



Ressources naturelles  
Canada

Natural Resources  
Canada



# L'INFRASTRUCTURE CANADIENNE DE DONNÉES GÉOSPATIALES PRODUIT D'INFORMATION 43f

## **Guide sur la qualité des données géospaciales**

GéoConnexions  
Intelli3 Inc.

2016



## **Guide sur la qualité des données géospatiales**

Préparé par:

Intelli<sup>3</sup> inc.  
3700 Wilfrid-Hamel, suite 80  
Quebec (Qc), Canada, G1P 2J2

Présenté à:

GéoConnexions  
Ressources naturelles Canada

2015

---

## REMERCIEMENTS

---

Les auteurs, le Dr Yvan Bédard, Sonia Rivest, M. Sc. d'Intelli<sup>3</sup>, et le Dr Marc Gervais, professeur à l'Université Laval à Québec, souhaiteraient remercier Jean Brodeur, de GéoConnexions à Ressources naturelles Canada, qui a supervisé le projet.

---

# Table des matières

<b>REMERCIEMENTS .....</b>	<b>1</b>
<b>Sommaire.....</b>	<b>5</b>
<b>Liste des figures.....</b>	<b>6</b>
<b>Liste des tableaux.....</b>	<b>8</b>
<b>Liste des acronymes .....</b>	<b>9</b>
<b>1. Introduction .....</b>	<b>10</b>
1.1 Pourquoi la qualité des données géospatiales?.....	10
1.2 Qui devrait se soucier de la qualité des données géospatiales? .....	12
1.3 Qualité des données géospatiales et gestion du risque .....	13
1.4 Pourquoi ce guide? .....	14
<b>2. Contexte.....</b>	<b>15</b>
2.1 L’incertitude inhérente des données géospatiales.....	15
2.2 Les perspectives de la qualité des données géospatiales.....	17
2.2.1 Qualité interne .....	17
2.2.2 Qualité externe .....	19
2.2.3 Qualité perçue.....	19
2.2.4 Métaqualité.....	19
2.3 Les domaines de la qualité des données géospatiales .....	20
2.4 La gestion des risques associés à la qualité des données géospatiales.....	22
2.5 La diffusion d’information sur la qualité des données géospatiales et les risques d’utilisation	23
2.6 Les normes sur lesquelles reposent la qualité des données géospatiales et la gestion du risque	24
2.6.1 Normes spécifiques.....	24
2.6.2 Normes génériques.....	26
<b>3. Principes fondamentaux de la gestion du risque et de la qualité des données géospatiales .....</b>	<b>28</b>
3.1 Gestion de la qualité des données géospatiales.....	29

3.2	Évaluation de la qualité des données géospatiales .....	31
3.2.1	Spécification formelle de la qualité des données géospatiales .....	32
3.2.2	Évaluation formelle de la qualité des données géospatiales.....	33
3.2.3	Production de rapports formels sur la qualité des données géospatiales .....	33
3.2.4	Évaluation informelle de la qualité des données géospatiales.....	34
3.3	Gestion du risque .....	34
3.3.1	Processus de gestion du risque.....	35
<b>4.</b>	<b>Mise en pratique de la théorie .....</b>	<b>40</b>
4.1	Mettre en pratique la gestion de la qualité des données géospatiales.....	40
4.1.1	Phase de conception .....	40
4.1.2	Phase de mise en œuvre .....	45
4.1.3	Phase de production .....	46
4.1.4	Phase de livraison .....	48
4.1.5	Phase d'utilisation.....	50
4.2	Mettre en pratique l'évaluation de la qualité des données géospatiales .....	51
4.2.1	Spécification formelle de qualité des données géospatiales.....	52
4.2.2	Évaluation formelle de la qualité des données géospatiales.....	53
4.2.3	Rapports formels sur la qualité des données géospatiales .....	55
4.2.4	Évaluation informelle de la qualité des données géospatiales.....	58
4.3	Gérer les risques de l'utilisation inappropriée des données géospatiales en pratique.....	60
4.3.1	Description du contexte de la gestion du risque .....	60
4.3.2	Évaluation des risques .....	60
4.3.3	Élaboration de réponses au risque (traitement du risque).....	63
4.3.4	Communication des risques.....	64
4.3.5	Surveillance et évaluation des risques.....	65
4.4	Communiquer au sujet de la qualité des données géospatiales et des risques liés à l'utilisation en pratique.....	66
<b>5.</b>	<b>Recommandations .....</b>	<b>67</b>
5.1	Recommandations dans un contexte E-E .....	67
5.2	Recommandations dans un contexte E-C .....	68
5.3	Recommandations dans un contexte C-C.....	69



---

<b>6. Conclusion .....</b>	<b>71</b>
<b>7. Bibliographie.....</b>	<b>72</b>



---

# Sommaire

La production, la distribution et l'utilisation des données géospatiales ont changé de façon spectaculaire au cours de la dernière décennie. D'importants nouveaux acteurs du domaine des TI sont venus rejoindre les rangs des fournisseurs de données de l'industrie traditionnelle et du gouvernement pour faciliter la diffusion et l'intégration des données géospatiales au bénéfice du grand public. Cette situation change la donne en ce qui a trait à l'évaluation de la qualité des données géospatiales et à la gestion du risque liés à l'utilisation inappropriée des données. On passe en effet de plus en plus d'une mentalité axée sur le commerce traditionnel d'entreprise à entreprise (E-E) à une mentalité axée sur le commerce d'entreprise à consommateur (E-C) et sur le commerce de consommateur à consommateur (C-C). Dans ces nouveaux contextes de E-C et de C-C, les utilisateurs typiques de données géospatiales ne saisissent pas la nature incertaine des données géospatiales et tiennent les données numériques pour acquises. Cela entraîne parfois de mauvaises décisions, qui peuvent avoir de grandes incidences sociales, politiques et économiques. Dans les marchés de masse, les consommateurs ont certains droits et les fournisseurs de biens et de services ont des obligations. Les réglementations et les décisions judiciaires indiquent que les consommateurs doivent être informés de la qualité d'un produit ou d'un service ainsi que des risques et des utilisations interdites. On peut trouver cette information dans les manuels d'utilisateur, les garanties et les mises en garde, par exemple.

Le présent guide explique comment gérer la qualité des données géospatiales et les risques entourant leur utilisation à chaque étape du cycle de vie d'un contenu informationnel : conception, mise en œuvre, production, distribution et utilisation. Il explique le processus d'évaluation des données géospatiales tel que décrit dans la norme internationale ISO 19157, Information géographique – Qualité des données. Il présente aussi le cadre général de gestion du risque de la norme internationale ISO 31000, Management du risque – Principes et lignes directrices, ainsi que de nombreux exemples associés à la gestion du risque associés à l'utilisation inappropriée des données géospatiales.

Les recommandations concernant la qualité des données géospatiales et la gestion du risque s'appliquent aux contextes E-E, E-C et C-C.

# Liste des figures

Figure 1. Les niveaux hiérarchiques de l'évaluation de la qualité des données géospatiales (adaptés de Devillers, 2004).....	21
Figure 2. La relation entre la qualité et le risque. ....	22
Figure 3. Cycle d'assurance de la qualité des données géospatiales pour une société œuvrant dans le domaine de la référence spatiale.....	28
Figure 4. Processus de gestion de la qualité des données géospatiales .....	29
Figure 5. Gestion de la qualité à chaque étape du cycle de vie du contenu informationnel .....	29
Figure 6. Méthodes d'évaluation de la qualité des données, selon le contexte .....	31
Figure 7. Processus formel et informel d'évaluation de la qualité des données géospatiales. ....	31
Figure 8. Composants de la qualité des données de la norme ISO 19157 (Organisation internationale de normalisation, 2013) .....	32
Figure 9. Gestion du risque : un équilibre entre les efforts et les impacts négatifs imprévus .....	35
Figure 10. Étapes du processus de gestion du risque de la norme ISO 31000 .....	36
Figure 11 : Schéma d'application de réseau routier enrichi de symboles d'avertissement pour la gestion du risque. ....	41
Figure 12 : Exemple de catalogue d'entités enrichi de renseignements sur la gestion du risque (tiré de Bédard, Chandler, Devillers et Gervais, 2009) .....	42
Figure 13 : Extrait du catalogue d'entités de CanVec+ (tiré de Ressources naturelles Canada, 2014). ....	43
Figure 14 : Exemple de schéma d'orthoimagerie numérique dans (Maitra, 2004). ....	44
Figure 15 : Extrait des spécifications de produit de données de CanVec+ (dans Ressources naturelles Canada, 2014).....	45
Figure 16 : Exemple de dépôt de contrainte (adapté de Normand, 1999). ....	46
Figure 17 : Extrait des métadonnées du produit de données 082C de CanVec+ (dans Ressources naturelles Canada, 2015).....	48
Figure 18 : Exemples de sections d'un manuel de l'utilisateur de données géospatiales (Gervais, 2004). ....	49
Figure 19 : Exemples de symboles pouvant être utilisés dans les avertissements.....	50
Figure 20 : Exemple d'utilisation de symboles pour faciliter la lecture d'un rapport sur la qualité des données. ....	50
Figure 21 : Exemple de (a) renseignements sur la qualité participatifs à l'aide d'un système de cote sur cinq étoiles (Koistinen, 2015), et (b) avertissement lié au contexte pour une recherche spatio-temporelle internationale couvrant une période commençant en 1990 comprenant les trois éléments recommandés par les normes ISO : niveau de risque, nature du problème et mesure à prendre pour résoudre le problème (Gervais, Bédard, Lévesque, Bernier et Devillers, 2009). ....	51
Figure 22 : Exemple de méthode d'échantillonnage régulier (non aléatoire) orienté par le secteur. ....	53



---

Figure 23 : Exemple de métadonnées sur la qualité (DQ_CompletenessCommission) tiré des métadonnées 082C de CanVec+ (Ressources naturelles Canada, 2015).....	56
Figure 24 : Résultats détaillés pour l'élément DQ_CompletenessCommission présentés dans un rapport sur la qualité des données (exemple tiré de la norme ISO 19157 (Organisation internationale de normalisation, 2013)). .....	56
Figure 25 : Utilisation d'un diagramme en radar pour comparer des éléments relatifs à la qualité des données d'un jeu de données au niveau de qualité requis. ....	57
Figure 26 : Sommaire d'évaluation de la qualité par catégorie d'entités, tiré d'un rapport sur la qualité des données (Gervais, Bédard, et Larrivée, 2007). ....	57
Figure 27 : Aperçu du tableau de bord pour des métadonnées sur la qualité (Devilleers, Bédard, Jeansoulin et Moulin, 2007).....	58
Figure 28 : Exemple de l'utilisation d'un système de cotation sur cinq étoiles pour coter la représentation d'immeubles dans un globe virtuel (Jones, 2011).....	59
Figure 29 : Exemple de l'utilisation d'un système de cotation sur cinq étoiles dans une infrastructure de données spatiales (Koistinen, 2015). ....	59
Figure 30 : Exemple de description de contexte de gestion du risque.....	60
Figure 31 : Exemple de matrice de classement du risque.....	62

---

# Liste des tableaux

Table 1. Exemples de niveaux hiérarchiques des données géospatiales pour l'évaluation de la qualité.....	22
Tableau 2 : Exemple de liste de risques de l'utilisation inappropriée des données géospatiales dans un contexte agricole (Gira, 2014).....	61
Tableau : 3 : Exemple de résultats d'une analyse des risques (Gira, 2014).....	62
Tableau 4 : Exemple de résultats d'une évaluation des risques (Gira, 2014).....	63
Tableau 5 : Exemple de registre des risques (Gira, 2014).....	66

# Liste des acronymes

Le tableau suivant indique la signification des acronymes employés dans ce document.

Acronyme/abréviation	Signification
E-E	Entreprise à entreprise
E-C	Entreprise à consommateur
C-C	Consommateur à consommateur
CSDGM	Norme du contenu des métadonnées géospatiales numériques
É.-U.	États-Unis d'Amérique
FAQ	Foire aux questions
FGDC	Federal Geographic Data Committee (anglais seulement)
GTQD	Groupe de travail sur la qualité des données
ICDG	Infrastructure canadienne de données géospatiales
IDS	Infrastructure des données spatiales
INSPIRE	Infrastructure d'information spatiale de la Communauté européenne
ISO	Organisation internationale de normalisation
NCCDS	National Committee on Digital Cartographic Data Standards
NSDI	Infrastructure nationale de données spatiales
OGC	Open Geospatial Consortium
PGF	Plateforme géospatiale fédérale
R.-U.	Royaume-Uni
RA	Renseignements d'affaires
SDTS	Norme spatiale de transfert des données
SHC	Service hydrographique du Canada
SNRC	Système national de référence cartographique
TI	Technologies de l'information
VGI	Cartographie participative
VQI	Information participative sur la qualité
XML	Langage de balisage extensible

---

# 1. Introduction

La production, la distribution et l'utilisation des données géospatiales ont changé de façon spectaculaire au cours de la dernière décennie. D'importants nouveaux acteurs du domaine des TI sont venus rejoindre les rangs des fournisseurs de données de l'industrie traditionnelle et du gouvernement pour faciliter la diffusion et l'intégration des données géospatiales au bénéfice du grand public. Grâce aux téléphones intelligents, aux systèmes de navigation, aux nombreuses sources de cartes numériques et d'imagerie et aux globes virtuels, les citoyens consomment et produisent de grandes quantités de données géospatiales. Alors que les années 1950 marquaient le début de l'ère du matériel et les années 1980 le début de l'ère des logiciels, la présente décennie est le début de l'ère des données.

Cette situation change la donne en ce qui a trait à l'évaluation de la qualité des données géospatiales et à la gestion des risques liés à l'utilisation inappropriée des données. On passe en effet de plus en plus d'une mentalité axée sur le commerce traditionnel d'entreprise à entreprise (E-E) à une mentalité axée sur le commerce d'entreprise à consommateur (E-C) et sur le commerce de consommateur à consommateur (C-C). Il y a aujourd'hui beaucoup plus de non-experts et de citoyens qui manipulent des données géospatiales que d'experts. En conséquence, les rôles et responsabilités des producteurs, des distributeurs et des utilisateurs de données et de services géospatiaux évoluent. De la même manière, les infrastructures nationales de données géospatiales s'adaptent aux nouveaux défis que représentent l'évaluation de la qualité des données, la communication de renseignements sur la qualité et la gestion du risque.

---

## 1.1 Pourquoi la qualité des données géospatiales?

Des données sont recueillies afin de répondre à des objectifs bien définis et leur qualité est habituellement définie afin de correspondre à ces objectifs. Il n'est pas aisé de recueillir des données pour répondre à un besoin précis en respectant les budgets et les délais établis compte tenu de la panoplie de méthodes et de technologies qui existe. Les sources, les utilisateurs et les utilisations des données géospatiales sont toutefois de nos jours plus variés que jamais.

Afin de réduire au minimum les coûts et les délais, la réutilisation et l'interopérabilité des données géospatiales sont de plus en plus répandues. Il n'est toutefois pas facile de trouver les données géospatiales qui répondent le mieux à un besoin particulier. Une fois qu'elles ont été trouvées, ces données sont souvent partagées, extraites, croisées et régulièrement transmises en mode sans fil vers des appareils ou utilisées sur demande à des fins différentes de celle pour laquelle elles ont été assemblées au départ.

---

---

Dans les nouveaux contextes de E-C et de C-C, les utilisateurs typiques de données géospatiales ne saisissent pas la nature incertaine des données géospatiales et tiennent les données numériques pour acquises. Cela entraîne de mauvaises décisions qui peuvent avoir de grandes incidences sociales, politiques et économiques (Devilleers, Bédard, et Gervais, 2013). Des journaux du Canada, des États-Unis, de la France, du Royaume-Uni, de l'Australie et d'autres pays ont rapporté des accidents, des blessures, des dommages matériels à des ponts et des résidences, des délais dans les services 911, des personnes perdues dans la nature et des victimes où l'utilisation de données géospatiales était en cause (voir projet GEOIDE (PIV-23), 2012). Les tribunaux rendent des décisions concernant des cartes défectueuses, des renseignements manquants, des mises en garde erronées, des responsabilités, etc. (Gervais, Bédard, Jeansoulin, et Cervelle, 2007), (Chandler, 2010), (Chandler et Levitt, 2011).

En lisant des codes d'éthique et des lignes directrices relatives aux bonnes pratiques de plusieurs professions, on constate qu'ils indiquent habituellement que les professionnels ont le devoir de protéger leurs clients. Dans les marchés de masse, les consommateurs ont certains droits et les fournisseurs de biens et de services ont des obligations. Les réglementations et les décisions judiciaires indiquent que les consommateurs doivent être informés de la qualité d'un produit ou d'un service ainsi que des risques et des utilisations interdites. On peut en trouver des exemples dans les manuels d'utilisateur, les garanties et les mises en garde. Les gouvernements influencent aussi ce qui est permis concernant le respect de la vie privée (voir le blogue <http://www.teresascassa.ca/>). Ces problèmes typiques des marchés de masse mûre s'immiscent dans les contextes des données géospatiales E-C et C-C.

Une récente étude réalisée auprès de l'industrie canadienne de la géomatique (Gervais, Bédard, Larrivée, Rivest et Roy, 2013) révèle que 70 % des répondants croient que les utilisateurs ne sont pas conscients des risques entourant l'utilisation de données géospatiales. Les répondants se disent aussi très préoccupés par la capacité des utilisateurs à gérer les risques et 81 % sont d'avis que l'industrie géospatiale pourrait en faire davantage pour réduire les risques pour les utilisateurs. Selon cette étude, les utilisateurs dénoncent surtout la documentation insuffisante et la qualité des données alors qu'une étude récente sur les besoins réalisée par (Hickling Arthurs Low Corporation, 2011) mentionnait que 15 % des problèmes soulevés par les répondants concernait la qualité des données, ce qui en fait l'un des problèmes les plus préoccupants en géomatique.

Par conséquent, l'Infrastructure canadienne des données géospatiales (ICDG) cherche à aider la communauté géospatiale à faire face au défi que représente la qualité des données puisque la qualité des données touche chaque maillon de la chaîne de service (GéoConnexions, 2013). Il sera donc plus facile de comparer les sources de données, d'analyser l'aptitude à l'emploi, d'estimer les coûts de préparation des données, d'évaluer les risques de l'utilisation inappropriée des données et d'améliorer la communication sur la qualité et les risques aux utilisateurs, en plus de contribuer à établir une communauté de bonnes pratiques.

## 1.2 Qui devrait se soucier de la qualité des données géospatiales?

Pour connaître la qualité des données géospatiales, chaque étape de la chaîne d’approvisionnement de données doit suivre des processus d’assurance et de contrôle de la qualité. Tous les acteurs y contribuent, de l’analyse des besoins et de la conception des bases de données à la collecte, l’intégration, la transformation et la diffusion des données.

L’industrie de la géomatique se soucie depuis longtemps de la qualité des données géospatiales parce qu’il s’agit d’un facteur déterminant pour établir les coûts d’un projet. Dans le contexte E-E, nous voyons souvent des contrats où la qualité est définie selon des spécifications et des responsabilités claires des parties désignées. Les pratiques sont très élaborées et tendent à se concentrer sur la précision spatiale, temporelle, sémantique et des attributs, l’exhaustivité, la cohérence logique, la traçabilité et la conformité aux normes. Ces renseignements sont habituellement communiqués à l’aide de métadonnées. Malgré l’existence de normes, il existe une culture dans le contexte E-E selon laquelle les experts s’adressent aux experts (Devillers, Stein, Bédard, Chrisman, Fisher et Shi, 2010). Dans les contextes E-C et C-C d’aujourd’hui, lorsqu’il est question de qualité des données et de renseignements adéquats pour les utilisateurs, il faut tenir compte du fait que les données (connues ou inconnues) peuvent être redistribuées et réutilisées.

Par conséquent, les acteurs suivants doivent se soucier de la qualité des données géospatiales :

- Tous les concepteurs et les développeurs de systèmes utilisant des données géospatiales, par exemple :
  - les experts des technologies de l’information géospatiale (matériel, logiciel et bases de données)
  - les amateurs qui développent des applications cartographiques basées sur le Web
  - les développeurs d’applications de localisation pour les téléphones intelligents
- Tous les fournisseurs de données géospatiales, par exemple :
  - les producteurs et les distributeurs de données sources
  - les producteurs et les distributeurs d’applications composites de données
  - les personnes qui participent à la production participative (p. ex. les « producteurs-utilisateurs » qui participent à la cartographie participative (VGI))
- Tous les experts qui traitent de données géographiques et diffusent des données géospatiales
- Tous les utilisateurs de données géospatiales :
  - les utilisateurs directs des données (c.-à-d. les principaux utilisateurs d’un système ou d’un jeu de données)
  - les utilisateurs indirects (c.-à-d. les personnes qui réutilisent des données, dans un système ou une application composite de données)
  - les utilisateurs experts ainsi que les non-experts et les citoyens

---

Ceci inclut non seulement l'industrie de la géomatique, mais tous les membres de la communauté géospatiale. Il peut par exemple s'agir d'un informaticien qui développe une carte basée sur le Web, d'un conducteur de voiture qui utilise un système de navigation, d'une personne qui cherche un lieu sur un globe virtuel, d'un randonneur qui utilise un système GPS pour trouver son chemin, d'un expert qui conçoit un système avec OpenStreetMap, d'un propriétaire foncier qui utilise une photographie aérienne sur le Web pour installer ses clôtures, etc. Les données et les services géospatiaux sont devenus des produits de masse et les règles générales en matière de qualité, d'information et de partage des responsabilités s'appliquent.

### 1.3 Qualité des données géospatiales et gestion du risque

Lorsqu'il est question de qualité des données géospatiales, certaines évaluations de cette qualité s'imposent. Le processus d'évaluation peut aboutir à de l'information générale ou détaillée, il peut s'appliquer à un jeu de données complet ou uniquement à un sous-ensemble (p. ex. à une région donnée ou un type de caractéristique), répondre à un seul ou plusieurs besoins, être bien évalué ou seulement perçu, etc. Selon la portée de l'évaluation, il peut être très complexe et robuste ou il peut rester général et comporter un degré d'incertitude plus élevé.

Les concepts et les procédures utilisés pour évaluer la qualité des données géospatiales sont décrits en détail dans la norme internationale ISO 19157 (Organisation internationale de normalisation, 2013). Cette norme est la principale source d'information pour effectuer une évaluation de la qualité des données géospatiales d'une façon communément comprise par les experts et qui permet l'interopérabilité. Pour cette raison, le présent guide s'appuie sur les concepts de cette norme.

La norme ISO 19157 a été élaborée dans un contexte E-E et elle est souple en ce qui concerne le niveau de détail des analyses de la qualité. Qui plus est, elle n'est pas destinée aux contextes E-C et C-C. Dans le contexte E-C, les experts servent des non-experts et ils ont le devoir professionnel de prévoir des stratégies de gestion du risque en vue de réduire les impacts négatifs potentiels pour ces utilisateurs non-experts. Un tel devoir émane du concept de précaution, de protection du consommateur et de responsabilités du code d'éthique des professionnels agréés. Dans les marchés matures, ce pouvoir s'exprime lorsque des personnes qui possèdent des connaissances offrent des produits et des services à d'autres personnes qui n'en possèdent pas ou au grand public. Dans le contexte C-C, les concepts de précaution, de protection du consommateur et de responsabilités comportent aussi la gestion du risque.

La gestion du risque aide à choisir les meilleures stratégies afin de minimiser les risques d'utilisation inappropriée des données géospatiales dans des situations d'incertitude (données incertaines, qualité des données incertaine, utilisations incertaines et expertise incertaine des utilisateurs). Ces stratégies permettent de sensibiliser les utilisateurs et les fournisseurs, de réduire le risque global et contribuent à reconnaître la responsabilité de tous les acteurs. Sans des mesures de gestion du risque, l'évaluation de la qualité des données géospatiales n'est pas complète d'un

---

point de vue social. Les résultats concernant la qualité des données géospatiales doivent également être bien communiqués aux utilisateurs.

Par conséquent, le présent guide explique les principales étapes à suivre pour évaluer la qualité des données géospatiales et gérer les risques potentiels découlant de l'utilisation inappropriée des données. C'est ainsi que les marchés matures peuvent protéger à la fois les consommateurs et les fournisseurs de biens et de services.

## 1.4 Pourquoi ce guide?

Le présent guide vise à aider les organisations et les personnes qui travaillent à procurer des capacités spatiales à la société canadienne dans les contextes E-E, E-C et C-C. Ce guide appuiera en particulier la communauté géospatiale canadienne dans ses efforts en vue d'évaluer la qualité des données à l'aide de normes internationales comme les normes ISO 19157, Qualité des données géospatiales, et ISO 31000, Management du risque.

Une telle approche globale et novatrice devrait encourager davantage l'utilisation généralisée des données géospatiales et l'Infrastructure canadienne des données géospatiales (ICDG). Cette approche devrait aider à repérer les données qui répondent le mieux à un besoin particulier, faciliter l'interopérabilité et stimuler l'adoption des bonnes pratiques observées dans les marchés matures. Enfin, le présent guide devrait permettre d'accroître la participation des acteurs qui se montrent hésitants à l'égard des initiatives de données ouvertes, de l'externalisation ouverte, des applications composites de données et des données géospatiales ubiquistes.

La section 2 du présent guide passe en revue les concepts sur lesquels repose la qualité des données géospatiales. La section 3 porte sur les aspects théoriques de l'évaluation de la qualité des données géospatiales et la gestion des risques associés à l'utilisation inappropriée des données géospatiales en plus de décrire les normes sur lesquelles reposent ces processus. La section 4 décrit les mesures détaillées qu'il faut prendre pour évaluer la qualité d'un jeu de données géospatiales ou d'un service et pour gérer les risques associés et fournit des exemples à l'appui. La section 5 comporte des recommandations pour chacun des contextes (E-E, E-C et C-C) décrits plus haut. Le guide se termine par des conclusions et une liste des documents de référence.



---

## 2. Contexte

L'industrie de la géomatique se soucie depuis longtemps de la qualité des données géospatiales. Il y a eu de plus en plus de travaux réalisés sur le sujet avec l'arrivée, au début des années 1980, de systèmes d'information géographique (SIG) capables d'intégrer des données spatiales et non spatiales (Devilleers, Stein, Bédard, Chrisman, Fisher et Shi, 2010). Quelques années plus tard, la notion « aptitude à l'emploi » pour définir la qualité, d'abord proposée par (Juran, Gryna et Bingham, 1974), a commencé à être utilisée dans la communauté géospatiale (Chrisman, 1983).

En 1987, le National Committee on Digital Cartographic Data Standards des États-Unis (NCDCCDS) s'est penché sur les éléments de la qualité des données géospatiales dans son rapport « A Draft Proposed Standard for Digital Cartographic Data » (National Committee for Digital Cartographic Data Standards, 1987). Ce rapport a servi de base pour la création de la norme spatiale de transfert des données (SDTS). En 1998, le Federal Geographic Data Committee (FGDC) a créé la norme de contenu des métadonnées géospatiales numériques (CSDGM) (Federal Geographic Data Committee, 1998) en complément de la SDTS. C'est au début des années 2000, parallèlement à l'élaboration des infrastructures nationales de données spatiales (NSDI), que les normes ISO/TC 211 sur les principes de qualité et les procédures d'évaluation de la qualité ont été proposées (ISO 19113:2002, Information géographique – Principes qualité, et ISO 19114:2003, Information géographique – Procédures d'évaluation de la qualité respectivement, maintenant remplacées par la norme ISO 19157:2013, Information géographique – Qualité des données). En 2007, l'Open Geospatial Consortium (OGC) a créé le Groupe de travail sur la qualité des données (GTQD), dont la mission consiste à mettre en place un forum permettant de décrire un « cadre ou modèle interopérable pour les mesures d'assurance de la qualité et les services Web de l'OGC pour faciliter l'accès à de l'information géospatiale de grande qualité et son partage, améliorer l'analyse des données et, en définitive, influencer les décisions politiques » (Trakas, 2008), (Open Geospatial Consortium, 2014). En 2012, l'ISO a publié la spécification technique ISO/TS 19158:2012, Information géographique – Assurance qualité relative à l'approvisionnement de données, qui fournit un cadre pour l'assurance de la qualité propre à l'information géographique.

Un aperçu de 30 années de recherche dans le domaine de la qualité des données géospatiales est présenté dans (Devilleers, Stein, Bédard, Chrisman, Fisher et Shi, 2010).

---

### 2.1 L'incertitude inhérente des données géospatiales

La cueillette de données géospatiales comporte un processus de sélection et d'abstraction visant un objectif particulier. Différents objectifs mèneront à des sélections différentes et, par conséquent,

---

à différents jeux de données possédant des attributs et des géométries différents, même s'ils représentent les mêmes entités du monde réel. Les modèles sur lesquels s'appuient les jeux de données géospatiales ne sont que des approximations du monde réel. Il est impossible de produire un modèle parfait de la réalité et c'est pourquoi tous les modèles comportent inévitablement un niveau d'incertitude (Longley, Goodchild, Maguire et Rhind, 2001). Il existe quatre facettes à l'incertitude (Bédard, 1988) :

- L'incertitude conceptuelle correspond au flou entourant l'identification d'une réalité observée (p. ex. le fait qu'il s'agisse ou non d'une rivière aura un impact sur l'existence ou la non-existence d'un objet; le fait qu'il s'agisse d'un ruisseau ou d'une crique aura un impact sur la catégorie de l'objet, sur les attributs à mesurer pour cet objet et sur les spécifications utilisées pour mesurer sa géométrie).
- L'incertitude descriptive correspond à l'incertitude relative aux valeurs de l'attribut d'une réalité observée (c.-à-d. l'imprécision dans les valeurs quantitatives et le flou dans les valeurs qualitatives).
- L'incertitude de localisation correspond au flou dans les valeurs qualitatives utilisées pour des références spatiales ou temporelles (p. ex. le nom d'une ville, le nom d'une rue, un jour de la semaine) et à l'imprécision dans les valeurs quantitatives utilisées pour situer une réalité observée dans l'espace et le temps (p. ex. précision de  $\pm 5$  mètres,  $\pm 1$  jour).
- La méta incertitude correspond au degré selon lequel les incertitudes précédentes sont inconnues (p. ex. ellipses d'erreur avec une probabilité de 95 % qu'il s'agisse bien d'une rivière de catégorie « 1 »).

Les quatre facettes à l'incertitude sont combinées pour établir l'incertitude totale du jeu de données. L'incertitude des données géospatiales peut varier dans l'espace et le temps. L'incertitude des données géospatiales peut être réduite à un certain niveau et l'incertitude résiduelle doit être absorbée. Plus le niveau d'incertitude sera réduit, moins il faudra absorber l'incertitude résiduelle (Bédard, 1988). On peut avoir recours à plusieurs moyens pour réduire ou absorber l'incertitude. La plupart des moyens utilisés pour réduire l'incertitude sont de nature technique alors que la plupart des moyens pour absorber l'incertitude résiduelle sont de nature institutionnelle. Voici certains moyens employés pour réduire l'incertitude : suivre des lignes directrices et des normes, améliorer les méthodes de cueillette de données, suivre de la formation, etc. Voici maintenant certains moyens employés pour absorber l'incertitude résiduelle : fournir une garantie pour le jeu de données, souscrire une assurance pour couvrir les dommages éventuels, établir des contrats ou des clauses mutuellement convenus, etc. La recherche d'un équilibre entre la réduction et l'absorption de l'incertitude constitue un exemple de processus de gestion des risques.

## 2.2 Les perspectives de la qualité des données géospatiales

Il existe plusieurs définitions de la qualité. (Juran, Gryna et Bingham, 1974) ont été les premiers à définir la qualité comme étant « l'aptitude à l'emploi ». Ce qui représente une bonne qualité pour un peut ne pas l'être pour un autre. Cette définition est maintenant largement reconnue et elle est utilisée dans la communauté géospatiale (Chrisman, 1983), (Veregin, 1999). La norme ISO 9000 définit la qualité comme l'ensemble des caractéristiques d'une entité qui lui confèrent l'aptitude à satisfaire des besoins implicites (Organisation internationale de normalisation, 2005). La norme ISO 19157 définit la qualité comme la mesure dans laquelle un ensemble de caractéristiques inhérentes satisfait les exigences (Organisation internationale de normalisation, 2013). Cette norme reconnaît qu'un producteur de données et un utilisateur de données peuvent avoir une perception différente de la qualité selon leur point de vue. Pour un producteur de données, la qualité des données est définie comme la mesure dans laquelle un jeu de données reflète son univers de discours, tel qu'établi dans les spécifications du contenu informationnel. Pour un utilisateur de données, la qualité des données est définie comme la capacité d'un jeu de données à satisfaire ses exigences d'utilisation. La perspective du producteur de données est souvent appelée la *qualité interne* d'un jeu de données ou ses propriétés intrinsèques qui découlent des méthodes de production de données. La perspective de l'utilisateur correspond à la définition de l'aptitude à l'emploi et est souvent appelée la *qualité externe* d'un jeu de données ou le niveau d'aptitude entre les caractéristiques des données et les besoins de l'utilisateur. Il existe autant de types d'évaluations de la qualité externe que de types d'utilisation d'un jeu de données.

En plus des concepts de qualité interne et de qualité externe, la littérature fait aussi état des concepts de qualité perçue et de métaqualité. Les différents aspects de la qualité des données géospatiales sont décrits dans les sections qui suivent.

### 2.2.1 Qualité interne

La qualité interne est décrite, selon la norme ISO 19157 (Organisation internationale de normalisation, 2013), à l'aide des cinq paramètres suivants : l'exhaustivité, la cohérence logique, la précision de positionnement, l'exactitude thématique et la qualité temporelle.

L'*exhaustivité* est définie comme la présence et l'absence d'entités, ainsi que leurs attributs et leurs relations. L'*exhaustivité* est décrite à l'aide des éléments de qualité des données suivants :

- commission – données excédentaires présentes dans un jeu de données (p. ex. un bâtiment redondant)
- omission – données absentes d'un jeu de données (p. ex. un bâtiment manquant)

---

*La cohérence logique* est décrite comme le degré d'adhésion aux règles logiques de structure de données, d'attribution et de relations. La cohérence logique comprend les cohérences suivantes :

- cohérence conceptuelle – adhésion aux règles du schéma conceptuel (p. ex. tous les ponts franchissent des rivières)
- cohérence de domaine – adhésion des valeurs aux domaines de valeur (p. ex. aucune année avant 1900)
- cohérence de format – degré auquel les données sont conservées dans la structure physique du jeu de données (p. ex. le format des mois est numérique (et non alphabétique), c.-à-d. 04 pour avril et non Avr.)
- cohérence topologique – justesse des caractéristiques topologiques explicitement codées dans un jeu de données (p. ex. il y a un nœud à toutes les intersections de rue)

*La précision de positionnement* est définie comme la précision de l'emplacement des entités dans un système de référence spatial. La précision de positionnement est décrite à l'aide des éléments suivants :

- précision absolue (ou précision externe) – proximité des valeurs de coordonnées reportées par rapport aux valeurs vraies ou acceptées comme telles (p. ex.  $\pm 5\text{m}$ )
- précision relative (ou précision interne) – proximité de la position relative et spatiale des entités d'un jeu de données par rapport à la position relative et spatiale respective acceptée comme étant vraie ou étant vraie (p. ex.  $\pm 1\text{m}$ )
- précision de la position des données maillées – proximité des valeurs de position spatiale des données maillées par rapport aux valeurs acceptées comme étant vraies ou étant vraies (p. ex.  $\pm 50\text{m}$ )

*L'exactitude thématique* est définie comme la précision des attributs quantitatifs et l'exactitude des attributs non quantitatifs et des classifications des entités et de leurs relations. L'exactitude thématique est décrite à l'aide des éléments suivants :

- exactitude de la classification – comparaison des catégories assignées à des entités ou à leurs attributs en fonction de l'univers du discours (c.-à-d. vérification au sol ou jeu de données de référence) (p. ex. 2 % des bâtiments à usage mixte peuvent être classés de façon erronée comme commerce ou résidence)
- exactitude des attributs non quantitatifs – mesure pour savoir si un attribut non quantitatif est exact ou inexact (p. ex. 5 % du style architectural des bâtiments peut être inexact)
- exactitude des attributs quantitatifs – proximité de la valeur d'un attribut quantitatif par rapport à une valeur acceptée comme étant vraie ou connue comme étant vraie (p. ex. la valeur commerciale est de  $\pm 10\%$ )

*La qualité temporelle* est définie comme la qualité des attributs temporels et des relations temporelles des entités. La qualité temporelle est décrite à l'aide des éléments suivants :

- précision d’une mesure du temps – proximité des mesures du temps reportées par rapport aux valeurs acceptées comme étant vraies ou connues comme étant vraies (p. ex.  $\pm 1$  minute)
- cohérence temporelle – exactitude de la chronologie des événements (p. ex. date de l’opération cadastrale de subdivision avant la date de construction de la résidence)
- validité temporelle – validité des données en ce qui a trait au format et au calendrier spécifiés pour le jeu de données (p. ex. aucun 30 février)

## 2.2.2 Qualité externe

La qualité externe d’un jeu de données géospatiales est le niveau de correspondance entre les caractéristiques des données (c.-à-d. la qualité interne) et les besoins explicites ou implicites d’un utilisateur pour une utilisation donnée dans un contexte donné. Les valeurs de qualité externe des données varient d’une utilisation à l’autre. La norme ISO 19157 (Organisation internationale de normalisation, 2013) introduit l’élément de qualité des données appelé utilité. *L’utilité* est fondée sur les exigences de l’utilisateur et tous les éléments de qualité interne de la norme ISO 19157 peuvent être utilisés pour évaluer l’utilité d’un jeu de données géospatiales.

## 2.2.3 Qualité perçue

Dans un environnement de collaboration, chaque utilisateur peut avoir une perception différente de la qualité externe d’un jeu de données. C’est ce qu’on appelle le concept de la qualité perçue (Gira, Bédard et Roche, 2009). Dans un tel contexte, le processus consistant à définir l’aptitude à l’emploi d’un jeu de données géospatiales peut être compliqué par le nombre d’utilisateurs concernés et la variété des besoins spécifiques. Chaque utilisateur a sa propre perception de la qualité du produit utilisé (comme la qualité d’un album sur Amazon ou l’Apple Store). Pour faciliter l’atteinte d’un consensus, chaque utilisateur peut noter le jeu de données selon sa perception (p. ex. utiliser un système de notation à cinq étoiles accompagné de commentaires, comme le font Amazon et Apple, voir (Jones, Devillers, Bédard et Schroth, 2013) et (Koistinen, 2015) pour voir des exemples). La qualité perçue globale est le résultat de l’agrégation des perceptions de tous les utilisateurs et elle est donc fondée sur une approche ascendante. Ces processus participatif ont été appelés information qualitative volontaire (VQI) (Bédard, 2012).

## 2.2.4 Métaqualité

La métaqualité fait référence à la qualité de l’information utilisée pour déterminer la qualité d’un objet, d’un concept ou d’un jeu de données (Organisation internationale de normalisation, 2013). Dans le contexte de la norme ISO 19157, la métaqualité est composée des éléments suivants : la confiance, la représentativité et l’homogénéité. La *confiance* est définie comme la fidélité d’un

---

résultat concernant la qualité des données (p. ex. un intervalle de confiance à un niveau donné de confiance). La *représentativité* est définie comme la mesure dans laquelle l'échantillon utilisé pour évaluer la qualité a produit un résultat qui est représentatif d'un jeu de données (p. ex. toutes les zones géographiques et les périodes concernées sont couvertes et la population est suffisamment importante). L'*homogénéité* est définie comme l'uniformité attendue ou testée des résultats obtenus pour une évaluation de la qualité des données (p. ex. comparaison des résultats de l'évaluation de plusieurs segments d'un jeu de données globales exprimés à l'aide d'erreurs quadratiques moyennes).

La connaissance de la qualité d'un résultat donné a souvent la même importance que le résultat lui-même. La métaqualité peut aider à quantifier le risque associé aux incertitudes des données géospatiales.

## 2.3 Les domaines de la qualité des données géospatiales

---

La qualité des données peut être évaluée selon différents niveaux de détails ou de granularité. La hiérarchie des niveaux de granularité peut être utilisée pour agréger de l'information de qualité à partir d'une valeur d'attribut d'une entité jusqu'à un jeu complet de données. La figure 1 présente les niveaux hiérarchiques de la qualité des données tels qu'ils sont définis dans la norme ISO 19157 : série de jeux de données, jeu de données, sous-ensemble d'un jeu de données, type d'entité, instance d'entité, type d'attribut et instance d'attribut (Organisation internationale de normalisation, 2013). Le tableau 1 présente des exemples de chaque niveau.

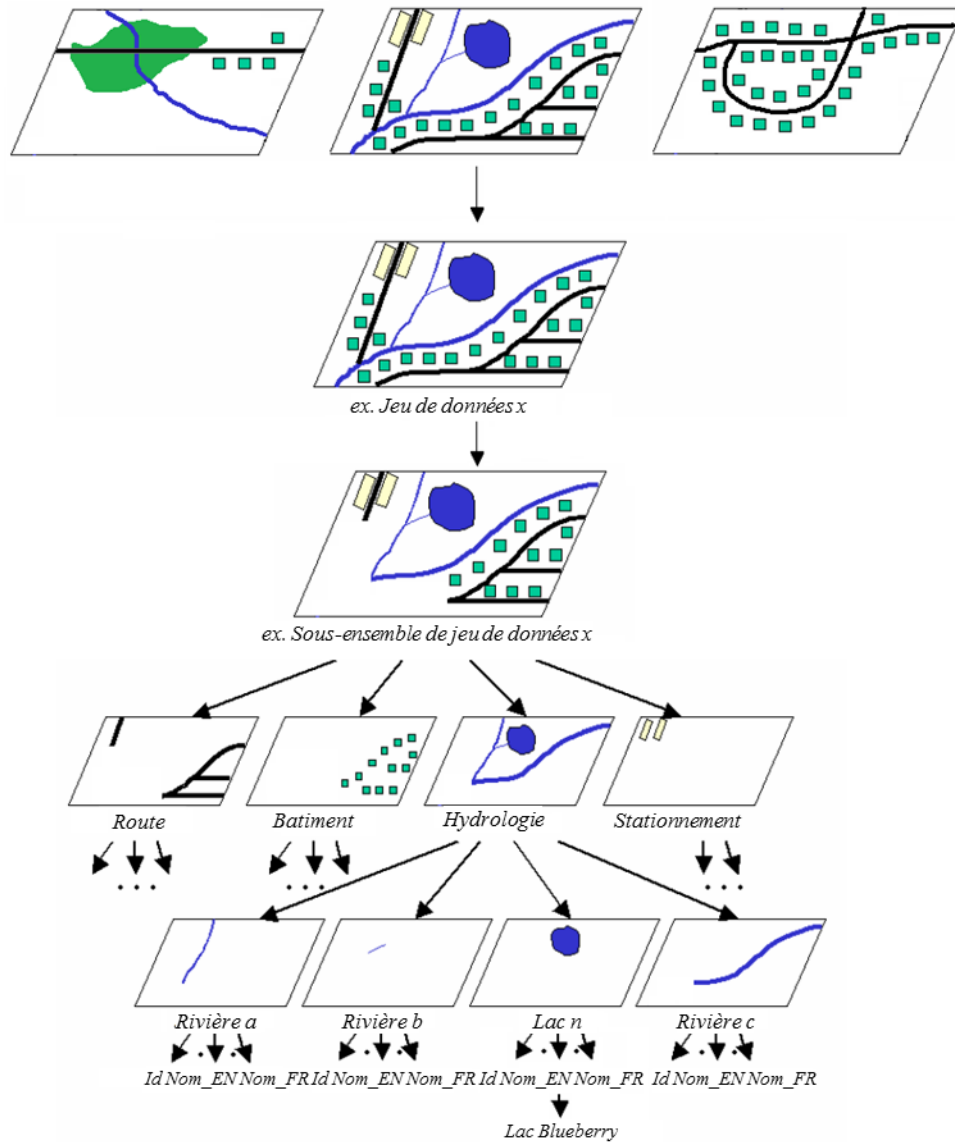


Figure 1. Les niveaux hiérarchiques de l'évaluation de la qualité des données géospatiales (adaptés de Devillers, 2004).

Niveau hiérarchique	Exemples
Série d'un jeu de données	<ul style="list-style-type: none"> <li>Le Système national de référence cartographique du Canada (SNRC)</li> <li>Les cartes hydrographiques du Service hydrographique du Canada (SHC)</li> </ul>
jeu de données	<ul style="list-style-type: none"> <li>Une carte du SNRC (p. ex. 031G Ottawa)</li> <li>Une carte hydrographique (p. ex. 1315 – Québec à Donnacona)</li> </ul>
Sous-ensemble d'un jeu de données	<ul style="list-style-type: none"> <li>Sous-ensemble de données d'une municipalité X</li> <li>Sous-ensemble de données pour les îles au nord du bras de mer</li> </ul>
Type d'entité	<ul style="list-style-type: none"> <li>Segment de route</li> </ul>

	<ul style="list-style-type: none"> <li>• Pont</li> </ul>
Instance d'entité	<ul style="list-style-type: none"> <li>• Route 417</li> <li>• Pont « Pierre-Laporte »</li> </ul>
Type d'attribut	<ul style="list-style-type: none"> <li>• Classification routière fonctionnelle</li> <li>• Hauteur libre</li> </ul>
Instance d'attribut	<ul style="list-style-type: none"> <li>• Code 2 (autoroute/route)</li> <li>• 35</li> </ul>

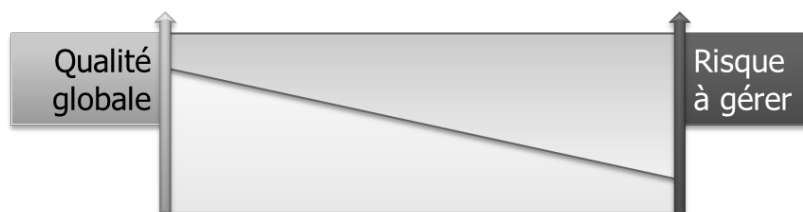
**Table 1. Exemples de niveaux hiérarchiques des données géospatiales pour l'évaluation de la qualité.**

## 2.4 La gestion des risques associés à la qualité des données géospatiales

Dans le contexte géospatiale évolutif, le nombre d'experts qui manipulent des données géospatiales a été largement dépassé par le nombre de non-experts et de citoyens. La connaissance des nouveaux utilisateurs des risques associés à l'utilisation des données géospatiales est limitée. Parallèlement, la quantité de données géospatiales mises à disposition augmente. Les données géospatiales sont maintenant partagées, échangées, intégrées, croisées et utilisées à des fins différentes de celles prévues par ceux qui les ont produites. Dans ce contexte, le risque que les données géospatiales soient utilisées de manière inappropriée augmente aussi.

Dans le domaine de la géomatique, on utilise un processus de gestion du risque comme guide pour expliquer comment éviter et gérer les impacts de l'incertitude : données géospatiales incertaines, qualité des données géospatiales incertaines, utilisations des données géospatiales incertaines et expertise des utilisateurs des données géospatiales incertaine. D'un point de vue juridique, l'utilisation d'un processus de gestion du risque est nécessaire pour protéger tant le producteur de données géospatiales que l'utilisateur (Gervais, Bédard, Jeansoulin et Cervelle, 2007).

Il est impossible d'avoir des données géospatiales d'une qualité irréprochable. La qualité générale englobe la qualité interne, la qualité externe, la qualité perçue et la métaqualité. De la même manière, le risque zéro n'existe pas. Les risques peuvent être réduits, mais rarement éliminés. La qualité et les risques sont interdépendants. Habituellement, plus la qualité est élevée, moins il y a de risques à gérer, comme le montre la figure 2.



**Figure 2. La relation entre la qualité et le risque.**



---

La qualité des données géospatiales sera abordée en détail aux sections 3.1 et 4.1. La gestion des risques associés à l'utilisation inappropriée des données géospatiales sera abordée en détail aux sections 3.2 et 4.2.

## **2.5 La diffusion d'information sur la qualité des données géospatiales et les risques d'utilisation**

L'information sur la qualité des données géospatiales et les risques de l'utilisation inappropriée des données géospatiales doit être communiquée aux différents acteurs qui interviennent dans la création et l'utilisation des données géospatiales. La façon traditionnelle de communiquer l'information sur la qualité consiste à utiliser des métadonnées. Toutefois, puisqu'elles sont conçues par des experts pour des experts (contexte E-E), les métadonnées sont moins appropriées pour d'autres types d'acteurs et particulièrement pour le grand public (marchés de masse E-C et C-C). Le processus utilisé pour communiquer de l'information sur la qualité des données et les risques doit être adapté pour tous les publics en employant un nouveau vocabulaire, de nouvelles méthodes et une nouvelle documentation relative aux produits. Puisque la qualité et les risques ont des valeurs différentes selon les utilisations, l'information variera en fonction de l'utilisateur et de l'utilisation.

La documentation proposée dans le présent guide, selon le contexte, comprend différentes méthodes conçues afin de mieux informer les utilisateurs et de réduire les risques associés à l'utilisation des données géospatiales qu'ils tenaient auparavant pour acquises. Inspirés de bonnes pratiques éthiques et d'exemples provenant des marchés de masse plus matures, ces produits d'information permettent de sensibiliser davantage tous les acteurs et peuvent prendre des formes variées (p. ex. du texte, une application axée sur la qualité, de la formation et des forums pour n'en nommer que quelques-unes).

Offrir de l'information appropriée permet d'aider les producteurs à respecter leurs obligations juridiques consistant à fournir de l'information, des conseils et des mises en garde. « Une bonne documentation sur la qualité et des avis de non-responsabilité bien rédigés permettront de réduire l'utilisation inappropriée et abusive des données » (National States Geographic Information Council, 2011).

La section 4 propose des exemples des produits de documentation répertoriés.

## 2.6 Les normes sur lesquelles reposent la qualité des données géospatiales et la gestion du risque

Un grand nombre de normes relatives à la qualité des données géospatiales et la gestion du risque ont été publiées. La présente section décrit les principales normes internationales, propres aux données géospatiales ou de portée générale, ainsi que les normes canadiennes pertinentes.

### 2.6.1 Normes spécifiques

L'Organisation internationale de normalisation (ISO) a publié des normes relatives à la qualité des données géospatiales : ISO 19115-1 et ISO 19115-2, ISO 19131, ISO 19157 et ISO 19158.

#### **ISO 19115-1 Information géographique – Métadonnées – Partie 1 – Principes de base**

La norme ISO 19115-1 (ISO 19115-1:2014 Information géographique – Métadonnées – Partie 1 – Principes de base), qui remplace la norme ISO 19115:2003, définit le schéma requis pour décrire des informations géographiques et des services au moyen de métadonnées (Organisation internationale de normalisation, 2014). Elle fournit de l'information concernant l'identification, l'étendue, la qualité (voir la norme ISO 19157), les aspects spatiaux et temporels, le contenu, la référence spatiale, la représentation des données, la distribution et d'autres propriétés des données géographiques numériques et des services. La norme ISO 19115-1 est applicable :

- au catalogage de tous les types de ressources, des activités des centres d'informations et à la description complète des jeux de données et des services;
- aux services géographiques, jeux de données géographiques, séries de jeux de données, entités géographiques individuelles et propriétés d'entités.

Le codage des métadonnées peut être fait en langage de balisage extensible (XML) à l'aide de la norme ISO 19139:2007 Information géographique – Métadonnées Implémentation de schémas XML (Organisation internationale de normalisation, 2007).

#### **ISO 19115-2 Information géographique – Métadonnées – Partie 2 – Extensions pour les images et les matrices**

La norme ISO 19115-2 (ISO 19115-2:2009 Information géographique – Métadonnées – Partie 2 – Extensions pour les images et les matrices) étend la norme existante sur les métadonnées en définissant le schéma requis de description des images et des matrices. Elle donne des informations relatives aux propriétés des équipements de mesure utilisés pour obtenir les données, la géométrie du processus de mesure utilisée par ces équipements et le processus de production utilisé pour numériser les données brutes (Organisation internationale de normalisation, 2009).

---

### **Le Profil nord-américain de la norme ISO 19115:2003**

Le Profil nord-américain de la norme ISO 19115:2003 a été conçu pour répondre aux besoins spécifiques des États-Unis et du Canada. Ce profil rend certains champs facultatifs de la norme ISO 19115 obligatoires, prend en charge plusieurs langues et transforme certains champs de texte libre en listes de codes. Il a été publié en tant que norme nationale américaine par l’American National Standards Institute (ANSI) et en tant que norme nationale canadienne par le Conseil canadien des normes (CCN).

### **Norme sur les données géospatiales du Secrétariat du Conseil du Trésor**

La Norme sur les données géospatiales du Secrétariat du Conseil du Trésor du Canada (Secrétariat du Conseil du Trésor du Canada, 2009) exige que les spécialistes du gouvernement du Canada responsables de créer ou d’utiliser des données géospatiales appliquent le Profil nord-américain de la norme ISO 19115 Information géographique - Métadonnées (PNA-Métadonnées).

### **ISO 19131 Information géographique – Spécifications de contenu informationnel**

La norme ISO 19131 (ISO 19131 Information géographique – Spécifications de contenu informationnel et ISO 19131:2007/Amd 1:2011 Exigences relatives à l’inclusion d’un schéma d’application et d’un catalogue d’objets géographiques et au traitement des couvertures dans un schéma d’application) décrit les exigences relatives à la spécification de contenu informationnel géographique, en s’appuyant sur les concepts présentés dans les autres Normes internationales ISO 19100, comme la norme relative à la qualité des données (ISO 19157) (Organisation internationale de normalisation, 2007), (Organisation internationale de normalisation, 2011). Elle apporte également une aide en matière de création de spécifications de contenu informationnel, de sorte qu’elles puissent être comprises et utilisées dans le cadre pour le lequel elles ont été prévues. Une spécification de contenu informationnel est une description détaillée d’un jeu de données ou d’une série de jeux de données comprenant des renseignements supplémentaires qui permet de créer un jeu de données ou une série de jeux de données et de le fournir à une autre partie qui pourra l’utiliser. Les utilisateurs peuvent aussi avoir recours à la norme ISO 19131 pour décrire leurs exigences.

### **ISO 19157 Information géographique – Qualité des données**

La norme ISO 19157 (ISO 19157 Information géographique – Qualité des données) (Organisation internationale de normalisation, 2013) établit les principes de description des données géographiques de la façon suivante :

- en définissant des composants destinés à décrire la qualité des données;

- en spécifiant des composants et la structure du contenu d'un registre de mesures de qualité des données;
- en décrivant des procédures générales d'évaluation de la qualité des données géographiques;
- en établissant les principes de la description de la qualité des données dans des rapports.

La norme ISO 19157 contient également un ensemble de mesures de qualité des données destinées à l'évaluation et à la mise en place de rapports sur la qualité des données. Elle s'applique aux producteurs de données pour évaluer la conformité d'un jeu de données par rapport à sa spécification de produit et aux utilisateurs cherchant à déterminer si des données géographiques spécifiques sont ou non aptes pour une application particulière.

La norme ISO 19157:2013 annule et remplace les normes ISO/TS 19138:2006, ISO 19114:2003 et ISO 19113:2002 qui ont fait l'objet d'une révision technique.

Cette norme est expliquée en détail aux sections 3.2 et 4.2.

#### **ISO/TS 19158 Information géographique – Assurance qualité relative à l'approvisionnement de données**

La norme ISO/TS 19158:2012 (ISO/TS 19158:2012, Information géographique – Assurance qualité relative à l'approvisionnement de données) (Organisation internationale de normalisation, 2012) fournit un cadre d'assurance qualité pour le producteur et l'utilisateur et leur relation de production. Cette norme technique traite des méthodes pour gérer de manière plus efficace et efficiente la qualité de la production. Elle favorise l'innovation et l'amélioration continue en ce qui concerne :

- les principes de qualité de l'information géographique et les procédures d'évaluation de la qualité en place;
- les systèmes existants de gestion de la qualité.

## **2.6.2 Normes génériques**

### **ISO 9000 – Management de la qualité**

La famille de normes ISO 9000 (ISO 9000:2005 Management de la qualité) couvre les divers aspects de la gestion de la qualité et elle offre des lignes directrices et des outils aux organismes qui veulent que leurs produits et services soient constamment en phase avec ce que leurs clients demandent et que la qualité ne cesse de s'améliorer. La famille de normes ISO 9000 comprend les normes suivantes (Organisation internationale de normalisation, 2005) :

- ISO 9001:2008 : établit les exigences relatives à un système de gestion de la qualité

- 
- ISO 9000:2005 : décrit les notions fondamentales et la terminologie
  - ISO 9004:2009 : montre comment augmenter l’efficacité et l’efficacité d’un système de gestion de la qualité
  - ISO 19011:2011 : établit des lignes directrices pour les audits internes et externes des systèmes de gestion de la qualité

Une nouvelle version de la norme ISO 9001 sera disponible d’ici la fin de 2015.

### **ISO 31000 Management du risque – Principes et lignes directrices**

La norme ISO 31000 (ISO 31000:2009 Management du risque – Principes et lignes directrices) fournit des principes, un cadre et un processus pour gérer les risques. Elle peut être utilisée par tout type d’organisme sans distinction de taille, d’activité ou de secteur. Les organismes qui ont recours à la norme ISO 31000 augmentent leurs chances d’atteindre leurs objectifs, sont mieux à même de cerner les occasions et les menaces et d’allouer et d’utiliser efficacement les ressources pour la gestion du risque (Organisation internationale de normalisation, 2009).

D’autres normes se rapportent également à la gestion du risque :

- le Guide ISO 73:2009 Management du risque - Vocabulaire fournit les termes et définitions relatifs à la gestion du risque
- la norme ISO/IEC 31010:2009 Gestion des risques – Techniques d’évaluation des risques est axée sur l’évaluation des risques

Cette norme est expliquée en détail aux sections 3.2 et 4.2.

### 3. Principes fondamentaux de la gestion du risque et de la qualité des données géospatiales

La qualité des données géospatiales repose fondamentalement sur la gestion de la qualité de toutes les étapes du cycle de vie du contenu informationnel. La gestion de la qualité nécessite également une évaluation de la qualité fondée sur des exigences ou des besoins déterminés. Les résultats d'une évaluation de la qualité peuvent être très complexes et fiables ou ils peuvent demeurer généraux et présenter un niveau d'incertitude plus élevé, selon la portée de l'évaluation. La gestion du risque permet de composer avec l'incertitude, mais aussi de sélectionner les meilleures stratégies pour atténuer le risque d'une utilisation inappropriée des données géospatiales. Les mesures de gestion du risque et les résultats des évaluations de la qualité des données géospatiales doivent être bien communiqués à tous les intervenants du processus. Toutes ces activités doivent être surveillées et examinées au besoin. Ce processus général est illustré à la figure 3.

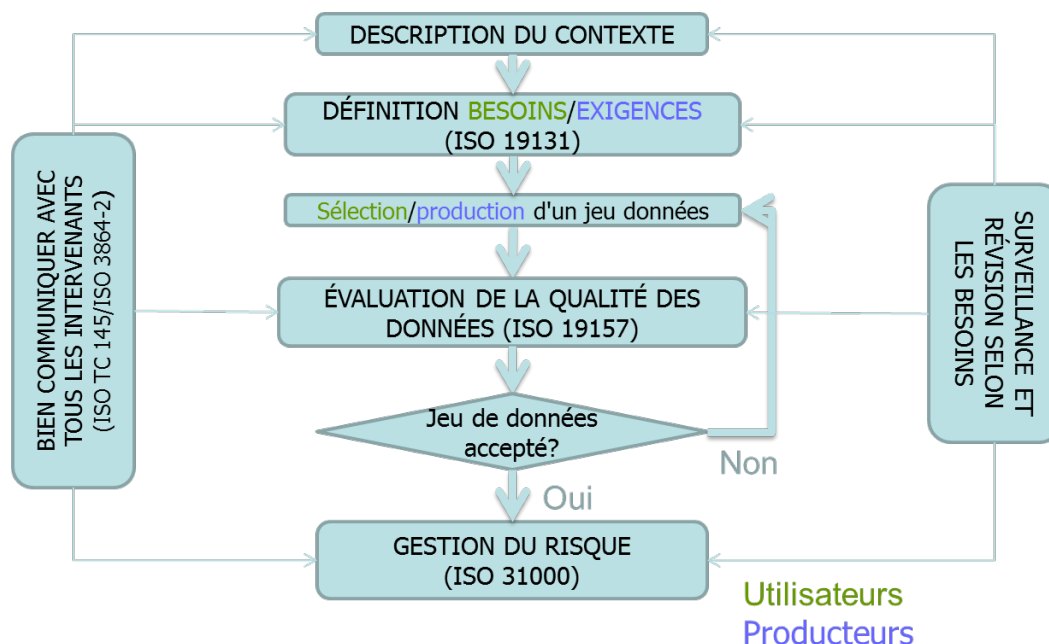


Figure 3. Cycle d'assurance de la qualité des données géospatiales pour une société œuvrant dans le domaine de la référence spatiale

La présente section explique les principales étapes d'un processus de gestion de la qualité des données géospatiales ainsi que des risques liés à l'utilisation inappropriée de ces données. Les marchés établis adoptent cette approche pour protéger à la fois les consommateurs et les fournisseurs de biens et de services. Des exemples détaillés de ces activités sont présentés à la section 4.

### 3.1 Gestion de la qualité des données géospatiales

La gestion de la qualité des données géospatiales est l'activité qui consiste à définir le niveau de qualité exigé des données requises, à définir, mettre en œuvre et contrôler les mesures nécessaires qui permettent de s'assurer que les critères de qualité sont respectés et à évaluer, documenter et diffuser de l'information sur la qualité (Figure 4).

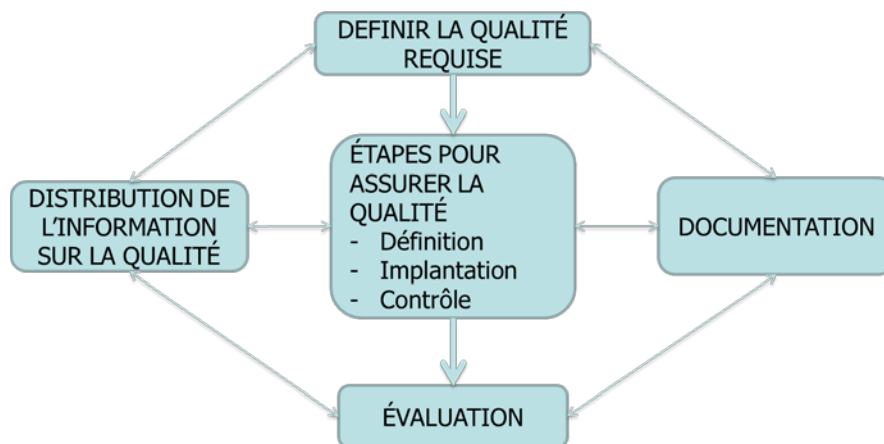


Figure 4. Processus de gestion de la qualité des données géospatiales

Du point de vue du producteur (contextes E-E ou E-C), la qualité des données géospatiales et les risques d'utilisation inappropriée doivent être gérés à chaque étape du cycle de vie du contenu informationnel (production ou processus de mise à jour), tel qu'illustré à la figure 5.

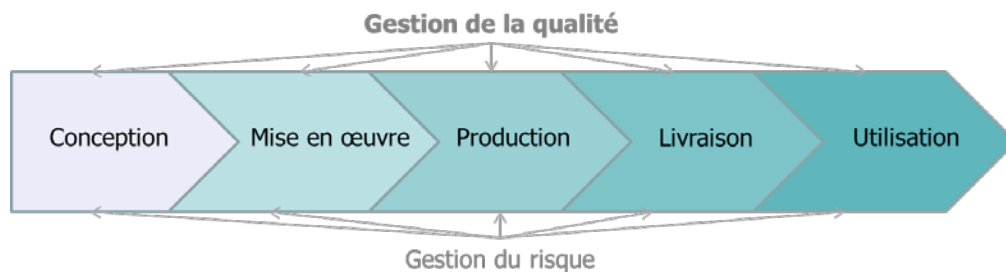


Figure 5. Gestion de la qualité à chaque étape du cycle de vie du contenu informationnel

---

À l'étape de la *conception*, la qualité des données peut être gérée pendant la modélisation des données (schéma d'application ou schéma de la géométrie et des fonctions de couverture), la création du dictionnaire de données (catalogue d'entités) ainsi que la conception des spécifications d'un produit, comme l'indique la section 4.1.1.

Les normes suivantes peuvent être utilisées à l'appui des activités de conception :

- ISO 19109:2005 Information géographique – Règles de schéma d'application
- ISO 19110:2005 Information géographique – Méthodologie de catalogage des entités
- ISO 19123:2005 Information géographique – Schéma de la géométrie et des fonctions de couverture
- ISO 19131:2007 Information géographique – Spécifications de contenu informationnel

À l'étape de la *mise en œuvre*, la qualité des données peut être gérée à l'aide de contraintes d'intégrité et de contraintes d'intégrité spatiale. Les contraintes d'intégrité sont des assertions qui limitent les données pouvant être stockées dans une base de données afin d'éviter les incohérences. Des exemples sont présentés à la section 4.1.2.

À l'étape de *production*, la qualité des données est évaluée ou contrôlée et les métadonnées sont gérées (consulter la section 4.1.3). Les normes suivantes peuvent être utilisées à l'appui de ces activités :

- ISO 19157:2013 Information géographique – Qualité des données
- ISO 19115-1:2014 Information géographique – Métadonnées – Partie 1 : Principes de base

Lors de l'étape de la *livraison*, la qualité des données est gérée en fournissant aux utilisateurs la documentation appropriée sur les jeux de données. Pour qu'un producteur de données puisse mieux s'acquitter de ses obligations légales, il est recommandé de joindre aux jeux de données et aux métadonnées connexes sur la qualité les mises en garde requises concernant les restrictions contenues dans le jeu de données et les risques possibles de l'utilisation inappropriée des données (Chandler & Levitt, 2011). La documentation peut également comprendre un manuel d'utilisation traditionnel ou interactif (contextuel), une liste des usages recommandés et non recommandés, du matériel de formation ou d'autres produits livrables mentionnés à la section 4.1.4. En d'autres termes, les stratégies de gestion du risque influencent directement la façon dont la qualité des données est gérée à cette étape.

Lors de l'étape de l'*utilisation*, on distingue le lien étroit qui unit la qualité des données et les risques d'utilisation inappropriée des données. Ces risques sont déterminés grâce aux mesures prises lors de l'étape de la prestation et ils peuvent être gérés en communiquant sur demande avec les utilisateurs, en leur fournissant des renseignements à jour et en leur proposant de la formation ainsi que des applications axées sur la qualité des données et d'autres stratégies, tel que décrit à la section 4.1.5.



## 3.2 Évaluation de la qualité des données géospatiales

L'évaluation de la qualité de l'information géospatiale peut être extrêmement complexe, même pour un expert de la géomatique. L'évaluation de la qualité des données géospatiales peut être définie comme un processus servant à déterminer si du contenu informationnel géospatial répond aux objectifs en ce qui concerne les spécifications du produit, du point de vue du producteur (c.-à-d. qualité interne), ou en ce qui concerne la nécessité du produit en vue d'une utilisation prévue, du point de vue de l'utilisateur (c.-à-d. convivialité ou qualité externe). Le processus d'évaluation peut être formel, selon la procédure décrite dans la norme ISO 19157, ou moins formel, selon le contexte défini à la première étape de l'analyse des risques (consulter la section 3.3.1). La figure 6 illustre l'étendue des méthodes d'évaluation de la qualité.



Figure 6. Méthodes d'évaluation de la qualité des données, selon le contexte

La figure 7 décrit les étapes du processus d'évaluation formel (fondé sur la norme ISO 19157), ainsi que le processus d'évaluation informel.

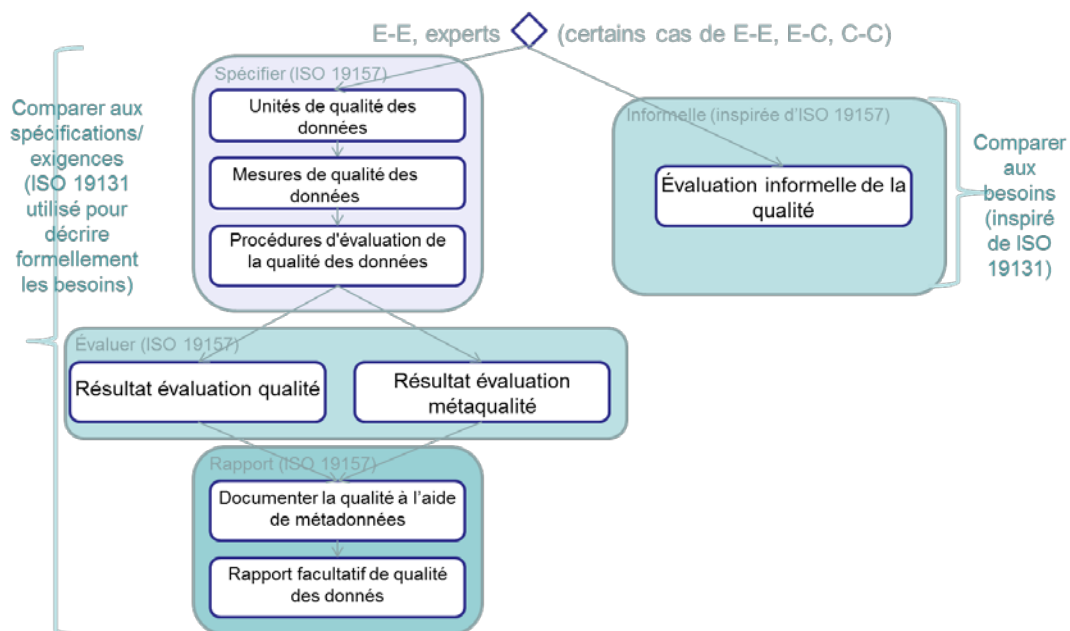
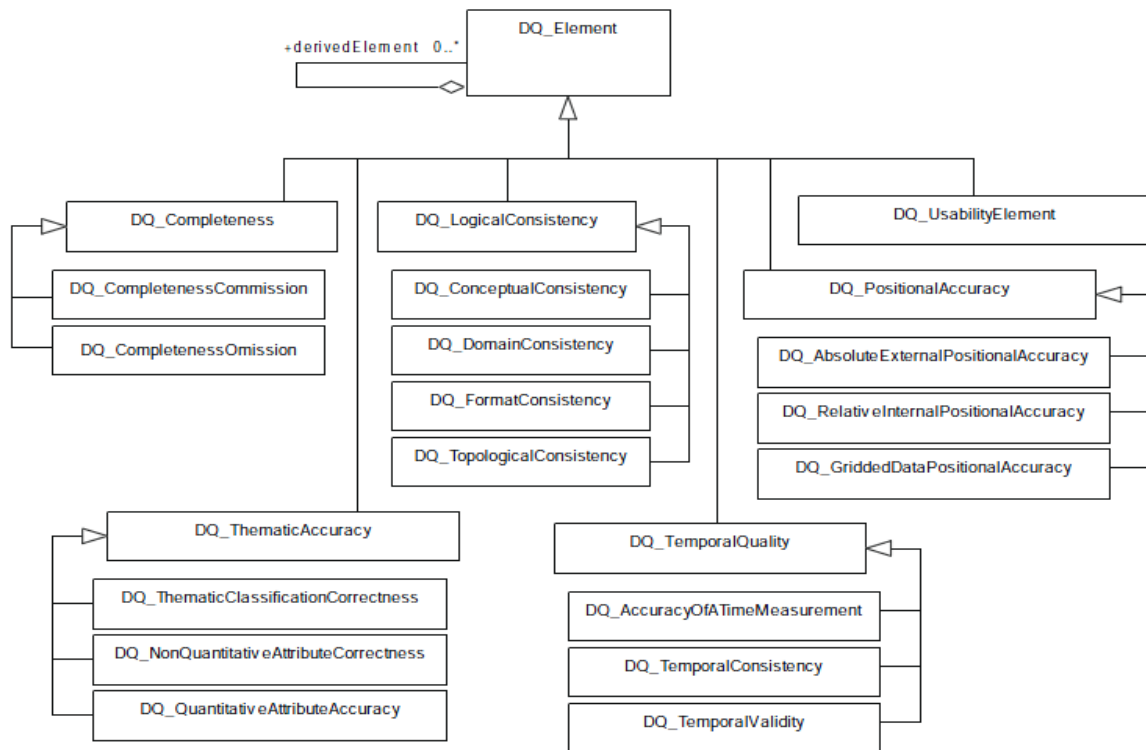


Figure 7. Processus formel et informel d'évaluation de la qualité des données géospatiales.

La norme ISO 19157 (Organisation internationale de normalisation, 2013) définit une procédure générale d'évaluation de la qualité des données géospatiales et des principes de production de rapports sur la qualité des données évaluées. Le processus d'évaluation proposé comprend les trois étapes suivantes : spécification, évaluation et production de rapports.

### 3.2.1 Spécification formelle de la qualité des données géospatiales

La première étape de la spécification consiste à définir les unités de qualité des données. Chaque unité de qualité des données englobe la portée (la portée précise l'étendue spatiale ou temporelle ou les caractéristiques communes qui définissent les données dont la qualité sera évaluée) et au moins un composant de la qualité des données. Les composants de la qualité des données de la norme ISO 19157 sont présentés à la figure 8.



**Figure 8. Composants de la qualité des données de la norme ISO 19157 (Organisation internationale de normalisation, 2013)**

La deuxième étape de la spécification consiste à définir les mesures de qualité des données pour comparer les caractéristiques de le jeu de données d'après des spécifications formelles (fondé sur la norme ISO 19131 Information géographique – Spécifications de contenu informationnel). Un exemple de mesure de la qualité des données pour le composant d'intégralité serait le nombre

---

d'éléments excédentaires. La norme ISO 19157 fournit une liste des mesures courantes. De nouvelles mesures peuvent également être créées.

La troisième étape de la spécification consiste à définir les procédures d'évaluation de la qualité des données servant à déterminer la valeur des mesures de la qualité des données sélectionnées. Une procédure d'évaluation de la qualité des données est réalisée au moyen d'une ou de plusieurs méthodes d'évaluation de la qualité des données. Les méthodes d'évaluation de la qualité des données peuvent être classées selon les deux grandes catégories suivantes : directe et indirecte. Les méthodes d'évaluation directe déterminent la qualité des données en comparant les données à des références internes ou externes. Les méthodes d'évaluation indirecte déduisent ou estiment la qualité des données à l'aide des renseignements sur les données, notamment leur origine. Des exemples sont fournis à la section 4.2.1.

### **3.2.2 Évaluation formelle de la qualité des données géospatiales**

La première étape de l'évaluation consiste à définir les extrants de l'évaluation de la qualité des données. Au moins un résultat doit être fourni pour chaque composant de la qualité des données. Le résultat de l'évaluation peut être un résultat quantitatif, un résultat de conformité, un résultat descriptif ou un résultat de couverture (des exemples de couverture sont fournis à la section 4.2.2).

La deuxième étape de l'évaluation consiste à évaluer la métaqualité. La métaqualité est définie comme l'information décrivant la qualité des données (Organisation internationale de normalisation, 2013). Les éléments de la métaqualité (fiabilité, représentativité et homogénéité) sont un jeu d'énoncés quantitatifs et qualitatifs sur une évaluation de la qualité et ses résultats. Les renseignements sur la qualité d'un résultat donné sont souvent aussi importants que le résultat en soi.

### **3.2.3 Production de rapports formels sur la qualité des données géospatiales**

Selon la norme ISO 19157, la première étape de la production de rapports consiste à rendre compte des résultats de l'évaluation de la qualité des données sous forme de métadonnées appropriées conformément aux normes ISO 19115-1:2014 Information géographique – Métadonnées – Partie 1 : Principes de base (Organisation internationale de normalisation, 2014) et ISO 19115-2:2009 Information géographique – Métadonnées – Partie 2 : Extensions pour les images et les matrices (Organisation internationale de normalisation, 2009).

La deuxième étape de la production de rapports consiste à créer un rapport sur la qualité des données (facultatif). Un rapport sur la qualité autonome peut être créé pour fournir des détails ou des nuances dans un format plus facile à comprendre que les métadonnées. Toutefois, selon la

---

norme ISO 19157, le rapport sur la qualité autonome ne vise pas à remplacer les métadonnées. Les métadonnées devraient constituer une référence du rapport sur la qualité autonome, le cas échéant.

Diverses méthodes présentant différents niveaux de complexité sont utilisées pour représenter la qualité. Les exemples comprennent l'utilisation de diagrammes en pour comparer les composants de qualité d'un jeu de données au niveau de qualité requis, aux tableaux sur la qualité, aux descriptions textuelles, etc. Ces méthodes adaptées pour simplifier la communication avec différents types d'utilisateurs de données géospatiales peuvent être développées davantage et mises à la disposition des utilisateurs sur les portails d'accès aux données du système de dissémination automatique de l'information, par exemple. Ces méthodes utilisent un langage compréhensible par la majorité des utilisateurs. Des exemples détaillés sont fournis à la section 4.2.4.

### 3.2.4 Évaluation informelle de la qualité des données géospatiales

Des processus d'évaluation de la qualité des données et de production de rapports moins officiels sont utilisés par l'industrie dans des contextes E-E ou E-C, selon le contexte défini au cours de la première étape de l'analyse des risques. Quoiqu'il en soit, il n'est pas surprenant de constater que leurs procédures et leur contenu sont semblables à l'approche de la norme ISO 19157, puisque cette norme repose sur une logique très rationnelle et qu'elle offre une certaine souplesse en ce qui concerne le niveau de détails. Il est tout à fait possible qu'aucune métadonnée ne soit fournie et que tous les renseignements soient transmis dans un rapport sur la qualité des données.

Dans les contextes E-E ou E-C, les utilisateurs de données géospatiales peuvent utiliser divers moyens pour évaluer si un jeu de données géospatiales répond à leurs besoins (c.-à-d. qualité externe). Sans le savoir, ils suivent habituellement le même raisonnement que celui de la norme ISO 19157, mais de façon moins rigoureuse et plus superficielle. Ils peuvent adapter leurs méthodes de mesure à leurs modes de représentation des données (p. ex. diagramme en toile d'araignée à cinq axes avec des unités de mesure allant de « non » à « tout à fait » en trois niveaux). En d'autres mots, ils peuvent combiner des méthodes plus officielles de représentation de la qualité avec des mesures de qualité plus qualitatives et moins rigoureuses. Lorsque les renseignements sur la qualité sont le fruit d'une production participative, des systèmes de communication volontaire de données peuvent être configurés d'après le système de cotation populaire à cinq étoiles (lequel indique la note moyenne, le nombre de votes par cote, le filtrage par date, région ou version de le jeu de données, etc.).

---

## 3.3 Gestion du risque

Le risque est « l'effet de l'incertitude sur les objectifs » (Canada, 2012), (Organisation internationale de normalisation, 2009). L'effet est tout écart par rapport à la situation attendue

(positif ou négatif). L'incertitude est liée aux lacunes en matière d'information. Tous les types de buts peuvent constituer des objectifs (économique, santé, environnement, etc.). Le risque est souvent exprimé comme une combinaison des conséquences positives ou négatives d'un événement et de leur probabilité d'occurrence. Le risque désigne l'effet de l'incertitude et il est donc prospectif. Dans le contexte des données géospatiales, le risque est considéré comme le risque d'utilisation inappropriée des données géospatiales.

Selon (Kerzner, 2009), la gestion du risque est un acte ou une pratique permettant de composer avec un risque. La gestion du risque consiste à diriger et à contrôler les mesures prises par une organisation pour minimiser les répercussions inattendues sur l'atteinte de ses objectifs. L'un des principes clés de la gestion du risque est que le risque zéro n'existe pas. La gestion du risque suppose d'équilibrer les efforts pour éviter des résultats inattendus qui pourraient avoir des effets négatifs sur l'atteinte des objectifs pour les intervenants (y compris les clients), comme l'indique la figure 9.

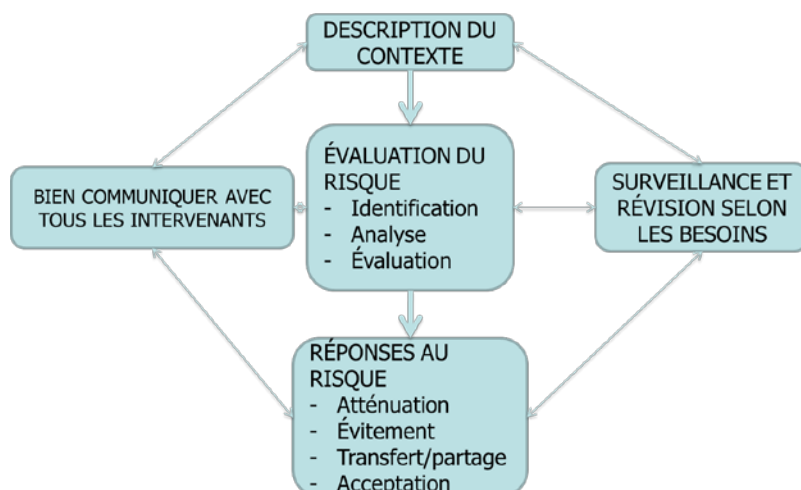


**Figure 9. Gestion du risque : un équilibre entre les efforts et les impacts négatifs imprévus**

La gestion du risque permet de réagir de manière préventive au changement en atténuant les menaces et en tirant profit des possibilités que l'incertitude présente pour les objectifs d'une organisation (Canada, 2012). D'un point de vue juridique, l'utilisation d'une approche de gestion du risque est nécessaire pour renforcer la protection du producteur et de l'utilisateur des données géospatiales.

### 3.3.1 Processus de gestion du risque

La gestion du risque est un processus systématique, continu et itératif qui permet successivement de cerner et d'évaluer les risques dans une situation donnée, d'élaborer des stratégies pour les gérer et d'en faire le suivi, de les documenter et de communiquer de l'information sur ceux-ci. La figure 10 présente le cadre de gestion du risque décrit dans la norme ISO 31000 – Management du risque – Principes et lignes directrices (Organisation internationale de normalisation, 2009).



**Figure 10. Étapes du processus de gestion du risque de la norme ISO 31000**

Les étapes générales du processus de gestion du risque de la norme ISO 31000 sont les suivantes :

- Description du contexte organisationnel et du contexte de gestion du risque
- Évaluation des risques
- Établissement des interventions
- Communication des risques
- Surveillance et examen des risques

#### **Description du contexte de gestion du risque**

Avant de commencer la conception et la mise en œuvre du processus de gestion du risque, il est important d'évaluer et de comprendre le contexte (général, interne, externe et particulier) de l'organisation, car cela peut grandement influencer la conception du processus (Organisation internationale de normalisation, 2009).

La description du contexte général comprend habituellement les objectifs de l'organisation (p. ex. sécurité, santé, économie) ainsi que la portée de la gestion du risque (p. ex. un projet, l'ensemble des activités de l'organisation, une collectivité donnée).

La description du contexte externe peut inclure les éléments suivants :

- position de l'organisation à l'échelle locale, régionale, nationale et internationale;
- besoins juridiques et de réglementation;
- perceptions des intervenants;
- microéconomie et macroéconomique;
- environnement sociopolitique;
- concurrence, tendances, etc.

---

La description du contexte interne peut inclure les éléments suivants :

- culture de l'organisation, gouvernance, normes, structure et stratégie;
- engagements et relations contractuelles;
- analyse FFPM (forces, faiblesses, possibilités et menaces).

La description du contexte correspondant à la portée définie peut inclure les éléments suivants :

- objectifs de gestion du risque pour la portée définie;
- ressources et temps requis, contraintes relatives à la gestion de projet;
- profondeur, ampleur, inclusions, exclusions, responsabilités;
- méthodologies;
- critères de risque, mesures, niveaux de tolérance, décisions à prendre.

### Évaluation des risques

Une fois que le contexte est défini, le processus d'évaluation des risques peut commencer. L'évaluation du risque est composée de trois activités principales : la détermination des risques, l'analyse des risques et l'évaluation des risques.

L'*identification du risque* consiste à établir une liste exhaustive des risques qui pourraient avoir une incidence (positive ou négative) sur l'atteinte des objectifs. La liste doit faire état des raisons pour lesquelles l'atteinte des objectifs pourrait être compromise. Pour chacun des risques identifiés, les éléments suivants doivent également être déterminés :

- les sources de risques, contrôlées ou non, connues, inconnues ou nouvelles (p. ex. économiques, sociales, politiques, naturelles, marchés, technologiques, opérationnelles, ressources humaines, juridiques);
- les répercussions et les effets cumulatifs;
- les scénarios possibles.

L'*analyse des risques* consiste à comprendre la nature des risques cernés en déterminant les sources, les causes, les conséquences (concrètes ou abstraites) sur les objectifs, les probabilités d'occurrence et l'interdépendance. L'analyse des risques peut :

- être fondée sur des données historiques, une extrapolation ou une prédiction;
- tenir compte des contrôles existants;
- être qualitative ou quantitative;
- être entreprise selon différents niveaux de détail.

---

Une fois que les conséquences et que la probabilité d'occurrence de chaque risque sont comprises, le niveau de risque peut être défini en combinant ces deux paramètres. Le degré de confiance relatif au niveau de risque est également déterminé.

L'*évaluation du risque* est la comparaison entre le niveau de risque établi lors de l'analyse des risques et les critères de risque établis lors de la définition du contexte. Cette comparaison permet aux décideurs de choisir des stratégies de traitement des risques et d'établir un ordre de priorité :

- Le risque ne peut être , son traitement est essentiel.
- Le risque peut être toléré, il doit être surveillé.
- Le risque est négligeable, mais il faut en demeurer conscient.

L'évaluation du risque doit tenir compte des risques juridiques et réglementaires et d'autres exigences. Les personnes ayant des connaissances pertinentes devraient participer à la détermination des risques opérationnels (p. ex. experts, utilisateurs, spécialistes des services de soutien, conseillers).

La norme d'accompagnement IEC 31010:2009 Gestion des risques – Techniques d'évaluation des risques fournit des directives sur les techniques d'évaluation des risques.

#### **Établissement des interventions (traitement du risque)**

Le traitement des risques consiste à choisir et à mettre en œuvre une ou une combinaison de stratégies visant à modifier un risque accepté pour atteindre les niveaux de tolérance acceptables. Quatre catégories de stratégies peuvent être utilisées :

- Atténuation : actions pour éliminer ou réduire les conséquences ou la probabilité des risques
- Évitement : renoncer à l'activité pour supprimer le risque associé
- Transfert/partage : déplacer les impacts vers un autre acteur, en totalité ou en partie
- Acceptation : acceptation volontaire et admission du risque (ignorer un risque équivaut à accepter le risque de façon informelle).

Il existe de nombreuses solutions de rechange pour gérer un risque donné; elles diffèrent sur le plan des coûts, des délais et de l'efficacité.

Cette étape consiste à équilibrer les efforts et les avantages de tous les intervenants à l'égard des objectifs énoncés dans la description du contexte. Des exemples détaillés de stratégies de risques liés à l'utilisation inappropriée des données géospatiales sont fournis à la section 4.3.3.



---

## Communication du risque

La communication du risque doit répondre aux questions relatives aux contextes et aux intérêts des intervenants internes et externes, aux causes et aux origines des risques cernés, à leurs conséquences positives et négatives, à leur niveau de risque, aux critères de risque et aux niveaux de tolérance, aux traitements qui existent déjà ou qui seront mis en œuvre, à leur surveillance, à leur examen, etc.

Des exemples de stratégies de communication sont fournis à la section 4.3.4, lesquelles consistent à intégrer des renseignements sur le risque dans les rapports et les métadonnées sur la qualité des données, dans les manuels d'utilisation, dans les garanties, les forums, etc. D'un point de vue juridique, les risques devraient être communiqués dans une langue que les utilisateurs visés seront en mesure de comprendre.

## Surveillance et examen du risque

La surveillance du risque est la surveillance régulière du risque et du succès de son traitement. Cette étape vise également à :

- détecter des changements dans les contextes;
- détecter les nouveaux risques;
- viser une amélioration constante.

---

## 4. Mise en pratique de la théorie

La présente section traite, de la gestion de la qualité des données géospatiales ainsi que de la gestion des risques d'utilisation inappropriée de données géospatiales, et donne des exemples pour chacune des étapes des deux processus.

### 4.1 Mettre en pratique la gestion de la qualité des données géospatiales

---

Du point de vue d'un producteur (contextes E-E ou E-C), la qualité des données géospatiales et le risque d'utilisation inappropriée doivent être gérés conjointement à chaque étape du cycle de vie d'un produit de données.

#### 4.1.1 Phase de conception

Lors de la conception d'un produit de données vectoriel, la norme ISO 19109:2005, Information géographique – Règles de schéma d'application, peut être utilisée pour définir correctement le modèle (appelé *schéma d'application*) du produit de données à créer ou mettre à jour (Organisation internationale de normalisation, 2005). Cette norme couvre :

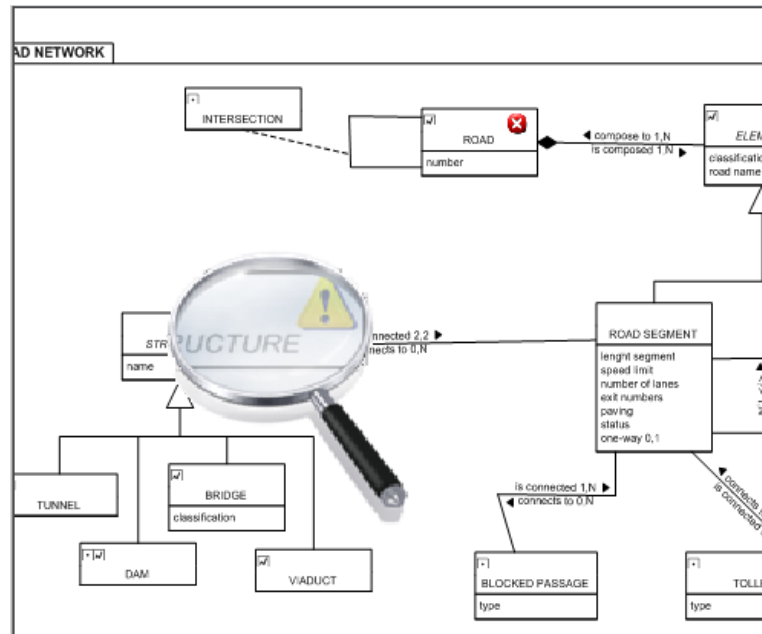
- la modélisation conceptuelle des entités et de leurs propriétés dans un univers du discours;
- la définition des schémas d'application;
- l'utilisation du langage de schéma conceptuel pour les schémas d'application;
- la transition des concepts du modèle conceptuel aux types de données du schéma d'application;
- l'intégration de schémas normalisés provenant d'autres normes ISO sur l'information géographique dans le schéma d'application.

Du point de vue de la qualité des données géospatiales, suivre une méthode formelle pour la conception du schéma d'application contribuera à améliorer la cohérence logique des données (élément de qualité des données DQ\_LogicalConsistency de la norme ISO 19157).

Du point de vue de la gestion du risque, le schéma d'application peut être enrichi d'éléments relatifs à la gestion du risque, comme des avertissements. La figure 11 présente un schéma d'application de réseau routier (extrait de Levesque, Bédard, Gervais et Devillers, 2007) enrichi

---

de symboles d'avertissement (fondé sur la norme ISO/TC 145/ISO 3864-2, Symboles graphiques – Couleurs de sécurité et signaux de sécurité – Partie 2 : Principes de conception pour l'étiquetage de sécurité des produits [Organisation internationale de normalisation, 2004]) afin de souligner les problèmes possibles relatifs à l'utilisation, la probabilité des risques et leurs répercussions.



**Figure 11 : Schéma d'application de réseau routier enrichi de symboles d'avertissement pour la gestion du risque.**

Lors de la conception d'un produit de données vectoriel, la norme ISO 19110:2005, Information géographique – Méthodologie de catalogage des entités, peut être utilisée pour définir correctement le dictionnaire de données (appelé *catalogue d'entités*) du produit de données à créer ou mettre à jour (Organisation internationale de normalisation, 2005). Cette norme décrit :

- une méthode de catalogage des types d'entités;
- la façon dont la classification des types d'entités est organisée en un catalogue d'entités et présentée aux utilisateurs d'un jeu de données géographiques.

Du point de vue de la qualité des données géospatiales, suivre une méthode formelle pour l'enregistrement des entités dans un catalogue améliorera la cohérence logique des données (élément de qualité des données DQ\_LogicalConsistency de la norme ISO 19157).

Du point de vue de la gestion du risque, le catalogue d'entités peut être enrichi par des éléments de gestion du risque, comme des descriptions détaillées du risque, la probabilité qu'il se présente, ses effets, des mesures recommandées et des avertissements se trouvant dans le schéma d'application correspondant. La figure 12 fournit un exemple de catalogue d'entités enrichi de renseignements sur la gestion du risque (Bédard, Chandler, Devillers et Gervais, 2009). La

figure 13 présente un extrait d'un catalogue d'entités de CanVec+ (Ressources naturelles Canada, 2014) enrichi d'une section d'avertissements qui décrit en détail les symboles d'avertissement utilisés dans le schéma d'application connexe.

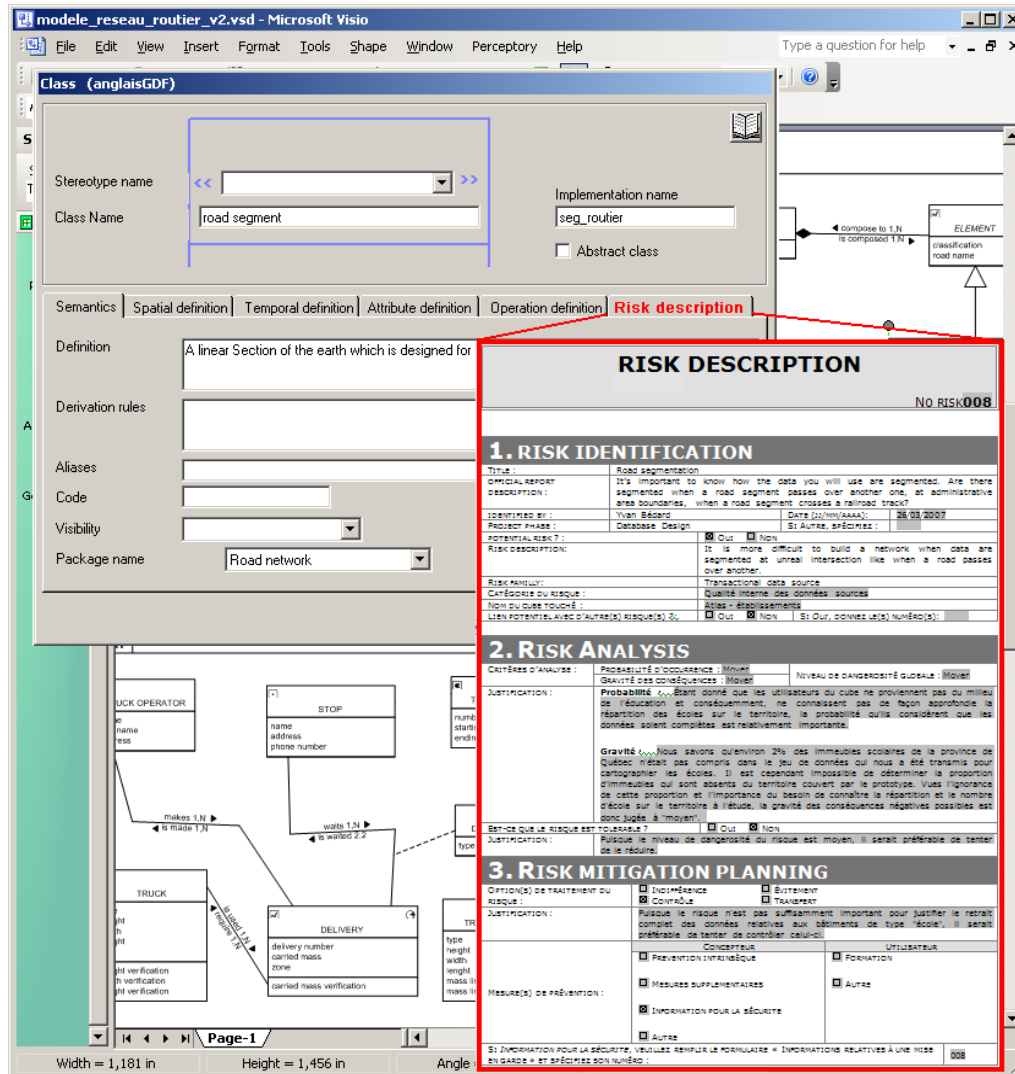


Figure 12 : Exemple de catalogue d'entités enrichi de renseignements sur la gestion du risque (tiré de Bédard, Chandler, Devillers et Gervais, 2009)

<b>Entité</b>						
Nom - ( Nom anglais )		<b>Parc d'amusement - ( Amusement park )</b>				
Définition		Zone où l'on se divertit avec jeux, manèges, etc.				
Thème		Code générique	Est abstrait	GéoBase	Géométrie	
Lieux d'intérêts		2260009	Non	Non	Surface	
<b>Attributs</b>						
Nom	<b>Code de spécification</b>					
Nom	<b>Date de validité</b>					
Nom	<b>Identifiant de l'élément</b>					
Nom	<b>Précision planimétrique</b>					
Nom	<b>Thème</b>					
<b>Combinaisons de valeurs d'attribut (Codes de spécification)</b>						
Combinaison d'attributs			Point	Ligne	Surface	
Parc d'amusement - ( Amusement park )					2260012	
<b>Dimensions minimales</b>						
Code de spécification	Distance latérale (Mètre)	Distance longitudinale (Mètre)	Longueur (Mètre)	Superficie (Mètre carré)	Tolérance d'angle d'anicroche (Degré)	Tolérance d'angle droit (Degré)
2260012	1.5	3	---	500	10	---
<b>Mises en garde</b>						

**Figure 13 : Extrait du catalogue d'entités de CanVec+ (tiré de Ressources naturelles Canada, 2014).**

Lors de la conception d'un produit de données matricielles, la norme ISO 19123:2005, Information géographique – Schéma de la géométrie et des fonctions de couverture, peut être utilisée pour définir correctement le modèle (appelé *schéma de la géométrie et des fonctions de couverture*) du produit de données à créer ou mettre à jour (Organisation internationale de normalisation, 2005). Cette norme décrit :

- le schéma conceptuel des caractéristiques spatiales des couvertures;
- la relation entre le domaine d'une couverture et une gamme d'attributs connexe.

Du point de vue de la qualité des données géospaciales, suivre une méthode formelle pour la conception du schéma de la géométrie et des fonctions de couverture améliorera la cohérence logique des données (élément de qualité des données DQ\_LogicalConsistency de la norme ISO 19157). La figure 14 présente un exemple de schéma d'orthoimagerie numérique (tiré de Maitra, 2004).

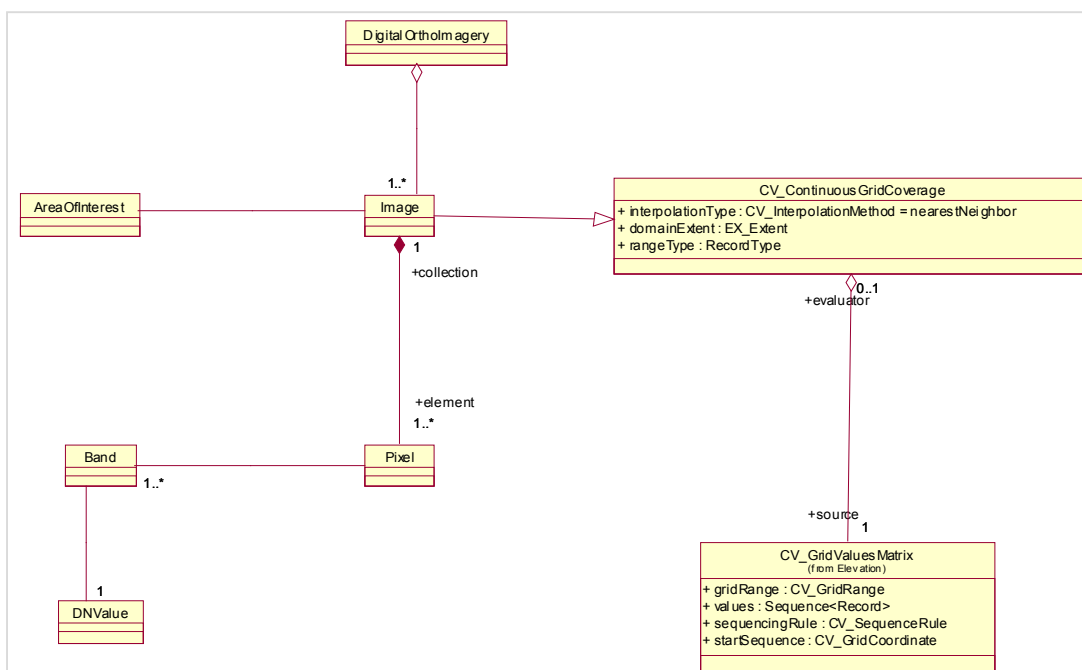


Figure 14 : Exemple de schéma d’orthoimagerie numérique dans (Maitra, 2004).

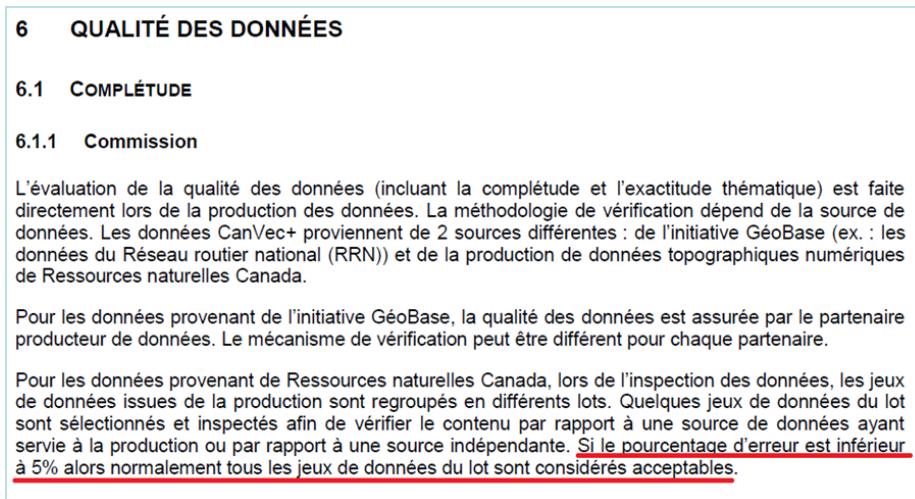
Du point de vue de la gestion du risque, le schéma d’application peut être enrichi d’éléments de gestion du risque comme l’identification des risques, la probabilité qu’ils se présentent, les répercussions, les mesures correctives et les avertissements suggérés.

La norme ISO 19131:2007, Information géographique – Spécifications de contenu informationnel (avec la norme ISO 19131:2007/Amd 1:2011 [Exigences relatives à l’inclusion d’un schéma d’application et d’un catalogue d’objets géographiques et au traitement des couvertures dans un schéma d’application]), peut être utilisée pour définir les spécifications détaillées du produit de données à créer ou mettre à jour. Cette norme :

- définit les exigences pour les spécifications de produits de données géographiques en fonction des concepts d’autres normes internationales ISO 19100;
- fait référence au schéma d’application, au catalogue d’entités ou au schéma de la géométrie de couverture.

Du point de vue de la qualité des données géospatiales, la description formelle des spécifications d’un produit de données à créer ou mettre à jour améliorera la cohérence logique des données (élément de qualité des données DQ\_LogicalConsistency de la norme ISO 19157). Les spécifications de produit de données sont également utiles pour l’ensemble du processus d’évaluation de la qualité des données, puisqu’elles précisent les valeurs attendues pour les éléments de qualité des données (DQ\_Element) de la norme ISO 19157, c’est-à-dire les critères de qualité. La figure 15 présente un extrait des spécifications de produit de données de CanVec+ qui

montre les valeurs attendues pour l'élément de qualité des données DQ\_CompletenessCommission (Ressources naturelles Canada, 2014).



**Figure 15 : Extrait des spécifications de produit de données de CanVec+ (dans Ressources naturelles Canada, 2014).**

Du point de vue de la gestion du risque, les normes sur les spécifications de produit ne comprennent pas explicitement de suggestions de métadonnées sur la gestion du risque (identification du risque, probabilité, incidences, mesures correctives et avertissements suggérés), mais il est possible de les ajouter aux spécifications.

#### **4.1.2 Phase de mise en œuvre**

Pendant la phase de mise en œuvre d'un produit de données, des contraintes d'intégrité supplémentaires sont appliquées (c.-à-d. en plus de celles déjà définies dans le schéma d'application et le catalogue d'entités) pour mieux contrôler la cohérence des données. Les contraintes d'intégrité peuvent être spatiales, temporelles ou descriptives. Elles peuvent renforcer les schémas d'application et les catalogues d'entités, mais, à cette phase, elles sont principalement orientées vers l'assurance de la qualité de la structure physique des données et vers le contrôle des erreurs potentielles de saisie de données. Différents niveaux de contraintes d'intégrité peuvent être définis :

- Intra-champ (p. ex. les valeurs d'un champ numérique doivent se trouver entre 0 et 1);
- Inter-champs (p. ex. si la valeur de l'attribut de classification des routes est « national », la valeur de l'attribut de vitesse maximale ne peut être nulle);
- Intra-entité (p. ex. la date d'une évaluation de maison mise à jour ne peut être antérieure à la date de l'évaluation précédente);

- Inter-entités (p. ex. la taille d'un « immeuble » doit être inférieure à celle de la « parcelle » sur laquelle il est construit);
- Intra-catégorie d'entité (p. ex. « immeuble » ne peut croiser « immeuble »);
- Inter-catégories d'entités (p. ex. « route » ne peut traverser « lac »);
- Intra-thème (p. ex. « rivière » peut être connecté à « canal »);
- Inter-thèmes (p. ex. « barrage » peut avoir les mêmes caractéristiques géométriques que « route »).

Du point de vue de la qualité des données géospatiales, l'utilisation de contraintes d'intégrité peut aider à contrôler tous les éléments relatifs à la qualité des données (DQ\_Element) de la norme ISO 19157. La figure 16 présente un exemple de dépôt de contraintes (adapté de Normand, 1999). Du point de vue de la gestion du risque, les contraintes d'intégrité pourraient contribuer à la réduction des incidences des données erronées lorsqu'elles sont appliquées à des entités ou attributs plus critiques (p. ex. limites des parcelles, valeur de la propriété et zonage).

Objet en relation			Figure
Intra-Thème			
Barrage(L)	Connecte	Barrage (L)	
Barrage(L)	Connecte	Conduit (L)	
Barrage(L)	Connecte	Cours d'eau (L)	
Barrage(L)	Connecte	Digue/Levée (L)	
Barrage(L)	Connecte	Étendue d'eau (S)	
Barrage(L)	Partage	Étendue d'eau (S)	
Inter-thème			
Barrage(L)	Partage	Chemin de fer (L)	

Figure 16 : Exemple de dépôt de contrainte (adapté de Normand, 1999).

### 4.1.3 Phase de production

Pendant la phase de production, l'évaluation de la qualité peut être utilisée pour contrôler les résultats de toute étape de l'acquisition ou du traitement de données ou pour contrôler le produit final. Cela permet de vérifier si la qualité évaluée respecte les critères de qualité définis dans la phase de conception.

La norme ISO 19157:2013, Information géographique – Qualité des données, peut être utilisée pour appuyer l'évaluation détaillée de la qualité (Organisation internationale de normalisation, 2013). Cette norme explique :

- les composantes à utiliser pour décrire la qualité des données;
- les composantes et la structure de contenu d'un registre pour les mesures de la qualité des données;



- les procédures générales d'évaluation de la qualité des données géographiques;
- les principes pour rapporter la qualité des données.

Les étapes détaillées de l'évaluation de la qualité des données géospatiales sont présentées à la section 4.2, avec des exemples.

La norme ISO 19115-1:2014, Information géographique – Métadonnées – Partie 1 : Principes de base, et la norme ISO 19115-2:2009, Information géographique – Métadonnées – Partie 2 : Extensions pour les images et les matrices, peuvent être utilisées pour transmettre des renseignements au sujet des données contenues dans le produit de données (c.-à-d. les métadonnées) (Organisation internationale de normalisation, 2014), (Organisation internationale de normalisation, 2009). Cette norme contient :

- des sections de métadonnées obligatoires et conditionnelles, des entités de métadonnées et des éléments de métadonnées;
- l'ensemble minimal de métadonnées requis pour servir la plupart des applications de métadonnées (découverte de données, tests d'utilisation, accès aux données, transfert des données et utilisation de données et de services numériques);
- des éléments de métadonnées optionnels pour permettre une description normalisée plus détaillée des ressources, au besoin;
- une méthode pour étendre les métadonnées de manière à leur permettre de combler des besoins spécialisés.

Le profil nord-américain de la norme ISO 19115:2003 est utilisé dans le contexte canadien pour décrire les métadonnées.

Les métadonnées peuvent être codées dans le format XML à l'aide de la norme ISO 19139:2007, Information géographique – Métadonnées – Implémentation de schémas XML.

Du point de vue de la qualité des données géospatiales, tous les éléments relatifs à la qualité des données (DQ\_Elements) de la norme ISO 19157 sont rapportés à l'aide de métadonnées. La figure 17 présente un extrait des métadonnées du jeu de données 082C de CanVec+, fondées sur le profil nord-américain de la norme ISO 19115, et présente des renseignements au sujet de l'élément relatif à la qualité des données DQ\_CompletenessCommission (dans Ressources naturelles Canada, 2015). Il est également possible de produire un rapport optionnel sur la qualité. Pour les contextes moins formels, au lieu ou en supplément des métadonnées ou de rapports sur la qualité, d'autres éléments pourraient être plus appropriés pour montrer les renseignements sur la qualité des données, comme l'explique la section suivante.

```

- <gmd:DQ_CompletenessCommission>
- <gmd:result>
- <gmd:DQ_ConformanceResult>
  <gmd:specification xlink:title="CanVec+ - Data Product Specifications, Edition 1.1"
  xlink:href="http://ftp2.cits.rncan.gc.ca/pub/canvec+/doc/CanVec+_product_specifications.pdf"/>
- <gmd:explanation xsi:type="gmd:PT_FreeText_PropertyType">
  <gco:CharacterString>The data quality assessment (including completeness and thematic accuracy) is directly performed during the
  data production process. The validation method applied depends on the data source used. CanVec+ data come from 2 different
  sources: GeoBase initiative (ex: National Road Network (NRN) data) and Natural Resources Canada digital topographic data. For
  GeoBase initiative data, data quality is ensured by the producer (and partner). The validation mechanism used may vary from
  one partner to another. For Natural Resources Canada data, during data inspection, datasets are produced into distinct batches.
  A few datasets are selected from each batch and inspected in order to check their contents and compare them to a data source
  used in production or another independent source. If the percentage of error detected is less than 5 %, then all datasets in the
  batch are normally considered acceptable.</gco:CharacterString>
- <gmd:PT_FreeText>
- <gmd:textGroup>
  <gmd:LocalisedCharacterString locale="#FR">L'évaluation de la qualité des données (incluant la complétude et l'exactitude
  thématique) est faite directement lors de la production des données. La méthodologie de vérification dépend de la source
  de données. Les données du produit CanVec+ proviennent de 2 sources différentes : de l'initiative GéoBase (ex. : les
  données du Réseau routier national RRN)) et de la production de données topographiques numériques de Ressources
  naturelles Canada. Pour les données provenant de l'initiative GéoBase, la qualité des données est assurée par le
  partenaire producteur de données. Le mécanisme de vérification peut être différent pour chaque partenaire. Pour les
  données provenant de Ressources naturelles Canada, lors de l'inspection des données, les jeux de données issus de la
  production sont regroupés en différents lots. Quelques jeux de données du lot sont sélectionnés et inspectés afin de
  vérifier le contenu par rapport à une source de données ayant servi à la production ou par rapport à une source
  indépendante. Si le pourcentage d'erreurs détecté est inférieur à 5%, alors tous les jeux de données du lot sont
  normalement considérés acceptables.</gmd:LocalisedCharacterString>
  </gmd:textGroup>
</gmd:PT_FreeText>
</gmd:explanation>
- <gmd:pass>
  <gco:Boolean>true</gco:Boolean>
</gmd:pass>

```

Figure 17 : Extrait des métadonnées du produit de données 082C de CanVec+ (dans Ressources naturelles Canada, 2015).

#### 4.1.4 Phase de livraison

À la phase de livraison, en plus du produit de données, de la documentation doit également être offerte à l'utilisateur. Habituellement, une communication convenable combine plusieurs produits et services d'information. Dans le contexte E-E, il s'agit de métadonnées sur la qualité avec des rapports optionnels sur la qualité, mais recommandés (potentiellement un par type d'utilisation/utilisateur).

Dans les contextes E-E et C-C, il est fortement recommandé de fournir un manuel de l'utilisateur comportant les éléments suggérés suivants (Gervais, 2004), selon des considérations juridiques :

- Licence;
- Garanties;
- Installation;
- Description du produit;
- Résolution (spatiale, temporelle, descriptive) des données;
- Conseils généraux;
- Spécifications fonctionnelles;
- Usages recommandés;
- Usages non recommandés;
- Avertissements et sécurité;
- Dépannage;
- Spécifications techniques.

Les métadonnées respectant la norme ISO 19115 contiennent déjà une partie de cette information, mais dans un jargon technique que la plupart des utilisateurs ne comprennent pas (voir (Gervais, 2004) pour connaître la correspondance entre les éléments du manuel de l'utilisateur et les métadonnées sur la qualité).

Du point de vue de la qualité des données géospatiales, les métadonnées sur la qualité sont obligatoires et les rapports sur la qualité sont optionnels. Dans tous les cas, la valeur de tous les éléments relatifs à la qualité des données respectant la norme ISO 19157 (DQ\_Elements) influencera la documentation du produit.

Du point de vue de la gestion du risque, le contenu des documents offerts est un résultat direct des stratégies de gestion du risque adoptées. Habituellement, ces documents contiennent un ensemble de stratégies d'atténuation des risques et de produits de communication (voir les sections 4.3.3 et 4.4.4 pour en savoir plus). Dans tous les cas, un bon manuel de l'utilisateur est recommandé, comme dans tous les marchés de masse développés. Différentes versions de ce manuel conçues pour chaque type d'usage peuvent être produites (p. ex. le « Guide de l'utilisateur, Enquête sur les dépenses des ménages, 2012 » (Statistiques Canada, 2012).

La figure 18 présente des exemples de sections d'un manuel de l'utilisateur de données géospatiales (dans Gervais, 2004).

<b>Garantie</b>	<i>Le manufacturier garantit que la topologie du jeu de données du produit permet l'utilisation appropriée des fonctionnalités du produit.</i>
<b>Mises en garde et sécurité</b>	<i>Mises en garde relatives aux données : Les données comprises dans le produit sont mises à jour mensuellement. Il est fortement recommandé de procéder aux mises à jour afin que le produit puisse continuer à fonctionner correctement et sécuritairement.</i>
<b>Support technique</b>	<i>Service à la clientèle: Pour toute question concernant l'utilisation de ce produit, svp contacter le Service à la clientèle au 1-800-***-****.</i>

Figure 18 : Exemples de sections d'un manuel de l'utilisateur de données géospatiales (Gervais, 2004).

Des avertissements peuvent être intégrés à la documentation fournie pour faciliter la lecture et informer les utilisateurs. La norme ISO/TC 145/ISO 3864-2 symboles graphiques (Organisation internationale de normalisation, 2004) établit les principes de la préparation, de la coordination et de l'application de symboles graphiques. Les symboles et étiquettes sont d'excellentes façons de transmettre la signification du risque :

- Type de risque (danger ou action positive);




- Niveau de risque;
- Description du risque;
- Mesures à prendre par rapport aux conséquences.

La figure 19 présente des exemples de symboles en fonction du niveau de danger (American National Standards Institute, 2006).



**Figure 19 : Exemples de symboles pouvant être utilisés dans les avertissements.**

La figure 20 présente un exemple d'utilisation de symboles pour faciliter la lecture d'un rapport sur la qualité (dans Gervais, Bédard, et Larrivée, 2007).

	Symbole	Description
	Avertissement	Symbole signifiant l'existence d'une situation problématique en regard de l'utilisation d'un jeu de données
	Interdiction	Symbole signifiant l'impossibilité d'utiliser un jeu de données afin de satisfaire un ou plusieurs besoins identifiés par le client
	Obligation	Symbole signifiant que des actions spécifiques devraient idéalement être entreprises avant d'utiliser un jeu de données

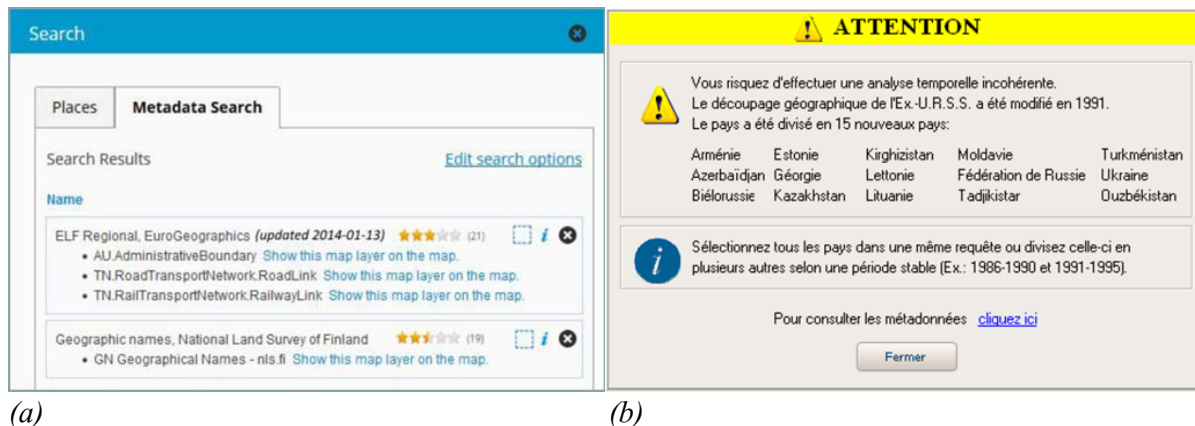
**Figure 20 : Exemple d'utilisation de symboles pour faciliter la lecture d'un rapport sur la qualité des données.**

#### 4.1.5 Phase d'utilisation

Pour la phase d'utilisation d'un produit de données, plusieurs stratégies de gestion du risque peuvent être mises en place pour aider les utilisateurs ayant des questions sur la qualité des données et l'utilisation appropriée des données. Dans de nombreux cas, la stratégie sera une poursuite des moyens mis en place pour transmettre les données tout en ayant le souci de la qualité. Parmi les exemples se trouvent une ligne téléphonique 1-800 gratuite, du clavardage en direct, un forum sur le Web à l'intention des utilisateurs, des webinaires réguliers et des listes de courriel pour informer les utilisateurs des nouveaux usages recommandés ou interdits ou des mises à jour de la qualité. Cela peut se rendre jusqu'à l'attestation, obligatoire ou volontaire, des utilisateurs.

Il est possible de mettre en place plus d'interaction à l'aide de l'information sur la qualité volontaire (VQI), qui proviennent des utilisateurs qui signalent volontairement les problèmes de qualité (p. ex. TomTom MapShare) ou qui utilisent le système de cote sur cinq étoiles et de commentaires

que l'on retrouve souvent sur le Web (p. ex. Amazon, Apple Store). Idéalement, une application sensible à la qualité paramétrée sur mesure pour différents profils d'utilisateurs serait le résultat des mesures précédentes. Il s'agit d'une application qui, selon l'interaction de l'utilisateur avec celle-ci, affiche des avertissements pour informer l'utilisateur au sujet des risques potentiels de l'utilisation inappropriée des données géospatiales.



**Figure 21 : Exemple de (a) renseignements sur la qualité participatifs à l'aide d'un système de cote sur cinq étoiles (Koistinen, 2015), et (b) avertissement lié au contexte pour une recherche spatio-temporelle internationale couvrant une période commençant en 1990 comprenant les trois éléments recommandés par les normes ISO : niveau de risque, nature du problème et mesure à prendre pour résoudre le problème (Gervais, Bédard, Lévesque, Bernier et Devillers, 2009).**

## 4.2 Mettre en pratique l'évaluation de la qualité des données géospatiales

Les étapes du processus d'évaluation de la qualité des données géospatiales ont été présentées à la figure 7. Les étapes du processus officiel sont fondées sur la norme ISO 19157 (Organisation internationale de normalisation, 2013) :

- Spécification :
  - Spécifier les unités de qualité des données;
  - Spécifier les mesures de la qualité des données;
  - Spécifier les procédures d'évaluation de la qualité des données;
- Évaluation :
  - Extrants de l'évaluation de la qualité des données;
  - Extrants de l'évaluation de la métaqualité;
- Rapports :
  - Signaler la qualité des données sous forme de métadonnées;

- Transmettre un rapport sur la qualité des données (facultatif).

## 4.2.1 Spécification formelle de qualité des données géospatiales

### Spécifier les unités de qualité des données

Une unité de qualité des données est composée d'une portée (MD\_Scope) et d'un à plusieurs éléments relatifs à la qualité des données (DQ\_Element, voir la figure 8). La portée précise l'étendue, spatiale ou temporelle, et les caractéristiques communes qui cernent les données en fonction desquelles la qualité des données doit être évaluée.

Exemple d'unité de qualité 1 :

- MD\_Scope : jeu de données 082C
- DQ\_Elements : DQ\_LogicalConsistency, DQ\_Completeness

Exemple d'unité de qualité 2 :

- MD\_Scope : type d'entité (borne d'incendie)
- DQ\_Element : DQ\_QuantitativeAttributeAccuracy

### Spécifier les mesures de la qualité des données

Un élément relatif à la qualité des données doit se rapporter à une mesure au moyen d'une référence (DQ\_MeasureReference). La référence est composée des éléments suivants :

- Identification de la mesure;
- Nom de la mesure;
- Description de la mesure.

La norme ISO 19157 comporte une liste de mesures normalisées. De nouvelles mesures peuvent être créées.

Voici des exemples de mesures pour l'élément relatif à la qualité des données DQ\_CompletenessCommission :

- Item en trop;
- Nombre d'items en trop;
- Taux d'items en trop;
- Nombre d'instances d'entité en double.

Voici des exemples de mesures pour l'élément relatif à la qualité des données DQ\_CompletenessOmission :

- Item manquant;
- Nombre d'items manquants;
- Taux d'items manquants.

### Spécifier les procédures d'évaluation de la qualité des données

Les valeurs de mesure de la qualité des données sont déterminées à l'aide de méthodes d'évaluation. Un ensemble de méthodes d'évaluation compose une procédure d'évaluation. Les méthodes d'évaluation de la qualité des données (DQ\_EvaluationMethod) peuvent être divisées en deux principales catégories : directes et indirectes. Les méthodes d'évaluation directes comportent une comparaison des données avec des données de référence internes ou externes. Ces méthodes peuvent être fondées sur des inspections complètes (c.-à-d. que chaque élément de la population est inspecté) des éléments connexes (DQ\_FullInspection), ou fondées sur un échantillon (DQ\_SampleBasedInspection). Les méthodes d'évaluation indirectes déduisent ou estiment la qualité des données à l'aide de renseignements sur les données, comme le lignage (DQ\_IndirectEvaluation).

Voici des exemples de méthodes d'évaluation fondées sur des échantillons :

- Échantillonnage orienté par les entités (échantillonnage non spatial) : fondé sur les attributs non spatiaux des entités
- Échantillonnage régulier (non aléatoire) orienté par le secteur
- Échantillonnage aléatoire orienté par le secteur

La figure 22 présente un exemple de méthode d'échantillonnage régulier orienté par le secteur.

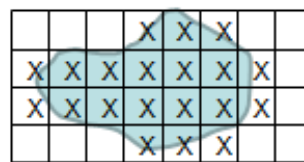


Figure 22 : Exemple de méthode d'échantillonnage régulier (non aléatoire) orienté par le secteur.

## 4.2.2 Évaluation formelle de la qualité des données géospatiales

### Évaluer la qualité des données géospatiales

Au moins un résultat en matière de qualité des données doit être fourni pour chaque élément relatif à la qualité des données. Le résultat de l'évaluation peut être un résultat quantitatif (DQ\_QuantitativeResult), un résultat de conformité (DQ\_ConformanceResult), un résultat descriptif (DQ\_DescriptiveResult) ou une couverture.

---

Exemple de DQ\_QuantitativeResult :

- (DQ\_CompletenessCommission), nombre d'articles en trop : 3

Exemple de DQ\_ConformanceResult :

- (DQ\_CompletenessCommission), nombre d'articles en trop : acceptable

Exemple de DQ\_DescriptiveResult :

- (DQ\_LogicalConsistency), Conformité au schéma conceptuel : « Les règles du schéma conceptuel CanVec+ ont toutes été enregistrées et validées dans la base de données contenant le produit CanVec+. Ceci permet d'assurer une cohérence conceptuelle entre le schéma conceptuel et le produit CanVec+. » (Tiré des métadonnées du jeu de données 082C de CanVec+ (Ressources naturelles Canada, 2015))

Un résultat de couverture est le résultat d'une évaluation de la qualité des données, organisé sous forme de couverture. Il est documenté dans la norme ISO 19115-2:2009 (Organisation internationale de normalisation, 2009).

### Évaluer la métaqualité d'une évaluation de la qualité

Les éléments de métaqualité sont un ensemble d'énoncés quantitatifs et qualitatifs au sujet de l'évaluation de la qualité et de ses résultats. La métaqualité peut être exprimée à l'aide des éléments suivants :

- Confiance (DQ\_Confidence) : fiabilité d'un résultat de qualité de données;
- Représentativité (DQ\_Representativity) : degré auquel l'échantillon utilisé a produit un résultat représentatif des données se trouvant dans la portée de la qualité des données;
- Homogénéité (DQ\_Homogeneity) : uniformité attendue ou testée des résultats obtenus pour une évaluation de la qualité des données.



Exemple de DQ\_Confidence :

- Écart-type ou intervalle de confiance pour un niveau de confiance donné

Exemple de DQ\_Representativity:

- Toutes les zones géographiques et les périodes concernées sont couvertes et la population est suffisamment large

Exemple de DQ\_Homogeneity:

- Comparaison des résultats de l'évaluation de plusieurs segments d'un jeu de données global exprimé à l'aide d'erreurs-types

### 4.2.3 Rapports formels sur la qualité des données géospatiales

Présenter la qualité sous forme de métadonnées

Selon la norme ISO 19157, la qualité des données est présentée sous forme de métadonnées conformément à la norme ISO 19115-1:2014, Information géographique – Métadonnées – Partie 1 : Principes de base (Organisation internationale de normalisation, 2014), et à la norme ISO 19115-2:2009, Information géographique – Métadonnées – Partie 2 : Extensions pour les images et les matrices (Organisation internationale de normalisation, 2009).

La figure 23 présente un exemple des résultats de l'évaluation de l'élément de qualité DQ\_CompletenessCommission.

```
- <gmd:DQ_CompletenessCommission>
- <gmd:result>
- <gmd:DQ_ConformanceResult>
  <gmd:specification xlink:title="CanVec+ - Data Product Specifications, Edition 1.1"
  xlink:href="http://ftp2.cits.rncan.gc.ca/pub/canvec+/doc/CanVec+_product_specifications.pdf"/>
- <gmd:explanation xsi:type="gmd:PT_FreeText_PropertyType">
  <gco:CharacterString>The data quality assessment (including completeness and thematic accuracy) is directly performed during the
  data production process. The validation method applied depends on the data source used. CanVec+ data come from 2 different
  sources: GeoBase initiative (ex: National Road Network (NRN) data) and Natural Resources Canada digital topographic data. For
  GeoBase initiative data, data quality is ensured by the producer (and partner). The validation mechanism used may vary from
  one partner to another. For Natural Resources Canada data, during data inspection, datasets are produced into distinct batches.
  A few datasets are selected from each batch and inspected in order to check their contents and compare them to a data source
  used in production or another independent source. If the percentage of error detected is less than 5 %, then all datasets in the
  batch are normally considered acceptable.</gco:CharacterString>
- <gmd:PT_FreeText>
- <gmd:textGroup>
  <gmd:LocalisedCharacterString locale="#FR">L'évaluation de la qualité des données (incluant la complétude et l'exactitude
  thématique) est faite directement lors de la production des données. La méthodologie de vérification dépend de la source
  de données. Les données du produit CanVec+ proviennent de 2 sources différentes : de l'initiative GéoBase (ex. : les
  données du Réseau routier national RRN)) et de la production de données topographiques numériques de Ressources
  naturelles Canada. Pour les données provenant de l'initiative GéoBase, la qualité des données est assurée par le
  partenaire producteur de données. Le mécanisme de vérification peut être différent pour chaque partenaire. Pour les
  données provenant de Ressources naturelles Canada, lors de l'inspection des données, les jeux de données issus de la
  production sont regroupés en différents lots. Quelques jeux de données du lot sont sélectionnés et inspectés afin de
  vérifier le contenu par rapport à une source de données ayant servi à la production ou par rapport à une source
  indépendante. Si le pourcentage d'erreurs détecté est inférieur à 5 %, alors tous les jeux de données du lot sont
  normalement considérés acceptables.</gmd:LocalisedCharacterString>
  </gmd:textGroup>
  </gmd:PT_FreeText>
</gmd:explanation>
- <gmd:pass>
  <gco:Boolean>true</gco:Boolean>
</gmd:pass>
```

**Figure 23 : Exemple de métadonnées sur la qualité (DQ\_CompletenessCommission) tiré des métadonnées 082C de CanVec+ (Ressources naturelles Canada, 2015).**

### Présenter la qualité dans un rapport sur la qualité des données

Afin de donner plus de détails que ceux transmis sous forme de métadonnées et dans un format plus facile à comprendre que les métadonnées, un rapport sur la qualité autonome peut être créé. Ce rapport sur la qualité est utilisé à titre de complément aux métadonnées. Les métadonnées doivent fournir une référence au rapport sur la qualité autonome si ce dernier existe. Si un jeu de données est destiné à différentes catégories d'utilisation, différentes versions du rapport sur la qualité pourraient être requises, car puisque les besoins varient, les qualités externes varient aussi.

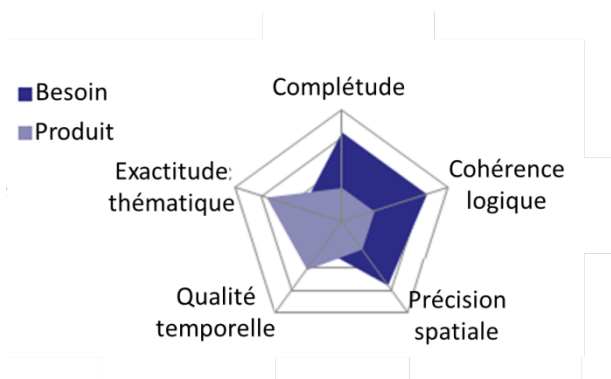
La figure 24 illustre un extrait d'un rapport sur la qualité des données présentant des résultats détaillés pour l'élément DQ\_CompletenessCommission pour chaque catégorie d'entité d'un jeu de données (tiré de la norme ISO 19157:2013) (Organisation internationale de normalisation, 2013).

Classe d'entités	Nombre d'instances attendu	Compte des items en trop	Pourcentage des items en trop <sup>a</sup>	Compte des items manquants	Pourcentage des items manquants <sup>b</sup>
Sentier	7	1	14	3	43
Route	5	2	40	0	0
Arbre	25	3	12	2	8
Édifice industriel	4	0	0	2	50
Maison	10	1	10	1	10

a pourcentage de commission = nombre d'items inclus / nombre d'items de l'univers de discours x 100  
 b pourcentage de omission = nombre d'items omis / nombre d'items de l'univers de discours x 100








**Figure 24 : Résultats détaillés pour l'élément DQ\_CompletenessCommission présentés dans un rapport sur la qualité des données (exemple tiré de la norme ISO 19157 (Organisation internationale de normalisation, 2013)).**

La figure 25 illustre un cadre de représentation fondé sur des diagrammes en radar qui aide à comparer des éléments relatifs à la qualité des données (DQ\_Elements) d'un jeu de données au niveau de qualité requis.



**Figure 25 : Utilisation d'un diagramme en radar pour comparer des éléments relatifs à la qualité des données d'un jeu de données au niveau de qualité requis.**

La figure 26 présente un autre exemple de contenu d'un rapport sur la qualité des données : un sommaire d'une évaluation de la qualité d'une catégorie d'entités (adapté d'un rapport privé par Gervais, Bédard, et Larrivée, 2007).

Classe d'entités	Résultat
Route	
Édifice	
Hydrologie	
Stationnement	
	Généralement bon, mais peut présenter des problèmes mineurs en relation avec les besoins identifiés
	Utilisable mais peut ne pas remplir tous les besoins identifiés
	La classe d'entités ne convient pas pour les besoins identifiés

**Figure 26 : Sommaire d'évaluation de la qualité par catégorie d'entités, tiré d'un rapport sur la qualité des données (Gervais, Bédard, et Larrivée, 2007).**

La figure 27 montre un aperçu du tableau de bord pour des métadonnées sur la qualité fondées sur les cinq éléments de qualité des normes ISO (indicateurs) au niveau de l'incident ainsi qu'un indicateur global de la qualité pour ce type d'entité (route) (Devillers, Bédard, Jeansoulin et Moulin, 2007).

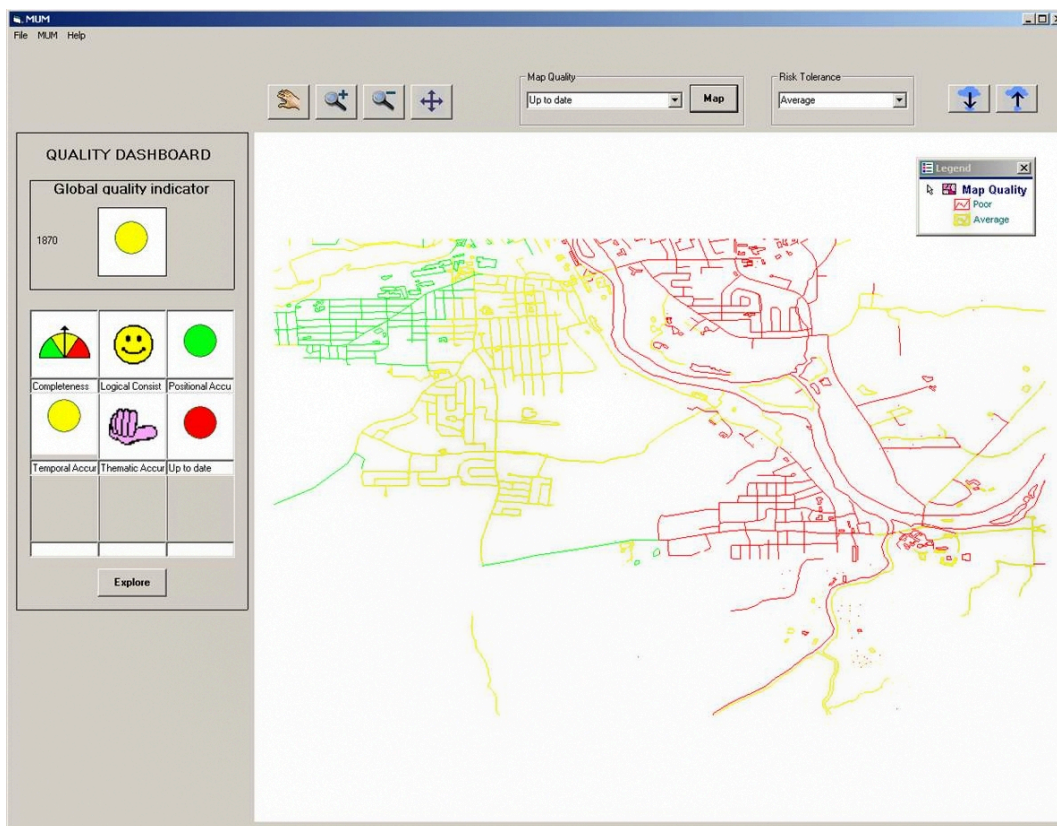


Figure 27 : Aperçu du tableau de bord pour des métadonnées sur la qualité (Devilleers, Bédard, Jeansoulin et Moulin, 2007).

#### 4.2.4 Évaluation informelle de la qualité des données géospatiales

Dans les contextes E-E et E-C, les processus moins formels d'évaluation et de présentation des données sur la qualité doivent être inspirés de l'approche de la norme ISO 19157 et bien se servir de la souplesse de cette dernière en ce qui a trait au niveau de résolution de l'analyse de la qualité. Dans le contexte E-E, il est possible qu'aucune métadonnée ne soit fournie séparément du rapport, mais, pour les rapports sur la qualité destinés à être réutilisés ou transmis par l'intermédiaire d'une infrastructure de données, l'utilisation de métadonnées sur la qualité est recommandée. En plus de faciliter la découverte de données ainsi que la communication et la compréhension entre les experts, les métadonnées facilitent la communication entre les appareils. Par conséquent, elles constituent un pas vers une interopérabilité axée sur la qualité à l'aide de fonctions d'association de métadonnées sur la qualité (p. ex. association de contexte comme le décrivent (Sboui et Bédard, 2012)). Dans le contexte E-C, les fournisseurs de données ne peuvent compter sur les métadonnées pour informer les utilisateurs au sujet de la qualité des données, puisqu'elles constituent un langage technique qu'ils ne comprennent habituellement pas. Les fournisseurs doivent plutôt utiliser des méthodes plus simples pour représenter la qualité des données, comme celles présentées dans les figures ci-dessus ou qui suivent.

Dans les contextes E-C et C-C, les utilisateurs de données géospatiales peuvent utiliser diverses méthodes pour évaluer par eux-mêmes la qualité des données externes. Ces méthodes sont souvent centrées sur le cadre de représentation de la qualité des données utilisé et visent à rassembler suffisamment d'information sur la qualité pour offrir une représentation adéquate. Des exemples de cadres de représentation ont été fournis à la section précédente.

Les figures 28 et 29 présentent l'utilisation d'un système de cotation sur cinq étoiles, dans un globe virtuel où la représentation en trois dimensions des immeubles peut être cotée de manière collaborative par les utilisateurs (Jones, 2011) et dans un contexte d'infrastructure de données spatiales, où chaque jeu de données disponible peut être coté par les utilisateurs (Koistinen, 2015), respectivement.



Figure 28 : Exemple de l'utilisation d'un système de cotation sur cinq étoiles pour coter la représentation d'immeubles dans un globe virtuel (Jones, 2011).

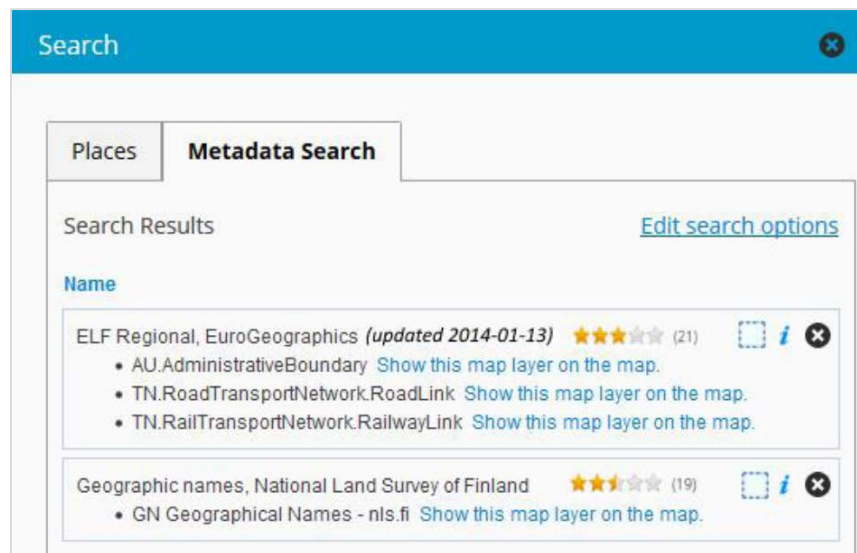


Figure 29 : Exemple de l'utilisation d'un système de cotation sur cinq étoiles dans une infrastructure de données spatiales (Koistinen, 2015).

## 4.3 Gérer les risques de l'utilisation inappropriée des données géospatiales en pratique

Le cadre de gestion du risque décrit dans la norme ISO 31000 (Management du risque – Principes et lignes directrices) (Organisation internationale de normalisation, 2009) a été présenté à la figure 10. Les étapes générales du cadre sont les suivantes :

- Décrire le contexte organisationnel et de gestion du risque;
- Évaluer les risques;
- Préparer les réponses aux risques;
- Communiquer les risques;
- Surveiller et examiner les risques.

### 4.3.1 Description du contexte de la gestion du risque

La gestion du risque commence par une description du contexte (général, externe, interne et spécifique) de l'organisation et du cadre de gestion du risque ciblé. La description du contexte aide à comprendre en quoi l'activité prévue cadre dans l'organisation plus large, avec le marché et la société et avec l'approche en matière de gestion du risque de l'organisation, afin de déterminer la portée de la stratégie de gestion du risque. La figure 30 présente un exemple de description du contexte.

**1. Contexte**

Il y a 3 autres sources de données publiques similaires à la nôtre. Leur coût et exactitude globale sont semblables aux nôtres et les clients ont de la difficulté à sélectionner la source qui répond le mieux à leurs besoins. L'évolution du marché est incertaine, tout comme les actions de la compétition. En mettant en place un nouveau système de gestion de la qualité et de gestion du risque, cela nous aidera à améliorer significativement la communication avec des clients potentiels et les aidera à comprendre comment nos données peuvent remplir leurs besoins. Un client potentiel qui comprend mieux le contexte a plus de chance de devenir un client réel. Notre objectif est d'augmenter notre part de marché de 10% en deux ans. Si nous mettons en place la stratégie, les coûts et les risques sont .... Si nous ne la mettons pas en place, les coûts et les risques sont ...

Figure 30 : Exemple de description de contexte de gestion du risque.

### 4.3.2 Évaluation des risques

Une fois que le contexte est défini, le processus d'évaluation du risque peut commencer. L'évaluation du risque comporte trois principales opérations : détermination du risque, analyse du risque et évaluation du risque.

L'identification du risque génère une liste détaillée de risques présentés par l'utilisation inappropriée des données géospatiales. Cette étape peut être réalisée à l'aide :

- de l'analyse de documents existants (p. ex. spécifications, contrats, organigrammes des tâches);
- d'entrevues;
- de séances de remue-méninges;
- d'une approche de collaboration avec les utilisateurs;
- ...

Le tableau 2 présente un exemple de liste de risques potentiels de l'utilisation inappropriée des données géospatiales en ce qui a trait aux entités/attributs présents dans un jeu de données géospatiales pour un contexte agricole (tiré de Grira, 2014).

Entité / Attribut	Risques identifiés
Parcelle cultivée	<b>R-1</b> : l'utilisateur peut considérer que toute la zone définie par les limites cadastrales est cultivée. Certaines zones peuvent ne pas l'être, comme les boisés, les sols rocheux ou les zones proches des limites cadastrales.
Plaine inondable	<b>R-2</b> : la notion de plaine inondable est une donnée vague. L'utilisateur peut penser que les limites indiquées sont précises, alors qu'elles sont floues et qu'il n'est pas possible de représenter des limites larges sur la carte.
Zone d'épandage de pesticide	<b>R-3</b> : les zones où le pesticide est appliqué ont des limites floues et larges et leur position dans la parcelle est floue (à cause des techniques et des méthodes d'application). L'utilisateur pourrait penser que la zone est précise, alors que les positions sont floues et ne sont pas prises en compte dans le calcul de la superficie. <u>Note</u> : l'incertitude de R-3 est liée à la zone d'épandage de pesticide, c.-à-d. à ses limites et à son emplacement. Toutefois, l'incertitude R-1 est liée à la parcelle (à ses limites) dans laquelle se trouve la zone d'épandage.

**Tableau 2 : Exemple de liste de risques de l'utilisation inappropriée des données géospatiales dans un contexte agricole (Grira, 2014).**

L'analyse du risque consiste en la compréhension de la nature des risques relevés en déterminant leurs causes, leurs sources, leurs conséquences et leur probabilité. L'étape de l'analyse des risques peut être réalisée à l'aide :

- de l'analyse des leçons tirées de projets précédents;
- de méthodes de simulation;
- d'une étude probabiliste;
- ...

Le tableau 3 présente les résultats de l'analyse des risques présentés au tableau 2 (extrait de Grira, 2014).

Entité / Attribut	Risques relevés	Impact du risque	Probabilité
Parcelle cultivé	<u>R-1</u>	Forte surestimation de la quantité de pesticide / hectare (risque élevé)	Moyenne
Plaine inondable	<u>R-2</u>	Forte sous-estimation de la quantité de pesticide qui peut être présente dans l'eau (élevé)	Moyenne
Zone d'épandage de pesticide	<u>R-3</u>	Forte sous-estimation de la quantité de pesticide qui peut être présente dans l'eau (élevé)	Moyenne

Tableau : 3 : Exemple de résultats d'une analyse des risques (Grira, 2014).

L'évaluation du risque consiste en l'établissement de l'ordre de priorité des risques analysés en fonction d'un niveau de tolérance (propre au contexte). L'évaluation du risque aide à sélectionner des stratégies pour les traitements du risque et doit tenir compte des exigences légales, réglementaires et autres. Cette étape est habituellement effectuée à l'aide d'une matrice de classement. La figure 31 présente un exemple de matrice de classement du risque formée de deux axes pour montrer la probabilité que le risque se présente ainsi que ses répercussions. Le croisement de ces deux axes montre la valeur globale du risque. Le tableau 4 présente les résultats de l'évaluation des risques présentée au tableau 2 (extrait de Grira, 2014).

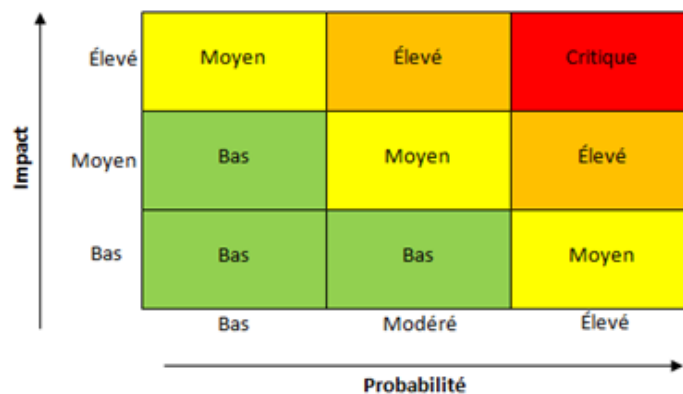


Figure 31 : Exemple de matrice de classement du risque.



Entité / Attribut	Risques relevés	Impact du risque	Probabilité	Évaluation globale du risque
Parcelle cultivée	<b>R-1</b>	Forte surestimation de la quantité de pesticide / hectare (risque élevé)	Moyenne	Élevée
Plaine inondable	<b>R-2</b>	Forte sous-estimation de la quantité de pesticide qui peut être présente dans l'eau (élevé)	Moyenne	Moyenne
Zone d'épandage de pesticide	<b>R-3</b>	Forte sous-estimation de la quantité de pesticide qui peut être présente dans l'eau (élevé)	Moyenne	Moyenne

Tableau 4 : Exemple de résultats d'une évaluation des risques (Griira, 2014).

### 4.3.3 Élaboration de réponses au risque (traitement du risque)

Le traitement du risque comporte la sélection et la mise en œuvre d'une stratégie ou d'un ensemble de stratégies afin de modifier un risque, en vue d'atteindre des niveaux de tolérance acceptés. Quatre catégories de stratégies peuvent être utilisées :

- Atténuation : Modifier les actions pour éliminer ou réduire les conséquences des risques ou leur probabilité
- Évitement : Éliminer une activité pour éliminer le risque
- Transfert/partage : Transférer les répercussions à une autre entité, en totalité ou en partie
- Acceptation : Accepter et prendre volontairement le risque

Voici des exemples de stratégies d'atténuation à appliquer lors de la gestion des risques d'utilisation inappropriée des données géospatiales :

- Améliorer la conception de le jeu de données/la structure du jeu de données;
- Améliorer le contrôle de la qualité du jeu de données (p. ex. ajouter des contraintes d'intégrité);
- Appliquer des normes (p. ex. pour l'interopérabilité de la qualité des données et de la gestion du risque);
- Bien informer les utilisateurs dans une langue qu'ils comprennent (fortement recommandé);
  - Fournir un manuel de l'utilisateur;
  - Offrir une ligne d'aide 1-800 ou une adresse de courriel de type [info@votrecompagnie.com](mailto:info@votrecompagnie.com);
  - Lister les utilisages recommandés et les utilisages non recommandés;
  - Fournir un guide des bonnes pratiques;
  - Former les utilisateurs;

- ...
- Effectuer des essais avec le jeu de données (utilisateurs);
- Comparer avec un autre jeu de données (utilisateurs).

Voici des exemples de stratégies d'évitement des risques :

- Cesser de distribuer ou d'utiliser le jeu de données ou une partie de celui-ci;
- Éliminer une catégorie d'utilisateurs;
- Éliminer un fournisseur de données;
- Interdire l'utilisation pour un usage particulier, clairement et explicitement;
- ...

Voici des exemples de transfert ou de partage des risques :

- Souscrire une assurance;
- Obtenir le jeu de données d'un courtier pouvant donner des conseils sur son utilisation;
- Choisir un jeu de données avec garantie qui explique clairement le partage de risques (c.-à-d., qui est responsable pour quoi) :
  - Le contenu d'une garantie pour des données géospatiales est décrit dans (Plante & Gervais, Geospatial Data Quality Garantie, 2015);
- Faire évaluer le jeu de données par un expert;
- Remplacer une stratégie E-C avec une stratégie E-E en contractant un courtier en données qui offrira lui-même la stratégie E-C;
- Faire évaluer la qualité des données par un expert externe pour les nouveaux usages.

Un utilisateur qui accepte un risque utilise le jeu de données peu importe sans égard au risque et sans apporter de modification, c'est-à-dire qu'il prend le risque. Pour le fournisseur de données, il est important de s'assurer que l'utilisateur le fasse en étant pleinement au courant du risque.

#### **4.3.4 Communication des risques**

La communication du risque comporte les éléments suivants :

- Communiquer avec les personnes en charge d'offrir les données/services, les personnes en charge de la production et les utilisateurs;
- Développer les produits d'information identifiés dans les stratégies de traitement du risque, notamment :
  - E-C et C-C : manuel à l'utilisateur en ligne ou en format papier avec le contenu recommandé précédemment;
  - E-E : intégrer l'information sur les risques dans les métadonnées sur la qualité et dans le rapport sur la qualité;

- Utiliser un vocabulaire adapté aux différents interlocuteurs;
- Informer les utilisateurs : distribuer les produits d’information, garder les utilisateurs informés des nouveaux éléments de contexte, des nouveaux contrôles de qualité, des nouvelles évaluations de qualité, des nouvelles stratégies de gestion du risque mises en place, des nouveaux usages, des nouvelles restrictions, des nouvelles bonnes pratiques, etc.;
- Promouvoir les comités conjoints incluant des expertises diverses;
- Mettre en place une ligne info 1-800, info@qualite.votreorganisation.com;
- Offrir des séminaires, webinaires, formations;
- ...

Des exemples sont offerts dans les sections précédentes.

### 4.3.5 Surveillance et évaluation des risques

La surveillance du risque est la surveillance régulière des risques et de l’efficacité de leur traitement. Cette étape compte les mesures suivantes :

- Surveiller et évaluer systématiquement à l’aide de métriques l’efficacité des mesures prises;
- Mettre à jour la liste initiale des risques identifiés et de leurs caractéristiques;
- Collecter de l’information utile en vue de développer ou de mettre à jour les stratégies de traitement des risques;
- Passer en revue certains aspects liés au processus de planification de la gestion des risques;
- Collecter des commentaires sur la qualité (p. ex., en utilisant un forum Web ou un système de qualité volontaire (VQI) avec barème 5 étoiles;
- Mettre en place de nouvelles restrictions, ajouter des contraintes d’intégrité, offrir de nouvelles formations, construire un registre de gestion de la qualité et du risque;
- ...

Le tableau 5 présente un exemple de registre des risques utilisé pour effectuer le suivi des risques (adapté de Grira, 2014).

Entité / Attribut	Risques identifiés	impact du risque	Propriétaire	Probabilité	Importance du risque	Réponse	Statut	Mesures à prévoir
Parcelle cultivée	<b>R-1</b>	Forte surestimation du rapport quantité pesticide / hectare (risque élevé)	Gestionnaire de projet	Moyen	Élevé	Atténuer (...)	(date) ouvert	Suivi auprès des utilisateurs

Plaine inondable	<b>R-2</b>	Forte sous-estimation de la quantité de pesticide qui peut contaminer l'eau (risque élevé)	Gestionnaire de projet	Moyen	Moyen	Atténuer (...)	(date) ouvert	Suivi auprès des utilisateurs
------------------	------------	--------------------------------------------------------------------------------------------	------------------------	-------	-------	----------------	---------------	-------------------------------

Tableau 5 : Exemple de registre des risques (Grira, 2014).

## 4.4 Communiquer au sujet de la qualité des données géospatiales et des risques liés à l'utilisation en pratique

Des exemples de documentation sur la qualité et les risques se trouvent dans les sections précédentes :

- Dans le contexte E-E (spécialistes) :
  - Produits et services de communication qui ont été relevés dans le contrat E-E;
  - Métadonnées sur la qualité des données (ISO 19115, ISO 19157), voir la figure 23;
  - Spécifications initiales en tant que renseignements supplémentaires pour les nouveaux utilisateurs externes spécialistes;
  - En fonction du contexte détaillé : rapport sur la qualité des données (ISO 19157) potentiel pour chaque type d'utilisation (pour compléter les métadonnées), voir la figure 24;
  - Facultatif : Manuel de l'utilisateur et autres produits/services de communication offerts dans les contextes E-C et C-C, voir la figure 18.
- Dans les contextes E-C et C-C (marché de masse) :
  - Manuel de l'utilisateur fortement recommandé (Gervais, 2004); il peut comprendre les éléments suivants :
    - Avertissements (y compris des symboles [ISO/TC 145/ISO 3864-2]), voir la figure 21;
    - Avis;
    - Utilisations recommandées et non recommandées;
    - Et beaucoup plus;
  - Licence;
  - Garantie;
  - ...

---

## 5. Recommandations

Dans cette section, nous présentons certaines recommandations pour la communauté géospatiale en ce qui a trait à la qualité des données géospatiales. Les recommandations ont été regroupées en fonction du contexte (E-E, E-C et C-C).

### 5.1 Recommandations dans un contexte E-E

---

Dans le contexte E-E, les difficultés liées à la qualité des données géospatiales visent l'utilisation de méthodes plus officielles pour évaluer la qualité des données et pour produire des rapports à ce sujet.

Les recommandations sont les suivantes :

- Faciliter la communication entre les spécialistes au sujet de la qualité des données en adoptant un langage commun fondé sur la norme ISO 19157 (Organisation internationale de normalisation, 2013), sur la norme ISO 19115-1 (y compris les métadonnées sur la qualité) (Organisation internationale de normalisation, 2014) ainsi que sur des rapports sommaires sur la qualité, des ensembles de renseignements sur la qualité et un système consultatif automatisé sous forme de questions et réponses.
  - Favoriser l'augmentation de l'efficacité de la réutilisation et de l'interopérabilité des données géospatiales en adoptant un cadre de référence commun relatif à la qualité (c.-à-d. un ensemble de concepts) fondé sur la norme ISO 19157 (Organisation internationale de normalisation, 2013) et la norme ISO/TS 19158 (Organisation internationale de normalisation, 2012).
  - Faciliter les ententes contractuelles en adoptant le cadre de référence commun et le langage commun proposés ci-dessus.
  - Réduire les risques d'utilisation inappropriée des données géospatiales à l'aide de concepts et de stratégies de gestion des risques fondés sur la norme ISO 31000 (Organisation internationale de normalisation, 2009). Ajouter à ces efforts une augmentation de la qualité des données.
  - Se familiariser graduellement avec les concepts et normes mentionnés ci-dessus en participant à une formation spécialisée.
  - Encourager la grande qualité des données par des processus d'assurance de la qualité et des contrôles de la qualité efficaces (c.-à-d. améliorer la métaqualité), en fonction de la norme ISO 9000 (Organisation internationale de normalisation, 2005), de la norme ISO 19157 (Organisation internationale de normalisation, 2013) et de la norme ISO 19158 (Organisation internationale de normalisation, 2012), réaliser des vérifications de la qualité et appliquer des certificats et l'accréditation.
-

- Élaborer un guide de bonnes pratiques pour représenter visuellement la qualité des données géographiques (cartes, radars, tableaux, symboles d’avertissement, etc. relatifs à la qualité).
- Élaborer un guide de bonnes pratiques pour atténuer les risques d’utilisation inappropriée des données géospatiales.
- Clarifier davantage les rôles et responsabilités entre les parties en intégrant des garanties de qualité.
- Mettre graduellement en place de bonnes pratiques avec l’aide de spécialistes de la qualité des données géospatiales.
- Élaborer une série de produits et services axés sur la qualité comme une garantie de qualité, un certificat de qualité, des vérifications de la qualité, des mécanismes de contrôle et d’assurance de la qualité et l’accréditation des spécialistes de la qualité.

## 5.2 Recommandations dans un contexte E-C

Dans un contexte E-C, la difficulté, en ce qui a trait à la qualité des données géospatiales, est la gestion des risques en vue d’une meilleure protection des consommateurs (et des fournisseurs).

Les recommandations sont les suivantes :

- Faciliter la sélection des données géospatiales en fonction des besoins des utilisateurs (qualité externe) en offrant :
  - une liste des utilisations recommandées et non recommandées;
  - une ligne 1-800 gratuite ou une adresse de courriel de type [info@votrecompagnie.com](mailto:info@votrecompagnie.com);
- Contribuer à l’avancement de la communauté géospatiale et des connaissances spatiales de la société en offrant :
  - des manuels de l’utilisateur écrits dans un langage que les utilisateurs cibles comprennent;
  - des conseils et avertissements clairs (utilisation de symboles, comme dans la norme ISO/TC 145/ISO 3864-2 (Organisation internationale de normalisation, 2004));
  - des garanties réelles (comme pour tout autre produit ou service sur un marché développé);
  - des guides sur les bonnes pratiques;
  - des rapports sommaires sur la qualité et des ensembles de renseignements sur la qualité;
- Stimuler l’analyse de la qualité et la sensibilisation des utilisateurs au moyen de ce qui suit :
  - Renseignements sur la qualité participatifs sur le Web (p. ex. système de cotation sur cinq étoiles avec commentaires);
  - Forums sur le Web à l’intention des utilisateurs;
  - Formation en ligne ou en personne;

- Réduire l’incertitude en ce qui concerne certains sujets touchant la loi en investissant dans des études afin de :
  - comprendre les nouvelles tendances et les nouveaux droits en ce qui touche la vie privée, la possession de données, le droit d’auteur, les données par rapport au service;
  - développer davantage le concept de garantie (Plante & Gervais, Geospatial Data Quality Garantie, 2015);
  - clarifier davantage la responsabilité des fournisseurs de données qui ne sont pas des spécialistes de la cartographie participative;
  - clarifier les responsabilités lorsque les services et données géospatiales traversent les frontières;
  - stimuler l’interopérabilité légale des données géospatiales (Uhlir, 2013);
- Prendre des mesures pour devenir rapidement un marché de masse développé :
  - Formation spécialisée, innovation, collaboration;
- Améliorer les métadonnées (p. ex. faciliter leur utilisation, nouveaux types de métadonnées axées sur la qualité et sur les risques conçus pour les utilisateurs finaux, illustrer de manière explicite les avertissements selon la norme ISO 3864-2:2004 (Symboles graphiques – Couleurs de sécurité et signaux de sécurité – Partie 2 : Principes de conception pour l’étiquetage de sécurité des produits);
- Augmenter la sensibilisation des fournisseurs et utilisateurs de données géospatiales au sujet des risques potentiels d’utilisation inappropriée des données en rassemblant, sur un site Web externe, des exemples de dommages (Remarque : Un site de ce type a été élaboré à l’aide de financement offert par GEOIDE et se trouve à l’adresse <http://dataquality.scg.ulaval.ca>).

## 5.3 Recommandations dans un contexte C-C

Dans le contexte C-C, la difficulté, en ce qui a trait à la qualité des données géospatiales, se trouve dans la sensibilisation.

Les recommandations sont les suivantes :

- Informer rapidement les développeurs de systèmes Web et d’applications pour téléphone intelligent destinés au public au sujet de leurs obligations et de leur responsabilité potentielle en ce qui concerne la qualité des données géospatiales;
- Faciliter la sélection des données géospatiales en fonction des besoins des utilisateurs (qualité externe) en offrant :
  - des listes des utilisations recommandées et non recommandées;
  - une ligne 1-800 gratuite ou une adresse de courriel de type [info@votrecompagnie.com](mailto:info@votrecompagnie.com);
  - des manuels de l’utilisateur écrits dans un langage que les utilisateurs cibles comprennent;

- 
- des conseils et avertissements clairs (utilisation de symboles, comme dans la norme ISO/TC 145/ISO 3864-2);
    - des guides sur les bonnes pratiques;
    - des activités de formation et de collaboration pour le consommateur agissant à titre de fournisseur (p. ex. pour la publication de jeux de données);
  - Réduire l’incertitude en ce qui concerne certains sujets touchant la loi en investissant dans des études afin de :
    - comprendre les nouvelles tendances et les nouveaux droits en ce qui touche la vie privée, la possession de données, le droit d’auteur;
    - clarifier davantage la responsabilité des fournisseurs de données, des responsables de l’intégration et des distributeurs de données;
    - clarifier les responsabilités lorsque les services et données géospatiales traversent les frontières.



---

## 6. Conclusion

Tandis que les données géospatiales sont de plus en plus produites, utilisées et réutilisées par de nouveaux types d'intervenants, la question de la qualité des données géospatiales devient de plus en plus importante. L'objectif du présent guide est d'appuyer la communauté géospatiale canadienne dans ses efforts visant à sensibiliser les membres de la société utilisant les données spatiales au sujet de la qualité des données et des risques liés à l'utilisation inappropriée de celles-ci. En se fondant sur des normes internationales comme la norme ISO 19157 (qualité des données géospatiales) et la norme ISO 31000 (gestion du risque), ce guide présente les concepts sous-jacents à la qualité des données géospatiales, la gestion de la qualité des données géospatiales, le processus détaillé d'évaluation de la qualité des données géospatiales (fondé sur la norme ISO 19157) et la gestion des risques liés à l'utilisation inappropriée des données géospatiales (fondée sur la norme ISO 31000). Des exemples détaillés de tâches de gestion de la qualité, d'évaluation et de gestion du risque à entreprendre dans les contextes E-E, E-C et C-C sont présentés.

---

## 7. Bibliographie

American National Standards Institute. (2006). *ANSI Z535.6 American National Standard for Product Safety Information in Product Manuals, Instructions, and Other Collateral Materials*. ANSI.

Bédard, Y. (18 mai 2012). « Geospatial Data Quality Awareness, The Next Challenge: Are We Ready? » Principal conférencier invité, GEOIDE IV-23 International Workshop on Geospatial Data Quality, Legal, Ethical and Technical Aspects. Québec (Québec), Canada.

Bédard, Y. (1988). « Uncertainties in Land Information System Databases », *Proceedings of the Eighth International Symposium on Computer Cartography*. Baltimore, Maryland.

Bédard, Y., J. Chandler, R. Devillers et M. Gervais. (22-27 mars 2009). « System Design Methods and Geospatial Data Quality », Association of American Geographers Professional Ethics - Session on Geographic Information Ethics and GIScience. Las Vegas, États-Unis.

Canada, Conseil du Trésor. (2012). *Guide de gestion intégrée du risque*. Secrétariat du Conseil du Trésor du Canada.

Chandler, J. A. (2010). « User generated legal issues: Potential legal issues raised by user-generated content for geographical information services ». *Proceedings GSDI 12 World Conference*. Singapour.

Chandler, J. A., et K. Levitt (2011). « Spatial Data Quality: The Duty to Warn Users of Risks Associated with Using Spatial Data ». *Alberta Law Review*, 49 (1), p. 79-106.

Chrisman, N. (1983). « The role of quality in the long-term functioning of a Geographic Information System », *Proceedings of AutoCarto 6*, (p. 303-321). Hull, Canada.

Conseil du Trésor du Canada. (2009). Norme sur les données géospatiales.

Devillers, R. (2004). *Conception d'un système multidimensionnel d'information sur la qualité des données géospatiales*. Thèse de doctorat présentée en cotutelle, Faculté de foresterie et géomatique, Université Laval.

Devillers, R., A. Stein, Y. Bédard, N. Chrisman, P. Fisher et W. Shi (2010). « 30 years of research on Spatial Data Quality : Achievements, failures and opportunities ». *Transactions in GIS*, 14 (4), p. 387-400.

Devillers, R., Y. Bédard et M. Gervais (9-13 avril 2013). « We need to adopt more ethical practices for geographic information data and products dissemination », Assemblée annuelle 2013 de l'AAG. Los Angeles, Californie, États-Unis.

Devillers, R., Y. Bédard, R. Jeansoulin et B. Moulin (2007). « Towards Spatial Data Quality Information Analysis Tools for Experts Assessing the Fitness for Use of Spatial Data ». *International Journal of Geographical Information Sciences (IJGIS)*, 21 (3), p. 261-282.

---

---

Federal Geographic Data Committee. (1998). *FGDC-STD-001-1998 Content Standard for Digital Geospatial Metadata*. FGDC.

GéoConnexions. (2013). Qualité des données de l'ICDG, affiche.

GEOIDE (PIV-23). (2012). Tiré de « Public protection and ethical dissemination of geospatial data - Social and legal aspects »: <http://dataquality.scg.ulaval.ca/>

Gervais, M. (2004). *La pertinence d'un manuel d'instructions au sein d'une stratégie de gestion du risque découlant de la fourniture de données géographiques numériques*. Thèse de doctorat, Département des sciences géomatiques, Université Laval.

Gervais, M., Y. Bédard, et S. Larrivée (2007). *Spatial Data Quality Audit*, 19 juin, Québec, Canada. Troisième conférence nationale des arpenteur-géomètres.

Gervais, M., Y. Bédard, M.-A. Lévesque, E. Bernier et R. Devillers, (2009). « Data Quality Issues and Geographic Knowledge Discovery ». Dans *Geographic Data Mining and Knowledge Discovery, Second Edition*. CRC Press.

Gervais, M., Y. Bédard, R. Jeansoulin et B. Cervelle, (2007). « Qualité des données géographiques. Obligations juridiques potentielles et modèle du producteur raisonnable ». *Revue internationale de géomatique*, 17 (1), p. 33-62.

Gervais, M., Y. Bédard, S. Larrivée, S. Rivest et T. Roy (3-4 octobre 2013). « Enquête canadienne sur la qualité des données géospatiales et la gestion du risque », Géomatique 2013, 3-4 octobre, Montréal, Canada.

Grira, J. (2014). *Improving knowledge about the risks of inappropriate uses of geospatial data by introducing a collaborative approach in the design of geospatial databases*. Thèse de doctorat, Département des sciences géomatiques, Université Laval.

Grira, J., Y. Bédard et S. Roche (2009). « Spatial Data Uncertainty in the VGI World: Going from Consumer to Producer ». *Geomatica*, 64 (1), p. 61-71.

Hickling Arthurs Low Corporation. (2011). *Final Report: CGDI Operational Policies Needs*. GéoConnexions.

Jones, K. (2011). *Communicating perceived geospatial data quality of 3D objects in virtual globes*. Thèse de maîtrise en sciences, Département de géographie, Université Memorial de Terre-Neuve.

Jones, K., R. Devillers, Y. Bédard et O. Schroth (2013). « Visualizing perceived spatial data quality of 3D objects within virtual globes ». *International Journal of Digital Earth*.

Juran, J. M., F. M. Gryna et R. S. Bingham (1974). *Quality Control Handbook*. New York: McGraw-Hill.

Kerzner, H. (2009). *Project Management, A Systems Approach to Planning, Scheduling and Controlling, 10th Edition*. New Jersey : John Wiley & Sons.

Koistinen, K. (2015). « Quality information representation in ELF Geo Product Finder ». *International Workshop on Spatial Data and Map Quality*. Malte.

---

Longley, P. A., M. F. Goodchild, D. J. Maguire et D. W. Rhind (2001). *Geographic Information Systems and Science*. Angleterre : John Wiley & Sons.

Maitra, J. B. (2004). « Applying ISO/TC 211 Standards in the Development of Data Content Standards ». *GSDI-7*.

National Committee for Digital Cartographic Data Standards. (1987). *A Draft Proposed Standard for Digital Cartographic Data*.

Normand, P. (1999). *Modélisation des contraintes d'intégrité spatiale, théorie et exemples d'applications*. Essai de M.Sc., Département des sciences géomatiques, Université Laval.

Open Geospatial Consortium. (2014). *Data Quality DWG*. Tiré de Open Geospatial Consortium : <http://www.opengeospatial.org/projects/groups/dqdwg>

Organisation internationale de normalisation. (2004). *ISO 3864-2:2004 Symboles graphiques -- Couleurs de sécurité et signaux de sécurité -- Partie 2 : Principes de conception pour l'étiquetage de sécurité des produits*. ISO.

Organisation internationale de normalisation. (2005). *ISO 19109:2005 Information géographique -- Règles de schéma d'application*. ISO.

Organisation internationale de normalisation. (2005). *ISO 19110:2005 - Information géographique -- Méthodologie de catalogage des entités*. ISO.

Organisation internationale de normalisation. (2005). *ISO 19123:2005 Information géographique -- Schéma de la géométrie et des fonctions de couverture*. ISO.

Organisation internationale de normalisation. (2005). *ISO 9000:2005 Systèmes de management de la qualité -- Principes essentiels et vocabulaire*. ISO.

Organisation internationale de normalisation. (2007). *ISO 19131:2007 Information géographique -- Spécifications de contenu informationnel*. ISO.

Organisation internationale de normalisation. (2007). *ISO 19139:2007 Information géographique -- Métadonnées -- Implémentation de schémas XML*. ISO.

Organisation internationale de normalisation. (2009). *ISO 19115-2:2009 Information géographique -- Métadonnées -- Partie 2 : Extensions pour les images et les matrices*. ISO.

Organisation internationale de normalisation. (2009). *ISO 31000:2009 Management du risque -- Principes et lignes directrices*. ISO.

Organisation internationale de normalisation. (2011). *ISO 19131:2007/Amd.1:2011 Information géographique -- Spécifications de contenu informationnel Amendement 1 Exigences relatives à l'inclusion d'un schéma d'application et d'un catalogue d'objets géographiques et au traitement des couvertures dans un schéma d'application*. ISO.

---

Organisation internationale de normalisation. (2012). *ISO/TS 19158:2012 Information géographique -- Assurance qualité relative à l'approvisionnement de données*. ISO.

Organisation internationale de normalisation. (2013). *ISO 19157:2013 Information géographique -- Qualité des données*. ISO.

Organisation internationale de normalisation. (2014). *ISO 19115-1:2014 Information géographique -- Métadonnées -- Partie 1 : Principes de base*. ISO.

Plante, K., et M. Gervais (mars 2015). « La garantie de qualité des données géospatiales ». *Geomatica*.

Ressources naturelles Canada. (2014). *CanVec+ Catalogue d'entités Édition 1.0.2*. [http://ftp2.cits.rncan.gc.ca/pub/canvec+/doc/CanVec+\\_catalogue\\_entites.pdf](http://ftp2.cits.rncan.gc.ca/pub/canvec+/doc/CanVec+_catalogue_entites.pdf)

Ressources naturelles Canada. (2014). *CanVec+ Spécifications de produit Édition 1.0*. [http://ftp2.cits.rncan.gc.ca/pub/canvec+/doc/CanVec+\\_specifications\\_produit.pdf](http://ftp2.cits.rncan.gc.ca/pub/canvec+/doc/CanVec+_specifications_produit.pdf)

Ressources naturelles Canada. (2015). *canvec\_082C\_pna*.

Sboui, T., et Y. Bédard (2012). « Universal Geospatial Ontology for the Semantic Interoperability of Data: What are the Risks and How to Approach them? ». Dans T. Podobnikar et M. Ceh (éd.), *Universal Ontology of Geographic Space: Semantic Enrichment for Spatial Data* (p. 1-27).

Statistique Canada. (2012). *Guide de l'utilisateur, Enquête sur les dépenses des ménages, 2012*.

Trakas, A. (2008). *Benefits of Open Standards - how to engage with OGC*.

Uhlir, P. F. (2013). « The Legal Interoperability of Data ». *NSGIC Midyear Meeting Archive*.

Veregin, H. (1999). « Data Quality Parameters ». Dans *Geographic Information Systems: Principles and Technical Issues* (p. 177-189). John Wiley & Sons, Inc.