation des mégadonnées par les villes ne se limite à son personnel. Des entreprises et d'autres organisations peuvent, par exemple, développer des outils ou des applications pour la ville suite dans le cadre d'événements initiés par la municipalité (Saxena et Kumar Sharma, 2016). C'est notamment l'exemple de l'application mobile « Info-Neige MTL » développée par une entreprise montréalaise au profit et suite à une initiative de la Ville de Montréal. Voir également la série d'initiatives menées par la ville d'Amsterdam dans Fitzgerald (2016).

Ainsi, la tendance actuelle consiste donc à innover en établissant des partenariats avec des acteurs de l'écosystème, ce qui est aligné avec les principes fondamentaux du gouvernement ouvert : transparence, participation et collaboration (Abella, Ortiz-de-Urbina-C. et De-Pablos-H., 2017). On retrouve dans la littérature plusieurs exemples de collaboration entre les gouvernements et les acteurs de leur écosystème afin d'offrir des solutions aux citoyens. Dobre et Xhafa (2014) par exemple indiquent que les initiatives de numérisation de la ville d'Amsterdam ont été menées en étroite collaboration avec les entreprises, les organismes gouvernementaux, les universités et d'autres organisations dans l'optique de mise en place de solutions intégrées.

Boyd et Crawford (2012) indiquent que l'écosystème actuel autour du Big Data crée un nouveau type de fracture numérique : le Big Data riche et le Big Data pauvre. Ainsi, l'accès aux données par les académiciens, par exemple, constitue une opportunité permettant une meilleure exploitation des données disponibles. C'est donc la participation des auteurs acteurs de l'environnement qui fait en sorte que nous sommes devant un cas de Big Data riche. Inversement, le manque de participation, d'implication ou de compétence des acteurs de l'environnement font en sorte que nous nous retrouverons dans une situation de Big Data pauvre. Cette collaboration avec les acteurs du milieu est d'autant plus importante que les villes peuvent rencontrer plusieurs défis quant aux fonds, aux capacités et aux compétences nécessaires à la mise en œuvre d'initiatives. Mergel, Rethemeyer et Isett (2016) soulignent d'ailleurs que les organismes gouvernementaux rencontrent un important défi lié au manque d'investissement dans les centres de données, les infrastructures de réseau à grande capacité, les sites informatiques performants, tels que les laboratoires ou les superordinateurs. Dans cette perspective, l'administration Obama a annoncé, en 2012, la mise en œuvre du « Big Data Research and Development Initiative », un investissement de 200 millions USD impliquant de multiples ministères et organismes (Kim, Trimi et Chung, 2014).

Par ailleurs, l'écosystème des villes prend également en compte les éléments réglementaires et législatifs auxquels les villes sont soumises. Mergel, Rethemeyer et Isett (2016) indiquent que les lois et les pratiques actuelles régissant la collecte de données au niveau individuel se concentrent uniquement sur les données collectées par l'administration. Ceci démontre que le gouvernement est assujéti à plus de contraintes réglementaires que d'autres organisations, même si certains secteurs d'activités sont plus réglementés que d'autres. Baesens *et al.* (2016 : 809), par exemple, cite les résultats d'un sondage effectué par Moges (2014) auprès de 50 institutions financières sur leurs investissements dans la qualité des données. L'auteur indique que ces entreprises ont principalement investi dans la qualité des données en raison de la pression réglementaire imposée par les directives de Bâle. Toutefois, la disponibilité d'une réglementation adéquate offrant un support réglementaire à la collecte, au traitement et à la diffusion des mégadonnées devrait favoriser l'utilisation des mégadonnées.

De ce fait, la richesse et le dynamisme de l'écosystème des mégadonnées au sein duquel les villes opèrent influencent l'utilisation des mégadonnées. Un écosystème adéquat devrait influencer positivement sur l'utilisation des mégadonnées par les villes.

H9

Le dynamisme de l'écosystème influencera positivement l'utilisation des mégadonnées par la ville

2.5.3 Variables de contrôle

Nous prendrons en compte quelques variables de contrôle. Il s'agit principalement de la taille de ville (l'organisation), de la taille de sa population et des ressources financières disponibles.

La taille organisationnelle représente plusieurs aspects d'une organisation tels que les ressources disponibles, la structure et la flexibilité de la prise de décision. À partir du moment qu'elle couvre des dimensions importantes de l'organisation (Rogers, 2003 : 411), la taille de l'organisation est l'un des antécédents les plus fréquents dans la littérature sur l'adoption et la diffusion des innovations. Toutefois, le rôle joué par la taille n'est pas clair, principalement en raison de l'attention existante entre la disponibilité des ressources et l'inertie structurelle (Zhu *et al.*, 2006).

La taille de la population représente plusieurs aspects sociaux économiques des villes tels que les ressources disponibles, le dynamisme de l'écosystème et le volume potentiel des données générées. Guillamón *et al.* (2016) souligne que la littérature a démontré une relation positive entre la population et la diffusion des applications Web par les villes. Par ailleurs, les travaux de Martens *et al.* (2016) suggèrent que la disponibilité de plus grands volumes de données devrait améliorer les modèles prédictifs, notamment lorsqu'il s'agit de l'analyse comportementale à haut degré de granularité.

Enfin, dans la mesure de la variable indépendante « support de la haute direction », nous avons pris en considération les ressources financières spécifiquement allouées aux initiatives de mégadonnées. Il s'agit cependant ici d'avoir une idée sur la disponibilité de ressources d'une manière générale. Ceci est pertinent, car au fur et à mesure que les technologies arrivent à maturité, leurs prix baissent (Gantz et Reinsel, 2012). Toutefois, les coûts liés à l'évaluation, l'analyse et la valorisation des données augmentent (Clarke, 2016).

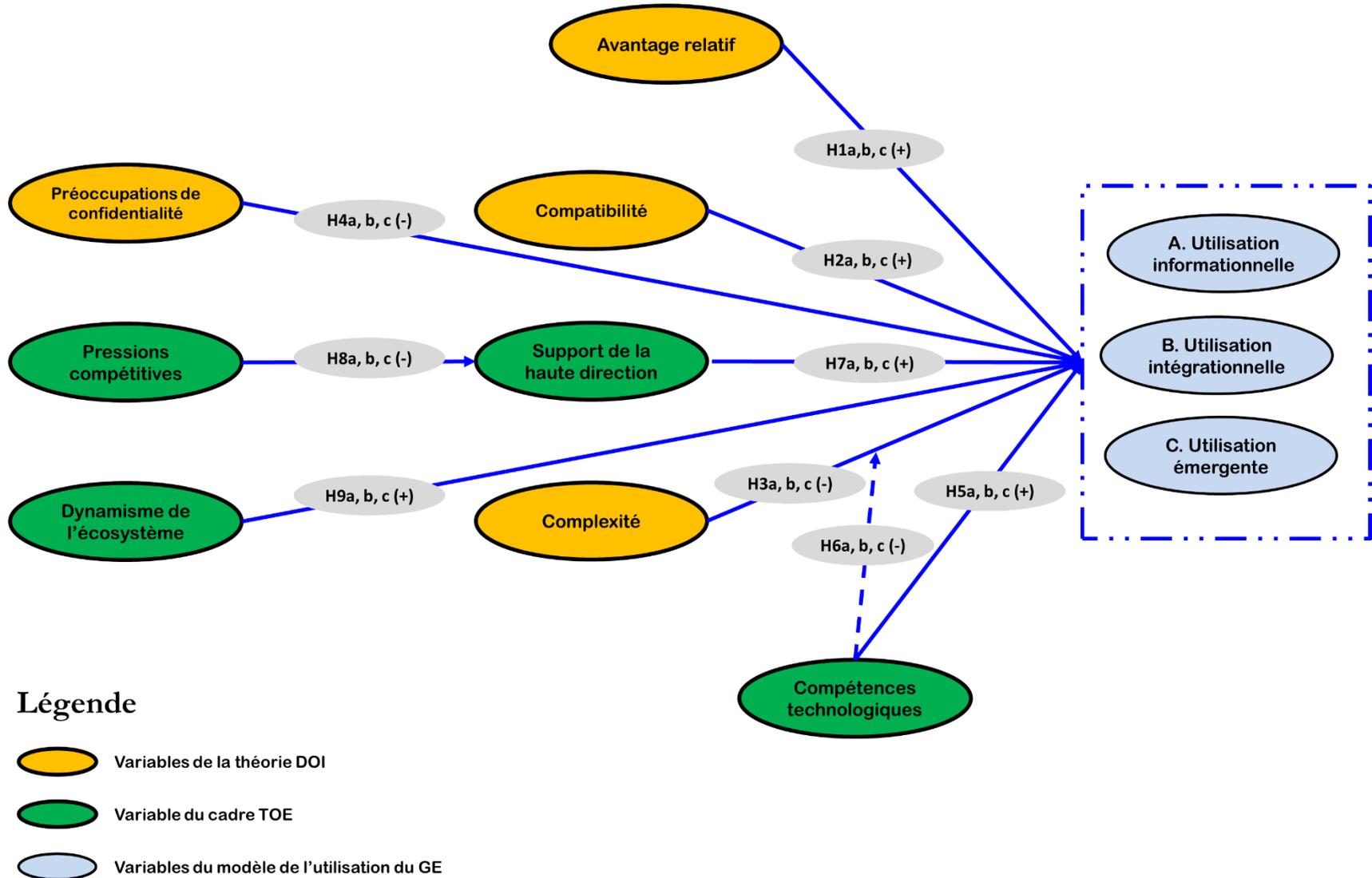
Synthèse

La liste des hypothèses du modèle est indiquée dans le [tableau 14](#) ci-dessous. Aussi, La [figure 7](#) illustre le modèle de recherche de la présente étude, couvrant les principaux construits retenus, les différentes hypothèses des effets entre ces construits, ainsi que sens des effets entre les construits. Le prochain chapitre vise à élaborer la méthodologie permettant de tester ce modèle de recherche.

Tableau 14 : Liste des hypothèses

No.	Hypothèse
H1	La perception des avantages relatifs des mégadonnées influencera positivement leur utilisation par les villes
H2	La compatibilité des mégadonnées influencera positivement leur utilisation par les villes
H3	La perception de complexité des mégadonnées influencera négativement leur utilisation par les villes
H4	Les préoccupations de confidentialité et d'atteinte à la vie privée liées aux mégadonnées influenceront négativement leur utilisation des par les villes
H5	Les compétences technologiques des villes influenceront positivement l'utilisation des mégadonnées
H6	Les compétences technologiques des villes modéreront négativement la perception de complexité des mégadonnées
H7	Le support de la haute direction de la ville influencera positivement l'utilisation des mégadonnées
H8	Les pressions compétitives influenceront positivement l'engagement de la haute direction des villes vis-à-vis de l'utilisation des mégadonnées
H9	Le dynamisme de l'écosystème influencera positivement l'utilisation des mégadonnées par la ville

Figure 7 : Modèle de recherche de l'utilisation des mégadonnées par les villes



**“ It’s easy to lie with statistics.
It’s hard to tell the truth without statistics ”**

Andrejs Dunkels

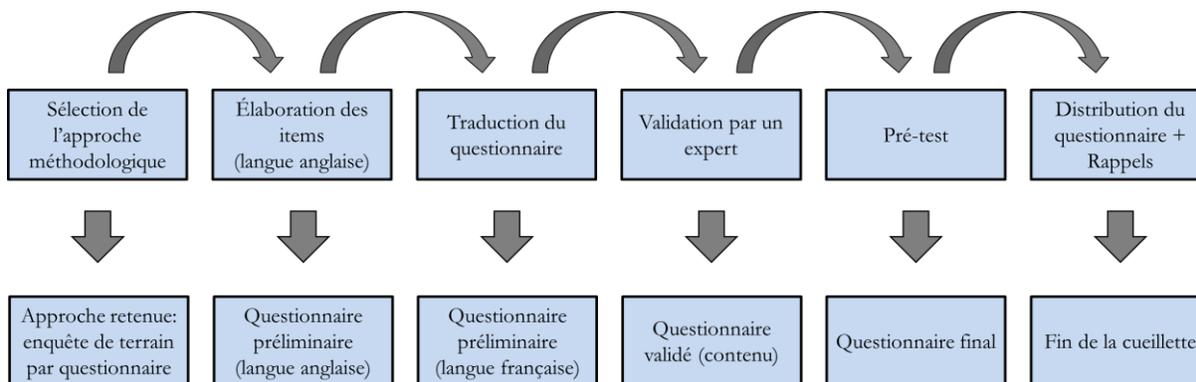
Chapitre 3 : Méthodologie

Ce chapitre vise à décrire la méthodologie retenue dans le cadre de la présente étude et qui sera utilisée afin de tester les hypothèses du modèle de recherche précédemment élaboré. Pour son élaboration, nous nous sommes notamment basés sur les lignes directrices suggérées par Kumar (2014).

La première section de ce chapitre justifie l'utilisation de l'enquête de terrain par questionnaire comme approche de validation du modèle de recherche et comme méthode de collecte de données. La deuxième section indiquera la démarche d'échantillonnage et de collecte de données. La troisième section présentera l'opérationnalisation des variables du modèle de recherche. Enfin, la dernière section décrira la manière dont s'effectuera l'évaluation des qualités psychométriques des mesures sélectionnées, de même que la démarche d'analyse des résultats.

La [figure 8](#) ci-dessous illustre le processus d'élaboration de l'approche méthodologique de l'étude.

Figure 8 : Design du protocole méthodologique



3.1 Sélection de l'approche méthodologique

Le choix de la méthodologie de recherche et de l'instrument de collecte de données est d'une importance majeure, car les forces et les faiblesses de la méthode retenue peuvent affecter la validité des résultats (Kumar, 2014). Les spécifications de l'étude, de même que les caractéristiques de la population cible sont essentielles dans ce choix. C'est ce que nous exposons dans la présente section.

La présente étude vise à répondre à la question suivante : quels sont les antécédents et les formes de l'utilisation des mégadonnées par les villes ? De ce fait, elle se caractérise par plusieurs aspects. D'abord, l'objectif de l'étude est d'analyser l'existence d'une relation d'interdépendance entre, d'une part, les caractéristiques des mégadonnées, le contexte technologique, organisationnel et externe des villes, et d'autre part, les différentes formes de l'utilisation des mégadonnées. Il s'agit donc d'une étude corrélative (Correlational Study) en fonction de la typologie proposée par Kumar (2014 : 12).

De plus, la présente étude se caractérise par son approche structurée et son aspect applicatif vu que les résultats de l'étude ont des implications pour la pratique et la recherche (Kumar, 2014). Enfin, le modèle de recherche de cette étude vise à étudier les effets entre divers construits et à prédire l'utilisation des mégadonnées en fonction des antécédents préalablement identifiés.

Considérant les éléments indiqués ci-dessus, l'approche la plus appropriée à cette étude est l'enquête de terrain par questionnaire auto-administré. Cette approche est d'ailleurs la plus utilisée dans la littérature sur l'étude des antécédents de l'utilisation organisationnelle des TI (annexe 5).

Le questionnaire comme approche de recherche et instrument de collecte de données présente plusieurs avantages (Kumar, 2014 : 181). D'abord, le questionnaire offre un haut niveau d'anonymat, ce qui est utile vu que certains aspects de l'étude peuvent se révéler sensibles ou délicats (par exemple : utilisation des mégadonnées pour des activités de surveillance des citoyens). De ce fait, le questionnaire devrait augmenter la probabilité d'obtenir de l'information précise. Deuxièmement, le questionnaire est un outil économique, d'autant plus que les répondants potentiels sont géographiquement répartis sur plusieurs régions/pays. Comme le questionnaire est administré en ligne, cela permet d'éliminer les coûts d'impression ainsi que les frais de manipulation et d'expédition des documents.

Toutefois, le questionnaire n'est pas sans limites. Plusieurs défis doivent être relevés, notamment en ce qui a trait à la clarté des questions, à la mise en forme du questionnaire, à la motivation (incitation à répondre) des répondants, au biais de non-réponse et à l'influence des réponses précédentes sur les questions restantes du questionnaire. Kumar (2014) indique quelques éléments à considérer lors de l'élaboration d'un questionnaire : 1) les questions doivent être claires et faciles à comprendre, car il n'y a pas d'occasion d'expliquer le sens des questions aux répondants. 2) la mise en forme du questionnaire doit être conviviale et doit faciliter la lecture des questions. 3) la séquence de questions devrait être facile à suivre. 4) le questionnaire doit être interactif (offrant un sentiment de dialogue avec le répondant. 5) préfacier les questions qui peuvent être considérées comme sensibles ou délicates en expliquant leur pertinence et en rappelant le caractère anonyme des réponses. L'ensemble des préconisations indiquées ci-dessus est pris en compte dans l'élaboration du questionnaire de la présente étude, et ce dans une optique de réduction des impacts des limites précédemment mentionnées.

Avant de décrire les étapes d'élaboration du questionnaire, nous présenterons les éléments liés à l'échantillon de l'étude.

3.2 Sélection de l'échantillon

La présente section vise à présenter les paramètres liés à l'échantillonnage, notamment l'identification des organisations visées et des répondants ciblés, de même que les aspects éthiques.

3.2.1 Organisations visées

La population visée par la présente étude est l'ensemble des villes utilisant les mégadonnées. Toutefois, les villes n'ayant toujours pas adopté cette innovation peuvent participer à l'étude et comme l'utilisation nécessite plusieurs technologies, le degré de maturité de l'implantation de ces technologies se manifesterà dans la sophistication de l'utilisation réelle (effective) des mégadonnées. Le Canada a été retenu pour la participation à cette étude, vu que c'est un pays ayant des stratégies gouvernementales de mégadonnées (Kim, Trimi et Chung, 2014) et que les chances à ce que les villes y participent sont grandes. L'échantillon de convenance est donc l'approche retenue pour la sélection de l'échantillon de la présente étude.

Pour des études ayant les mêmes caractéristiques que la présente, et notamment celles utilisant la technique PLS, Gefen, Straub et Boudreau (2000 : 28) suggèrent une taille d'échantillon qui correspondrait à dix fois le nombre d'items de l'échelle mesurant le construit le plus complexe (c'est-à-dire la variable vers laquelle pointent le plus grand nombre de chemins, voir également Barclay, Higgins et Thompson (1995)). Dans notre modèle de recherche, le construit le plus complexe est celui portant sur « l'utilisation émergente » des mégadonnées dont l'échelle de mesure est composée d'une trentaine d'items. Ainsi, la taille de l'échantillon adéquate pour la présente étude est d'environ 300, ce qui devrait être suffisant pour générer des estimations statistiquement fiables à propos des effets entre les divers construits.

3.2.2 Répondants

L'instrument de recherche est destiné à un seul répondant et ayant des connaissances approfondies sur l'utilisation des données municipales. Le répondant idéal est la personne occupant le poste de CDO (Chief Data Officer) (Yang *et al.*, 2014) ou exerçant des fonctions liées à la gestion ou à la gouvernance des données au sein de la ville (par exemple Senior Data Scientist ou IT Manager). Afin d'aider les répondants à répondre plus efficacement aux questions, le questionnaire commencera par la présentation des principaux termes, notamment ce que nous entendons par « mégadonnées ».

Les invitations de participation au questionnaire sont envoyées par courrier électronique aux répondants. Afin d'améliorer le taux de réponse, trois rappels sont programmés et sont transmis à deux semaines d'intervalle. Dans le but d'améliorer la qualité des informations collectées, nous allons examiner les réponses et procéderons à l'élimination des questionnaires des répondants n'occupant pas de fonctions liées à la gestion ou à la gouvernance des données au sein des villes. Enfin, et afin de vérifier la présence de biais de non-réponse, nous avons procédé à la comparaison des principales variables du modèle des réponses survenues après les rappels avec celles obtenues avant la transmission des rappels.

3.2.3 Considérations éthiques

Étant donné que la présente étude implique des êtres humains et recueil de données de nature potentiellement confidentielle, cette démarche est soumise pour approbation au comité d'éthique de la recherche (CER) de HEC Montréal.

Le CER a pour mission de définir et de mettre en œuvre une politique qui satisfait aux exigences énoncées par les trois conseils subventionnaires canadiens, soit : le Conseil de la recherche en sciences humaines du Canada (CRSH), l'Institut de recherche en santé du Canada (IRSC) et le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG). Le CER de HEC Montréal a le mandat d'assurer que l'ensemble des projets de recherche réalisés sous les auspices de l'École respecte l'Énoncé de politique des trois conseils: Éthique de la recherche avec des êtres humains 2014 ([EPTC2](#)).

Un engagement de respect des exigences éthiques est pris auprès du comité qui a délivré un certificat d'approbation éthique en date du 22 juin 2017 ([annexe 3](#)). L'[annexe 4](#) indique la mise en application des exigences éthiques pertinentes pour la présente étude.

3.3 Opérationnalisation des variables

Afin de tester les hypothèses du modèle de recherche élaboré dans le précédent chapitre, nous avons besoin de données primaires, car, à notre connaissance, il n'y a aucune base de données qui contient les données requises pour la réalisation cette étude. Dans la présente section, nous allons donc concevoir les échelles de l'instrument de mesure (questionnaire).

3.3.1 Démarche d'élaboration du questionnaire

Le questionnaire est conçu sur la base d'une profonde compréhension de la littérature identifiée dans le précédent chapitre. Afin de renforcer la validité de construit et dans le but d'être en cohérence et de faciliter la comparaison avec d'autres études, nous avons, dans la mesure du possible, utilisé des items de mesure qui ont déjà été testés dans la littérature sur la diffusion des innovations. Pour des fins de traçabilité des mesures, nous utilisons, pour la construction du questionnaire, des tableaux ayant le format illustré dans la [figure 9](#) suivante (voir l'

Figure 9 : Illustration du tableau d'extraction des items

Source	Item original	Échelle Fiabilité Validité	Items retenus	Échelle
Source 1	Item original 1	Likert-10 LISREL Cronbach 0.79	Item adapté 1	Likert-7
Source 1	Item original 2	Likert-10 LISREL Cronbach 0.79	Non pertinent	-
Source 2

Dans le cas où les mesures identifiées dans la littérature n'étaient pas pertinentes pour la présente étude, ou qu'elles n'existaient pas, nous avons procédé à l'élaboration de nouveaux items afin d'obtenir une mesure précise des construits. Avant de procéder aux pré-tests, le questionnaire a été amélioré grâce à des échanges avec le professeur encadrant les présents travaux.

Par ailleurs, comme les villes visées par les présents travaux se retrouvent dans des pays utilisant le français ou l'anglais comme, le questionnaire est diffusé en ces deux langues. Dans la mesure du possible, les items originaux retenus, en langue anglaise, n'ont pas été modifiés. Ainsi, la version originale du questionnaire est développée en langue anglaise et est, par la suite, traduite vers la langue française.

Toutes les variables du modèle sont mesurées par des items rédigés sous la forme d'un énoncé avec lequel le répondant est en accord ou en désaccord, ou d'un énoncé mesurant le degré d'utilisation des mégadonnées, sur une échelle de type Likert à 7 points de réponse. Cette échelle offre des réponses avec un degré de précision élevé et permet une bonne expérience utilisateur notamment pour un usage avec de petits appareils mobiles.

A. L'option « Ne sais pas »

Afin d'améliorer la qualité des données recueillies avec le questionnaire, nous avons évalué les deux options suivantes : ajouter un bouton « Ne sais pas » à la liste des choix offerte et forcer les participants à répondre aux questions. D'abord, nous anticipons qu'il y aurait des questions que certains répondants ne pourront pas répondre. L'ajout d'une option « Ne sait pas » permet donc d'obtenir une information précise et devrait permettre de réduire le nombre de réponses aléatoires et le taux d'abandon du questionnaire. Toutefois, il est possible que cette option soit également utilisée par les personnes qui ne désirent pas faire l'effort de penser à la réponse adéquate (Krosnick, 1991 : 214). De plus, le fait d'obliger les répondants à répondre aux questions sans offrir cette option pourrait augmenter le taux d'abandon du questionnaire et pourrait également entraîner des problèmes dans la qualité et dans la représentativité des données recueillies (Couper, 2008). Le fait de permettre aux répondants d'ignorer des questions peut entraîner le risque d'avoir beaucoup plus de valeurs manquantes que celles qui correspondent aux personnes qui, en fait, ne savent pas quoi répondre ou ne veulent pas répondre (Couper, 2008).

de Leeuw, Hox et Boevé (2016) proposent d'exploiter l'interactivité des formulaires en ligne en permettant aux répondants de sauter une réponse et de préciser clairement une option « ne sait pas ». De plus, les auteurs suggèrent que lorsqu'un répondant essaie une de ces options, un message apparaît et lui demande de confirmer ce choix tout en soulignant l'importance de sa réponse pour l'étude. Les auteurs ajoutent que la présence d'un message de confirmation augmente la fiabilité des données et indiquent qu'il est préférable de ne pas forcer les participants à répondre et de ne proposer l'option « Ne sait pas » que lorsqu'on s'attend à ce que les répondants ne sachent pas quoi répondre.

En fonction des éléments indiqués précédemment, aucune question du questionnaire ne nécessite de réponse obligatoire et l'option « Ne sais pas » est ajoutée aux choix de réponse. Un message de confirmation sera programmé afin de confirmer les réponses ignorées de même que celles portant sur l'option « Ne sais pas ».

B. Nature des échelles de mesure

Lors de la construction des échelles, plusieurs décisions ont été prises quant à la nature des items ayant constitué les échelles de mesure, de même que concernant la nature même des échelles.

Tout d'abord, la richesse des construits composant le modèle recherche de la présente étude fait en sorte que leurs mesures ne peuvent être évaluées par un seul item. En outre, Diamantopoulos et Winklhofer (2001 : 270) indiquent que l'utilité des échelles de mesure d'un concept théorique comportant un seul item est limitée, car plusieurs problèmes peuvent survenir lors de l'évaluation statistique de l'erreur de mesure. De ce fait, l'ensemble des échelles développées pour les fins de cette étude sont de nature multi-items.

Par ailleurs, la majorité des travaux portant sur les étapes post-adoption analysés dans la présente étude utilisent des échelles réflexives. Les échelles formatives constituent une autre perspective de mesure et « les items formatifs sont supposés avoir une relation causale avec une variable latente, alors que dans le cas des items réfléchifs, c'est la variable latente qui provoque les variables observées » (Bollen, 1989 : 65). Dans cette perspective, si l'une des mesures augmente, la variable augmenterait (même si les autres indicateurs ne changeaient pas). Inversement, si la variable latente augmente, cela ne s'accompagnera pas nécessairement d'une augmentation dans l'ensemble des items.

Le choix entre une échelle formative ou une échelle réflexive dépend principalement de la relation causale entre les items et la variable latente (Diamantopoulos et Winklhofer, 2001). Ainsi, « si l'objectif de l'étude est d'expliquer un construit abstrait ou non observé, les items formatifs donneraient un plus grand pouvoir explicatif » (Fornell et Bookstein, 1982 : 442). Par ailleurs, Fornell et Bookstein (1982) indiquent que lorsque les construits sont formés de plus d'une variable, les items devraient être de nature formative. Ainsi, « un concept est supposé être défini par ses mesures » (Bagozzi et Fornell (1982 :34) cité par Diamantopoulos et Winklhofer (2001 : 270)). Ces éléments correspondent aux objectifs de l'étude de même qu'à la nature des construits qui composent le modèle de recherche. Il s'avère donc cohérent de construire les échelles de recherche grâce aux items formatifs. Ceci est également cohérent avec le modèle d'utilisation du GE au niveau municipal (Titah, 2010) duquel découlent les formes d'utilisation retenues dans le modèle de recherche de la présente étude.

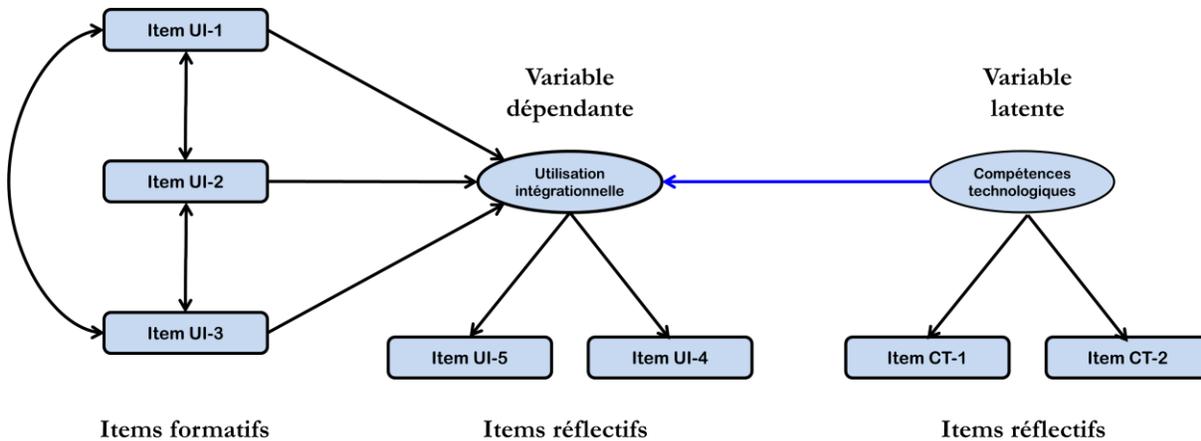
Diamantopoulos et Winklhofer (2001) indiquent les principales propriétés des échelles de mesure de type formatif. Ils mentionnent que les items formatifs ne sont pas interchangeables, que l'omission d'un item revient à l'omission d'une partie du construit et que la corrélation entre les indicateurs formatifs n'est pas expliquée par le modèle de mesure. Ils indiquent également que le construit ainsi mesuré ne peut être estimé que s'il est placé dans un modèle qui intègre les conséquences de la variable latente. Les auteurs soulignent qu'en raison de ces caractéristiques, plusieurs défis doivent être relevés. Premièrement, les procédures conventionnelles utilisées pour évaluer la validité et la fiabilité des échelles réflexives ne peuvent convenir aux échelles formatives. Deuxièmement, d'un point de vue théorique, il est nécessaire de s'assurer que les items couvrent toute la portée de la variable latente. Troisièmement, d'un point de vue pratique, un nombre élevé d'items n'est pas souhaité en raison des exigences liées à la collecte des données et à la complexité du modèle.

En réponse à ces défis, nous avons posé une série d'actions dans une optique d'élaboration d'une échelle de mesure valide à partir d'items formatifs. Premièrement, les items de mesure ont été adaptés à partir d'un examen approfondi de la littérature portant sur l'étude des étapes post-adoption des innovations, ainsi que suite à une revue systématique de la littérature sur les mégadonnées dans un contexte municipal. Deuxièmement, deux entrevues semi-structurées ont été menées avec des employés municipaux (un directeur de l'informatique et un conseiller en informatique/géomatique). Cette étape était nécessaire afin de valider les principaux aspects du questionnaire, et notamment les formes possibles de l'utilisation des mégadonnées et les facteurs favorisant leur utilisation. Suite à ces entrevues, des ajustements ont été apportés aux items du questionnaire qui a ensuite été validé par le professeur supervisant les présents travaux. Troisièmement, nous avons ajouté des items réflexifs aux items formatifs afin d'obtenir un modèle de type Multiple indicators and multiple causes (MIMIC). Ceci permet une validation satisfaisante des échelles tout en évaluant les items comme un ensemble vu que la validité du modèle globale peut être prise comme preuve de soutien pour l'ensemble des items qui forment l'échelle de mesure. De plus, il sera possible d'évaluer la contribution et l'importance de chaque item individuellement (Diamantopoulos et Winklhofer, 2001 : 272). Enfin, chaque construit est relié à au moins un autre construit (Antécédent/conséquence), ce qui permet ainsi de constituer un réseau nomologique. Ceci offre une validation particulièrement pertinente pour la présente étude, vu que les échelles existantes dans la littérature ont été modifiées en éliminant les items les moins pertinents et en ajoutant d'autres items.

Enfin, en cohérence avec les travaux de Titah et Barki (2011), l'élaboration des échelles en s'appuyant sur des items réflexifs et formatifs (modèle de type MIMIC) répond à l'un des principaux objectifs de l'étude qui est d'offrir une riche conceptualisation de l'utilisation des mégadonnées par les villes en

élicitant une liste d'utilisations éventuelles, permettant ainsi de « définir » chaque type d'utilisation. Des items réfléchitifs ont été rajoutés à chaque échelle de mesure permettant d'obtenir un modèle réfléchitif qui pourrait être testé séparément des items formatifs et qui serait également plus facile à tester dans un réseau nomologique vu le petit nombre d'items qu'il contient Titah et Barki (2011 : 16). La figure 10 illustre la nature des échelles de mesure et la différence entre les items réfléchitifs et les items formatifs.

Figure 10 : Illustration de la nature des échelles de mesure



C. Les items inversés (Reverse Codes Items)

L'incorporation d'items inversés pour l'identification de biais de réponse est une procédure courante dans le développement d'échelles de mesure multi-items (Herche et Engelland, 1996 : 366). Plusieurs auteurs ont recommandé l'utilisation de ces items avec des échelles de type Likert (voir notamment Churchill Jr (1979) et Paulhus (1991)). Swain, Weathers et Niedrich (2008 : 116) indiquent que l'objectif de ces items est d'alerter le chercheur sur les questionnaires émanant de répondants inattentifs et de réduire ainsi les biais possibles qui peuvent se produire dans les résultats. Selon Churchill Jr (1979 : 68), l'incorporation d'items inversés devrait permettre de réduire les réponses linéaires (yea-saying ou nay-saying), la logique étant que ces éléments devraient exiger des répondants de s'engager dans un effort cognitif contrôlé.

Toutefois, la recherche a montré la limite de cette technique dans la réduction des réponses linéaires. Schmitt et Stults (1986) indiquent que les items inversés peuvent ne pas atteindre leur but vu qu'une fois que les répondants établissent un schéma de réponse à un questionnaire, ils peuvent ne pas répondre ou faire attention aux alternations des deux types d'items. De plus, la fiabilité de l'échelle de mesure s'affecte considérablement même dans le cas où il y a très peu de répondants qui n'ont pas été attentifs aux items inversés (Krosnick et Presser, 2010).

De ce qui précède, nous n'avons pas eu recours aux items inversés. Ceci devrait éviter de confondre les répondants, d'autant plus que ceux visés dans cette étude sont répartis sur plusieurs pays et peuvent ne pas avoir une parfaite maîtrise d'une des deux langues du questionnaire. Par ailleurs, ceci a également l'avantage de réduire la manipulation des données lors de leur importation et de leur traitement dans les applications statistiques. Toutefois, il est à noter que la formulation de certains items nécessitera l'inversion de l'échelle de mesure.

3.3.2 Mesure des variables indépendantes

Les variables indépendantes sont au nombre de huit (8). Les quatre premières variables découlent de la théorie DOI et portent sur les caractéristiques des mégadonnées préalablement identifiées, soit : l'avantage relatif, la complexité, la compatibilité et les préoccupations de confidentialité. Les quatre autres variables découlent du cadre TOE et concernent le contexte d'adoption et d'utilisation des mégadonnées, soit : les compétences technologiques (contexte technologique), le support de la haute direction (contexte organisationnel) et les pressions concurrentielles et le dynamisme de l'écosystème (environnement externe). Les items de mesure de ces variables ont principalement été identifiés dans la littérature pertinente portant sur l'utilisation organisationnelle post-adoption, analysée dans la section 2.4.2 ([annexe 5](#)) et suite à la revue systématique de la littérature (RSL) portant sur les mégadonnées et les villes ([annexe 2](#)).

- **Avantage relatif** : les items de mesure de l'avantage relatif ont été adaptés de ceux utilisés par D. Chen, Preston et Swink (2015), Iacovou, Benbasat et Dexter (1995) et Premkumar et Roberts (1999). D'autres items ont dû être élaborés à partir de la revue de la littérature.
- **Compatibilité** : les items de mesure de la compatibilité ont été adaptés de ceux utilisés par D. Chen, Preston et Swink (2015), Zhu *et al.* (2006) et Lin (2008). D'autres items ont dû être élaborés à partir de la revue de la littérature.
- **Complexité** : les items de mesure de la complexité ont été adaptés de ceux utilisés par Premkumar et Roberts (1999) et Lin (2008). D'autres items ont dû être élaborés à partir de la revue de la littérature.
- **Préoccupations de confidentialité** : les items de mesure des préoccupations de confidentialité ont principalement été élaborés à partir de la revue de la littérature. D'autres items ont été adaptés de ceux utilisés par Zhu *et al.* (2006) et Kendall *et al.* (2001).
- **Compétence technologique** : les items de mesure des compétences technologiques ont été adaptés de ceux utilisés par D. Chen, Preston et Swink (2015) et Zhu *et al.* (2006).
- **Support de la haute direction** : les items de mesure du support de la haute direction ont été adaptés de ceux utilisés par D. Chen, Preston et Swink (2015), par Premkumar et Roberts (1999) et Lin (2008). D'autres items ont dû être élaborés à partir de la revue de la littérature.
- **Pressions compétitives** : les items de mesure des pressions compétitives ont été adaptés de ceux utilisés par D. Chen, Preston et Swink (2015), Iacovou, Benbasat et Dexter (1995) et Premkumar et Roberts (1999). D'autres items ont dû être élaborés à partir de la revue de la littérature.
- **Dynamisme de l'écosystème** : les items de mesure du dynamisme de l'écosystème ont été adaptés de ceux utilisés par Premkumar et Roberts (1999), Zhu *et al.* (2006) et Zhu et Kraemer (2005).

L'[annexe 6](#) indique la démarche utilisée pour l'identification des items de mesure des variables indépendantes, tandis que l'[annexe 7](#) indique la liste finale des items utilisés pour la mesure des variables du modèle de recherche. Comme le questionnaire a été élaboré en deux langues, les flèches indiquent le sens de traduction des items originaux.

3.3.2 Mesure de la variable dépendante

Les diverses formes d'utilisation des TI constituent les variables dépendantes du modèle de recherche, conformément au modèle d'utilisation du gouvernement électronique au niveau municipal (Titah, 2010: 21). En fonction de ce modèle, les formes d'utilisation des mégadonnées sont de cinq types : transactionnelle, relationnelle, informationnelle, intégrationnelle et émergente. Suite à l'analyse des utilisations des mégadonnées dans la littérature, nous avons retenu les trois dernières formes d'utilisation. Une attention particulière a été portée à l'utilisation émergente des mégadonnées.

La majorité des items de mesure de l'utilisation des mégadonnées ont été élaborés à partir de la revue de la littérature. D'autres items ont été adaptés de ceux utilisés par Titah et Barki (2011), D. Chen, Preston et Swink (2015) et Boynton, Zmud et Jacobs (1994).

L'utilisation informationnelle est mesurée par les services d'information à l'intention des citoyens offerts par les municipalités. L'utilisation intégrationnelle sera principalement mesurée par le partage de l'information et des données par les villes (Gil-Garcia, Zhang et Puron-Cid, 2016; Malomo et Sena, 2017; van Loenen, Kulk et Ploeger, 2016; Welch, Feeney et Park, 2016) notamment dans une optique de réutilisation des données (van Loenen, Kulk et Ploeger, 2016).

En ce qui concerne l'utilisation émergente, on retrouve dans la littérature les aspects clés où les villes désirent avoir le plus d'impact (Nam et Pardo, 2014). Pour Washburn et al. (2010), les villes concentrent leurs investissements dans les domaines de l'administration de la ville, l'éducation, les soins de santé, la sécurité publique, l'immobilier, le transport et les services publics. Pour Giffinger et al. (2007) il s'agit des personnes (capital social et humain), l'économie (compétitivité), la gouvernance (participation, transparence et fonctionnement de l'administration), la mobilité (transport et technologies), l'environnement (ressources naturelles, protection de l'environnement et gestion des ressources) et la vie urbaine (qualité de vie, accès à la culture, à la santé et à la sécurité). Ces aspects offrent une perspective de sous-catégorisation des utilisations des mégadonnées au sein des villes.

L'[annexe 6](#) indique la démarche utilisée pour l'identification des items de mesure des variables indépendantes, tandis que l'[annexe 7](#) indique la liste finale des items utilisés pour la mesure des variables du modèle de recherche. Comme le questionnaire a été élaboré en deux langues, les flèches indiquent le sens de traduction des items originaux.

3.4 Démarche d'évaluation des mesures et d'analyse des données

La présente section vise à décrire la démarche retenue pour l'évaluation psychométrique des mesures élaborées dans la précédente section, de même que l'approche statistique d'analyse des données recueillies. Enfin, nous décrirons l'approche de codification des variables du modèle. Le [tableau 15](#) résume l'approche d'évaluation des mesures du questionnaire et d'analyse statistique des données.

Tableau 15 : Approche d'évaluation des mesures et d'analyse des données

Analyse	Approche/Type	Avant la collecte	Après la collecte
Fiabilité des mesures	Fiabilité de consistance interne (Internal Consistency Reliability)	-	Cronbach $\alpha \geq 0.707$ Composite Reliability ≥ 0.70
	Validité de contenu (Face/Content Validity)	Jugement d'un expert	-
Validité des mesures	Validité convergente de construit (Convergent Validity)	Pré-test	Coefficients de saturation (loadings ≥ 0.50 AVE > 0.50 CR ≥ 0.70
	Validité discriminante de construit (Discriminant Validity)		Analyse de colinéarité (items formatifs) : les coefficients d'inflation de variance < 10 Racine carrée d'AVE d'une variable $>$ les corrélations de cette variable avec les autres HTMT ≤ 0.9 Analyse factorielle : les items devant avoir un chargement plus grand sur la variable cible que sur les autres variables.
Analyse statistique	Technique des modèles d'équations structurelles (SEM, Structural Equation Modeling)	-	Méthode des moindres carrés partiels (PLS) Seuil de confiance de 10% Qualité ajustement modèles MIMIC : $R^2 > 0.30$ Importance relative (loadings) > 0.6 Validité nomologique : $R^2 > 0.30$

3.4.1 Approche d'évaluation des qualités psychométriques des mesures

La section précédente a permis d'établir les items nécessaires à la mesure des variables du modèle de recherche. Si certaines de ces mesures ont été reprises de la littérature existante, d'autres mesures ont été produites dans le cadre des présents travaux. De ce fait, il s'avère important de procéder à la vérification de leur validité et de leur fiabilité. Comme la fiabilité d'une mesure est un préalable à sa validité, nous allons dans un premier temps présenter la démarche d'analyse de la fiabilité des mesures retenues avant de procéder à la présentation de la démarche d'analyse de leur validité. Enfin, et comme ne nous pouvons modifier le questionnaire une fois qu'il a été transmis aux répondants, nous avons procédé à des pré-tests avant d'entamer la collecte des données.

A. Démarche d'analyse de fiabilité des mesures

Dans les sciences sociales, il est impossible d'avoir un instrument de recherche offrant une précision parfaite, vu qu'il n'est pas possible de contrôler les facteurs qui affectent sa fiabilité (Kumar, 2014 : 216). Il existe deux principales approches de détermination de la fiabilité d'un instrument: la fiabilité de consistance interne (Internal Consistency Reliability) et la fiabilité de consistance externe (External Consistency Reliability) (Kumar, 2014 : 217).

Pour les items réfléchitifs des échelles de mesure, nous nous limiterons à l'analyse de la fiabilité de consistance interne, vu que c'est le type de fiabilité le plus utilisé dans la littérature sur l'utilisation post-adoption des TI. Pour ainsi faire, nous utiliserons l'Alpha de Cronbach qui varie entre 0 et 1, avec un seuil de fiabilité de $\geq .707$ (Hair *et al.*, 1998; Nunnally, 1994).

Il est à noter que l'Alpha de Cronbach dépend de la corrélation moyenne entre les items, de même que le nombre d'items (Carmines et Zeller, 1979 : 45). Ainsi, à mesure que la corrélation moyenne entre les items augmente, le nombre d'items augmente et l'Alpha de Cronbach augmente. Les auteurs soulignent que l'ajout d'items qui n'affectent la corrélation moyenne entre les items devrait augmenter la fiabilité des mesures (Carmines et Zeller, 1979 : 46). Toutefois, les auteurs indiquent trois principales limites. D'abord, l'ajout progressif d'items ne va avoir aucun effet sur la validité de l'échelle. En outre, plus une échelle comporte d'items plus importantes sont les ressources (temps et effort) nécessaires à sa construction. Enfin, l'ajout d'items à une échelle peut, dans certains cas, réduire la fiabilité de l'échelle si les items ajoutés font baisser la corrélation moyenne entre les items. Comme l'Alpha de Cronbach ne s'ajuste pas au nombre d'items de l'échelle (Ortiz de Guinea et Webster, 2015), Hair *et al.* (1998) recommandent de le compléter par un autre indicateur de fiabilité. Ainsi, nous calculerons également le Composite Reliability qui varie entre 0 et 1, avec un seuil de fiabilité de $\geq .70$ (Chin, Marcolin et Newsted, 2003 : 200).

B. Démarche d'analyse de la validité des mesures

La validité d'une mesure couvre trois dimensions : la validité de contenu (Face or Content Validity), la validité de construit (Construct Validity) et la Validité de critère (Criterion-related validity) (Carmines et Zeller, 1979 : 17; Kumar, 2014 : 214). Comme l'évaluation de la validité des échelles de mesure des variables formatives diffère de celle employée pour les items réfléchitifs, nous allons établir la validité des deux premières dimensions : validité de contenu pour l'ensemble des items et validité de construit pour les items formatifs.

B.1 Validité de contenu

Le jugement selon lequel un instrument mesure ce qui est supposé mesurer repose principalement sur le lien logique entre les questions et les objectifs de l'étude (Kumar, 2014). Par conséquent, plus le lien entre les mesures est fort, plus grand sera le niveau de validité de contenu. Toujours selon Kumar, les mesures doivent également couvrir l'ensemble des dimensions de la question de recherche. De plus, la distribution des mesures sur les différents aspects du modèle de recherche doit être équilibrée (cohérence de la représentativité de chaque dimension par les mesures). Enfin, l'auteur indique que la validité de contenu peut être démontrée grâce au jugement et à la validation d'un expert dans le domaine.

Comme il a été mentionné précédemment, l'élaboration des mesures utilisées a été faite suite à une revue systématique de la littérature. Les mesures retenues sont évaluées et commentées par le professeur qui dirige les présents travaux et ont fait l'objet d'un jugement d'expert lors du prétest (voir section « C » ci-dessous). Le contenu des mesures ajustées suite à ces révisions sera donc considéré comme valide.

B.2 Validité de construit

La validité de construit couvre la validité convergente et la validité discriminante (Esposito Vinzi *et al.*, 2010 : 182). La validité convergente mesure la corrélation d'une mesure avec d'autres mesures, conceptuellement semblables. La validité discriminante mesure la corrélation d'une mesure avec d'autres mesures conceptuellement différentes (Barki et Hartwick, 1994 : 71).

Suite à la cueillette des données, l'analyse de la validité convergente des mesures réflexives est effectuée avec la méthode AVE (Average Variance Extracted). Nous considérerons que les mesures sont valides si $AVE > .50$ ou que la racine carrée d'AVE $> .707$ (Fornell et Larcker, 1981b : 46).

L'analyse de la validité discriminante des mesures réflexives est effectuée avec la méthode stipulant que la racine carrée d'AVE d'une variable est plus grande que les corrélations de cette variable avec les autres (Fornell et Larcker, 1981a). Nous procéderons à une analyse factorielle selon laquelle les items devant avoir un chargement plus grand sur la variable cible que sur les autres variables.

En ce qui concerne les items formatifs, une analyse de colinéarité sera effectuée suite à la cueillette des données. Elle permet de s'assurer que les coefficients d'inflation de variance soient inférieurs à 10, prévenant ainsi le risque de multi-colinéarité (Diamantopoulos et Winklhofer, 2001).

C. Pré-test

Suite aux précédentes vérifications, une validation du questionnaire est effectuée grâce à un prétest. L'objectif général du prétest est d'améliorer la clarté et la compréhension des énoncés composant le questionnaire et d'avoir une estimation approximative de la durée de sa passation.

Le prétest est mené auprès de trois gestionnaires en TI, œuvrant dans deux différentes villes québécoises. Il est également mené auprès d'un professeur du département des technologies de l'information de HEC Montréal et de deux professionnels travaillant dans le domaine des TI et de l'intelligence d'affaires et ayant un niveau scolaire de deuxième cycle. Les participants sont invités à remplir le questionnaire en se mettant dans la peau d'un gestionnaire TI d'une municipalité en portant une attention particulière aux divers énoncés ainsi qu'aux libellés des choix de réponse. Tout en complétant le questionnaire, les participants sont invités à indiquer les éléments ambigus ou ceux qui manquent de clarté ou qui nécessitent des précisions. Chaque énoncé du questionnaire comporte un numéro unique auquel les participants peuvent se référer dans leurs commentaires. En bas de chaque page du questionnaire, un espace réservé aux remarques et commentaires est prévu. Une fois les participants ont complété le formulaire, une entrevue téléphonique a été conduite avec chacun d'eux afin de recueillir plus d'éléments permettant d'améliorer le questionnaire. Ces pré-tests ont permis d'apporter des ajustements au questionnaire et d'estimer la durée de sa passation à 20 minutes.

L'instrument est développé avec l'outil Qualtrics, une plateforme infonuagique de création et d'administration de sondages et de questionnaires. L'outil permet la collecte, le stockage et l'analyse des données. Il permet également d'exporter les données recueillies dans des formats qui facilitent leur exploitation avec des logiciels d'analyse statistique tel que SPSS (Statistical Package for the Social Sciences) ou avec les outils de modélisation des équations structurelles tels que SmartPLS. Il est à noter que le compte Qualtrics utilisé pour le développement du questionnaire a été obtenu gratuitement grâce à la licence détenue par HEC Montréal.

3.4.2 Approche statistique d'analyse des données

Cette section vise à présenter l'approche prévue pour l'analyse statistique des données recueillies grâce à l'enquête par questionnaire.

Nous utiliserons la technique des modèles d'équations structurelles (SEM, Structural Equation Modeling) pour l'analyse statistique des données. Plusieurs éléments démontrent la pertinence de cette technique pour la présente étude. Premièrement, SEM est largement connue comme une puissante technique d'analyse multivariée de deuxième génération permettant de tester les relations hypothétiques entre les variables d'un même modèle (Fornell, 1982; Fornell et Bookstein, 1982). Deuxièmement, SEM est plus puissante que les techniques de première génération, telles que la régression linéaire, Logit, ANOVA et MANOVA du fait que SEM permet l'analyse simultanée des relations hypothétiques entre les variables indépendantes et dépendantes (Gefen, Straub et Boudreau, 2000). Enfin, les mesures des variables dépendantes du modèle de recherche comme celles des mesures variables indépendantes sont de type ordinal et SEM est la technique la plus appropriée pour ce type de variables.

De plus, nous avons choisi d'utiliser la méthode des moindres carrés partiels (PLS, Partial Least Squares) comme outil de modélisation des équations structurelles. Ce choix repose sur les raisons suivantes. Premièrement, PLS traite à la fois des mesures réflectives (Reflective Items) et formatives (Chin, 1998b). Deuxièmement, PLS est plus appropriée vu que l'objectif de recherche de la présente étude est l'élaboration d'une théorie et que PLS permet justement de tester et de rejeter les hypothèses nulles portant sur des variables n'ayant pas d'effet dans le modèle (Gefen, Straub et Boudreau, 2000). Troisièmement, PLS permet d'effectuer des analyses à la fois sur des modèles de recherche confirmatoire et exploratoire (Gefen, Straub et Boudreau, 2000). Enfin, et contrairement à la technique LISREL, il est démontré que la SEM s'applique bien aux petits échantillons et semble donner des résultats satisfaisants même pour les modèles complexes avec de nombreux construits (Fornell et Bookstein, 1982). Gefen, Straub et Boudreau (2000) suggèrent tout de même une taille d'échantillon qui correspond à dix fois le nombre d'items utilisés pour mesurer la variable la plus complexe du modèle.

L'analyse des données démographiques et l'élaboration des statistiques descriptives sont effectuées avec le logiciel d'analyse statistique SPSSv 23 (Statistical Package for the Social Sciences). La version 3 du logiciel SmartPLS a été utilisée pour l'analyse des données avec la technique SEM-PLS. SmartPLS permet d'effectuer la modélisation des construits réflectifs et formatifs, la régression PLS, l'analyse multi-groupe et offre des options avancées de bootstrapping (voir le site Web de l'éditeur à l'adresse suivante : www.smartpls.com).

3.4.3 Approche de codification des items de mesure

Nous décrivons ici l'approche de codification utilisée dans la présente étude. La codification des items dépend principalement de la manière dont les variables ont été mesurées (Kumar, 2014). Dans la présente étude, les variables indépendantes, la variable médiatrice de même que les variables dépendantes sont mesurées grâce à une échelle ordinale de type Likert. En ce qui concerne les variables de contrôle, des échelles ordinales, nominales et de ratio sont utilisées. L'échelle ordinale est utilisée pour mesurer disponibilité des ressources. L'échelle nominale est utilisée pour capturer

l'information liée au pays et au poste occupé par le répondant. Enfin, l'échelle de ratio est utilisée pour mesurer la taille de la municipalité et la taille de sa population.

Ainsi, nous avons procédé à la création de la structure de codification des items de mesures, illustrée dans l'[annexe 8](#). Afin d'accueillir les données recueillies, la structure de codification des items a été reproduite dans le logiciel d'analyse statistique SPSS, de même que dans l'outil de modélisation des équations structurelles SmartPLS.

Les données collectées ont été extraites de la plateforme Qualtrics vers ces deux applications en générant un fichier en format SAV (format adapté au logiciel SPSS). Suite à l'importation des données dans SPSS, nous avons procédé au recodage des items inversés et au calcul des autres variables (notamment les variables calculées et la conversion des « Numeric Logs ») dans le fichier CSV.

Le prochain chapitre vise à présenter les résultats de l'analyse des données, effectuée grâce à la démarche indiquée dans le présent chapitre. Les données collectées grâce au questionnaire permettront de tester le modèle de recherche dans l'optique de mieux comprendre les antécédents de l'utilisation des mégadonnées au sein des villes ainsi que la manière dont ces mégadonnées sont utilisées par les villes.

“ Torture the data, and it will confess to anything ”
Ronald Coase

Chapitre 4 : Analyse des résultats

Le présent chapitre a trait aux résultats et à l'analyse des données. Les données collectées grâce au questionnaire permettront de tester le modèle de recherche dans l'optique de mieux comprendre les antécédents de l'utilisation des mégadonnées au sein des villes ainsi que la manière dont ces mégadonnées sont utilisées par les villes.

4.1 Analyse statistique des données collectées

La présente section vise à présenter l'analyse statistique des données collectées. Nous présentons d'abord le travail de préparation des données suivi par l'analyse des données démographiques. Nous évaluerons par la suite les qualités psychométriques pour les items réfléchitifs, de même que pour les items formatifs. Enfin, nous procéderons à l'analyse de régression du modèle de recherche.

4.1.1 Collecte et préparation des données

Entre le 17 janvier et le 28 février 2018, 111 questionnaires ont été reçus. Près d'un tiers des questionnaires (33) ont été reçus suite à l'invitation initiale et plus de la moitié (59) suite au premier rappel. Le dernier rappel a permis de collecter 19 nouvelles réponses.

Afin de s'assurer de la qualité des réponses obtenues, un examen préliminaire des questionnaires reçus a été effectué. Suite à cet examen, un seul questionnaire a été rejeté parce que le répondant a coché l'option « Ne sais pas » pour l'ensemble ou la majorité des items et quatre (4) autres questionnaires ont été rejetés vu que les répondants n'ont répondu à aucune question. En tout, 106 questionnaires ont été retenus pour la présente analyse.

Par ailleurs, des corrections ont été apportées à certains questionnaires. Il s'agit par exemple d'un questionnaire dans lequel le répondant indiquait que sa municipalité est située aux États-Unis. Dans un autre questionnaire, le répondant indiquait que la population de sa municipalité comptait moins de 10 000 habitants et indiquait en même temps que la municipalité emploie entre 1001 et 5000 employés, ce qui est très peu probable. Par ailleurs, nous avons effectué un rapprochement entre les différents postes occupés par les répondants. Par exemple, nous avons regroupé sous « IT Director/Manager » les fonctions de : IT Manager, IT Director, GIS Manager et Manager of Geospatial Data and Analytics. Enfin, nous avons procédé à l'inversion des échelles pour les items inversés (voir [annexe 8](#)).

4.1.2 Données démographiques

Le questionnaire a été transmis à 797 villes canadiennes. L'analyse des données démographiques, détaillées à l'[annexe 9](#), permet de constater que l'échantillon est composé de 106 villes, ce qui correspond à un taux de réponse de 13%. La catégorisation de ces villes en fonction du nombre d'habitants démontre que deux tiers des villes participants abritent une population de plus de 10 000 habitants. Plus d'un cinquième des villes a une population de plus de 100 000 habitants.

La majorité des villes participantes emploient moins de 1 000 fonctionnaires municipaux (76,4%). Sept (7) des villes participantes (6.6%) emploient plus de 5 000 personnes.

Les répondants occupent principalement des postes en technologies de l'information : 45% sont des gestionnaires des services de l'informatique ou des CIO et 10% des professionnels en TI (chargé de projet, conseiller aux projets, architecte de données, Lead Open Government et analyste). Les directeurs généraux représentent près d'un tiers (30.5%) des répondants. Enfin, 54.7% des répondants ont utilisé la version anglaise du questionnaire contre 45.3% pour sa version française.

En ce qui concerne la disponibilité des ressources financières, la majorité (74.6%) des répondants estiment que leur ville dispose de peu ou de très peu de ressources financières à consacrer à l'utilisation des mégadonnées. La capacité financière, mesurée par une échelle Likert à 5 niveaux, présente ainsi une moyenne de 1.91, une médiane de 2 et un écart-type de 0,92.

La majorité des répondants (80%) indiquent que l'unité utilisée pour la mesure du volume des données dans leurs villes est le Térabit (1000 Gigabits) alors que seules 9 villes (8.8%) utilisent le Pétabite pour mesurer le volume des données disponibles. Moins de la moitié des répondants (47.2%) indiquent que leur ville utilise des technologies d'analyse des données massives.

4.1.3 Évaluation des qualités psychométriques des items réfléchifs

Comme indiqué précédemment, nous procédons dans un premier temps à l'évaluation psychométrique des items réfléchifs du modèle de recherche. L'objectif de cette analyse est d'examiner les caractéristiques générales des construits et l'homogénéité de leur échelle de mesure.

A. Statistiques descriptives

L'analyse des statistiques descriptives permet essentiellement de s'assurer de la cohérence interne des échelles de mesure des variables du modèle de recherche. Pour rappel, les construits du modèle sont : l'avantage relatif (RA), la compatibilité (CT), les préoccupations de confidentialité (PC), la complexité (CX), les compétences technologiques (CT), les pressions compétitives (CP), le dynamisme de l'écosystème (ED), le support de la haute direction (TS), l'utilisation informationnelle (UN), l'utilisation intégrationnelle (UT) et l'utilisation émergente (UM).

Comme indiqué dans le [tableau 16](#) ci-dessous, la majorité des coefficients Alpha se situent au-dessus du seuil d'acceptation ($\alpha > 0,707$), sauf pour la complexité qui en est assez proche ($\alpha = 0.637$) et le dynamisme de l'écosystème des mégadonnées ($\alpha = 0,497$). Le retrait de l'item CT10 devrait permettre d'améliorer considérablement la fiabilité de la mesure de la complexité des mégadonnées ($\alpha = 0.895$).

Toutefois, comme le coefficient α ne s'ajuste pas au nombre d'items utilisés pour mesurer le construit (Chin, 1998b), Hair *et al.* (1998) recommandent d'effectuer des vérifications complémentaires avant de procéder aux retraits d'items. C'est ce que nous effectuerons dans les prochaines sections.

Tableau 16 : Mesures de fiabilité des items réfléchifs

Items	Observations	Moyenne	Écart-type	Coefficient α
RA08	91	5,09	1,427	$\alpha = 0.895$
RA09	91	4,60	1,505	
RA10	91	4,57	1,586	
CT08	93	2,72	1,338	$\alpha = 0.637$
CT09	93	2,86	1,434	
CT10	93	3,32	1,752	
CX05	90	4,33	1,703	$\alpha = 0.837$
CX06	90	4,86	1,562	
CX07	90	4,36	1,763	
PC07	85	5,47	1,666	$\alpha = 0.777$
PC09	85	5,75	1,388	
PC10	85	4,85	1,680	
TC09	106	3,57	2,182	$\alpha = 0.866$
TC10	106	2,93	1,853	
TC11	106	2,73	1,870	
CP01	68	2,21	1,345	$\alpha = 0.744$
CP02	68	2,26	1,356	
CP03	68	3,68	1,880	
CP04	68	3,81	1,814	
CP05	68	4,76	1,971	
ED09	61	3,27	1,661	$\alpha = 0.497$
ED10	61	3,26	1,653	
TS10	101	2,17	1,550	$\alpha = 0.770$
TS11	101	1,85	1,410	
UN09	105	2,18	1,440	$\alpha = 0.947$
UN10	105	2,33	1,511	
UT09	101	2,08	1,391	$\alpha = 0.938$
UT10	101	2,04	1,469	
UM31	102	1,99	1,346	$\alpha = 0.903$
UM32	102	2,15	1,410	
UM33	102	2,27	1,642	

B. Analyse de la validité convergente

L'analyse de la validité convergente a pour but de s'assurer que les items sont plus fortement corrélés à la variable qu'ils mesurent qu'aux autres variables du modèle (Campbell et Fiske, 1959). Ainsi, nous analyserons la saturation (loading) individuelle des items qui indiquent la corrélation des items avec leurs construits respectifs. Le maintien d'éléments à faible saturation diminuerait la corrélation entre les items d'un même construit (Nunnally, 1994). La fiabilité des items mesure également le niveau d'erreur aléatoire pour chaque construit. Ainsi, des chargements élevés indiquent la fiabilité des mesures de la variable latente. Plus la saturation de l'item est faible, plus le niveau d'erreur aléatoire est élevé. Par conséquent, cette procédure pourrait identifier et éliminer les items qui pourraient augmenter le niveau d'erreur aléatoire d'un construit (Fornell et Larcker, 1981a).

La littérature suggère des seuils de fiabilité pour les saturations des items : Hair *et al.* (1998) ont suggéré que les saturations supérieures à 0.3 étaient significatives et celles supérieures à 0.5 étaient très significatives. Falk et Miller (1992) suggèrent un seuil de 0.55, Chin (1998a) estimait que les saturations devraient être supérieures à 0.6. Carmines et Zeller (1979) ont maintenu une limite de fiabilité de 0.7 tandis que Barclay, Higgins et Thompson (1995) et Fornell (1982) ont spécifié 0.707 comme limite minimale. En tenant compte de l'ensemble des éléments indiqués ci-dessus, et afin de maximiser la capacité du modèle de mesure à satisfaire aux exigences de la validité convergente, une saturation de 0.5 a été déterminée comme valeur minimale.

Les contributions factorielles croisées sont indiquées dans l'annexe 10. Les calculs ont été réalisés par SmartPLS en utilisant trois modèles réflectifs correspondant aux trois types d'utilisations. La fonction « Consistent PLS Algorithm » a été utilisée avec les options indiquées dans le tableau 17 ci-dessous.

Tableau 17 : Configuration des calculs dans SmartPLS

Subsamples	5000
Initial Calculation	Connect all LVs Initial Calculation
Weighting Scheme	Factor/Path
Maximum Iteration	300
Stop Criterion	7
Missing Values	Mean Replacement

L'analyse des structures factorielles permet de constater que les items CT10 et ED09 ont des coefficients de corrélation inférieurs au seuil de 0.5 et ce dans les trois types d'utilisation. De plus, l'item CP05 a des saturations proches du seuil de 0.5 en lien avec l'utilisation informationnelle et l'utilisation intégrationnelle.

Nunnally (1994) a fait valoir que dans le cas d'un soutien théorique solide, d'autres examens des articles à faibles saturations étaient justifiés. Cela serait particulièrement pertinent si les éléments à faible saturation ajoutaient au pouvoir explicatif du modèle. Ainsi, et avant de procéder à un éventuel retrait de ces items, nous allons analyser la fiabilité de consistance interne de même que la variance moyenne extraite (AVE). La cohérence interne a été mesurée en calculant le composite reliability (CR) qui n'est pas influencé par le nombre d'indicateurs contrairement au coefficient α par exemple (Ortiz de Guinea et Webster, 2015). Comme suggéré par Barclay, Higgins et Thompson (1995), les construits ayant un coefficient CR de 0.70 ou plus sont retenus pour une analyse plus poussée. Par ailleurs, nous avons utilisé la variance moyenne extraite (AVE) pour mesurer la robustesse de

l'analyse statistique à travers l'évaluation de la validité convergente. Fornell et Larcker (1981a) et Campbell et Fiske (1959) ont spécifié que l'AVE devrait être d'au moins 0.5 pour que la validité convergente soit acceptable.

Les résultats de l'analyse, indiqués au [tableau 18](#), démontrent que seule la variable « Dynamisme de l'écosystème » pour l'utilisation informationnelle possède un coefficient CR en dessous du seuil de 0.7. Ceci est également confirmé par le bas niveau de fiabilité du coefficient α . Toutefois, cette variable présente un coefficient CR au-dessus du seuil d'acceptation pour les deux autres types d'utilisation et une AVE au-dessus du seuil d'acceptation pour les trois types d'utilisation. Par ailleurs, la variable « Pressions compétitives » est la seule ayant une AVE en dessous du seuil de 0.5 pour les trois types d'utilisation. Enfin, la variable « Préoccupations de confidentialité » a une AVE en dessous du seuil d'acceptation pour l'utilisation intégrationnelle et l'utilisation émergente.

Tableau 18 : Validité convergente des items réflectifs

Types d'utilisation	Construits	α	CR	AVE
Utilisation informationnelle	Avantage relatif	0.887	0.891	0.738
	Compatibilité	0.669	0.748	0.531
	Complexité	0.821	0.827	0.620
	Compétences technologiques	0.873	0.877	0.705
	Dynamisme de l'écosystème	0.464	0.689	0.596
	Pressions compétitives	0.718	0.734	0.360
	Préoccupations de confidentialité	0.740	0.744	0.502
	Support de la haute direction	0.769	0.800	0.675
	Utilisation informationnelle	0.948	0.950	0.906
Utilisation intégrationnelle	Avantage relatif	0.887	0.892	0.740
	Compatibilité	0.669	0.745	0.523
	Complexité	0.821	0.826	0.617
	Compétences technologiques	0.873	0.878	0.707
	Dynamisme de l'écosystème	0.464	0.797	0.735
	Pressions compétitives	0.718	0.736	0.361
	Préoccupations de confidentialité	0.740	0.740	0.490
	Support de la haute direction	0.769	0.799	0.673
	Utilisation intégrationnelle	0.927	0.930	0.869
Utilisation émergente	Avantage relatif	0.887	0.893	0.743
	Compatibilité	0.669	0.744	0.520
	Complexité	0.821	0.827	0.619
	Compétences technologiques	0.873	0.877	0.705
	Dynamisme de l'écosystème	0.464	0.744	0.667
	Pressions compétitives	0.718	0.729	0.351
	Préoccupations de confidentialité	0.740	0.742	0.495
	Support de la haute direction	0.769	0.793	0.663
	Utilisation émergente	0.906	0.910	0.772

De ce fait, et afin d'augmenter la fiabilité de l'échelle de mesure de la variable « Compatibilité » des mégadonnées, il est possible de procéder au retrait de l'item CT10 (effet positif sur le coefficient α , CR et AVE) et de l'item CP05 de l'échelle de mesure de la variable « Pressions compétitives » (effet positif sur AVE qui devient très proche de 0.5). Toutefois, nous avons décidé de maintenir ces items, de même que ceux de l'échelle de mesure du « Dynamisme de l'écosystème » des mégadonnées, même si le chargement de l'item ED09 reste très bas et le coefficient α reste faible. Cette décision est en cohérence avec les recommandations de Nunnally (1994) qui fait valoir que les items doivent être maintenus s'ils pouvaient ajouter au pouvoir explicatif du modèle.

En guise de conclusion, nous pouvons avancer que l'analyse des saturations individuelles des items, de même que les coefficients α , CR et AVE, indique que le modèle de recherche, mesuré avec des items réflexifs, démontre une validité convergente acceptable.

C. Analyse de la validité discriminante

La validité discriminante est une preuve de validité de construit qui se définit comme le degré auquel un construit est différent de toute autre construit du modèle (Barclay, Higgins et Thompson, 1995). L'évaluation de la validité discriminante peut s'effectuer au niveau du construit de même qu'au niveau de l'item. Nous avons procédé à l'évaluation de la validité discriminante au niveau du construit en comparant la racine carrée de l'AVE avec les corrélations inter-construits (Fornell et Larcker, 1981a). Ainsi, la racine carrée de l'AVE, présentée en diagonale, devrait être plus grande que les corrélations inter-construit. Le [tableau 19](#), ci-dessus, indique que les valeurs de la racine carrée de l'AVE sont supérieures aux autres facteurs de corrélation, confirmant ainsi la validité discriminante de l'ensemble des construits du modèle.

Henseler, Ringle et Sarstedt (2015) démontrent que les approches traditionnelles (AVE/Fornell-Larcker Criterion et l'analyse des saturations croisées) peuvent, dans certaines situations, ne pas détecter de manière fiable le manque de validité discriminante. Ils proposent donc le ratio Heterotrait-Monotrait (HTMT) qui se base sur l'étude de simulation Monte-Carlo. La grande majorité des valeurs de l'indicateur HTMT, indiquées dans le [tableau 20](#), sont inférieures ou proches de 0.9 (voir sur le sujet : Gold, Malhotra et Segars (2001) ou Thompson SH Teo, Srivastava et Jiang (2008)) ce qui confirme que les mesures sont toutes différentes.

Au niveau des items, la validité discriminante peut être évaluée en analysant les matrices de saturations croisées (construits en colonnes et items en lignes) qui permet de vérifier la corrélation item-construit à tout moment (Barclay, Higgins et Thompson, 1995). la saturation des items pour un construit doit ainsi être supérieur à la saturation des autres items pour le même construit (colonne) afin de prouver la validité discriminante des construits. L'analyse des saturations des items en fonction de leurs construits respectifs, détaillée dans l'[annexe 10](#), démontre que la majorité des items réflexifs du modèle démontrent une corrélation élevée avec leurs construits respectifs. Il est à noter que les items du construit « Avantage relatif » ont des saturations importantes sur le construit « Pressions compétitives », mais demeurent inférieurs à ceux de leur construit. Toutefois, leurs chargements sont généralement du même ordre que ceux des items CP01 à CP05. Par ailleurs, le chargement de l'item CP04 dans « Avantage relatif » est plus généralement plus important que son chargement dans son construit « Pressions compétitives » et est plus important que les chargements de RA09 dans « Avantage relatif ».

Tableau 19 : Validité discriminante des items réflectifs (Fornell-Larcker Criterion)

	Variables	RA	CT	CX	TC	ED	CP	PC	TS	UN UT UM
Utilisation informationnelle	Avantage relatif	0.858								
	Compatibilité	0.238	0.892							
	Complexité	0.062	-0.395	0.788						
	Compétences technologiques	0.027	0.405	-0.559	0.840					
	Dynamisme de l'écosystème	0.260	0.443	-0.200	0.312	0.756				
	Pressions compétitives	0.613	0.371	-0.298	0.115	0.309	0.705			
	Préoccupations de confidentialité	0.085	-0.048	0.198	-0.190	0.142	-0.061	0.701		
	Support de la haute direction	0.153	0.258	-0.108	0.362	0.445	0.256	-0.196	0.823	
	Utilisation informationnelle	0.341	0.458	-0.279	0.219	0.469	0.418	-0.018	0.335	0.952
Utilisation intégrationnelle	Avantage relatif	0.859								
	Compatibilité	0.238	0.891							
	Complexité	0.062	-0.395	0.786						
	Compétences technologiques	0.027	0.403	-0.561	0.841					
	Dynamisme de l'écosystème	0.248	0.396	-0.173	0.273	0.839				
	Pressions compétitives	0.604	0.374	-0.306	0.120	0.286	0.705			
	Préoccupations de confidentialité	0.062	-0.057	0.204	-0.190	0.096	-0.083	0.713		
	Support de la haute direction	0.153	0.258	-0.106	0.365	0.405	0.259	-0.200	0.821	
	Utilisation intégrationnelle	0.279	0.337	-0.249	0.246	0.400	0.418	-0.123	0.432	0.932
Utilisation émergente	Avantage relatif	0.860								
	Compatibilité	0.238	0.891							
	Complexité	0.059	-0.395	0.787						
	Compétences technologiques	0.029	0.404	-0.560	0.840					
	Dynamisme de l'écosystème	0.255	0.418	-0.186	0.292	0.797				
	Pressions compétitives	0.601	0.374	-0.306	0.121	0.298	0.704			
	Préoccupations de confidentialité	0.085	-0.048	0.198	-0.190	0.133	-0.065	0.701		
	Support de la haute direction	0.154	0.260	-0.107	0.367	0.428	0.260	-0.198	0.815	
	Utilisation émergente	0.273	0.411	-0.221	0.243	0.394	0.406	-0.050	0.442	0.879

Ceci démontre qu'il y a une certaine convergence entre les deux construits « Avantage relatif » et « Pressions compétitives » vu que les items de ce dernier construit sont rédigés d'une manière à mesurer les perceptions des avantages compétitifs procurés par les mégadonnées. Ainsi, une corrélation élevée entre les deux construits est justifiée et ne semble pas problématique vu notamment que les saturations des items sur leurs construits respectifs sont élevées et que la validité discriminante au niveau du construit est vérifiée. De ce fait, nous concluons que le modèle de recherche, mesuré avec des items réflectifs, démontre une validité discriminante acceptable.

Tableau 20 : Validité discriminante des items réflectifs (ratio Heterotrait-Monotrait)

Variables	RA	CT	CX	TC	ED	CP	PC	TS	UN	UT	UM
Avantage relatif											
Compatibilité	0.322										
Complexité	0.125	0.542									
Compétences technologiques	0.078	0.516	0.565								
Dynamisme de l'écosystème	0.494	0.591	0.345	0.478							
Pressions compétitives	0.629	0.488	0.343	0.294	0.706						
Préoccupations de confidentialité	0.134	0.223	0.190	0.183	0.332	0.282					
Support de la haute direction	0.155	0.315	0.147	0.372	0.499	0.335	0.199				
Utilisation informationnelle	0.339	0.520	0.278	0.223	0.604	0.499	0.094	0.340			
Utilisation intégrationnelle	0.276	0.430	0.253	0.248	0.517	0.502	0.112	0.441	0.907		
Utilisation émergente	0.269	0.508	0.213	0.237	0.512	0.533	0.068	0.456	0.838	0.947	

4.1.4 Évaluation des qualités psychométriques des items formatifs

L'évaluation de la fiabilité et de la validité des échelles de mesure constituées d'items formatifs diffère de celle employée pour les échelles réflectives (Diamantopoulos et Winklhofer (2001) et Cenfetelli et Bassellier (2009)). Même si la construction de mesures basée sur des items formatifs en est encore aux premiers stades du développement méthodologique, l'évaluation de leur validité de contenu est importante. Bien qu'il n'existe pas de consensus concernant l'analyse des échelles formatives, la littérature recommande certaines procédures permettant de mieux s'assurer de leur validité. Dans la présente étude, nous nous concentrerons sur les aspects liés à la manière dont les mesures ont été élaborées de même que sur l'analyse de colinéarité des items.

Contrairement aux variables réflectives pour qui les items doivent être fortement corrélés, les items formatifs doivent contribuer à former leurs variables latentes respectives et ne doivent pas présenter une corrélation trop forte (Cenfetelli et Bassellier, 2009).

La validité des échelles de mesure d'une mesure formative repose notamment sur la spécification de son contenu et de ses items (Diamantopoulos et Winklhofer, 2001). Les contenus et les items de mesure des construits du modèle de recherche ont été élaborés à partir d'une revue systématique de la littérature. Certains des items de mesures utilisés ont fait l'objet d'une validation par leurs auteurs. De plus, des entrevues avec trois professionnels TI, œuvrant dans deux différentes villes québécoises ont été effectuées. Enfin, les mesures retenues ont été évaluées, corrigées et validées par le professeur qui dirige les présents travaux et ont fait l'objet d'un jugement d'expert lors du prétest. Le contenu des mesures ajustées suite à ces révisions est donc considéré comme valide.

Par ailleurs, comme la colinéarité est une menace potentielle à l'interprétation des construits formatifs (Titah et Barki, 2011), nous avons procédé à une analyse de colinéarité afin de s'assurer que les coefficients d'inflation de variance (VIF) soient inférieurs au seuil d'acceptation (Diamantopoulos et Sigauw, 2006), prévenant ainsi le risque de multi-colinéarité (Diamantopoulos et Winklhofer, 2001).

Le [tableau 21](#), ci-dessous, permet de constater que les valeurs VIF de tous les construits formatifs sont inférieures à la valeur limite de 10 recommandée par Hair *et al.* (1998). Toutefois, il est important de signaler que certains auteurs recommandent un seuil maximal de 3.33 (Diamantopoulos et Sigauw, 2006) tandis que Mathieson, Peacock et Chin (2001) suggèrent de porter une attention particulière aux items ayant un VIF allant de 5 à 10. Compte tenu de la nature exploratoire de la présente étude, l'absence de chevauchement conceptuel entre les mesures, de même que les éventuelles conséquences négatives liées au retrait des items sur la définition et la nature des construits, tous les items formatifs du modèle ont été conservés (Cenfetelli et Bassellier, 2009).

Enfin, dans l'approche MIMIC, un niveau d'ajustement du modèle est considéré comme une preuve de soutien pour l'ensemble des items formatifs d'un construit (voir Titah et Barki (2011) et Diamantopoulos et Winklhofer (2001)). Ainsi, dans la première partie de la prochaine section nous allons procéder à l'évaluation des modèles MIMIC pour chacun des construits du modèle de recherche (sauf pour « Pressions compétitives » qui est uniquement mesuré par des items réfléchitifs).

Tableau 21 : Coefficients d'inflation de variance (VIF)

RA		ED		UN		UM	
RA01	2.346	ED01	1.573	UN01	1.459	UM01	4.885
RA02	3.183	ED02	2.030	UN02	2.039	UM02	9.835
RA03	2.152	ED03	1.689	UN03	1.345	UM03	8.906
RA04	2.139	ED04	1.271	UN04	1.099	UM04	6.731
RA05	2.351	ED05	1.900	UN05	1.570	UM05	3.387
RA06	2.720	ED06	1.801	UN06	1.751	UM06	4.739
RA07	2.363	ED07	1.575	UN07	1.510	UM07	7.437
		ED08	1.500	UN08	1.572	UM08	4.900
						UM09	4.898
						UM10	5.244
						UM11	3.488
						UM12	4.539
						UM13	6.863
						UM14	8.671
						UM15	5.419
						UM16	7.120
						UM17	4.389
						UM18	4.136
						UM19	4.571
						UM20	3.870
						UM21	3.786
						UM22	4.196
						UM23	3.742
						UM24	2.936
						UM25	6.998
						UM26	6.289
						UM27	4.155
						UM28	3.283
						UM29	5.520
						UM30	4.237

CT		TS		UT	
CT03	2.069	TS01	5.420	UT01	1.858
CT04	3.363	TS02	7.216	UT02	3.421
CT05	1.563	TS03	3.721	UT03	3.377
CT06	1.765	TS04	2.337	UT04	3.685
CT07	1.856	TS05	3.055	UT05	3.476
		TS06	2.286	UT06	3.118
		TS07	2.230	UT07	3.410
		TS08	2.555	UT08	3.634
		TS09	1.778		

CX		TC	
CX01	3.825	TC01	2.333
CX02	5.735	TC02	2.564
CX03	5.430	TC03	3.063
CX04	2.263	TC04	1.703
		TC05	1.268
		TC06	1.931
		TC07	1.780
		TC08	1.280

PC	
PC01	3.752
PC02	4.344
PC03	2.629
PC04	3.093
PC05	1.796
PC06	1.361
PC08	1.161

4.1.5 Analyse de régression

Le test des modèles est réalisé à travers une analyse de régression par moindres carrés partiels (PLS). Dans un premier temps, nous allons évaluer les modèles MIMIC pour chacun des construits du modèle de recherche (sauf pour « Pressions compétitives » qui est uniquement mesuré par des items réfléchitifs). Nous procéderons, par la suite, à l'évaluation des modèles structurels afin de valider les différentes hypothèses de l'étude. Le niveau de significativité des coefficients est constamment examiné par une procédure de « Bootstrap » consistant à répliquer l'estimation des modèles sur un grand nombre de sous-échantillons constitués aléatoirement dans l'échantillon principal (Efron et Gong, 1983). Précisément, les modèles sont ici testés sur 5 000 échantillons de 106 observations et pour un seuil de confiance de 10% ($p < 0.1$).

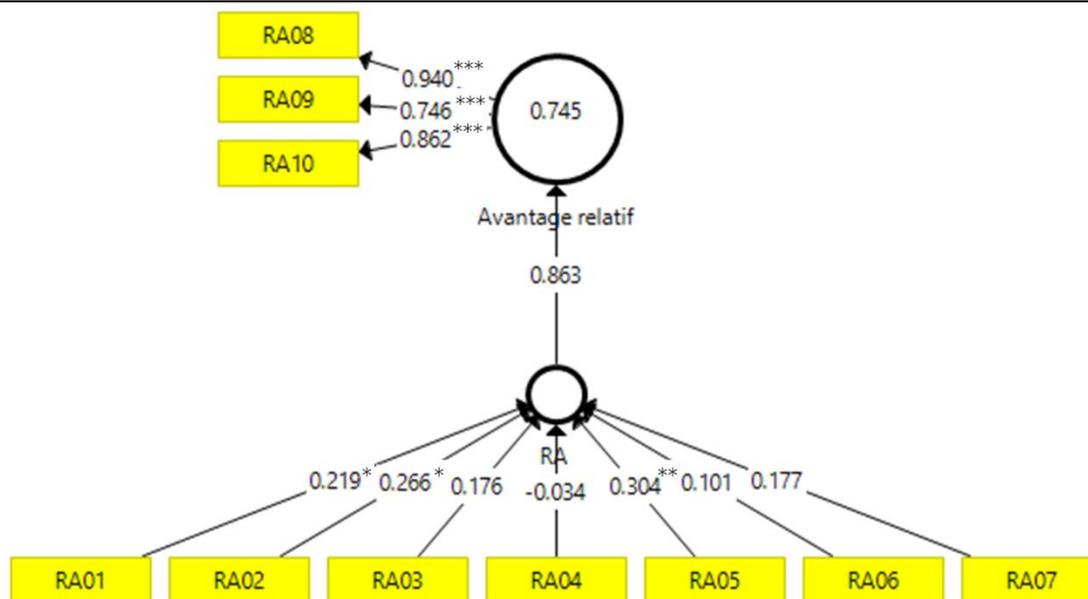
A. Test des modèles MIMIC

L'ensemble des construits du modèle de recherche est testé à travers une approche MIMIC, sauf « Pressions compétitives ». Les modèles MIMIC sont notamment recommandés lorsqu'une variable latente est mesurée par un ensemble d'indicateurs formatifs et expliquée par différentes variables (Diamantopoulos et Winklhofer, 2001).

L'analyse portera dans un premier temps sur la significativité des items à travers l'évaluation du T statistique et de la valeur de p. Pour les items qui présentent de faibles résultats, nous procéderons à l'analyse de leurs coefficients de saturation (Cenfetelli et Bassellier, 2009). En effet, ces auteurs recommandent, qu'en plus d'analyser l'importance absolue des indicateurs (Weights), de procéder à l'évaluation de leur importance relative (loadings). Les indicateurs qui ont une contribution relativement faible peuvent avoir une contribution absolue importante si ces indicateurs sont indépendamment évalués. Les corrélations bivariées (loadings) entre les indicateurs et leurs construits sont indiquées dans l'[annexe 11](#).

Avantage relatif : le modèle MIMIC de l'avantage relatif des mégadonnées est présenté à la [figure 11](#). Tous les items réfléchitifs sont significatifs ($\gamma_8 = 0.940$, $\gamma_9 = 0.746$ et $\gamma_{10} = 0.862$ avec $p < 0.001$). Des 7 indicateurs formatifs, seuls 3 sont significatifs ($\gamma_1 = 0.219$ et $\gamma_2 = 0.266$ avec $p < 0.1$ et $\gamma_5 = 0.304$ avec $p < 0.01$). Ceci indique que, pour les villes participantes, l'avantage relatif des mégadonnées réside essentiellement dans l'amélioration de la qualité de vie des citoyens (RA01), l'amélioration de la qualité des services offerts aux citoyens (RA02) et dans l'amélioration de la qualité de l'information (RA05). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que les items suivants démontrent une importance relative supérieure à 0.6 : l'amélioration des capacités de prévision des villes (RA07), l'obtention de l'information en temps opportun pour la prise de décision (RA06), possibilité d'offrir de nouveaux produits ou services aux citoyens (RA03) et l'amélioration de l'efficacité des opérations municipales (RA04). L'analyse des coefficients d'inflation de variance (VIF) de ces indicateurs de même que leur pertinence théorique indique qu'ils ne présentent pas de potentiels chevauchements conceptuels. Nous décidons donc de les conserver. Les indicateurs formatifs du modèle expliquent 74,5% de la variance ($R^2 = 0.745$).

Figure 11 : Modèle MIMIC de l'avantage relatif



Compatibilité : le modèle MIMIC de la compatibilité des mégadonnées est présenté à la figure 12. Tous les items réflectifs sont significatifs ($\gamma_8 = 0.962$ et $\gamma_9 = 0.826$ avec $p < 0.001$). Des 7 indicateurs formatifs, seuls 2 sont significatifs ($\gamma_2 = 0.532$ et $\gamma_4 = 0.494$ avec $p < 0,1$). Ceci indique que, pour les répondants, la compatibilité des mégadonnées réside essentiellement dans l'adéquation de son utilisation avec la culture organisationnelle (CT02) et sa cohérence avec le système de prise de décision (CT04). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que les items suivants démontrent une importance relative supérieure à 0.6 : la conformité de l'utilisation des mégadonnées aux pratiques (CT01) et aux valeurs (CT03) de la municipalité. L'analyse des coefficients d'inflation de variance (VIF) de ces indicateurs de même que leur pertinence théorique indique qu'ils ne présentent pas de potentiels chevauchements conceptuels. Nous décidons donc de les conserver. Les indicateurs formatifs du modèle expliquent 36.8% de la variance ($R^2 = 0.368$).

Complexité : le modèle MIMIC de la complexité des mégadonnées est présenté à la figure 13. Tous les items réflectifs sont significatifs ($\gamma_5 = 0.802$, $\gamma_6 = 0.790$ et $\gamma_7 = 0.744$ avec $p < 0.001$). Des 4 indicateurs formatifs, un seul est significatif ($\gamma_3 = 0.623$ avec $p < 0.1$). Ceci indique que, pour les répondants, la complexité des mégadonnées réside essentiellement dans l'exploitation de données de formes différentes et provenant de sources variées (CX03). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que les items suivants démontrent une importance relative supérieure à 0.6 : les défis liés au traitement de grandes quantités de données (CX01), les défis liés au traitement de données en temps réel (CX02) et les défis liés à la nécessité de s'assurer de la qualité des données (CX04). L'analyse des coefficients d'inflation de variance (VIF) de ces indicateurs de même que leur pertinence théorique indique qu'ils ne présentent pas de potentiels chevauchements conceptuels. Nous décidons donc de les conserver. Les indicateurs formatifs du modèle expliquent 48.5% de la variance ($R^2 = 0.485$).

Figure 12 : Modèle MIMIC de la compatibilité

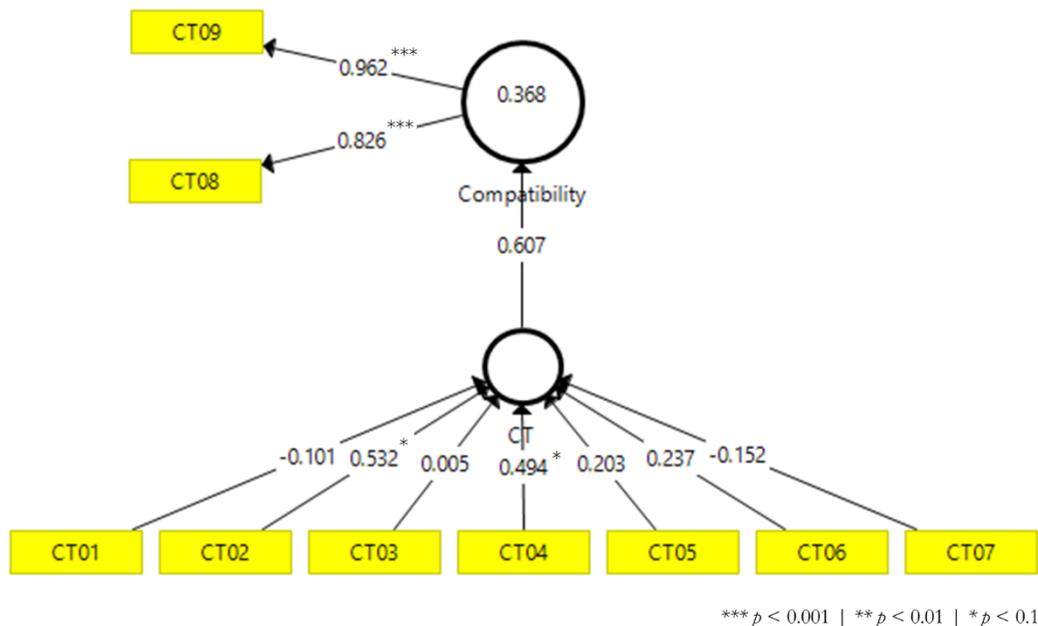
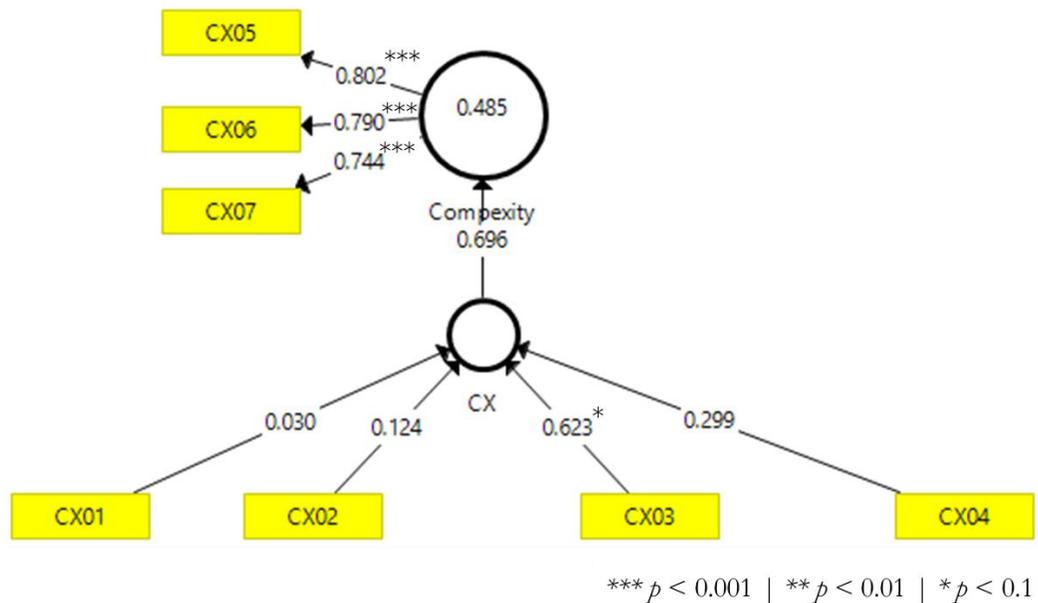


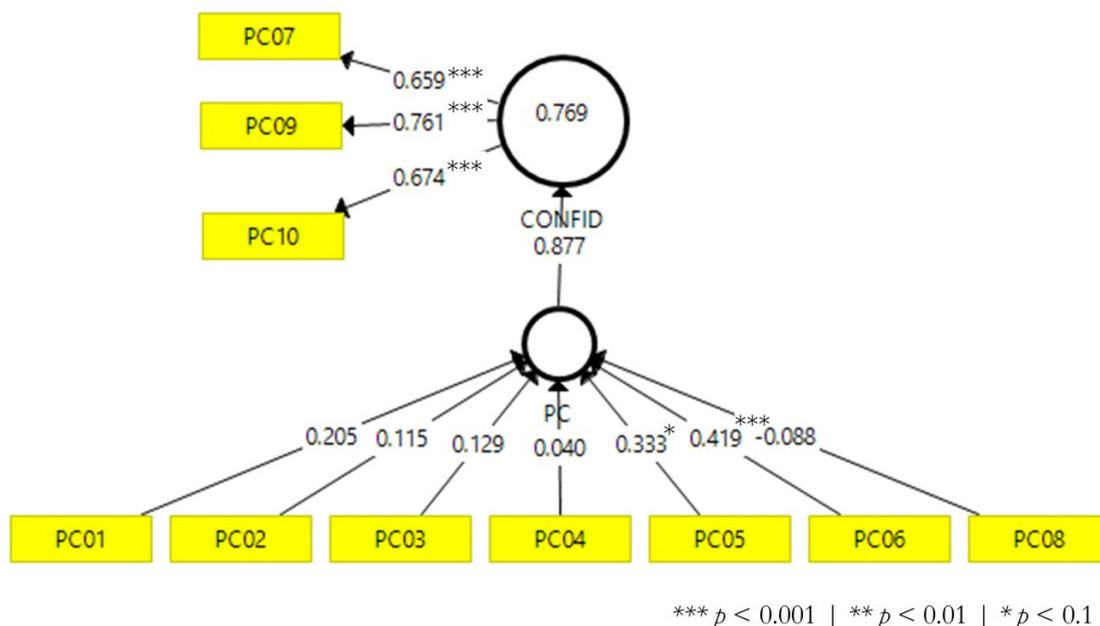
Figure 13 : Modèle MIMIC de la complexité



Préoccupations de confidentialité : le modèle MIMIC des préoccupations de confidentialité est présenté à la figure 14. Tous les items réflectifs sont significatifs ($\gamma_7 = 0.659$, $\gamma_9 = 0.761$ et $\gamma_{10} = 0.674$ avec $p < 0.001$). Des 7 indicateurs formatifs, deux sont significatifs ($\gamma_5 = 0.333$ avec $p < 0.1$ et $\gamma_6 = 0.419$ avec $p < 0.001$). Ceci indique que, pour les répondants, les préoccupations de confidentialité

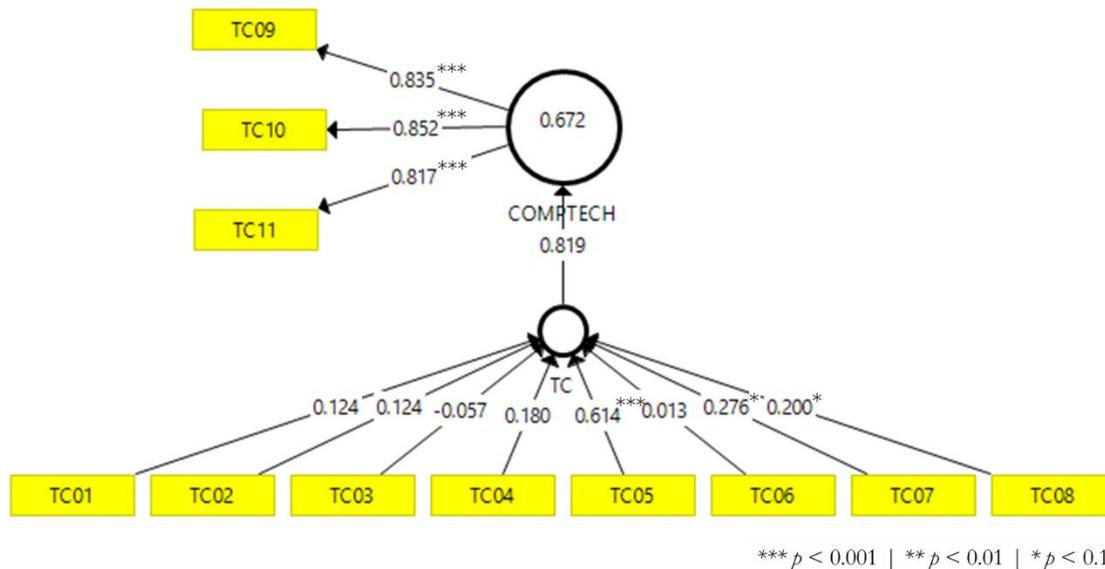
résident essentiellement dans le fait que les citoyens des villes sont préoccupés par la sécurité et la confidentialité de leurs données (PC05) et le fait qu'il existe des restrictions légales qui limitent la façon dont les villes collectent, stockent ou partagent des données personnelles (CP06). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que les items suivants démontrent une importance relative supérieure à 0.6 : les préoccupations liées à la sécurité des données des citoyens (PC01), les informations confidentielles contenues dans les jeux de données (PC02), les droits de propriété des jeux de données (PC03) et dans la révélation d'informations exclusives ou de données sensibles (PC04). L'analyse des coefficients d'inflation de variance (VIF) de ces indicateurs de même que leur pertinence théorique indique qu'ils ne présentent pas de potentiels chevauchements conceptuels. Nous décidons donc de les conserver. Les indicateurs formatifs du modèle expliquent 76.9% de la variance ($R^2 = 0.769$).

Figure 14 : Modèle MIMIC des préoccupations de confidentialité



Compétences technologiques : le modèle MIMIC des compétences technologiques est présenté à la figure 15. Tous les items réflectifs sont significatifs ($\gamma_9 = 14.638$, $\gamma_{10} = 19.538$ et $\gamma_{11} = 14.153$ avec $p < 0.001$). Des 8 indicateurs formatifs, trois sont significatifs ($\gamma_5 = 5.999$ avec $p < 0.001$ et $\gamma_7 = 2.204$ et $\gamma_8 = 1.908$ avec $p < 0.1$). Ceci indique que, pour les répondants, les compétences technologiques de leurs villes résident essentiellement dans l'accès à une infrastructure adaptée à l'utilisation des mégadonnées (TC05), dans le fait d'avoir recruté du personnel spécialisé dans les mégadonnées (TC07) et dans l'accès à des consultants externes spécialisés dans les mégadonnées (TC08). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que l'item suivant démontre une importance relative supérieure à 0.6 : possibilité d'avoir accès à des outils permettant la gestion de grands ensembles de données (TC01). L'analyse du coefficient d'inflation de variance (VIF) de cet indicateur de même que sa pertinence théorique indique qu'il ne présente pas de potentiels chevauchements conceptuels. Nous décidons donc de le conserver. Les indicateurs formatifs du modèle expliquent 67.2% de la variance ($R^2 = 0.672$).

Figure 15 : Modèle MIMIC des compétences technologiques



Dynamisme de l'écosystème : le modèle MIMIC du dynamisme de l'écosystème des mégadonnées est présenté à la figure 16. Tous les items réflectifs sont significatifs ($\gamma_9 = 0.558$ et $\gamma_{10} = 0.541$ avec $p < 0.001$). Des 8 indicateurs formatifs, deux sont significatifs ($\gamma_7 = 0.394$ et $\gamma_8 = 0.657$ avec $p < 0.1$). Ceci indique que, pour les répondants, le dynamisme de l'écosystème dans leurs villes réside essentiellement dans la mise en œuvre d'une stratégie ou d'une politique de mégadonnées par le gouvernement central (ED07) de même que dans la disponibilité d'un support législatif à favorisant l'accès aux données dont la ville a besoin (ED08). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique qu'aucun item ne démontre une importance relative supérieure à 0.6. Les indicateurs formatifs du modèle expliquent 56.3% de la variance ($R^2 = 0.563$).

Support de la haute direction : le modèle MIMIC du dynamisme de l'écosystème des mégadonnées est présenté à la figure 17. Tous les items réflectifs sont significatifs ($\gamma_{10} = 12.054$ et $\gamma_{11} = 9.503$ avec $p < 0.001$). Des 9 indicateurs formatifs, trois sont significatifs ($\gamma_6 = 1.948$ avec $p < 0.1$, $\gamma_7 = 2.857$ avec $p < 0.01$ et $\gamma_9 = 4.236$ avec $p < 0.001$). Ceci indique que, pour les répondants, le support de la haute direction dans leurs villes réside essentiellement dans le fait que la haute direction considère que l'analyse statistique des données est plus fiable que l'intuition (TS06), dans le fait que la haute direction nomme un responsable principal des données, chargé de s'assurer de la qualité des données (TS07) et dans le fait que des politiques-cadres ont été mises en place afin de favoriser l'utilisation des mégadonnées (TS09). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que l'item suivant démontre une importance relative supérieure à 0.6 : l'allocation de ressources adéquates à l'utilisation des technologies des mégadonnées (TS04), dans la création d'une unité responsable de l'utilisation et de la gestion des données (TS08), dans le soutien offert aux initiatives impliquant l'utilisation des mégadonnées (TS02) de même que dans le fait que la haute direction considère les mégadonnées comme une priorité stratégique pour l'organisation (TS03). L'analyse du coefficient d'inflation de variance (VIF) de cet indicateur de même que sa pertinence théorique indique qu'il ne présente pas de potentiels chevauchements conceptuels. Nous décidons donc de le conserver. Les indicateurs formatifs du modèle expliquent 86.3% de la variance ($R^2 = 0.863$).

Figure 16 : Modèle MIMIC du dynamisme de l'écosystème

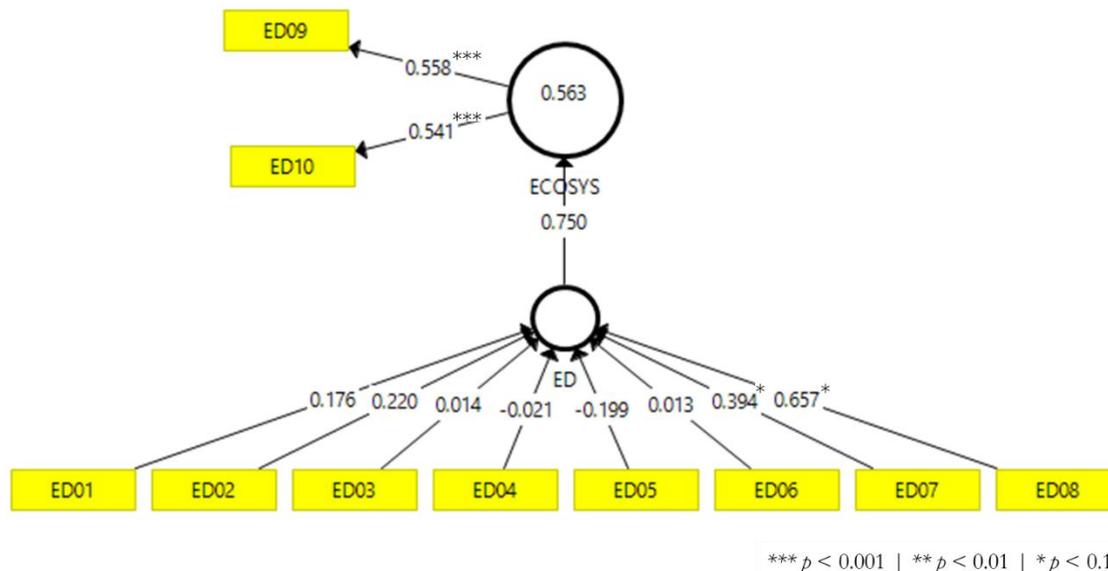
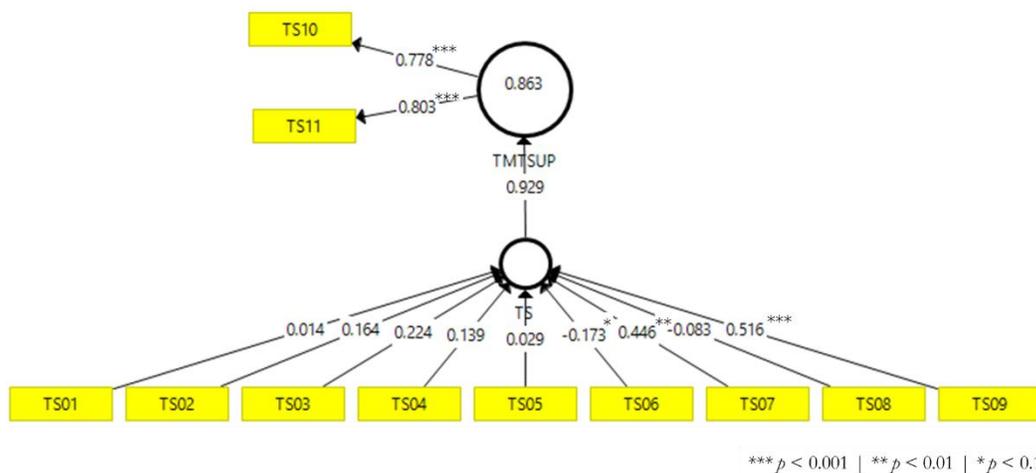


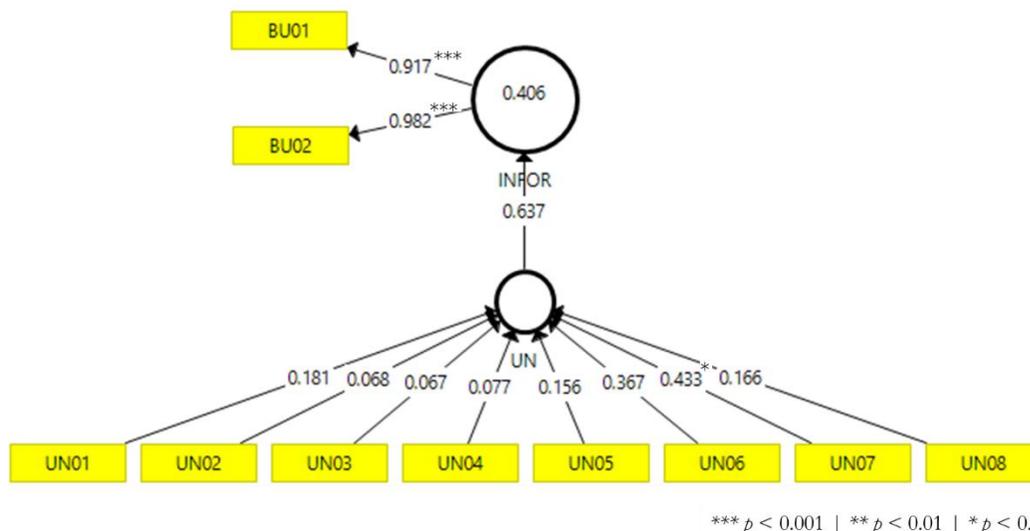
Figure 17 : Modèle MIMIC du support de la haute direction



Utilisation informationnelle : le modèle MIMIC de l'utilisation informationnelle des mégadonnées est présenté à la figure 18. Tous les items réflectifs sont significatifs ($\gamma_{BU01} = 0.917$ et $\gamma_{BU02} = 0.982$ avec $p < 0.001$). Des 8 indicateurs formatifs, un seul est significatif ($\gamma_7 = 0.433$ avec $p < 0.1$). Ceci indique que, pour les répondants, l'utilisation informationnelle des mégadonnées réside essentiellement dans la mise à disposition d'information à propos des services touristiques, tels que les attractions, les événements et les festivals (UN07). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que l'item suivant démontre une importance relative supérieure à 0.6 : informer les citoyens sur la circulation et les conditions routières (UN02), sur l'état d'avancement des opérations de nettoyage des rues ou du déneigement (UN05), sur les activités de

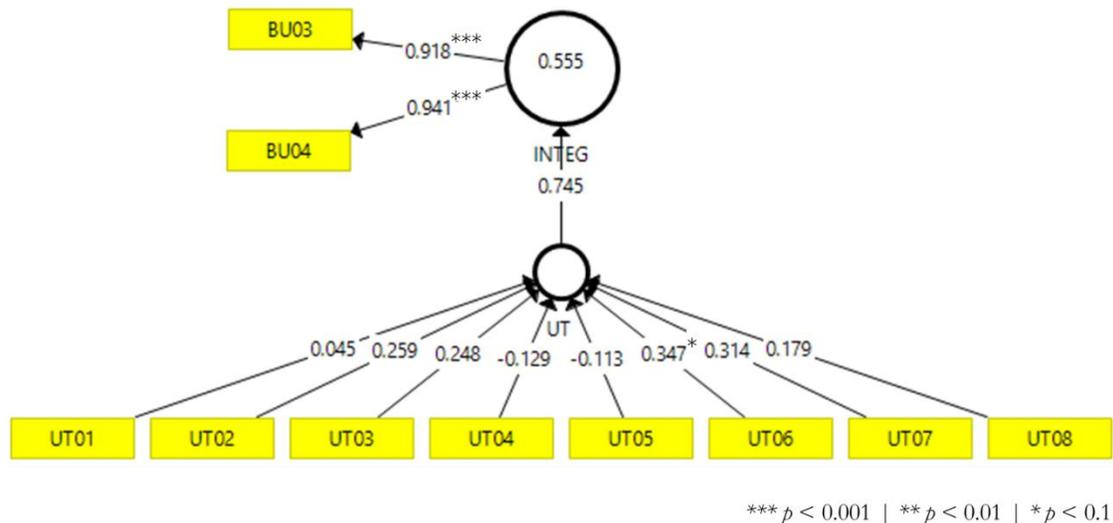
collecte, telles que les ordures ménagères et les matières recyclables (UN06) et sur les actes criminels commis sur le territoire de la ville (UN08). L'analyse du coefficient d'inflation de variance (VIF) de cet indicateur de même que sa pertinence théorique indique qu'il ne présente pas de potentiels chevauchements conceptuels. Nous décidons donc de le conserver. Les indicateurs formatifs du modèle expliquent 40.6% de la variance ($R^2 = 0.406$).

Figure 18 : Modèle MIMIC de l'utilisation informationnelle



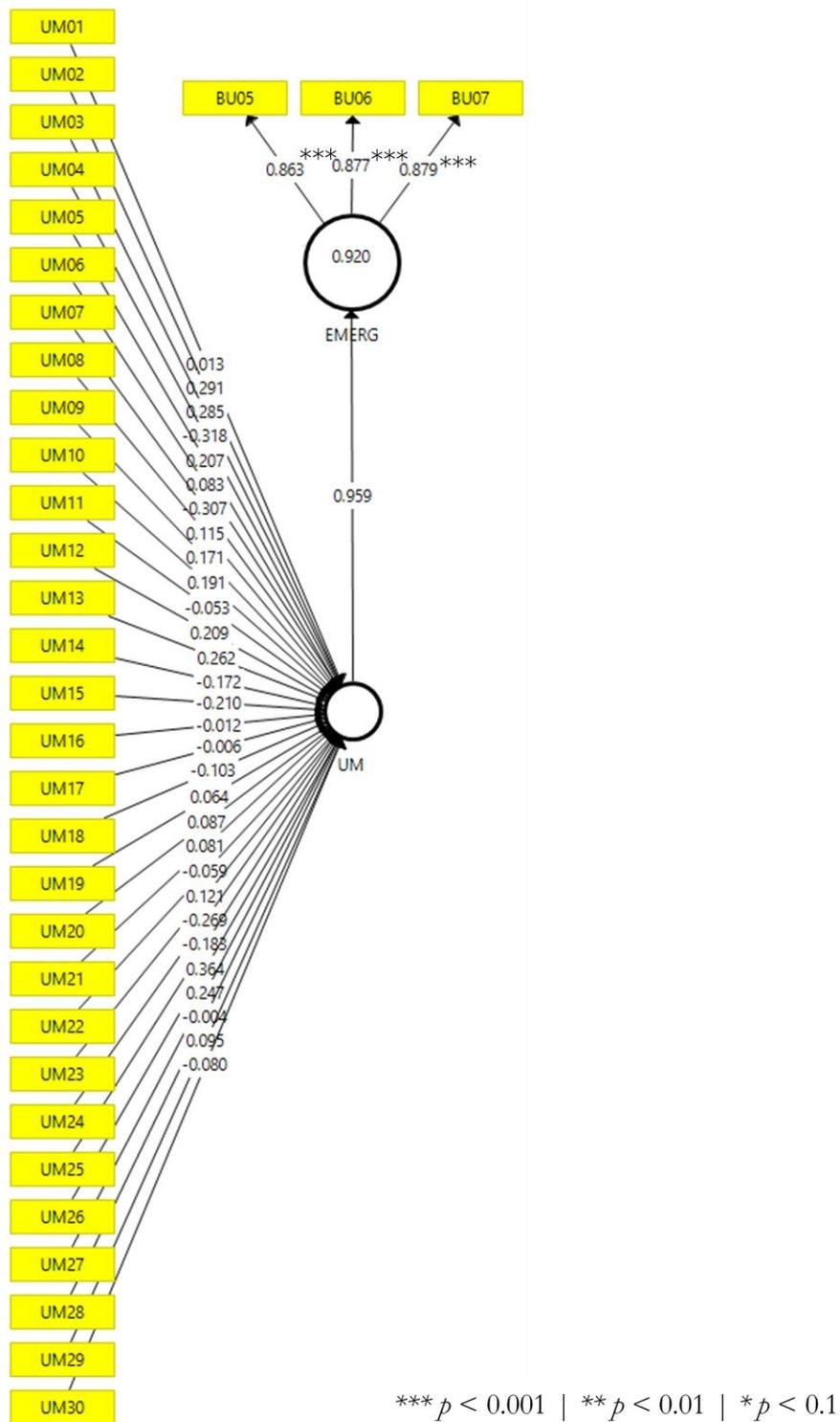
Utilisation intégrationnelle : le modèle MIMIC de l'utilisation intégrationnelle des mégadonnées est présenté à la figure 19. Tous les items réflectifs sont significatifs ($\gamma_{BU03} = 0.918$ et $\gamma_{BU04} = 0.941$ avec $p < 0.001$). Des 8 indicateurs formatifs, un seul est significatif ($\gamma_6 = 0.347$ avec $p < 0.1$). Ceci indique que, pour les répondants, l'utilisation intégrationnelle des mégadonnées réside essentiellement dans l'intégration des services offerts aux citoyens, par le biais de portails citoyens par exemple (UT06). L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que l'item suivant démontre une importance relative supérieure à 0.6 : alimentation d'entrepôts de données afin de disposer d'une source unique d'information (UT01), intégration des services avec les partenaires ou les organismes gouvernementaux (UT07), le désir de devenir un gouvernement ouvert et transparent (UT08), le partage des données avec les différents services/départements municipaux (UT02), avec d'autres municipalités (UT03), avec le gouvernement provincial ou central (UT04) ou encore avec les partenaires de la ville (UT05). L'analyse du coefficient d'inflation de variance (VIF) de cet indicateur de même que sa pertinence théorique indique qu'il ne présente pas de potentiels chevauchements conceptuels. Nous décidons donc de le conserver. Les indicateurs formatifs du modèle expliquent 55.5% de la variance ($R^2 = 0.555$).

Figure 19 : Modèle MIMIC de l'utilisation intégrationnelle



Utilisation émergente : le modèle MIMIC de l'utilisation émergente des mégadonnées est présenté à la figure 20. Tous les items réflectifs sont significatifs ($\gamma_{BU05} = 25.638$, $\gamma_{BU06} = 28.857$ et $\gamma_{BU07} = 24.696$ avec $p < 0.001$). Des 30 indicateurs formatifs, aucun ne semble être significatif. L'évaluation des coefficients de saturation des autres indicateurs de l'échelle de mesure indique que l'item suivant démontre une importance relative supérieure à 0.6 : les mégadonnées comme intrant aux tableaux de bord en temps réel (UM01), pour établir des prévisions, la planification ou la stratégie (UM03), l'amélioration de la prise de décision (UM02), le suivi des processus ou des équipements (UM04), la localisation des véhicules (de collecte des déchets, autobus, etc.) (UM05), l'utilisation efficace des ressources (UM08), l'automatiser des processus (UM09), le développement de nouveaux services ou produits (UM12), la promotion de l'entrepreneuriat et de l'innovation (UM13), la stimulation du développement économique local (UM14), l'amélioration de l'infrastructure cyclable (UM19), le soutien à la planification et le développement urbain (UM25) de même qu'aux services d'urgence (tels que les services médicaux, de sauvetage et de lutte contre les incendies) (UM26) et la sécurité du public et des lieux publics (UM29). L'analyse du coefficient d'inflation de variance (VIF) de cet indicateur de même que sa pertinence théorique indique qu'il ne présente pas de potentiels chevauchements conceptuels. Nous décidons donc de le conserver. Les indicateurs formatifs du modèle expliquent 92% de la variance ($R^2 = 0.920$).

Figure 20 : Modèle MIMIC de l'utilisation émergente



Conclusion :

De ce qui précède, nous constatons que l'analyse de la significativité des coefficients, réalisée par une procédure de Bootstrap, confirme la pertinence de l'ensemble des items réfléchifs des modèles. Toutefois, plusieurs indicateurs formatifs ne semblent pas significatifs même si la qualité de l'ajustement des modèles MIMIC est validée, du moins en partie, avec l'analyse de régression PLS. Cette analyse a permis de constater que les coefficients de détermination (R^2) présentent tous des valeurs satisfaisantes situées au-dessus de 0.30 (Chin, Marcolin et Newsted, 2003) et allant de 0.360 à 0.920.

Comme indiqué ci-dessus, Cenfetelli et Bassellier (2009 : 693) recommandent, en plus d'analyser l'importance absolue des indicateurs (Weights), d'évaluer leur importance relative (Loadings). En suivant les recommandations de ces auteurs, plusieurs indicateurs présentent une importance absolue faible, mais une importance relative importante. L'analyse des coefficients d'inflation de variance (VIF) de ces indicateurs, de même que leur pertinence théorique indique qu'ils ne présentent pas de potentiels chevauchements conceptuels, ce qui justifie donc leur maintien (Cenfetelli et Bassellier, 2009).

En ce qui concerne les indicateurs qui présentent une faible importance absolue conjuguée à une faible importance relative, Cenfetelli et Bassellier (2009) recommandent d'analyser leur pertinence théorique. Cette analyse nous permet de confirmer qu'il n'existe pas de potentiels chevauchements conceptuels entre les indicateurs en question (CT05, CT06, CT07, PC08, TC02, TC03, TC04, TC06, ED01, ED02, ED03, ED04, ED05, ED06, TS01, TS05, UN01, UN03, UN04, UM06, UM07, UM10, UM11, UM15, UM16, UM17, UM18, UM20, UM21, UM22, UM23, UM24, UM27, UM28, UM30).

Ainsi, et en prenant en considérant le contexte des villes participantes, la taille de l'échantillon de l'étude, la nature exploratoire de la présente étude, les procédures de validation des construits formatifs recommandées par la littérature et les effets négatifs du retrait des indicateurs formatifs sur la signification et le pouvoir prédictif des construits du modèle (Diamantopoulos et Winklhofer, 2001), nous choisissons de conserver l'ensemble des indicateurs des modèles MIMIC. Ces indicateurs pourront être testés lors de futurs travaux.

B. Test des modèles structurels

La validation des hypothèses sous-jacentes au modèle de recherche est opérée avec l'évaluation des trois (3) modèles structurels (utilisation informationnelle, intégrationnelle et émergente) représentant les relations présumées entre les variables indépendantes et dépendantes du modèle de recherche. Vu la taille de l'échantillon de la présente étude, nous avons choisi d'effectuer les tests sur des modèles réflectifs uniquement, le test du modèle formatif nécessitant une taille d'échantillon d'environ 300 réponses. De plus, la validité de la présente procédure est confirmée par l'analyse effectuée à la section 4.1.3 qui démontre que le modèle de recherche, mesuré avec des items réflectifs, présente une validité (convergente et discriminante) acceptable. Les modèles structurels ont été testés afin d'examiner le degré et l'ampleur des relations entre les variables endogènes et exogènes du modèle de recherche. Plus précisément, le test des modèles structurels, utilisant la technique SEM-PLS, permet d'examiner les relations hypothétiques à travers un ensemble d'indicateurs : les coefficients de régression qui indiquent les forces et la direction des relations entre les variables, les tests de Student qui indiquent l'importance de l'influence et, enfin, les valeurs de R^2 qui déterminent le pouvoir explicatif des modèles et la validité nomologique des variables endogènes. Comme suggéré par Falk et Miller (1992), nous utiliserons 0.1 comme valeur seuil pour la validité nomologique des variables endogènes. La procédure adoptée par l'étude est le ré-échantillonnage bootstrap en utilisant 5 000 rééchantillons. Le [tableau 22](#) synthétise les résultats des tests des trois modèles structurels.

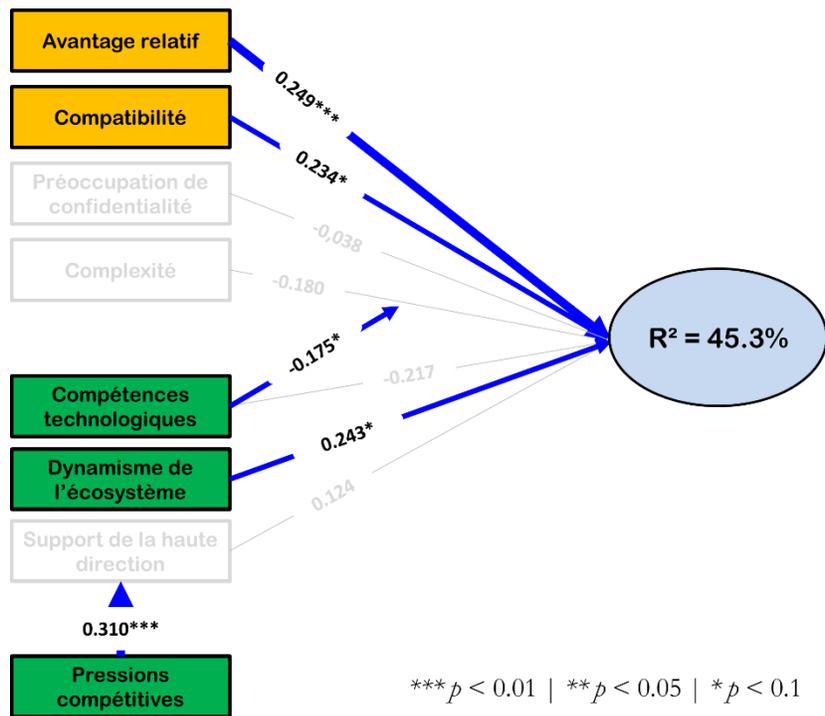
Nous avons dans un premier temps effectué le test du modèle complet (effet simultané des antécédents sur les trois formes d'utilisation). Dans un second moment, nous avons effectué le test des modèles structurel en fonction du type d'utilisation. Les résultats de ces deux tests étant sensiblement les mêmes, nous présenterons ci-dessous les résultats détaillés de chaque modèle structurel.

Tableau 22 : Synthèse des résultats des tests des modèles structurels

	Informationnelle $R^2 = 45.3\%$		Intégrationnelle $R^2 = 37.6\%$		Émergente $R^2 = 39.7\%$	
	T-Student	P Value	T-Student	P Value	T-Student	P Value
Avantage relatif \Rightarrow Utilisation	2.948	0.003	1.917	0.055	1.913	0.056
Compatibilité \Rightarrow Utilisation	1.823	0.068	1.024	0.306	2.235	0.025
Complexité \Rightarrow Utilisation	1.495	0.135	1.155	0.248	0.588	0.557
Préoccupations de confidentialité \Rightarrow Utilisation	0.737	0.461	0.535	0.592	0.229	0.819
Compétences technologiques \Rightarrow Utilisation	1.353	0.176	0.825	0.410	0.553	0.580
CT x CX \Rightarrow Utilisation	1.856	0.064	1.289	0.197	0.842	0.400
Support de la haute direction \Rightarrow Utilisation	1.327	0.185	2.071	0.038	2.037	0.042
Pressions compétitives \Rightarrow Support HD	2.876	0.004	2.975	0.003	2.966	0.003
Dynamisme de l'écosystème \Rightarrow Utilisation	1.943	0.052	1.574	0.115	1.079	0.281

Modèle structurel de l'utilisation informationnelle : le modèle structurel de l'utilisation informationnelle, illustré à la figure 21, présente un coefficient de détermination de 45.3% ($R^2 = 0,453$). Aussi, le coefficient de détermination du support de la haute direction s'établit à 12.6% ($R^2 = 0,126$). Les coefficients de régression des variables de l'avantage relatif ($\gamma_1 = 0,249$ avec $p \leq 0,01$), de la compatibilité ($\gamma_2 = 0,234$ avec $p \leq 0,1$), de l'effet modérateur des compétences technologiques sur la complexité ($\gamma_6 = -0,175$ avec $p \leq 0,1$), des pressions compétitives ($\gamma_8 = 0,310$ avec $p \leq 0,01$) et du dynamisme de l'écosystème ($\gamma_9 = 0,243$ avec $p \leq 0,1$) sont significatifs et valident ainsi les hypothèses H1a, H2a, H6a, H8a et H9a.

Figure 21 : Les antécédents de l'utilisation informationnelle



Modèle structurel de l'utilisation intégrationnelle : le modèle structurel de l'utilisation intégrationnelle, illustré à la figure 22, présente un coefficient de détermination de 37.6% ($R^2 = 0,376$). Aussi, le coefficient de détermination du support de la haute direction s'établit à 12.6% ($R^2 = 0,126$). Les coefficients de régression des variables de l'avantage relatif ($\gamma_1 = 0,198$ avec $p \leq 0,1$), du support de la haute direction ($\gamma_7 = 0,264$ avec $p \leq 0,05$) et des pressions compétitives ($\gamma_8 = 0,313$ avec $p \leq 0,01$) sont significatifs et valident ainsi les hypothèses H1b, H7b et H8b.

Modèle structurel de l'utilisation émergente : le modèle structurel de l'utilisation émergente, illustré à la figure 23, présente un coefficient de détermination de 39.7% ($R^2 = 0,397$). Aussi, le coefficient de détermination du support de la haute direction s'établit à 12.6% ($R^2 = 0,126$). Les coefficients de régression des variables de l'avantage relatif ($\gamma_1 = 0,160$ avec $p \leq 0,05$), de la compatibilité ($\gamma_2 = 0,270$ avec $p \leq 0,05$), du support de la haute direction ($\gamma_7 = 0,303$ avec $p \leq 0,05$) et des pressions compétitives ($\gamma_8 = 0,320$ avec $p \leq 0,01$) sont significatifs et valident ainsi les hypothèses H1c, H2c, H7c et H8c.

Figure 22 : Les antécédents de l'utilisation intégrationnelle

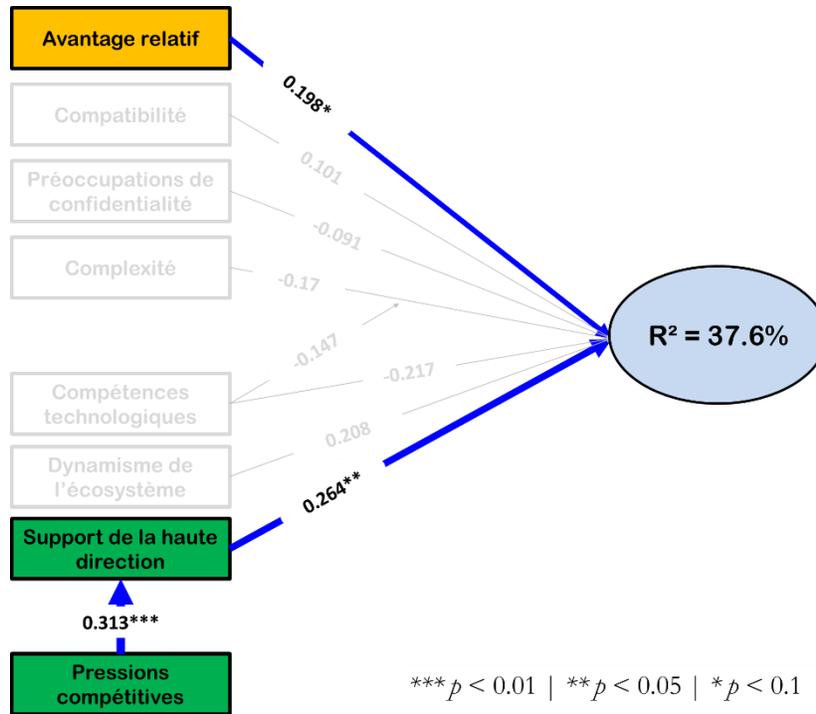
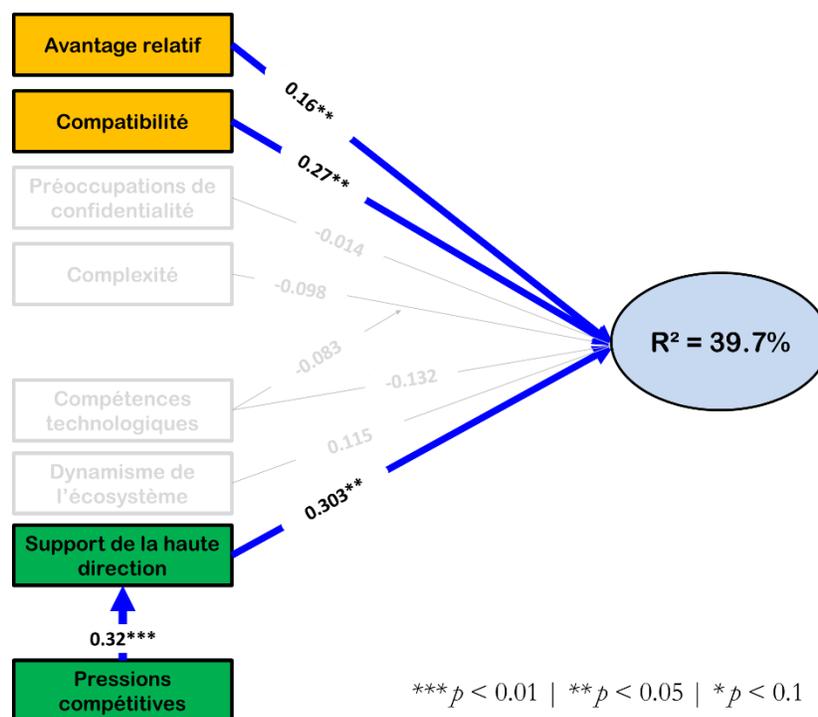


Figure 23 : Les antécédents de l'utilisation intégrationnelle



Nous avons effectué des tests afin de déterminer si les variables de contrôle ont un effet sur l'utilisation des mégadonnées. Ainsi, la taille des villes, mesurée par le nombre d'employés et la population, de même que la disponibilité de ressources financières ne se sont pas avérés significatifs sur les trois types d'utilisation. Nous avons également mesuré la force de l'effet de ces variables sur les trois types d'utilisation des mégadonnées et les résultats, synthétisé dans le [tableau 23](#) n'ont pas été satisfaisants indiquant que ces variables ont un très faible effet sur les variables endogènes (Chin, 1998b). De ce fait, il ne semble pas pertinent d'effectuer des analyses sur des sous-échantillons en se basant sur ces trois variables.

En définitive, l'analyse de régression des 3 modèles structurels permet de valider 12 des 27 hypothèses formulées à travers le modèle de recherche (tableau 8). Les trois modèles structurels d'intention d'utilisation informationnelle ($R^2 = 0.453$), intégrationnelle ($R^2 = 0.376$) et émergente ($R^2 = 0.397$) présentent des coefficients de détermination satisfaisants. Le test des modèles permet de constater que les antécédents sont différents en fonction de la forme d'utilisation des mégadonnées. La section suivante vise à discuter de l'ensemble des résultats de l'analyse des données collectées.

Tableau 23 : Résultats des tests sur les variables de contrôle

Variables de contrôle	Utilisation		
	Informationnelle	Intégrationnelle	Emergente
Nombre d'employés	$F^2 = 0.000 < 0.02$ $R^2 443 \Rightarrow 445$	$F^2 = 0.001 < 0.02$ $R^2 .368 \Rightarrow .372$	$F^2 = 0.003 < 0.02$ $R^2 .366 \Rightarrow .371$

Population	$F^2 = 0.006 < 0.02$ $R^2 .443 \Rightarrow .447$	$F^2 = 0.000 < 0.02$ $R^2 .368 \Rightarrow .370$	$F^2 = 0.015 < 0.02$ $R^2 .366 \Rightarrow .379$
Aspects financiers	$F^2 = 0.048 < 0.15$ $R^2 .443 \Rightarrow .437$	$F^2 = 0.02 < 0.15$ $R^2 .368 \Rightarrow .365$	$F^2 = 0.03 < 0.15$ $R^2 .366 \Rightarrow .365$

Tableau 24 : Synthèse de l'analyse des modèles structurels

	Variables	Coefficient de détermination (R ²)	Coefficient de régression	Validation des hypothèses
Utilisation informationnelle	Avantage relatif ⇒ Utilisation	45.3%	2.948***	H1a supportée
	Compatibilité ⇒ Utilisation		1.823*	H2a supportée
	Complexité ⇒ Utilisation		1.495	H3a non supportée
	Préoccupations de confidentialité ⇒ Utilisation		0.737	H4a non supportée
	Compétences technologiques ⇒ Utilisation		1.353	H5a non supportée
	CT x CX ⇒ Utilisation		1.856*	H6a supportée
	Support de la haute direction ⇒ Utilisation		1.327	H7a non supportée
	Pressions compétitives ⇒ Support HD		2.876***	H8a supportée
	Dynamisme de l'écosystème ⇒ Utilisation		1.943*	H9a supportée
Utilisation intégrationnelle	Avantage relatif ⇒ Utilisation	37.6%	1.917*	H1b supportée
	Compatibilité ⇒ Utilisation		1.024	H2b non supportée
	Complexité ⇒ Utilisation		1.155	H3b non supportée
	Préoccupations de confidentialité ⇒ Utilisation		0.535	H4b non supportée
	Compétences technologiques ⇒ Utilisation		0.825	H5b non supportée
	CT x CX ⇒ Utilisation		1.289	H6b non supportée
	Support de la haute direction ⇒ Utilisation		2.071**	H7b supportée
	Pressions compétitives ⇒ Support HD		2.975***	H8b supportée
	Dynamisme de l'écosystème ⇒ Utilisation		1.574	H9b non supportée
Utilisation émergente	Avantage relatif ⇒ Utilisation	39.7%	1.913*	H1c supportée
	Compatibilité ⇒ Utilisation		2.235**	H2c supportée
	Complexité ⇒ Utilisation		0.588	H3c non supportée
	Préoccupations de confidentialité ⇒ Utilisation		0.229	H4c non supportée
	Compétences technologiques ⇒ Utilisation		0.553	H5c non supportée
	CT x CX ⇒ Utilisation		0.842	H6c non supportée
	Support de la haute direction ⇒ Utilisation		2.037**	H7c supportée
	Pressions compétitives ⇒ Support HD		2.966***	H8c supportée
	Dynamisme de l'écosystème ⇒ Utilisation		1.079	H9c non supportée

4.2 Discussion des résultats

La précédente section a permis d'analyser les résultats d'une collecte de données réalisée par le biais d'une enquête de terrain par questionnaire, administrée à 106 municipalités canadiennes. L'évaluation des qualités psychométriques des échelles de mesures élaborées dans le cadre de la présente étude a démontré leur fiabilité et leur validité. En outre, le test des modèles MIMIC offre une conceptualisation précise des différents types d'utilisation des mégadonnées par les villes. Enfin, le test des modèles structurels a permis de tester les hypothèses du modèle de recherche. Nous discuterons, dans un premier temps, les antécédents de l'utilisation des mégadonnées et, dans un second temps, les formes de l'utilisation des mégadonnées par les villes.

Le modèle structurel explique 45.3% de la variance dans l'utilisation informationnelle, 37.6% de la variance dans l'utilisation intégrationnelle et 39.7% de la variance dans l'utilisation émergente. Les antécédents en fonction du type de l'utilisation sont indiqués dans le [tableau 25](#).

Tableau 25 : Comparatif des antécédents de l'utilisation des mégadonnées par les villes

	Informationnelle	Intégrationnelle	Émergente
Avantage relatif ⇒ Utilisation	✓***	✓*	✓*
Compatibilité ⇒ Utilisation	✓*		✓**
Complexité (CT) ⇒ Utilisation			
Compétences technologiques (TC) ⇒ Utilisation			
CT x TC ⇒ Utilisation	✓*		
Support de la haute direction (HD) ⇒ Utilisation		✓**	✓**
Dynamisme de l'écosystème ⇒ Utilisation	✓*		
Pressions compétitives ⇒ Support HD	✓***	✓***	✓***
Préoccupations de confidentialité ⇒ Utilisation			

Le modèle structurel explique 45.3% de la variance dans l'utilisation informationnelle. Les antécédents significatifs sont l'avantage relatif, la compatibilité, le dynamisme de l'écosystème et l'effet modérateur des compétences technologiques sur la complexité des mégadonnées. De plus, l'effet des pressions compétitives sur le support de la haute direction s'est avéré significatif.

Le modèle structurel explique 37.6% de la variance dans l'utilisation intégrationnelle. Les antécédents significatifs sont l'avantage relatif, le support de la haute direction qui lui-même est expliqué par les pressions compétitives.

Le modèle structurel explique 39.7% de la variance dans l'utilisation émergente. Les antécédents significatifs sont l'avantage relatif, la compatibilité et le support de la haute direction qui lui-même est expliqué par les pressions compétitives.

Les antécédents de l'utilisation des mégadonnées résultent de l'intégration de deux cadres conceptuels : le cadre technologie-organisation-environnement (TOE) de Tornatzky et Fleischer (1990), la théorie de la diffusion de l'innovation (DOI) de Rogers (2003).

L'avantage relatif des mégadonnées s'est avéré être un antécédent significatif pour les trois types d'utilisation et plus particulièrement pour l'utilisation informationnelle ($p < 0.01$). La perception de l'avantage relatif des mégadonnées par les villes permet donc de prédire significativement les trois formes d'utilisation de leur utilisation.

Par ailleurs, la compatibilité des mégadonnées s'est avérée être un antécédent significatif pour deux types d'utilisation (utilisation informationnelle et utilisation émergente). L'analyse de la puissance de l'effet de la compatibilité sur l'utilisation confirme qu'elle a un effet modéré sur l'utilisation émergente et l'utilisation informationnelle ($0.02 > F^2 > 0.15$) et aucun effet sur l'utilisation intégrationnelle ($F^2 < 0.02$). Le fait que cet antécédent n'a pas été significatif pour l'utilisation intégrationnelle peut suggérer que cette utilisation est motivée par des facteurs tels que la réciprocité dans l'échange de données avec les parties prenantes et la nécessité de répondre à certaines obligations notamment envers les autres paliers du gouvernement. D'une manière générale, la pertinence de la compatibilité des mégadonnées suggère que leurs caractéristiques peuvent ne pas être jugées compatibles avec les processus, les pratiques et les systèmes de valeurs existants au sein de certaines organisations. Ce manque de compatibilité peut entraîner une résistance organisationnelle et/ou populaire ce qui influe négativement sur leur utilisation.

Le support de la haute direction s'est avéré être un antécédent significatif pour deux types d'utilisation (utilisation intégrationnelle et utilisation émergente). L'analyse de la puissance de l'effet du support de la haute direction sur l'utilisation des mégadonnées confirme que ce construit possède un effet modéré sur l'utilisation intégrationnelle et émergente ($0.02 > F^2 > 0.15$) et aucun effet sur l'utilisation informationnelle ($F^2 < 0.02$). Toutefois, le T-statistique de l'effet de l'engagement de la haute direction sur l'utilisation informationnelle est proche du seuil d'acceptation ($T = 1.324$) ce qui indique qu'un échantillon de plus grande taille permettrait d'obtenir des résultats satisfaisants. D'une manière générale, la compréhension et l'appréciation des capacités des mégadonnées par les membres de la haute direction, de même que le degré de leur implication dans les initiatives impliquant les mégadonnées est un facteur clé dans l'utilisation des mégadonnées. La création de valeur découle donc d'une étroite collaboration du personnel d'analyste avec les décideurs (Sharma, Mithas et Kankanhalli, 2014). Le support de la haute direction à l'utilisation des mégadonnées devrait ainsi se traduire par des actions concrètes et se cristallise avec deux principaux mécanismes : l'engagement financier envers les mégadonnées et le niveau décisionnel liés aux mégadonnées.

Le dynamisme de l'écosystème des mégadonnées s'est avéré être un antécédent significatif pour l'utilisation informationnelle des mégadonnées. L'analyse de la puissance de l'effet du dynamisme de l'écosystème sur l'utilisation des mégadonnées indique que ce construit possède un effet modéré sur l'utilisation informationnelle et intégrationnelle ($0.02 > F^2 > 0.15$) et aucun effet sur l'utilisation émergente ($F^2 < 0.02$). Ceci indique qu'un échantillon de plus grande taille permettrait d'obtenir des résultats satisfaisants concernant l'effet de l'engagement de la haute direction sur l'utilisation intégrationnelle. La valeur du T-statistique, proche du seuil d'acceptation ($T = 1.588$) confirme cette hypothèse. L'utilisation des mégadonnées par les villes ne se limite pas au personnel municipal et la disponibilité d'un écosystème d'entreprises, d'organismes, d'incubateurs et de communautés de

développeurs dans le domaine des mégadonnées est nécessaire pour la réalisation d'initiatives utilisant les mégadonnées. Pour les villes participantes, le dynamisme de l'écosystème dans leurs villes réside essentiellement dans la mise en œuvre d'une stratégie ou d'une politique de mégadonnées par le gouvernement central de même que dans la disponibilité d'un support législatif à favorisant l'accès aux données dont la ville a besoin.

Les pressions compétitives se sont avérées être un antécédent significatif pour le support de la haute direction, et ce pour les trois types d'utilisation des mégadonnées. Ceci suggère que les contraintes de l'environnement, exercée par les pairs en adoptant et en utilisant les mégadonnées amènent les villes à embrasser les technologies des mégadonnées adoptées par les concurrents et partenaires. De plus en plus de villes accentuent leurs efforts de compétitivité en menant des initiatives numérisation (Lee et Lee, 2014).

Les compétences technologiques de même que la complexité des mégadonnées ne se sont pas avérées significatives pour les trois types d'utilisation des mégadonnées.

- Même si les mégadonnées détiennent des caractéristiques qui indiquent qu'il s'agit d'une innovation complexe à comprendre et à utiliser, la complexité ne s'est pas avérée être un antécédent significatif pour leur utilisation. Ceci peut se justifier par deux aspects. D'une part, l'effet négatif de la complexité sur l'adoption a été significatif dans l'étude de Thong (1999) alors que son effet n'a pas été significatif sur la diffusion du commerce électronique dans l'étude de (Lin, 2008). Ainsi, ce construit ne semble pas avoir une base théorique solide dans la littérature post-adoption des innovations. D'autre part, il est probable que certains répondants ne détiennent pas les connaissances nécessaires des enjeux et défis que posent les mégadonnées.
- Le construit des «compétences technologiques» des villes capte la capacité des infrastructures technologiques des villes, ainsi que les compétences liées aux mégadonnées des employés. Même si la littérature a démontré la pertinence de la compétence technologique, il ne s'est pas avéré pertinent pour la présente étude. De même que pour la complexité, nous pensons que le manque de connaissances de certains répondants des capacités technologiques nécessaires à l'utilisation des mégadonnées aurait affecté la significativité de ce construit.
- Enfin, l'effet modérateur des compétences technologiques sur la complexité des mégadonnées s'est avéré significatif pour l'utilisation informationnelle seulement ($p < 0.1$). Ceci suggère que les compétences technologiques atténuent la complexité des mégadonnées lorsqu'il s'agit d'une utilisation relativement simple et n'ont pas d'effet lorsque l'utilisation visée présente une plus grande complexité technique (utilisation intégrationnelle ou émergente).

Par ailleurs, les préoccupations de confidentialité ne se sont pas avérées être un antécédent significatif pour l'utilisation des mégadonnées. Ceci suggère que les gestionnaires publics ne sont probablement pas conscients des questions juridiques, éthiques et techniques affectant les cadres réglementaires existants et que l'utilisation des mégadonnées peut comporter des risques liés à la possibilité de relever des informations confidentielles ou sensibles, ou des données exclusives. Par ailleurs, comme la présente étude est la seule, à notre connaissance, à intégrer ce construit, il se peut

que l'échelle qui a été élaborée pour le mesurer mérite d'être révisée et que des exemples de risques de confidentialité soient intégrés dans de futurs questionnaires.

Ainsi, ces résultats offrent un support à la théorie DOI et au cadre TOE et renforcent leur pouvoir explicatif tout en enrichissant leur contexte d'application (ville) et l'innovation étudiée (mégadonnées). D'une manière générale, les données ne sont d'aucune utilité jusqu'à ce qu'ils soient dans une forme utilisable Ackoff (1989). Cette règle s'applique également aux mégadonnées qui, sans l'application de techniques d'analyse, ce n'est que du bruit (Boyd et Crawford, 2012 : 667). Ainsi, le fait de disposer de jeux de données ne suffit pas, et qu'un processus de création de la valeur à partir des données brutes doit avoir lieu. C'est de ce processus de transformation qu'il a été question dans la présente section. Le prochain chapitre vise à synthétiser les apprentissages acquis dans le cadre de la présente étude.

**“ Numbers have an important story to tell.
They rely on you to give them a voice ”**
Stephen Few

Chapitre 5 : Conclusion

Le présent chapitre dresse un portrait des principaux apprentissages réalisés dans le cadre de la présente étude. Dans un premier temps, un rappel des objectifs de recherche et de la méthodologie utilisée sera présenté. Il sera suivi de la présentation des principaux résultats de l'étude, des contributions pratiques et théoriques, de l'identification des limites et, enfin, de la présentation des perspectives de futures recherches.

5.1 Rappel des objectifs de l'étude et de l'approche méthodologique

En dépit des avantages que peuvent procurer les mégadonnées, et en dépit du fait que l'utilisation de l'information est profondément ancrée dans les fonctions essentielles des gouvernements, et même si les organisations ont de plus en plus accès à de grandes quantités de données, quelques-unes d'entre elles seulement les exploitent (Martens et al., 2016; Tachizawa, Alvarez-Gil et Montes-Sancho, 2015). Ainsi, l'objectif de la présente étude était de comprendre les mécanismes sous-jacents à l'utilisation des mégadonnées par les villes et de conceptualiser les différentes formes de leur utilisation aux étapes post-adoption du cadre de Kwon et Zmud (1987). Découlant de ce principal objectif, l'étude visait également le développement d'un cadre conceptuel et d'un modèle de recherche, la clarification de la notion de mégadonnées dans un contexte municipal et la création d'échelles permettant de mesurer à la fois les antécédents influençant l'utilisation des mégadonnées et l'utilisation des mégadonnées elle-même.

Afin d'atteindre ces objectifs, une revue systématique de la littérature (RSL) portant sur les mégadonnées dans un contexte de ville a été effectuée. Les résultats de cette revue ont alimenté les éléments constitutifs d'un modèle conceptuel, qui a été traduit en modèle de recherche. Afin de tester les hypothèses du modèle de recherche, nous avons besoin de concevoir les échelles de l'instrument de mesure et de collecter des données primaires. Un questionnaire est conçu sur la base d'une profonde compréhension de la littérature et une enquête de terrain a été menée auprès de 106 municipalités canadiennes.

5.2 Synthèse des principaux résultats

Chacune des formes d'utilisation présente un ensemble d'antécédents différent. D'une manière générale, et en cohérence avec la littérature sur les antécédents organisationnels portant sur la diffusion et l'utilisation des innovations, le support de la haute direction, l'avantage relatif, la compatibilité des mégadonnées et le dynamisme de l'écosystème des mégadonnées se sont avérés des antécédents significatifs positifs pour l'utilisation des mégadonnées par les villes.

La présente étude a également permis de constater que les éléments de l'environnement des villes sont des antécédents significatifs positifs pour l'utilisation des mégadonnées à travers l'effet modérateur du support de la haute direction. Ainsi, les pressions compétitives ont un effet positif sur le support de la haute direction qui, à son tour, affecte positivement l'utilisation des mégadonnées. Ces résultats sont en cohérence avec les travaux de D. Chen, Preston et Swink (2015) et de Liang *et al.* (2007). Curieusement, la complexité des mégadonnées et les compétences technologiques des

villes n'ont pas été des antécédents significatifs dans la présente étude même si la littérature sur le sujet est abondante. De plus, l'effet modérateur des compétences technologiques est significatif uniquement pour l'utilisation informationnelle. Enfin, les « préoccupations de confidentialités » n'a pas été un antécédent significatif dans la présente étude.

En somme, deux des quatre caractéristiques des mégadonnées provenant de la théorie de diffusion de l'innovation de Rogers ont été validées. Il s'agit de l'avantage relatif et de la compatibilité. En ce qui concerne les éléments contextuels du cadre TOE de Tornatzky et Fleischer (1990), deux ont été validés dans la présente étude : le contexte organisationnel (support de la haute direction) et le contexte environnemental (pressions compétitives).

5.3 Apports de l'étude

La présente étude présente plusieurs contributions théoriques et pratiques.

En effet, en dépit de l'importance de la conceptualisation de l'utilisation des technologies de l'information au niveau des villes, la littérature sur le sujet reste peu abondante (Titah et Barki, 2006). À notre connaissance, la présente étude est la première à étudier les antécédents et l'utilisation des mégadonnées par les villes, durant les phases post-adoption du cadre de Kwon et Zmud (1987). De plus, cette étude figure parmi les rares études empiriques portant sur la valorisation des mégadonnées au sein des villes, alors la majorité de la littérature est de nature conceptuelle (Chauhan, Agarwal et Kar, 2016).

À travers une revue systématique de la littérature, la présente étude offre une contribution pertinente aux connaissances sur les mégadonnées dans un contexte urbain. Cette revue nous a ainsi permis de constater que très peu d'études portent sur l'utilisation des mégadonnées par les villes et que les mégadonnées demeurent un sujet en émergence. La clarification des caractéristiques des mégadonnées permet de mettre en perspective les enjeux à prendre en considération lors de la gestion des mégadonnées dans les villes. La revue de littérature a également permis de catégoriser l'ensemble des antécédents de l'utilisation des diverses innovations (Big Data Analytics, e-business, EDI, internet). Cette catégorisation permet de disposer d'un ensemble complet d'antécédents à prendre en compte dans les initiatives de diffusion des innovations aux phases de post-adoption du cadre de Kwon et Zmud (1987). Cet ensemble peut être pertinent pour les futurs travaux de même que pour les praticiens. Les antécédents de l'utilisation des mégadonnées peuvent donc être pris en compte par les gestionnaires municipaux lors de la conception de stratégies de transformation de leurs villes. Ces antécédents constituent donc des facteurs clés de succès des projets et initiatives mobilisant les mégadonnées.

La présente étude a également permis de mettre en perspectives les avantages des mégadonnées dans un contexte de ville. Aussi, si les mégadonnées procurent plusieurs avantages aux villes, plusieurs défis doivent être relevés pour une gestion efficace des mégadonnées et cette étude devrait donc aider les praticiens, et notamment les gestionnaires des TI au sein des villes, à prendre conscience de ces défis. Aussi, elle constitue une tentative de proposition de solutions pratiques aux gestionnaires publics et devrait contribuer à offrir un aperçu de la façon dont certains de ces défis devraient être abordés.

En contribuant à la compréhension de la manière dont les mégadonnées sont utilisées dans un contexte où les actions prises peuvent avoir un impact sur la vie et le quotidien de milliers de citoyens, cette étude devrait également inspirer et aider les praticiens, et notamment les gestionnaires des TI au sein des villes, à innover dans l'élaboration des stratégies de transformation de leurs villes.

Par ailleurs, la présente étude offre une contribution pertinente en conceptualisant, d'une manière précise, le dynamisme de l'écosystème des mégadonnées qui est une des dimensions de l'environnement des villes. La pertinence de cette conceptualisation réside également dans l'intégration de deux antécédents étudiés dans la littérature : la disponibilité d'un support réglementaire qui est réside dans les actions prises par le gouvernement afin d'encourager l'adoption et l'utilisation d'une technologie (Zhu, Kraemer et Dedrick, 2004; Zhu et Kraemer, 2005) et la prédisposition des partenaires d'affaires, qui est la mesure dans laquelle les partenaires disposent de technologies similaires ou compatibles avec l'innovation en question (Mishra, Konana et Barua, 2007; Zhu et al., 2006).

Par ailleurs, l'étude contribue à combler un manque constaté dans la littérature sur les systèmes d'information. En effet, contrairement à la riche littérature qui se concentre sur la prédiction de l'adoption des TI par les organisations à partir de l'analyse des intentions au niveau individuel, les études sur les comportements de post-adoption au niveau organisationnel restent relativement rares, malgré l'importance de ce phénomène (D. Chen, Preston et Swink, 2015 ; Thong, 1999). Traitant de l'utilisation des mégadonnées au niveau organisationnel, la présente étude a permis de fournir une profonde compréhension des enjeux qui découlent des caractéristiques des mégadonnées et des implications de la diffusion de cette innovation au sein des municipalités. Aussi, en mettant l'accent sur l'utilisation post-adoption des mégadonnées, la présente étude a permis d'obtenir une meilleure compréhension de la manière dont les mégadonnées peuvent être utilisées dans le processus de transformation des villes.

La présente étude s'inscrit également en réponse aux appels de la littérature à constituer des assises solides au phénomène des mégadonnées (Goes, 2014) et constitue une tentative visant à contribuer au renforcement des fondements théoriques du domaine de l'étude des systèmes d'information. Elle constitue aussi une réponse aux appels de plusieurs auteurs à ce que la littérature sur les systèmes d'information prenne le leadership en étudiant les implications des mégadonnées dans les organisations et en dirigeant la manière dont les transformations organisationnelles doivent être menées (Goes, 2014 ; Mergel, Rethemeyer et Isett, 2016).

Pour les besoins de l'étude, nous avons intégré trois perspectives théoriques afin d'analyser et expliquer les forces qui favorisent les initiatives organisationnelles d'utilisation des mégadonnées : le cadre technologie-organisation-environnement (TOE) de Tornatzky et Fleischer (1990), la théorie de la diffusion de l'innovation (DOI) de Rogers (2003) et le modèle d'utilisation du gouvernement électronique de Titah (2010). Cette combinaison a permis l'obtention d'un modèle cohérent et complet de l'utilisation organisationnelle des mégadonnées par les villes. Ce modèle offre un cadre théorique pertinent pour de futurs travaux qui peuvent porter sur une variété d'innovations, au sein d'organisations opérant dans des contextes différents. Toutefois, ce cadre est fortement recommandé pour l'étude des innovations qui transforment les lignes d'affaires des organisations, qui permettent d'étendre les produits et services offerts et dont les implications de leur utilisation dépassent les frontières organisationnelles.

Enfin, sur le plan psychométrique, l'opérationnalisation du modèle de recherche en fait un instrument fiable et valide, permettant de prédire significativement les antécédents pour les trois (3) formes de l'utilisation des mégadonnées par les villes.

5.4 Limites de l'étude et perspectives de recherche

La présente étude n'est pas sans limites. Ces limites jettent les bases d'une série de recommandations pour les futurs travaux de recherche.

Premièrement, l'étude a principalement porté sur les étapes de post-adoption du processus d'implantation des mégadonnées au sein des villes du cadre de Kwon et Zmud (1987). Toutefois, nous ne pouvons déterminer à quelle phase se retrouve chacune des municipalités participantes (acceptation, routinisation ou infusion) et nous ne pouvons évaluer avec précision le degré de maturité des technologies utilisées par les villes participantes à l'étude. La sophistication de l'utilisation est le seul indicateur qui permet d'évaluer cette maturité. Ceci n'empêche pas le fait que certaines villes soient à divers stades (adaptation, acceptation, routinisation ou infusion) du processus d'implantation.

De plus, même si des définitions ont été intégrées au questionnaire, et en dépit des efforts déployés pour que le questionnaire soit le plus clair et le plus précis possible, les répondants peuvent avoir des perceptions différentes de ce que sont les « mégadonnées », « l'utilisation des mégadonnées » et des entités utilisatrices des mégadonnées. De plus, les mégadonnées peuvent être utilisées par différents départements et unités rattachées aux municipalités, ou même par des entités privées. La connaissance des répondants en lien avec l'utilisation des mégadonnées peut donc se limiter à certains départements de leur municipalité et peut ne pas couvrir l'ensemble des utilisations effectuées au sein de la ville.

Aussi, la conceptualisation des construits du modèle de recherche reste perfectible. Précisément, plusieurs des indicateurs formatifs présentent ainsi une marge d'erreur supérieure au seuil de confiance établi. Bien que la littérature n'offre pas de consensus sur les procédures de validation des échelles formatives, et que le retrait d'indicateurs affecte le pouvoir prédictif du modèle, les échelles de mesure auraient pu être améliorées à travers des entrevues avec des experts du domaine.

En outre, le questionnaire a été l'approche retenue pour la collecte de données et pour la validation du modèle de recherche. Il limite la profondeur de l'analyse vu qu'il ne permet pas de recueillir toute la richesse de l'expérience détenue par les répondants. Une étude avec une approche qualitative ou combinant plusieurs approches peut donc être considérée pour pallier cette limite. Aussi, une recherche longitudinale permettrait de mesurer l'influence des antécédents à différentes phases de post-adoption et de prédire avec plus de précision l'utilisation des mégadonnées par les villes.

Aussi, une des prémisses de l'étude a été de considérer que les villes ont accès aux données dont elles ont besoin. Cet accès a été mesuré par l'item ED09, avec une échelle Likert à 7 niveaux. Les résultats indiquent une moyenne 3.33, une médiane de 3 et un écart-type de 1,7. Ceci peut signifier que certaines villes peuvent ne pas avoir accès aux mégadonnées.

Par ailleurs, même si une série d'antécédents se sont révélés significatifs dans le cadre de la présente étude et expliquent une variance de plus de 0.3 pour les trois formes d'utilisation, nous estimons que le modèle de recherche reste également perfectible.

- Bien que son effet ne s'est pas avéré significatif pour l'échantillon à l'étude, nous pensons que la complexité des mégadonnées mérite une attention particulière dans les futurs travaux (Thong, 1999). Ils devraient également s'attarder sur l'étude des effets sur les personnes et leurs environnements institutionnels et sociaux (Marchand et Peppard, 2013 : 7). Nous pensons que la réussite des initiatives de mégadonnées dépend de la prise en compte de l'ensemble de ces aspects, et sur lesquels de futurs travaux peuvent être conduits (Mergel, Rethemeyer et Isett, 2016).
- Une attention particulière doit également être prêtée aux problèmes de confidentialité qui ont souvent été négligés dans la littérature sur les étapes post-adoption des innovations. Pourtant, l'utilisation des mégadonnées comporte des risques liés à la possibilité de relever des informations confidentielles ou sensibles, ou des données exclusives (van Zoonen, 2016). Les gestionnaires publics doivent être conscients des questions juridiques, éthiques et techniques affectant les cadres réglementaires existants.
- Même si la variable « Compétences technologiques » n'a pas été significative pour l'échantillon à l'étude, nous estimons qu'une attention particulière doit y être portée. En cohérence avec les études de Zhu et Kraemer (2005) et de Zhu et al. (2006), nous avons intégré sous ce construit deux aspects qu'on retrouve dans la littérature : l'impact de la sophistication de l'infrastructure TI (Armstrong et Sambamurthy, 1999) et le niveau de connaissances en TI des employés (Thong, 1999). Une des pistes à explorer est de séparer ce construit selon ces deux aspects et de prendre en compte les conclusions des travaux de Li, Hsieh et Rai (2013), qui discutent de l'utilisation individuelle post-adoption dans un contexte d'intelligence d'affaires, dans l'élaboration du second construit. D'ailleurs, Pearlson et Saunders (2004) et Awad et Ghaziri (2004) suggèrent que plus on monte dans les niveaux supérieurs de la hiérarchie DIKW plus l'apport de l'humain augmente et plus le rôle de la technologie baisse.
- L'une des prémisses de l'étude est que l'engagement de la haute direction détermine les ressources financières allouées aux utilisations des mégadonnées. L'analyse de la variable de contrôle portant sur les ressources financières disponibles pour l'utilisation des mégadonnées indique que la majorité (74.6%) des répondants estime que leur ville dispose de peu ou de très peu de ressources financières à consacrer pour l'utilisation des mégadonnées. La capacité financière, mesurée par une échelle Likert inversée à 5 niveaux, présente ainsi une moyenne et une médiane de 4 et un écart-type de 0.92. Ceci indique que l'introduction du construit « Engagement financier » pourrait être pertinente pour les futurs travaux.
- Il existe plusieurs raisons pour lesquelles la valeur générée par l'utilisation des mégadonnées peut être limitée (D. Chen, Preston et Swink, 2015) et un rapport de Gartner (2013) indique qu'une grande confusion entoure les mégadonnées et une incertitude existe en ce qui trait à leur capacité à générer des bénéfices tangibles pour les organisations. L'observabilité, qui est la mesure dans laquelle les résultats d'une innovation sont visibles par les autres, pourrait être une des caractéristiques des mégadonnées à introduire dans les futurs travaux.

- D'autres dimensions plus spécifiques peuvent être introduites dans les futurs travaux afin d'expliquer l'utilisation des mégadonnées par les villes aux étapes de post-adoption. Il s'agit par exemple de la participation électronique des citoyens (Guillamón et al. (2016), de la connaissance organisationnelle du processus de valorisation des données (Mishra, Konana et Barua, 2007) ou de la vision des politiciens en lien avec les mégadonnées.

La présente étude s'est concentrée sur l'étude des antécédents de l'utilisation des mégadonnées, et qui correspond au modèle d'utilisation du cadre conceptuel d'utilisation et d'impact du GE (voir Figure X?). Comme la présente étude a permis de conceptualiser l'utilisation des mégadonnées par les villes, de futurs travaux pourront porter sur l'étude de la performance et de la réalisation de bénéfices découlant de l'utilisation des mégadonnées. Ces travaux pourraient examiner pourquoi et comment les organisations peuvent générer de la valeur à partir des mégadonnées et proposer les approches qui doivent être suivies.

Enfin, sur le plan méthodologique, notre échantillon constitué de municipalités canadiennes limite la généralisation des résultats de cette étude. Nous pensons que l'utilisation des mégadonnées présente de grandes disparités entre les différentes régions du monde. Ainsi, et dans le cadre de futurs travaux, un échantillon constitué de villes à travers le monde pourrait donner des résultats statistiquement généralisables. L'échantillon pourrait être constitué, par exemple, de villes ayant des initiatives ou ayant reçu des prix de « villes intelligentes ».

Annexes

Annexe 1 : Listes des requêtes effectuées pour la revue de la littérature

Annexe 2 : Liste des articles retenus pour la revue systématique de la littérature

Annexe 3 : Certificat d'approbation éthique

Annexe 4 : Mise en application des principales exigences éthiques

Annexe 5 : Littérature sur les antécédents organisationnels de l'utilisation des TI

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Annexe 7 : Items de mesure des variables du modèle de recherche

Annexe 8 : Structure de codification des items de mesure

Annexe 9 : Données démographiques

Annexe 10 : Contributions factorielles croisées

Annexe 11 : Significativité des items des modèles MIMIC

Annexe 1 : Listes des requêtes effectuées pour la revue de la littérature

A. Liste des requêtes effectuées (Basket of 8 et revues spécialisées en Administration publique (en date du 2017-03-02))

No	Requête	Base de données	Restrictions
1	AB (big data OR Huge Data OR Massive DATA OR Urban data OR City Data OR Cities Data OR "internet of things" OR IoT) AND SO MIS Quarterly		MIS Quarterly
2	(Smart Cit* OR Digital Cit* OR Ubiquitous Cit* OR Connected Cit* OR Intelligent Cit* OR Virtual Cit* OR Real-time Cit*) AND MIS Quarterly	Business Source Complete (EBSCO)	
3	AB (big data OR Huge Data OR Massive DATA OR Urban data OR City Data OR Cities Data OR "internet of things" OR IoT) AND SO Journal of management information systems		Journal of management information systems
4	AB (Smart Cit* OR Digital Cit* OR Ubiquitous Cit* OR Connected Cit* OR Intelligent Cit* OR Virtual Cit* OR Real-time Cit*) AND SO Journal of management information systems		
5	TITLE-ABSTR-KEY(big data OR Huge Data OR Massive DATA OR Urban data OR City Data OR Cities Data OR "internet of things" OR IoT) and JOURNAL-NAME(Journal of Strategic Information Systems)	ScienceDirect	Journal of Strategic Information Systems
6	TITLE-ABSTR-KEY (Smart Cit* OR Digital Cit* OR Ubiquitous Cit* OR Connected Cit* OR Intelligent Cit* OR Virtual Cit* OR Real-time Cit*) AND SO Journal of Strategic Information Systems		
7	ANYWHERE ("big data" OR "Huge Data" OR "Massive DATA" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT)	INFORMS PubsOnLine	Information Systems Research
8	"Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City"		
9	AB (big data OR Huge Data OR Massive DATA OR Urban data OR City Data OR Cities Data OR "internet of things" OR IoT) AND JN "Journal of the Association for Information Systems"	Computers & Applied Sciences Complete	Journal of the Association for Information Systems
10	JN "Journal of the Association for Information Systems" AND AB ("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Virtual Cities" OR "Real-time Cities")	Computers & Applied Sciences Complete	
11	("big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT)	Palgrave Macmillan Journals et	Journal of Information Technology
12	"Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Virtual Cities" OR "Real-time Cities"	ABI/INFORM Global	

Formes et antécédents de l'utilisation des mégadonnées par les villes

No	Requête	Base de données	Restrictions
13	"big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT in Abstract AND Information Systems Journal in Publication Titles	Wiley-Blackwell Full Collection	Information Systems Journal
14	("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Virtual Cities" OR "Real-time Cities") in Abstract AND Information Systems Journal in Publication Titles		
15	("big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT) AND PUBID(30520)	European Business Database + Palgrave Macmillan Journals	European Journal of Information Systems
16	("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Virtual Cities" OR "Real-time Cities") AND PUBID(30520)		
17	("big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT)	Directory of Open Access Journals	Electronic Journal of e-Government
18	("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Virtual Cities" OR "Real-time Cities")		
19	("big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT) AND PUBID(406306)	ProQuest Central	Electronic Journal of e-Government
20	("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Virtual Cities" OR "Real-time Cities") AND PUBID(406306)		
21	TITLE-ABSTR-KEY(("big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT)) and JOURNAL-NAME(Government Information Quarterly)	ScienceDirect Journals	Government Information Quarterly
22	TITLE-ABSTR-KEY("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "real-time Cities") and JOURNAL-NAME("Government Information Quarterly").		
23	"Public Administration Review" in Publication Titles AND "big data" OR "Huge Data" OR "Massive Data" OR "Urban data" OR "City Data" OR "Cities Data" OR "internet of things" OR IoT in Article Titles	Wiley-Blackwell Full Collection	Public Administration Review
24	"Public Administration Review" in Publication Titles AND "Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Virtual City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "real-time Cities" in All Fields		

B. Liste des requêtes effectuées sur les principales bases de données identifiées (en date du 2017-03-17)

No.	Requête	Base de données	Restrictions
25	AB ("Big Data" OR "Huge Data" OR "Massive Data" OR "Large Data" OR "Urban data" OR "City Data" OR "Municipality Data") AND AB ("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Real-time Cities" OR "e-Government ")	Business Source Complete (EBSCO)	Revue académiques (relues par un comité de lecture)
26	all("Big Data" OR "Huge Data" OR "Massive Data" OR "Large Data" OR "Urban data" OR "City Data" OR "Municipality Data") AND all("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Real-time Cities" OR "e-Government")	ProQuest	Revu par les pairs Publications académiques
27	TITLE-ABSTR-KEY("Big Data" OR "Huge Data" OR "Massive Data" OR "Large Data" OR "Urban data" OR "City Data" OR "Municipality Data") and TITLE-ABSTR-KEY("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Real-time Cities" OR "e-Government")	ScienceDirect	/
28	("Big Data" OR "Huge Data" OR "Massive Data" OR "Large Data" OR "Urban data" OR "City Data" OR "Municipality Data") in Abstract AND ("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Real-time Cities" OR "e-Government") in Keywords	Wiley Online Library (John Wiley & Sons)	/
29	"International Conference on Information Systems " AND "proceeding" "Big data" "Americas Conference on Information Systems" AND "proceeding" "Big data" "European Conference on Information Systems " AND "proceeding" "Big data" "Americas Conference on Information Systems" AND "proceeding" "Big data"	IEEEExplore	/
30	Abstract: ("Big Data" OR "Huge Data" OR "Massive Data" OR "Large Data" OR "Urban data" OR "City Data" OR "Municipality Data") AND Abstract:("Smart City" OR "Digital City" OR "Ubiquitous City" OR "Connected City" OR "Intelligent City" OR "Real-time City" OR "Smart Cities" OR "Digital Cities" OR "Ubiquitous Cities" OR "Connected Cities" OR "Intelligent Cities" OR "Real-time Cities" OR "e-Government")	Emerald	/

Annexe 2 : Liste des articles retenus pour la revue systématique de la littérature

- Abbasi, Ahmed, Suprateek Sarker et Roger H. L. Chiang (2016). « Big Data Research in Information Systems: Toward an Inclusive Research Agenda » [Article], *Journal of the Association for Information Systems*, vol. 17, no 2, p. i-xxxii.
- Abella, Alberto, Marta Ortiz-de-Urbina-C. et Carmen De-Pablos-H. (2017). « A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems » [Article], *Cities*, vol. 64, p. 47-53.
- Baesens, Bart, Ravi Bapna, James R. Marsden, Jan Vanthienen et J. Leon Zhao (2016). « Transformational issues of Big Data and analytics in networked business » [Article], *MIS Quarterly*, vol. 40, no 4, p. 807-818.
- Belanche-Gracia, Daniel, Luis V. Casalo-Ariño et Alfredo Pérez-Rueda (2015). « Determinants of multi-service smartcard success for smart cities development: A study based on citizens' privacy and security perceptions », *Government Information Quarterly*, vol. 32, no 2, p. 154-163.
- Bergamaschi, Sonia, Emanuele Carlini, Michelangelo Ceci, Barbara Furletti, Fosca Giannotti, Donato Malerba, et al. (2016). « Big Data Research in Italy: A Perspective », *Engineering*, vol. 2, no 2, p. 163-170.
- Bertot, John Carlo, Ursula Gorham, Paul T. Jaeger, Lindsay C. Sarin et Heeyoon Choi (2014). « Big data, open government and e-government: Issues, policies and recommendations » [Article], *Information Polity: The International Journal of Government & Democracy in the Information Age*, vol. 19, no 1/2, p. 5-16.
- Brynjolfsson, Erik, Tomer Geva et Shachar Reichman (2016). « Crowd-squared: amplifying the predictive power of search trend data » [Article], *MIS Quarterly*, vol. 40, no 4, p. 941-A936.
- Chauhan, Sumedha, Neetima Agarwal et Arpan Kumar Kar (2016). « Addressing big data challenges in smart cities: a systematic literature review » [Article], *Info*, vol. 18, no 4, p. 73-90.
- Chen, D., David Preston et Morgan Swink (2015). « How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management » [Article], *Journal of Management Information Systems*, vol. 32, no 4, p. 4-39.
- Clarke, Roger (2016). « Big data, big risks », *Information Systems Journal*, vol. 26, no 1, p. 77-90.
- Constantiou, Ioanna D. et Jannis Kallinikos (2015). « New games, new rules: big data and the changing context of strategy », *Journal of Information Technology*, vol. 30, no 1, p. 44-57.
- Dobre, C. et F. Xhafa (2014). « Intelligent services for Big Data science » [Article], *Future Generation Computer Systems*, vol. 37, p. 267-281.
- Fitzgerald, Michael (2016). « Data-Driven City Management: A Close Look at Amsterdam's Smart City Initiative », *MIT Sloan Management Review*, vol. 57, no 4, p. n/a.
- Foulonneau, Muriel, Slim Turki, Gérardine Vidou et Sébastien Martin (2014). « Open data in Service design », *Electronic Journal of E-Government*, vol. 12, no 2, p. 99-107.
- Gil-Garcia, J. Ramon, Jing Zhang et Gabriel Puron-Cid (2016). « Conceptualizing smartness in government: An integrative and multi-dimensional view », *Government Information Quarterly*, vol. 33, no 3, p. 524-534.

- Hashem, Ibrahim Abaker Targio, Victor Chang, Nor Badrul Anuar, Kayode Adewole, Ibrar Yaqoob, Abdullah Gani, *et al.* (2016). « The role of big data in smart city » [Article], *International Journal of Information Management*, vol. 36, no 5, p. 748-758.
- Höchtel, Johann, Peter Parycek et Ralph Schöllhammer (2016). « Big data in the policy cycle: Policy decision making in the digital era », *Journal of Organizational Computing and Electronic Commerce*, vol. 26, no 1-2, p. 147-169.
- Janssen, Marijn et George Kuk (2016a). « Big and Open Linked Data (BOLD) in research, policy, and practice », *Journal of Organizational Computing and Electronic Commerce*, vol. 26, no 1-2, p. 3-13.
- Janssen, Marijn et George Kuk (2016b). « The challenges and limits of big data algorithms in technocratic governance », *Government Information Quarterly*, vol. 33, no 3, p. 371-377.
- Janssen, Marijn et Jeroen van den Hoven (2015). « Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? », *Government Information Quarterly*, vol. 32, no 4, p. 363-368.
- Janssen, Marijn, Haiko van der Voort et Agung Wahyudi (2017). « Factors influencing big data decision-making quality », *Journal of Business Research*, vol. 70, p. 338-345.
- Jara, Antonio J., Dominique Genoud et Yann Bocchi (2015). « Big data for smart cities with KNIME a real experience in the SmartSantander testbed », *Software: Practice and Experience*, vol. 45, no 8, p. 1145-1160.
- Kallinikos, Jannis et Ioanna D. Constantiou (2015). « Big data revisited: a rejoinder », *Journal of Information Technology*, vol. 30, no 1, p. 70-74.
- Kitchin, Rob (2014). « The real-time city? Big data and smart urbanism », *GeoJournal*, vol. 79, no 1, p. 1-14.
- Lavertu, Stéphane (2016). « We All Need Help: “Big Data” and the Mismeasure of Public Administration », *Public Administration Review*, vol. 76, no 6, p. 864-872.
- Lee, Jungwoo et Hyejung Lee (2014). « Developing and validating a citizen-centric typology for smart city services », *Government Information Quarterly*, vol. 31, Supplement 1, p. S93-S105.
- Li, Xixi, J. J. Po-An Hsieh et Arun Rai (2013). « Motivational Differences Across Post-Acceptance Information System Usage Behaviors: An Investigation in the Business Intelligence Systems Context », *Information Systems Research*, vol. 24, no 3, p. 659-682.
- Loebbecke, Claudia et Arnold Picot (2015). « Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda », *The Journal of Strategic Information Systems*, vol. 24, no 3, p. 149-157.
- Luftman, Jerry, Barry Derksen, Rajeev Dwivedi, Martin Santana, Hossein S Zadeh et Eduardo Rigoni (2015). « Influential IT management trends: an international study » [journal article], *Journal of Information Technology*, vol. 30, no 3, p. 293-305.
- Martens, David, Foster Provost, Jessica Clark et Enric Junqué de Fortuny (2016). « Mining massive fine-grained behavior data to improve predictive analytics » [Article], *MIS Quarterly*, vol. 40, no 4, p. 869-888.

- Mehmood, Rashid, Royston Meriton, Gary Graham, Patrick Hennelly et Mukesh Kumar (2017). « Exploring the influence of big data on city transport operations: a Markovian approach » [Article], *International Journal of Operations & Production Management*, vol. 37, no 1, p. 75-104.
- Menon, Syam et Sumit Sarkar (2016). « Privacy and big data: scalable approaches to sanitize large transactional databases for sharing » [Article], *MIS Quarterly*, vol. 40, no 4, p. 963-982.
- Mergel, Ines, R. Karl Rethemeyer et Kimberley Isett (2016). « Big Data in Public Affairs », *Public Administration Review*, vol. 76, no 6, p. 928-937.
- Müller, Oliver, Iris Junglas, Jan Vom Brocke et Stefan Debortoli (2016). « Utilizing big data analytics for information systems research: challenges, promises and guidelines », *European Journal of Information Systems*, vol. 25, no 4, p. 289-302.
- Newell, Sue (2015). « Managing knowledge and managing knowledge work: what we know and what the future holds » [journal article], *Journal of Information Technology*, vol. 30, no 1, p. 1-17.
- Pan, Yunhe, Yun Tian, Xiaolong Liu, Dedao Gu et Gang Hua (2016). « Urban Big Data and the Development of City Intelligence », *Engineering*, vol. 2, no 2, p. 171-178.
- Popescu, Daniela et Laura Diana Radu (2016). « Data Security in Smart Cities: Challenges and Solutions », *Informatica Economica*, vol. 20, no 1, p. 29-38.
- Saboo, Alok R., V. Kumar et Insu Park (2016). « Using big data to model time-varying effects for marketing resource (re)allocation » [Article], *MIS Quarterly*, vol. 40, no 4, p. 911-940.
- Saxena, Stuti et Sujeet Kumar Sharma (2016). « Integrating Big Data in “e-Oman”: opportunities and challenges », *Info : the Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media*, vol. 18, no 5, p. 79-97.
- Spiekermann, Sarah et Jana Korunovska (2016). « Towards a value theory for personal data » [journal article], *Journal of Information Technology*.
- Šuh, Jelena, Vladimir Vujin, Dušan Barać, Zorica Bogdanović et Božidar Radenković (2015). « Designing cloud infrastructure for big data in e-government » [Article], *RUO: Revija za Univerzalno Odlicnost*, vol. 4, no 1, p. A26-A38.
- Tachizawa, Elcio M., María J. Alvarez-Gil et María J. Montes-Sancho (2015). « How “smart cities” will change supply chain management », *Supply Chain Management: An International Journal*, vol. 20, no 3, p. 237-248.
- van Loenen, Bastiaan, Stefan Kulk et Hendrik Ploeger (2016). « Data protection legislation: A very hungry caterpillar: The case of mapping data in the European Union », *Government Information Quarterly*, vol. 33, no 2, p. 338-345.
- van Zoonen, Liesbet (2016). « Privacy concerns in smart cities », *Government Information Quarterly*, vol. 33, no 3, p. 472-480.
- Zuboff, Shoshana (2015). « Big other: surveillance capitalism and the prospects of an information civilization », *Journal of Information Technology*, vol. 30, no 1, p. 75-89.

Annexe 3 : Certificat d'approbation éthique

HEC MONTRÉAL

Comité d'éthique de la recherche

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de notre politique en cette matière.

Projet # : 2018-2736

Titre du projet de recherche : Conceptualisation de l'utilisation des mégadonnées (Big data) par les villes

Chercheur principal :
Hamza Ali,
Étudiant M. Sc. - HEC Montréal

Directeur/codirecteurs :
Ryad Titah
Professeur - HEC Montréal

Date d'approbation du projet : 22 juin 2017

Date d'entrée en vigueur du certificat : 22 juin 2017

Date d'échéance du certificat : 01 juin 2018



Maurice Lemelin
Président du CER de HEC Montréal

Annexe 4 : Mise en application des principales exigences éthiques

Exigences (EPTC2)	Mise en pratique
<p>Consentement libre et éclairé</p> <ul style="list-style-type: none"> ▪ Le respect d'autrui signifie le respect de l'exercice du consentement individuel ▪ Le consentement doit être éclairé et donné volontairement avant la collecte ou l'accès aux données. ▪ Le consentement doit être un processus continu. 	<ul style="list-style-type: none"> ▪ Liberté des répondants de décider de ne pas participer à l'étude. ▪ Liberté des répondants de se retirer de l'étude ou de refuser de répondre à des questions. ▪ Contact des répondants potentiels grâce à l'envoi de courriels individuels. ▪ Consentement implicite des répondants en procédant au remplissage du questionnaire
<p>Respect de la vie privée et des renseignements personnels</p> <ul style="list-style-type: none"> ▪ Le respect de la confidentialité des données recueillies. ▪ La prise de mesures de protection et de sécurité pour les informations collectées. ▪ Le consentement pour l'utilisation secondaire des données à des fins de recherche. 	<ul style="list-style-type: none"> ▪ Accès aux données brutes recueillies restreint au chercheur et au directeur de recherche. ▪ Conservation sécurisée des données recueillies (authentification nécessaire). ▪ Consentement implicite des répondants pour l'utilisation des données à d'autres fins. ▪ Questionnaire anonyme, aucun renseignement personnel n'a été recueilli.
<p>Réduction des risques pour le répondant et pour les institutions</p> <ul style="list-style-type: none"> ▪ La suppression ou la réduction des inconvénients que pourraient subir les répondants. ▪ La réduction des conflits d'intérêts institutionnels. 	<ul style="list-style-type: none"> ▪ Questionnaire anonyme, aucune donnée permettant d'identifier l'organisation pour laquelle travaille le répondant n'a été recueillie. ▪ Absence de conflit d'intérêt financier (emploi ou compensation).

Annexe 5 : Littérature sur les antécédents organisationnels de l'utilisation des TI

Source	Variabes et méthodologie	Conclusions
Armstrong et Sambamurthy (1999)	<p>Antécédents: Quality of senior leadership, Sophistication of IT infrastructures, Organizational size</p> <p>Variable dépendante: IT assimilation</p> <p>Méthodologie: questionnaire (N=265)</p>	<p>Les connaissances d'affaires et en TI des CIO et le support de la haute direction influent de manière significative l'assimilation technologique de l'organisation.</p> <p>La sophistication de l'infrastructure technologique a une influence notable sur l'assimilation informatique.</p> <p>La taille de l'organisation et les connaissances TI des membres de la haute direction n'ont pas d'influence significative.</p>
Boynton, Zmud et Jacobs (1994)	<p>Antécédents: IT management Climate, Managerial IT Knowledge, IT management Process effectiveness.</p> <p>Variable dépendante: utilisation des TI ($R^2 = ?$)</p> <p>Méthodologie : questionnaire (N= 132)</p>	<p>Les connaissances TI des gestionnaires et le climat de gestion des TI influencent positivement l'utilisation des TI.</p> <p>Le climat de gestion des TI a également une influence positive sur l'efficacité du processus de gestion des TI.</p> <p>L'efficacité du processus de gestion des TI n'a pas d'effet significatif.</p>
D. Chen, Preston et Swink (2015)	<p>Antécédents: Expected Benefits, Technology Compatibility, Organizational Readiness, Competitive Pressure</p> <p>Variable dépendante: Big Data Analytics use ($R^2 = 40\%$)</p> <p>Méthodologie: questionnaire (N = 161)</p>	<p>Les facteurs technologiques influent directement sur l'utilisation organisationnelle de la BDA.</p> <p>Les facteurs organisationnels (prédisposition organisationnelle) et environnementaux (pression concurrentielle) influent indirectement l'utilisation par l'intermédiaire du soutien de la haute direction.</p>
Guillamón <i>et al.</i> (2016)	<p>Antécédents: E-participation Index, Multiculturalism, income, Population, Internet, education, Financial Autonomy, Leverage, Gender of Mayor, Political ideology, Type of municipality, Country</p> <p>Variable dépendante: Utilisation des médias sociaux [Facebook index] ($R^2 = ?$)</p> <p>Méthodologie : questionnaire (N = 217)</p>	<p>Le niveau de participation électronique, la taille de la population, le niveau de revenu des citoyens et le niveau d'endettement ont une incidence sur l'utilisation de Facebook par les gouvernements locaux.</p> <p>Un haut niveau d'utilisation de Facebook suggère que les gouvernements locaux ont tendance à permettre aux citoyens de surveiller le gouvernement afin de rendre plus transparentes les informations, les données et les processus municipaux.</p>

Annexe 5 : Littérature sur les antécédents organisationnels de l'utilisation des TI

Source	Variables et méthodologie	Conclusions
Iacovou, Benbasat et Dexter (1995)	<p>Antécédents: Niveau de prédisposition, Pressions externes, Bénéfices perçus</p> <p>Variable dépendante : Adoption et intégration de l'EDI</p> <p>Méthodologie : entrevue (N=7)</p>	<p>Les pressions externes (partenaires d'affaires) constituent le facteur le plus important dans l'<u>adoption</u> de l'EDI.</p> <p>Une grande prédisposition organisationnelle et la volonté de concrétiser les bénéfices escomptés sont des facteurs importants pour l'<u>intégration</u> et la concrétisation des impacts de l'adoption de l'EDI.</p>
Lin (2008)	<p>Antécédents : Perceived relative advantage, Complexity, Compatibility, Managerial commitment, Systems orientation, Knowledge acquisition, Knowledge dissemination</p> <p>Variable dépendante: e-business implementation success ($R^2 = ?$)</p> <p>Méthodologie: Questionnaire (N=163)</p>	<p>Deux caractéristiques d'innovation (avantage relatif et compatibilité) et quatre capacités d'apprentissage organisationnel (engagement managérial, orientation des systèmes, acquisition de connaissances et diffusion de la connaissance) ont un effet important sur la réussite de la mise en œuvre du commerce électronique.</p> <p>L'effet de la complexité de l'innovation n'est pas significatif.</p>
Massetti et Zmud (1996)	<p>Antécédents : caractéristiques de l'EDI (volume, breadth, diversity, and depth)</p> <p>Variable dépendante : EDI extend usage</p> <p>Méthodologie : étude de cas (N=7)</p>	<p>Approche proposée pour mesurer la valeur créée par l'utilisation de l'EDI. Il s'agit d'une approche à quatre dimensions : le volume, la diversité, l'ampleur et la profondeur des initiatives EDI d'une organisation. Chacune de ces facettes est définie et décrite par son application dans les contextes de sept sites de cas.</p>
Mishra, Konana et Barua (2007)	<p>Antécédents: Procurement-process digitalization, Diversity of organizational procurement knowledge, Suppliers' sales-process digitalization, Organizational perceptions of technological uncertainty, Organizational perceptions of volume uncertainty</p> <p>Variable dépendante: impact de l'utilisation d'Internet ($R^2 = 0.11$, $R^2 = 0.38$)</p> <p>Méthodologie : questionnaire (N=412)</p>	<p>La numérisation des processus d'approvisionnement et la diversité des connaissances organisationnelles sur l'approvisionnement influent sur l'utilisation de l'Internet.</p> <p>Les déterminants de l'utilisation organisationnelle de l'Internet dans le processus d'approvisionnement n'ont pas la même influence (fonction du processus).</p> <p>L'infrastructure technologique et la numérisation des processus de vente des fournisseurs ont un impact positif sur l'utilisation de l'Internet.</p> <p>Les perceptions sur l'incertitude technologique et sur l'incertitude liée aux volumes sont mitigées.</p>

Annexe 5 : Littérature sur les antécédents organisationnels de l'utilisation des TI

Source	Variabiles et méthodologie	Conclusions
Thong (1999)	<p>Antécédents : <u>Caractéristiques du dirigeant</u> (innovation, niveau de connaissance TI) de l'<u>innovation</u> (avantage relative, compatibilité et complexité), de l'<u>organisation</u> (taille, niveau de connaissance TI des employés, intensité de l'information) et de l'<u>environnement externe</u> (compétition).</p> <p>Variable dépendante : Adoption des SI ($R^2 = 43\%$)</p> <p>Méthodologie : questionnaire (N=166)</p>	<p>Alors que les caractéristiques du dirigeant (innovation et niveau de connaissance TI) et celles de l'innovation (avantage relatif, compatibilité et complexité) sont des déterminants importants de l'adoption, elles n'ont pas d'influence sur l'étendue de l'adoption.</p> <p>L'étendue de l'adoption est principalement déterminée par les caractéristiques organisationnelles (taille de l'entreprise et niveau de connaissance des employés).</p> <p>La concurrence n'a aucun effet direct sur l'adoption des SI.</p>
Zhu, Kraemer et Dedrick (2004)	<p>Antécédents: Technology readiness, firm size, global scope, financial resources, competition intensity, and regulatory environment</p> <p>Variable dépendante: Création de la valeur ($R^2 = 58\%$)</p> <p>Méthodologie : questionnaire (N=612)</p>	<p>La prédisposition technologique est le facteur le plus significatif, suivi de la disponibilité de ressources financières, la portée mondiale et le support réglementaire.</p> <p>La taille a un effet négatif.</p> <p>L'effet des pressions compétitives n'est pas significatif.</p>
Zhu et Kraemer (2005)	<p>Antécédents: Technology Competence, Organisation Size, International scope, Financial commitment, Competitive pressure, Regulatory support.</p> <p>Méthodologie: E-Business use ($R^2 = 20\%$)</p> <p>Méthodologie: questionnaire (N=624)</p>	<p>La compétence technologique, l'engagement financier, la pression concurrentielle et le soutien réglementaire ont une influence importante.</p> <p>La compétence technologique semble être le facteur le plus important (suivi du soutien réglementaire).</p> <p>La taille a un effet négatif tandis que la portée internationale est positive, mais statistiquement insignifiante.</p>
Zhu <i>et al.</i> (2006)	<p>Antécédents : caractéristiques de l'innovation (Avantage relatif, compatibilité, coûts, préoccupations de sécurité) et facteurs contextuels (compétences technologiques, taille, pressions compétitives et prédisposition des partenaires).</p> <p>Variable dépendante : Impact du e-business ($R^2 = 58\%$)</p> <p>Méthodologie : questionnaire (N = 1415)</p>	<p>Parmi les caractéristiques d'innovation, la compatibilité est le facteur (positif) le plus significatif, et les préoccupations de sécurité l'emportent sur le coût (négatif).</p> <p>Parmi les variables contextuelles, la compétence technologique, la prédisposition des partenaires et les pressions concurrentielles influencent considérablement l'utilisation du commerce électronique (positif). L'inertie structurelle des grandes entreprises tend à ralentir sa pénétration (négatif).</p>

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
Items de mesure de l'avantage relatif (bénéfices escomptés)				
D. Chen, Preston et Swink (2015)	Using Big Data Analytics enables/will enable our organization to Improve the quality of work	Likert-5 1 = strongly disagree 5 = strongly agree	/	/
	Using BDA enables/will enable our organization to Make work more efficient	Component factor analysis (CFA) in Partial least squares (PLS)	/	/
	Using BDA enables/will enable our organization to Lower costs	Composite reliability = 0.913	Using Big Data enables/will enable our City to help lower costs	Likert-7
	Using BDA enables/will enable our organization to Improve customer service/patient care		Using Big Data enables/will enable our City to improve citizen service	Likert-7
	Using BDA enables/will enable our organization to Grow sales to new customers or new markets		/	/
	Using BDA enables/will enable our organization to Identify new product/service opportunities		Using Big Data enables/will enable our city to offer new product/service to the citizens	Likert-7
Zhu <i>et al.</i> (2006)	The degree to which your company expected e-business to help <ul style="list-style-type: none"> • reduce costs • increase sales 	Likert-5 Partial least squares (PLS)	/	/
Iacovou, Benbasat et Dexter (1995)	Direct Benefits: Reduced transaction costs		/	/
	Direct Benefits: Improved cash flow		/	/
	Direct Benefits: Reduced inventory levels		/	/
	Direct Benefits: Higher information quality		Using Big Data enables/will enable our city to higher the information quality	Likert-7
	Indirect Benefits : Increased operational efficiency	/	Using Big Data enables/will enable our city to increase operational efficiency	Likert-7
	Indirect Benefits : Better customer service		/	/
	Indirect Benefits : Improved trading partner relationships		/	/
	Indirect Benefits : Increased ability to compete		Using Big Data enables/will enable our city to increase its ability to compete	Likert-7

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
Lin (2008)	Following implementation of e-business, my firm will . . . <ul style="list-style-type: none"> • provide better products or services. • enhance business efficiency. • enhance staff productivity. • improve coordination with trading partners. 	Likert-5 1 = strongly disagree 5 = strongly agree CFA in LISREL	/	/
Moore et Benbasat (1991)	Using a Personal Work Station (PWS) <ul style="list-style-type: none"> • enables me to accomplish tasks more quickly* • improves the quality of work I do.* • makes it easier to do my job. * • improves my job performance. • enhances my effectiveness on the job.* • gives me greater control over my work. * • increases my productivity Overall, I find using a PWS to be advantageous in my job.	Likert-7 1 = Extremely disagree 7 = Extremely agree Card sorting CFA Cronbach's $\alpha = 0.9$	/	/
Zhu, Kraemer et Dedrick (2004)	Impact on internal efficiency: Internat processes more efficient		/	/
	Impact on internal efficiency: Staff productivity increased	Likert-5	/	/
	Impact on coordination: Transaction costs with business partners decreased		/	/
	Impact on coordination: Coordination with business partners or suppliers improved		/	/
Premkumar et Roberts (1999)	The technology will allow us to better communicate with our business partners.	Likert-5 1 = strongly disagree 5 = Strongly agree	/	/
	The technology will allow us to cut costs in our operations.	Component factor analysis (CFA)	/	/
	Implementing the technology will increase the profitability of our business.	Cronbach's $\alpha = 0.72$	/	/
	Adoption of the technology will provide timely information for decision making.		Using Big Data enables/will enable our city to get timely information for decision making	Likert-7

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
Items de mesure de la compatibilité				
D. Chen, Preston et Swink (2015)	Using Big Data Analytics is consistent with our business practices	Likert-5 1 = strongly disagree 5 = strongly agree	Using Big Data is consistent with our City practices	Likert-7
	Using Big Data Analytics fits our organizational culture		Using Big Data fits our organizational culture	Likert-7
	Overall, it is/will be easy to incorporate Big Data Analytics into our SCM practices	CFA in PLS Composite reliability = 0.857	Overall, it is/will be easy to incorporate Big Data into our practices	Likert-7
Zhu <i>et al.</i> (2006)	Selling over the Internet is compatible with your company's current selling process	Likert-5	/	/
	Buying over the Internet is compatible with your company's current procurement process		/	/
	Conducting transactions over the Internet is compatible with existing distribution channels		/	/
	Doing e-business is compatible with your company's corporate culture and value system		Using Big Data is not compatible with our value system	Likert-7
Lin (2008)	Implementation of e-business . . . <ul style="list-style-type: none"> • is acceptable to corporate culture and value system. • does not contradict the current internal IS applications. • is supported by the existing IS infrastructure. • is supported by the organizational IT human resources. 	Likert-5 1 = strongly disagree 5 = strongly agree CFA in LISREL	Using Big Data... <ul style="list-style-type: none"> • is supported by the existing IS infrastructure. • is supported by the organizational IT human resources. 	Likert-7
Moore et Benbasat (1991)	Using a PWS is compatible with all aspects of my work.		/	/
	Using a PWS is completely compatible with my current situation.		/	/
	I think that using a PWS fits well with the way I like to work	/	/	/
	Using a PWS fits into my work style.		/	/
Premkumar et Roberts (1999)	Implementing the changes caused by adoption of these new technologies is not compatible with our firm's values and beliefs	Likert-5 Analyse factorielle (CFA) Cronbach's $\alpha = /$	The changes that the Big Data adoption caused/will cause are not compatible with our City's values and beliefs	Likert-7
Kendall <i>et al.</i> (2001)	Changing of company's policy and organizational structure are necessary if my company is to do business over the internet	Likert-6	Changing of our city policy and organizational structure are/were necessary to use Big Data	Likert-7
Items de mesure de la complexité				

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
Lin (2008)	Implementation of e-business will lead to difficulties in <ul style="list-style-type: none"> controlling work quality. integrating the overall business process. protecting the security of data and transactions over the Internet. building long-term relationships between my firm and trading partners. 	Likert-5 1 = strongly disagree 5 = strongly agree CFA in LISREL	Implementation of Big Data lead/ will lead to difficulties in <ul style="list-style-type: none"> controlling work quality. integrating the overall process. protecting the security and privacy of citizen data 	Likert-7
Premkumar et Roberts (1999)	The skills required to use these technologies are too complex for our employees. Integrating these technologies in our current work practices will be very difficult.	Likert-5 Analyse factorielle (CFA) Cronbach's $\alpha = 0.56$	The skills required to use the Big Data technologies are too complex for our employees. Integrating these technologies in our current work practices will be very difficult.	Likert-7 Likert-7
Items de mesure des préoccupations de confidentialité				
Zhu <i>et al.</i> (2006)	The degree to which your company is concerned about the security of data and transactions over the Internet The degree to which your customers are concerned about the security of data and privacy over the Internet	Likert-5	Our city is concerned about the security of its citizen data Our citizens are concerned about the security and privacy of their data	Likert-7 Likert-7
Kendall <i>et al.</i> (2001)	My company is concerned about the security of payment over the internet My company is concerned that information involved in a transaction over the internet is not private	Likert-6	/	/
Items de mesure des compétences technologiques				
D. Chen, Preston et Swink (2015)	To what extent are the following factors preventing your business unit from fully exploiting BDA? <ul style="list-style-type: none"> Lacking capital/financial resources Lacking needed IT infrastructure Lacking analytics capability Lacking skilled resources 	Likert-5 (R) CFA in PLS Composite reliability = 0.847	To what extent are the following factors preventing your city from fully exploiting Big Data? <ul style="list-style-type: none"> Lacking needed IT infrastructure Lacking analytics capability Lacking skilled resources 	Likert-7
Zhu et Kraemer (2005)	Number of PCs per employee IT professionals, as percent of total employees Number of items the establishment has in the following list: Use of e-mail, website accessible by public, use of intranet, use of extranet, use of EDI, use of EFT, use of call center	Continuous variable Continuous variable Continuous variable	/	/
Zhu <i>et al.</i> (2006)	IT infrastructure: the strength of existing IT infrastructure, as measured by related technologies that your company has in place, including EDI, intranet, extranet, LAN, WAN	Likert-5	/	/

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
	Internet skills: The extent to which the majority of your employees are capable of using the following applications: Web browser, intranet, online order processing		/	/
Zhu <i>et al.</i> (2006)	Skill development – Has your company done the following to help employees develop e-business skills: (a) in-house training? (b) participating in IT training such as courses and seminars by third parties? (c) legitimizing certain work time for IT learning/training? (d) establishing self-learning or e-learning programs? (e) recruiting staff with special IT skills?	Multiple choices	Has your city done the following to help employees develop Big Data skills: (a) in-house training? (b) participating in a training such as courses and seminars by third parties? (c) legitimizing certain work time for learning/training? (d) establishing self-learning or e-learning programs? (e) recruiting staff with special Big Data skills?	Likert-7
Thong (1999)	My employees were all computer-literate	Likert-5 1 = Strongly disagree 5 = Strongly agree	/	/
	There was at least one employee who was a computer expert		/	/
	I would rate my employees' understanding of computers as very good compared with other small companies in the same industry.		/	/
Kendall <i>et al.</i> (2001)	My company does not have the technical knowledge to do business over the internet	Likert-6	Our city does not have the technical knowledge to use Big Data	Likert-7
Items de mesure de l'engagement de la haute direction				
D. Chen, Preston et Swink (2015)	Does the Top Management Team (TMT, e.g., CIO, COO, CSCO) support the use of Big Data Analytics?	Likert-5 1 = not at all 3 = somewhat 5 = it is among the highest priorities	To what extent does the Top Management Team (TMT, e.g., Mayor, CTO, CIO, CDO, General Director) support the use of Big Data?	Likert-7
	To what extent does the TMT promote the use of Big Data Analytics in your organization?		To what extent does the TMT promote the use of Big Data in your organization?	Likert-7
	To what extent does the TMT create support for Big Data Analytics initiatives within your organization?	Partial least squares (PLS)	To what extent does the TMT create support for Big Data Analytics initiatives within your organization?	Likert-7
	To what extent has the TMT promoted Big Data Analytics as a strategic priority within your organization?		To what extent does the TMT promote Big Data use as a strategic priority within your organization?	Likert-7
Premkumar et Roberts (1999)	The owner or manager enthusiastically supports the adoption of these new technologies.	Likert-5 Analyse factorielle (CFA) Cronbach's $\alpha = 0.86$	To what extent does the TMT enthusiastically support the adoption of these new technologies.	Likert-7
	The owner or manger has allocated adequate resources to adoption of these new technologies.		To what extent does the TMT allocate adequate resources to adoption of these new technologies.	Likert-7
	Top management is aware of the benefits of these new technologies.		To what extent the TMT is aware of the benefits of these new technologies.	Likert-7

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
	Top management actively encourages employees to use the new technologies in their daily tasks.		To what extent does the TMT actively encourage employees to use the new technologies in their daily tasks.	Likert-7
Lin (2008)	My firm views employee training as an investment, not an expense.		To what extent does the TMT view employee training as an investment, not an expense.	Likert-7
[Managerial commitment]	Managers frequently involve their staff in important decision-making processes.	Likert-5 1 = strongly disagree 5 = strongly agree	/	/
	Employee learning capability is considered as a key factor in my firm.		To what extent does the TMT consider employee learning capability as a key factor.	Likert-7
	Innovation ideas that work are rewarded in my firm.	CFA in LISREL	/	/
Zhu et Kraemer (2005)	Financial commitment: IS operating budget, as percent of total revenue	Continuous variable	/	/
	Financial commitment : Web-based spending, as percent of total revenue		/	/
Zhu, Kraemer et Dedrick (2004)	Web-based spending, as percentage of total revenue	Continuous variable	/	/
Items de mesure des pressions compétitives				
D. Chen, Preston et Swink (2015)	How many others in your industry have been implementing Big Data Analytics?	Likert-5 0 = don't know	/	/
	To what extent have your competitors implemented Big Data Analytics?	1 = none have adopted 5 = almost all have adopted	To what extent have the cities near yours implemented Big Data technologies?	Likert-7
	To what extent have your suppliers implemented Big Data Analytics?		/	/
	To what extent have your customers implemented Big Data Analytics?	CFA in PLS CR = 0.847	/	/
Zhu et Kraemer (2005)	Degree affected by competitors in the local market	Likert-5	/	/
	Degree affected by competitors in the national market		/	/
Zhu <i>et al.</i> (2006)	Percentage of competitors in your industry that have conducted Internet-based selling	Likert-5	/	/
	Percentage of competitors in your industry that have conducted Internet-based procurement and coordination		/	/
	Percentage of competitors in your industry that have conducted Internet-based services		/	/

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
Zhu, Kraemer et Dedrick (2004)	Degree affected by competitors in the local markets	Likert-5	/	/
	Degree affected by competitors nationwide		/	/
	Degree affected by competitors worldwide		/	/
Thong (1999)	Ease for a customer to switch to a competitor	Likert-5	/	/
	level of rivalry among businesses in the same industry	1 = Strongly disagree 5 = Strongly agree	/	/
	effect of substitutable products and services.		/	/
(Iacovou, Benbasat et Dexter, 1995)	The estimated proportion of a firm's competitors and partners that are EDI-capable	/	/	/
	Perceptual measures about possible competitive disadvantage due to the lack of EDI-capability.		Perception about possible competitive disadvantage due to the lack of Big Data-capability.	Likert-7
Premkumar et Roberts (1999)	We believe we will lose our customers to our competitors if we do not adopt these new technologies	Likert-5	/	/
	We feel it is a strategic necessity to use these technologies to compete in the marketplace	Analyse factorielle (CFA) Cronbach's $\alpha = 0.81$	We feel it is a strategic necessity to use these technologies to compete in the marketplace	Likert-7
Items de mesure du dynamisme de l'écosystème				
(Zhu <i>et al.</i> , 2006)	The extent to which downstream customers have e-business systems ready to support Internet-based Selling	Likert-5	/	/
	The extent to which upstream partners have e-business systems ready to support Internet-based procurement		The extent to which upstream partners have technologies ready to support Big Data use	Likert-7
	The extent to which e-business systems owned by trading partners are interoperable with yours		/	/
Zhu, Kraemer et Dedrick (2004)	Government provided incentive	Likert-5	/	/
	Required for government purchase		/	/
(Zhu et Kraemer, 2005)	Business laws support electronic business		The municipal laws support Big Data use	Likert-7
	Legal protection for consumer purchase on the Internet		There is a legal protection for security and privacy of citizen data	Likert-7
Mishra, Konana et Barua (2007)	Suppliers' sales-process digitization	Likert-7	/	/
	Our suppliers have computer systems in place to quickly respond to our product inquiries.	1=Strongly disagree 7=Strongly agree	/	/
	Our suppliers can electronically process business documents (e.g., invoices, designs, POs).	Cronbach's $\alpha=0.81$	/	/

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
	Our suppliers have computerized their ordermanagement process.	Composite reliability=0.81	/	/
Premkumar et Roberts (1999) [External Pressure]	Our suppliers require the use of these technologies to do business with them.	Likert-5	/	/
	Our customers are demanding the use of these technologies for doing business with them.	Analyse factorielle (CFA) Cronbach's $\alpha = 0.75$	Our citizens are demanding products / services that need the use of these Big Data	Likert-7
Premkumar et Roberts (1999) [External Support]	There are businesses in the community which provide technical support for effective use of these technologies.		There are businesses in the community which provide technical support for effective use of Big Data	Likert-7
	Community agencies provide incentives for adoption of these technologies.		/	/
	There are agencies in the community who provide training on these new technologies.	Likert-5 Analyse factorielle (CFA) Cronbach's $\alpha = 0.78$	There are agencies in the community who provide training on Big Data technologies.	Likert-7
	Technology vendors actively market these new technologies by providing incentives for adoption.		Technology vendors actively market Big Data technologies by providing incentives for adoption or offering free training sessions.	Likert-7
	Technology vendors promote these new technologies by offering free training sessions.		/	/

Items de mesure de la variable dépendante (utilisation des mégadonnées)

Titah (2010)	In our municipality, we use e-Government systems to INFORM citizens (individuals and businesses) or the public in general about... Recreational programs; Tenders and requests for proposals; Business opportunities; Employment opportunities; Council / committee meetings; Emergency measures; Municipal services and contact details.	Likert-10 Not at all to Very much	In our municipality, we use Big Data technologies to INFORM citizens (individuals and businesses) or the public in general about... <ul style="list-style-type: none"> • Travel time • Weather conditions • Traffic and road conditions
	Reflective items: Overall, using e-Government systems to inform citizens is a common practice of our municipality; Overall, our municipality uses e-Government systems to inform citizens; When informing your citizens, what percentage of the time does your municipality use e-Government systems?	Likert-10 Disagree completely to Agree completely	
	In our municipality we use e-Government systems to... Integrate our different internal tasks; Share information/data with other municipalities; Share information/data with the provincial government; Integrate services with provincial government; Manage our partnerships (e.g., with other municipalities and/or organizations);	Likert-10 Not at all to Very much	In our municipality we use Big Data technologies to... Integrate our different internal tasks; Share information/data with other municipalities; Share information/data with the provincial government; Integrate services with provincial government; Manage our

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
	Share information/data with our partners.		partnerships (e.g., with other municipalities and/or organizations); Share information/data with our partners/citizens.	
	<p>Reflective items: Overall, using e-Gov. systems to coordinate/integrate our tasks/activities is a common practice; Overall, using e-Government systems to enhance our external linkages is a common practice; When coordinating/integrating your tasks/activities, what percentage of the time does your municipality use the e-Government systems?</p>	Likert-10 Disagree completely to Agree completely		
	<p>In our municipality we use e-Government systems in NOVEL AND INNOVATIVE ways such as... Generating alternative and original revenue sources (other than taxes and fines); Creating a paperless working environment; Allowing remote attendance to our meetings.</p>	Likert-10 Not at all to Very much	In our municipality we use e-Government systems in NOVEL AND INNOVATIVE ways such as... Generating alternative and original revenue sources (other than taxes and fines); Creating a paperless working environment;	
	<p>Reflective items: Overall, our municipality has been successful in finding new ways of using e-Government systems; Overall, we always try to use e-Government systems in novel ways; Overall, using e-Government systems has allowed us to accomplish tasks that were not feasible or identified before.</p>	Likert-10 Disagree completely to Agree completely		
D. Chen, Preston et Swink (2015)	<p>To what extent has your organization implemented Big Data Analytics in Sourcing analysis</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in Purchasing spend analytics</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in CRM/customer/patient analysis</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in Network design/optimization</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in Warehouse operations improvements</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in Process/equipment monitoring</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in Production run optimization</p> <hr/> <p>To what extent has your organization implemented Big Data Analytics in Logistics improvements</p>	Likert-5 1 = little or no usage 3 = moderate usage 5 = heavy usage Partial least squares (PLS)	<p>To what extent has your organization implemented Big Data in citizen relationship management analysis</p> <hr/> <p>To what extent has your organization implemented Big Data in processes design/optimization</p> <hr/> <p>To what extent has your organization implemented Big Data in operations improvements</p> <hr/> <p>To what extent has your organization implemented Big Data in Process/equipment monitoring</p> <hr/> <p>To what extent has your organization implemented Big Data in Logistics or procurement improvements</p>	

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
	To what extent has your organization implemented Big Data Analytics in Forecasting/demand management		To what extent has your organization implemented Big Data in Forecasting/planning	
	To what extent has your organization implemented Big Data Analytics in Inventory optimization			
Zhu et Kraemer (2005)	Number of items in the following list : Providing information online, making sales online, providing service online, joining electronic intermediaries for online sales, making purchases online, joining electronic intermediaries for online purchase.	Continuous variable		
	Percent of consumer sales conducted online			
	Percent of business-to-business sales conducted online			
	Percent of goods for resale ordered online			
	Percent of supplies and equipment for doing business ordered online			
Zhu <i>et al.</i> (2006)	Percentage of sales to businesses conducted online	Likert-5		
	Percentage of sales to consumers conducted online			
	Percentage of customer services conducted online			
	Percentage of procurement conducted online			
Boynton, Zmud et Jacobs (1994)	Cost reduction: IS developed to reduce the cost of business activities	Likert-5	Big Data is used to reduce the cost of city activities	
	Management support: IS developed to assist in monitoring, controlling, and designing business activities	(1) no use at all, (2) just starting	Big Data is used to assist in monitoring, controlling, and designing city activities	
	Strategic planning: IS developed to assist in formulating business strategies	(3) used to some extent, (4) used to a great extent, (5) "industry leader."	Big Data is used to assist in formulating city strategies	
	Competitive thrust: IS developed to establish a competitive advantage in the market	Coefficient alpha = 0.81	Big Data is used to establish a competitive advantage in the market	

Items de mesure des variables de contrôle

Zhu et Kraemer (2005)	Size: Number of employees at establishment	y/n (Dummy variable)		
Zhu <i>et al.</i> (2006)	Organization size: Number of employees	Log transformed	Organization size: Number of employees	
Zhu, Kraemer et Dedrick (2004)	IT spending, as percentage of total revenue	Continuous variable		

Annexe 6 : Traçabilité des items de mesure des variables du modèle de recherche

Source	Items originaux	Échelle Fiabilité Validité	Items retenus après adaptation	Échelle
Thong (1999)	Number of employees	Because the values of size were highly skewed, they were subjected to a logarithmic transformation to reduce the variance		
Armstrong et Sambamurthy (1999)	Organizational Size (Sales revenue, net income, and number of employees)	Continuous variable		
Guillamón <i>et al.</i> (2016)	Number of inhabitants	Logarithm		
Guillamón <i>et al.</i> (2016)	Type of municipality	Province:1 Municipality: 0		
Titah et Barki (2011)	Population size	less than < 10,000 between 10,001 and 50,000 between 50,001 and 100,000 between 100,001 and 500,000 between 500,001 and 1,000,000 more than 1,000,001	Population size	
Titah (2010)	Dans l'ensemble, jusqu'à quel point considérez-vous que votre municipalité dispose des ressources financières nécessaires pour utiliser les systèmes de gouvernement électronique ?	Likert-10 1 = Peu 10 = Beaucoup	Dans l'ensemble, jusqu'à quel point considérez-vous que votre municipalité dispose des ressources financières nécessaires pour l'utilisation des mégadonnées ?	Likert-5

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL			
No.	Scé	Items en langue anglaise	Items en langue française
		Section 1: Please indicate your perceptions about the <u>expected benefits</u> of using Big Data in your municipality	Section 1: Veuillez indiquer vos perceptions sur les <u>avantages escomptés</u> de l'utilisation du Big data par votre municipalité
		Using Big data would enable our municipality to	L'utilisation du Big data permettrait à notre municipalité
RA01	[1]	▪ improve citizens' quality of life	⇒ ▪ d'améliorer la qualité de vie des citoyens
RA02	[1]	▪ improve citizens' services	⇒ ▪ d'améliorer la qualité des services offerts aux citoyens
RA03	[1]	▪ offer new products/services to the citizens	⇒ ▪ d'offrir de nouveaux produits/services aux citoyens
RA04	[2]	▪ increase operational efficiency	⇒ ▪ d'améliorer l'efficacité de ses opérations
RA05	[3]	▪ increase information quality	⇒ ▪ d'améliorer la qualité de l'information
RA06	[2]	▪ get timely information for decision making	⇒ ▪ d'obtenir de l'information en temps opportun pour la prise de décision
RA07	[8]	▪ improve its predictive capabilities	⇐ ▪ d'améliorer ses capacités de prévision
RA08	[8]	Overall, using Big data is advantageous to our municipality	⇒ Dans l'ensemble, l'utilisation du Big data est avantageuse pour notre municipalité
RA09	[8]	Overall, Big data is particularly useful to developing successful urban policies	⇒ Dans l'ensemble, le Big data est particulièrement utile pour la réussite des politiques municipales
RA10	[8]	Overall, Big data will strengthen the competitive capabilities of our municipality	⇒ Dans l'ensemble, le Big data renforcera les capacités compétitives de notre municipalité
		Section 2: Please indicate your opinion about the <u>compatibility</u> of Big Data with your municipality characteristics	Section 2: Veuillez indiquer votre avis sur la <u>compatibilité</u> du Big data avec les diverses caractéristiques de votre municipalité
		Using Big data...	⇒ L'utilisation du Big data ...
CT01	[1]	▪ is consistent with the practices of our municipality	⇒ ▪ est conforme aux pratiques de notre municipalité
CT02	[1]	▪ fits our organizational culture	⇒ ▪ est adaptée à notre culture organisationnelle
CT03	[4]	▪ is compatible with our values	⇒ ▪ est compatible avec nos valeurs
CT04	[8]	▪ fits into our decision making system	⇒ ▪ est cohérente avec notre système de prise de décision
CT05	[8]	▪ is socially acceptable among our citizens	⇐ ▪ est socialement acceptable par les citoyens de notre municipalité
CT06	[8]	In our municipality, we have concerns about the danger that data will drive public questions (R)	⇒ Dans notre municipalité, on est préoccupés par le danger de se baser sur le Big data dans la conduite des politiques publiques (R)
CT07	[8]	In our municipality, we have concerns about the disruptive impacts of Big data on various jobs (R)	⇒ Dans notre municipalité, on est préoccupés par les effets du Big data sur la transformation des emplois (R)
CT08	[1]	Overall, it is easy to incorporate Big data into our practices	⇒ Dans l'ensemble, il est facile d'intégrer le Big data dans nos pratiques
CT09		Overall, Big data is fully compatible with our current work practices	Dans l'ensemble, le Big data est entièrement compatible avec nos pratiques de travail actuelles
CT10		Using Big data required/requires an overall change in the values, norms and culture within our municipality	L'utilisation du Big data a nécessité/nécessitera un changement dans les valeurs, les normes et la culture au sein de notre municipalité
		Section 3: Please indicate your opinion about the <u>complexity</u> related to the use of Big Data in your municipality	Section 3: Veuillez indiquer votre opinion sur la <u>complexité</u> liée à l'utilisation du Big Data dans votre municipalité

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL				
No.	Scé	Items en langue anglaise		Items en langue française
CX01	[8]	We find it difficult to process huge amounts of data	⇒	Le traitement de grandes quantités de données est un défi important pour l'exploitation du Big data
CX02	[8]	We find it difficult to process data in real time	⇒	Le traitement de données en temps réel est un défi important pour l'exploitation du Big data
CX03	[8]	We find it difficult to process a variety of data types and sources	⇒	L'exploitation de données de formats différents et provenant de sources variées est un défi majeur pour l'exploitation du Big data
CX04	[8]	We find it difficult to ensure Big data quality	⇒	S'assurer de la qualité des données est un défi majeur pour l'exploitation du Big data
CX05	[3]	Overall, the skills required to use Big data are too complex for our employees/colleagues	⇒	Dans l'ensemble, les compétences requises pour utiliser le Big data sont trop complexes pour nos employés/collègues
CX06	[8]	Overall, leveraging Big data to deliver value is complex	⇒	Dans l'ensemble, la création de la valeur à partir de l'utilisation du Big data est complexe
CX07	[8]	Overall, Big data is more difficult to use than traditional data sets	⇒	Dans l'ensemble, le Big data est plus difficile à utiliser comparativement aux jeux de données traditionnels
		Section 4: Please indicate your opinion about the <u>privacy concerns</u> related to the use of Big Data in your municipality		Section 4: Veuillez indiquer votre opinion sur les <u>préoccupations de confidentialité</u> liées à l'utilisation du Big Data dans votre municipalité
		Our municipality is concerned about	⇒	Notre municipalité est préoccupée par
PC01	[4]	▪ the security of citizens' data	⇒	▪ la sécurité des données des citoyens
PC02	[6]	▪ private information embedded in Big data sets	⇒	▪ les informations confidentielles contenues dans les jeux de données
PC03	[8]	▪ the property rights of datasets	⇒	▪ les droits de propriété des jeux de données
PC04	[8]	▪ revealing proprietary information or sensitive data	⇒	▪ la révélation d'informations exclusives ou de données sensibles
PC05	[4]	Our citizens are significantly concerned about the security and privacy of their data	⇒	Nos citoyens sont de plus en plus préoccupés par la sécurité et la confidentialité de leurs données
PC06	[8]	Legal restrictions limit the way our municipality collects, stores or shares personal data	⇒	Les restrictions légales limitent la façon dont notre municipalité collecte, stocke ou partage des données personnelles
PC07	[8]	When analyzing citizens' data, privacy is a major issue	⇒	Lors de l'analyse de données des citoyens, la vie privée est un problème majeur
PC08	[8]	Our municipality has mechanisms to ensure the safe storage and transmission of its citizens' information	⇒	Notre municipalité dispose de mécanismes de stockage et de transmission sécurisés des données des citoyens R
PC09	[8]	Overall, privacy and digital anonymity are major concerns with Big data use	⇒	Dans l'ensemble, la confidentialité et l'anonymat numérique sont des préoccupations majeures lors de l'utilisation des données
PC10	[8]	Overall, existing regulation concerning privacy and data protection limits the use of Big data	⇒	Dans l'ensemble, les lois sur la vie privée et la sécurité de l'information limitent l'utilisation du Big data par notre municipalité
		Section 5: Please indicate your opinion about the <u>technology context</u> of your municipality		Section 5: Veuillez indiquer votre opinion sur le <u>contexte technologique</u> de votre municipalité
		In our municipality...		Dans notre municipalité...
TC01	[8]	▪ we have access to tools that allow us to handle large data sets	⇒	▪ nous avons accès à des outils permettant la gestion de grands ensembles de données
TC02	[8]	▪ we have access to an IT environment where sensitive data can be processed	⇒	▪ nous disposons d'un environnement où des données sensibles peuvent être traitées
TC03	[8]	▪ we have the software applications that make data use more efficient	⇒	▪ nous avons des logiciels qui permettent une utilisation efficace des données

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL				
No.	Scé	Items en langue anglaise		Items en langue française
TC04	[8]	▪ we have all the data needed for relevant use	⇒	▪ nous détenons toutes les données utiles à une utilisation pertinente
TC05	[8]	▪ an upgrade of the IT infrastructure is necessary to use Big data	⇒	▪ une mise à niveau de l'infrastructure informatique est nécessaire pour utiliser le Big data
TC06	[4]	▪ we developed a plan to upgrade our employees' Big data skills	⇒	▪ un plan a été élaboré pour renforcer les compétences en Big data de nos employés
TC07	[4]	▪ we have recruited staff with specialized skills in Big data	⇒	▪ nous avons recruté du personnel spécialisé dans le Big data
TC08	[4]	▪ we have access to external consultants specialized in Big data	⇒	▪ nous avons accès à des consultants externes spécialisés en Big data
		Overall, to what extent are the following factors preventing your municipality from fully using Big data? (R)		Dans quelle mesure les facteurs suivants empêchent-ils votre municipalité d'utiliser pleinement le Big data ? (R)
TC09	[1]	▪ Inadequate IT infrastructure	⇒	▪ Infrastructure IT inadéquate
TC10	[1]	▪ Limited analytical capabilities	⇒	▪ Capacités d'analyse limitées
TC11	[1]	▪ Lack of skilled resources	⇒	▪ Manque de ressources qualifiées
		Section 6: Please indicate your opinion about the <u>competitive pressures</u> of the external environment of your municipality		Section 6: Veuillez indiquer votre avis sur les <u>pressions compétitives</u> existantes au sein de l'environnement externe de votre municipalité
CP01	[1]	Almost all of the cities nearby are using Big data technologies	⇒	La majorité des villes voisines utilisent les technologies du Big Data
CP02	[8]	Almost all of the cities we compete with are using Big data technologies	⇒	La majorité des villes avec lesquelles nous sommes en concurrence utilisent les technologies du Big data
CP03	[2]	A lack of Big data-capabilities is a competitive disadvantage	⇒	Ne pas disposer de capacités d'utilisation du Big Data constitue un désavantage concurrentiel
CP04	[3]	Using Big data is a strategic necessity to compete with other cities	⇒	L'utilisation du Big data est une nécessité stratégique pour la concurrence avec les autres villes
CP05	[8]	Our municipality aims to become a smart/digital city	⇐	Notre municipalité est engagée dans une démarche de ville intelligente/numérique
		Section 7: Please indicate your opinion about the <u>dynamism of the Big data ecosystem</u> in your city		Section 7: Veuillez indiquer votre opinion sur le <u>dynamisme de l'écosystème du Big data</u> dans votre ville
ED01	[4]	Our partners have technologies ready to support Big data use	⇒	Nos partenaires disposent de technologies supportant l'utilisation du Big data
ED02	[3]	There are businesses in the community which provide training or technical support for Big data use	⇒	Il existe des entreprises dans notre communauté qui peuvent fournir de la formation ou du soutien technique pour l'utilisation du Big data
ED03	[3]	Technology vendors actively market Big data technologies by providing incentives such as free training sessions	⇒	Les vendeurs commercialisent activement les technologies Big data en offrant des incitations telles que des sessions de formation gratuites
ED04	[3]	Our citizens are demanding products/services that need the use of Big data	⇒	Nos citoyens exigent des produits/services qui nécessitent l'utilisation du Big data
ED05	[8]	In our community, there are companies with whom we can launch Big data initiatives through public-private partnerships	⇒	Dans notre communauté, il y a des partenaires avec lesquels notre municipalité peut réaliser des projets de Big data à travers des partenariats public-privé
ED06	[7]	The municipal laws support Big data use	⇒	Les lois municipales offrent un support pertinent à l'utilisation du Big data
ED07	[8]	The central government is implementing a big data strategy/agenda	⇒	Le gouvernement central met en œuvre une stratégie/politique de Big data
ED08	[8]	The legislation supports the access to the data needed by our municipality	⇒	La législation en vigueur facilite l'accès aux données dont notre municipalité a besoin
ED09	[8]	Overall, in our municipality, we have access to all the data we need for	⇒	Dans l'ensemble, nous avons accès à toutes les données dont nous avons besoin pour une prise de

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL			
No.	Scé	Items en langue anglaise	Items en langue française
		actionable insights	décision éclairée
ED10	[8]	Overall, the environment in which our municipality operates is favorable to the use of Big data	↳ Dans l'ensemble, l'environnement dans lequel évolue notre municipalité est favorable pour l'utilisation du Big data
		Section 8: Please indicate your opinion about the <u>top management support to Big data use in your municipality</u>	Section 8: Veuillez indiquer votre opinion concernant le support de la haute direction de votre municipalité pour l'utilisation du Big data
		To what extent does the Top Management Team [TMT] (e.g., Mayor, General Director, CTO, CIO, CDO)	Dans quelle mesure l'équipe de la haute direction (par exemple le maire, directeur général, CTO, CIO, CDO)
TS01	[1]	▪ promote the use of Big data in your municipality?	⇒ ▪ soutient l'utilisation du Big data dans votre organisation?
TS02	[1]	▪ create support for Big data initiatives within your municipality?	⇒ ▪ soutient les initiatives Big data au sein de votre organisation?
TS03	[1]	▪ promote Big data use as a strategic priority within your organization?	⇒ ▪ considère le Big data comme une priorité stratégique pour l'organisation?
TS04	[3]	▪ allocate adequate resources to use Big data technologies?	⇒ ▪ alloue des ressources adéquates à l'utilisation des technologies du Big data?
TS05	[3]	▪ know about the benefits of Big data use?	⇒ ▪ est bien consciente des avantages et bénéfices de l'utilisation du Big data ?
TS06	[8]	▪ consider that data and analytics insights are more reliable than intuition?	⇒ ▪ considère que l'analyse statistique du Big data est plus fiable que l'intuition ?
		In our municipality...	Dans notre municipalité...
TS07	[8]	▪ we appointed a chief data officer (or another title), charged with ensuring high quality data	⇒ ▪ nous avons nommé un responsable principal des données (ou un autre titre) chargé de s'assurer de la qualité des données
TS08	[8]	▪ we established an office that is responsible for the use and management of data	⇒ ▪ nous avons créé une structure qui est responsable de l'utilisation et de la gestion des données
TS09	[8]	▪ we have policy frameworks that support the use of Big data	⇒ ▪ nous avons des politiques-cadres favorisant l'utilisation du Big data
TS10	[8]	Overall, Big data receives active support from top managers	Dans l'ensemble, le Big data reçoit un support actif de la part de la haute direction
TS11	[8]	Overall, top management allocates sufficient resources for Big data use	Dans l'ensemble, la haute direction alloue des ressources suffisantes pour l'utilisation du Big data
		Section 9: Please indicate your opinion about Big data use in your municipality	Section 9: Veuillez indiquer votre opinion sur l'utilisation du Big data par votre municipalité
		In our municipality, we use Big data to inform citizens about...	Dans notre municipalité, nous utilisons le Big data pour informer les citoyens sur ...
UN01	[8]	▪ Travel time estimation on public transport	⇒ ▪ Les délais de trajet sur les transports en commun
UN02	[8]	▪ Traffic and road conditions	⇒ ▪ La circulation et les conditions routières
UN03	[8]	▪ The availability of parking spots	⇒ ▪ La disponibilité des espaces de stationnement
UN04	[8]	▪ The location of towed vehicles	⇒ ▪ La localisation des véhicules remorqués
UN05	[8]	▪ Status of street sweeping or snow removal operations	⇒ ▪ L'état d'avancement des opérations de nettoyage des rues ou du déneigement
UN06	[8]	▪ Collection activities (such as household garbage and recyclable materials)	⇒ ▪ Les activités de collecte (telles que les ordures ménagères et les matières recyclables)
UN07	[8]	▪ Tourist services (such as attractions, events and festivals)	⇒ ▪ Les services touristiques (telles que les attractions, les événements et les festivals)
UN08	[8]	▪ Crimes committed on the city territory	⇒ ▪ Les actes criminels commis sur le territoire de la ville
		In our municipality, we use Big data to...	Dans notre municipalité, nous utilisons le Big data pour ...
UT01		▪ Feed a data warehouse and provide a single source of information	⇒ ▪ Alimenter un entrepôt de données afin de disposer d'une source unique d'information

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL			
No.	Scé	Items en langue anglaise	Items en langue française
UT02	[8]	▪ Share data with the various departments of our municipality	⇒ ▪ Partager les données avec les différents services/départements de notre municipalité
UT03	[9]	▪ Share information/data with other municipalities	- ▪ Partager des informations/données avec d'autres municipalités
UT04	[9]	▪ Share information/data with the provincial/central government	- ▪ Partager des informations/données auprès du gouvernement provincial/central
UT05	[9]	▪ Share information/data with our partners	- ▪ Partager des informations/données avec nos partenaires
UT06	[8]	▪ Integrate services with the citizens (through citizen portals, for example)	⇒ ▪ Intégrer les services offerts aux citoyens (par le biais de portails citoyens par exemple)
UT07	[9]	▪ Integrate services with partners or other government agencies	- ▪ Intégrer nos services avec les partenaires ou les organismes gouvernementaux
UT08	[8]	▪ Become an open and transparent government	⇒ ▪ Devenir un gouvernement ouvert et transparent
		In our municipality, we use Big data in novel and innovative ways such as...	Dans notre municipalité, nous utilisons le Big data de manière novatrice et innovante, comme ...
UM01	[8]	▪ Feeding real time dashboards	⇒ ▪ Intrans aux tableaux de bord en temps réel
UM02	[8]	▪ Improving decision-making	⇒ ▪ Améliorer la prise de décision
UM03	[8]	▪ Forecasting, planning or strategy making	⇒ ▪ Intrans pour établir des prévisions, la planification ou la stratégie
UM04	[1]	▪ Monitoring processes/equipments	⇒ ▪ Suivi des processus / équipements
UM05	[8]	▪ Vehicle tracking (such as buses, garbage collectors, etc.)	⇒ ▪ Localisation des véhicules (de collecte des déchets, autobus, etc.)
UM06	[8]	▪ Intelligent inventory management	⇒ ▪ Gestion intelligente des inventaires
UM07	[8]	▪ Intelligent infrastructure management	▪ Gestion intelligente des infrastructures
UM08	[8]	▪ Efficient resource utilization	⇒ ▪ Utilisation efficace des ressources
UM09	[8]	▪ Process automation	⇒ ▪ Automatiser des processus
UM10	[9]	▪ Generating alternative revenue sources (other than taxes and fines)	- ▪ Générer des sources de revenus alternatives (autres que les taxes et les amendes)
UM11	[8]	▪ Data monetization (i.e., the act of exchanging information-based products or services for legal tender or something of perceived equivalent value)	⇒ ▪ La monétisation des données (l'acte d'échanger un produit/service basé sur des données ou de l'information contre un produit/service de valeur équivalente).
UM12	[8]	▪ Developing new services/products	⇒ ▪ Développement de nouveaux services / produits
UM13	[8]	▪ Promoting entrepreneurship and innovation	⇒ ▪ Promouvoir l'entrepreneuriat et l'innovation
UM14	[8]	▪ Boosting local economic development	⇒ ▪ Stimuler le développement économique local
UM15	[8]	▪ Increasing the efficiency of the transportation management system	⇒ ▪ Accroître l'efficacité du système de transport public
UM16	[8]	▪ Regulating real-time traffic flows (such as traffic light synchronization and suggesting the speed)	⇒ ▪ Réguler en temps réel les flux de circulation (comme la synchronisation des feux de circulation et la suggestion des limites de vitesse)
UM17	[8]	▪ Forecasting transportation demand	⇒ ▪ Établir des prévisions sur la demande en termes de transport
UM18	[8]	▪ Providing better taxi services to citizens	⇒ ▪ Offrir un meilleur service de taxi au citoyens
UM19	[8]	▪ Improving the cycling infrastructure	⇒ ▪ Améliorer l'infrastructure cyclable
UM20	[8]	▪ Establishing a better planning of the urban transportation network	⇒ ▪ Établir une meilleure planification du réseau de transport urbain
UM21	[8]	▪ Monitoring air quality and pollution	⇒ ▪ Surveiller la qualité de l'air et l'état de pollution atmosphérique

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL			
No.	Scé	Items en langue anglaise	Items en langue française
UM22	[8]	▪ Monitoring water or land contamination	⇒ ▪ Surveiller la contamination de l'eau ou des sols
UM23	[8]	▪ Monitoring energy consumption	⇒ ▪ Surveiller la consommation de l'énergie
UM24	[8]	▪ Managing daily waste services	⇒ ▪ Gérer les opérations de traitement des déchets
UM25	[8]	▪ Supporting urban planning and development	⇒ ▪ Soutenir la planification et le développement urbains
UM26	[8]	▪ Supporting emergency services (such as medical, firefighting, rescue, etc.)	⇒ ▪ Soutenir les services d'urgence (tels que les services médicaux, de sauvetage et de lutte contre les incendies)
UM27	[8]	▪ Performing social media analytics for crowd movement monitoring	⇒ ▪ Analyser les médias sociaux pour surveiller le mouvement des foules
UM28	[8]	▪ Performing criminal investigations or fraud detection	⇒ ▪ Effectuer des enquêtes criminelles ou détecter des cas de fraude
UM29	[8]	▪ Ensuring public Safety	⇒ ▪ Assurer la sécurité du public et des lieux publics
UM30	[8]	▪ Designing tactical and strategic security plans	⇒ ▪ Concevoir des plans tactiques et stratégiques de sécurité
		Overall, ...	Dans l'ensemble, ...
BU01	[9]	Overall, using Big data to inform citizens is a common practice for our municipality	- utiliser le Big data pour informer les citoyens est une pratique commune au sein de notre municipalité
BU02	[9]	Overall, our municipality uses Big data to inform citizens	- notre municipalité utilise le Big data pour informer les citoyens
BU03	[9]	Overall, using Big data to coordinate/integrate our activities is a common practice	- utiliser le Big data pour coordonner/intégrer nos activités est une pratique commune au sein de notre municipalité
BU04	[9]	Overall, using Big data to enhance our external linkages is a common practice	- utiliser le Big data pour consolider nos liens externes est une pratique commune au sein de notre municipalité
BU05	[9]	Overall, our municipality has been successful in finding new ways of using Big data	- notre municipalité a su trouver de nouvelles façons d'utiliser le Big data
BU06	[9]	Overall, we always try to use Big data in novel ways	- nous essayons toujours de trouver des façons innovantes d'utiliser le Big data
BU07	[9]	Overall, using Big data has allowed/will allow us to accomplish tasks that were not feasible or identified before	- l'utilisation du Big data nous a permis d'accomplir des tâches qui n'étaient pas faisables ou reconnues auparavant
		Section 10: This section encompasses general information about your municipality to help us classify your answers.	Section 10: Cette section comprend des questions générales sur votre municipalité. Ces données nous aideront dans le classement de vos réponses.
CQ01		In which country your municipality is located? <ul style="list-style-type: none"> ▪ Australia ▪ Belgium ▪ Canada ▪ Japan ▪ Netherlands ▪ Sweden ▪ United-Kingdom ▪ United-States ▪ Other (please, specify) 	Dans quel pays votre municipalité est située? <ul style="list-style-type: none"> ▪ Australie ▪ Belgique ▪ Canada ▪ Japon ▪ Pays-Bas ▪ Suède ▪ Royaume-Uni ▪ Etats-Unis ▪ Autre (veuillez préciser)

Annexe 7 : Items de mesure des variables du modèle de recherche

[1] D. Chen, Preston et Swink (2015) [2] Iacovou, Benbasat et Dexter (1995) [3] Premkumar et Roberts (1999) [4] Zhu <i>et al.</i> (2006) [5] Lin (2008) [6] Kendall <i>et al.</i> (2001) [7] Zhu et Kraemer (2005) [8] RSL			
No.	Scé	Items en langue anglaise	Items en langue française
CQ02		Please indicate the category that best describes the size of your city in terms of population <ul style="list-style-type: none"> ▪ less than 10,000 ▪ between 10,001 and 50,000 ▪ between 50,001 and 100,000 ▪ between 100,001 and 500,000 ▪ between 500,001 and 1,000,000 ▪ more than 1,000,001 	Veillez indiquer la catégorie qui décrit le mieux votre taille en termes de population <ul style="list-style-type: none"> ▪ moins de 10,000 ▪ entre 10,001 et 50,000 ▪ entre 50,001 et 100,000 ▪ entre 100,001 et 500,000 ▪ entre 500,001 et 1,000,000 ▪ plus de 1,000,001
CQ03		Please indicate the category that best describes the size of your municipality in terms of number of employees <ul style="list-style-type: none"> ▪ less than 1000 ▪ between 1001 and 5000 ▪ between 5001 and 10,000 ▪ between 10,001 and 25,000 ▪ more than 25,001 	Veillez indiquer la catégorie qui décrit le mieux la taille de votre municipalité en fonction du nombre d'employés <ul style="list-style-type: none"> ▪ moins de 1000 ▪ entre 1001 et 5000 ▪ entre 5001 et 10,000 ▪ entre 10,001 et 25,000 ▪ plus de 25,001
CQ04		Overall, to what extent your municipality has the financial resources to use Big Data? (1: Far too much – 5: Far too little)	Dans l'ensemble, jusqu'à quel point considérez-vous que votre municipalité dispose de ressources financières pour utiliser le Big data? (1 : Beaucoup – 5 : Aucune)
CQ05		Is your municipality currently using any kind of Big data technologies? (Yes/No)	Votre municipalité utilise-t-elle des technologies du Big data? (Oui/Non)
CQ06		In our municipality, the unit that we use to measure the data volume is <ul style="list-style-type: none"> ▪ Terabyte (1000 Gigabytes) ▪ Petabyte (1000 Terabytes) ▪ Exabyte (1000 Petabytes) ▪ Zettabyte (1000 Exabytes) ▪ Yottabyte (1000 Zettabytes) ▪ Other (please specify): 	Les données que nous manipulons dans notre municipalité se mesurent en <ul style="list-style-type: none"> ▪ Téaoctet (1000 Gigaoctets) ▪ Pétaoctet (1000 Téraoctets) ▪ Exaoctet (1000 Pétaoctets) ▪ Zettaoctet (1000 Exaoctets) ▪ Yottaoctet (1000 Zettaoctets) ▪ Autre (veuillez spécifier)
CQ07		Job title (please specify):	Veillez indiquer votre fonction :
CQ08		Do you have any comments or suggestions?	Avez-vous des commentaires ou des suggestions?

Annexe 8 : Structure de codification des items de mesure

No.	Variable	Code item	Code Qualtrics	Échelle inversée	Nature de l'item	Pattern de réponse	Valeur
1	Avantage relatif [RA]	RA01	4	Non	Formatif	Don't know	-
		RA02	4	Non	Formatif	Disagree completely	1
		RA03	4	Non	Formatif	.	.
		RA04	4	Non	Formatif	.	.
		RA05	4	Non	Formatif	Agree completely	7
		RA06	4	Non	Formatif		
		RA07	4	Non	Formatif		
		RA08	5	Non	Réflexif		
		RA09	5	Non	Réflexif		
		RA10	5	Non	Réflexif		
	Compatibilité [CT]	CT01	6	Non	Formatif	Don't know	-
		CT02		Non	Formatif	Disagree completely	1
		CT03		Non	Formatif	.	.
		CT04		Non	Formatif	.	.
		CT05		Non	Formatif	Agree completely	7
		CT06		Oui	Formatif		
		CT07		Oui	Formatif		
		CT08		Non	Réflexif		
		CT09		Non	Réflexif		
		CT10		Oui	Réflexif		
	Complexité [CX]	CX01	8	Non	Formatif	Don't know	-
		CX02		Non	Formatif	Disagree completely	1
		CX03		Non	Formatif	.	.
		CX04		Non	Formatif	.	.
		CX05		Non	Réflexif	Agree completely	7
		CX06		Non	Réflexif		
		CX07		Non	Réflexif		
	Préoccupations de confidentialité [PC]	PC01	9	Non	Formatif	Don't know	-
		PC02		Non	Formatif	Disagree completely	1
		PC03		Non	Formatif	.	.
		PC04		Non	Formatif	.	.
		PC05		Non	Formatif	Agree completely	7
		PC06		Non	Formatif		
		PC07		Non	Réflexif		
		PC08		Oui	Formatif		
		PC09		Non	Réflexif		
		PC10		Non	Réflexif		
	Compétences technologiques [TC]	TC01	11	Non	Formatif	Don't know	-
		TC02	11	Non	Formatif	Disagree completely	1
		TC03	11	Non	Formatif	.	.
		TC04	11	Non	Formatif	.	.
		TC05	11	Oui	Formatif	Agree completely	7
		TC06	11	Non	Formatif		
		TC07	11	Non	Formatif		
		TC08	11	Non	Formatif		
		TC09	12	Oui	Réflexif		
		TC10	12	Oui	Réflexif		
		TC11	12	Oui	Réflexif		
	Pressions compétitives [CP]	CP01	13	Non	Réflexif	Don't know	-
		CP02		Non	Réflexif	Disagree completely	1
		CP03		Non	Réflexif	.	.
		CP04		Non	Réflexif	.	.
		CP05		Non	Réflexif	Agree completely	7
	Dynamisme de l'écosystème [ED]	ED01	15	Non	Formatif	Don't know	-
		ED02		Non	Formatif	Disagree completely	1
		ED03		Non	Formatif	.	.
		ED04		Non	Formatif	.	.
		ED05		Non	Formatif	Agree completely	7
		ED06		Non	Formatif		
		ED07		Non	Formatif		
		ED08		Non	Formatif		
		ED09		Non	Réflexif		
		ED10		Non	Réflexif		

Annexe 8 : Structure de codification des items de mesure

No.	Variable	Code item	Code Qualtrics	Échelle inversée	Nature de l'item	Pattern de réponse	Valeur
	Support de la haute direction [TS]	TS01	16	Non	Formatif	Don't know	-
		TS02	16	Non	Formatif	Disagree completely	1
		TS03	16	Non	Formatif	.	.
		TS04	16	Non	Formatif	.	.
		TS05	16	Non	Formatif	Agree completely	7
		TS06	16	Non	Formatif		
		TS07	17	Non	Formatif		
		TS08	17	Non	Formatif		
		TS09	17	Non	Formatif		
		TS10	17	Non	Réflexif		
		TS11	17	Non	Réflexif		
	Utilisation informationnelle [UN]	UN01	19	Non	Formatif	Don't know	-
		UN02		Non	Formatif	Not at all	1
		UN03		Non	Formatif	.	2
		UN04		Non	Formatif	.	3
		UN05		Non	Formatif	.	4
		UN06		Non	Formatif	.	5
		UN07		Non	Formatif	.	6
		UN08		Non	Formatif	Very much	7
	Utilisation intégrationnelle [UT]	UT01	20	Non	Formatif	Don't know	-
		UT02		Non	Formatif	Not at all	1
		UT03		Non	Formatif	.	2
		UT04		Non	Formatif	.	3
		UT05		Non	Formatif	.	4
		UT06		Non	Formatif	.	5
		UT07		Non	Formatif	.	6
		UT08		Non	Formatif	Very much	7
	Utilisation émergente [UM]	UM01	22	Non	Formatif	Don't know	-
		UM02		Non	Formatif	Not at all	1
		UM03		Non	Formatif	.	2
		UM04		Non	Formatif	.	3
		UM05		Non	Formatif	.	4
		UM06		Non	Formatif	.	5
		UM07		Non	Formatif	.	6
		UM08		Non	Formatif	Very much	7
		UM09		Non	Formatif		
		UM10		Non	Formatif		
		UM11		Non	Formatif		
		UM12		Non	Formatif		
		UM13		Non	Formatif		
		UM14		Non	Formatif		
		UM15		Non	Formatif		
		UM16		Non	Formatif		
		UM17		Non	Formatif		
		UM18		Non	Formatif		
		UM19		Non	Formatif		
		UM20		Non	Formatif		
		UM21		Non	Formatif		
		UM22		Non	Formatif		
		UM23		Non	Formatif		
		UM24		Non	Formatif		
		UM25		Non	Formatif		
		UM26		Non	Formatif		
		UM27		Non	Formatif		
		UM28		Non	Formatif		
		UM29		Non	Formatif		
		UM30		Non	Formatif		
		BU01	23	Non	Réflexif	Don't know	-
		BU02		Non	Réflexif	Not at all	1
		BU03		Non	Réflexif	.	2
		BU04		Non	Réflexif	.	3
		BU05		Non	Réflexif	.	4
		BU06		Non	Réflexif	.	5
		BU07		Non	Réflexif	.	6
					Very much	7	

Annexe 8 : Structure de codification des items de mesure

No.	Variable	Code item	Code Qualtrics	Échelle inversée	Nature de l'item	Pattern de réponse	Valeur
	Variables de contrôle [CQ]	CQ01	26	Non		Australie	1
			26	Non		Belgique	2
			26	Non		Canada	3
			26	Non		Japon	4
			26	Non		Pays-Bas	5
			26	Non		Suède	6
			26	Non		Royaume-Uni	7
			26	Non		États-Unis	8
			26_9_TEXT	Non		Autre (veuillez préciser)	
		CQ02	27	Non		Moins de 10,000	1
				Non		Entre 10,001 et 50,000	2
				Non		Entre 50,001 et 100,000	3
				Non		Entre 100,001 et 500,000	4
				Non		Entre 500,001 et 1,000,000	5
				Non		Plus de 1,000,000	6
		CQ03	28	Non		Moins de 1000	1
				Non		Entre 1001 et 5000	2
				Non		Entre 5001 et 10,000	3
				Non		Entre 10,001 et 25,000	4
				Non		Plus de 25,000	5
		CQ04	29	Oui		Far too much	1
						.	2
						.	3
						.	4
						Far too little	5
		CQ05	30	Oui		Yes	1
				Oui		No	2
		CQ06	Q29	Non		Téraoctet (1000 Gigaoctets)	1
			Q29	Non		Pétaoctet (1000 Téraoctets)	2
			Q29	Non		Exaoctet (1000 Pétaoctets)	3
			Q29	Non		Zettaoctet (1000 Exaoctets)	4
			Q29	Non		Yottaoctet (1000 Zettaoctets)	5
			Q29_6_TEXT	Non		Autre (veuillez spécifier)	
		CQ07	31	Non		Texte libre (fonction du répondant)	
		CQ08	32	Non		Texte libre (commentaires et remarques)	

Annexe 9 : Données démographiques

	Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Catégorisation des villes en fonction de la population				
Moins de 10 000	35	33.0	33.0	33.0
Entre 10 001 et 50 000	36	34.0	34.0	67.0
Entre 50 001 et 100 000	14	13.2	13.2	80.2
Entre 100 001 et 500 000	14	13.2	13.2	93.4
Entre 500 001 et 1 000 000	4	3.8	3.8	97.2
Plus de 1 000 000	3	2.8	2.8	100.0
Total	106	100.0	100.0	
Catégorisation des villes en fonction du nombre d'employés				
Moins de 1000	81	76.4	76.4	76.4
Entre 1001 et 5000	18	17.0	17.0	93.4
Entre 5001 et 10 000	2	1.9	1.9	95.3
Entre 10 001 et 25 000	3	2.8	2.8	98.1
Plus de 25 001	2	1.9	1.9	100.0
Total	106	100.0	100.0	
Fonctions occupées par les répondants				
CAO/City Manager	33	31.1	31.1	31.1
CIO	7	6.6	6.6	37.7
Director of Corporate Services	6	5.7	5.7	43.4
IT Director/Manager	39	36.8	36.8	80.2
IT Professional	10	9.4	9.4	89.6
Mayor	2	1.9	1.9	91.5
Autre	2	1.9	1.9	93.4
Données manquantes	7	6.6	6.6	100.0
Total	106	100.0	100.0	
Disponibilité des ressources financières				
Beaucoup	1	0.9	0.9	0.9
	4	3.8	3.8	4.7
	22	20.8	20.8	25.5
	36	34.0	34.0	59.4
Aucune	43	40.6	40.6	100.0
Total	106	100.0	100.0	
Unité de mesure de la quantité des données utilisée par les villes participantes				
Térambit (1000 Gigabits)	82	77.4	80.4	80.4
Pérambit (1000 Térabits)	9	8.5	8.8	89.2
Autre	11	10.4	10.8	100.0
Total	102	96.2	100.0	
Données manquantes	4	3.8		
Total	106	100.0		
Utilisation des technologies du Big Data				
Oui	50	47.2	47.2	47.2
Non	56	52.8	52.8	100.0
Total	106	100.0	100.0	
Version du questionnaire utilisée				
Version anglaise	58	54.7	54.7	54.7
Version française	48	45.3	45.3	100.0
Total	106	100.0	100.0	

Annexe 10 : Contributions factorielles croisées

Tableau 26 : Contributions factorielles croisées (utilisation informationnelle)

	Avantage relatif	Compatibilité	Complexité	Compétences technologiques	Dynamisme de l'écosystème	Pressions compétitives	Préoccupations de confidentialité	Support de la haute direction	Utilisation information.
BU01	0.350	0.483	-0.273	0.231	0.465	0.461	0.032	0.330	1.001
BU02	0.297	0.423	-0.260	0.185	0.405	0.454	-0.012	0.309	0.900
CP01	0.154	0.325	-0.381	0.231	0.198	0.554	-0.148	0.237	0.289
CP02	0.124	0.277	-0.399	0.247	0.249	0.553	-0.112	0.222	0.304
CP03	0.618	0.199	-0.048	-0.015	0.226	0.650	0.059	0.147	0.251
CP04	0.717	0.191	-0.086	-0.071	0.197	0.731	0.107	0.142	0.337
CP05	0.199	0.109	0.055	0.275	0.328	0.482	0.282	0.231	0.263
CT08	0.204	0.849	-0.356	0.367	0.358	0.409	0.001	0.182	0.395
CT09	0.221	0.879	-0.349	0.357	0.414	0.302	-0.040	0.277	0.421
CT10	-0.193	0.316	-0.330	0.250	0.063	-0.109	-0.168	0.120	0.172
CX05	0.000	-0.441	0.920	-0.475	-0.230	-0.265	0.117	-0.195	-0.240
CX06	0.087	-0.337	0.785	-0.468	-0.068	-0.173	0.157	-0.010	-0.289
CX07	0.071	-0.257	0.631	-0.376	-0.150	-0.216	0.123	-0.018	-0.117
ED09	-0.192	0.193	-0.219	0.233	0.287	-0.048	0.155	0.057	0.198
ED10	0.345	0.435	-0.159	0.282	1.053	0.451	0.169	0.479	0.465
PC07	0.015	-0.155	0.055	-0.145	0.120	0.050	0.652	-0.142	-0.015
PC09	0.220	0.055	0.092	-0.082	0.256	0.169	0.886	-0.057	0.084
PC10	-0.031	-0.056	0.243	-0.165	-0.042	-0.167	0.544	-0.200	-0.087
RA08	1.042	0.227	-0.055	0.100	0.364	0.626	0.150	0.160	0.353
RA09	0.690	0.064	0.146	-0.043	0.102	0.523	0.049	0.099	0.225
RA10	0.807	0.180	0.113	-0.014	0.152	0.534	0.106	0.128	0.287
TC09	0.053	0.438	-0.373	0.759	0.318	0.199	-0.022	0.127	0.216
TC10	0.054	0.429	-0.481	0.897	0.243	0.157	-0.151	0.367	0.176
TC11	-0.033	0.253	-0.548	0.857	0.216	0.120	-0.244	0.398	0.166
TS10	0.147	0.255	-0.157	0.325	0.417	0.330	-0.145	0.965	0.335
TS11	0.099	0.190	0.011	0.272	0.291	0.162	-0.134	0.647	0.201

Tableau 27 : Contributions factorielles croisées (utilisation intégrationnelle)

	Avantage relatif	Compatibilité	Complexité	Compétences technologiques	Dynamisme de l'écosystème	Pressions compétitives	Préoccupations de confidentialité	Support de la haute direction	Utilisation intégrationnelle
BU03	0.252	0.358	-0.197	0.209	0.353	0.384	-0.093	0.400	0.881
BU04	0.268	0.340	-0.266	0.248	0.376	0.508	-0.121	0.407	0.981
CP01	0.154	0.327	-0.384	0.231	0.172	0.582	-0.172	0.237	0.334
CP02	0.124	0.279	-0.401	0.246	0.217	0.572	-0.126	0.223	0.332
CP03	0.616	0.197	-0.048	-0.015	0.217	0.646	0.013	0.147	0.240
CP04	0.714	0.187	-0.086	-0.071	0.189	0.704	0.071	0.142	0.287
CP05	0.199	0.107	0.056	0.273	0.300	0.477	0.234	0.231	0.268
CT08	0.204	0.843	-0.356	0.364	0.320	0.411	-0.014	0.183	0.305
CT09	0.221	0.860	-0.349	0.356	0.368	0.305	-0.072	0.277	0.296
CT10	-0.193	0.343	-0.332	0.250	0.053	-0.104	-0.141	0.120	0.213
CX05	-0.001	-0.446	0.900	-0.474	-0.202	-0.271	0.153	-0.195	-0.162
CX06	0.086	-0.343	0.778	-0.468	-0.053	-0.180	0.171	-0.010	-0.266
CX07	0.070	-0.263	0.660	-0.377	-0.129	-0.223	0.144	-0.018	-0.167
ED09	-0.190	0.194	-0.220	0.232	0.255	-0.039	0.152	0.057	0.104
ED10	0.345	0.437	-0.159	0.280	1.186	0.449	0.118	0.480	0.463
PC07	0.015	-0.160	0.055	-0.145	0.106	0.047	0.651	-0.142	-0.058
PC09	0.220	0.052	0.093	-0.083	0.226	0.162	0.637	-0.058	-0.012
PC10	-0.031	-0.055	0.244	-0.166	-0.043	-0.170	0.801	-0.200	-0.155
RA08	1.055	0.223	-0.054	0.099	0.336	0.617	0.103	0.160	0.306
RA09	0.685	0.059	0.147	-0.043	0.104	0.514	0.021	0.099	0.171
RA10	0.799	0.178	0.114	-0.014	0.148	0.525	0.083	0.128	0.225
TC09	0.054	0.442	-0.374	0.741	0.277	0.204	-0.023	0.128	0.192
TC10	0.054	0.436	-0.483	0.904	0.214	0.162	-0.176	0.368	0.209
TC11	-0.032	0.258	-0.551	0.870	0.190	0.126	-0.264	0.399	0.219
TS10	0.147	0.257	-0.156	0.326	0.378	0.333	-0.179	0.961	0.419
TS11	0.099	0.193	0.013	0.273	0.264	0.162	-0.144	0.650	0.276

Tableau 28 : Contributions factorielles croisées (utilisation émergente)

	Avantage relatif	Compatibilité	Complexité	Compétences technologiques	Dynamisme de l'écosystème	Pressions compétitives	Préoccupations de confidentialité	Support de la haute direction	Utilisation intégrationnelle
BU05	0.222	0.449	-0.237	0.287	0.353	0.404	-0.061	0.429	0.931
BU06	0.261	0.462	-0.258	0.267	0.349	0.468	-0.009	0.378	0.951
BU07	0.243	0.261	-0.064	0.056	0.311	0.475	-0.045	0.361	0.739
CP01	0.154	0.328	-0.382	0.232	0.184	0.569	-0.153	0.238	0.342
CP02	0.124	0.280	-0.400	0.247	0.231	0.554	-0.115	0.223	0.322
CP03	0.613	0.196	-0.048	-0.015	0.222	0.619	0.051	0.148	0.213
CP04	0.711	0.186	-0.087	-0.071	0.193	0.687	0.101	0.142	0.282
CP05	0.199	0.107	0.056	0.275	0.313	0.520	0.278	0.233	0.364
CT08	0.204	0.841	-0.356	0.366	0.337	0.415	-0.004	0.184	0.372
CT09	0.221	0.853	-0.349	0.357	0.389	0.308	-0.048	0.279	0.359
CT10	-0.192	0.353	-0.330	0.250	0.057	-0.106	-0.168	0.122	0.246
CX05	-0.003	-0.448	0.917	-0.475	-0.215	-0.268	0.125	-0.195	-0.181
CX06	0.084	-0.346	0.770	-0.468	-0.059	-0.179	0.160	-0.008	-0.219
CX07	0.069	-0.264	0.649	-0.377	-0.138	-0.221	0.127	-0.016	-0.118
ED09	-0.188	0.195	-0.219	0.233	0.269	-0.039	0.156	0.058	0.136
ED10	0.345	0.437	-0.160	0.282	1.123	0.459	0.162	0.484	0.424
PC07	0.016	-0.161	0.055	-0.145	0.112	0.055	0.674	-0.144	-0.068
PC09	0.220	0.051	0.092	-0.083	0.239	0.171	0.832	-0.059	0.010
PC10	-0.030	-0.055	0.244	-0.165	-0.043	-0.170	0.582	-0.201	-0.045
RA08	1.072	0.222	-0.055	0.100	0.350	0.623	0.143	0.161	0.323
RA09	0.669	0.057	0.146	-0.044	0.104	0.515	0.044	0.100	0.142
RA10	0.794	0.178	0.114	-0.014	0.151	0.526	0.102	0.130	0.213
TC09	0.055	0.443	-0.373	0.759	0.296	0.214	-0.023	0.130	0.234
TC10	0.054	0.438	-0.482	0.890	0.227	0.168	-0.158	0.371	0.173
TC11	-0.031	0.259	-0.549	0.863	0.202	0.132	-0.250	0.402	0.210
TS10	0.147	0.257	-0.157	0.325	0.396	0.339	-0.153	0.941	0.406
TS11	0.099	0.193	0.011	0.272	0.276	0.167	-0.137	0.664	0.309

Annexe 11 : Significativité des items des modèles MIMIC

Construit	Item	Sample Mean	Standard Deviation	T Statistics	P Values	Coefficient de saturation
Avantage relatif	RA01 -> RA	0.207	0.126	1.736	0.083	0.807
	RA02 -> RA	0.272	0.143	1.868	0.062	0.879
	RA03 -> RA	0.173	0.120	1.465	0.143	0.780
	RA04 -> RA	-0.030	0.129	0.265	0.791	0.701
	RA05 -> RA	0.294	0.113	2.686	0.007	0.833
	RA06 -> RA	0.096	0.129	0.778	0.437	0.800
	RA07 -> RA	0.182	0.121	1.466	0.143	0.800
	RA08 <- RA	0.933	0.035	26.898	0.000	0.940
	RA09 <- RA	0.751	0.060	12.409	0.000	0.746
	RA10 <- RA	0.859	0.040	21.812	0.000	0.862
Compatibilité	CT01 -> CT	-0.076	0.252	0.401	0.689	0.666
	CT02 -> CT	0.504	0.238	2.230	0.026	0.896
	CT03 -> CT	0.021	0.163	0.029	0.977	0.645
	CT04 -> CT	0.439	0.198	2.489	0.013	0.917
	CT05 -> CT	0.173	0.172	1.181	0.238	0.481
	CT06 -> CT	0.240	0.202	1.173	0.241	0.171
	CT07 -> CT	-0.155	0.203	0.752	0.452	0.019
	CT08 <- CT	0.835	0.064	12.800	0.000	0.826
	CT09 <- CT	0.953	0.061	15.832	0.000	0.962
Complexité	CX01 -> CX	0.027	0.234	0.127	0.899	0.839
	CX02 -> CX	0.093	0.273	0.453	0.650	0.915
	CX03 -> CX	0.609	0.287	2.171	0.030	0.972
	CX04 -> CX	0.321	0.196	1.524	0.127	0.857
	CX05 <- CX	0.797	0.068	11.824	0.000	0.802
	CX06 <- CX	0.789	0.075	10.510	0.000	0.790
	CX07 <- CX	0.739	0.093	8.010	0.000	0.744
Préoccupations de confidentialité	PC01 -> PC	0.197	0.144	1.420	0.156	0.782
	PC02 -> PC	0.124	0.163	0.704	0.481	0.802
	PC03 -> PC	0.118	0.165	0.780	0.436	0.722
	PC04 -> PC	0.040	0.180	0.220	0.826	0.749
	PC05 -> PC	0.312	0.135	2.469	0.014	0.816
	PC06 -> PC	0.417	0.123	3.393	0.001	0.774
	PC07 <- PC	0.665	0.085	7.733	0.000	0.659
	PC08 -> PC	-0.089	0.088	1.004	0.315	-0.330
	PC09 <- PC	0.759	0.068	11.233	0.000	0.761
	PC10 <- PC	0.668	0.073	9.242	0.000	0.674
Compétences technologiques	TC01 -> TC	0.115	0.115	1.077	0.281	0.674
	TC02 -> TC	0.114	0.159	0.780	0.435	0.575
	TC03 -> TC	-0.043	0.182	0.312	0.755	0.534
	TC04 -> TC	0.176	0.161	1.120	0.263	0.519

Formes et antécédents de l'utilisation des mégadonnées par les villes

Construit	Item	Sample Mean	Standard Deviation	T Statistics	P Values	Coefficient de saturation
	TC05 -> TC	0.601	0.106	5.798	0.000	0.827
	TC06 -> TC	0.008	0.145	0.088	0.930	0.541
	TC07 -> TC	0.273	0.126	2.185	0.029	0.566
	TC08 -> TC	0.194	0.104	1.923	0.055	0.556
	TC09 <- TC	0.822	0.057	14.744	0.000	0.835
	TC10 <- TC	0.857	0.044	19.397	0.000	0.852
	TC11 <- TC	0.823	0.055	14.845	0.000	0.817
Dynamisme de l'écosystème	ED01 -> ED	0.133	0.235	0.747	0.455	0.534
	ED02 -> ED	0.193	0.210	1.047	0.295	0.249
	ED03 -> ED	0.022	0.215	0.067	0.947	0.296
	ED04 -> ED	-0.067	0.196	0.107	0.915	0.143
	ED05 -> ED	-0.132	0.247	0.808	0.419	0.179
	ED06 -> ED	0.044	0.283	0.046	0.963	0.591
	ED07 -> ED	0.341	0.233	1.694	0.090	0.732
	ED08 -> ED	0.546	0.266	2.472	0.013	0.898
	ED09 <- ED	0.787	0.142	5.728	0.000	0.814
	ED10 <- ED	0.799	0.107	7.456	0.000	0.800
Support de la haute direction	TS01 -> TS	0.002	0.144	0.097	0.923	0.532
	TS02 -> TS	0.176	0.182	0.900	0.368	0.615
	TS03 -> TS	0.232	0.158	1.418	0.156	0.606
	TS04 -> TS	0.123	0.147	0.944	0.345	0.701
	TS05 -> TS	0.017	0.132	0.217	0.828	0.481
	TS06 -> TS	-0.163	0.089	1.935	0.053	0.336
	TS07 -> TS	0.427	0.156	2.853	0.004	0.698
	TS08 -> TS	-0.073	0.158	0.525	0.600	0.617
	TS09 -> TS	0.837	0.053	16.060	0.000	0.857
	TS10 <- TS	0.781	0.064	12.077	0.000	0.778
	TS11 <- TS	0.798	0.084	9.596	0.000	0.803
Utilisation informationnelle	BU01 <- UN	0.921	0.038	23.962	0.000	0.917
	BU02 <- UN	0.976	0.035	27.752	0.000	0.982
	UN01 -> UN	0.160	0.200	0.902	0.367	0.394
	UN02 -> UN	0.035	0.224	0.303	0.762	0.613
	UN03 -> UN	0.052	0.172	0.390	0.696	0.400
	UN04 -> UN	0.076	0.135	0.567	0.571	0.278
	UN05 -> UN	0.160	0.166	0.939	0.348	0.609
	UN06 -> UN	0.351	0.251	1.466	0.143	0.778
	UN07 -> UN	0.403	0.208	2.083	0.037	0.804
	UN08 -> UN	0.159	0.209	0.795	0.427	0.664
Utilisation intégrationnelle	BU03 <- UT	0.916	0.054	17.090	0.000	0.918
	BU04 <- UT	0.943	0.036	26.420	0.000	0.941
	UT01 -> UT	0.051	0.190	0.237	0.813	0.652
	UT02 -> UT	0.251	0.213	1.216	0.224	0.839

Formes et antécédents de l'utilisation des mégadonnées par les villes

Construit	Item	Sample Mean	Standard Deviation	T Statistics	P Values	Coefficient de saturation
	UT03 -> UT	0.266	0.205	1.211	0.226	0.732
	UT04 -> UT	-0.121	0.210	0.613	0.540	0.701
	UT05 -> UT	-0.123	0.227	0.501	0.617	0.656
	UT06 -> UT	0.330	0.190	1.826	0.068	0.888
	UT07 -> UT	0.278	0.209	1.501	0.133	0.856
	UT08 -> UT	0.175	0.198	0.905	0.366	0.889
Utilisation émergente	BU05 <- UM	0.877	0.034	25.188	0.000	0.863
	BU06 <- UM	0.889	0.030	28.916	0.000	0.877
	BU07 <- UM	0.856	0.036	24.537	0.000	0.879
	UM01 -> UM	0.080	0.196	0.067	0.947	0.688
	UM02 -> UM	0.241	0.213	1.364	0.173	0.838
	UM03 -> UM	0.249	0.244	1.167	0.243	0.858
	UM04 -> UM	-0.230	0.219	1.451	0.147	0.715
	UM05 -> UM	0.164	0.136	1.519	0.129	0.610
	UM06 -> UM	0.063	0.236	0.350	0.726	0.532
	UM07 -> UM	-0.281	0.261	1.178	0.239	0.547
	UM08 -> UM	0.140	0.192	0.599	0.549	0.764
	UM09 -> UM	0.108	0.198	0.861	0.390	0.685
	UM10 -> UM	0.210	0.230	0.831	0.406	0.503
	UM11 -> UM	-0.043	0.155	0.342	0.732	0.344
	UM12 -> UM	0.248	0.197	1.062	0.288	0.699
	UM13 -> UM	0.135	0.263	0.993	0.321	0.704
	UM14 -> UM	-0.060	0.234	0.735	0.462	0.766
	UM15 -> UM	-0.230	0.220	0.953	0.341	0.392
	UM16 -> UM	-0.023	0.242	0.050	0.960	0.367
	UM17 -> UM	-0.001	0.202	0.029	0.977	0.382
	UM18 -> UM	-0.110	0.183	0.564	0.573	0.241
	UM19 -> UM	0.065	0.217	0.293	0.769	0.641
	UM20 -> UM	0.118	0.211	0.413	0.680	0.509
	UM21 -> UM	0.030	0.183	0.443	0.658	0.381
	UM22 -> UM	-0.044	0.147	0.399	0.690	0.487
	UM23 -> UM	0.101	0.186	0.652	0.515	0.582
	UM24 -> UM	-0.248	0.172	1.567	0.117	0.354
	UM25 -> UM	-0.156	0.251	0.727	0.467	0.690
	UM26 -> UM	0.334	0.264	1.377	0.169	0.732
	UM27 -> UM	0.155	0.183	1.349	0.177	0.562
UM28 -> UM	0.004	0.173	0.025	0.980	0.369	
UM29 -> UM	0.099	0.256	0.370	0.711	0.620	
UM30 -> UM	-0.097	0.211	0.378	0.706	0.472	

Bibliographie :

- Aaltonen, Aleksi et Niccoló Tempini (2014). « Everything counts in large amounts: A critical realist case study on data-based production » [journal article], *Journal of Information Technology*, vol. 29, no 1, p. 97-110.
- Abbasi, Ahmed, Suprateek Sarker et Roger H. L. Chiang (2016). « Big data research in information systems: Toward an inclusive research agenda » [Article], *Journal of the Association for Information Systems*, vol. 17, no 2, p. i-xxxii.
- Abella, Alberto, Marta Ortiz-de-Urbina-C. et Carmen De-Pablos-H. (2017). « A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems » [Article], *Cities*, vol. 64, p. 47-53.
- Ackoff, Russell L (1989). « From data to wisdom », *Journal of applied systems analysis*, vol. 16, no 1, p. 3-9.
- Agarwal, Ritu, Corey M Angst, Catherine M DesRoches et Michael A Fischer (2010). « Technological viewpoints (frames) about electronic prescribing in physician practices », *Journal of American Medical Information Association*, vol. 17, no 4, p. 425-431.
- Agarwal, Ritu et Elena Karahanna (2000). « Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage », *MIS Quarterly*, vol. 24, no 4, p. 665-694.
- Aguilera, Unai, Oscar Peña, Oscar Belmonte et Diego López-de-Ipiña « Citizen-centric data services for smarter cities », *Future Generation Computer Systems*.
- Ajzen, Icek (1991). « The theory of planned behavior », *Organizational behavior and human decision processes*, vol. 50, no 2, p. 179-211.
- Alkhalil, Adel, Reza Sahandi et David John (2017). « An exploration of the determinants for decision to migrate existing resources to cloud computing using an integrated toe-doi model », *Journal of Cloud Computing*, vol. 6, no 1, p. 1.
- Anderson, Chris (2008). « The end of theory: The data deluge makes the scientific method obsolete », *Wired magazine*, vol. 16, no 7, p. 16-07.
- Anderson, Kim Viborg et Helle Zinner Henriksen (2005). *The first leg of e-government research: Domains and application areas 1998-2003*, IGI Global.
- Armstrong, Curtis P et Vallabh Sambamurthy (1999). « Information technology assimilation in firms: The influence of senior leadership and it infrastructures », *Information systems research*, vol. 10, no 4, p. 304-327.
- Awad, Elias M et Hassan M Ghaziri (2004). « Knowledge management, 2004 », ed: Prentice-Hall, Upper Saddle River, New Jersey.
- Baesens, Bart, Ravi Bapna, James R. Marsden, Jan Vanthienen et J. Leon Zhao (2016). « Transformational issues of big data and analytics in networked business » [Article], *MIS Quarterly*, vol. 40, no 4, p. 807-818.
- Barclay, Donald, Christopher Higgins et Ronald Thompson (1995). « The partial least squares (pls) approach to causal modeling: Personal computer adoption and use as an illustration », *Technology studies*, vol. 2, no 2, p. 285-309.
- Barki, Henri et Jon Hartwick (1994). « Measuring user participation, user involvement, and user attitude », *MIS Quarterly*, vol. 18, no 1, p. 59-82.

- Barki, Henri, Guy Paré et Claude Sicotte (2008). « Linking it implementation and acceptance via the construct of psychological ownership of information technology », *Journal of Information Technology*, vol. 23, no 4, p. 269-280.
- Barki, Henri, Ryad Titah et Céline Boffo (2007). « Information system use-related activity: An expanded behavioral conceptualization of individual-level information system use », *Information Systems Research*, vol. 18, no 2, p. 173-192.
- Beatty, Robert C., J. P. Shim et Mary C. Jones (2001). « Factors influencing corporate web site adoption: A time-based assessment », *Information & Management*, vol. 38, no 6, p. 337-354.
- Blanche-Gracia, Daniel, Luis V. Casalo-Ariño et Alfredo Pérez-Rueda (2015). « Determinants of multi-service smartcard success for smart cities development: A study based on citizens' privacy and security perceptions », *Government Information Quarterly*, vol. 32, no 2, p. 154-163.
- Bellinger, Gene, Durval Castro et Anthony Mills (2004). « Data, information, knowledge, and wisdom ».
- Bergamaschi, Sonia, Emanuele Carlini, Michelangelo Ceci, Barbara Furletti, Fosca Giannotti, Donato Malerba, *et al.* (2016). « Big data research in Italy: A perspective », *Engineering*, vol. 2, no 2, p. 163-170.
- Bhattacharjee, Anol et Chieh-peng Lin (2015). « A unified model of IT continuance: Three complementary perspectives and crossover effects », *European Journal of Information Systems*, vol. 24, no 4, p. 364-373.
- Bhimani, Alnoor (2015). « Exploring big data's strategic consequences », *Journal of Information Technology*, vol. 30, no 1, p. 66-69.
- Bollen, Kenneth A. (1989). *Structural equations with latent variables*, New York, Wiley.
- Boudreau, Marie-Claude et L. Seligman (2006). « Quality of use of a complex technology: A learning-based model », *Contemporary Issues in End User Computing*, vol. 248.
- Boyd, Danah et Kate Crawford (2012). « Critical questions for big data : Provocations for a cultural, technological, and scholarly phenomenon », *Information, Communication & Society*, vol. 15, no 5, p. 662-679.
- Boynton, Andrew C, Robert W Zmud et Gerry C Jacobs (1994). « The influence of IT management practice on IT use in large organizations », *MIS Quarterly*, p. 299-318.
- Brynjolfsson, Erik, Tomer Geva et Shachar Reichman (2016). « Crowd-squared: Amplifying the predictive power of search trend data » [Article], *MIS Quarterly*, vol. 40, no 4, p. 941-A936.
- Bryson, Steve, David Kenwright, Michael Cox, David Ellsworth et Robert Haines (1999). « Visually exploring gigabyte data sets in real time », *Commun. ACM*, vol. 42, no 8, p. 82-90.
- Burton-Jones, Andrew et Camille Grange (2013). « From use to effective use: A representation theory perspective », *INFORMATION SYSTEMS RESEARCH*, vol. 24, no 3, p. 632-658.
- Burton-Jones, Andrew et Detmar W. Straub, Jr. (2006). « Reconceptualizing system usage: An approach and empirical test », *Information Systems Research*, vol. 17, no 3, p. 228-246.
- Campbell, Donald T et Donald W Fiske (1959). « Convergent and discriminant validation by the multitrait-multimethod matrix », *Psychological bulletin*, vol. 56, no 2, p. 81.
- Carmines, Edward G. et Richard A. Zeller (1979). *Reliability and validity assessment*, vol. 17;no. 07-017;, Beverly Hills, Sage.

- Carter, Lemuria et France Bélanger (2005). « The utilization of e- government services: Citizen trust, innovation and acceptance factors », *Information systems journal*, vol. 15, no 1, p. 5-25.
- Cenfetelli, Ronald T. et Geneviève Bassellier (2009). « Interpretation of formative measurement in information systems research », *MIS Quarterly*, vol. 33, no 4, p. 689-707.
- Chaffey, Dave et Gareth White (2010). *Business information management: Improving performance using information systems*, Pearson Education.
- Chau, Patrick Y. K. et Kar Yan Tam (1997). « Factors affecting the adoption of open systems: An exploratory study », *MIS Quarterly*, vol. 21, no 1, p. 1-24.
- Chauhan, Sumedha, Nectima Agarwal et Arpan Kumar Kar (2016). « Addressing big data challenges in smart cities: A systematic literature review » [Article], *Info*, vol. 18, no 4, p. 73-90.
- Chen, D., David Preston et Morgan Swink (2015). « How the use of big data analytics affects value creation in supply chain management » [Article], *Journal of Management Information Systems*, vol. 32, no 4, p. 4-39.
- Chen, H. C., R. H. L. Chiang et V. C. Storey (2012). « Business intelligence and analytics: From big data to big impact », *MIS QUARTERLY*, vol. 36, no 4, p. 1165-1188.
- Chichernea, Virgil (2015). « Smart cities communities and smart ict platform », *Journal of Information Systems & Operations Management*, p. 1-11.
- Chin, Wynne W (1998b). « The partial least squares approach to structural equation modeling », *Modern methods for business research*, vol. 295, no 2, p. 295-336.
- Chin, Wynne W., Abhijit Gopal et W. David Salisbury (1997). « Advancing the theory of adaptive structuration: The development of a scale to measure faithfulness of appropriation », *Information Systems Research*, vol. 8, no 4, p. 342-367.
- Chin, Wynne W., Barbara L. Marcolin et Peter R. Newsted (2003). « A partial least squares latent variable modeling approach for measuring interaction effects: Results from a monte carlo simulation study and an electronic-mail emotion/adoption study », *Information Systems Research*, vol. 14, no 2, p. 189-217.
- Choo, Chun Wei (1996). « The knowing organization: How organizations use information to construct meaning, create knowledge and make decisions », *International journal of information management*, vol. 16, no 5, p. 329-340.
- Churchill Jr, Gilbert A (1979). « A paradigm for developing better measures of marketing constructs », *Journal of marketing research*, p. 64-73.
- Clarke, Roger (2016). « Big data, big risks », *Information Systems Journal*, vol. 26, no 1, p. 77-90.
- Cleveland, Harlan (1982). « Information as a resource », *Futurist*, vol. 16, no 6, p. 34-39.
- Constantiou, Ioanna D. et Jannis Kallinikos (2015). « New games, new rules: Big data and the changing context of strategy », *Journal of Information Technology*, vol. 30, no 1, p. 44-57.
- Cooley, Mike (1987). *Architect or bee? The human price of technology. Revised edition*, London: The Hogarth Press.
- Cooper, Randolph B et Robert W Zmud (1990). « Information technology implementation research: A technological diffusion approach », *Management science*, vol. 36, no 2, p. 123-139.

- Couper, Mick P (2008). *Designing effective web surveys*, vol. 75, Cambridge University Press New York, NY.
- Covi, L. et R. Kling (1996). « Organizational dimensions of effective digital library use: Closed rational and open natural systems models », *Journal of the american society for information science*, vol. 47, no 9, p. 672-689.
- Davenport, Thomas H. (2014). *Big data at work: Dispelling the myths, uncovering the opportunities*, harvard Business review Press.
- Davenport, Thomas H., Paul Barth et Randy Bean (2012). « How big data is different », *MIT Sloan Management Review*, vol. 54, no 1, p. 43-46.
- Davenport, Thomas H. et Jinho Kim (2013). *Keeping up with the quants: Your guide to understanding and using analytics*, Harvard Business Review Press.
- Davis, Fred D., Richard P. Bagozzi et Paul R. Warshaw (1989). « User acceptance of computer technology: A comparison of two theoretical models », *Management Science*, vol. 35, no 8, p. 982-1003.
- De Gennaro, Michele, Elena Paffumi et Giorgio Martini (2016). « Big data for supporting low-carbon road transport policies in europe: Applications, challenges and opportunities », *Big Data Research*, vol. 6, p. 11-25.
- de Leeuw, Edith D., Joop J. Hox et Anja Boevé (2016). « Handling do-not-know answers: Exploring new approaches in online and mixed-mode surveys », *Social Science Computer Review*, vol. 34, no 1, p. 116-132.
- DeLone, William H. et Ephraim R. McLean (2003). « The delone and mclean model of information systems success: A ten-year update », *Journal of Management Information Systems*, vol. 19, no 4, p. 9-30.
- Derman, Emanuel (2011). *Models. Behaving. Badly.: Why confusing illusion with reality can lead to disaster, on wall street and in life*, Simon and Schuster.
- DeSanctis, Gerardine et Marshall Scott Poole (1994). « Capturing the complexity in advanced technology use: Adaptive structuration theory », *Organization Science*, vol. 5, no 2, p. 121-147.
- Diamantopoulos, Adamantios et Judy A Siguaw (2006). « Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration », *British Journal of Management*, vol. 17, no 4, p. 263-282.
- Diamantopoulos, Adamantios et Heidi M. Winklhofer (2001). « Index construction with formative indicators: An alternative to scale development », *Journal of Marketing Research*, vol. 38, no 2, p. 269-277.
- Dobre, C. et F. Xhafa (2014). « Intelligent services for big data science » [Article], *Future Generation Computer Systems*, vol. 37, p. 267-281.
- Duhigg, Charles (2012). « How companies learn your secrets », *The New York Times*, vol. 16, p. 2012.
- Efron, Bradley et Gail Gong (1983). « A leisurely look at the bootstrap, the jackknife, and cross-validation », *The American Statistician*, vol. 37, no 1, p. 36-48.
- Esposito Vinzi, Vincenzo, Wynne W Chin, J. Henseler et H. Wang (2010). *Handbook of partial least squares: Concepts, methods and applications*, New York;Berlin;, Springer.
- Falk, R Frank et Nancy B Miller (1992). *A primer for soft modeling*, University of Akron Press.

- Fietkiewicz, Kaja J., Agnes Mainka et Wolfgang G. Stock (2017). « Egovernment in cities of the knowledge society. An empirical investigation of smart cities' governmental websites », *Government Information Quarterly*, vol. 34, no 1, p. 75-83.
- Fitzgerald, Michael (2016). « Data-driven city management: A close look at amsterdam's smart city initiative », *MIT Sloan Management Review*, vol. 57, no 4, p. n/a.
- Fornell, Claes (1982). *A second generation of multivariate analysis. 2. Measurement and evaluation*, vol. 2, Praeger Publishers.
- Fornell, Claes et Fred L. Bookstein (1982). « Two structural equation models: Lisrel and pls applied to consumer exit-voice theory », *Journal of Marketing Research*, vol. 19, no 4, p. 440-452.
- Fornell, Claes et David F. Larcker (1981a). « Evaluating structural equation models with unobservable variables and measurement error », *Journal of Marketing Research*, vol. 18, no 1, p. 39-50.
- Fornell, Claes et David F. Larcker (1981b). « Structural equation models with unobservable variables and measurement error: Algebra and statistics », *Journal of Marketing Research*, vol. 18, no 3, p. 382-388.
- Foulonneau, Muriel, Slim Turki, Gérardine Vidou et Sébastien Martin (2014). « Open data in service design », *Electronic Journal of E-Government*, vol. 12, no 2, p. 99-107.
- Foundation, National Science (2012). *Core techniques and technologies for advancing big data science & engineering (bigdata)*. https://nsf.gov/events/event_summ.jsp?cntn_id=124058&org=NSF
- Fung, Archon (2015). « Putting the public back into governance: The challenges of citizen participation and its future », *Public Administration Review*, vol. 75, no 4, p. 513-522.
- Gagliardi, Dimitri, Laura Schina, Marco Lucio Sarcinella, Giovanna Mangialardi, Francesco Niglia et Angelo Corallo « Information and communication technologies and public participation: Interactive maps and value added for citizens », *Government Information Quarterly*.
- Galizia, Thomas, Trevor Gee et Ken Landis (2011). « “Gold rush: The scramble to claim and protect value in the digital world” deloitte ».
- Gantz, John et David Reinsel (2011). « Extracting value from chaos », *IDC iView*, vol. 1142, no 2011, p. 1-12.
- Gantz, John et David Reinsel (2012). « The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east », *IDC iView: IDC Analyze the future*, vol. 2007, no 2012, p. 1-16.
- Gartner (2013). *Gartner predicts business intelligence and analytics will remain top focus for cios through 2017*. Récupéré le 3 mai 2017 de <http://www.gartner.com/newsroom/id/2637615>
- Gefen, David, Detmar Straub et Marie-Claude Boudreau (2000). « Structural equation modeling and regression: Guidelines for research practice », *Communications of the association for information systems*, vol. 4, no 1, p. 7.
- Giffinger, Rudolf, Christian Fertner, Hans Kramar et Evert Meijers (2007). « City-ranking of european medium-sized cities », *Cent. Reg. Sci. Vienna UT*.
- Gil-Garcia, J. Ramon, Jing Zhang et Gabriel Puron-Cid (2016). « Conceptualizing smartness in government: An integrative and multi-dimensional view », *Government Information Quarterly*, vol. 33, no 3, p. 524-534.
- Goes, Paulo B. (2014). « Big data and is research » [Article], *MIS Quarterly*, vol. 38, no 3, p. iii-viii.

- Gold, Andrew H, Arvind Malhotra et Albert H Segars (2001). « Knowledge management: An organizational capabilities perspective », *Journal of management information systems*, vol. 18, no 1, p. 185-214.
- Greenfeld, Norman et James T. Kuznicki (1975). « Implied competence, task complexity, and imitative behavior », *The Journal of Social Psychology*, vol. 95, no 2, p. 251-261.
- Grönlund, Åke et Thomas A Horan (2005). « Introducing e-gov: History, definitions, and issues », *Communications of the association for information systems*, vol. 15, no 1, p. 39.
- Gschwender, Antonio, Marcela Munizaga et Carolina Simonetti (2016). « Using smart card and gps data for policy and planning: The case of transantiago », *Research in Transportation Economics*, vol. 59, p. 242-249.
- Guillamón, María-Dolores, Ana-María Ríos, Benedetta Gesuele et Concetta Metallo (2016). « Factors influencing social media use in local governments: The case of italy and spain », *Government Information Quarterly*, vol. 33, no 3, p. 460-471.
- Hair, Joseph F, William C Black, Barry J Babin, Rolph E Anderson et Ronald L Tatham (1998). « Multivariate data analysis », *Upper Saddle River*.
- Hashem, Ibrahim Abaker Targio, Victor Chang, Nor Badrul Anuar, Kayode Adewole, Ibrar Yaqoob, Abdullah Gani, *et al.* (2016). « The role of big data in smart city » [Article], *International Journal of Information Management*, vol. 36, no 5, p. 748-758.
- Hayashi, Alden M. (2014). « Thriving in a big data world », *MIT Sloan Management Review*, vol. 55, no 2, p. 35-39.
- Heeks, Richard et Savita Bailer (2007). « Analyzing e-government research: Perspectives, philosophies, theories, methods, and practice », *Government information quarterly*, vol. 24, no 2, p. 243-265.
- Henseler, J., Christian M. Ringle et Marko Sarstedt (2015). « A new criterion for assessing discriminant validity in variance-based structural equation modeling », *Journal of the Academy of Marketing Science*, vol. 43, no 1, p. 115-135.
- Herche, Joel et Brian Engelland (1996). « Reversed-polarity items and scale unidimensionality », *Journal of the Academy of Marketing Science*, vol. 24, no 4, p. 366-374.
- Hill, Kashmir (2012). *How target figured out a teen girl was pregnant before her father did*. Récupéré le 2017-06-05 2017 de <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#288302b46668>
- Höchtel, Johann, Peter Parycek et Ralph Schöllhammer (2016). « Big data in the policy cycle: Policy decision making in the digital era », *Journal of Organizational Computing and Electronic Commerce*, vol. 26, no 1-2, p. 147-169.
- Howard, Rick, Cathleen E. Blanton, Rick Holgate, Neville Cannon et Bettina Tratz-Ryan (2016). « The top 10 strategic technology trends for government in 2016 », *Gartner*, no G00308910.
- Iacovou, Charalambos L, Izak Benbasat et Albert S Dexter (1995). « Electronic data interchange and small organizations: Adoption and impact of technology », *MIS quarterly*, p. 465-485.
- International Electrotechnical Commission (2014). *Orchestrating infrastructure for sustainable smart cities* Geneva, Switzerland.
- Janssen, Marijn et George Kuk (2016a). « Big and open linked data (bold) in research, policy, and practice », *Journal of Organizational Computing and Electronic Commerce*, vol. 26, no 1-2, p. 3-13.

- Janssen, Marijn et George Kuk (2016b). « The challenges and limits of big data algorithms in technocratic governance », *Government Information Quarterly*, vol. 33, no 3, p. 371-377.
- Janssen, Marijn et Jeroen van den Hoven (2015). « Big and open linked data (bold) in government: A challenge to transparency and privacy? », *Government Information Quarterly*, vol. 32, no 4, p. 363-368.
- Jara, Antonio J., Dominique Genoud et Yann Bocchi (2015). « Big data for smart cities with knime a real experience in the smartsantander testbed », *Software: Practice and Experience*, vol. 45, no 8, p. 1145-1160.
- Jashapara, Ashok (2004). *Knowledge management: An integrated approach*, Pearson Education.
- Kallinikos, Jannis et Ioanna D. Constantiou (2015). « Big data revisited: A rejoinder », *Journal of Information Technology*, vol. 30, no 1, p. 70-74.
- Kendall, Jon D., Lai Lai Tung, Khoon Hui Chua, Chia Hong Dennis Ng et Suan Meng Tan (2001). « Receptivity of singapore's smes to electronic commerce adoption », *Journal of Strategic Information Systems*, vol. 10, no 3, p. 223-242.
- Kim, Gang-Hoon, Silvana Trimi et Ji-Hyong Chung (2014). « Big-data applications in the government sector », *Commun. ACM*, vol. 57, no 3, p. 78-85.
- Kiron, David, Rebecca Shockley, Nina Kruschwitz, Glenn Finch et Michael Haydock (2012). « Analytics: The widening divide », *MIT Sloan Management Review*, vol. 53, no 2, p. 1.
- Kitchin, Rob (2014). « The real-time city? Big data and smart urbanism », *GeoJournal*, vol. 79, no 1, p. 1-14.
- Klopp, Jacqueline M. et Danielle L. Petretta (2017). « The urban sustainable development goal: Indicators, complexity and the politics of measuring cities », *Cities*, vol. 63, p. 92-97.
- Koppenjan, Joop F. M. et Bert Enserink (2009). « Public–private partnerships in urban infrastructures: Reconciling private sector participation and sustainability », *Public Administration Review*, vol. 69, no 2, p. 284-296.
- Kourtit, Karima, Peter Nijkamp et John Steenbruggen « The significance of digital data systems for smart city policy », *Socio-Economic Planning Sciences*.
- Krosnick, Jon A (1991). « Response strategies for coping with the cognitive demands of attitude measures in surveys », *Applied cognitive psychology*, vol. 5, no 3, p. 213-236.
- Krosnick, Jon A et Stanley Presser (2010). « Question and questionnaire design », *Handbook of survey research*, vol. 2, no 3, p. 263-314.
- Kumar, Ranjit (2014). *Research methodology: A step-by-step guide for beginners*, vol. 4th, Los Angeles, SAGE.
- Kuneva, Meglena (2009). *Keynote speech* [SPEECH/09/156], European Consumer Commission.
- Kwon, Tae H et Robert W Zmud (1987). « Unifying the fragmented models of information systems implementation », communication présentée au *Critical issues in information systems research*,
- Laney, Doug (2001). « 3d data management: Controlling data volume, velocity and variety », *Application Delivery Strategies*, vol. 6, p. 70.

- LaValle, Steve, Eric Lesser, Rebecca Shockley, Michael S. Hopkins et Nina Kruschwitz (2011). « Big data, analytics and the path from insights to value », *MIT Sloan Management Review*, vol. 52, no 2, p. 21-32.
- Lazer, David, Ryan Kennedy, Gary King et Alessandro Vespignani (2014). « Big data. The parable of google flu: Traps in big data analysis », *Science*, vol. 343, no 6176, p. 1203.
- Lee, Jungwoo et Hyejung Lee (2014). « Developing and validating a citizen-centric typology for smart city services », *Government Information Quarterly*, vol. 31, Supplement 1, p. S93-S105.
- LeRouge, Cynthia, Alan R Hevner et Rosann Webb Collins (2007). « It's more than just use: An exploration of telemedicine use quality », *Decision Support Systems*, vol. 43, no 4, p. 1287-1304.
- Liang, Huigang, Nilesh Saraf, Qing Hu et Yajiong Xue (2007). « Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management », *MIS quarterly*, p. 59-87.
- Limayem, Moez, Sabine Gabriele Hirt et M. K. Cheung Christy (2007). « How habit limits the predictive power of intention: The case of information systems continuance », *MIS Quarterly*, vol. 31, no 4, p. 705-737.
- Lin, Hsiu-Fen (2008). « Empirically testing innovation characteristics and organizational learning capabilities in e-business implementation success », *Internet Research*, vol. 18, no 1, p. 60-78.
- Loebbecke, Claudia et Arnold Picot (2015). « Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda », *The Journal of Strategic Information Systems*, vol. 24, no 3, p. 149-157.
- Luftman, Jerry, Barry Derksen, Rajeev Dwivedi, Martin Santana, Hossein S Zadeh et Eduardo Rigoni (2015). « Influential it management trends: An international study » [journal article], *Journal of Information Technology*, vol. 30, no 3, p. 293-305.
- Lycett, Mark (2013). « 'Datafication': Making sense of (big) data in a complex world », *European Journal of Information Systems*, vol. 22, no 4, p. 381-386.
- Malomo, Fola et Vania Sena (2017). « Data intelligence for local government? Assessing the benefits and barriers to use of big data in the public sector », *Policy & Internet*, vol. 9, no 1, p. 7-27.
- Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, *et al.* (2011). « Big data: The next frontier for innovation, competition, and productivity ».
- March, Hug et Ramon Ribera-Fumaz (2014). « Smart contradictions: The politics of making barcelona a self-sufficient city », *European Urban and Regional Studies*, vol. 23, no 4, p. 816-830.
- Markus, M. Lynne (2015). « New games, new rules, new scoreboards: The potential consequences of big data », *Journal of Information Technology*, vol. 30, no 1, p. 58-59.
- Martens, David, Foster Provost, Jessica Clark et Enric Junqué de Fortuny (2016). « Mining massive fine-grained behavior data to improve predictive analytics » [Article], *MIS Quarterly*, vol. 40, no 4, p. 869-888.
- Massetti, Brenda et Robert W Zmud (1996). « Measuring the extent of edi usage in complex organizations: Strategies and illustrative examples », *MIS quarterly*, p. 331-345.

- Mathieson, Kieran, Eileen Peacock et Wynne W Chin (2001). « Extending the technology acceptance model: The influence of perceived user resources », *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, vol. 32, no 3, p. 86-112.
- Mehmood, Rashid, Royston Meriton, Gary Graham, Patrick Hennelly et Mukesh Kumar (2017). « Exploring the influence of big data on city transport operations: A markovian approach » [Article], *International Journal of Operations & Production Management*, vol. 37, no 1, p. 75-104.
- Menon, Syam et Sumit Sarkar (2016). « Privacy and big data: Scalable approaches to sanitize large transactional databases for sharing » [Article], *MIS Quarterly*, vol. 40, no 4, p. 963-982.
- Mergel, Ines, R. Karl Rethemeyer et Kimberley Isett (2016). « Big data in public affairs », *Public Administration Review*, vol. 76, no 6, p. 928-937.
- Mishra, Abhay Nath, Prabhudev Konana et Anitesh Barua (2007). « Antecedents and consequences of internet use in procurement: An empirical investigation of us manufacturing firms », *Information Systems Research*, vol. 18, no 1, p. 103-120.
- Moges, Helen Tadesse (2014). « A contextual data quality analysis for credit risk management in financial institutions ».
- Moore, Gary C. et Izak Benbasat (1991). « Development of an instrument to measure the perceptions of adopting an information technology innovation », *Information Systems Research*, vol. 2, no 3, p. 192-222.
- Müller, Oliver, Iris Junglas, Jan Vom Brocke et Stefan Debortoli (2016). « Utilizing big data analytics for information systems research: Challenges, promises and guidelines », *European Journal of Information Systems*, vol. 25, no 4, p. 289-302.
- Nam, Taewoo et Theresa A. Pardo (2014). « The changing face of a city government: A case study of philly311 », *Government Information Quarterly*, vol. 31, Supplement 1, p. S1-S9.
- Nations Unies (2014). *Progrès accomplis dans la mise en oeuvre des textes issus de la deuxième conférence des nations unies sur les établissements humains (habitat ii) et nouveaux défis à relever dans le domaine du développement urbain durable* (n° A/CONF.226/PC.1/5). Récupéré de <http://unhabitat.org/wp-content/uploads/2014/07/Progress-to-date-outcome-Habitat-II-FRENCH.pdf>
- Nelson, R Ryan (2005). « Project retrospectives: Evaluating project success, failure, and everything in between », *MIS Quarterly Executive*, vol. 4, no 3, p. 361-372.
- Newell, Sue (2015). « Managing knowledge and managing knowledge work: What we know and what the future holds » [journal article], *Journal of Information Technology*, vol. 30, no 1, p. 1-17.
- Nunnally, Jum C. (1994). *Psychometric theory*, vol. 3th, New York, McGraw-Hill.
- O'Malley, Martin (2014). « Doing what works: Governing in the age of big data », *Public Administration Review*, vol. 74, no 5, p. 555-556.
- Ortiz de Guinea, Ana et M. Lynne Markus (2009). « Why break the habit of a lifetime? Rethinking the roles of intention, habit, and emotion in continuing information technology use », *MIS Quarterly*, vol. 33, no 3, p. 433-444.
- Ortiz de Guinea, Ana et Jane Webster (2015). « The missing links: Cultural, software, task and personal influences on computer self-efficacy », *The International Journal of Human Resource Management*, vol. 26, no 7, p. 905-931.

- Ostlund, Lyman E. (1974). « Perceived innovation attributes as predictors of innovativeness », *Journal of Consumer Research (pre-1986)*, vol. 1, no 2, p. 23.
- Pan, Yunhe, Yun Tian, Xiaolong Liu, Dedao Gu et Gang Hua (2016). « Urban big data and the development of city intelligence », *Engineering*, vol. 2, no 2, p. 171-178.
- Paré, Guy, Mary Tate, David Johnstone et Spyros Kitsiou (2016). « Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews », *European Journal of Information Systems*, vol. 25, no 6, p. 493-508.
- Paulhus, Delroy L (1991). « Measurement and control of response bias ».
- Pavlou, Paul A. et Omar A. El Sawy (2006). « From it leveraging competence to competitive advantage in turbulent environments: The case of new product development », *Information Systems Research*, vol. 17, no 3, p. 198-227.
- Pearlson, Keri et Carol S Saunders (2004). *Managing and using information systems: A strategic approach*, Wiley New York, NY.
- Picone, Marco, Michele Amoretti, Marco Martalò, Francesco Zanichelli et Gianluigi Ferrari (2015). « Combining geo-referencing and network coding for distributed large-scale information management », *Concurrency and Computation: Practice and Experience*, vol. 27, no 13, p. 3295-3315.
- Popescul, Daniela et Laura Diana Radu (2016). « Data security in smart cities: Challenges and solutions », *Informatica Economica*, vol. 20, no 1, p. 29-38.
- Premkumar, G., K. Ramamurthy et Sree Nilakanta (1994). « Implementation of electronic data interchange: An innovation diffusion perspective », *Journal of Management Information Systems*, vol. 11, no 2, p. 157-186.
- Premkumar, G. et Margaret Roberts (1999). « Adoption of new information technologies in rural small businesses », *Omega*, vol. 27, no 4, p. 467.
- Press, Gil (2013). « A very short history of big data », *Forbes Tech Magazine, May*, vol. 9.
- Price, Derek J (1961). *Science since babylon*, New Haven, CT, New Haven and London Yale University Press.
- PricewaterhouseCoopers (2013). « Capitalizing on the promise of big data ».
- Ransbotham, Sam, Robert G. Fichman, Ram Gopal et Alok Gupta (2016). « Special section introduction—ubiquitous it and digital vulnerabilities », *Information Systems Research*, vol. 27, no 4, p. 834-847.
- Rathore, M. Mazhar, Awais Ahmad, Anand Paul et Seungmin Rho (2016). « Urban planning and building smart cities based on the internet of things using big data analytics » [Article], *Computer Networks*, vol. 101, p. 63-80.
- Re Calegari, Gloria, Irene Celino et Diego Peroni (2016). « City data dating: Emerging affinities between diverse urban datasets », *Information Systems*, vol. 57, p. 223-240.
- Rogers, Everett Mitchell (1995). « Diffusion of innovations », *New York*, vol. 12.
- Rogers, Everett Mitchell (2003). *Diffusion of innovations*, vol. 5th --, New York, Free Press.
- Rowley, Jennifer (2007). « The wisdom hierarchy: Representations of the dikw hierarchy », *Journal of Information Science*, vol. 33, no 2, p. 163-180.

- Rowley, Jennifer et Frances Slack (2009). « Conceptions of wisdom », *Journal of Information Science*, vol. 35, no 1, p. 110-119.
- Saboo, Alok R., V. Kumar et Insu Park (2016). « Using big data to model time-varying effects for marketing resource (re)allocation » [Article], *MIS Quarterly*, vol. 40, no 4, p. 911-940.
- Saga, V. L. et R. W. Zmud (1994). « The nature and determinants of it acceptance, routinization, and infusion », dans, vol 45, AMSTERDAM, ELSEVIER SCIENCE PUBL B V, p. 67-86.
- Sambamurthy, Vallabh et Robert W. Zmud (1996). *Information technology and innovation: Strategies for success*, Financial Executives Research Foundation.
- Saxena, Stuti et Sujeet Kumar Sharma (2016). « Integrating big data in “e-oman”: Opportunities and challenges », *Info : the Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media*, vol. 18, no 5, p. 79-97.
- Schmitt, Neal et Daniel M Stults (1986). « Methodology review: Analysis of multitrait-multimethod matrices », *Applied Psychological Measurement*, vol. 10, no 1, p. 1-22.
- Schroeck, Michael, Rebecca Shockley, Janet Smart, Dolores Romero-Morales et Peter Tufano (2012). « Analytics: The real-world use of big data », *IBM Global Business Services*, p. 1-20.
- Seddon, Peter B. (1997). « A respecification and extension of the delone and mclean model of is success », *Information Systems Research*, vol. 8, no 3, p. 240-253.
- Shareef, Mahmud Akhter, Vinod Kumar, Uma Kumar et Yogesh K Dwivedi (2011). « E-government adoption model (gam): Differing service maturity levels », *Government Information Quarterly*, vol. 28, no 1, p. 17-35.
- Sharma, Rajeev, Sunil Mithas et Atreyi Kankanhalli (2014). « Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations », *European Journal of Information Systems*, vol. 23, no 4, p. 433-441.
- Spiekermann, Sarah et Jana Korunovska (2016). « Towards a value theory for personal data » [journal article], *Journal of Information Technology*.
- Sugimoto, Cassidy R, Hamid R Ekbia et Michael Mattioli (2016). *Big data is not a monolith*, MIT Press.
- Šuh, Jelena, Vladimir Vujin, Dušan Barać, Zorica Bogdanović et Božidar Radenković (2015). « Designing cloud infrastructure for big data in e-government » [Article], *RUO: Revija za Univerzalno Odlicnost*, vol. 4, no 1, p. A26-A38.
- Swain, Scott D, Danny Weathers et Ronald W Niedrich (2008). « Assessing three sources of misresponse to reversed likert items », *Journal of Marketing Research*, vol. 45, no 1, p. 116-131.
- Swanson, E. Burton (1994). « Information systems innovation among organizations », *Management Science*, vol. 40, no 9, p. 1069.
- Tachizawa, Elcio M., María J. Alvarez-Gil et María J. Montes-Sancho (2015). « How “smart cities” will change supply chain management », *Supply Chain Management: An International Journal*, vol. 20, no 3, p. 237-248.
- Teo, H. H., K. K. Wei et I. Benbasat (2003). « Predicting intention to adopt interorganizational linkages: An institutional perspective », *MIS Quarterly*, vol. 27, no 1, p. 19-49.

- Teo, Thompson SH, Shirish C Srivastava et Li Jiang (2008). « Trust and electronic government success: An empirical study », *Journal of management information systems*, vol. 25, no 3, p. 99-132.
- Thompson, N., R. Ravindran et S. Nicosia (2015). « Government data does not mean data governance: Lessons learned from a public sector application audit », *GOVERNMENT INFORMATION QUARTERLY*, vol. 32, no 3, p. 316-322.
- Thong, James Y. L. (1999). « An integrated model of information systems adoption in small businesses », *Journal of Management Information Systems*, vol. 15, no 4, p. 187-214.
- Titah, Ryad (2010). *Utilisation et impact du gouvernement électronique au niveau municipal*, Dissertation/Thesis.
- Titah, Ryad et Henri Barki (2006). « E-government adoption and acceptance: A literature review », *International Journal of Electronic Government Research (IJEGR)*, vol. 2, no 3, p. 23-57.
- Titah, Ryad et Henri Barki (2011). *The impact of e-government in municipalities*, vol. no 11-03, Montréal, HEC Montréal, Chaire de recherche du Canada en implantation et gestion des technologies de l'information.
- Tornatzky, L. G. et M Fleischer (1990). « The processes of technological innovation ».
- Tornatzky, L. G. et K. J. Klein (1982). « Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings », *IEEE Transactions on Engineering Management*, vol. EM-29, no 1, p. 28-45.
- van der Graaf, Shenja et Carina Veeckman (2014). « Designing for participatory governance: Assessing capabilities and toolkits in public service delivery », *info*, vol. 16, no 6, p. 74-88.
- van Loenen, Bastiaan, Stefan Kulk et Hendrik Ploeger (2016). « Data protection legislation: A very hungry caterpillar: The case of mapping data in the european union », *Government Information Quarterly*, vol. 33, no 2, p. 338-345.
- van Zoonen, Liesbet (2016). « Privacy concerns in smart cities », *Government Information Quarterly*, vol. 33, no 3, p. 472-480.
- Venkatesh, Viswanath, Michael G. Morris, Gordon B. Davis et Fred D. Davis (2003). « User acceptance of information technology: Toward a unified view », *MIS Quarterly*, vol. 27, no 3, p. 425-478.
- Washburn, Doug, Usman Sindhu, Stephanie Balaouras, Rachel A Dines, N Hayes et Lauren E Nelson (2010). « Helping cities understand “smart city” initiatives », *Growth*, vol. 17, no 2, p. 1-17.
- Webster, Jane et Richard T. Watson (2002). « Analyzing the past to prepare for the future: Writing a literature review », *MIS Quarterly*, vol. 26, no 2, p. xiii-xxiii.
- Welch, Eric W., Mary K. Feeney et Chul Hyun Park (2016). « Determinants of data sharing in u.S. City governments », *Government Information Quarterly*, vol. 33, no 3, p. 393-403.
- Wersig, Gernot et Ulrich Neveling (1975). « The phenomena of interest to information science », *The information scientist*, vol. 9, no 4, p. 127-140.
- Woerner, Stephanie L. et Barbara H. Wixom (2015). « Big data: Extending the business strategy toolbox », *Journal of Information Technology*, vol. 30, no 1, p. 60-62.
- Yang, Lee, Stuart Madnick, Richard Wang, Forea Wang et Zhang Hongyun (2014). « A cubic framework for the chief data officer: Succeeding in a world of big data » [Article], *MIS Quarterly Executive*, vol. 13, no 1, p. 1-13.

- Zeleny, Milan (1987). « Management support systems: Towards integrated knowledge management », *Human systems management*, vol. 7, no 1, p. 59-70.
- Zhang, Qingchen, Laurence T. Yang et Zhikui Chen (2016). « Privacy preserving deep computation model on cloud for big data feature learning » [Article], *IEEE Transactions on Computers*, vol. 65, no 5, p. 1351-1362.
- Zhou, Zuojian, Wanchun Dou, Guochao Jia, Chunhua Hu, Xiaolong Xu, Xiaotong Wu, *et al.* (2016). « A method for real-time trajectory monitoring to improve taxi service using gps big data », *Information & Management*, vol. 53, no 8, p. 964-977.
- Zhu, Kevin, Shutao Dong, Sean Xin Xu et Kenneth L. Kraemer (2006). « Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of european companies », *European Journal of Information Systems*, vol. 15, no 6, p. 601-616.
- Zhu, Kevin, Kenneth L Kraemer et Jason Dedrick (2004). « Information technology payoff in e-business environments: An international perspective on value creation of e-business in the financial services industry », *Journal of management information systems*, vol. 21, no 1, p. 17-54.
- Zhu, Kevin et Kenneth L. Kraemer (2005). « Post-adoption variations in usage and value of e-business by organizations: Cross-country evidence from the retail industry », *Information Systems Research*, vol. 16, no 1, p. 61-84.
- Zhu, Kevin, Kenneth L. Kraemer et Sean Xu (2003). « Electronic business adoption by european firms: A cross-country assessment of the facilitators and inhibitors », *European Journal of Information Systems*, vol. 12, no 4, p. 251-268.
- Zuboff, Shoshana (2015). « Big other: Surveillance capitalism and the prospects of an information civilization », *Journal of Information Technology*, vol. 30, no 1, p. 75-89.