

Les données administratives publiques dans l'espace numérique

N° de la lecture individuelle :

5

Étudiant

Laurent Térence

Sujet

Open Data

Source

<https://univ.scholarvox.com/catalog/book/docid/88928800?searchterm=donn%C3%A9es%20ouvertes>



Table des matières

Table des illustrations	3
Introduction générale.....	4
Introduction.....	4
La description des données administratives	5
Les défis de l'utilisation des données administratives	6
Comprendre le besoin en données administratives de l'administration publique.....	6
Etude de cas de la Drug Enforcement Administration (DEA).....	7
Résumé du chapitre.....	8
Les données administratives : de la création à l'usage	8
La préparation des données	9
L'exploration des données :	9
Résumé du chapitre.....	10
Les approches, les outils et les processus	10
La gouvernance	11
La collecte des données et leur organisation	12
Liens avec données ouvertes.....	12
Les manipulations et les combinaisons	13
La diffusion et la communication	15
Résumé du chapitre.....	16
L'exploitation des données administratives.....	17
La visualisation et représentation des données	17
Quelques approches.....	18
Résumé du chapitre.....	21

Table des illustrations

Figure 1 Chaîne de valeur des données administratives	5
Figure 2 Chemin critique pour une gouvernance des données administratives	11
Figure 3 Données administratives couplées	14
Figure 4 Vocabulaire du catalogue de données (DCAT)	15
Figure 5 Pipeline de visualisation des données	18
Figure 6 Graphique linéaire	19
Figure 7 Infographie	21
Figure 8 Outils d'information spatiale	21

Introduction générale

Ce document vise à présenter une synthèse du livre "Les données administratives publiques dans l'espace numérique" rédigé par Pierre Desrochers. L'objectif principal de cette lecture est d'explorer le thème des données ouvertes, tel que discuté dans l'ouvrage. Les extraits cités dans ce document proviennent du livre mentionné et sont utilisés à des fins internes uniquement. L'auteur, Pierre Desrochers, offre une perspective approfondie sur l'intégration des données administratives publiques dans le contexte du monde numérique, mettant particulièrement l'accent sur les implications et les opportunités liées à l'ouverture des données. Dans cette synthèse, nous explorerons les idées clés présentées par l'auteur, en mettant en lumière les concepts fondamentaux liés aux données ouvertes dans le contexte administratif public.

Introduction

Le monde contemporain est façonné par l'omniprésence des données administratives publiques, un phénomène d'une importance croissante amplifié par l'évolution rapide des outils statistiques et technologiques. Pierre Desrochers, dans son ouvrage "Les données administratives publiques dans l'espace numérique", souligne la prééminence des données dans les débats parlementaires et au sein des communautés de pratique, lesquelles envisagent des visions de gouvernements algorithmiques. Cette mise en avant des données s'inscrit dans un contexte où les leaders actuels exigent un contrôle accru et une responsabilité quant à la collecte et à l'exploitation des données, ainsi que des institutions et des cadres réglementaires qui les encadrent (Balsillie, 2019).

La montée en puissance des données administratives n'est pas un phénomène isolé, mais s'inscrit dans une tendance plus large observée tant dans le secteur privé que dans certaines sphères étatiques. Les entreprises utilisent depuis longtemps les données pour établir des portraits détaillés des consommateurs, guidées par des objectifs de marketing et de manipulation. Des États également, orientent leurs efforts vers l'influence comportementale et sociétale, exploitant des données pour prédire des comportements précis (Bollier et Firestone, 2010; Curran et Smart, 2021; Dubois de Prisque, 2019).

Cet ouvrage s'inscrit dans un paysage où de nombreux livres traitent de l'exploitation des données, mais peu offrent une vue d'ensemble spécifique sur les données administratives au sein des organismes gouvernementaux. Pierre Desrochers propose non seulement une exploration approfondie de ce que sont les données administratives, mais également un guide pratique destiné à soutenir les gestionnaires dans leurs démarches auprès de chercheurs, d'analystes, de praticiens et du public.

Le contenu de l'ouvrage est structuré selon une approche en chaîne de valeur, illustrée dans la figure 1. Cette approche s'inspire du concept de chaîne d'approvisionnement, qui représente le système et les ressources nécessaires pour amener un produit ou un service du fournisseur au client. La chaîne de valeur considère comment la valeur est ajoutée à chaque étape de ce processus, tant pour le produit que pour le service, en impliquant divers acteurs.

Chaque étape de cette chaîne de valeur est détaillée avec ses objectifs spécifiques et les principales activités qui la composent. De plus, l'ouvrage aborde les enjeux, les défis, et les particularités liés à la création et à l'utilisation des données administratives dans le contexte gouvernemental. Ces éléments sont présentés comme un guide pratique, fournissant des informations supplémentaires en annexe pour faciliter la compréhension et l'application des concepts discutés.

FIGURE 1.1
Chaîne de valeur des données administratives

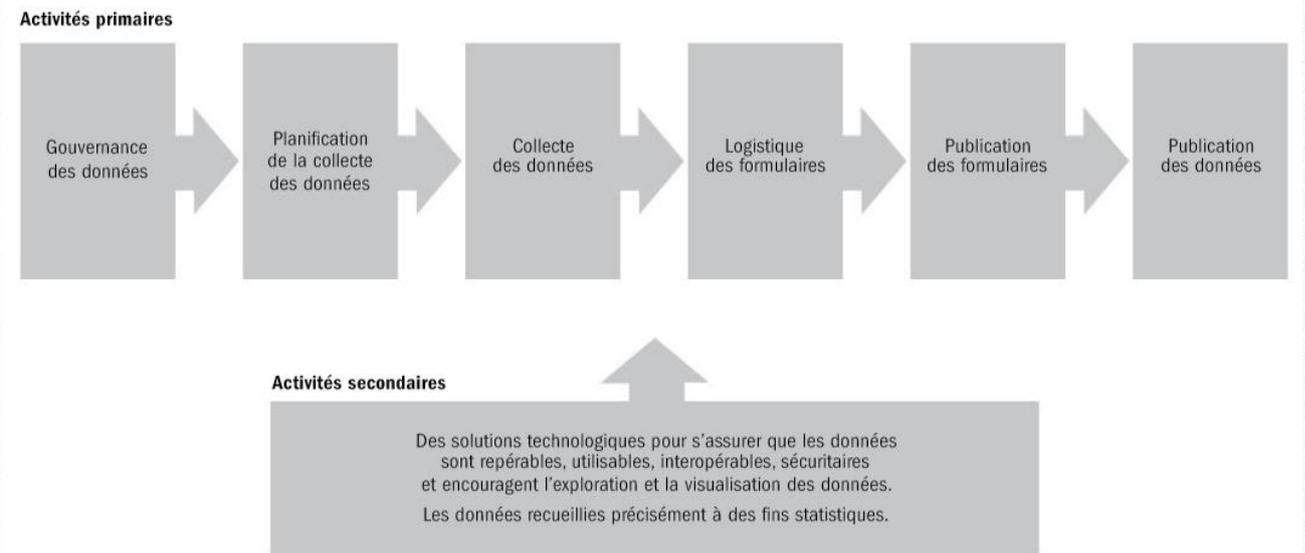


Figure 1 Chaîne de valeur des données administratives

Dans la suite de cette synthèse, nous explorerons les différentes sections de l'ouvrage, allant de la définition des données administratives à leur utilisation pratique, tout en mettant en lumière les défis, les approches, et les enjeux soulevés par Pierre Desrochers.

La description des données administratives

Ce chapitre explore la relation fondamentale entre les données et l'information, en soulignant leur interdépendance dans le contexte de l'administration publique. Les données, définies de manière technique, sont présentées comme des variables liées à des faits, sujets, ou informations, utilisées pour la prise de décision, le traitement, la communication, ou la réinterprétation.

L'information, quant à elle, est décrite comme des données présentées de manière compréhensible, auxquelles une signification a été attribuée dans le contexte de leur utilisation. L'information administrative est générée ou reçue par les organisations, servant à prendre des décisions et permettant de réutiliser et de transmettre les connaissances dans les systèmes informatiques de manière officielle ou non, structurée ou non.

En se penchant sur le concept de données administratives, le texte souligne leur collecte systématique lors d'enregistrements ou de transactions, destinées principalement à des fins administratives. Cependant, ces données peuvent également alimenter des enquêtes statistiques, créant une dualité entre leur usage primaire et leur potentiel usage secondaire.

L'histoire des données administratives, remontant à la Commission Glassco de 1962, met en lumière l'importance des formulaires dans la collecte et le partage d'informations et de données administratives. L'avènement de l'automatisation à la fin des années 1950 a introduit de nouveaux défis, notamment le contrôle de la qualité des données, l'inférence, et la valorisation. Ce contexte a également suscité des préoccupations liées à la publication des données administratives sous un

régime de gouvernement ouvert, incitant les citoyens à s'engager dans les processus décisionnels de l'administration publique. L'automatisation est présentée comme un soutien aux discours sur les défis liés à l'utilisation des données administratives, les considérant comme un actif stratégique dans divers secteurs.

Les défis de l'utilisation des données administratives

Ce passage examine les défis inhérents à l'utilisation des données administratives, soulignant l'impact de la numérisation sur la gestion de ces données. Les données administratives, provenant des processus administratifs, sont cruciales pour évaluer les progrès et améliorer les activités gouvernementales.

L'avènement des technologies numériques a introduit une complexité accrue en raison du volume considérable de données générées par les interactions avec la clientèle. Cette augmentation du volume et de la complexité des données a des répercussions sur leur gestion, affectant divers aspects tels que la prise de décision, la gestion des flux de données, la qualité des données, et la sécurité de l'information.

Les défis organisationnels comprennent la difficulté à inventorier les données dans les systèmes informatiques, le manque de documentation sur la source d'autorité des données, et la compréhension variable du contenu et du sens des données. Des enjeux de responsabilité et de qualité des données sont également soulevés, incluant des questions sur la sécurité, la confidentialité, et le partage des données.

Sur le plan technologique, des problèmes comme la duplication inutile des données, l'inconsistance dans l'intégrité des données, et la modélisation des données par les gestionnaires de programmes sont identifiés. De plus, l'absence fréquente d'une architecture d'entreprise ou de référence appropriée contribue à des défis supplémentaires dans le développement des systèmes.

Le passage conclut en soulignant que, dans le contexte actuel, de nombreuses organisations gèrent un ensemble varié et expansif de données administratives, nécessitant une coordination approfondie entre professionnels et gestionnaires pour rester à la pointe dans la planification et la mise en œuvre de programmes et de services.

Comprendre le besoin en données administratives de l'administration publique

Ce passage examine le besoin croissant en données administratives dans le contexte d'une société de plus en plus numérique. Il remet en question la métaphore traditionnelle selon laquelle les données sont le nouveau pétrole, les comparant plutôt au plutonium en raison de leur puissance, de leur potentiel danger et de leurs conséquences graves en cas de mauvaise utilisation.

Pour définir le besoin en données administratives de l'administration publique, la première étape consiste à comprendre leur rôle dans la planification, la gestion des opérations, et le suivi des performances des programmes. Le gouvernement moderne, axé sur les technologies numériques, doit exploiter ces données pour favoriser le progrès économique, social, international et gouvernemental.

L'extraction de sens des données administratives devient cruciale, impliquant le traitement, l'analyse, et la communication des données de manière régulière et rigoureuse. Les gestionnaires et les professionnels recherchent des données fiables, faciles à traiter et à analyser, pertinentes, et disponibles dans des délais courts.

Les données administratives doivent être structurées et organisées pour en tirer profit dans l'interprétation du contexte administratif. L'exemple de la Drug Enforcement Administration des États-Unis est cité, démontrant comment des données administratives bien analysées peuvent révéler des informations cruciales, comme dans le cas des opioïdes.

En conclusion, les organisations sont encouragées à revoir leurs processus, politiques administratives, et solutions technologiques pour recueillir, traiter, contrôler et interpréter les données de manière plus stratégique, reconnaissant ainsi les données administratives comme des actifs stratégiques essentiels.

Etude de cas de la Drug Enforcement Administration (DEA)

L'étude de cas de la Drug Enforcement Administration (DEA) des États-Unis offre un exemple remarquable de l'utilisation stratégique des données administratives dans le contexte de la crise des opioïdes.

Contexte :

Au cours des années 2000 et 2010, les États-Unis ont été confrontés à une crise majeure liée à l'abus d'opioïdes, avec des conséquences dévastatrices sur la santé publique. Pour comprendre l'ampleur du problème, la DEA a entrepris une initiative ambitieuse pour recueillir, analyser et interpréter les données administratives liées à la fabrication, à la distribution et à la vente d'opioïdes.

Objectif :

L'objectif principal de cette démarche était d'obtenir une vision complète du flux des opioïdes, depuis leur production par les fabricants jusqu'à leur distribution aux pharmacies, afin de mieux comprendre les pratiques commerciales et d'identifier d'éventuelles activités illégales.

Méthodologie :

La DEA a compilé une vaste quantité de données administratives provenant de diverses sources, notamment des fabricants, des distributeurs et des pharmacies. Ces données comprenaient des informations détaillées sur la quantité d'opioïdes produite, distribuée et vendue. L'analyse a porté sur une période significative, couvrant plusieurs années.

Résultats :

Les résultats de cette analyse ont fourni des informations cruciales. Parmi les conclusions marquantes, on a découvert que quelques fabricants étaient responsables d'une proportion très importante des opioïdes produits, et que certains États recevaient des quantités disproportionnées par rapport à leur population. Par exemple, une ville de Virginie-Occidentale a reçu un nombre exceptionnellement élevé de comprimés d'opioïdes par habitant.

Impacts :

Cette analyse approfondie a eu des conséquences significatives. Elle a permis d'identifier des schémas alarmants dans la distribution d'opioïdes, conduisant à des enquêtes, des actions réglementaires et des poursuites judiciaires à l'encontre des entreprises impliquées. Les données administratives ont ainsi servi de base à des actions visant à atténuer la crise des opioïdes.

Leçons tirées :

L'étude de cas de la DEA souligne l'importance d'une utilisation stratégique des données administratives pour résoudre des problèmes sociaux complexes. Elle met en lumière la capacité des

données à révéler des tendances, à identifier des responsabilités, et à informer des actions gouvernementales efficaces. Cette approche démontre également la nécessité pour les gouvernements de moderniser leurs méthodes de collecte, de traitement et d'analyse des données administratives pour répondre aux défis contemporains de manière proactive.

Liens avec données ouverte :

Dans l'étude de cas de la Drug Enforcement Administration (DEA), bien que le terme "données ouvertes" ne soit pas explicitement mentionné, il y a un lien indirect avec le concept de données ouvertes. Les données administratives utilisées par la DEA pour comprendre la crise des opioïdes et prendre des mesures stratégiques représentent une forme de données gouvernementales.

Le concept de données ouvertes implique la mise à disposition de données gouvernementales de manière transparente, accessible et utilisable par le public. Bien que l'étude de cas ne traite pas spécifiquement de la diffusion publique de ces données, les principes de transparence et de responsabilité inhérents aux données ouvertes peuvent être pertinents.

Si les résultats de cette analyse étaient rendus accessibles au public, cela pourrait contribuer à l'objectif de transparence gouvernementale. Les citoyens, les chercheurs, et d'autres parties prenantes pourraient ainsi accéder à ces données, les analyser indépendamment, et participer au processus de surveillance et de responsabilisation du gouvernement dans la gestion de la crise des opioïdes.

En résumé, bien que l'étude de cas ne traite pas explicitement des données ouvertes, elle illustre la manière dont l'utilisation stratégique des données administratives par le gouvernement peut être alignée sur les principes de transparence et de responsabilité, qui sont des aspects clés du mouvement des données ouvertes.

Résumé du chapitre

Le chapitre explore la diversité des définitions attribuées au terme "données" dans diverses disciplines. Dans le cadre de l'administration publique, les données sont définies comme des variables liées à des faits, des sujets ou des éléments d'information, destinées à soutenir la prise de décision, le traitement, la communication et la réinterprétation. Une attention particulière est accordée aux données administratives, qui sont systématiquement collectées lors de l'enregistrement de personnes, de la réalisation de transactions, ou générées au cours des activités liées à la prestation de services.

Le chapitre souligne l'importance cruciale d'établir des méthodes et des normes pour la création, la protection et la gestion de l'information et des données en tant qu'actifs stratégiques. Il met également en lumière le lien entre ces données administratives et le concept de données ouvertes, encourageant une approche transparente et accessible à ces actifs stratégiques. En mettant en avant la nécessité de traiter, analyser et communiquer régulièrement et rigoureusement ces données, le chapitre s'inscrit dans le contexte plus large de l'ouverture des données pour maximiser leur utilité et leur interprétation opportune.

Les données administratives : de la création à l'usage

Le chapitre 2 explore le parcours des données administratives, de leur création à leur utilisation. Il met en évidence deux types principaux de dossiers : les dossiers d'exploitation, qui englobent les informations générées dans le cadre des fonctions normales de l'organisation, y compris des données confidentielles provenant d'autres entités; et les dossiers administratifs, liés aux activités internes de

l'administration. Ces derniers englobent divers services tels que les finances, les ressources humaines et la gestion des biens.

L'importance de la transparence et de la reddition de comptes au Parlement et à la population est soulignée, avec une explication détaillée des méthodes utilisées pour mesurer la performance organisationnelle, notamment la fréquence, la durée et l'intensité. Le chapitre met également en garde contre la quantification excessive de la performance, soulignant la nécessité de revoir de manière cohérente la compréhension de la gestion publique.

Les gestionnaires de programmes sont confrontés à des défis lors de la collecte de données, notamment la diversité des données recueillies et les contraintes liées aux technologies de l'information et à la culture organisationnelle. Le chapitre souligne la nécessité d'une approche collaborative, de la patience et de l'utilisation de techniques d'analyse approfondies pour surmonter ces défis. Bien que de nombreuses organisations recueillent des données, la normalisation et la cohérence dans la diffusion de ces données entre les entités participantes restent des obstacles à surmonter.

La préparation des données

Ce chapitre explore le défi que représentent les données administratives pour de nombreuses organisations. Elle souligne le besoin d'une approche globale de la gestion de l'information en tant qu'actif stratégique, mettant en garde contre la prolifération non maîtrisée de données à plusieurs niveaux de l'organisation.

La nécessité de procédures et de politiques administratives pour garantir la qualité des données dès leur collecte est soulignée. Une revue par un comité d'experts ou un gestionnaire public est recommandée pour établir un équilibre entre le coût de la collecte, son exhaustivité et la fiabilité des données. L'imputabilité de la validation de la collecte de données est attribuée à un individu au sein de l'organisation.

Des questions clés sont suggérées comme balises dans les processus de collecte et de préparation des données, couvrant des aspects tels que la source des données, la qualité des réponses, la pertinence des données recueillies, et la possibilité de réutiliser l'information dans le futur.

Enfin, la section souligne l'existence de modèles pour mesurer la maturité de l'organisation en matière de gouvernance des données, ainsi que la nécessité d'une adaptation aux besoins de l'organisation et de la validation par des utilisateurs externes conformément à des directives ou normes spécifiques. Elle illustre ces principes par des exemples concrets, notamment dans le secteur de l'exploitation du pétrole et du gaz, soulignant l'impact des obligations légales sur les politiques administratives encadrant la collecte et la consignation des données administratives.

L'exploration des données :

Le chapitre sur "L'exploration des données" examine le rôle crucial de cette phase dans la chaîne de production des données administratives, suivant les procédures de collecte de données et les principes d'assurance qualité présentés dans le premier chapitre. La gestion des données, concernant la consignation, l'indexation, le traitement, la diffusion, et la veille stratégique des données administratives, est positionnée comme un élément clé dans la détermination de l'utilité des données pour atteindre les objectifs organisationnels.

Le chapitre souligne l'importance de la documentation via un plan de gestion des données, défini a priori lors de la collecte des données. Ce plan, généralement centralisé, joue un rôle essentiel en fournissant des directives pour la normalisation, l'intégration, la documentation des métadonnées,

les standards de collecte, la consignation, le nettoyage, et d'autres fonctions de traitement des données. Ces plans sont essentiels pour assurer la réutilisation des données par d'autres organismes ou instituts.

L'exemple du réseau Portage est cité comme un outil bien connu pour l'intendance partagée des données de recherche au Canada. Il souligne également l'importance de créer des plans de gestion des données conformes aux politiques de diffusion et de communication de la recherche scientifique subventionnée par l'État, en fournissant des éléments clés tels que la collecte de données, la documentation et les métadonnées, le stockage et la sauvegarde, la conservation, le partage et la réutilisation, les responsabilités et les ressources, ainsi que la conformité aux lois et à l'éthique.

En conclusion, le chapitre met en avant l'existence de différentes disciplines techniques ayant développé des pratiques, vocabulaires et conventions spécialisés pour la gestion des données. Il suggère que l'adaptation de modèles pour des domaines techniques spécifiques peut être nécessaire, et prône une approche agile pour simplifier la planification et la mise en œuvre du plan de gestion des données, soutenant ainsi la création des données administratives dans leurs contextes particuliers d'usage.

Résumé du chapitre

Le chapitre aborde la création et l'utilisation des données administratives, soulignant l'importance de la réflexion approfondie dans le processus de collecte ou création de données. La collecte est guidée par la problématique identifiée, liée à une théorie, permettant ainsi une compréhension de l'enjeu de politique publique et la définition de la portée de la collecte de données.

L'utilisation des données administratives varie selon les organisations, qui déterminent la portée et la qualité nécessaires pour les opérations, établissant ainsi les modalités de collecte. Ces données sont souvent constituées à la fois des traces d'activités et de transactions opérationnelles, ainsi que des actifs informationnels tels que les dossiers de cas, les dossiers d'exploitation et les dossiers administratifs.

Le chapitre souligne l'importance pour le gestionnaire public de remettre en question ce qui est observable et observé, mettant en avant la qualité des données et interrogeant la culture organisationnelle et les processus de collecte pour garantir leur adaptation à l'usage. Il mentionne la formulation d'outils visuels pour soutenir la prise de décision et la socialisation afin de comprendre les valeurs, la mentalité, les concepts et les définitions préexistantes influençant la création d'indicateurs.

La collecte des données est présentée comme reposant sur trois assises : la fréquence, la durée dans le temps et l'intensité. La préparation des données s'appuie sur deux principes, à savoir les procédures de collecte de données et la qualité des données. Enfin, l'exploration des données est intégrée dans la chaîne de production et de valeur des données administratives, suivant les procédures de collecte de données et les principes d'assurance qualité.

Les approches, les outils et les processus

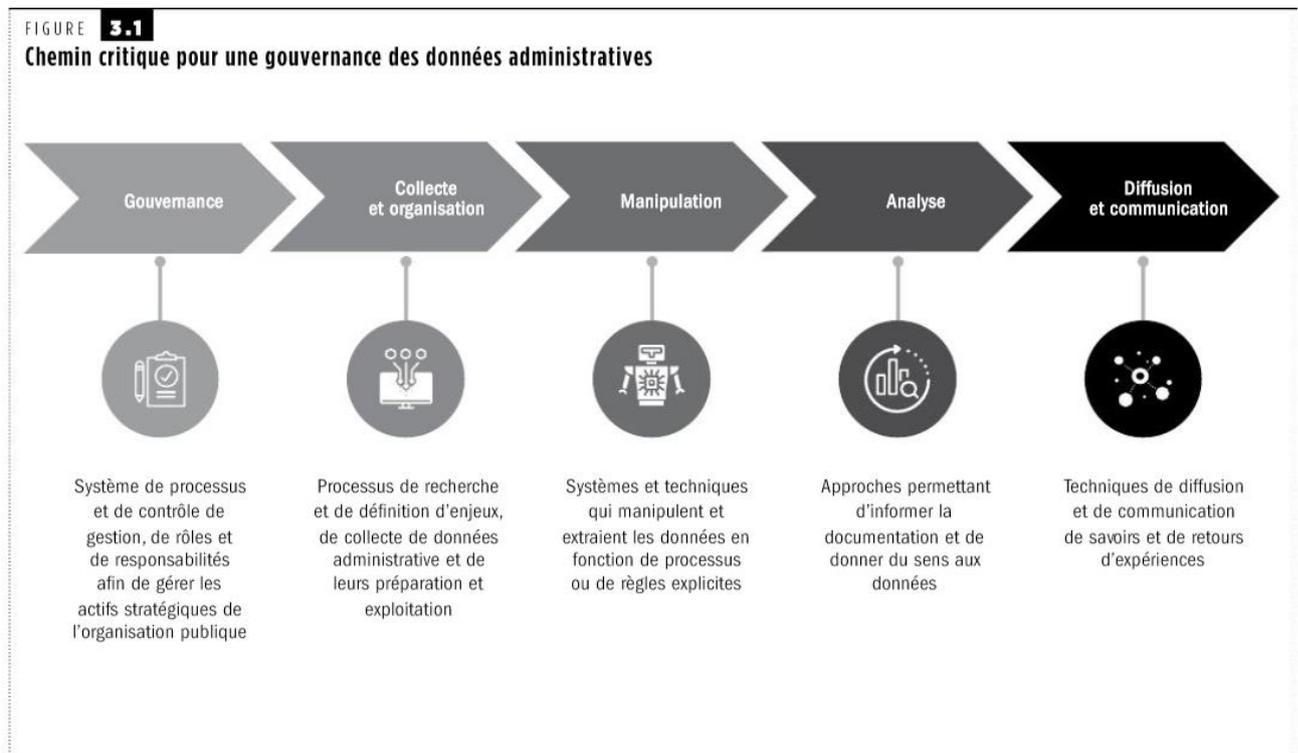
Le chapitre met en lumière les défis et les obstacles liés à l'utilisation des données administratives dans les organisations publiques. Il souligne que, bien que ces données soient désormais considérées comme une ressource stratégique essentielle, de nombreuses organisations reconnaissent les difficultés liées à leur utilisation et à leur partage. Un exemple illustratif de ces défis est le scandale Windrush au Royaume-Uni, qui a mis en évidence les conséquences néfastes de l'utilisation de

données administratives inexactes ou incomplètes sur le statut d'immigrants, entraînant des problèmes administratifs et des conséquences graves.

Le chapitre souligne également le manque d'informations sur la gestion des données par les différents ministères, ce qui rend les évaluations de leur utilisation difficile. On suggère que peu d'organisations publiques ont réussi à mettre en place une gestion saine et un contrôle efficace des données administratives.

Le gouvernement fédéral, conscient de la valeur potentielle des données administratives, a lancé une initiative en 2018 pour promouvoir leur réutilisation dans le but d'améliorer les services et de les rendre disponibles sur un portail de données ouvertes. Cependant, de nombreuses organisations publiques rencontrent des difficultés pour exploiter pleinement les données sous leur contrôle en raison d'outils, d'approches et de processus d'analyse obsolètes.

Le chapitre conclut en soulignant la nécessité d'une approche stratégique avec un cadre précis de gouvernance des données, notamment en période de transformation numérique, pour remodeler l'approche organisationnelle en matière de données et reconnaître la relation entre la gouvernance des données, la gouvernance de l'information et la gestion des données.



Icônes du Noun Project, <<https://thenounproject.com>>, par Justin Blake, WEBTECHOPS LLP, Komkrit Noenpoempisut, Aneeqe Ahmed Wenjie / CC BY 3.0.

Figure 2 Chemin critique pour une gouvernance des données administratives

La gouvernance

Le chapitre sur la gouvernance met en avant la nécessité d'établir un cadre rigoureux pour la valorisation des données administratives. La gouvernance, première étape de ce cadre, est définie comme reposant sur une vision globale pour l'organisation, mettant l'accent sur les processus et offrant un cadre pour mieux exploiter les connaissances.

Le texte souligne la confusion courante entre les concepts d'information, de données, de gouvernance de l'information et de gouvernance des données. Il insiste sur l'importance de clarifier les limites entre ces concepts pour extraire de la valeur des données administratives.

La gouvernance repose sur la mise en place d'un cadre composé de normes, de processus, de rôles et de paramètres pour répondre aux exigences réglementaires, juridiques, opérationnelles et en matière de risques. La gouvernance de l'information et la gouvernance des données, situées à un niveau élevé dans l'organisation, font partie de la gouvernance d'entreprise et nécessitent un leadership fort.

La gouvernance des données est définie comme un système de processus et de contrôles de gestion, de politiques administratives, de droits de décision et de responsabilités visant à assurer la mise en œuvre de contrôles de gestion officiels. Elle opère au niveau intermédiaire des objectifs de l'organisation et partage des enjeux de gouvernance similaires à ceux de l'information.

Le texte explique que la gestion des données repose fortement sur les technologies de l'information, qui soutiennent la planification, l'élaboration, la mise en œuvre et l'administration des systèmes pour gérer efficacement les flux de données d'une organisation. La gestion des données se concentre principalement sur les aspects opérationnels ou techniques, et différents gestionnaires dans l'organisation peuvent être responsables de cette fonction.

La collecte des données et leur organisation

Le chapitre souligne que l'Indice du bien-être des collectivités, principal outil pour mesurer le bien-être socioéconomique dans les réserves, était incomplet car il ne prenait pas en compte des aspects essentiels tels que la santé, l'environnement, la langue et la culture des membres des Premières Nations. Pour combler ces lacunes, le texte propose des réponses techniques, mettant en avant l'accessibilité croissante des données administratives aux gestionnaires de programmes.

Il mentionne la mise en place de catalogues de données administratives comme une solution technique pour gérer les actifs organisationnels. Ces catalogues, reposant sur des standards de documentation technique, facilitent les échanges entre organisations publiques et contiennent des informations techniques et des métadonnées essentielles. Les métadonnées, définies comme des informations descriptives ou contextuelles associées à une ressource, sont cruciales pour la gestion efficace des données administratives.

Le chapitre insiste sur la construction des métadonnées, les classant en trois catégories : descriptives, techniques et administratives. Les métadonnées descriptives facilitent l'identification, la localisation et la récupération des données, tandis que les métadonnées techniques décrivent les processus techniques utilisés pour produire les données. Les métadonnées administratives sont utilisées pour gérer des aspects administratifs liés aux données.

Le catalogue de données est présenté comme un outil essentiel permettant de promouvoir la gestion des données externes, de conserver une trace des usages internes, des changements dans le temps et des relations entre les actifs. Il offre également des avantages tels que l'analyse de l'information personnelle, la gestion temporelle des données et le support des efforts de dépersonnalisation.

En résumé, le catalogue de données est considéré comme une composante fondamentale de l'exploitation des données administratives, offrant une plateforme de gestion des flux informationnels, des systèmes de processus et de contrôle de qualité.

Liens avec données ouvertes

Le texte mentionne brièvement l'initiative du Partenariat pour un gouvernement ouvert, dans laquelle le gouvernement canadien participe. Cette initiative suggère que l'ensemble des données

administratives peuvent être publiées sur un portail de données ouvertes, permettant aux partenaires et aux citoyens de tirer profit des sources de données des organisations publiques. Cela s'aligne avec le mouvement mondial en faveur de la transparence et de l'accès ouvert aux données du secteur public.

Le lien avec les données ouvertes réside dans la possibilité de rendre accessibles au public certaines données administratives. Les catalogues de données administratives, mentionnés dans le chapitre, peuvent également être considérés comme des outils qui favorisent la mise en œuvre de pratiques de données ouvertes en rendant les informations accessibles, compréhensibles et utilisables.

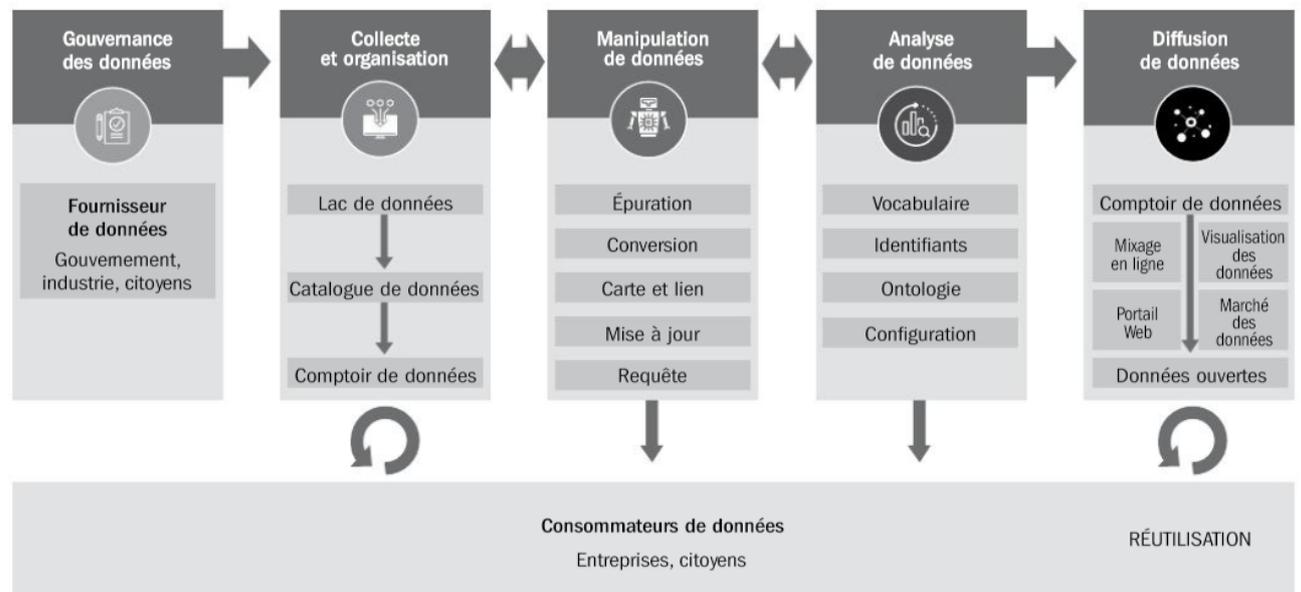
En résumé, il existe un lien potentiel entre la gestion des données administratives, telle que discutée dans le chapitre, et l'initiative des données ouvertes. Les principes de transparence, de partage et d'accès libre à l'information du secteur public peuvent être intégrés dans la gouvernance et la gestion des données administratives.

Les manipulations et les combinaisons

Le chapitre aborde la troisième étape du processus de valorisation des données administratives, intitulée "Les manipulations et les combinaisons". Voici les points importants à retenir avec un accent sur le lien avec les données ouvertes :

1. **Master Data Management (MDM)** : Le chapitre introduit les techniques de manipulation des données, notamment le MDM, un système visant à capter, définir, gérer et communiquer une vue d'ensemble des données faisant autorité au sein d'un système informatique. L'objectif est d'obtenir des données de qualité, pertinentes, exactes, actuelles, accessibles, intelligibles, cohérentes et issues d'une seule source.
2. **Évolution des Techniques** : Le texte souligne que les techniques de manipulation des données ont évolué, mettant en avant l'utilisation d'outils technologiques qui s'écartent du paradigme des bases de données relationnelles. Le couplage des données administratives provenant de diverses sources est mentionné comme une approche pour nettoyer les données, identifier les anomalies, les doublons, les données manquantes, et standardiser les valeurs, améliorant ainsi l'exactitude et la cohérence des données.

FIGURE 3.3
Données administratives couplées



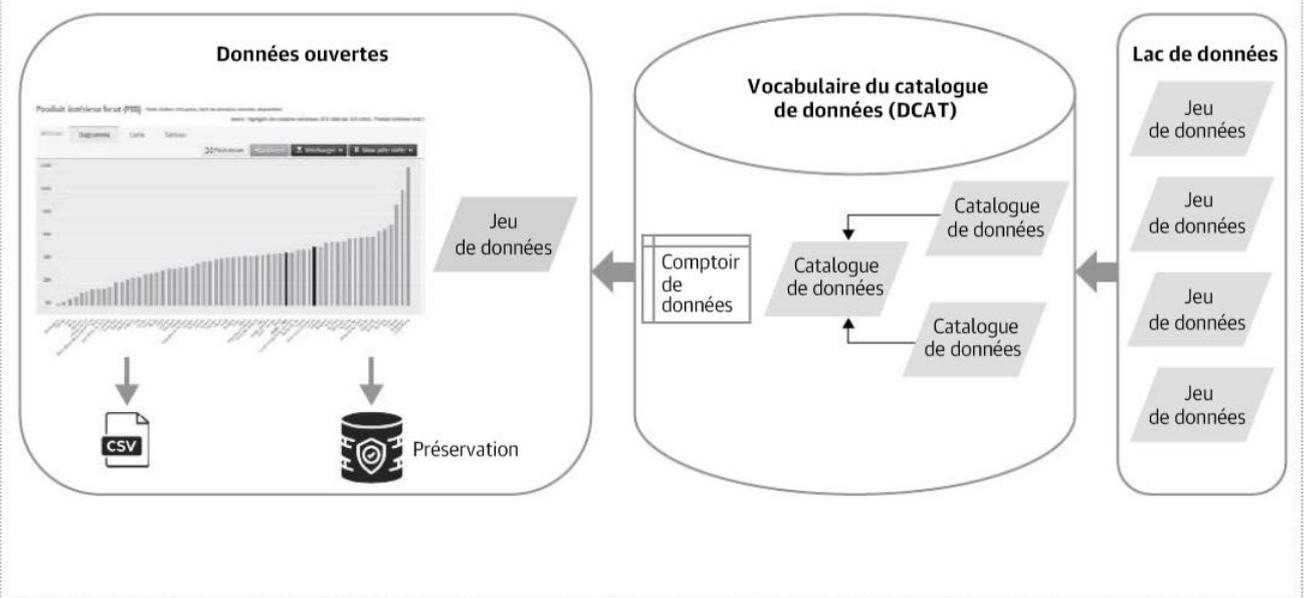
Icônes du Noun Project, <<https://thenounproject.com>>, par Justin Blake, WEBTECHOPS LLP, Komkrit Noenpoempisut, Aneeqe Ahmed, Wenjie, Anna Sophie / CC BY 3.0.
Source : Inspiré de Desrochers, 2012 ; Ding, Peristeras et Hausenblas, 2012.

Figure 3 Données administratives couplées

3. Apprentissage Automatique : L'apprentissage automatique est présenté comme une méthode qui peut effectuer des évaluations sur les données manquantes, accélérant considérablement les activités de manipulation des données. Cependant, le texte souligne les risques liés à l'opacité des algorithmes de prise de décision automatique, notamment en termes de préjugés et de conséquences potentielles sur les droits fondamentaux.

4. Lien avec les Données Ouvertes : Le texte établit un lien avec les données ouvertes en mettant en avant l'importance de l'interopérabilité et la capacité technique à coopérer, comprendre et communiquer l'information. Il présente le vocabulaire du catalogue de données (DCAT) comme un exemple facilitant l'interopérabilité entre les catalogues de données, favorisant ainsi la recherche et l'utilisation des données par les chercheurs et le public à travers les portails de données ouvertes.

FIGURE 3.4
Vocabulaire du catalogue de données (DCAT)



icônes du Noun Project, <<https://thenounproject.com>>, par Ilham Fitrotul Hayat, Ben Davis / CC BY 3.0.
Source: Inspiré de Lisowska, 2016.

Figure 4 Vocabulaire du catalogue de données (DCAT)

5. Évolution Historique : L'évolution historique des initiatives de données ouvertes est brièvement mentionnée, depuis l'invention du World Wide Web par Tim Berners-Lee jusqu'aux mouvements de données ouvertes lancés dans différentes juridictions à partir de 2009. Le texte souligne également l'importance des vocabulaires de métadonnées, tels que SDMX, ADMS, CCMS, et la norme DCIP, pour faciliter l'interopérabilité des données.

6. Plateformes Collaboratives : Enfin, le chapitre mentionne l'existence de plateformes collaboratives permettant aux gestionnaires, chercheurs, analystes et praticiens de repérer différents catalogues de données, soulignant ainsi l'importance de la collaboration dans le contexte des données ouvertes.

En résumé, le chapitre explore les techniques de manipulation des données, mettant en avant l'évolution technologique, les risques liés à l'apprentissage automatique, et établissant un lien significatif avec les principes des données ouvertes, notamment en termes d'interopérabilité et de partage de métadonnées.

La diffusion et la communication

Le chapitre souligne l'importance de revendiquer un enjeu de politique publique avec autorité, soulignant qu'une résolution efficace dépend de la facilité d'exploitation et de communication des données, particulièrement dans l'espace public numérique. Voici les points clés, mettant en lumière le lien avec les données ouvertes :

1. Numérisation de l'Espace Public : L'arrivée du numérique remet en question la manière dont les institutions publiques présentent et diffusent l'information dans l'espace public. La transition vers le numérique nécessite une réévaluation des méthodes de construction et de diffusion des connaissances.

2. **Modèles de Gestion des Connaissances** : Les modèles de gestion des connaissances, tels que ceux de Dalkir (2005) et de Mayer et Zack (1996), sont explorés. Les données administratives doivent être accompagnées de guides ou d'informations complémentaires pour faciliter leur interprétation et les mettre en contexte.

3. **Normalisation des Données** : La normalisation des données dans l'appareillage technologique de soutien est cruciale pour déterminer comment le contenu sera stocké, manipulé et récupéré. Des normes et politiques administratives sont nécessaires pour garantir une utilisation saine et complète des données tout en respectant des balises telles que la confidentialité.

4. **Objectif de l'Utilisation des Données** : La pleine utilisation des données n'implique pas nécessairement une image parfaite et complète de la population, mais plutôt une analyse nécessaire à la formulation et à la résolution d'un problème public. Il est crucial de comprendre les limites d'interprétation des données administratives.

5. **Gouvernement Ouvert** : La diffusion proactive de données, par le biais de portails de données ou d'initiatives de gouvernement ouvert, est présentée comme une pratique alignée avec des éléments fondamentaux de bonne gouvernance, de transparence, d'intégrité et de participation des parties prenantes.

6. **Écosystème de Données Ouvertes** : Le gouvernement du Québec a développé un "écosystème de données ouvertes" pour décloisonner l'information, maximiser la valorisation de l'information entre les partenaires, et mettre en place des structures de gouvernance, des normes, des processus et des règles administratives pour l'exploitation et la diffusion des données

7. **Accessibilité des Données** : La capacité des citoyens ou des acteurs publics à repérer, interpréter et traiter uniformément les données est limitée par la difficulté de trouver ces données. Un portail rassemblant les catalogues de données administratives peut réduire ces barrières institutionnelles.

8. **Principes de Fidélité et Validité** : La Sunlight Foundation propose 10 principes de fidélité et de validité pour les données ouvertes, soulignant leur complétude, caractère primaire, disponibilité rapide, accessibilité, lisibilité par machine, non-discrimination, encadrement par des normes communes, publication sous licence ouverte, et disponibilité sans frais.

En résumé, le chapitre met en avant la nécessité d'adapter les pratiques de gestion des connaissances et de normalisation des données à l'ère numérique, en soulignant le rôle crucial des données ouvertes dans la gouvernance transparente et participative.

Résumé du chapitre

Ce chapitre met en lumière les différentes dimensions de la gouvernance de l'information, la gouvernance des données et la gestion des données au sein d'une organisation. La gouvernance de l'information et la gouvernance des données sont axées sur la vision stratégique, tandis que la gestion des données se concentre sur la mise en œuvre technologique. La gestion des données repose largement sur les technologies de l'information pour administrer les systèmes d'architecture, les catalogues de données, ainsi que les pratiques informatiques, garantissant ainsi une gestion efficace des données. La création d'un catalogue de données administratives émerge comme une pratique essentielle pour gérer les actifs stratégiques de l'organisation.

Le programme de gestion des données Master Data Management (MDM) est souligné comme une méthode pour garantir l'obtention de données de qualité, bien que ses limites soient reconnues face à l'évolution des nouvelles technologies. Le couplage des données administratives est présenté comme une alternative ou un complément au programme MDM, visant à améliorer l'exactitude et la cohérence des données.

Enfin, le chapitre souligne l'importance de la conceptualisation et de la présentation des données et de l'information pour la diffusion et la communication. Ces approches guident la manière dont les concepts sont abordés, compris et utilisés dans l'organisation, soulignant ainsi l'importance de la communication efficace dans la gestion des données.

L'exploitation des données administratives

Ce chapitre se concentre sur les défis et les opportunités liés à l'exploitation des données administratives, offrant aux gestionnaires des perspectives pour maximiser leur utilité. L'importance des données administratives dans l'alimentation des technologies d'apprentissage automatique et de l'intelligence artificielle est soulignée. En plus de leur rôle crucial dans la prise de décision, le chapitre explore comment ces données peuvent être utilisées efficacement en relation avec ces technologies numériques.

La visualisation des données administratives est présentée comme une première approche, permettant aux gestionnaires de tirer parti de graphiques statiques, de tableaux de bord et de graphiques interactifs. Ces outils visuels facilitent une prise de décision éclairée en aidant les organisations à explorer et à comprendre plus profondément le sens de leurs données, voire à anticiper des résultats potentiels.

En poursuivant, le chapitre examine la contribution potentielle des données administratives dans le domaine du développement et de l'utilisation de l'apprentissage automatique et de l'intelligence artificielle. Il explore comment ces technologies peuvent être optimisées grâce à des données bien structurées, ouvrant ainsi la voie à des analyses plus avancées et à des prédictions plus précises pour soutenir les processus décisionnels.

La visualisation et représentation des données

Le chapitre explore l'exploitation des données administratives, mettant en avant la puissance de la visualisation et de la représentation des données. Il souligne que des représentations visuelles efficaces permettent aux gestionnaires et praticiens d'explorer et de comprendre plus profondément leurs données, identifiant des tendances, des modèles, et des éléments significatifs. Cette exploration génère une connaissance approfondie, facilitant ainsi des prises de décision fondées sur des données pertinentes, exactes, intègres, et actuelles.

Selon les travaux d'Azzam et ses collaborateurs, la visualisation des données repose sur trois critères : elle est basée sur des données qualitatives ou quantitatives, elle offre une représentation graphique de l'information numérique, et elle est lisible par le lecteur, facilitant ainsi l'exploitation, l'examen, et la communication des données.

Le chapitre explore également comment la visualisation des données peut politiser un enjeu de politique publique en influençant l'intérêt des parties prenantes et en suscitant des débats. Il souligne le rôle essentiel de la visualisation dans la simplification de la compréhension, permettant aux gestionnaires de créer une conversation autour d'un enjeu de politique publique, même pour des politiciens non-experts.

En conclusion, le chapitre souligne que la visualisation des données peut contribuer à activer des enjeux de politique publique, promouvoir l'usage de données probantes, et réduire le poids des opinions non fondées dans le débat public. Cependant, il avertit que la visualisation doit rester neutre, évitant d'établir une certaine idéologie dans l'univers politique, et rappelle l'importance d'utiliser des méthodes préservant le côté factuel des données. Le processus de construction de la visualisation est présenté comme la clé du succès, soulignant la nécessité de faciliter l'accès à l'information et de transmettre le message de manière simple tout en préservant la complexité pertinente des données.



Figure 5 Pipeline de visualisation des données

Quelques approches

Le chapitre explore diverses approches de construction de la visualisation des données administratives. La méthode la plus courante consiste à présenter un dénombrement des informations dans des diagrammes ou tableaux, codant visuellement ces informations en fonction d'une ou plusieurs dimensions. Le nombre de dimensions détermine le niveau de complexité de la visualisation, influençant son accessibilité et sa compréhension par le lecteur.

Le chapitre souligne que des représentations visuelles, comme les graphiques linéaires, peuvent simplifier ou complexifier la visualisation en fonction du public cible et de son expertise dans le domaine. L'objectif principal est de faciliter la réception et la compréhension de l'information par le lecteur.

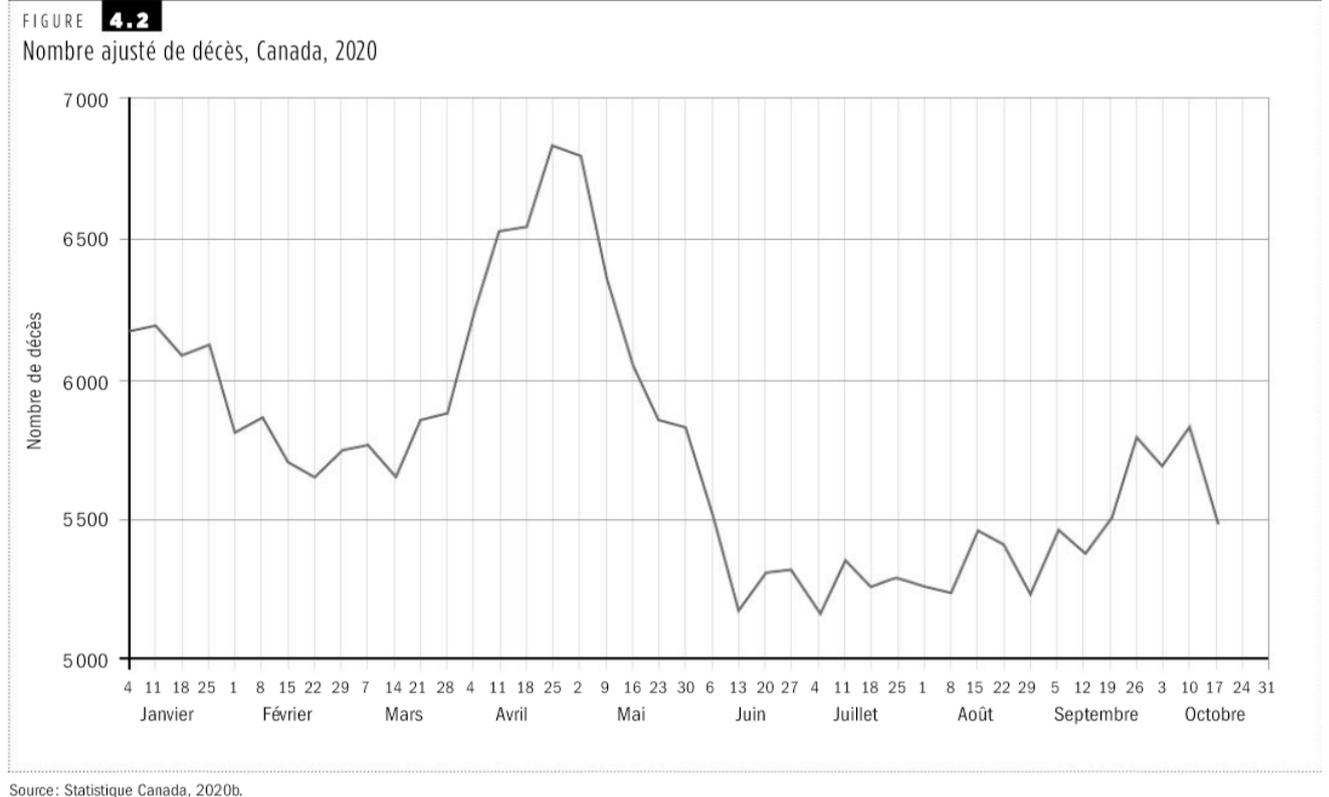


Figure 6 Graphique linéaire

D'autres types de représentations, tels que les infographies (Figure 6) sont présentés comme des moyens destinés à attirer l'attention du lecteur, en particulier pour explorer ou expliquer un enjeu spécifique. Les tableaux de bord interactifs offrent une exploration approfondie des sous-variables, la modification de la chronologie et l'intégration de cartes et d'autres fonctionnalités pertinentes.

FIGURE 4.4

Répercussions de la COVID-19 sur les entreprises au Canada

LES RÉPERCUSSIONS DE LA COVID-19 SUR LES ENTREPRISES AU CANADA

Résultats de l'Enquête canadienne sur la situation des entreprises

LE TÉLÉTRAVAIL AVANT, PENDANT ET APRÈS LA PANDÉMIE

Pourcentage d'entreprises ayant déclaré que 10 % ou plus de leur effectif faisait du télétravail

Avant le
1^{er} février 2020

16,6 %

Le 29 mai 2020

32,6 %



22,5 %

des entreprises s'attendent à ce que 10 % ou plus de leur effectif continue de faire du télétravail une fois la pandémie terminée.

LA MAJORITÉ DES ENTREPRISES PRÉVOIENT METTRE EN ŒUVRE DES MESURES POUR PROTÉGER LEUR PERSONNEL ET LEUR CLIENTÈLE APRÈS LE RETOUR AU TRAVAIL SUR PLACE

79,0 % ont indiqué qu'il était probable ou très probable qu'elles fournissent des masques, des gants et d'autres types d'équipement de protection individuelle à leur personnel.



91,2 % ont indiqué qu'il était probable ou très probable qu'elles insistent pour que le personnel présentant des signes de maladie reste à la maison.



87,0 % ont indiqué qu'il était probable ou très probable qu'elles accroissent les mesures sanitaires sur le lieu de travail.



ALORS QUE LES MESURES DE DISTANCIATION PHYSIQUE S'ASSOUPLISSENT, 80,5 % DES ENTREPRISES ONT BESOIN OU S'ATTENDENT À AVOIR BESOIN D'ÉQUIPEMENT DE PROTECTION INDIVIDUELLE (EPI)



PRINCIPAUX SECTEURS AYANT BESOIN D'EPI
Soins de santé et assistance sociale (93,0 %)
Services d'hébergement et de restauration (92,9 %)
Commerce de détail (91,9 %)
Transport et entreposage (90,3 %)



PRINCIPAUX BESOINS
Masques et équipement de protection pour les yeux (68,1 %)
Produits de nettoyage (64,4 %)
Gants (56,6 %)

COMMENT LES ENTREPRISES CANADIENNES S'EN TIRENT-ELLES PENDANT LA PANDÉMIE DE COVID-19?

- 63,7 %** ont obtenu du financement auprès du gouvernement ou du crédit auprès d'autres sources
 - 23,6 %** des entreprises qui paient un loyer ou une hypothèque ont obtenu le report de leurs paiements
 - 39,9 %** ont réduit les heures de travail ou les quarts de travail
 - 28,4 %** ont effectué des mises à pied
- Pourcentage d'entreprises

LES ENTREPRISES EN MODE SURVIE

19,3 % des entreprises peuvent continuer de fonctionner à leur niveau actuel de revenus et de dépenses pendant **moins de six mois** avant de devoir envisager d'autres mesures relatives au personnel, la fermeture ou la faillite.

Source: Statistique Canada, mai 2020, Enquête canadienne sur la situation des entreprises: les répercussions de la COVID-19 sur les entreprises au Canada.

© Sa Majesté la Reine du chef du Canada, représentée par le ministre de l'Industrie, 2020

Figure 7 Infographie

Enfin, le chapitre aborde la visualisation des données sous forme d'outils d'information spatiale basés sur des systèmes d'information géographique (SIG) (Figure 7). Ces outils fournissent un profil géographique des régions, permettant une meilleure compréhension des données en précisant leur origine géographique. Des exemples concrets, tels que les infographies de Statistique Canada et les données sur la pandémie de COVID-19 d'ESRI Canada, illustrent comment ces approches contribuent à informer le public et à soutenir les autorités sanitaires dans la prise de mesures appropriées. La construction de ces visualisations implique l'agrégation de données administratives provenant de diverses sources et niveaux de gouvernement.



Figure 8 Outils d'information spatiale

Résumé du chapitre

Le chapitre met en lumière le rôle crucial de la visualisation des données, définie comme un ensemble de technologies et de processus transformant notre perception de la réalité. Il souligne que

cette visualisation procure du sens, influençant ainsi la compréhension des enjeux et des processus décisionnels. L'utilisation conjointe des données administratives et des techniques de visualisation alimente l'échange de connaissances entre le gouvernement et l'administration.

Le processus de visualisation et de représentation des données est présenté comme un ensemble d'activités en pipeline, illustrant la séquence d'actions nécessaires pour créer des visualisations informatives. Par ailleurs, l'intelligence artificielle (IA) est définie comme l'ensemble des technologies facilitant l'extraction d'informations à partir de vastes ensembles de données, permettant ainsi des récurrences et des prédictions basées sur ces informations. Une des étapes cruciales pour l'IA consiste en la collecte, l'organisation et la manipulation des données.

Le chapitre souligne également l'importance pour les institutions publiques de mettre en place un cadre de gouvernance assurant la sécurité des données, notamment dans l'utilisation de l'IA. Il encourage l'adoption de mesures administratives et technologiques visant à garantir une protection adéquate de la vie privée. Enfin, l'application de la dimension éthique des données est abordée, mettant en avant la nécessité d'un questionnement critique, d'une signalisation des problématiques éthiques liées à la vie privée, de l'engagement avec les parties concernées et du développement de solutions socialement acceptables.

Bibliographie

Desroches, P. (2022). *Les données administratives publiques dans l'espace numérique*. Presses de l'université du Québec.