

N° 11-633-X au catalogue — N° 022
ISSN 2371-3437
ISBN 978-0-660-32181-3

Études analytiques : méthodes et références

Littératie des données et comment la mesurer dans la fonction publique

par Aneta Bonikowska, Claudia Sanmartin et Marc Frenette

Date de diffusion : le 14 août 2019

 Statistique
Canada

Statistics
Canada

Canada 

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à STATCAN.infostats-infostats.STATCAN@canada.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|-----------------------------------------------------------------------------|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Programme des services de dépôt

- | | |
|-----------------------------|----------------|
| • Service de renseignements | 1-800-635-7943 |
| • Télécopieur | 1-800-565-7757 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté la Reine du chef du Canada, représentée par le ministre de l'Industrie 2019

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Littératie des données et comment la mesurer dans la fonction publique

par

Aneta Bonikowska, Claudia Sanmartin et Marc Frenette

Division des études analytiques
Statistique Canada

11-633-X
2019XXX
ISSN 2371-3437
ISBN 978-0-660-32181-3

Août 2019

Études analytiques : méthodes et références

Les documents de cette série traitent des méthodes utilisées pour produire des données qui seront employées pour effectuer des études analytiques à Statistique Canada sur l'économie, la santé et la société. Ils ont pour but de renseigner les lecteurs sur les méthodes statistiques, les normes et les définitions utilisées pour élaborer des bases de données à des fins de recherche. Tous les documents de la série ont fait l'objet d'un examen par les pairs et d'une révision institutionnelle, afin de veiller à ce qu'ils soient conformes au mandat de Statistique Canada et qu'ils respectent les normes généralement reconnues régissant les bonnes pratiques professionnelles.

Les documents peuvent être téléchargés gratuitement de www.statcan.gc.ca.

Table des matières

Résumé.....	5
1 Introduction.....	6
2 Méthodologie	7
3 Définition de la littératie des données.....	7
4 Cadres de compétences.....	8
5 Mesure de la littératie des données — outils disponibles	11
5.1 Évaluation des compétences individuelles en littératie des données — autoévaluation	11
5.1.1 Databilities	11
5.1.2 Qlik	12
5.1.3 Bureau de l'administrateur scolaire du comté de Maricopa	12
5.1.4 The Open Data Institute	12
5.2 Évaluation des compétences individuelles en littératie des données — mesures objectives	13
5.2.1 Department of Education des États-Unis.....	13
5.2.2 WestEd	13
5.3 Évaluation de la littératie des données des organisations — autoévaluation	13
5.3.1 Outil d'évaluation du modèle de maturité de la littératie des données	13
5.3.2 Sondage sur la littératie des données de PoliVisu.....	14
5.3.3 Gartner	14
5.3.4 Projet de littératie des données.....	14
6 Mesure de la littératie des données dans la fonction publique.....	15
7 Discussion	16
8 Annexe.....	18
Références	24

Résumé

Le présent rapport donne un aperçu des définitions et des cadres de compétence de la littératie des données, ainsi que des outils d'évaluation utilisés pour la mesurer. Tout cela est fondé sur la littérature existante et les pratiques actuelles dans le monde. La littératie des données, ou la capacité de tirer des renseignements utiles des données, est un concept relativement nouveau. Cependant, on la reconnaît de plus en plus comme un ensemble de compétences essentielles à l'ère de l'information. Les approches existantes à l'égard de la mesure de la littératie des données (des outils d'autoévaluation aux mesures objectives et des évaluations individuelles aux évaluations organisationnelles) sont analysées dans le présent rapport afin d'orienter l'élaboration d'un outil d'évaluation pour la littératie des données au sein de la fonction publique canadienne.

1 Introduction

Les organisations sont de plus en plus nombreuses à découvrir la façon dont les données peuvent améliorer la livraison des produits et services aux clients, les aider à mieux gérer leur effectif et leur fournir des explications utiles à partir de leurs ressources de données. Pour dégager toute cette puissance, les employés doivent posséder les compétences requises pour tirer des explications utiles des données. La littératie des données est souvent décrite comme une compétence essentielle pour le 21^e siècle (Bryla 2018). D'ici 2020, Gartner estime que 80 % des organisations commenceront à déployer des initiatives de littératie des données pour améliorer les compétences de leurs employés (Gartner 2018). Cette réalité touche aussi la fonction publique.

Reconnaissant le pouvoir des données pour favoriser une meilleure prise de décision, concevoir de meilleurs programmes et offrir des services plus efficaces aux citoyens, le gouvernement du Canada a diffusé la Feuille de route pour la Stratégie de données de la fonction publique fédérale (gouvernement du Canada 2018). L'un des obstacles possibles à la concrétisation de cette vision est le niveau de compétences en littératie des données chez les fonctionnaires. Pour mieux comprendre l'écart actuel, une évaluation de l'état actuel de la littératie des données est requise.

Comme première étape vers l'atteinte de cet objectif, le rapport présente une analyse environnementale du concept de littératie des données et des approches permettant de la mesurer. Le concept de littératie des données est encore relativement nouveau. La plus grande partie du travail pour définir systématiquement les compétences pertinentes à la littératie des données a été effectuée au cours des cinq dernières années.

2 Méthodologie

La recherche a été menée en ligne au moyen de Google, Google Scholar et des bases de données de publication disponibles dans la bibliothèque de Statistique Canada. Les termes de recherche comprenaient « littératie des données », « littératie des données secteur public », « littératie des données gouverner* », « enquête sur la littératie des données », « mesure* de la littératie des données », « évaluation de la littératie des données », « outil d'évaluation de la littératie des données », « alphabétisation de données », « littératie de données ». La recherche a couvert les articles publiés dans les revues scientifiques, les produits des conférences, les livres blancs et la littérature grise, et a inclus le matériel des sites Web de diverses organisations à but lucratif et à but non lucratif, complété dans certains cas par des courriels ou des conversations téléphoniques avec des représentants de ces organisations. La recherche a couvert l'information du domaine public présentée principalement en anglais.

3 Définition de la littératie des données

La littératie est historiquement associée à la capacité de lire et d'écrire. Une définition plus complète soulignant les nombreuses dimensions du concept a été élaborée par l'UNESCO en 2003 :

La littératie est la capacité d'identifier, de comprendre, d'interpréter, de créer, de communiquer et d'utiliser du matériel imprimé et écrit, dans des contextes variables. La littératie suppose une continuité de l'apprentissage pour permettre aux individus d'atteindre leurs objectifs, de développer leurs connaissances et leur potentiel et de participer pleinement à la vie de leur communauté et de la société tout entière. (Robinson 2005, p. 13)

Plus récemment, le terme « littératie » a aussi été utilisé plus largement pour décrire la compétence dans un domaine particulier, comme la littératie statistique, la littératie informatique et la littératie financière. De même, la plupart des définitions connues de la littératie des données renvoient à la compétence dans les nombreuses dimensions de l'interaction avec les données. Plus précisément, la littératie des données est la « capacité de dériver de l'information utile et significative des données » (Sperry 2018). Certains auteurs voient la littératie des données selon son chevauchement avec d'autres types de littératie, comme la littératie statistique, la littératie en matière d'information et la littératie numérique (Wolff et coll. 2016; Prado et Marzal 2013; Schield 2004; Bhargava et coll. 2015). Toutefois, il se pourrait que les définitions existantes deviennent inadéquates avec le temps, tandis que les types de données disponibles évoluent pour devenir plus imposants et plus complexes (Wolff et coll. 2016), et les technologies émergentes telles que l'intelligence artificielle modifient notre façon de concevoir et d'utiliser les données (Bhargava et coll. 2015). C'est pourquoi Bhargava et coll. (2015) affirment qu'au lieu de promouvoir la littératie des données et les autres sous-types de littératie, il faudrait plutôt se concentrer sur la promotion de la « littératie à l'ère des données ».

Après avoir examiné la littérature disponible, Wolff et coll. (2016) ont défini la littératie des données comme suit :

La littératie des données est la capacité de poser des questions concrètes et d'y répondre à partir de vastes et de petits ensembles de données au moyen d'un processus de requête, en tenant compte de l'utilisation éthique des données. Elle repose sur des compétences de base pratiques et créatives, avec une capacité d'étendre les connaissances des compétences de traitement des données spécialisées selon les objectifs. Cela comprend les capacités de sélectionner, de trier, d'analyser, de visualiser, de critiquer et d'interpréter les données, de même que de communiquer

les histoires à partir des données et d'utiliser les données dans le cadre d'un processus de conception (p. 23).

Une revue distincte de la littérature menée par une équipe interdisciplinaire de chercheurs de l'Université de Dalhousie, au Canada (Ridsdale et coll. 2015) a donné lieu à une définition plus concise : « La littératie des données est la capacité de recueillir, de gérer, d'évaluer les données et de les mettre en pratique de manière critique » (p. 2). Les chercheurs notent que la définition devrait pouvoir changer et évoluer avec les commentaires et les observations des intervenants.

À l'aide de la revue effectuée par Ridsdale et coll., *Data to the People* met de l'avant une définition encore plus concise de la littératie des données, où cette dernière est « notre capacité de lire, d'écrire et de comprendre les données, tout comme la littératie est notre capacité de lire, d'écrire et de comprendre notre langue maternelle » (*Data to the People* 2018).

On comprend, dans certaines parties de la littérature, qu'un « alphabète des données » ne devrait pas être une étiquette réservée aux scientifiques et aux spécialistes des données. On devrait concevoir la littératie des données comme « la capacité de non-spécialistes d'utiliser les données » (Frank et coll. 2016, résumé) et elle devrait mesurer « la capacité d'une personne de lire, d'analyser les données, de travailler et de défendre des points de vue avec les données » (Qlik 2018, p. 3), en utilisant probablement des statistiques simples comme des moyennes et des pourcentages.

Pour résumer, une personne qui est alphabète des données serait au minimum capable de comprendre l'information tirée de données et résumée en statistiques simples, de faire des calculs supplémentaires à l'aide de ces statistiques et d'utiliser les statistiques pour éclairer les décisions. Cependant, cette définition dépend du contexte, conclusion qui sera illustrée ci-dessous.

4 Cadres de compétences

La mesure et la formation exigent une feuille de route des compétences pertinentes au concept de littératie des données. Des efforts visant à compiler une liste ou un modèle des compétences ont été entrepris dans le milieu universitaire, particulièrement dans les domaines suivants : bibliothéconomie pour les bibliothécaires en formation; éducation, mathématiques, science informatique et affaires pour les étudiants en enseignement et les enseignants; et les secteurs des organisations non gouvernementales (ONG) et privé pour promouvoir la prise de décision éclairée par les données.

Les cadres de compétences actuellement disponibles présentent des différences pour de nombreuses dimensions. Certains tentent de déterminer des domaines de compétences, en donnant des exemples uniquement des compétences qui pourraient en faire partie. D'autres précisent les compétences à divers niveaux de détails. Certains tentent de préciser les aptitudes qui seraient requises à différents niveaux de maîtrise à l'intérieur d'une compétence donnée, tandis que d'autres ne le font pas. Voici une brève description de certains cadres de compétences. Le tableau 1 résume les compétences couvertes par chaque cadre.

L'une des listes de compétences et d'aptitudes les plus détaillées a été compilée par Ridsdale et coll. (2015) à la suite d'une vaste analyse environnementale sur la littératie des données, qui avait pour but d'éclairer l'éducation sur la littératie des données. Même si aucun niveau de maîtrise n'est précisé, les chercheurs regroupent les compétences en compétences conceptuelles, compétences de base et compétences avancées. Le tableau matriciel complet des connaissances et des tâches associées aux compétences est présenté à l'annexe A.

À l'aide du cadre de Ridsdale et coll., [Data to the People](#) a mis au point une échelle de capacités pour accompagner les compétences, dans le but de créer un outil d'évaluation de la littératie des données. L'échelle consiste en un maximum de six niveaux, tandis que certaines compétences sont évaluées uniquement aux niveaux de capacités supérieurs. La description des aptitudes démontrées à chaque niveau est comparable dans les compétences. Au niveau inférieur (après ne posséder aucune aptitude dans une compétence donnée), une personne peut accomplir certaines tâches avec des conseils et une orientation. Au niveau supérieur, une personne peut aider les autres à accomplir de telles tâches et les enseigner. Le cadre de compétences qui en résulte, appelé [Databilities](#), comprend 15 compétences (un sous-ensemble des compétences compilées par Ridsdale et coll. [2015]), avec un total de 25 sous-catégories.

Wolff et coll. (2016) ont construit un cadre de compétences en littératie des données autour du processus d'enquête « problème, plan, données, analyse et conclusion » (PPDAC), fondé sur les travaux menés sur la littératie des données et la littératie statistique (consulter l'annexe B). Les auteurs ne précisent pas les aptitudes qui pourraient être attendues à différents niveaux de compétence, mais ils identifient quatre types de citoyens qui auraient besoin de différents niveaux de complexité des aptitudes compte tenu de leur interaction attendue avec les données : lecteur, communicateur, fabricant et scientifique.

Grillenberger et Romeike (2018) ont construit leur cadre de compétences selon des travaux menés sur la gestion des données et la science des données. Leur objectif était de documenter l'enseignement de la science informatique dans les écoles et ils décrivent comment cet objectif les a menés à inclure ou exclure des facteurs qui étaient pris en compte dans d'autres articles sur la littératie des données. L'interaction entre les quatre domaines de contenu et les quatre domaines de processus en résultant crée un tableau matriciel de quatre par quatre des compétences. Les auteurs énumèrent des exemples d'aptitudes qui pourraient être prises en compte dans chaque cellule du tableau (consulter l'annexe C). Les niveaux de compétence ne sont pas précisés dans ce modèle; ils le seront dans le cadre de futurs travaux.

Sternkopf (2017) (aussi résumé par Sternkopf et Mueller 2018) a élaboré un modèle de maturité de la littératie des données au niveau organisationnel, en insistant plus particulièrement sur les ONG. Il décrit les étapes par lesquelles une organisation pourrait passer dans sa progression de l'état de non-utilisation des données dans ses activités à l'utilisation des données à des niveaux avancés. À ce titre, non seulement ce modèle identifie les compétences pertinentes liées à la littératie des données, mais il comprend aussi quatre niveaux de capacités pour chacune (consulter l'annexe D).

En dépit des approches différentes, il y a un degré élevé de chevauchement en ce qui a trait aux compétences incluses dans les différents cadres, comme le résume le tableau 1. Ils comprennent tous, à un niveau individuel ou organisationnel, les aptitudes nécessaires pour accéder aux données, les manipuler, évaluer leur qualité, mener des analyses, interpréter les résultats et (dans la majorité des cadres) utiliser les données de manière éthique. Ensemble, les cadres montrent clairement que le concept de littératie des données est complexe et met en cause une variété de compétences avec un continuum de capacités.

Tableau 1
Comparaison de certains cadres de compétences en littératie des données

	Risdale et coll. 2015	Databilities par Data To The People, 2018 ¹	Wolff et coll. 2016	Sternkopf et Mueller 2018	Grillenberger et Romeike 2018
Planifier, mettre en œuvre et surveiller les plans d'action	sans objet	sans objet	Oui	sans objet	sans objet
Entreprendre le processus de requête axée sur les données	sans objet	sans objet	Oui	sans objet	sans objet
Connaissance et compréhension des données, de leurs utilisations et de leurs applications	Oui	sans objet	Oui	sans objet	sans objet
Pensée critique	Oui	sans objet	sans objet	sans objet	Oui ††
Culture des données	Oui	sans objet	sans objet	Oui	sans objet
Éthique des données (p. ex., sécurité, enjeux de protection des renseignements personnels)	Oui	sans objet	Oui	Oui	Oui
Outils de données	Oui	sans objet	Oui	sans objet	Oui
Découverte des données (capacité de trouver les données et d'y accéder)	Oui †	Oui	Oui	Oui	Oui
Collecte de données	Oui †	Oui	Oui	sans objet	Oui
Gestion et organisation des données	Oui	Oui	sans objet	sans objet	Oui
Manipulation des données	Oui	Oui	Oui	Oui	Oui
données et des sources	Oui	Oui	Oui	Oui	Oui
Citation des données	Oui	sans objet	sans objet	sans objet	sans objet
Analyse des données de base	Oui	Oui	Oui	Oui	Oui
Visualisation des données	Oui	Oui	sans objet	Oui	Oui
Présentation des données oralement	Oui	Oui	sans objet	Oui	sans objet
Interprétation des données (comprendre les données)	Oui	Oui	Oui	Oui	Oui
Identification des problèmes en utilisant les données	Oui	Oui	Oui	Oui	sans objet
Prises de décisions fondées sur les données (PDFD)	Oui	Oui	Oui	sans objet	sans objet
Évaluation des décisions et des conclusions fondée sur les données	Oui	Oui	sans objet	sans objet	sans objet
Création et utilisation des métadonnées	Oui	Oui	sans objet	sans objet	sans objet
Durée et réutilisation des données	Oui	Oui	sans objet	sans objet	Oui
Partage de données	Oui	sans objet	sans objet	sans objet	Oui
Conservation des données	Oui	sans objet	sans objet	sans objet	Oui
Conversion des données (d'un format à l'autre)	Oui	Oui	Oui	Oui	sans objet
Formuler des hypothèses	sans objet	sans objet	Oui	sans objet	sans objet
Travailler avec de vastes ensembles de données	sans objet	sans objet	Oui	sans objet	sans objet

† La découverte et la collecte des données sont mentionnées comme une seule compétence dans le cadre de Risdale et coll.

†† Grillenberger et Romeike ne mentionnent pas la pensée critique comme une compétence, mais cela est implicite dans leur cadre.

1. Les sondages sur la littératie des données menés par Data to the People comprennent des compétences supplémentaires adaptées à l'organisation du client. Ce tableau énumère uniquement les compétences qui font partie du cadre Databilities accessible au public.

Note : Adapté par Statistique Canada aux fins d'accessibilité.

Source : Statistique Canada, compilation de l'auteur.

5 Mesure de la littératie des données — outils disponibles

Des outils d'évaluation de la littératie des données mis à la disposition du public (ou annoncés) sont offerts et utilisés principalement par les entreprises du secteur privé et les organisations à but non lucratif qui offrent également de la formation en littératie des données. Ce sont en général des autoévaluations des capacités de littératie des personnes en matière de données, bien que certaines enquêtes comprennent des questions sur la mesure dans laquelle les données sont utilisées dans le processus décisionnel d'une organisation dans son ensemble. Des évaluations plus objectives de la littératie des données ont été utilisées dans le secteur de l'éducation aux États-Unis; elles font appel à des tests fondés sur des scénarios. Ces évaluations correspondent davantage à la façon utilisée par l'Organisation de coopération et de développement économiques (OCDE) pour évaluer la littératie et les compétences en numératie, notamment dans le Programme pour l'évaluation des étudiants internationaux et le Programme pour l'évaluation internationale des compétences des adultes.

5.1 Évaluation des compétences individuelles en littératie des données — autoévaluation

Les personnes peuvent autoévaluer leurs niveaux de maîtrise de différentes compétences en littératie des données au moyen de sondages. Voici quelques outils d'autoévaluation disponibles.

5.1.1 Databilities

Peut-être l'outil d'évaluation le plus complet de la littératie des données d'une personne, il a été mis au point par l'entreprise Data to the People, établie en Australie. Il est fondé sur la cartographie des compétences produite par Risdale et coll. (2015). L'outil, appelé [myDatabilities](#), est un sondage d'autoévaluation individuelle en ligne. Il convient de noter que l'outil d'évaluation mis à la disposition du public sur le site Web de l'entreprise ne mesure pas l'éthique des données, une compétence incluse dans la plupart des modèles de compétences en littératie des données.

Chacune des questions demande aux répondants de sélectionner la réponse qui les décrit le mieux. Les réponses sont présentées en ordre croissant, du moins compétent dans une compétence donnée au plus compétent, et suivent le même format pour chaque compétence. Les questions et réponses prennent la forme suivante :

Parmi les énoncés suivants, quel est celui qui vous décrit le mieux?

- a. Avec des conseils, je peux décrire l'analyse que je dois effectuer.
- b. Je peux élaborer un plan simple d'analyse en utilisant les cadres qui me sont fournis.
- c. Je peux élaborer un plan simple d'analyse.
- d. Je peux élaborer un plan d'analyse afin de mieux comprendre divers problèmes.
- e. Je peux aider les autres à élaborer un plan d'analyse.
- f. Je peux enseigner aux autres et les aider à élaborer un plan d'analyse afin de mieux comprendre divers problèmes.
- g. Aucun de ces énoncés ne me décrit.

L'entreprise mène aussi des évaluations organisationnelles qui, en plus du sondage de base décrit ci-dessus, comprennent des questions supplémentaires adaptées à l'organisation, au ministère ou service et à la description de poste de la personne qui répond au sondage. Une telle évaluation déterminerait les domaines au sein de l'organisation présentant des concentrations élevées et faibles de compétences en littératie des données et déterminerait également les groupes qui profiteraient d'une formation sur les compétences en littératie des données. L'entreprise propose des suggestions quant au type de formation dont pourraient profiter des groupes d'individus au sein de l'organisation qui est évaluée. Data to the People a été retenue pour mener des sondages sur la littératie des données dans diverses entreprises de partout dans le monde, y compris auprès de certains ministères gouvernementaux.

5.1.2 Qlik

Le fabricant de logiciels [Qlik](#) (ainsi que le [Projet de littératie des données](#) dont il fait partie) offre sur son site Web un [sondage](#) de 10 questions accessible au public sur les compétences individuelles en littératie des données. Chaque question comporte quatre catégories de réponse. À première vue, les catégories ne semblent pas représenter une progression graduelle des compétences et, dans certains cas, elles mesurent les sentiments du répondant par rapport aux données plutôt que les compétences directement. Les résultats sont regroupés en quatre niveaux de compétences, avec des titres comme Rêveur de données et Chevalier de données, et donnent une brève description du niveau de compétence et des domaines qui peuvent être améliorés. Des liens vers des formations proposées sont fournis. Voici un exemple de question et de ses catégories de réponse :

À quelle fréquence vous arrive-t-il de vous sentir dépassé par les données?

- a. Assez souvent, avec toutes les demandes que je reçois.
- b. Je ne me laisse pas vraiment déranger par cela. Ce ne sont que des chiffres et des stats.
- c. Tous les mois, quand les factures arrivent.
- d. Jamais... elles sont utiles, pourquoi seraient-elles envahissantes?

L'entreprise a commandé une étude sur la littératie des données parmi les employés et les décideurs d'entreprises de divers pays dans le monde.

5.1.3 Bureau de l'administrateur scolaire du comté de Maricopa

Le Bureau de l'administrateur scolaire du comté de Maricopa a publié un [sondage](#) sur son site Web intitulé « Quel est votre QI de littératie des données? » qui a pour but d'aider les enseignants à évaluer leur niveau de littératie des données. Il comprend sept énoncés à propos de diverses compétences en matière de données et demande aux répondants de noter chaque énoncé sur une échelle de cinq points, 0 signifiant jamais vrai et 4 signifiant toujours vrai. Par exemple, l'un des énoncés se lit comme suit :

Agir : Je prends mes décisions et j'agis en fonction de données.

5.1.4 The Open Data Institute

The Open Data Institute (ODI) est une entreprise à but non lucratif établie au Royaume-Uni qui offre des services de consultation et de formation en développement et en utilisation de données ouvertes. L'entreprise a récemment mené un sondage auprès d'employés gouvernementaux de partout dans le monde, en mettant l'accent sur les compétences que les employés disent vouloir apprendre (au lieu de tenter de mesurer les capacités individuelles dans le domaine de la littératie

des données), sur leur méthode de formation privilégiée (p. ex. en personne, en ligne) et sur les facteurs ayant une incidence sur l'utilisation des données ouvertes dans l'emploi du répondant. On peut trouver les questions du sondage sur le [site Web d'ODI](#).

L'entreprise ODI a aussi élaboré ce qu'elle appelle un [cadre des compétences](#), qui est lié à la littératie des données et qui vise à permettre aux personnes d'évaluer leur progrès en vue de devenir alphabètes des données. Chacune des cases de compétences est liée à une formation appropriée offerte par ODI. Les compétences sont regroupées en quatre catégories afin de décrire le niveau de la personne : explorateur, stratège, praticien et pionnier. Les capacités individuelles sont représentées par des icônes, mais il n'y a aucune description précise des compétences individuelles. Le cadre des compétences et le sondage ne sont pas reliés.

5.2 Évaluation des compétences individuelles en littératie des données — mesures objectives

Une évaluation pratique est une approche de rechange pour mesurer les compétences individuelles en littératie des données.

5.2.1 Department of Education des États-Unis

Une étude de la littératie des données auprès des enseignants aux États-Unis, commandée par le Department of Education des États-Unis, a évalué un échantillon d'enseignants dans le cadre d'interviews sur place sur leurs compétences en littératie des données et leur façon de concevoir les données (Means et coll. 2011). Les enseignants, individuellement ou en petits groupes, ont été soumis à des scénarios hypothétiques à propos de leurs étudiants et des données présentées dans un tableau ou un diagramme. Les questions leur demandaient de trouver l'information appropriée dans les données et d'effectuer des calculs simples. Leurs réponses complètes, y compris les étapes pour y arriver, ont été enregistrées et analysées. Le rapport énumère les scénarios utilisés ainsi que les compétences mises à l'essai dans chaque scénario. Les chercheurs laissent supposer que des scénarios comme ceux-ci pourraient être utilisés dans le cadre d'une formation en littératie des données pour les enseignants.

5.2.2 WestEd

Une organisation à but non lucratif aux États-Unis appelée [WestEd](#) a élaboré quatre évaluations fondées sur des scénarios pour mesurer les compétences en littératie des données des enseignants.

5.3 Évaluation de la littératie des données des organisations — autoévaluation

En plus des évaluations individuelles, il est aussi possible de mener des évaluations organisationnelles. Ces dernières servent à déterminer la mesure dans laquelle les données sont utilisées dans le cadre des activités quotidiennes au sein de l'organisation dans son ensemble, c.-à-d. la culture entourant l'utilisation des données en plus de la littératie des données des employés. À l'échelle de l'organisation, il est difficile de déterminer qui répond au sondage et comment la personne résume les compétences en littératie des employés.

5.3.1 Outil d'évaluation du modèle de maturité de la littératie des données

Le modèle de maturité de la littératie des données élaboré par Sternkopf (2017) et résumé par Sternkopf et Mueller (2018), comme décrit dans la section précédente, est accompagné d'un [outil d'évaluation](#) accessible sur le site Web de l'organisation Datenschule [École des données], située en Allemagne. L'outil comporte des questions à propos de neuf compétences. Il y a quatre

réponses à chaque question, décrivant quatre niveaux de compétence, et le répondant doit en noter un. Voici un exemple de question et de ses catégories de réponse :

Comment évaluez-vous votre capacité de formuler des questions afin de trouver des réponses utiles dans les données?

- a. Capacité de poser des questions auxquelles il est possible de répondre à l'aide de simples requêtes de données.
- b. Aucun sentiment à l'égard des questions auxquelles il est possible de répondre à l'aide de données.
- c. Des projets entiers sont fondés sur des questions multidimensionnelles qui exigent des requêtes de données complexes et de multiples itérations pour y répondre.
- d. Des projets entiers sont fondés sur des questions multidimensionnelles qui exigent des requêtes de données complexes et répondent à toutes les questions principales et sous-questions. Aucune personne à elle seule ne peut gérer ces demandes.

Cet outil d'évaluation a été utilisé par Datenschule dans le cadre de ses travaux avec des ONG et des administrations locales.

5.3.2 Sondage sur la littératie des données de PoliVisu

Fondé par le programme de recherche et innovation Horizon 2020 de l'Union européenne, ce sondage a recueilli des données probantes sur la connaissance et l'utilisation des mégadonnées par les autorités publiques de l'Union européenne. Il s'adressait principalement aux organisations qui œuvraient dans le domaine de la mobilité et du transport. Les questions du sondage demandaient si des mégadonnées étaient utilisées ou produites au sein de l'organisation, à quelles fins elles étaient principalement utilisées et qui, au sein de l'organisation, les utilisaient. Le sondage comprenait une question sur le niveau de compétence, sur une échelle de 1 à 8, concernant des compétences telles que l'analyse ou l'interprétation des données. Il convient de noter que les liens vers le sondage en ligne ne sont plus disponibles.

5.3.3 Gartner

L'entreprise d'experts-conseils Gartner annonce sa « Trousse à outils : pour faire de la littératie des données et de l'information une langue seconde », que l'on peut se procurer auprès de l'entreprise. Elle semble inclure un outil d'évaluation de la littératie des données au sein d'une organisation.

5.3.4 Projet de littératie des données

Le Projet de littératie des données est une collaboration entre de nombreuses firmes d'experts-conseils, y compris Data to the People et Qlik. Qlik, qui met au point des logiciels d'informatique décisionnelle et de visualisation des données, et a commandé un système de mesure appelé le résultat de la littératie des données de l'entreprise (LDE). Cet indice a été élaboré en collaboration avec des universitaires de la Wharton School de la University of Pennsylvania. Les entreprises sont évaluées selon trois dimensions : les compétences des employés en littératie des données, la prise de décisions fondées sur les données et la circulation des compétences en matière de données. La note totale varie de 0 à 100. Qlik a recueilli les données d'entreprises privées œuvrant dans une variété de secteurs industriels de partout dans le monde et a démontré dans une analyse subséquente que les notes supérieures à l'évaluation de la LDE étaient associées à une valeur supérieure de l'entreprise (Qlik 2018). Toutefois, une main-d'œuvre alphabète des données ne s'accompagnait pas toujours d'une utilisation efficace des données au sein de

l'entreprise (Qlik 2018). Le Projet de littératie des données vise à créer un outil d'évaluation permettant aux entreprises de mesurer leur littératie en matière de données par rapport au résultat de la LDE et d'offrir l'outil sur le site Web du Projet de littératie des données en 2019.

6 Mesure de la littératie des données dans la fonction publique

L'OCDE place la littératie des données parmi les six compétences essentielles à l'innovation dans la fonction publique (OCDE 2017). Son rapport établit une distinction claire entre les spécialistes et les non-spécialistes des données et met en évidence la nécessité pour les non-spécialistes de devenir alphabètes des données. Il présente également une échelle de progression de la littératie des données dans quatre domaines : (a) les données utilisées dans la prise de décision, (b) les données utilisées pour gérer des services publics, (c) la capacité des non-spécialistes des données d'engager des spécialistes des données, et (d) la capacité des spécialistes des données de communiquer efficacement avec les non-spécialistes à propos des données et des résultats d'analyses.

Sur le plan international, la littératie des données dans la fonction publique est encore un concept nouveau, comme on a pu le constater dans les documents accessibles au public. La fonction publique australienne possède une stratégie bien articulée sur la littératie des données (Gouvernement australien 2016), mais ne semble pas avoir mené une évaluation des niveaux de littératie des données de ses employés. Le gouvernement du R.-U. semble aussi s'intéresser à la littératie des données, et son Service numérique a élaboré un programme de littératie des données à l'intention des fonctionnaires (Cattell 2016; Duhaney 2018). Le Bureau de la statistique nationale (BSN) travaillait aussi à l'élaboration d'une échelle des capacités de littératie des données similaire à l'échelle des capacités de littératie numérique utilisée au sein du gouvernement. Un billet de blogue du BSN en avril 2018 présentait une ébauche du cadre des compétences (Knight 2018). Cependant, il ne semble pas y avoir de modèle pour mesurer la littératie des données dans l'ensemble d'une fonction publique.

7 Discussion

La littératie des données est un concept complexe. Selon certains des cadres de compétences décrits dans le présent rapport, il pourrait sembler qu'une personne doit devenir experte dans toutes les compétences requises pour extraire et utiliser l'information contenue dans les données afin d'être qualifiée d'alphabète des données. Cela pourrait être un résultat attendu pour certaines professions, pour les diplômés de certains cours ou programmes, ou pour des équipes. Cependant, au sein d'une organisation (ou plus précisément, dans le cadre du présent rapport, au sein de l'ensemble de la fonction publique), il est plus raisonnable de présumer que la gestion de données spécialisées et les outils d'analyse continueront d'être le domaine d'un groupe relativement restreint d'emplois, si ce n'est qu'en raison de la formation, de l'expérience et de l'amélioration périodique des compétences requises pour accomplir de telles tâches avec compétence. L'objectif d'améliorer la littératie des données dans la fonction publique pourrait plutôt ressembler davantage aux lignes directrices précisées par l'OCDE (2017), soit de bâtir une culture de prise de décisions et de prestation de services éclairée par des données et établir des relations de travail efficaces entre les spécialistes des données et les non-spécialistes. Cela exigerait à tout le moins que les non-spécialistes soient en mesure d'apprécier la valeur des données dans la prise de décisions et qu'ils aient une compréhension intuitive de la manière d'interpréter les données. Bien entendu, cela exigerait aussi des compétences avancées en littératie des données chez les spécialistes.

Pour ce qui est de mesurer la littératie des données dans la fonction publique, cela laisse supposer qu'une évaluation approfondie de toutes les facettes de la littératie des données requises par les spécialistes et les non-spécialistes serait idéale. La plupart des outils d'évaluation utilisés sont subjectifs; c'est-à-dire, les répondants évaluent leurs propres compétences en littératie des données ou leurs attitudes à l'égard des données. Tandis qu'il pourrait être raisonnable de présumer qu'une autoévaluation des attitudes à l'égard des données donne des résultats fiables, il se pourrait qu'il en soit autrement dans le cas d'une autoévaluation des compétences en littératie des données. Bien qu'étant moins coûteuses que les évaluations objectives, les autoévaluations pourraient produire un portrait déformé de la répartition réelle des compétences. En effet, des études antérieures ont démontré que les personnes présentant un faible rendement surestiment grandement leur rendement dans des tâches intellectuelles (Ehrlinger et coll. 2008). La déclaration des compétences en informatique diffère aussi selon le sexe. Hargittai et Shafer (2006) ont montré que les femmes ont tendance à sous-évaluer leurs capacités à trouver de l'information en ligne par rapport aux hommes. Cependant, dans une évaluation objective, les compétences démontrées des hommes et des femmes n'étaient pas statistiquement différenciables les unes des autres.

Il se pourrait que certaines des compétences requises pour une littératie des données ne soient pas faciles à mettre à l'essai dans une évaluation objective. Les tâches telles que lire et convertir des données avec un logiciel ou produire des outils de visualisation des données dynamiques pourraient tout simplement être trop coûteuses à évaluer objectivement et peut-être trop onéreuses pour les répondants (sauf dans le contexte de concours pour des emplois). Dans ce cas, une évaluation subjective pourrait bien être l'option la plus réaliste. Bien entendu, il conviendrait de garder à l'esprit les biais décrits ci-dessus lors de la mesure des compétences autoévaluées à un certain moment. Si, en revanche, l'objectif est de suivre l'évolution de ces compétences dans la fonction publique, il est possible (mais pas nécessairement le cas) que les biais demeurent cohérents au fil du temps. Dans ce cas, l'autoévaluation pourrait être une approche pratique pour suivre les tendances à l'égard de la littératie des données au fil du temps.

En revanche, l'évaluation de la compréhension des données et des statistiques pourrait se prêter à une évaluation objective. L'évaluation n'a pas à être longue (l'Enquête canadienne sur les capacités financières évalue la littératie financière des Canadiens à l'aide d'un instrument de 14 questions). En principe, un outil similaire pourrait être élaboré pour évaluer la littératie statistique de base des fonctionnaires. Cela serait particulièrement utile pour évaluer les compétences des consommateurs de données (les fonctionnaires qui se fient à la lecture de rapports statistiques ou d'études de recherche liés à leur programme). Ces travailleurs pourraient simplement avoir besoin de savoir comment interpréter les résultats plutôt que comment les produire.

8 Annexe

Annexe A

Cadre de compétences en littératie des données élaboré par Ridsdale et coll. (2015)

Domaine clé de capacité/connaissance	Compétence	Connaissance/tâches		
Cadre conceptuel	Présentation des données	Connaissance et compréhension des données	Connaissance et compréhension des utilisations et applications des données	
Collecte des données	Découverte et collecte des données	Effectuer une exploration des données	Identifie les données utiles	Recueille des données
	Évaluation et assurance de la qualité des données et des sources	Évalue de façon critique les sources de données pour en garantir la fiabilité	Évalue de façon critique la qualité des ensembles de données pour y détecter les erreurs ou les problèmes	
Gestion des données	Organisation des données	Connaissance des méthodes et outils d'organisation de base des données	Évalue les exigences en organisation des données	Organise les données
	Manipulation des données	Évalue les méthodes pour nettoyer les données	Détecte les valeurs aberrantes et les anomalies	Nettoie les données
	Conversion des données (d'un format à l'autre)	Connaissance de différents types de données et de différentes méthodes de conversion	Convertit les données d'un type de format ou de fichier à l'autre	
	Création et utilisation des métadonnées	Crée des descripteurs de métadonnées	Attribue les descripteurs de métadonnée appropriés aux ensembles de données d'origine	
	Conservation, sécurité et réutilisation des données	Évalue les exigences de conservation des données (p. ex., échéancier de conservation, stockage, accessibilité, conditions de partage, etc.)	Évalue les exigences en matière de sécurité des données (p. ex., accès restreint, répertoires protégés, etc.)	Conserve les données
	Préservation des données	Évalue les exigences en matière de préservation	Évalue les méthodes et les outils de préservation des données	Préserve les données

Note : Adapté par Statistique Canada aux fins d'accessibilité.

Source : Ridsdale et coll. (2015), p. 38 Annexe 1 – Tableau matriciel des compétences en littératie des données

LÉGENDE :

Compétences conceptuelles	Compétences de base	Compétences avancées
---------------------------	---------------------	----------------------

Annexe A

Cadre de compétences en littératie des données élaboré par Ridsdale et coll. (2015) (suite)

Domaine clé de capacité/connaissance	Compétence	Connaissance/tâches				
Évaluation des données	Outils de données	Connaissance des outils et des techniques d'analyse des données	Sélectionne l'outil ou la technique d'analyse des données approprié	Met en pratique les outils et techniques d'analyse des données		
	Analyse des données de base	Élabore des plans d'analyse	Met en pratique des méthodes et outils d'analyse	Mène des analyses exploratoires	Évalue des résultats d'analyse	Compare les résultats d'analyse avec d'autres résultats
	Interprétation des données (comprendre les données)	Lit et comprend des graphiques, des tableaux et des diagrammes	Détermine les principaux points à retenir et les intègre à d'autres renseignements importants	Détecte les divergences à l'intérieur des données		
	Identification des problèmes en utilisant les données	Utilise les données pour identifier les problèmes dans des situations pratiques (p. ex., efficacité en milieu de travail)	Utilise les données pour identifier les problèmes de niveau plus élevé (p. ex., politique, environnement, expérimentation scientifique, marketing, économique, etc.)			
	Visualisation des données	Crée des tableaux utiles pour organiser et présenter visuellement les données	Crée des représentations graphiques utiles des données	Évalue l'efficacité des représentations graphiques	Évalue avec un esprit critique les représentations graphiques aux fins d'exactitude et de représentation fidèle des données	
	Présentation des données (oralement)	Évalue le(s) résultat(s) souhaité(s) pour présenter les données	Évalue les besoins du public cible et sa connaissance du sujet	Planifie la réunion ou le type de présentation appropriés	Utilise des tableaux et des visualisations utiles pour communiquer les données	Présente les arguments ou les résultats clairement et de façon cohérente
	Prises de décisions fondées sur les données (PDFD) (Prendre des décisions en fonction des données)	Accorde la priorité à l'information recueillie à partir des données	Convertit les données en renseignements permettant de prendre des mesures	Évalue le mérite et les répercussions des solutions et décisions possibles	Met en œuvre les décisions et solutions	
Application des données	Pensée critique	Au courant des enjeux et des défis de haut niveau associés aux données	Réfléchit de façon critique lorsqu'il travaille avec les données			
	Culture des données	Reconnaît l'importance des données	Appuie un environnement qui favorise l'utilisation critique des données pour l'apprentissage, la recherche et la prise de décisions			
	Éthique des données	Au courant des enjeux légaux et éthiques associés aux données	Applique les données et travaille avec ces dernières de façon éthique			
	Citation des données	Connaissance des méthodes de citation des données largement utilisées	Crée de bonnes citations pour les ensembles de données secondaires			
	Partage de données	Évalue les méthodes et les plateformes de partage des données	Partage les données de façon légale et éthique			
	Évaluation des décisions fondée sur les données	Recueille des données de suivi pour évaluer l'efficacité des décisions ou des solutions fondées sur les données	Mène des analyses des données de suivi	Compare les résultats d'analyse avec d'autres résultats	Évalue les décisions ou les solutions fondées sur les données	Retient les conclusions ou décisions d'origine, ou met en œuvre de nouvelles décisions/solutions

Note : Adapté par Statistique Canada aux fins d'accessibilité.

Source : Ridsdale et coll. (2015), p. 38 Annexe 1 – Tableau matriciel des compétences en littératie des données

LÉGENDE :

Compétences conceptuelles	Compétences de base	Compétences avancées
---------------------------	---------------------	----------------------

Annexe B

Cadre de compétences en littératie des données compilé par Wolff et coll. (2016)

Compétences	Compétence de base	PPDAC
Processus de requête		
Planifier, mettre en œuvre et surveiller les plans d'action	Sans objet	Sans objet
Entreprendre le processus de requête axée sur les données	Sans objet	Sans objet
Connaissance de base		
Comprendre l'éthique d'utiliser les données	Éthique	Sans objet
Utiliser les données pour résoudre des problèmes (réels)	Contexte de résolution de problèmes dans un monde réel	Sans objet
Comprendre le rôle et l'impact des données dans la société dans différents contextes		
Déterminer les problèmes ou les questions qui peuvent être solutionnés à l'aide des données	Poser des questions à partir des données	Problème
Formuler des hypothèses, Identifier les données	Formuler des hypothèses et identifier les sources de données potentielles	Plan
Recueillir ou acquérir des données, <i>Critiquer les données</i>	Recueillir ou acquérir des données	Données
Transformer les données en information et en dernier ressort, en connaissances permettant de prendre des mesures, Créer des explications à partir des données, <i>Accéder aux données, Analyser les données, Comprendre les types de données, Convertir les données, Préparer les données aux fins d'analyse, Combiner les données quantitatives et qualitatives, Utiliser les outils appropriés, Travailler avec de vastes ensembles de données, Résumer les données</i>	Analyser les données et créer des explications à partir de ces données	Analyse
Interpréter l'information dérivée des ensembles de données, Critiquer les interprétations des données présentées	Évaluer la validité des explications selon les données et formuler de nouvelles questions	Conclusion

Note : PPDAC signifie Problème, Plan, Données, Analyse et Conclusion. Adapté par Statistique Canada aux fins d'accessibilité.

Source : Wolff et coll. (2016), p. 13-14. Tableau 1. Catégorisation des compétences en littératie des données à travers de multiples définitions.

Annexe C

Cadre de compétences en littératie des données et exemples d'aptitudes correspondant aux compétences élaborés par Grillenberger et Romeike (2018)

	P1 Collecte, modélisation et nettoyage	P2 Mise en œuvre et optimisation	P3 Analyse, visualisation et interprétation	P4 Partage, archivage et effacement
C1 Information sur les données	<ul style="list-style-type: none"> - Choisir les capteurs appropriés pour recueillir l'information souhaitée comme données - Structurer les données recueillies de manière appropriée aux fins d'analyse subséquente - Évaluer si les données saisies représentent correctement l'information initiale 	<ul style="list-style-type: none"> - Mettre en œuvre des algorithmes pour recueillir les données souhaitées - Mettre en œuvre des algorithmes simples pour télécharger les données à partir des applications Web - Discuter des optimisations et des limites de la collecte de données 	<ul style="list-style-type: none"> - Combiner les données pour obtenir de nouveaux renseignements - Insister sur l'information souhaitée dans la visualisation - Interpréter les données et les résultats d'analyse pour obtenir de nouveaux renseignements 	<ul style="list-style-type: none"> - Décider de partager les données initiales ou pas - Déterminer les données initiales à stocker pour conserver l'information requise - Déterminer une manière appropriée de supprimer des données particulières
C2 Stockage des données et accès	<ul style="list-style-type: none"> - Sélectionner un modèle de données approprié - Structurer les données recueillies de manière appropriée aux fins de stockage - Visualiser les modèles de données de manière appropriée 	<ul style="list-style-type: none"> - Déterminer un stockage approprié des données et stocker les données - Utiliser les possibilités de permettre un accès efficace aux données - Améliorer l'efficacité de stockage en utilisant la compression 	<ul style="list-style-type: none"> - Accéder aux données de manière appropriée aux fins d'analyse - Utiliser des formats de données appropriés pour les données à analyser - Stocker leurs résultats d'analyse de manière appropriée 	<ul style="list-style-type: none"> - Déterminer à qui donner accès aux données stockées - Déterminer les droits d'accès aux données - Discuter des enjeux liés à la validité des données lors de l'effacement des données
C3 Analyse des données	<ul style="list-style-type: none"> - Déterminer si des données particulières influencent les résultats d'analyse - Structurer les données de manière appropriée aux fins d'analyse - Relier les données provenant de différentes sources aux fins d'analyse 	<ul style="list-style-type: none"> - Mettre en œuvre des algorithmes d'analyse simples - Déterminer des vis de réglage aux fins d'analyse - Optimiser les analyses de données afin d'obtenir des résultats de meilleure qualité 	<ul style="list-style-type: none"> - Choisir les méthodes d'analyse appropriées - Visualiser les données et les résultats d'analyse - Interpréter les résultats d'analyse 	<ul style="list-style-type: none"> - Décider des résultats d'analyse à partager et avec qui - Réfléchir à la nécessité de stocker les données initiales après les avoir analysées - Déterminer s'il est raisonnable de partager l'information à propos du processus d'analyse
C4 Éthique et protection des données	<ul style="list-style-type: none"> - Tenir compte des enjeux éthiques lors de la cueillette d'information - Déterminer s'il est raisonnable de combiner différentes sources de données dans des contextes particuliers - Discuter des répercussions sur la vie privée lors de la saisie continue des données 	<ul style="list-style-type: none"> - Discuter de la façon d'anonymiser adéquatement les données - Exclure les données du stockage permanent en fonction de considérations d'ordre éthique - Choisir les droits d'accès aux données en fonction des enjeux liés à la protection de la vie privée 	<ul style="list-style-type: none"> - Discuter des répercussions éthiques des analyses de données menées et de leurs résultats - Déterminer si les résultats d'analyse sont suffisamment anonymisés - Déterminer si l'analyse de données particulières soulève des enjeux liés à la protection de la vie privée 	<ul style="list-style-type: none"> - Déterminer si le stockage de données pour utilisations futures devrait être permis d'un point de vue éthique - Trouver des façons appropriées d'effacer les données initiales et les résultats d'analyse de façon sécuritaire - Trouver des façons de retirer de façon appropriée les attributs qui entraînent des enjeux de protection de la vie privée

Note : Adapté par Statistique Canada aux fins d'accessibilité.

Source : Grillenberger et Romeike (2018), 8^e page. Tableau 1 : Tableau matriciel de compétences exemplaires pour les différentes combinaisons de processus (P1–P4) et les différents domaines de contenu (C1–C4).

Annexe D

Cadre de compétences en littératie des données élaboré par Sternkopf (2017)

	1 Incertitude	2 Révélation	3 Certitude	4 Facilité avec les données
	Les organisations ne sont pas au courant de la nécessité d'avoir des compétences en littératie des données et comprennent très peu ou pas du tout ce qui est requis. Les personnes pourraient manifester un certain intérêt pour les données et le travail informatique, mais ne connaissent pas vraiment les différentes étapes qui existent lorsqu'elles travaillent avec les données.	Les organisations expérimentent la mise en pratique de sujets liés aux données. Décrit un état où beaucoup d'information à propos des données a déjà été comprise théoriquement, mais cette information ne peut pas être mise en pratique dans de nombreuses situations et une formation supplémentaire est requise.	Les organisations accomplissent les étapes de manipulation des données avec confiance et ont des activités axées sur les données qui sont intégrées à leurs processus de routine lorsque cela est logique. Les procédures et normes génériques sur la manière de manipuler les données sont officialisées et largement utilisées, et les avantages sont compris à tous les échelons de l'organisation.	Les organisations ont établi une culture éclairée par les données à travers tous les échelons. Les données sont utilisées activement pour améliorer les processus et créer des flux de production.
Culture des données	Les données sont perçues comme un terme ambigu qui entraîne des insécurités.	Les données sont perçues comme un concept intéressant et les avantages sont appréciés. Des insécurités demeurent quant aux cas d'utilisation et aux résultats à attendre.	Les données ne sont pas perçues comme une source d'insécurité, mais plutôt comprises comme un outil permettant de progresser et d'appuyer les activités existantes et planifiées. La haute direction et les chefs de file appuient les initiatives en matière de données.	Les barrières psychologiques des données ont été abaissées (p. ex., insécurités, crainte, résignation) et l'aisance entourant les données est encouragée. Les questionnaires de niveau supérieur et les gestionnaires de projet comprennent et appuient l'importance des ressources consacrées aux échanciers, au budget et aux ressources humaines dans le cadre de la manipulation et de la conversion des données.
Éthique et sécurité des données	Aucune connaissance des directives qui assurent la confidentialité, l'intégrité et la disponibilité des données.	Améliorer la sensibilisation et tentatives non coordonnées de promouvoir l'importance de l'utilisation responsable des données. Absence de directives définies.	Connaissance des répercussions de l'utilisation des données. Les directives sur la manipulation responsable des données sont définies et incorporées aux activités.	Des processus sont en place pour assurer la confidentialité, l'intégrité et la disponibilité des données. Seules les données qui sont nécessaires sont recueillies et utilisées. Les politiques cohérentes à l'échelle de l'entreprise pour la manipulation sécuritaire et éthique des données sont constamment redéfinies et mises à jour.
Poser des questions/définir	Capacité insuffisante de formuler des questions afin de trouver des réponses utiles dans les données. Indifférence quant aux questions auxquelles il est possible de répondre à l'aide de données.	Des questions peuvent être posées aux données dans un nombre limité de situations, et les réponses sont fournies au moyen de requêtes simples.	Les questions aux données sont formulées précisément et orientées pour trouver des réponses significatives dans la plupart des cas.	Des projets en entier sont fondés sur des questions multidimensionnelles. On peut systématiquement trouver les réponses aux besoins en information dans les données, en raison de la connaissance élevée des questions auxquelles on peut trouver des réponses grâce aux données (aucune interprétation abusive).
Rechercher	Compréhension limitée des sources de données possibles. Utilisation de moteurs de recherche de base pour trouver les données. Aucune expérience de l'identification et de la sélection des sources de données les plus pertinentes.	Connaissance limitée à seulement quelques sources de données. Utilisation avancée des moteurs de recherche. L'utilisation des sources de données internes et les recherches de données auprès des institutions publiques sont de pratique courante.	Vaste compréhension des différentes sources de données et les plus pertinentes peuvent être choisies à partir d'une sélection de sources de données. Connaissance et utilisation de portails de données pour des sujets particuliers.	Compréhension profonde des divers types de sources de données possibles. Des critères d'évaluation pour sélectionner les plus pertinentes à un besoin en information sont formulés. Capacité de détecter lorsqu'un problème ou un besoin donné ne peut être réglé à l'aide des données existantes et connaissance des techniques de recherche pour obtenir de nouvelles données (p. ex., requêtes complexes).
Obtenir	Les données sont dérivées de textes intégraux et utilisées comme base pour de plus amples traitements.	Utilisation de téléchargements et de formats de données comme .csv. Utilisation fréquente de programmes internes pour accéder aux données (p. ex., CRM).	Il est possible d'accéder aux données en utilisant des formats de données plus complexes (p. ex., JSON, XML). Utilisation d'interfaces de protocoles d'applications pour obtenir les données.	Accès aux données à l'aide de méthodes perfectionnées (p. ex., grattages/scripts automatisés des données). Capacité de convertir un format d'entrée en format qui peut être utilisé pour mener des activités de traitement et d'analyse plus poussées.

Note : Adapté par Statistique Canada aux fins d'accessibilité.

Source : Sternkopf et Mueller (2018), p. 5051-5052. Tableau 3. Grille de maturité de la littératie des données.

Annexe D

Cadre de compétences en littératie des données élaboré par Sternkopf (2017) (suite)

	1 Incertitude	2 Révélation	3 Certitude	4 Facilité avec les données
	Les organisations ne sont pas au courant de la nécessité d'avoir des compétences en littératie des données et comprennent très peu ou pas du tout ce qui est requis. Les personnes pourraient manifester un certain intérêt pour les données et le travail informatique, mais ne connaissent pas vraiment les différentes étapes qui existent lorsqu'elles travaillent avec les données.	Les organisations expérimentent la mise en pratique de sujets liés aux données. Décrit un état où beaucoup d'information à propos des données a déjà été comprise théoriquement, mais cette information ne peut pas être mise en pratique dans de nombreuses situations et une formation supplémentaire est requise.	Les organisations accomplissent les étapes de manipulation des données avec confiance et ont des activités axées sur les données qui sont intégrées à leurs processus de routine lorsque cela est logique. Les procédures et normes génériques sur la manière de manipuler les données sont officialisées et largement utilisées, et les avantages sont compris à tous les échelons de l'organisation.	Les organisations ont établi une culture éclairée par les données à travers tous les échelons. Les données sont utilisées activement pour améliorer les processus et créer des flux de production.
Vérifier	Il n'existe aucune évaluation critique des données. Les données ne sont pas mises en doute. Les critères d'évaluation des données ne peuvent pas être décrits.	Vérification critique de mesures simples de la qualité des données.	De multiples couches de vérification des données sont implantées dans les procédures normalisées à travers l'organisation.	Capacité de faire des vérifications de la qualité des données de manière indépendante. Les critères d'évaluation des données relativement à la paternité d'une œuvre, à la méthode utilisée pour obtenir et analyser les données, à la comparabilité et à la qualité sont définis avec précision.
Nettoyer	Aucune sensibilité au fait que certaines données pourraient devoir être vérifiées, nettoyées ou normalisées. Les données sont traitées telles quelles.	Sensibilité au fait que certaines données sont le plus souvent imparfaites. Connaissance de certains critères de qualité des données (p. ex., champs vides, doublons) et correction manuelle des données.	Les enregistrements non valides peuvent être détectés et sont retirés au moyen de programmes qui appuient le nettoyage des données (p. ex., OpenRefine). Connaissance approfondie des critères de qualité des données (p. ex., traitables par machine, champs vides, détection des doublons).	Capacité indépendante de retirer les enregistrements non valides et de traduire toutes les colonnes pour utiliser un ensemble de valeurs au moyen d'un script automatisé. Capacité de combiner différents ensembles de données en un seul tableau, de retirer les enregistrements en double ou d'appliquer toute quantité d'autres normalisations.
Analyser	Graphiques à barres et à secteurs, utilisation simple de tableaux de données et sommaires de base des données.	Capacité à travailler avec des statistiques descriptives de base. Tableaux croisés dynamiques pour agréger l'information, histogrammes et diagrammes à surface.	Capacité à travailler avec des statistiques avancées (p. ex., vue inférentielle des données, régression linéaire, arbres décisionnels).	Suite complète d'outils d'apprentissage automatique (p. ex., catégorisation des résultats, prévision, renforcement, apprentissage par ensemble).
Visualiser	Aucune connaissance de la multiplicité des manières de présenter les données. Aucune compréhension des situations où les visualisations normalisées sont choisies; décisions fondées sur ce qui présente la meilleure apparence (essais et erreurs).	Capacité à trouver des produits précis selon l'information qui doit être représentée (p. ex., dans Excel).	Création de graphiques/tableaux de bord interactifs. On constate toujours des incertitudes avec les données.	Connaissance approfondie des diverses formes qui peuvent être utilisées pour présenter les données (écrite, numérique ou graphique). Des visualisations spécialisées sont programmées et couplées et des tableaux de bord dynamiques qui anticipent les demandes des utilisateurs sont conçus.
Communiquer	Les interprétations des données ne sont pas communiquées ou mises dans un plus vaste contexte.	Capacité limitée à trouver des produits précis. Des textes simples appuient les visualisations statistiques et chiffres clés (p. ex., rapports présentés aux partenaires financiers, bulletins).	Ses projets sont appuyés par des visualisations interactives et des textes plus complexes dans un plus vaste contexte (p. ex., histoire des données, conférences, discussions, mises à jour mensuelles, billets de blogue).	Capacité de résumer et de communiquer de façons adaptées à la nature des données, à leur but et au public cible (p. ex., histoire des données, campagnes axées sur les données, ateliers, conférences, mises à jour mensuelles, billets de blogue, recherche reproductive).
Évaluer et interpréter	Les produits des données sont utilisés sans remettre en question leur exactitude et leur message.	Sensibilisation accrue à l'évaluation critique des produits des données et à l'interprétation des résultats. Insécurités quant à déterminer les éléments qui doivent retenir l'attention.	Les produits et les résultats des données sont interprétés avec confiance et de manière critique. Les critères d'évaluation sont assimilés.	Les produits et les résultats des données sont systématiquement remis en question, l'interprétation va au-delà de l'évidence et l'information est traduite avec succès en connaissances pouvant être mises en œuvre.

Note : Adapté par Statistique Canada aux fins d'accessibilité.

Source : Sternkopf et Mueller (2018), p. 5051-5052. Tableau 3. Grille de maturité de la littératie des données.

Références

Gouvernement australien. 2016. « Data skills and capability in the Australian public service ». *Department of the Prime Minister and Cabinet*. Disponible au lien suivant : <https://www.pmc.gov.au/resource-centre/public-data/data-skills-and-capability-australian-public-service> (consulté le 15 novembre 2018).

Bhargava, R., E. Deahl, E. Letouzé, A. Noonan, D. Sangokoya et N. Shoup. 2015. « Beyond data literacy: Reinventing community engagement and empowerment in the age of data ». *Data-Pop Alliance*. Disponible au lien suivant : <https://datapopalliance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/> (consulté le 14 novembre 2018).

Bryla, M. 2018. « Data literacy: A critical skill for the 21st century ». Tableau Blog. Disponible au lien suivant : <https://www.tableau.com/about/blog/2018/9/data-literacy-critical-skill-21st-century-94221> (consulté le 16 janvier 2019).

Cattell, J.A. 2016. « Planning a data literacy programme in government ». *Public Sector Staff*. 29 janvier. Disponible au lien suivant : <https://jacattell.wordpress.com/2016/01/29/planning-a-data-literacy-programme-in-government/> (consulté le 15 janvier 2019).

Data to the People. 2018. *Databilities: A Data Literacy Competency Framework*. Disponible au lien suivant : <https://www.datatothepeople.org/databilities> (consulté le 23 mai 2019).

Duhaney, D. 2018. « Data literacy—improving conversations about data ». *Data in Government*. 21 février. Disponible au lien suivant : <https://dataingovernment.blog.gov.uk/> (consulté le 15 janvier 2019).

Ehrlinger, J., K. Johnson, M. Banner, D. Dunning et J. Kruger. 2008. « Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent ». *Organizational Behavior and Human Decision Processes* 105 : 98 à 121.

Frank, M., J. Walker, J. Attard et A. Tygel. 2016. « Data literacy: What is it and how can we make it happen? » *The Journal of Community Informatics* 12 (3) : 4 à 8.

Gartner. 2018. « Gartner keynote: Do you speak data? » *Smarter with Gartner*. Disponible au lien suivant : <https://www.gartner.com/smarterwithgartner/gartner-keynote-do-you-speak-data/> (consulté le 16 janvier 2019).

Gouvernement du Canada. 2018. *Rapport au greffier du Conseil Privé : Feuille de route de la Stratégie de données pour la fonction publique fédérale*. Ottawa : Gouvernement du Canada. Disponible au lien suivant : https://www.canada.ca/content/dam/pco-bcp/documents/clk/Data_Strategy_Roadmap_FRA.pdf (consulté le 16 janvier 2019).

Grillenberger, A. et R. Romeike. 2018. « Developing a theoretically founded data literacy competency model ». Dans *Produits du 13^e atelier en enseignement informatique primaire et secondaire*, Potsdam, Allemagne, du 4 au 6 octobre 2018. New York : ACM.

Hargittai, E. et S. Shafer. 2006. « Differences in actual and perceived online skills: The role of gender ». *Social Science Quarterly* 87 (2) : 432 à 448.

Knight, M. 2018. « A data literacy scale? » *ONS Digital: Blog for the @ONSdigital Team*. 5 avril. Disponible au lien suivant : <https://digitalblog.ons.gov.uk/2018/04/05/a-data-literacy-scale/> (consulté le 3 janvier 2019).

Means, B., E. Chen, A. DeBarger et C. Padilla. 2011. *Teachers' Ability to Use Data to Inform Instruction: Challenges and Supports*. Washington, D.C. : Office of Planning, Evaluation and Policy Development, Département de l'Éducation des États-Unis.

OCDE (Organisation de coopération et de développement économique). 2017. « Core skills for public sector innovation. » Dans *Skills for a High Performing Civil Service*. Paris : Publication de l'OCDE.

Prado, J.C. et M.A. Marzal. 2013. « Incorporating data literacy into information literacy programs: Core competencies and contents ». *Libri* 63 (2) : 123 à 134.

Qlik. 2018. *The Data Literacy Index*. Disponible au lien suivant : <https://thedata literacyproject.org/> (consulté le 3 janvier 2019).

Ridsdale, C., J. Rothwell, M. Smit, H. Ali-Hassan, M. Bliemel, D. Irvine, D. Kelley, S. Matwin et B. Wuetherick. 2015. *Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report*. Disponible au lien suivant : <https://dalspace.library.dal.ca/xmlui/handle/10222/64578> (consulté le 15 novembre 2018).

Robinson, C. 2005. *Aspects of Literacy Assessment: Topics and Issues from the UNESCO Expert Meeting*. Paris : UNESCO.

Schild, M. 2004. « Information literacy, statistical literacy and data literacy ». *IASSIST Quarterly* Été/automne : 6 à 11.

Sperry, J., 2018. « Data literacy: Exploring economic data ». *Série de webinaires sur les programmes économiques de 2018*. U.S. Census Bureau.

Sternkopf, H. 2017. *Doing Good with Data: Development of a Maturity Model for Data Literacy in Non-Governmental Organizations*. Mémoire de maîtrise, Berlin School of Economics and Law.

Sternkopf, H. et R.M. Mueller. 2018. « Doing Good with Data: Development of a Maturity Model for Data Literacy in Non-Governmental Organizations ». Dans les *Produits de la 51^e Conférence internationale sur les sciences des systèmes à Hawaï*. Disponible au lien suivant : <http://hdl.handle.net/10125/50519> (consulté le 3 janvier 2019).

Wolff, A., D. Gooch, J.J. Cavero Montaner, U. Rashid et G. Kortuem. 2016. « Creating an understanding of data literacy for a data-driven society ». *The Journal of Community Informatics* 12 (3) : 9 à 26.