

# La gouvernance des données et de l'information : élaboration d'un cadre conceptuel intégrateur

## An integrative conceptual framework for data and information governance

Jacky Akoka<sup>1</sup>, Isabelle Comyn-Wattiau<sup>2</sup>

<sup>1</sup> Laboratoire CEDRIC-CNAM, 2 Rue Conté, 75003 PARIS, [jacky.akoka@lecnam.net](mailto:jacky.akoka@lecnam.net)

<sup>2</sup> ESSEC Business School, 3 Av. B. Hirsch, 95021 CERGY Cedex, [wattiau@essec.edu](mailto:wattiau@essec.edu)

**RÉSUMÉ.** La gouvernance des données et de l'information est une activité essentielle pour les organisations qui cherchent à exploiter les données comme un atout stratégique. Son objectif est de maximiser la valeur tout en minimisant les coûts et les risques. Dans cet article, nous présentons un cadre conceptuel pour la gouvernance des données et de l'information offrant une vue holistique enrichissant les cadres et modèles académiques et professionnels existants. En utilisant des techniques bibliométriques, nous analysons la littérature existante afin d'identifier les éléments clés de la gouvernance des données et de l'information, notamment sa structure intellectuelle, ses thèmes de recherche et les articles les plus influents qui en forment la colonne vertébrale (« backbone »). Nous comparons les sphères tant communes que spécifiques de la gouvernance de la donnée et de la gouvernance de l'information. Dans un deuxième temps, nous proposons un cadre conceptuel enrichi fondé sur la théorie des systèmes. Ce cadre englobe cinq dimensions primordiales : le but, la structure, les activités, l'environnement et le résultat. Il permet également de prendre en considération l'interaction entre ces dimensions. Afin d'illustrer ce cadre conceptuel, nous décrivons comment celui-ci a permis de structurer les questions d'un baromètre dédié à l'évaluation de la maturité de la gouvernance des données et de l'information dans les organisations. Nous présentons ensuite quelques cas d'usage, puis nous discutons des implications pour les chercheurs et les praticiens.

**ABSTRACT.** Data and information governance is an important activity for organizations seeking to leverage data as a strategic asset. Its goal is to maximize value while minimizing costs and risks. In this article, we present a conceptual framework for data and information governance that offers a holistic view, enriching existing academic and professional frameworks and models. Using bibliometric techniques, we analyze the existing literature to identify the key elements of data and information governance, including its intellectual structure, research themes, and the most influential articles that form its backbone. We compare the common and specific spheres of data governance and information governance. In a second step, we propose an enriched conceptual framework based on systems theory. This framework encompasses five key dimensions: purpose, structure, activities, environment, and outcome. It also allows for consideration of the interaction between these dimensions. To illustrate this conceptual framework, we describe how it was used to structure the questions in a barometer designed to assess the maturity of data and information governance in organizations.

**MOTS-CLÉS.** Gouvernance des données, gouvernance de l'information, théorie des systèmes, cadre conceptuel, valeur, risque, coût, étude bibliométrique.

**KEYWORDS.** Data governance, information governance, systems theory, conceptual framework, value, risk, cost, bibliometric study.

## 1. Introduction

Les données sont un atout stratégique et une ressource pour les organisations, car elles conditionnent la prise de décision. La gouvernance des données permet la gestion des données de manière efficace pour améliorer la performance de l'organisation tout en respectant les obligations réglementaires et en minimisant les coûts et les risques. La gouvernance des données désigne l'ensemble des processus, politiques et outils qui permettent de déterminer qui est responsable de la prise de décision d'une organisation au sujet de ses données [KHA 10]. En cela, ils distinguent la gouvernance des données de leur management, lequel consiste à prendre et mettre en œuvre ces décisions. Smallwood la définit comme un ensemble de structures, politiques, procédures, processus et technologies utilisés pour collecter, organiser, utiliser et sécuriser les données [SMA 19]. Plusieurs revues de la littérature sur la

gouvernance des données ont été publiées [NGU 16, ABR 19, MER 19]. La gouvernance des données n'est pas simplement une question de technologie ou de management des données ; il s'agit fondamentalement d'établir les règles et les responsabilités relatives au management des données. Il manque, à ce jour, une définition acceptée par tous, tant chercheurs que praticiens. De plus, de nombreux cadres de référence (« frameworks ») ou modèles ont été proposés sans qu'aucun ne fasse consensus.

Donaldson et Walker ont lancé une initiative de gouvernance de l'information au sein de National Health Service [DON 04]. C'est la première mention de ce concept. Il est repris ensuite dans de nombreux articles, parfois comme un synonyme de gouvernance de la donnée. Quand il n'est pas synonyme, il est plutôt repris dans le monde du document et de la gestion des enregistrements (« records management »). Tallon définit la gouvernance de l'information comme ayant pour objectif de maximiser la valeur de l'information tout en minimisant les risques et les coûts [TAL 13a].

Cette cohabitation de plusieurs communautés conceptualisant deux thématiques en grande partie similaires nous a conduit à initier une étude bibliométrique afin d'identifier les articles les plus influents, la zone de recouvrement de ces deux domaines mais aussi l'ensemble des contributions afin de mieux comprendre ce concept et de fournir un cadre de référence fondé sur une théorie. Notre recherche est ainsi guidée par les questions suivantes :

QR1 : Quelle est la structure intellectuelle, les thèmes de recherche et les articles influents dans le domaine de la gouvernance des données et de l'information ?

QR2 : Quel cadre théorique peut-on proposer pour structurer les composants de ce domaine ?

Pour répondre à ces questions, nous effectuons une analyse rétrospective complète de la littérature relative à la gouvernance des données et de l'information à partir de la base de données Scopus (QR1). Pour ce faire, trois méthodes d'analyse bibliométrique quantitative sont employées : l'analyse des co-citations (CCA), l'analyse du couplage bibliographique (BCA) et l'analyse des chemins principaux (MPA). Grâce à cette étude bibliométrique, nous identifions les dimensions les plus saillantes et les lacunes, ce qui nous permet de développer un nouveau cadre conceptuel (QR2). Notre cadre n'est pas destiné à remplacer les cadres existants, mais plutôt à fournir une carte conceptuelle qui montre comment les différents éléments s'assemblent pour définir la gouvernance des données et de l'information d'une manière multidimensionnelle.

Dans les deux sections qui suivent, nous présentons un état de l'art structuré autour d'une étude bibliométrique. D'abord le périmètre de la gouvernance des données est étudié en détail. Ensuite, le périmètre de la gouvernance de l'information et celui des deux domaines réunis sont étudiés et comparés au premier. Alimenté par les principaux résultats de l'étude bibliométrique, le cadre conceptuel fondé sur la théorie des systèmes est décrit dans la section 4. Dans la section 5, nous illustrons l'utilisation de ce cadre pour générer la structure d'un baromètre dédié à l'étude de la maturité des organisations en gouvernance des données et de l'information et décrivons quelques cas d'usages. La section 6 conclut et suggère des voies de recherche future.

## 2. Etude bibliométrique de la gouvernance des données

L'objet principal de cette section est de dresser un état de l'art de la littérature scientifique sur la gouvernance des données et de l'information. Au-delà, nous visons à cartographier cette littérature à l'aide de techniques bibliométriques. A cette fin, nous utilisons trois techniques d'analyse des citations, après une étude descriptive du domaine. Dans la suite, nous présentons le dispositif méthodologique fondé sur trois étapes : la constitution du jeu de données, l'analyse descriptive et l'analyse par les citations.

## 2.1. Constitution du jeu de données

La constitution du jeu de données requiert le choix d'une base de données bibliographique, la construction d'une requête, l'extraction des résultats et le nettoyage des données. Dans cette recherche, nous avons soumis à la base de données bibliographiques Scopus, réputée pour sa pertinence dans le domaine de l'informatique et des systèmes d'information, la requête ci-dessous :

```
TITLE-ABS-KEY("data governance") AND (EXCLUDE( DOCTYPE,"ch") OR EXCLUDE ( DOCTYPE,"bk" ) OR EXCLUDE ( DOCTYPE,"no" ) OR EXCLUDE ( DOCTYPE,"ed" ) OR EXCLUDE ( DOCTYPE,"sh" ) OR EXCLUDE ( DOCTYPE,"er" ) OR EXCLUDE ( DOCTYPE,"le" ) OR EXCLUDE ( DOCTYPE,"tb" ) ) AND ( LIMIT-TO ( LANGUAGE,"English" ) )
```

La chaîne de recherche choisie est « data governance » limitée aux titres, résumés et mots-clés des publications en anglais dans des revues ou des conférences, ce qui nous retourne 3449 documents. Certains auteurs utilisent, de manière interchangeable, l'expression « information governance ». L'ajout de cette chaîne dans la requête n'a pas apporté de modifications significatives pertinentes. Après exportation des données, à l'aide du logiciel CRExplorer [THO 21], nous avons procédé à la déduplication des références. En effet, Scopus ne produit pas de documents en double mais, dans ces documents, de nombreuses références sont mal codées, rendant difficile leur rapprochement par les logiciels d'analyse bibliométrique. CRExplorer dispose d'un outil de rapprochement et d'harmonisation de ces références. Un nettoyage « manuel » a aussi été nécessaire pour éliminer quelques références incomplètes qui perturbaient les analyses, en l'occurrence une référence, ce qui nous a conduit à un jeu de données de 3448 articles.

## 2.2. Analyse descriptive

Avant d'utiliser les techniques d'analyse des citations, nous avons procédé à une analyse descriptive du jeu de données résultant de l'étape précédente. La distribution des articles obtenus est représentée à la figure 1. Le terme « gouvernance des données » apparaît de façon erronée dans Scopus en 1992 dans un article de la revue IEEE Security & Privacy, alors qu'il est en réalité publié en 2010 [BRU 10]. Les premiers articles véritables datent de 2003 et le phénomène reste embryonnaire jusqu'en 2011 environ.

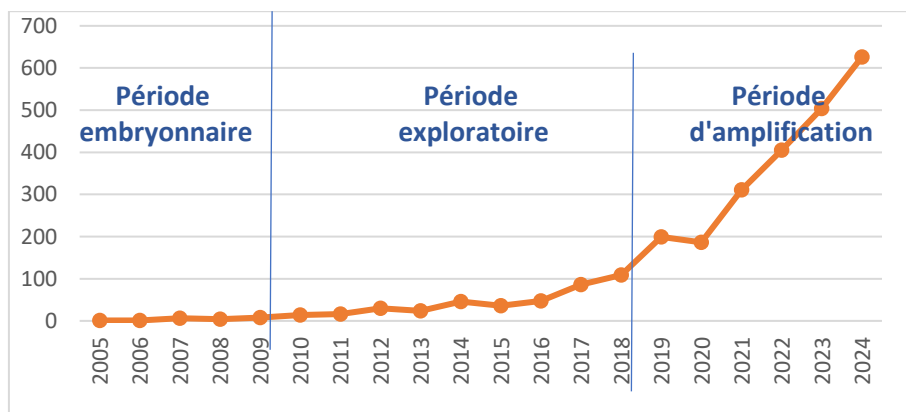
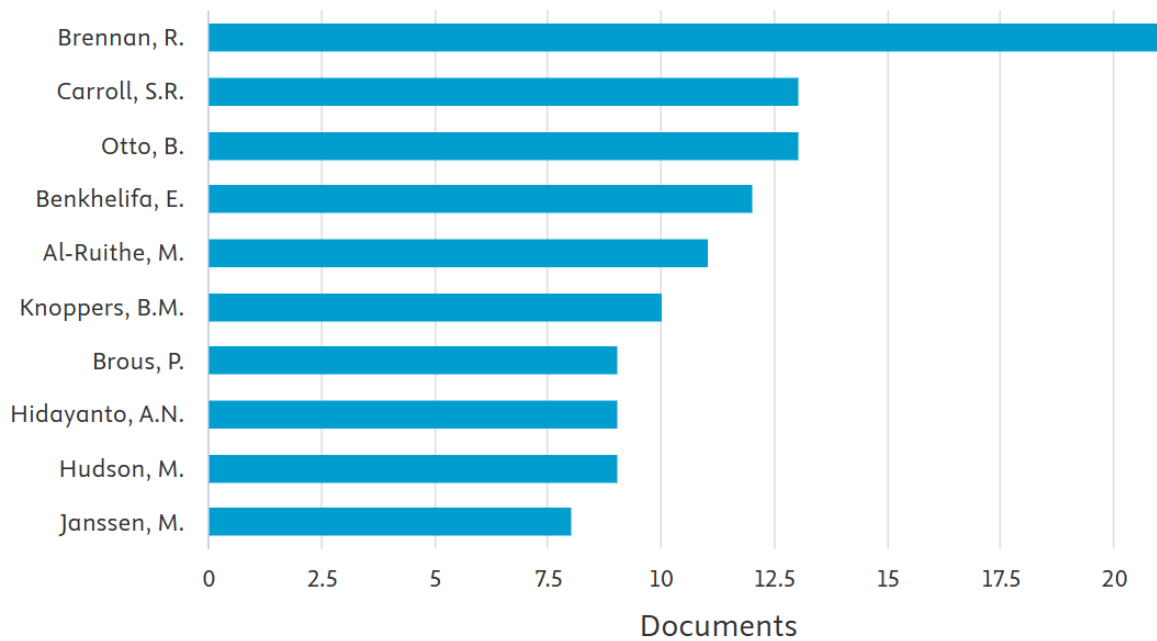


Figure 1. Nombre de publications par an.

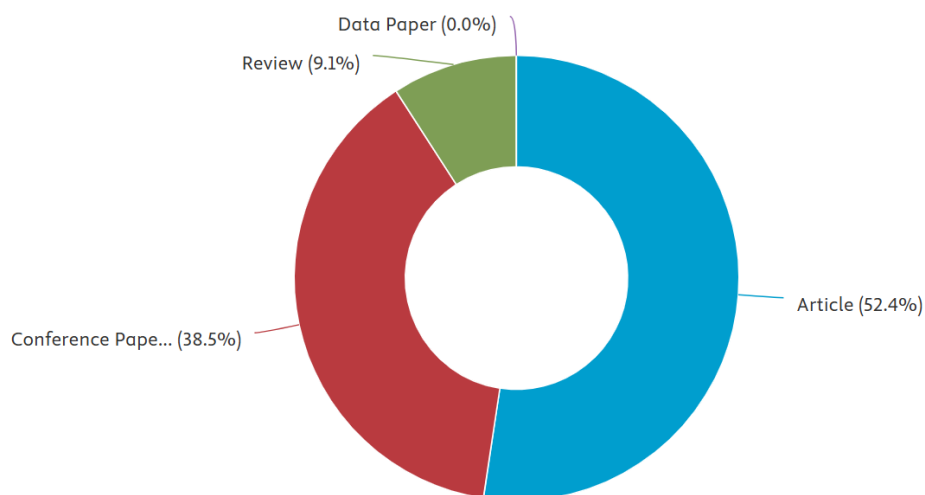
La période 2011 à 2018, qu'on peut qualifier d'exploratoire, révèle une augmentation régulière passant de 16 publications en 2011 à 109 en 2018. La période de 2018 à 2024, ou période d'amplification, voit le nombre de publications augmenter beaucoup plus rapidement jusqu'à atteindre 626 en la seule année 2024. Le terme de gouvernance de données est ainsi adopté progressivement. Le domaine de la gouvernance des données fait ainsi l'objet d'une recherche plus abondante. Scopus fournit un ensemble de tableaux et de graphiques décrivant certains traits saillants du jeu de données renvoyé par la requête. Nous avons sélectionné les plus pertinents. Ainsi, en ce qui concerne les auteurs les plus prolifiques (Fig.

2), Rob Brennan<sup>1</sup> se détache avec plus de 20 publications centrées autour de la valeur de la donnée, sa qualité et la mesure de sa valeur [BRE 18]. Les auteurs ne sont sélectionnés ici que sur leur nombre de publications dans le jeu de données, c'est-à-dire dans les articles « étiquetés » gouvernance de données.



**Figure 2.** Auteurs les plus prolifiques.

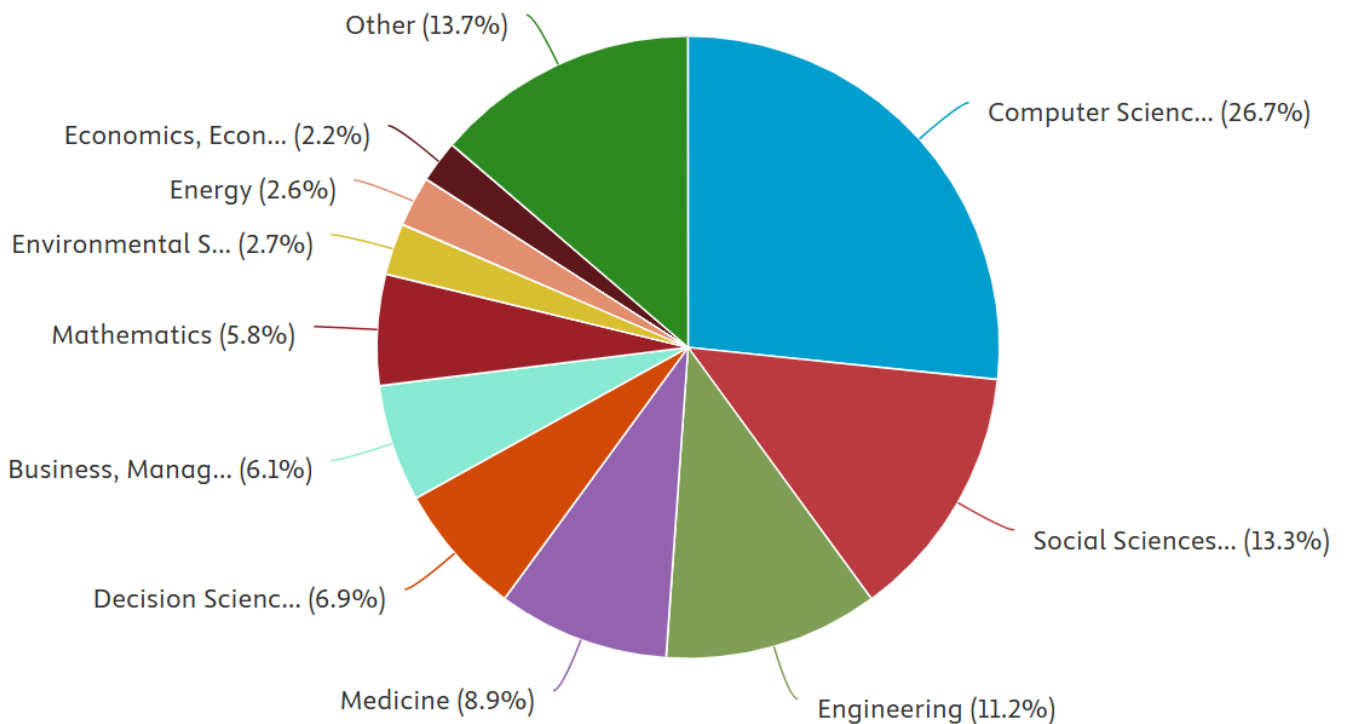
La distribution des documents par type montre deux pôles assez équilibrés entre les articles publiés dans des revues et ceux parus dans les actes de conférence (Fig. 3).



**Figure 3.** Publications par type

La recherche relative à la gouvernance des données est multidisciplinaire, avec toutefois près de 30% pour l'informatique (« computer science ») (Fig. 4). A noter qu'un domaine d'application, la médecine, émerge nettement (près de 9% des publications y sont rattachées).

<sup>1</sup> Rob Brennan n'apparaît cependant dans les graphes parce que seul le nom du premier auteur apparaît.



**Figure 4.** Documents par discipline.

En résumé, cette analyse permet d'identifier une recherche en croissance en gouvernance des données, avec quelques auteurs très présents et des sujets comme la qualité de la donnée, sa confidentialité, sa valeur et donc le besoin de gouvernance.

### 2.3. Analyse par les citations

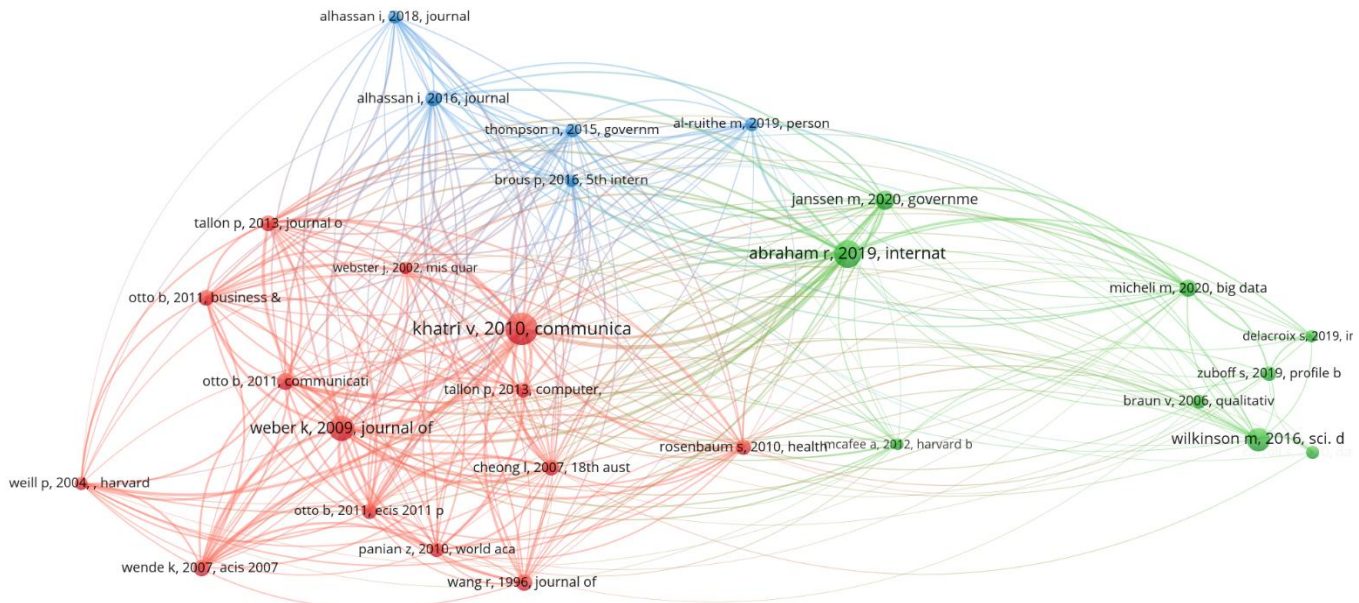
Les publications extraites sur la gouvernance des données sont trop nombreuses pour en faire une analyse individuelle. Dès lors, le recours aux techniques bibliométriques est judicieux. Dans cette section, nous présentons l'application des techniques de co-citation (« co-citation analysis ») pour identifier la structure intellectuelle du domaine. Puis le couplage bibliographique (« bibliographic coupling analysis ») nous permet de repérer les thèmes principaux. Enfin, au moyen du calcul de chemins principaux (« main path analysis »), nous mettons en évidence les articles les plus influents et les liens de citation qui les relient.

#### 2.3.1. L'analyse des co-citations (CCA)

L'analyse des co-citations consiste à construire des classes (clusters) d'articles (ou références) cités par un ensemble d'articles citants. Plus deux références sont citées ensemble, plus elles sont proches et susceptibles d'appartenir au même cluster. Le logiciel utilisé ici est Vosviewer (pour Visualization of Similarities) qui met en œuvre une similarité appelée force d'association (association strength) pour répartir les articles dans le plan [VAN 13]. A l'aide de Vosviewer, nous avons catégorisé trois domaines de références regroupant les 28 articles cités plus de 30 fois dans notre ensemble de publications (Fig. 5). Le choix du seuil de 30 est empirique et dicté par le souci de ne retenir que des articles influents et d'obtenir une visualisation de taille raisonnable.

La première catégorie (14 articles en rouge) regroupe les publications proposant des modèles et des cadres conceptuels (« frameworks ») sur le thème de l'organisation, les rôles, les responsabilités et la prise de décision relatifs à la gouvernance des données. Ainsi, Khatri et Brown [KHA 10] différencient

la gouvernance des données de son management et proposent un cadre distinguant les cinq domaines de décision relative à la gouvernance des données : les principes qui gouvernent l'usage des données, la qualité des données, les méta-données, l'accès aux données et le cycle de vie de la donnée. L'article [WEB 09] introduit un modèle matriciel de gouvernance de la donnée qui croise les rôles et les activités. Un autre article séminal décrit une étude empirique de la gouvernance de l'information concernant trente organisations [TAL 13]. Ils en déduisent un modèle composé d'antécédents, facilitateurs ou inhibiteurs, des mécanismes structurels, procéduraux ou relationnels de la gouvernance de l'information avec des résultats en termes de performance et de mitigation des risques. A noter qu'il s'agit de la gouvernance de l'information et non de la gouvernance des données, bien que ces deux termes soient souvent utilisés de manière interchangeable.



**Figure 5.** Analyse des co-citations.

La seconde catégorie (9 articles en vert) regroupe les publications proposant des cadres conceptuels de gouvernance de données intégrant davantage de dimensions. Ces articles paraissent dans des revues relevant de domaines très différents (informatique, juridique, politique, psychologique, managérial). Le plus cité est [ABR 19] qui étudie 145 sources afin d'identifier les composants de la gouvernance des données selon six dimensions : le domaine concerné (la qualité des données, la sécurité, les méta-données, raffinant les cinq domaines de [KHA 10]), le périmètre organisationnel concerné (un département, toute l'organisation, l'inter-organisationnel), le type de données concerné (traditionnelles ou massives), les mécanismes de gouvernance (structurels, procéduraux ou relationnels), les antécédents tant internes qu'externes et les conséquences (performances et risques). Pour les données scientifiques, Wilkinson et al. proposent les quatre principes FAIR (pour facile à trouver, accessible, interopérable et réutilisable) pour un bon management des données et une bonne intendance (stewardship) [WIL 16]. Janssen et al. en proposent treize à appliquer aux systèmes algorithmiques des big data, dont beaucoup reprennent les règles relatives aux données personnelles (par exemple la minimisation de l'accès aux données ou encore la nécessité d'informer) [JAN 20].

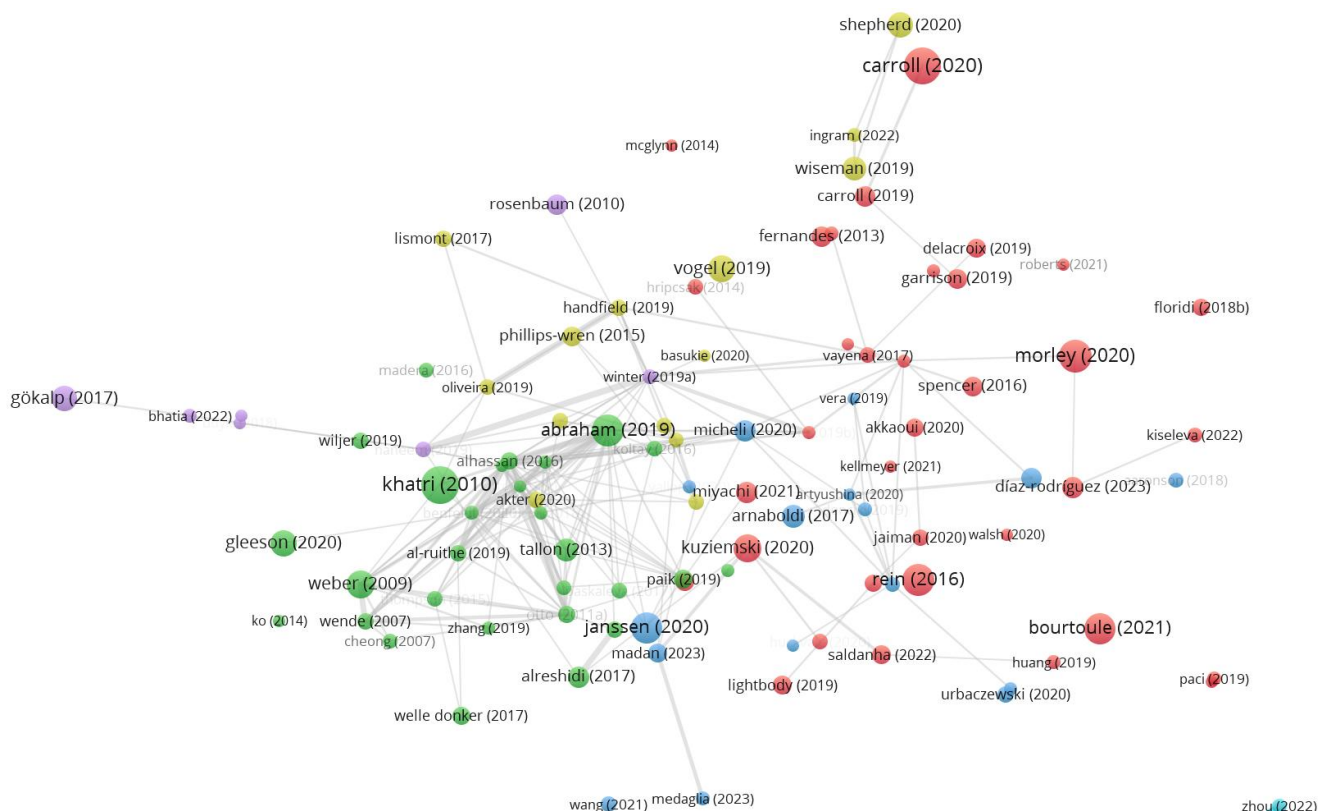
La troisième catégorie (5 articles en bleu) est dominée par des analyses de la littérature naissante sur la gouvernance de données. Ainsi, Alhassan et al. [ALH 16] proposent une analyse de fréquence des activités de gouvernance de données selon l'action (définir, mettre en œuvre et surveiller) effectuée dans un espace (standards liés aux données, exigences liées aux données, stratégie en matière de données, politique des données, etc.) pour un des cinq domaines définis dans [KHA 10]. Al-Ruithe et al. présentent une analyse systématique de la littérature qui compare la gouvernance des données en général à celle liée au cloud [ALR 19]. Six dimensions sont ainsi comparées : la politique de gouvernance de la donnée,

son administration, sa structure organisationnelle, la dimension technologique, la dimension réglementaire et ses outils de mesure et de surveillance.

En conclusion, cette analyse nous a permis d'identifier la structure intellectuelle de la gouvernance des données, une première composante principalement orientée vers la dimension organisationnelle du sujet, une seconde élargissant la dimension à des cadres conceptuels plus complets et l'appliquant à des domaines variés, la troisième composante s'attachant à l'étude de la littérature naissante sur le sujet.

### 2.3.2. L'analyse des couplages bibliographiques (BCA).

L'analyse des couplages bibliographiques consiste à construire des classes d'articles considérés proches parce qu'ils citent un même ensemble de références. Plus cet ensemble commun est grand, plus ils sont susceptibles d'appartenir au même cluster. Elle permet de rapprocher les articles travaillant sur des thématiques proches. D'autre part, elle complète l'analyse des co-citations puisqu'elle permet d'intégrer aussi des articles récents, lesquels ne sont pas encore assez cités pour apparaître dans l'analyse des co-citations.



**Figure 6.** Analyse des couplages bibliographiques

L'analyse des couplages bibliographiques produite par Vosviewer catégorise en 6 classes les 128 articles cités plus de 50 fois et connectés (Fig. 6). Le seuil de 50 est défini de façon empirique pour retenir les articles les plus influents. Vosviewer permet aussi de choisir le nombre de classes souhaité en modifiant la valeur de la résolution. Nous avons effectué ce calibrage en examinant, par essais et erreurs, le contenu des classes et surtout leur interprétabilité.

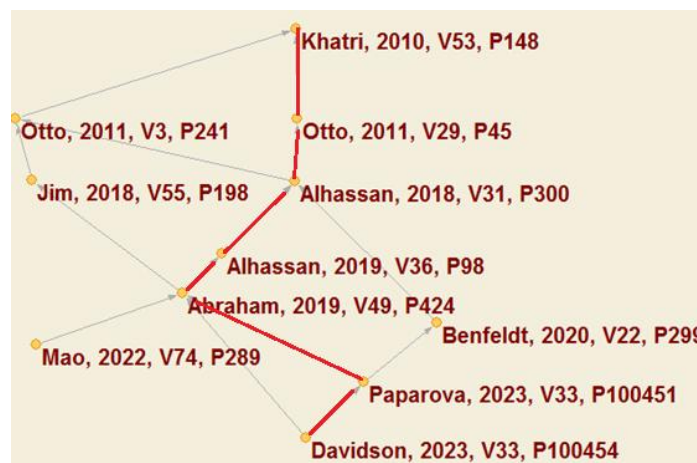
La première classe (35 nœuds en rouge) regroupe des articles étudiant les problématiques de gouvernance de données dans le domaine de l'apprentissage automatique, de la santé et de la « blockchain ». La deuxième classe (28 nœuds en vert) contient de nombreux articles déjà mentionnés dans l'analyse des co-citations. Elle regroupe donc les articles fondateurs sur la gouvernance de données, notamment les différents cadres conceptuels. La troisième classe (16 nœuds en bleu) traite de la gouvernance de données appliquée à différents domaines dont les « smart cities », l'aspect juridique, les plates-formes de données, le « big data ». La quatrième classe (14 nœuds en jaune) est aussi caractérisée

par le big data mais aussi d'autres domaines d'application, comme l'agriculture. La cinquième classe (7 nœuds en violet) cible le domaine de l'industrie 4.0 et celui des « smart cities ». Enfin, la sixième classe (2 nœuds en bleu clair) porte sur l'interdépendance entre bases de données et intelligence artificielle. A noter que le thème des « smart cities » commun aux classes 3 et 5 est traité sur un plan organisationnel dans la première et plus technologique dans la seconde. En conclusion, on a une seule classe (la deuxième) qui concentre les articles sur les fondamentaux de la gouvernance de données tandis que les autres les appliquent à différents domaines.

Ainsi, deux thèmes principaux émergent. Le premier est lié aux fondements de la gouvernance de données. Le second, regroupant le reste des classes, est dédié aux applications dans différents domaines.

### 2.3.3. L'analyse des chemins principaux (MPA).

L'analyse des chemins principaux s'effectue sur le graphe de citations entre les articles du jeu de données. Elle consiste à rechercher les trajectoires principales de ce graphe. La métrique utilisée ici est SPLC (Search Path Link Count) qui compte le nombre de fois où un arc est traversé si l'on parcourt tous les chemins possibles depuis tous les ancêtres du nœud d'origine vers tous les puits du graphe. A l'aide de cette métrique, on peut ensuite calculer les chemins les plus influents au moyen de différents algorithmes. Le logiciel Pajek propose plusieurs algorithmes. Celui utilisé ici est le global key-route (ici 10 key-routes) qui consiste à retenir les dix arcs ayant la valeur de SPLC la plus importante et, pour chacun de ces arcs, à calculer les chemins passant par cet arc.



**Figure 7.** Analyse des chemins principaux de la gouvernance des données.

On obtient le graphe de la figure 7 qui représente une vision synthétique de la structuration du domaine de la gouvernance des données et les différents canaux de diffusion de la connaissance. Parmi ces chemins, le plus impactant est en rouge sur la figure. Le nœud racine est l'article [KHA 10] déjà mentionné dans la CCA. Otto compare l'organisation de la gouvernance de données dans deux entreprises de télécommunications selon trois dimensions (le but, la structure et les processus) [OTT 10]. Alhassan *et al.* [ALH 18] comparent la littérature scientifique et la presse professionnelle selon un cadre qui croise les cinq domaines de décision définis dans [KHA 10] et les trois actions Définir, Mettre en œuvre et Surveiller. Les mêmes auteurs identifient aussi les facteurs clés de succès de la gouvernance des données au travers de la littérature [ALH 19]. Abraham *et al.* proposent un cadre de référence de la gouvernance des données qui combine trois dimensions (types de données, niveau organisationnel, processus de mise en œuvre) pour analyser les mécanismes de gouvernance, leurs antécédents et les résultats de la gouvernance des données [ABR 19]. Paparova *et al.* introduisent le concept d'espace de gouvernance de données qui combine une logique verticale et une logique horizontale en matière de rôles et responsabilités [PAP 23]. Cet article fait partie d'un numéro spécial sur la gouvernance des données dont l'éditorial porte sur les défis de la gouvernance des données à l'ère

de l'innovation digitale [DAV 23]. Ce chemin principal constitue une sorte de colonne vertébrale (« backbone ») de diffusion de la connaissance du domaine naissant de la gouvernance de données.

### 3. Etude bibliométrique de la gouvernance de l'information.

Les deux termes « gouvernance des données » et « gouvernance de l'information » sont souvent utilisés de manière interchangeable. Cependant, ils ne recouvrent pas exactement la même réalité, ce qui a conduit l'organisme ISO à proposer deux normes différentes. Le tableau 1 résume les traits saillants qui les définissent. ISO place la gouvernance de l'information à un niveau plus stratégique, tant dans la définition que dans les acteurs ou le rattachement hiérarchique. Nos observations plus récentes semblent indiquer que, de plus en plus, les organisations tendent à gérer de façon plus unifiée les deux aspects.

	<b>Gouvernance de la donnée</b>	<b>Gouvernance de l'information</b>
<b>Source</b>	ISO38505	ISO24143
<b>Définition</b>	Principes, responsabilités, structures et processus utilisés pour évaluer, diriger et surveiller l'utilisation des données dans une organisation	Cadre stratégique pour la gestion transverse des actifs informationnels de l'ensemble d'une organisation
<b>Fait partie de</b>	Gouvernance des systèmes d'information	Gouvernance d'entreprise
<b>Objectif</b>	S'assurer que les données sont utilisées efficacement, de manière efficiente et acceptable, tout en gérant les risques, en garantissant la conformité et en créant de la valeur pour les parties prenantes	Soutenir les résultats opérationnels et garantir que les risques relatifs à ces informations, et par conséquent les capacités opérationnelles et l'intégrité de l'organisation, sont correctement identifiés et gérés
<b>Acteurs</b>	Direction des systèmes d'information	Direction stratégique ou direction générale
<b>Exemple</b>	Politique de sécurité des données	Politique de conservation et de destruction de l'information

**Tableau 1.** *Gouvernance de la donnée vs. gouvernance de l'information (selon ISO)*

Bien que beaucoup d'auteurs utilisent de façon interchangeable les termes de gouvernance des données et de gouvernance de l'information, il reste certaines spécificités à chacune des deux qu'il nous a semblé utile d'étudier. Nous avons donc constitué deux jeux de données supplémentaires, le premier est propre à la gouvernance de l'information. Quant au second, il résulte de la fusion des deux jeux de données initiaux. Le tableau 2 résume les caractéristiques des trois jeux de données ainsi constitués et les paramétrages et résultats principaux des analyses opérées sur chacun. Le deuxième jeu de données relatif à la gouvernance de l'information résulte de la transformation de la première requête (« data governance ») dans laquelle la chaîne de recherche est « information governance » et conduit à un jeu de 683 articles. Ainsi le terme « gouvernance de l'information » est beaucoup moins utilisé que celui de « gouvernance des données ». Parfois les deux sont utilisés de façon interchangeable. C'est la raison pour laquelle il existe quelques articles communs aux deux jeux de données.

La fusion et le dédoublonnage des deux jeux d'articles conduisent à un ensemble de 4067 articles qui contiennent ainsi la chaîne « data governance » ou la chaîne « information governance » ou les deux dans le titre, le résumé ou les mots-clés. Dans les deux jeux de données, les premiers articles datent de 2003.

	Gouvernance des données	Gouvernance de l'information	Gouvernance des données et de l'information
Taille	3448	683	4067
Période	2003-2025	2003-2025	2003-2025
<b>Analyse descriptive</b>			
Auteurs les plus présents (nb d'articles)	Brennan R. (20), Otto B. (13), Carroll S.R. (12), Al-Ruithe M. (11), Benkhelifa E. (11)	Ford D.V. (7), Jones K.H. (7), Lyons R.A. (7), Sigmon P.W. (7), Washington L. (7)	Brennan R. (20), Otto B. (13), Carroll S.R. (12), Al-Ruithe M. (11), Benkhelifa E. (11)
Disciplines les plus représentées	Informatique, Sciences sociales, Ingénierie	Médecine, Informatique, Sciences sociales	Informatique, Sciences sociales, Ingénierie (ex aequo avec Médecine)
<b>Analyse des co-citations</b>			
Nb articles et nb citations	28 articles cités plus de 30 fois	15 articles cités plus de 10 fois	32 articles cités plus de 30 fois
Résolution	1	0,8	1
Nb clusters	3	2	3
<b>Analyse des couplages bibliographiques</b>			
Nb articles et nb citations	128 articles cités plus de 50 fois	55 articles cités plus de 20 fois	114 articles cités plus de 50 fois
Résolution	0,5	0,4	0,4
Nb clusters	6	6	6
<b>Analyse des chemins principaux</b>			
Métrique et algorithme	SPLC, global key-route (10 key-routes)		
Nb nœuds résultants	11	11	18

**Tableau 2.** Comparaison des trois jeux de données

Les auteurs les plus prolifiques en gouvernance de l'information ne sont pas les mêmes que ceux mentionnés plus haut pour la gouvernance des données, en raison de la taille relative des deux jeux d'articles. Les cinq premiers ont 7 articles sur ce thème recensés dans Scopus à comparer aux 11 à 20 articles des auteurs les plus prolifiques en gouvernance des données.

Scopus rattache les publications à des domaines disciplinaires. Ainsi, les trois domaines les plus représentés dans les publications relatives à la gouvernance des données sont respectivement l'informatique, les sciences sociales et l'ingénierie. Pour la gouvernance de l'information, on obtient dans l'ordre la médecine, l'informatique et les sciences sociales. Quand on fusionne les deux ensembles, c'est l'informatique qui est la première discipline présente suivie par les sciences sociales. L'ingénierie et la médecine sont troisièmes ex aequo. La présence du domaine de la médecine pour la gouvernance de l'information reflète l'intérêt de communautés différentes pour cette problématique dans son acception plus organisationnelle qui inclut notamment les aspects liés à la sécurité et à la réglementation.

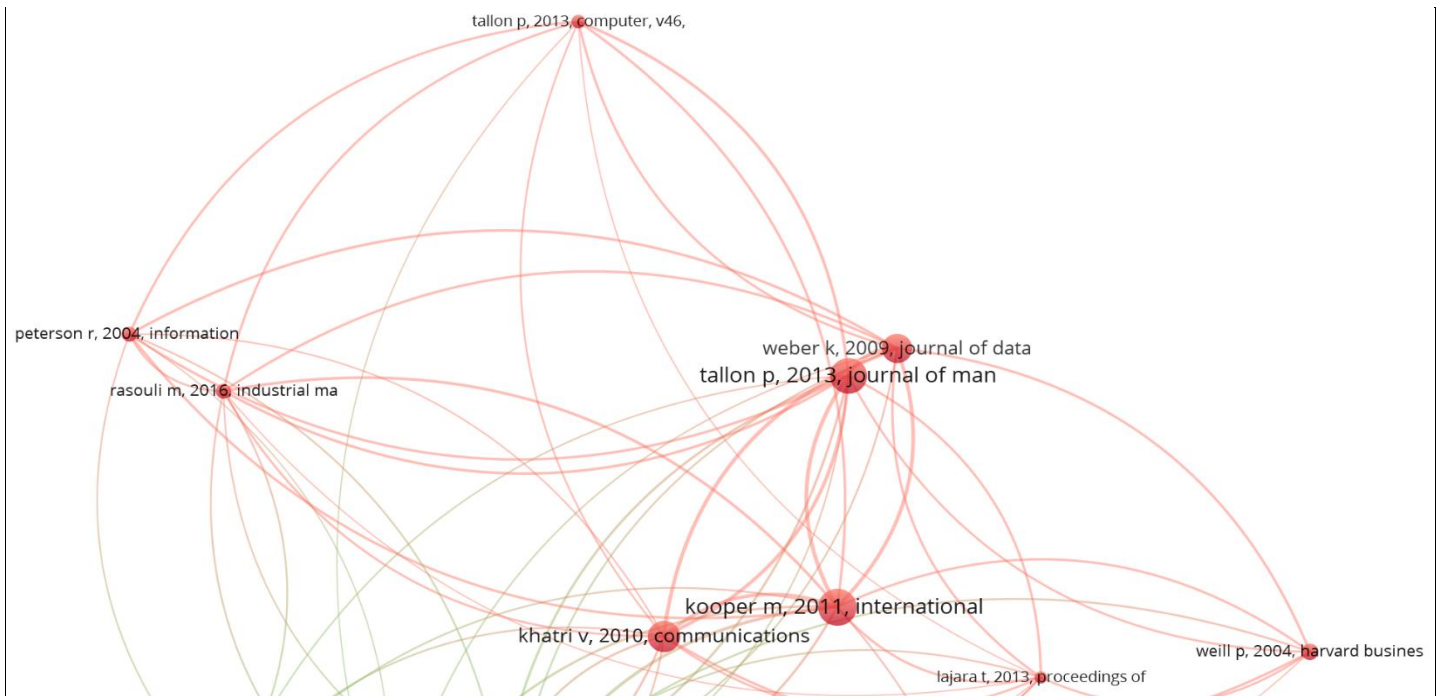
La suite du tableau 2 synthétise les paramétrages (nombre de citations, résolution, algorithmes, métriques) des analyses par les citations effectuées sur les trois jeux de données. Ces paramétrages prennent en compte la taille relative des trois jeux de données. Pour rendre cohérente la comparaison des

trois analyses des chemins principaux, le même algorithme (global search avec 10 key-routes) et la même métrique (SPLC) ont été retenus.

L'analyse des co-citations conduit à des clusters très similaires dans les trois cas (Tableau 3). Ainsi, il existe un cluster décrivant des cadres conceptuels et organisationnels dans les trois cas. Celui issu de la requête « Gouvernance de l'information » (Figure 8) contient 5 nœuds communs [KHA 10, WEB 09, TAL 13, TAL 13a, WEI 04] avec le cluster comparable dans le jeu d'articles relatif à la gouvernance des données. Ce recouvrement important traduit le caractère souvent interchangeable des deux concepts de gouvernance des données et gouvernance de l'information.

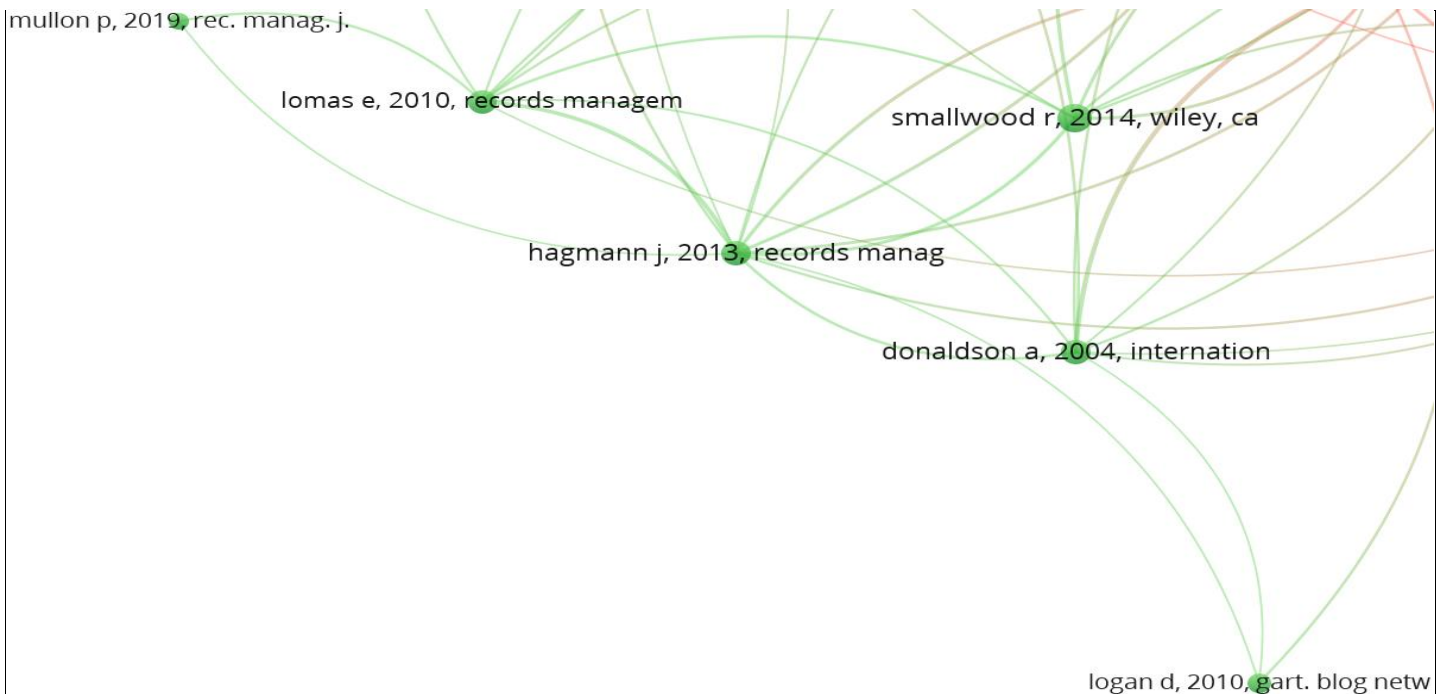
	<b>Gouvernance de l'information</b>	<b>Gouvernance des données</b>	<b>Gouvernance des données et de l'information</b>
<b>Clusters CCA</b>	Cadres conceptuels et organisationnels		
		Cadres conceptuels élargis et applications	
		Revue et analyses de la littérature	
	Etudes de cas et bonnes pratiques de gouvernance de l'information et gestion des enregistrements		
<b>Clusters BCA</b>	Fondements gouvernance information et big data	Fondements et cadres théoriques	Fondements et cadres théoriques et applications
	5 clusters dédiés aux applications en santé, blockchain et construction	4 clusters dédiés aux applications en IA, santé et blockchain, smart cities, droit et plates-formes de données, big data, analytique et agriculture, industrie 4.0	4 clusters dédiés aux applications en santé, intelligence artificielle et secteur public, agriculture et smart cities et bâtiment
		Interdépendance bases de données et IA	
<b>Chemins MPA</b>	Figure 10	Figure 7	Figure 11

**Tableau 3.** Comparaison des analyses bibliométriques



**Figure 8.** CCA *gouvernance de l'information* - cluster « *Cadres conceptuels et organisationnels* ».

En *gouvernance de l'information*, il n'y a pas de cluster qui peut être comparé à celui des cadres conceptuels élargis et applications, décelé dans l'analyse par les co-citations en *gouvernance des données*. En revanche le second cluster identifié comme caractérisant la structure intellectuelle de la *gouvernance de l'information* (Figure 9) contient des études de cas et bonnes pratiques. Il est à remarquer que 3 des six articles du cluster ont été publiés dans la revue *Records Management Journal* qui combine les contributions professionnelles autant que celles des chercheurs du monde du document.

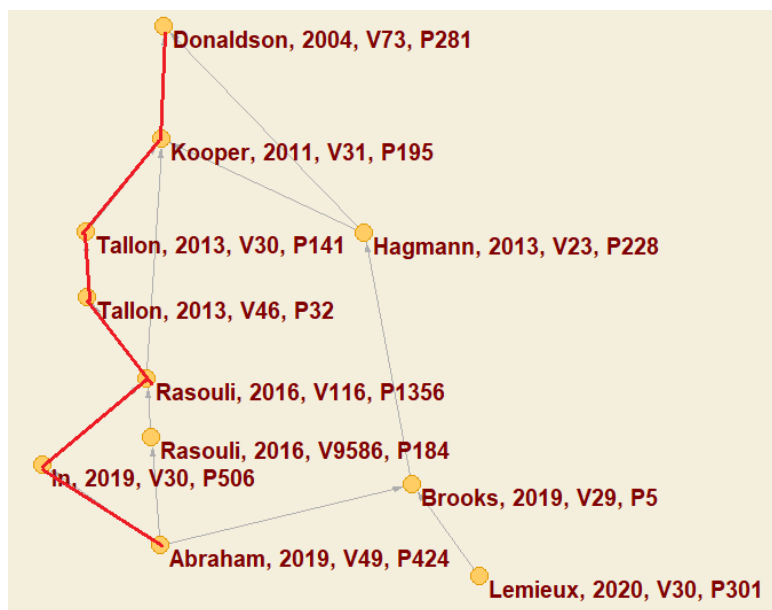


**Figure 9.** CCA *gouvernance de l'information* – cluster « *Etudes de cas et bonnes pratiques* ».

L'analyse des couplages bibliographiques fait ressortir 6 clusters dans les trois jeux d'articles. Dans les trois cas, on retrouve un cluster dédié aux fondements et cadres théoriques. Dans le jeu d'articles fusionnés (*gouvernance des données* et de *l'information*), on retrouve le cluster « *interdépendance bases*

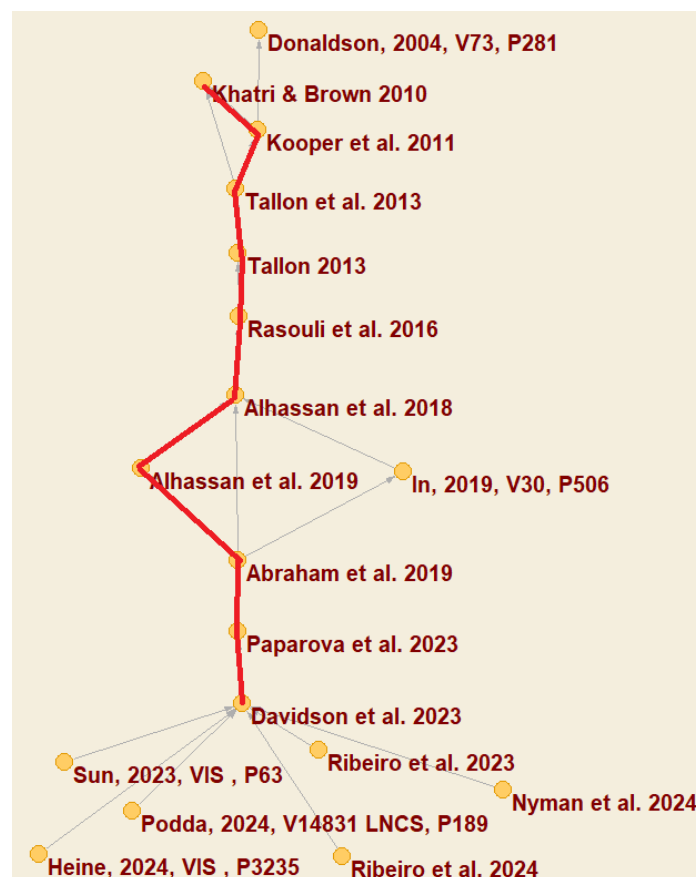
de données et IA » déjà remarqué dans l'analyse relative à la seule gouvernance des données. Les autres clusters, dans tous les jeux d'articles, sont dédiés à des applications dans différents domaines plus ou moins représentés (santé, smart cities, construction, etc.).

La figure 10 représente le résultat du MPA pour le domaine de la gouvernance de l'information. Il contient 11 nœuds. Au vu du nombre de citations, le chemin en rouge de Donaldson [DON 04] à Abraham [ABR 19] en passant par Kooper [KOO 11] et In [IN 19] est celui à plus grand impact. Centré sur les articles mentionnant la gouvernance de l'information, au sein d'un jeu d'articles beaucoup plus restreint que celui relatif à la gouvernance des données, il fait apparaître des contributions qui couvrent les deux thèmes (données et information) comme dans Tallon, 2013, V30 [TAL 13] (qui considère les termes donnée et information comme synonymes) et d'autres qui sont spécifiques au sujet de la gouvernance de l'information dans son acception plus restreinte au monde du document, comme Donaldson [DON 04] et Kooper [KOO 11].



**Figure 10.** Analyse des chemins principaux – Gouvernance de l'information.

La figure 11 représente le MPA du jeu d'articles fusionnés. Le chemin rouge relie les articles les plus cités dans ce graphe. Nous l'avons finalisé en 2023 pour éviter d'y intégrer des articles trop récents pour qu'on puisse en mesurer l'impact réel. On retrouve, dans ce chemin rouge, les articles présents dans les deux « backbones » précédents. L'analyse du graphe montre la prédominance des articles mentionnant la gouvernance de l'information avant 2018 (à l'exception de Khatri et Brown 2010) supplantés ensuite par ceux mentionnant la gouvernance des données.



**Figure 11** : Analyse des chemins principaux – Gouvernance des données et de l'information.

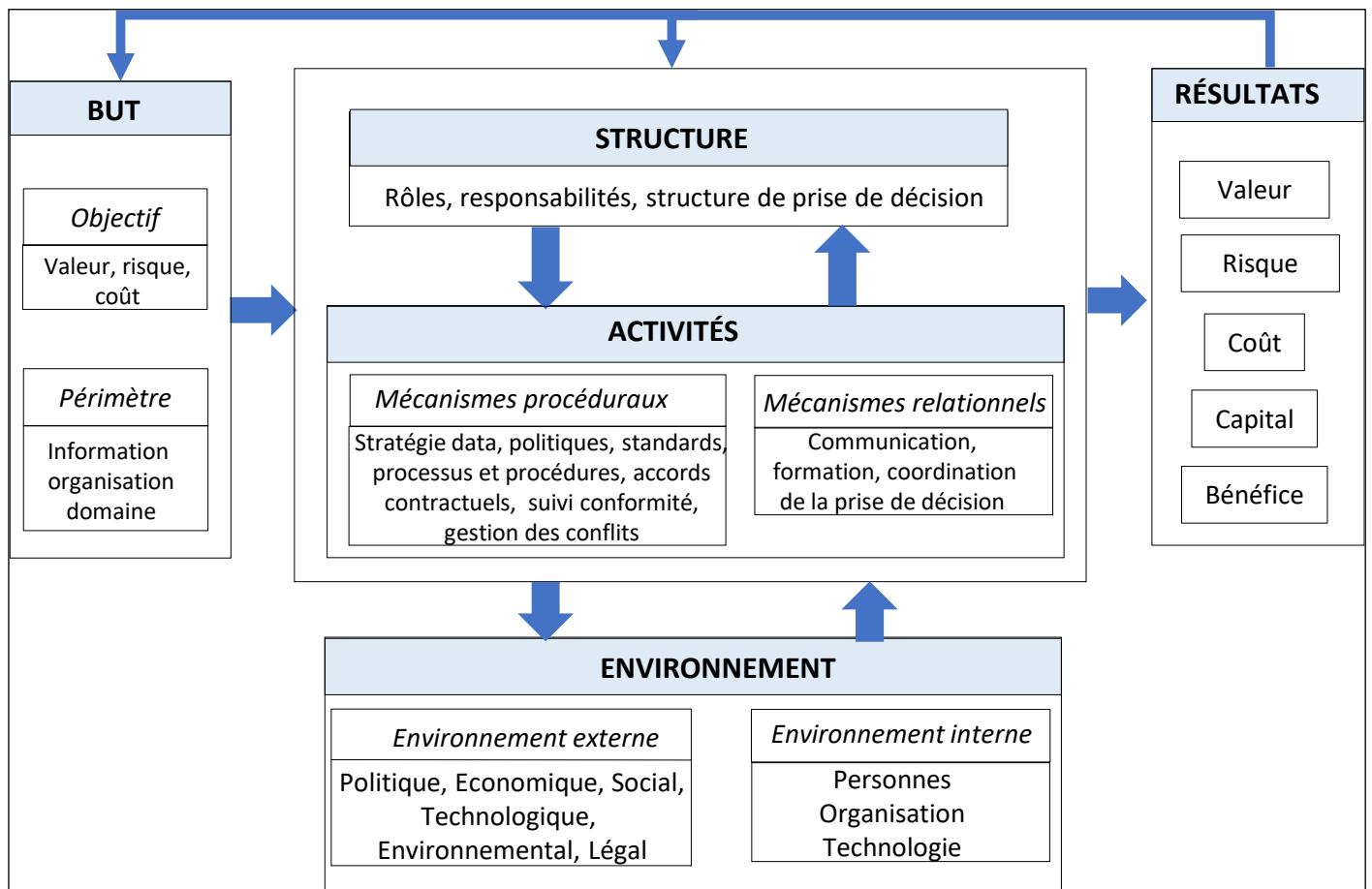
Si l'on compare les trois MPA (Figures 7, 10 et 11), le seul nœud commun est Abraham et al. 2019, qui semble ainsi faire converger les deux domaines. Par son nombre de citations, cet article rassemble une revue de littérature très large et en déduit un cadre conceptuel qui réunit les objectifs de gouvernance de la donnée et de l'information et capitalise sur les précédents cadres conceptuels. L'analyse comparative des figures 7 et 10 semble faire émerger deux domaines très différents. Toutefois quand ces deux domaines sont considérés ensemble (Figure 11), leur taille relative minimise cette différence et leur fusion donne du relief à plusieurs articles qui ont fait le choix de ne pas différencier les deux concepts de gouvernance des données d'une part et de gouvernance de l'information d'autre part.

Les différentes analyses menées font ressortir le concept de but de la gouvernance de donnée et celui de périmètre. Certains articles mettent l'accent sur les activités liées à la gouvernance des données, par exemple les efforts de mise en conformité. D'autres pointent sa structure qui comprend les rôles et les responsabilités. D'autres encore identifient des éléments facilitateurs ou des facteurs clés de succès. Enfin, certains auteurs structurent tout ou partie de ces éléments en un cadre conceptuel. L'étude bibliométrique menée plus haut a fait ressortir le rôle important de l'article [ABR19] qui réunit les deux domaines parfois distincts de gouvernance des données et de gouvernance de l'information et propose un tel cadre conceptuel. Il est à noter que ce cadre emprunte principalement au modèle proposé par Tallon et al. [TAL 13]. A notre connaissance, aucun cadre relatif à la gouvernance des données ou de l'information n'utilise un fondement théorique ni ne considère une vue holistique de cette gouvernance. Les éléments principaux (but, périmètre, activités, structure, etc.), constituant la gouvernance des données dans la littérature et retrouvés dans notre étude bibliométrique, nous conduisent à proposer un cadre fondé sur la théorie des systèmes qui est structurée autour de ces dimensions. C'est un tel cadre que nous proposons dans la section suivante, reprenant les éléments principaux des articles influents identifiés dans l'étude bibliométrique.

#### 4. Vers un cadre de référence pour la gouvernance des données et de l'information.

Notre cadre de référence est fondé sur la théorie des systèmes [SKY 96] et reprend en les complétant les éléments structurants issus des articles du « backbone » de la figure 11. L'utilisation de la théorie des

systèmes découle du fait que, comme nous l'avons établi dans [AKO 19], la gouvernance des données et de l'information est un artefact de conception que l'organisation développe et met en œuvre pour atteindre ses objectifs. Simon caractérise les artefacts en termes de fonctions (*activités*), d'*objectifs* et d'adaptation (*évolution*) [SIM 96]. Il distingue également la *structure* de l'artefact de *l'environnement* dans lequel il opère. De plus, la gouvernance des données et de l'information est au cœur d'un processus de pilotage dont elle ne peut être définie indépendamment. Son objet est la prise de décision, à la fois pour fixer les orientations de l'action (les *objectifs*), et pour définir et ajuster le cadre de fonctionnement (*structure* et *environnement*) et les *activités* correspondantes. En fixant les objectifs, les modalités et les règles du jeu des activités de management de données, la gouvernance des données et de l'information est également amenée à les faire *évoluer* si nécessaire, en fonction des *résultats* obtenus : c'est là son rôle de pilote principal de la donnée. La gouvernance des données et de l'information trouve sa signification dans la mise en relation des éléments qui la composent [LEM 94, SKY 96]. Ces dimensions forment un ensemble indissociable donnant lieu à une évaluation des *résultats* bénéficiant d'une *rétroaction*. À ce titre, la gouvernance des données et de l'information possède toutes les caractéristiques d'un système. Nous présentons ci-dessous les composants de ce système et leur interaction (Fig. 12).



**Figure 12.** Le cadre de référence.

La dimension **But** de notre système comporte deux sous-dimensions. La première sous-dimension se rapporte à l'objectif de la gouvernance des données. Bien sûr son but est dicté par l'alignement avec la stratégie mais il reste celui de maximiser sa valeur en minimisant les coûts et les risques [TAL 13a]. La deuxième sous-dimension précise le périmètre auquel s'applique la gouvernance. Trois composants définissent ce périmètre, à savoir le périmètre « donnée », celui de « l'organisation » et celui du « domaine », tel que suggéré dans [ABR 19]. Dans notre cadre, le périmètre « donnée » intègre les données traditionnelles (structurées), les données massives (semi-structurées ou non structurées) mais aussi les données synthétiques générées par l'intelligence artificielle. Le périmètre « organisation » englobe les niveaux intra- et inter-organisationnels, mais peut aussi ne s'appliquer qu'à une entité de l'organisation. Enfin, le périmètre « domaine » est décrit par sept éléments, à savoir la qualité des

données, la sécurité, l'architecture, le cycle de vie, les métadonnées, le stockage et l'infrastructure, ainsi que la gestion des documents et du contenu.

La deuxième dimension a trait à la **Structure**, en liaison avec les mécanismes structurels proposés dans [TAL 13]. Ces mécanismes englobent les rôles et les responsabilités ainsi que l'attribution du pouvoir de décision.

La troisième dimension, appelée **Activités**, identifie les mécanismes procéduraux et les mécanismes relationnels de la gouvernance tels qu'ils sont proposés dans [TAL 13]. Les premiers englobent la stratégie d'information, les politiques, les règles, les normes, les processus et les procédures, les accords contractuels, le contrôle de la conformité et la gestion des problèmes. Les mécanismes relationnels font référence à la communication, à la formation et à la coordination de la prise de décision.

La quatrième dimension ou **Environnement** du système est composée de facteurs externes et internes à l'organisation et ayant un impact sur sa gouvernance des données. Tallon *et al.* identifient un facteur externe, celui de la réglementation [TAL 13]. Nous proposons d'enrichir cet ensemble de facteurs externes à l'aide du modèle PESTLE [RAS 16], incluant tant l'impact de la politique, que l'économie et le social, mais aussi les éléments technologiques, légaux et environnementaux. Comme le montrent Tallon *et al.*, l'environnement est également composé de différents facteurs internes à l'organisation, ayant un impact potentiel sur le système de gouvernance de la donnée et de l'information [TAL 13]. Plusieurs auteurs mentionnent une liste d'éléments (culture, stratégie informatique, soutien de la direction générale, etc.) Nous proposons de structurer ces facteurs en trois catégories : les personnes, l'organisation et la technologie [HEV 04].

La cinquième dimension **Evolution** décrit les changements du système de gouvernance des données au fil du temps. Elle n'apparaît pas dans la figure 12 qui représente la gouvernance à un moment donné. Ces changements peuvent affecter toutes les autres dimensions. Elle n'est pas décrite dans les cadres de référence de la littérature, autres que les modèles de maturité. Elle constitue une dimension fondamentale permettant aux entreprises de construire et d'alimenter ces modèles de maturité.

Enfin, le système comprend une dimension **Résultats** (valeur, risque, coût, bénéfice, capital investi) qui permet d'exercer la boucle de rétroaction pour faire évoluer le système.

En conclusion, en mobilisant la théorie des systèmes comme fondement de ce cadre conceptuel et les principaux résultats de l'analyse bibliométrique, nous offrons une vue holistique de la gouvernance des données et de l'information composée de cinq dimensions, permettant d'identifier les interactions existant entre elles. Notre contribution consolide les différents axes de la recherche en matière de gouvernance des données et de l'information en un modèle unique enrichi avec les composants d'autres domaines [HEV 04, RAS 16].

Il existe des cadres de référence proposés par les praticiens<sup>2</sup> (DAMA-DMBoK, COBIT, DGI, SAS, BCG). Malgré plusieurs points communs (principes, rôles et responsabilités organisationnels, politiques, règles et normes, processus de base, etc.), chaque cadre a un centre d'intérêt et son propre champ d'application : certains se concentrent sur la gouvernance stratégique, d'autres sur le management en termes de pratiques détaillées de gestion des données, d'autres encore sur des objectifs spécifiques (qualité, conformité, valeur). Notre cadre conceptuel offre une vue globale qui peut être déclinée par niveau et par domaine d'application, par exemple données massives et intelligence artificielle, secteur public et villes intelligentes.

---

<sup>2</sup> <https://www.kellton.com/kellton-tech-blog/popular-data-governance-frameworks#:~:text=,activities%20like%20metadata%20and%20architecture>

## 5. Applications du cadre de référence.

Ce cadre a été utilisé pour générer la structure d'un baromètre d'évaluation de la maturité des entreprises et organisations françaises en matière de gouvernance des données, appelé MetraData<sup>3</sup>. Cette enquête a été menée auprès de 150 entités. Nous avons généré un questionnaire en ligne comprenant environ 40 questions auprès de dirigeants d'entreprise, directeurs des systèmes d'information, directeurs de données et directeurs fonctionnels dans tous les secteurs (privé, public, parapublic, association), tous les domaines d'activité (assurance, technologie, banque, etc.) et toutes les tailles d'entreprises. On a recueilli la perception du concept de gouvernance des données, son opérationnalisation dans l'organisation, ses acteurs-clés, le contexte et les plans d'évolution. Le baromètre met en lumière l'importance des facteurs organisationnels et humains permettant d'améliorer la gouvernance des données. La première colonne du tableau 4 reprend chaque dimension du cadre de référence. La seconde colonne contient un exemple de question fondée sur cette dimension. La troisième colonne fournit les modalités de réponse proposées.

Dimension/ Sous-dimension	Question	Réponses
But/objectif	Quels sont les objectifs principaux de la gouvernance des données au sein de votre organisation ?	Maximiser la valeur métier tirée des données Assurer la conformité aux lois et réglementations Minimiser les risques liés aux données Minimiser les coûts de gestion des données
But/périmètre/ donnée	Quels sont les principaux périmètres d'intervention de la gouvernance des données dans votre organisation ?	Bases de données classiques, documents, "big data", emails, sites web, supports numériques (réseaux sociaux), supports papier, données synthétiques de l'IA
Structure/ Rôles et responsabilités	Quels sont les principaux rôles et responsabilités de la gouvernance des données identifiés dans votre organisation ?	Gestionnaire de données, architecte de données, propriétaire de données, directeur de données, analyste des données, délégué à la protection des données
Structure/Prise de décision	Quelles instances de votre organisation ont en charge tout ou partie de la gouvernance des données ?	Comité stratégique, comité risques et sécurité, comité de direction, comité de pilotage dédié, aucun
Activités/Mécanismes procéduraux	Les mécanismes ci-dessous sont-ils mis en place dans votre organisation pour opérationnaliser les activités de gouvernance des données ?	Gestion de la conformité Preuve et signature électronique Sécurité des données Architecture d'entreprise et modélis. données et flux Architecture des données Gestion de la qualité Définition des rôles liés aux données Investigation informatique (e-discovery) Archivage électronique Gestion des méta-données et catalogage Gestion des données maîtres et de référence Analyse des risques informationnels Gestion des documents et contenus

<sup>3</sup> <https://chairestratgouvinfo.essec.edu/newsroom/barom%C3%A8tre-metradata>

Activités/Mécanismes relationnels	Comment qualifiez-vous le dialogue entre les responsables de la gouvernance des données et les métiers dans votre organisation ?	Excellent, bon, améliorabile, tendu, inexistant
Environnement/Externe	Quel a été l'élément déclencheur qui a suscité une préoccupation pour la gouvernance des données ?	Fusion/acquisition Cyber-attaque Fuite de donnée Nouvelle réglementation (e.g. RGPD), etc.
Environnement/ Interne/ Technologie	Quels sont les outils spécifiques que vous avez mis en place dans votre organisation afin d'opérationnaliser les activités de gouvernance des données ?	Master Data Management Mise en qualité des données Gestion documentaire Modélisation et architecture de données Classification et catalogage de données Registre de traitement RGPD, etc.
Environnement/ Interne/ Personnes et Organisation	Quels sont les éléments qui freinent ou pourraient freiner l'implantation de la gouvernance des données dans votre organisation ?	Manque de moyens humains Résistance au changement Culture d'entreprise éloignée de ces préoccupations Manque de dialogue métiers et équipes gouv. donnée Manque d'engagement de la direction Manque de dialogue entre la DSI et les équipes gouvernance donnée La data n'est pas un enjeu clé
Résultats	Quels éléments de la gouvernance des données sont suivis dans le reporting régulier de l'organisation ?	Avancement des projets Budget Niveau de satisfaction des services data Gestion des compétences

**Tableau 4.** Correspondance entre le cadre de référence et le baromètre MetraData

D'autres applications sont à l'étude. Premièrement, le cadre fournit un modèle théorique pour les recherches futures sur la gouvernance des données. De nouvelles questions de recherche peuvent émerger en croisant les principales dimensions. À titre d'illustration, l'étude de l'évolution de la gouvernance des données dans une organisation ainsi qu'au niveau sectoriel est un domaine de recherche prometteur. Un cadre offre de nombreux avantages pour l'élaboration d'approches de recherche. La théorie des systèmes en constitue une base solide.

Deuxièmement, le cadre peut être utilisé comme modèle pour structurer la charte de gouvernance des données d'une organisation. Certaines chartes sont disponibles sur des sites web, principalement dans le domaine de la santé. Toutefois, elles se concentrent sur les obligations légales et les questions de conformité. Fournir un tel cadre aux experts chargés de rédiger une charte facilite la recherche de l'exhaustivité. Nous avons vérifié sur certaines chartes existantes que notre cadre couvre tous les sujets qu'elles contiennent.

Troisièmement, le cadre de référence fournit une base pour l'évaluation de la gouvernance des données. Dans cette optique, chaque dimension peut être associée à des tests d'audit et à des mesures. Nous avons commencé à définir les questions d'audit avec un panel de praticiens. À titre d'exemple, nous avons affiné les objectifs de valeur, de risque et de coût en élaborant des typologies pour chacun d'entre eux. La valeur englobe plusieurs composantes : commerciale, intrinsèque, de performance, économique, etc. Chacune d'entre elles peut être mesurée à l'aide d'un indicateur spécifique.

Enfin, le cadre de référence peut servir de point de départ au développement d'un modèle de maturité de la gouvernance des données. De nombreux modèles de maturité sont proposés dans la littérature. Ils

sont utilisés pour évaluer, par exemple, la manière dont les parties prenantes comprennent et adoptent les concepts et les méthodes. Pour développer un modèle de maturité, chaque dimension du cadre de référence doit être déclinée en niveaux de maturité et associée à un ensemble de mesures.

## 6. Conclusion et recherche future.

Cette recherche a pour but de structurer le domaine de la gouvernance des données et de l'information en développant un cadre conceptuel fondé sur la théorie des systèmes qui définit ses dimensions clés et alimenté grâce à une étude bibliométrique large incluant les sphères d'influence de la littérature relative à la gouvernance des données et celle ayant trait à la gouvernance de l'information. Ce faisant, nous avons répondu au besoin d'une vision holistique qui puisse guider à la fois les chercheurs dans l'élaboration d'hypothèses et les praticiens dans l'organisation de leur gouvernance des données. Cette recherche a abouti à un cadre conceptuel pour la gouvernance des données et de l'information mais également mis en évidence plusieurs axes dans lesquels des recherches supplémentaires sont nécessaires. Ainsi, toutes les recherches relatives aux sujets comme l'intelligence artificielle, le « cloud », la « blockchain » soulèvent des questions importantes quant à la gouvernance des données. De plus, une validation empirique du cadre conceptuel auprès de praticiens est en cours, ainsi que l'opérationnalisation d'un modèle de maturité de la gouvernance des données et de l'information. L'applicabilité du cadre conceptuel à d'autres types de gouvernance (informatique, connaissance, entreprise, etc.) est une autre voie de recherche future.

## Références

- [ABR 19] ABRAHAM, R., SCHNEIDER, J., & VOM BROCKE, J. (2019). « Data governance: A conceptual framework, structured review, and research agenda. » *International journal of information management*, 49, 424-438.
- [AKO 19] AKOKA, J., & COMYN-WATTIAU, I. (2019). « Évaluation de la gouvernance de l'information : une approche holistique. » In *AIM 2019: 24ème Conférence de l'Association Information et Management*, Nantes, France.
- [ALH 18] ALHASSAN, I. SAMMON, D., & DALY, M. (2018). « Data governance activities: a comparison between scientific and practice-oriented literature. » *Journal of Enterprise Information Management*, 31(2), 300-316. <https://doi.org/10.1108/JEIM-01-2017-0007>
- [ALH 16] ALHASSAN, I. SAMMON, D., & DALY, M. (2016). « Data governance activities: an analysis of the literature. » *Journal of Decision Systems*, 25(sup1), 64-75.
- [ALH 19] ALHASSAN, I. SAMMON, D., & DALY, M. (2019). « Critical success factors for data governance: a theory building approach. » *Information systems management*, 36(2), 98-110.
- [ALR 19] AL-RUITHE, M., BENKHELIFA, E., & HAMEED, K. (2019). « A systematic literature review of data governance and cloud data governance. » *Personal and ubiquitous computing*, 23, 839-859.
- [BRE 18] BRENNAN, R., QUIGLEY, S., DE LEENHEER, P. & MALDONADO, A. (2018). « Automatic extraction of data governance knowledge from slack chat channels. » In *OTM Confederated International Conferences- On the Move to Meaningful Internet Systems" (555-564)*. Cham: Springer International Publishing.
- [BRU 10] BRUENING, P.J., & WATERMAN, K.K. (2010). « Data tagging for new information governance models. » *IEEE Security & Privacy*, 8(5), 64-68.
- [DAV 23] DAVIDSON, E., WESSEL, L., WINTER, J.S., & WINTER, S. (2023). « Future directions for scholarship on data governance, digital innovation, and grand challenges. » *Information and Organization*, 33(1), 100454.
- [DON 04] DONALDSON, A., & WALKER, P. (2004). « Information governance—a view from the NHS. *International journal of Medical informatics*. » 73(3), 281-284.
- [HEV 04] HEVNER, A.R., MARCH, S.T., PARK, J., & RAM, S. (2004). « Design science in information systems research. » *MIS Quarterly*, 75-105.
- [IN 19] IN, J., BRADLEY, R., BICHESCU, B.C., & AUTRY, C.W. (2019). « Supply chain information governance: Toward a conceptual framework. » *The International Journal of Logistics Management*, 30(2), 506-526.
- [JAN 20] JANSSEN, M., BROUS, P., ESTEVEZ, E., BARBOSA, L.S., & JANOWSKI, T. (2020). « Data governance: Organizing data for trustworthy Artificial Intelligence. » *Government information quarterly*, 37(3), 101493.

- [KHA 10] KHATRI, V., & BROWN, C.V. (2010). « Designing data governance. » *Communications of the ACM*, 53(1), 148-152.
- [KOO 11] KOOPER, M.N., MAES, R., & LINDGREEN, E.R. (2011). « On the governance of information: Introducing a new concept of governance to support the management of information. » *International journal of information management*, 31(3), 195-200.
- [LEM 94] LE MOIGNE, J. L. (1994). *La théorie du système général : théorie de la modélisation*. PUF.
- [MER 19] MERKUS, J., HELMS, R., & KUSTERS, R.J. (2019). « Data Governance and Information Governance: Set of Definitions in Relation to Data and Information as Part of DIKW. » In *ICEIS* (2) (pp. 143-154).
- [NGU 16] NGUYEN, T.C. (2016). « Information governance and management in the context of Gov 2.0 (Doctoral dissertation, Swinburne).
- [OTT 11] OTTO, B. (2011). « Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. » *Communications of the Association for Information Systems*, 29(1), 3.
- [PAP 23] PAPAROVA, D., AANESTAD, M., VASSILAKOPOULOU, P., & BAHUS, M.K. (2023). « Data governance spaces: the case of a national digital service for personal health data. » *Information and Organization*, 33(1), 100451.
- [RAS 16] RASTOGI, N., & TRIVEDI, M.K. (2016). « PESTLE technique—a tool to identify external risks in construction projects. » *International Research Journal of Engineering and Technology*, 3(1), 384-388.
- [SIM 96] SIMON, H.A. (1996). *The architecture of complexity: hierarchic systems*.
- [SKY 96] SKYTTNER, L. (1996). *General systems theory: An introduction*. London: Macmillan Press.
- [SMA 19] SMALLWOOD, R. F. (2019). *Information Governance: Concepts, Strategies and Best Practices*, Wiley.
- [TAL 13] TALLON, P.P., RAMIREZ, R.V., & SHORT, J.E. (2013). « The information artifact in IT governance: Toward a theory of information governance. » *Journal of management information systems*, 30(3), 141-178.
- [TAL 13a] TALLON, P.P. (2013). « Corporate governance of big data: Perspectives on value, risk, and cost. » *Computer*, 46(6), 32-38.
- [THOR 21] THOR, A., BORNMANN, L., HAUNSCHILD, R., & LEYDESDORFF, M. (2021). « Which are the influential publications in the Web of Science subject categories over a long period of time? » CRExplorer software used for big-data analyses in bibliometrics. *Journal of Information Science*, 47(3), 419-428.
- [VAN 13] VAN ECK, N.J. & WALTMAN, L. (2013). VOSviewer manual. Leiden: Universiteit Leiden, 1(1), 1-53.
- [WEB 09] WEBER, K., OTTO, B., & OSTERLE, H. (2009). « One size does not fit all---a contingency approach to data governance. » *Journal of Data and Information Quality*, 1(1), 1-27.
- [WEI 04] WEILL, P., & ROSS, J. W. (2004). « IT governance: How top performers manage IT decision rights for superior results. » *Harvard Business Press*.
- [WIL 16] WILKINSON, M. D., DUMONTIER, M., AALBERSBERG, I. J., APPLETON, G., AXTON, M., BAAK, A., ... & MONS, B. (2016). « The FAIR Guiding Principles for scientific data management and stewardship. » *Scientific data*, 3(1), 1-9.