



Ressources naturelles
Canada

Natural Resources
Canada



L'INFRASTRUCTURE CANADIENNE DE DONNÉES GÉOSPATIALES PRODUIT D'INFORMATION 36f

Guide d'introduction à la préservation des données géospatiales

GéoConnexions
Hickling Arthurs Low

2013

©Sa Majesté la Reine du chef du Canada, représentée par le ministre de Ressources naturelles Canada, 2013

Canada 

REMERCIEMENTS

GéoConnexions tient à reconnaître les contributions de Hickling Arthurs Low Corporation. – Tracey P. Lauriault et Ed Kennedy (recherche, rédaction et révision), Yvette Hackett (spécialiste en la matière) et Marcel Fortin (réviseur externe). Le personnel de GéoConnexions a aussi travaillé à la gestion, aux commentaires et à l’orientation pour ce projet : Cindy Mitchell, Simon Riopel, ainsi que Jenna Findlay et Nadia Eckardt qui ont apporté leur soutien lors de la création de ce document.

Table des matières

1. Préambule.....	1
2. Introduction	3
2.1 Archivage et préservation des données.....	3
2.1.1 Terminologie de base	4
2.1.2 Dépôts numériques fiables	5
2.1.3 L'importance des métadonnées.....	7
2.1.4 Les effets du changement technologique	7
2.2 Lignes directrices à l'intention des créateurs de données géospatiales	8
2.3 Lignes directrices à l'intention des préservateurs de données géospatiales	10
3. Lois et politiques influant sur l'archivage et la préservation	14
3.1 Responsabilités liées à l'archivage et à la préservation.....	14
3.2 Lois et règlements fédéraux.....	16
3.3 Politiques et directives fédérales	18
4. Cadres d'archivage et de préservation	19
4.1 Introduction	19
4.2 Modèle de référence du Système ouvert d'archivage d'information.....	20
4.3 European Long Term Data Preservation Common Guidelines	22
4.4 Trustworthy Repositories Audit & Certification: Criteria and Checklist	25
5. Exemples de préservation de données géospatiales	28
5.1 Étude de cas : Système de gestion des données d'observation de la Terre (SGDOT).....	28
5.1.1 Introduction.....	28
5.1.2 Modèle opérationnel en cours d'utilisation – Mise en œuvre à ce jour	29
5.1.3 Défis éprouvés.....	35
5.1.4 Leçons retenues.....	37
5.2 Profil : Système d'archivage de l'information géographique de l'Ontario (SAIG).....	37
5.2.1 Introduction.....	37
5.2.2 Modèle opérationnel en cours d'utilisation.....	38
5.2.3 Pratiques exemplaires.....	39

5.3	Profil : Gestion des données scientifiques intégrées (GDSI).....	41
5.3.1	Introduction.....	41
5.3.2	Modèle opérationnel en cours d'utilisation.....	41
5.3.3	Pratiques exemplaires.....	42
5.4	Profil : Préservation des données de l'Année polaire internationale (API)	43
5.4.1	Introduction.....	43
5.4.2	Modèle opérationnel en cours d'utilisation.....	43
5.4.3	Défis éprouvés	45
5.4.4	Leçons retenues.....	46
6.	Création d'un système de préservation des données géospatiales	48
6.1	Introduction	48
6.2	Déterminer la portée et les objectifs du système	49
6.3	Définir la communauté des utilisateurs du système.....	52
6.4	Obtenir et gérer les ressources.....	54
6.5	Planifier la préservation.....	55
6.6	Élaborer des politiques et des procédures.....	58
6.7	Évaluer la valeur de préservation des documents d'archives	60
6.8	Acquérir et saisir les documents d'archives	61
6.9	Préserver les documents d'archives.....	62
6.10	Décrire les métadonnées d'archives	65
6.11	Gérer et maintenir les documents d'archives.....	66
6.12	Fournir un accès aux documents d'archives.....	67
7.	Défis et solutions	70
8.	Conclusions.....	74
A.	Ouvrages de référence.....	76
B.	Glossaire des termes	84
C.	Lois et règlements propres à la préservation et à la gestion de l'information du gouvernement	90
D.	Politiques et directives du Secrétariat du Conseil du Trésor (SCT) sur la gestion de l'information.....	92
E.	Faits saillants du Geoarchiving Business Planning Guidebook	95

1. Préambule

Le présent guide d'introduction fait partie d'une série de documents sur les politiques opérationnelles que GéoConnexions prépare actuellement. Il est destiné à offrir aux intervenants de l'Infrastructure canadienne de données géospatiales (ICDG) des renseignements sur la nature et la portée de l'archivage et de la préservation des données géospatiales numériques et les réalités, les défis et les pratiques exemplaires des politiques opérationnelles connexes.

L'essor des applications géospatiales en ligne et l'avalanche de données, combinées à la complexité croissante de l'archivage et de la préservation des données numériques, ont révélé une lacune importante dans la couverture de la politique opérationnelle de l'infrastructure canadienne de données géospatiales (ICDG).

Actuellement, il n'existe pas de directives communément acceptées pour les intervenants de l'ICDG qui sont désireux ou chargés de préserver leurs actifs de données géospatiales pour un accès et une utilisation à long terme. Plus précisément, il y a peu ou pas de directives disponibles pour éclairer les décisions en matière de politique opérationnelle sur la façon de gérer, préserver et fournir l'accès à une collection de données géospatiales numériques. La préservation des données géospatiales sur une période de temps est particulièrement importante lorsque les ensembles de données doivent informer des applications de modélisation telles que les prévisions sur les effets des changements climatiques, les prévisions de crues et la gestion de l'utilisation des terres. De plus, les gestionnaires de données peuvent avoir à la fois une responsabilité légale et morale de mettre en œuvre des programmes d'archivage et de préservation efficaces.

Selon la recherche et l'analyse du cadre législatif canadien et des pratiques internationales en vigueur en matière d'archivage et de préservation des données numériques, ce guide d'introduction fournit des indications sur les facteurs à considérer et les mesures à prendre dans la planification et la mise en œuvre d'un programme d'archivage et de préservation des données. Il décrit une approche pour établir un centre d'archives de données géospatiales axé sur les pratiques exemplaires de la littérature et des études de cas canadiennes.

Le programme *GéoConnexions* est une initiative nationale sous l'égide de Ressources naturelles Canada. GéoConnexions soutient l'intégration et l'utilisation de l'Infrastructure canadienne de données géospatiales (ICDG).

L'ICDG est une ressource en ligne qui permet d'améliorer l'échange, l'accessibilité et l'utilisation de l'information géospatiale canadienne – l'information sur des lieux géographiques du Canada. Elle aide les décideurs de tous les paliers de gouvernement, du secteur privé, des organisations non gouvernementales et du monde universitaire à prendre de meilleures décisions au sujet des priorités sociales, économiques et environnementales.

Ce guide d'introduction fournira aux intervenants de l'ICDG des renseignements sur la façon d'intégrer les facteurs en matière d'archivage et de préservation dans un processus efficace de gestion de données qui couvre l'ensemble du cycle de vie (DCC, 2013) (BAC, 2006) de leurs actifs de données géospatiales (c.-à-d., la création et la réception, la distribution, l'utilisation, le maintien et la disposition). Il vise à informer les intervenants de l'ICDG sur l'importance de la préservation des données à long terme, et leur fournir les renseignements et les outils nécessaires pour prendre des décisions stratégiques pour la création d'un centre d'archives et la préservation des données géospatiales numériques.

Ce guide d'introduction aborde également des sujets d'ordre juridique, courants au moment de la publication, à des fins d'information générale seulement. Il s'appuie sur l'*Archivage, gestion et préservation des données géospatiales – Rapport sommaire et recommandations* de GéoConnexions et HAL (2011). Les éléments d'information qu'il contient ne s'appliquent peut-être pas à toutes les administrations. GéoConnexions n'est pas responsable de l'utilisation de tout renseignement ou contenu du présent guide d'introduction. Le contenu du présent guide d'introduction ne constitue pas un avis juridique et ne doit pas être considéré comme tel.

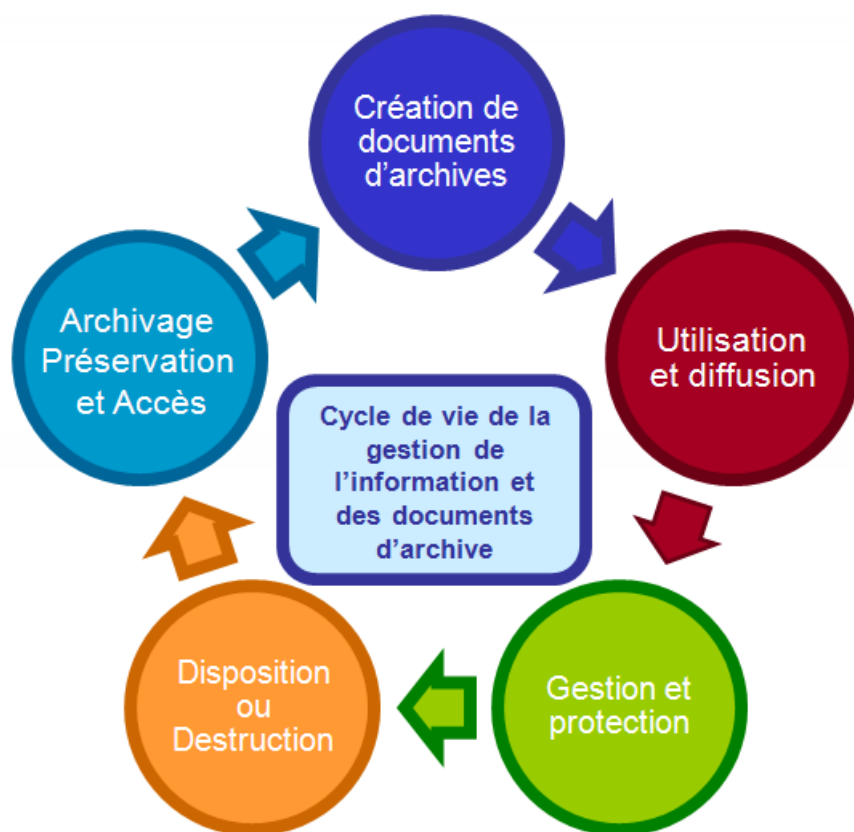
2. Introduction

Le but du présent chapitre est d'introduire le lecteur aux concepts, aux processus et à la terminologie de base concernant la préservation et l'archivage des données numériques. Il est important de reconnaître que certains termes peuvent ne pas avoir la même signification dans les différentes communautés de pratique, et certaines de ces différences sont référencées. De plus, des indications particulières pour les créateurs et les préservateurs de données géospatiales sont brièvement abordées.

2.1 Archivage et préservation des données

La préservation des données (ou documents d'archives, terme le plus souvent utilisé par la communauté archivistique) fait partie intégrante du cycle de vie de la gestion de l'information, comme l'illustre la figure 1. Au début de ce cycle, il importe que les créateurs de données songent à la possibilité de les préserver en vue d'un accès et d'une utilisation à long terme.

Figure 1 : Cycle de vie de la gestion des documents d'archives et de l'information



Source : <http://slcoarchives.wordpress.com/2012/04/13/managing-records-now-for-the-future/>

La préservation des données peut se faire à court, à moyen ou à long terme et est déterminée au cours du processus d'évaluation des données de l'organisation, qui peut être documenté dans des énoncés de mandat, des politiques et des ententes d'acquisition, tel que déterminé par les créateurs, les préservateurs et les utilisateurs. Bien qu'il puisse encore y avoir une exigence de préserver des copies papier des données sous forme de cartes ou de cartes spécialisées, de telles activités de préservation commencent généralement longtemps après la création et incluent rarement la participation des créateurs de données. Avec les ressources numériques, cependant, il est nécessaire de gérer activement la ressource à chaque étape de son cycle de vie, de reconnaître les interdépendances entre chaque étape et de commencer les activités de préservation le plus tôt possible (RLG et OCLC, 2002) (Lauriault, Craig, Pulsifer et Taylor, 2008). Étant donné qu'une grande partie de l'information à l'appui nécessaire à la préservation de l'information archivée est plus facilement disponible ou uniquement disponible au moment où l'information originale est produite, les résultats sont meilleurs lorsque ces organisations participent à l'effort de préservation. Ce guide d'introduction met l'accent sur les facteurs que les organisations doivent prendre en compte dans la conception et l'élaboration des politiques et des processus de préservation et d'archivage des données numériques.

2.1.1 Terminologie de base

Au départ, il est important de reconnaître que les termes d'archivage et de préservation sont souvent utilisés sans distinction, par exemple : **archives**, **préservation**, **système de sauvegarde et stockage**. Par souci de cohérence, ce guide d'introduction fera référence à la base de données terminologique du projet International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2 (InterPARES 2, 2013) et à la terminologie utilisée dans les cadres de préservation examinés (se reporter au chapitre 4). Un glossaire des termes clés est fourni à l'annexe B.

Tout d'abord, « l'organisme ou l'établissement chargé de préserver et de communiquer des documents d'archives [p. ex., données et métadonnées] préservés en permanence » est connu comme un **centre d'archives** (InterPARES 2, 2013). D'autre part, la **préservation** est l'ensemble des activités gérées nécessaires pour assurer un accès continu aux documents d'archives numériques pour aussi longtemps que nécessaire, au-delà des limites de la défaillance d'un support ou d'un changement technologique (Digital Preservation Coalition, 2008). Plus précisément, la préservation est « l'ensemble des principes, des politiques, des règles et des stratégies visant à prolonger l'existence d'un objet [ensemble de données, base de données, logiciel] en le maintenant dans un état convenable pour l'utilisation, que ce soit dans son format d'origine ou dans un format plus durable, tout en laissant intact la forme intellectuelle de l'objet » (InterPARES 2, 2013). Un **système de sauvegarde** est la technologie utilisée pour faire une copie d'un fichier de données en vue de la remise en état d'un système, tandis que le **stockage** est le placement des données dans un système de stockage sur support numérique (p. ex., ruban de stockage). Les administrateurs des technologies de l'information (TI) considèrent parfois leurs systèmes de stockage et de sauvegarde comme des archives, même s'ils ne sont pas permanents et que les bandes sont souvent écrasées. Un système d'archives de données géospatiales

comprendra un système de sauvegarde et un stockage des données à l'intérieur de son infrastructure opérationnelle, et ceux-ci feront partie intégrante de son système de préservation des documents d'archives.

La communauté archivistique se réfère à la création de **documents d'archives**, qui sont des « documents [données et métadonnées] produits ou reçus, dans le cadre d'une activité pratique, à titre d'instruments ou de produits dérivés d'une telle activité, et préservés pour un usage futur ou comme source de référence » (InterPARES 2, 2013). Dans certains cas, les organisations créent et mettent de côté les documents d'archives qui demeurent une partie active de leurs processus opérationnel en cours. Par exemple, les données brutes d'Observation de la Terre (OT) sont recueillies à partir de stations de réception satellite et saisies dans un système de préservation, ce qui constitue un élément normal du processus du cycle de vie du document. Ces données restent actives et sont préservées pour les raisons suivantes : la loi oblige souvent les créateurs de données d'OT (p. ex., *Loi sur les systèmes de télédétection spatiale* [L.C. 2005, ch. 45]) à le faire, une analyse de rentabilisation justifie de les préserver et on s'attend à ce que ces données répondent aux demandes futures.

Les principaux acteurs dans le processus d'archivage des données, selon le *modèle de référence du Système ouvert d'archivage d'information (SOAI)* (CCSDS, 2012) qui sera décrit au chapitre 4, sont les producteurs, les consommateurs et la direction. Les producteurs (c.-à-d., les créateurs de documents) peuvent être les créateurs individuels de l'ensemble de données géospatiales ou la personne morale responsable de sa création. Les producteurs peuvent avoir leur propre système de préservation interne ou être des organisations externes fournissant des données aux archives par mandat ou de façon volontaire. La direction supervise le processus, mais ne participe pas aux opérations quotidiennes des archives, qui sont normalement effectuées par une entité fonctionnelle administrative. Les consommateurs sont les utilisateurs des archives, des intervenants ou une communauté désignée. Une communauté désignée peut être le public canadien, ou un groupe spécifique et distinct comme des professionnels de la géomatique, des spécialistes des sciences de la Terre, ou des océanographes, chacun ayant des besoins qui lui sont propres, nécessitant des fonctionnalités et support différents. Ce guide d'introduction mettra l'accent sur les créateurs de données géomatiques et les professionnels en général, mais présentera, dans les études de cas et les profils ci-dessous, certaines communautés désignées distinctes.

2.1.2 Dépôts numériques fiables

Une archives de données géospatiales numériques devra démontrer qu'elle dispose d'un système de préservation fiable devant être considéré comme digne de confiance par ses communautés désignées (InterPARES 2, 2002) (MacNeil, 2000). Les archives sont considérées comme digne de confiance lorsqu'elles peuvent démontrer que les données numériques préservées dans sa collection se révéleront exactes, fiables et authentiques et que l'unité d'affaires effectuant la préservation peut démontrer qu'elle « n'a aucune raison de modifier les documents d'archives conservés ou de permettre à d'autres de les modifier et est en mesure de mettre en œuvre toutes

les exigences pour la préservation de copies authentiques » (InterPARES 2, 2013). Bien que ces trois concepts soient connus des scientifiques et des praticiens de la géomatique, ils sont compris différemment par les archivistes (Roeder, Eppard, Underwood et Lauriault, 2008). Il est important de reconnaître que, pour les archivistes, ces termes font référence aux attributs du document d'archives et non aux aspects scientifiques et méthodologiques qui affectent la qualité des données sur lesquelles repose le document d'archives. Un **dépôt numérique fiable** (DNF) est une **archive numérique** « dont la mission est de fournir à sa communauté désignée un accès fiable et à long terme aux ressources numériques gérées, maintenant et dans le futur » (RLG et OCLC, 2002, p. 5).

Un groupe multidisciplinaire et multisectoriel de préservateurs et d'experts en accréditation et en vérification d'archives internationaux a élaboré un processus de certification. Le document *Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist* [Critères et liste de contrôle de la vérification et de l'accréditation des dépôts fiables (VADF)] (OCLC et CRL 2007) a été créé pour aider la communauté de la préservation, en particulier les responsables de la mise en œuvre des archives numériques. La *Liste de contrôle VADF* est un outil indispensable pour les gestionnaires et les administrateurs afin d'évaluer leurs pratiques actuelles, de déterminer les lacunes et d'élaborer des solutions. En 2012, la *Liste de contrôle VADF* est devenue la pratique recommandée de la norme ISO¹ 6363:2012 *Systèmes de transfert des informations et des données spatiales – Audit et certification des référentiels numériques de confiance* (ISO 2012). Bien que la *Liste de contrôle VADF* soit conçue dans le but de procéder à la vérification des systèmes de préservation et qu'elle soit longue et généralisée à l'ensemble des archives numériques, elle est

PRATIQUE EXEMPLAIRE

Le rapport *Trusted Digital Repositories : Attributes and Responsibilities* (2002), produit par le Research Library Group (RLG) et le Online Computer Library Centre (OCLC) en collaboration avec les archivistes, accorde les attributs qui suivent au DNF :

- assume la responsabilité du
- maintien à long terme des ressources numériques pour le compte de ses déposants (p. ex., créateurs ou producteurs) et au bénéfice des utilisateurs actuels et futurs, des consommateurs ou des collectivités désignées;
- dispose d'un système organisationnel qui prend en charge non seulement la viabilité à long terme du système de préservation, mais aussi l'information numérique dont il a la responsabilité;
- démontre la responsabilité financière et la durabilité;
- conçoit ses systèmes en conformité avec les conventions et les normes communément acceptées pour assurer sur une base permanente la gestion, l'accès et la sécurité des documents qui y sont déposés;
- établit des méthodes pour l'évaluation du système qui répondent aux attentes de fiabilité de la communauté;
- dépend de l'accomplissement de ses responsabilités à long terme auprès des déposants et des utilisateurs ouvertement et explicitement;
- a des politiques, des pratiques et un rendement qui peuvent être vérifiés et mesurés.

¹ Organisation internationale de normalisation

un outil indispensable pour les organisations produisant des données géospatiales qui souhaitent élaborer un centre d'archives de données géospatiales ou qui souhaitent autoévaluer leurs systèmes. Elle sera réexaminée plus loin, à la section 4.3.

2.1.3 L'importance des métadonnées

Dans l'*Archivage, gestion et préservation des données géospatiales – Rapport sommaire et recommandations* (GéoConnexions et HAL, 2011), les portails de données géospatiales ont été reconnus comme étant plus que des outils d'exploration. Les portails sont également des points d'accès aux collections de données géospatiales qui ont été évaluées comme ayant une valeur opérationnelle ou scientifique par l'institution qui héberge le portail, les contributeurs de données et leurs communautés d'utilisateurs. De plus, ces données sont détectables au moyen de vastes métadonnées géospatiales, et dans certains cas, les portails de données géospatiales ont des politiques en place pour gérer ces données à long terme. Ces portails qui adhèrent aux spécifications ouvertes, aux normes d'interopérabilité et à la norme *ISO 19115 Information géographique - Métadonnées* et adoptent des logiciels libres ont davantage de chances de résister à l'épreuve du temps par rapport à ceux qui utilisent des systèmes exclusifs (J. Roeder, Eppard, Underwood et Lauriault, 2008, p. 44, 45). Enfin, comme prévu, il a été découvert que les données géospatiales sont **complexes, dynamiques et interactives** et qu'elles existent dans de multiples formats, sont conservées dans des systèmes spécialisés et sont accessibles et diffusées dans les systèmes distribués selon les disciplines pratiques spécifiques. Toutefois, puisque les données géospatiales sont recueillies selon des modèles scientifiques bien définis et normalisés, et des méthodes réglementaires, elles doivent être bien décrites par les créateurs de métadonnées.

2.1.4 Les effets du changement technologique

Le **changement technologique** est un facteur particulièrement important dans la préservation des données, en particulier si les données doivent être préservées et demeurer accessibles à long terme. Peu importe la façon dont les archives conservent leurs ressources documentaires actuelles, il faudra un jour faire migrer une grande partie de son contenu vers différents supports (qui peut ou non comprendre le changement des séquences de bits), ou vers une configuration matérielle ou un environnement logiciel différents pour les garder accessibles. Surmonter l'obsolescence technologique du matériel et des logiciels peut être réalisé par différents moyens, notamment les techniques d'émulation et de migration. Le reformatage et l'actualisation de l'information peuvent également être nécessaires pour déplacer des données entre des supports de stockage de données. L'article de Denise Bleakly intitulé *Long-Term Spatial Data Preservation and Archiving: What are the Issues?* donne un aperçu de ces questions technologiques (Bleakly, 2002). Les cadres examinés au chapitre 4 abordent le changement technologique pour les créateurs et les préservateurs de données géospatiales, et donnent des directives sur la façon de planifier en conséquence.

De plus, les formats de données géospatiales sont complexes, changent fréquemment et deviennent obsolètes. Par conséquent, les registres de formats ont été créés pour permettre la

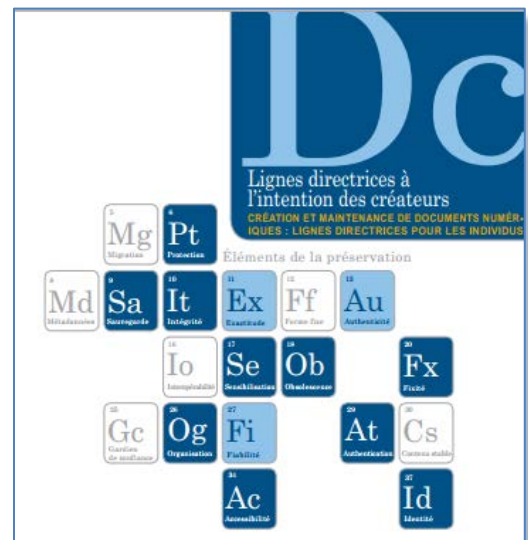
préservation à long terme et l'accès aux données. Les registres renforcent les efforts de migration de format, notamment ceux qui visent à ce que les environnements d'émulation demeurent portables (Erwin & Sweetking-Singer, 2009). La Library of Congress [Bibliothèque du Congrès des É.-U.] possède le registre *Sustainability of Digital Formats*, qui comprend des données géospatiales (Bibliothèque du Congrès, 2013). Ce registre est un outil utile pour les créateurs et les préservateurs de données géospatiales, et les créateurs devraient enregistrer leurs formats de fichiers dans ce registre, alors que les descriptions de formats de fichiers doivent faire partie de la description archivistique des métadonnées (Hoebelheinrich & Munn, *Assessing the Utility of Current Format Registry Efforts for Geospatial Formats*, 2009).

Même si les données géospatiales sont uniques, il reste quelques règles générales qui s'appliquent à tous les créateurs et préservateurs de documents d'archives (InterPARES 2, 2007), qui sont abordées dans les sections qui suivent.

2.2 Lignes directrices à l'intention des créateurs de données géospatiales

Les *Lignes directrices à l'intention des créateurs* du projet InterPARES 2 (InterPARES 2, 2007) fournissent des recommandations qui seront familières aux professionnels de la géomatique. Les logiciels choisis doivent être rétrocompatibles (c.-à-d. qu'ils peuvent fonctionner avec des entrées générées par un produit plus ancien) et interopérables à travers le temps et l'espace, largement adoptées et conformes aux normes. Cela est particulièrement vrai si les créateurs souhaitent que leurs données soient lues dans le futur et veulent que l'information créée à partir de leurs données soit comprise et visualisée comme ils l'entendaient. Pour éclairer les stratégies de migration, tous les logiciels spécialisés doivent être entièrement documentés notamment les

personnalisations de logiciels. Les notes dans le code du logiciel sont particulièrement utiles à cet égard. La création du système géospatial, notamment la structure et les fonctions, le matériel et les logiciels, le système d'exploitation et comment ces derniers fonctionnent les uns avec les autres, doit être documentée dans les spécifications de base, car ces dernières informeront les mises à jour. Enfin, des formats non exclusifs, indépendants des plateformes, non compressés et largement acceptés, avec accès aux spécifications et aux versions et encodages documentés, doivent être utilisés. Les *Lignes directrices à l'intention des créateurs* recommandent également le regroupement logique des documents d'archives et la détermination des stratégies de conservation pour ces groupes au moment de la création, car c'est plus efficace.



Comme les données géospatiales sont souvent rendues dans des présentations interactives, accessibles par le biais des services Web ou des portails de données et peuvent être générées dynamiquement avec des capteurs en temps quasi réel, il est important de comprendre les concepts de forme fixe et de variabilité restreinte (InterPARES 2, 2007). Afin de veiller à ce que l'apparence d'un document soit la même à chaque fois qu'il est récupéré, le contenu du document (p. ex., les données pour créer la carte et l'algorithme utilisé pour rendre celle-ci sur l'écran) doit être fixe. La forme documentaire du document d'archives (p. ex., spécifications et logiciel utilisé pour créer et visualiser la carte interactive) doit être inaltérable, de sorte que sa présentation reste toujours la même. Des règles fixes doivent être instaurées guidant le choix et la présentation du contenu, permettant ainsi une variabilité stable de la carte interactive ou du modèle. Cela garantit qu'un document fiable et authentique de la carte interactive est accessible de la façon dont les créateurs souhaitaient. L'acceptation de la norme *ISO 19128:2005 Information géographique – Interface de carte du serveur Web* (ISO 2005) par le Secrétariat du Conseil du Trésor du Canada (SCT) pour l'ensemble de l'information géographique du gouvernement du Canada (GC) (SCT, 2012) favorise une approche commune des normes ouvertes pour la production des cartes qui permettront de simplifier le travail des préservateurs.

PRATIQUE EXEMPLAIRE

Une analyse plus détaillée des éléments de métadonnées pour la préservation à long terme des ensembles de données géospatiales est disponible dans les publications de *Geospatial Multistate Archive and Preservation Partnership (GeoMAPP)*, créées sous l'égide du National Digital Information Infrastructure and Preservation Program (NDIIPP) de la Bibliothèque du Congrès des É.-U. (GeoMAPP, 2011).

Les professionnels de la géomatique sont bien versés dans les **métadonnées** ou la **documentation de création** tel que mentionné dans le *modèle de référence du SOAI*. Pour les archivistes, les métadonnées aident à maintenir l'**identité** du document qui est la qualité à partir de laquelle un document peut être identifié et distingué d'un autre tandis que l'**intégrité** est la qualité d'un dossier indiquant qu'il est inchangé et complet (InterPARES 2, 2013). Les éléments de métadonnées sur l'**identité des archives** sont très similaires à la plupart des éléments de métadonnées géomatiques nécessaires, mais certains éléments supplémentaires peuvent être nécessaires, tels que les descriptions des formats de fichiers, les liens des composantes et l'environnement de préservation, pour n'en nommer que quelques-uns.

Les *Lignes directrices à l'intention des créateurs* font également référence à l'**authentification**, qui est une déclaration à un moment donné de l'authenticité d'un document. Cela se fait par l'insertion d'un élément ou l'ajout d'une déclaration au document par une personne autorisée qui a le droit de le faire. L'authentification sans recours à la technologie est une méthode, et la plupart des recommandations formulées dans les *Lignes directrices* aident à déterminer si un ensemble de données ou un groupe de données peuvent être présumés authentiques, à partir de faits connus sur la façon dont ils ont été créés. De plus, les politiques et pratiques administratives qui sont technologiquement indépendantes ou neutres, comme la protection des données contre toute manipulation en contrôlant l'accès à l'endroit où les données sont stockées en restreignant l'accès physique, sont recommandées. L'authentification dépendante de la technologie peut également être réalisée avec les technologies de transmission (p. ex., la cryptographie et les

signatures numériques), mais celles-ci sont assujetties à l'obsolescence technologique. Restreindre l'accès aux données avec des mots de passe et d'autres mesures de protection et de sécurité et élaborer des protocoles d'autorisation d'accès constituent d'autres approches. Peu importe les méthodes employées, il est important d'être en mesure de démontrer que les documents d'archives ne peuvent pas être altérés.

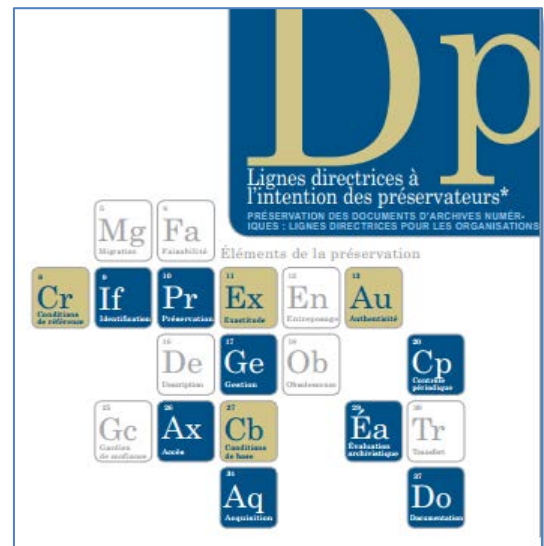
Enfin, les *Lignes directrices à l'intention des créateurs* abordent la nécessité de mesures de sécurité, de sauvegardes régulières des données opérationnelles et de protection contre les obsolescences du matériel et des logiciels. De plus, il est recommandé que les créateurs de données élaborent une stratégie de préservation, considèrent les questions de préservation à long terme et désignent un gestionnaire de confiance.

2.3 Lignes directrices à l'intention des préservateurs de données géospatiales

Les *Lignes directrices à l'intention des préservateurs* (InterPARES 2, 2007) ont été conçues pour offrir des directives concrètes aux divers groupes responsables de la préservation à long terme des documents d'archives numériques et sont axées sur un cadre de la **chaîne de préservation**, qui est «un système de contrôles qui s'étend sur tout le cycle de vie des documents d'archives afin d'en préserver l'identité et l'intégrité dans le temps» (InterPARES 2, 2013). Ce cadre comprend les politiques, les stratégies et les méthodes nécessaires à la gestion des documents d'archives numériques. La portée et les objectifs du programme de préservation de données géospatiales

doivent inclure une décision sur

quelles données géospatiales doivent être préservées, comment celles-ci doivent être rendues accessibles, à quelle communauté désignée et pour répondre à quels besoins et exigences techniques spécifiques. Les *Lignes directrices à l'intention des préservateurs* dirigent également vers un document de *cadre stratégique* (InterPARES 2, 2008), qui comprend les principes de la politique pour les créateurs et les préservateurs, et elles recommandent d'utiliser le *modèle de référence du SOAI* (CCSDS, 2012) et la *Liste de contrôle de la VADF* (OCLC et CRL 2007) pour élaborer les aspects fonctionnels des archives, qui seront examinés au chapitre 4.



PRATIQUE EXEMPLAIRE

Un bon exemple d'une ligne directrice pour l'élaboration d'un plan d'activités pour une initiative d'archivage de données géospatiales est le *Geoarchiving Business Planning Guidebook* (GeoMAPP, 2011a). Ce guide propose une approche détaillée à la description de la manière dont les objectifs d'archivage de données seront atteints, ainsi que la façon de décrire les justifications nécessaires pour l'initiative. Un résumé du contenu du guide est fourni à l'annexe F.

Des archives de données géospatiales nécessitent des ressources technologiques, humaines et financières et il est impératif que ces dernières soient garanties et durables par le biais de l'élaboration d'un plan d'activités solide. Ceci inclut une stratégie de communication claire pour convaincre les bailleurs de fonds potentiels, et la capacité d'attirer de nouvelles ressources une fois le programme de préservation en place. Les bailleurs de fonds peuvent être plus réceptifs à une collaboration et à une stratégie progressive d'acquisition des ressources et à une collaboration pour la mutualisation des ressources. Le *Geoarchiving Business Cost-Benefit Analysis Guidance Document* de Geospatial Multistate Archives and Preservation Partnership (GeoMAPP) peut aussi aider à justifier les initiatives en matière d'archivage et de préservation aux bailleurs de fonds (GeoMAPP, 2012).

Une fois les archives géospatiales établies, elles doivent avoir une fonction consultative envers le créateur du document pour s'assurer que la préservation fait partie du processus de création du document et qu'elle respecte les *Requirements for Assessing and Maintaining the Authenticity of Electronic Records* d'InterPARES 2 (InterPARES 2, 2002), qui fournissent des *Conditions de référence* pour assurer un bon environnement de préservation. Les *Lignes directrices à l'intention des préservateurs* recommandent que les préservateurs mettent en place des mesures de contrôle pour le « transfert, la maintenance et la reproduction des documents d'archives, y compris des procédures et des systèmes pour transférer des documents d'archives à leur organisation ou à un programme au sein de l'organisation, les maintenir et les reproduire » conformément aux *Conditions de référence* d'InterPARES 2 afin de garantir l'**identité** et l'**intégrité** des documents d'archives. De plus, la mise en œuvre de stratégies de maintenance claires est la clé d'une stratégie de préservation (voir l'encadré).

PRATIQUE EXEMPLAIRE

Les *Lignes directrices à l'intention des préservateurs* d'InterPARES recommandent les stratégies de maintenance qui suivent pour les données préservées (InterPARES 2, 2007) :

- la répartition claire des responsabilités;
- la mise en place d'une infrastructure technique appropriée;
- la mise en œuvre d'un plan pour la maintenance, le soutien et le remplacement des systèmes;
- la mise en œuvre d'un plan pour le transfert régulier de documents d'archives vers de nouveaux supports de stockage;
- le respect des conditions de stockage et de manipulation appropriées aux supports de stockage;
- la redondance et la sauvegarde régulière des entités numériques;
- l'établissement des mesures de sécurité des systèmes;
- la planification en cas de catastrophe.

PRATIQUE EXEMPLAIRE

Le projet National Geospatial Digital Archive (NGDA) des É.-U. partage ses politiques d'élaboration de collection, ses ententes de fournisseurs de données et un manuel de procédures, qui décrit les plans, les pratiques et les procédures de transfert (NGDA, 2009). Consulter le site <http://www.ngda.org/policies.html>

L'**évaluation** des documents d'archives fait partie de tous les plans de conservation, et les *Lignes directrices à l'intention des préservateurs* recommandent que les préservateurs collaborent avec les créateurs de documents d'archives afin de procéder à une évaluation archivistique et de déterminer les méthodes de transfert, dès le début du processus de création de données géospatiales. De plus, les préservateurs peuvent parfois trouver utile de participer à la conception des systèmes de création et de maintenance des documents d'archives, afin de renforcer la préservation autant que possible. L'évaluation consiste également à déterminer les

propriétaires des données géospatiales afin d'évaluer les ramifications de la préservation, ce qui est complexe lorsqu'il s'agit de données distribuées ou de données accessibles au moyen d'un portail de données géospatiales. L'évaluation de l'**authenticité** des données géospatiales est essentielle et peut être enregistrée dans un rapport d'évaluation en fonction des exigences de référence en matière d'authenticité. Les données géospatiales qui ont été retenues pour la préservation doivent également être contrôlées périodiquement afin de s'assurer qu'elles ne sont pas supprimées par accident, que les mises à jour logicielles ne changent pas leurs attributs, que les changements organisationnels n'ont pas de répercussions sur les décisions de conservation antérieures, et que les pratiques de gestion des documents d'archives sont respectées dans l'environnement du créateur. De plus, les composantes numériques doivent être définies et les relations implicites doivent être formulées explicitement dans les métadonnées et les composantes (p. ex., formats, conteneurs de fichiers, fichiers de formes ESRI) avant le transfert, afin de s'assurer qu'une fois—que les données sont extraites du système qui les a créés, elles peuvent être recrées de la manière prévue par le créateur. L'évaluation consiste également à déterminer la faisabilité de la préservation en étudiant attentivement et en évaluant les exigences de préservation techniques et les coûts connexes au cours de la période sélectionnée, que ce soit à court, à moyen ou à long terme.

Une fois les données géospatiales évaluées, elles doivent être acquises aux fins de préservation ce qui nécessite le déplacement des données de la garde du créateur à celle du préservateur, que ce rôle existe à l'intérieur ou à l'extérieur de l'organisation du créateur. Les exigences comprennent : un plan de transfert où les deux parties conviennent d'un format logique et d'un format physique, un examen des systèmes à l'intérieur desquels se trouvent les documents d'archives, et un processus convenu pour assurer leur transfert en toute sécurité dans un nouveau système. Cela inclura également le contrôle périodique et la mise en application des procédures de transfert et des tests préalables. Autant que possible, il est recommandé que le plus ancien format logique disponible soit maintenu. De plus, il est important d'éviter d'insérer des doublons en s'assurant que les documents d'archives sont transférés une seule fois, tout comme la documentation de tous les processus tels que la vérification des virus, la validation des données de contrôle et la confirmation de l'identité du document.

Les *Lignes directrices à l'intention des préservateurs* suggèrent qu'une fois que les données géospatiales versées dans un système de préservation de confiance, des règles et des procédures doivent être établies pour la production continue de copies authentiques, à mesure que les systèmes de préservation deviennent obsolètes et que les technologies doivent être améliorées. La description archivistique des données géospatiales est utile dans ce cas, et cela inclut la façon dont les données ont été recueillies, évaluées et traitées, ainsi que l'accès et la propriété intellectuelle et les droits de la protection des renseignements personnels. Comprendre les aspects de la gestion des droits numériques des données géospatiales est également important, car il peut y avoir des conséquences juridiques si vous contournez les environnements exclusifs durant l'extraction des données aux fins de la préservation. Il est également important d'analyser la stratégie de préservation sélectionnée pour s'assurer qu'elle est efficace et afin de maintenir un stockage approprié.

Le facteur final pour les préservateurs est de rendre accessibles les données géospatiales préservées. Les *Lignes directrices à l'intention des préservateurs* recommandent que les préservateurs fournissent de la documentation sur les processus de reproduction, de transfert et de contrôle périodique des données, qui permettent aux utilisateurs d'évaluer l'authenticité du document d'archives et de décider s'il provient d'une source fiable. De plus, le préservateur doit fournir les moyens technologiques afin de permettre aux utilisateurs d'avoir accès aux données géospatiales (p. ex., des données brutes par le biais d'un portail, à l'intérieur d'un nouvel environnement de cartographie, ou tout autre outil de visualisation). Les choix et les méthodes d'accès à la technologie doivent être fondés sur les compétences et les besoins des communautés désignées de l'archives.

3. Lois et politiques influant sur l'archivage et la préservation

Le but de ce chapitre est d'inciter le lecteur à considérer les mandats législatifs, notamment les règlements, les politiques et les directives qui devraient les obliger à examiner attentivement leurs pratiques de préservation. Alors que la discussion de ce chapitre est limitée au cadre législatif fédéral, des obligations similaires dans les lois, les politiques, les directives et les règlements provinciaux peuvent également être prises en considération.

3.1 Responsabilités liées à l'archivage et à la préservation

Les gouvernements sont responsables de la gestion et de la préservation des données et de l'information qu'ils produisent. Le gouvernement de l'Ontario, par exemple, doit se conformer à sa *Loi sur les Archives publiques et la conservation des documents*, et les Archives de l'Ontario fournissent une série de lignes directrices, de calendriers de conservation et de fiches de renseignements pour aider les bureaux du gouvernement avec leur tenue de documents d'archives (ministère des Services gouvernementaux de l'Ontario, 2011). Une des fiches de renseignements correspond aux *périodes de conservation, à l'accès aux archives et à la Loi sur l'accès à l'information*. La *Loi sur les archives* des Territoires du Nord-Ouest (T.N.-O.) prescrit la manière dont les documents d'archives sont gérés et une politique sur les archives des T.N.-O. oriente le processus (gouvernement des T.N.-O., 1993). Les T.N.-O. fournissent une liste des organismes et des ministères qui ont des calendriers spécifiques de conservation et de disposition des documents d'archives qui dicte également les décisions sur l'acquisition des documents d'archives (Prince of Wales Northern Heritage College, 2013).



Les organisations du gouvernement du Canada (GC) sont également touchées par la législation et la réglementation obligatoires qui peuvent dicter explicitement la gestion, la préservation ou le dépôt de certains documents d'archives qu'ils produisent. La première d'entre elles est la *Loi sur la Bibliothèque et les Archives du Canada (2004)* qui interdit la destruction de documents d'archives par toutes les institutions du GC sans l'obtention préalable d'une autorisation écrite. De plus, certaines lois et certains règlements font explicitement référence aux données géospatiales et à leur logiciel connexe, qui peuvent justifier la gestion des documents d'archives ou les mesures de préservation. Enfin, un certain nombre de directives et de politiques rendent obligatoires l'archivage et la préservation. C'est du Secrétariat du Conseil du Trésor (SCT) que

relève principalement la responsabilité de la politique sur la gestion de l'information au sein du gouvernement du Canada. Dans d'autres cas, les données sont préservées, car elles peuvent être considérées d'une importance capitale aux activités quotidiennes d'un organisme du GC (p. ex., les images du Centre canadien de télédétection discuté à la section 5.1), être utilisées pour éclairer les décisions du gouvernement (p. ex., le processus de l'Ontario sur l'information géographique déterminant quelles données sont officielles tel que discuté dans la section 5.2), avoir une valeur scientifique importante (p. ex., les données du ministère des Pêches et des Océans discutées à la section 5.3), ou être définies par les intervenants comme importantes, comme c'est le cas avec les données de l'Année polaire internationale discutées à la section 5.4.

Le rapport de 2011 de GéoConnexions *Archivage, gestion et préservation des données géospatiales – Rapport sommaire et recommandations* (HAL, 2011) mentionnait qu'une base de données géospatiales, un ensemble de données ou une carte peuvent être un document d'archives (InterPARES 2, 2013), à mettre de côté pour référence future et que les décisions liées à la préservation dépendraient des exigences législatives, réglementaires, politiques et en matière de gestion de l'information des organisations créant le document. Une institution créant des documents d'archives comme le ministère des Pêches et des Océans (MPO), par exemple, peut choisir des ensembles de données pour la préservation en fonction de la façon dont ceux-ci sont pris en compte, ou mis de côté afin que des mesures soient prises. Par exemple, les données du dénombrement des poissons utilisées pour établir les quotas de pêche au fil du temps peuvent être préservées, en raison de leur mérite scientifique ou parce qu'elles ont éclairé une décision stratégique clé du GC. Autrement, le *Règlement sur les cartes marines et les publications nautiques* (JC, 2007), conformément à la *Loi sur la prévention de la pollution des eaux arctiques* (JC, 2010), fait référence au Système électronique de visualisation des cartes marines (SEVCM) et sa base de données connexe de la carte électronique de navigation (CEN), et dans ce cas, le logiciel et les données peuvent être considérés comme dignes d'être préservés.

De plus, lorsque l'on considère la disposition des données, des questions comme « l'information contenue dans la base de données protège-t-elle les droits des citoyens? ou les intérêts du GC? » doivent également être prises en compte (BAC, 2009). Les lignes directrices pour aider les créateurs à décider quelles ressources d'information ont une valeur opérationnelle pour le Ministère sont décrites dans la trousse d'outils en tenue de documents (TD) de Bibliothèque et Archives Canada (BAC) (BAC, 2012). Les décisions visant à déterminer le moment et la fréquence de la saisie des instantanés de ces documents d'archives et le moment de les verser dans les archives doivent être prises entre les créateurs de documents d'archives, leurs intervenants et les préservateurs et ces décisions s'inscrivent dans le cadre des politiques de préservation de l'organisation. De plus, comme expliqué dans la section 2.2, un document est considéré stable lorsque sa forme est fixe et qu'il est mis de côté. Cette situation est problématique dans le contexte de bases de données et de cartes multiples qui sont continuellement mises à jour. Les données et le logiciel du MPO qui viennent d'être mentionnés pourraient rester actifs tout au long de leur cycle de vie et être également préservés. Il existe de nombreux modèles de cycle de vie à choisir et le *Review of Data Management Lifecycles Models*

produit par l'Université de Bath au Royaume-Uni explique huit modèles de façon tangible et concise (Ball, 2012).

La législation, la réglementation, les directives et les politiques discutés brièvement dans les sections qui suivent doivent être prises en compte lors de l'élaboration de la mission, du mandat et des objectifs des archives. Les provinces et les territoires peuvent avoir une législation et une réglementation similaires, et les villes et les municipalités peuvent avoir besoin de les respecter ainsi que toutes les résolutions, les politiques et les directives locales.

3.2 Lois et règlements fédéraux

Les organisations produisant des données géospatiales du gouvernement du Canada sont mandatées pour se conformer aux obligations de plusieurs lois et règlements primordiaux. Les obligations, les limitations et les défis liés à la législation qui suit sont résumées à l'annexe C et sont davantage examinées dans l'*Archivage, gestion et préservation des données géospatiales – Rapport sommaire et recommandations* (GéoConnexions et HAL, 2011).

- *Loi sur la Bibliothèque et les Archives du Canada* (Loi sur la BAC) : Cette loi stipule « que le patrimoine documentaire du Canada soit préservé pour les générations présentes et futures », que le Canada se dote d'une institution qui soit une source de savoir permanent accessible à tous, que cette institution puisse faciliter au Canada la concertation des divers milieux intéressés à l'acquisition, à la préservation et à la diffusion du savoir et qu'elle soit la mémoire du GC. De plus, la Loi sur la BAC stipule que « le ministre peut constituer un comité consultatif qu'il charge de conseiller l'administrateur général sur la façon de faire connaître le patrimoine documentaire aux Canadiens et à quiconque s'intéresse au Canada, et de le rendre accessible » (JC, 2004).
- *Loi sur le droit d'auteur* : Cette loi stipule que les créateurs de documents d'archives peuvent exercer un contrôle sur leurs œuvres, les œuvres créés par les employés du GC contenant des données géospatiales appartiennent à Sa Majesté, et que ces œuvres peuvent être autorisées sous licence (JC, 1985). Il est important de noter que le droit d'auteur du Canada ne subsiste pas dans les données en elles-mêmes, mais peut subsister dans la sélection ou la disposition d'*origine* des données. Pour qu'une compilation des données soit protégée par le droit d'auteur, il doit y avoir un auteur qui a créé quelque chose « d'original ». La Cour suprême du Canada dans sa décision de CCH 2004 a déclaré que l'*originalité* exige qu'une œuvre « doit être le produit de l'exercice du talent et du jugement d'un l'auteur » qui « ne doit pas être négligeable au point qu'on puisse le qualifier d'entreprise purement mécanique » ([2004] I R.C.S. 339). En date d'aujourd'hui, il n'y a pas de licences universelles pour les données géospatiales ou autres utilisées au sein du GC, mais il y a des lignes directrices telles que *La diffusion de données géographiques gouvernementales au Canada – Guide des pratiques exemplaires* (GéoConnexions, 2008). La Directive sur le gouvernement ouvert devrait produire une licence de « données libres », intitulée la licence du gouvernement ouvert, qui offrira une licence commune pour obtenir

de l'information et des données du GC, notamment les données géospatiales, dans un proche avenir. D'autres administrations qui possèdent des droits d'auteur de la Couronne adoptent des licences interopérables internationales telles que des licences Creative Commons et d'utilisation des données ouvertes (GéoConnexion et HAL, 2011).

- La *Loi sur l'accès à l'information* « a pour objet d'élargir l'accès aux documents de l'administration fédérale en consacrant le principe du droit du public à leur communication, les exceptions indispensables à ce droit étant précises et limitées et les décisions quant à la communication étant susceptibles de recours indépendants du pouvoir exécutif » (JC, 1985).
- *Loi sur la protection des renseignements personnels* et *Règlement sur la protection des renseignements personnels* : L'esprit de cette loi est de protéger « des renseignements personnels relevant des institutions fédérales et de droit d'accès des individus aux renseignements personnels qui les concernent » (JC, 1985). La *Loi sur la protection des renseignements personnels* énonce également comment les données doivent être maintenues, et la qualité des données ainsi que les règles de disposition. De plus, les renseignements personnels ne doivent être utilisés qu'aux fins pour lesquelles ils ont été initialement recueillis et des règles strictes s'appliquent sur la façon dont les renseignements personnels peuvent et ne peuvent être divulgués.
- *Loi sur la protection des renseignements personnels et les documents électroniques* : Le but de cette loi, qui s'applique au secteur privé au Canada, est de « fixer, dans une ère où la technologie facilite de plus en plus la circulation et l'échange de renseignements, des règles régissant la collecte, l'utilisation et la communication de renseignements personnels d'une manière qui tient compte du droit des individus à la vie privée à l'égard des renseignements personnels qui les concernent et du besoin des organisations de recueillir, d'utiliser ou de communiquer des renseignements personnels à des fins qu'une personne raisonnable estimerait acceptables dans les circonstances » (JC, 2000).
- *Loi sur la preuve au Canada* : Cette loi prévoit des déclarations claires sur la fourniture de la preuve documentaire et, entre autres choses, la façon d'évaluer l'authenticité, l'intégrité et l'accréditation des documents d'archives numériques déposés en preuve (JC, 1985).

La législation qui fait référence aux documents d'archives géospatiaux a été examinée en détail dans l'*Archivage, gestion et préservation des données géospatiales – Rapport sommaire et recommandations* (GéoConnexions et HAL, 2011) et un résumé des recommandations se trouve à l'annexe E de ce rapport. Dans la plupart des cas, il est recommandé que les données, les cartes et les logiciels soient créés et gérés selon les dispositions de la *Loi sur la BAC*, en gardant à l'esprit les lois primordiales comme la *Loi sur le droit d'auteur* et les politiques et directives du SCT. Dans certains cas, des documents d'archives spécifiques doivent être préservés dans l'éventualité où ils seraient requis en preuve (p. ex., la *Loi de mise en œuvre de l'Accord atlantique Canada – Terre-Neuve*).

3.3 Politiques et directives fédérales

Le Secrétariat du Conseil du Trésor du Canada (SCT) a élaboré des politiques et des directives pour les créateurs de documents d'archives du GC en général et pour les producteurs de données géospatiales du GC en particulier. Celles-ci sont énumérées ci-dessous alors que l'annexe D fournit un résumé des objectifs et des obligations des créateurs de données géospatiales. Les politiques et directives du SCT ont une incidence sur la façon dont les ressources d'information sont gérées au cours du cycle de vie d'un document et si elles sont respectées, faciliteront le processus de préservation des données géospatiales. La norme sur les données géospatiales du SCT fait référence à la description et à la diffusion; les producteurs de données géospatiales du GC doivent adopter la norme *ISO 19115 : 2005 Information géographique - Métadonnées* et la norme *ISO 19128 Information géographique - Interface de carte du serveur Web*. Les autres politiques et directives sont des pratiques exemplaires générales pour tous les types de documents d'archives, notamment les données, les métadonnées et les technologies telles que les logiciels, y compris :

- Cadre stratégique pour l'information et la technologie
- Politique sur la gestion de l'information (Politique sur la GI)
- Politique sur la gestion des technologies de l'information
- Directive sur les rôles et responsabilités en matière de gestion de l'information
- Directive sur la tenue de documents
- Norme pour les systèmes de gestion électronique des documents et des dossiers (SGEDD)
- Normes sur les métadonnées
- Norme sur les données géospatiales

De plus, BAC, conformément à son mandat et à la Directive sur la tenue de documents du SCT, a produit un certain nombre de produits, d'outils, de lignes directrices et de méthodologies d'information pour aider les créateurs de documents d'archives, qui sont disponibles sur son site Internet (BAC, 2011).

4. Cadres d'archivage et de préservation

4.1 Introduction

Ce chapitre a pour but de présenter les deux cadres importants suivants et une liste de contrôle qui fournissent des directives aux créateurs et aux préservateurs de données géospatiales pour leurs processus de planification de la préservation et d'en faire un bref profil.

- Le *modèle de référence du Système ouvert d'archivage d'information (SOAI)* (CCSDS, 2012), élaboré par le Consultative Committee for Space Data Systems (CCSDS).
- Le *European Long Term Preservation of Earth Observation Space Data: European LTDP Common Guidelines* (LTDP Working Group, 2012), élaboré par le Long Term Data Preservation (LTDP) Working Group of the Ground Segment Coordination Body (GSCB).
- Le document *Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC Checklist)* [Critères et liste de contrôle de la vérification et de l'accréditation des dépôts fiables (VADF)] (OCLC et CRL, 2007), élaboré par le Center for Research Libraries et le Online Computer Library Center, Inc.

Le *modèle de référence du SOAI* et le *Cadre du LTDP* sont utilisés pour orienter le développement fonctionnel d'un système de préservation et, en même temps que la *Liste de contrôle de la VADF*, faire référence aux mêmes activités de préservation, à des niveaux de détail différents et dans un ordre différent en raison de leurs principes d'organisation adoptés et leur accent. Le *modèle de référence du SOAI* examine les éléments fonctionnels d'un système de préservation et ne tient pas compte de l'environnement du créateur. Le *Cadre du LTDP* est un système de préservation du SOAI adaptée aux environnements de données de la mission de l'OT. La *Liste de contrôle de la VADF* a été conçue comme un outil de vérification et d'accréditation de la conformité au modèle du SOAI, et est également utilisé pour autoévaluer et orienter la planification de la préservation. Comme la terminologie varie entre ces trois documents clés, dans la mesure du possible, le lecteur sera renvoyé à la terminologie sur les archives et aux concepts d'InterPARES 2.

4.2 Modèle de référence du Système ouvert d'archivage d'information

Reconnaissant que l'obsolescence rapide des technologies numériques créait un risque accru d'être incapable de restaurer, de rendre ou d'interpréter l'information préservée, le CCSDS a élaboré le *modèle de référence du SOAI* afin de fournir un cadre commun permettant de consulter les défis archivistiques. L'accent a été mis sur l'information numérique, aussi bien comme formes primaires de l'information détenue que comme information supplémentaire pour les documents préservés



numériquement et physiquement. Alors que le modèle adapte l'information qui est intrinsèquement non numérique (p. ex., une carte physique), la modélisation et la préservation de cette information n'est pas abordée en détail. Le modèle s'applique aux archives qui doivent accueillir des flux réguliers d'information (p. ex., réseaux de capteurs) ainsi que ceux qui expérimentent principalement des entrées irrégulières. Le modèle a été adopté par l'Organisation internationale de normalisation (ISO) comme la norme *ISO 14721:2012 Systèmes de transfert des informations et données spatiales -- Système ouvert d'archivage d'information (SOAI) -- Modèle de référence*.

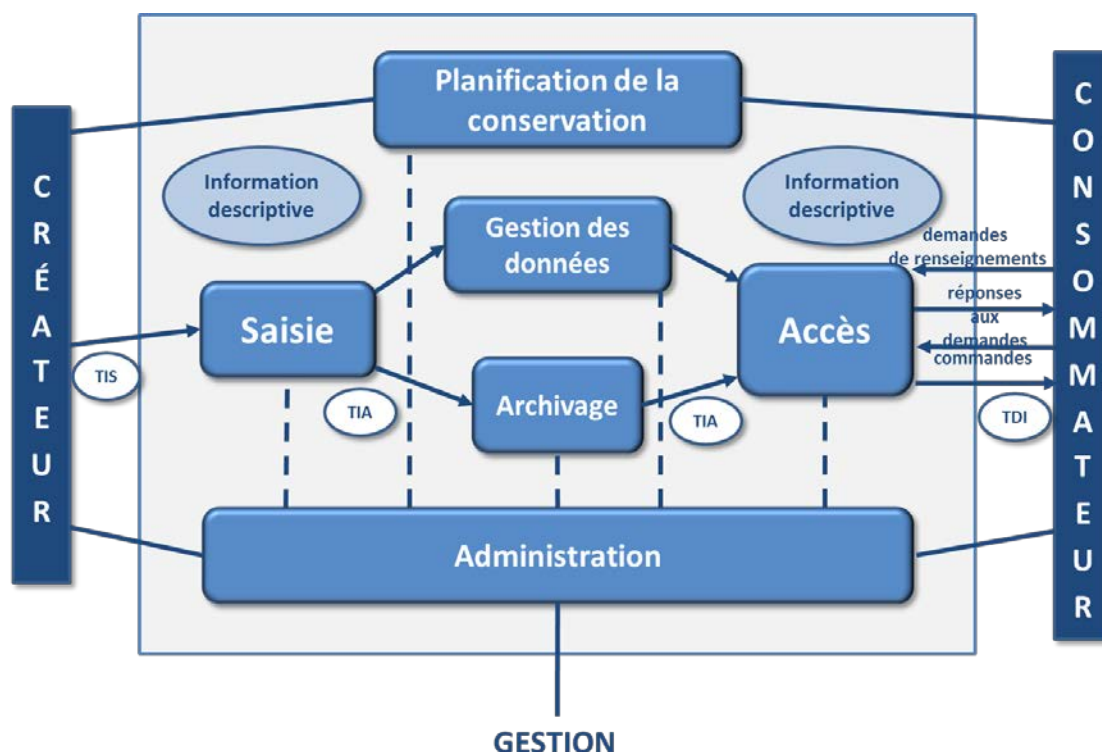
Un *Système ouvert d'archivage d'information* (SOAI) est défini comme étant un centre d'archives « constitué d'une organisation, qui peut faire partie d'une plus grande organisation, de personnes et de systèmes, qui a accepté la responsabilité de préserver l'information et de la rendre disponible à une communauté désignée » (CCSDS 2012, p. 1-13). Un « centre d'archives » dans le contexte d'un SOAI est analogue à un système de préservation des documents d'archives comme défini par le projet InterPARES 2. Le *Modèle de référence du SOAI* tient compte du caractère hautement réparti des banques d'information numérique et de la nécessité de mise en œuvre locales de politiques et de procédures efficaces soutenant la préservation de l'information. Cela permet une grande variété de dispositions organisationnelles, y compris les archives traditionnelles et les archives numériques réparties, ce qui est particulièrement adapté pour les environnements collaboratifs produisant des données géospatiales. L'information circule entre trois intervenants principaux du *Modèle de référence du SOAI* qui interagissent avec le système de préservation.

- *Créateur* : fournit l'information (fichiers, objets de données numériques) qui doit être préservée (c.-à-d. le créateur de données).
- *Gestionnaire* : définit les politiques générales des archives et les gère dans l'environnement du système de préservation (c.-à-d. le préservateur).
- *Consommateur* : interagit avec les services du système de préservation pour rechercher et acquérir des informations d'intérêt préservées, et pour y avoir accès (c.-à-d. l'utilisateur ou la communauté désignée).

Le modèle englobe les six entités fonctionnelles suivantes telles qu'illustrées à la figure 2.

- *Saisie* : L'interface entre le créateur et le conservateur. Cette entité reçoit les données des créateurs sous forme de trousse d'information sur les soumissions (TIS), qui peuvent être des données, une base de données, leurs logiciels et métadonnées connexes, et toute autre information descriptive liée au document qui pourrait être enregistrée dans un rapport d'évaluation. Pendant la saisie, une trousse d'information archivée (TIA) est créée, contenant la TIS convertie et une description archivistique ou l'information sur les descriptions de préservation.
- *Archivage* : Cette entité reçoit les TIA de la saisie pour un stockage permanent et la récupération et gère les processus de stockage de l'information pour le système de préservation.
- *Gestion des données* : Cette entité gère les documents d'archives du système de préservation et toutes les données administratives concernant la gestion du système de conservation.
- *Administration* : Cette entité contient les services et les fonctions pour contrôler l'ensemble des activités des autres entités du système de préservation. Cela comprend la négociation des ententes de dépôt avec les créateurs, la vérification des dépôts, la gestion des configurations matérielles et logicielles, et les activités de surveillance.
- *Planification de la préservation* : Cette entité s'assure que l'intégrité et l'identité des documents d'archives sont maintenues pendant les transformations telles que la migration et les mises à jour. Elle s'assure que les documents d'archives sont accessibles et compréhensibles à travers le temps et l'espace - qu'ils sont authentiques. Elle surveille également les besoins technologiques des consommateurs. C'est là que des plans de migration, des prototypes de logiciels et des plans d'analyse sont élaborés à des fins de migration.
- *Accès* : Cette entité permet d'accéder aux TIA en les récupérant du système, reçoit les demandes et y répond, et crée des trousse de diffusion de l'information (TDI) et les distribue aux consommateurs.

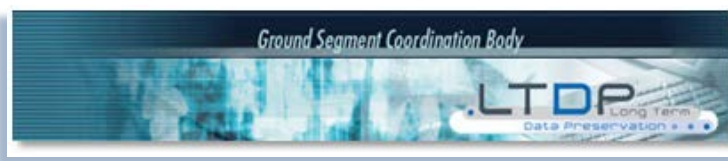
Bien que la direction ne joue pas un rôle opérationnel dans le système de préservation, elle fournit au système sa charte et sa portée et peut jouer un rôle de soutien très important. Par exemple, la direction peut mettre en œuvre des politiques qui exigent que toutes les activités financées dans sa sphère d'influence soumettent des produits de données au système de préservation et se conforment également aux normes et aux procédures. La direction fournit également des ressources de financement, évalue le rendement du système et peut participer à la résolution des conflits impliquant les créateurs, les consommateurs et l'administration interne, au besoin.

Figure 2 : Modèle fonctionnel du SOAI²

Source : CCSDS (2012. 4-1)

4.3 European Long Term Data Preservation Common Guidelines

En 2006, l'Agence spatiale européenne a entrepris une action de coordination pour élaborer une approche commune pour la préservation des données sur l'observation spatiale de la Terre. Le but était d'assurer la préservation à long terme de toutes les données européennes et canadiennes sur l'observation spatiale de la Terre, et d'en favoriser l'accès et la convivialité au moyen d'une approche commune, concertée et harmonisée qui est empruntée par les propriétaires de données sur l'observation spatiale de la Terre. Un document de synthèse sur les *European LTDP Common Guidelines* a été produit et aborde huit principaux « thèmes » sur la LTDP définissant pour chacun, le « principe directeur » et un ensemble de « lignes directrices clés » qui doivent être



appliquées pour garantir la préservation des données d'observation spatiale de la Terre à long terme.

² Les lignes reliant les entités dans le modèle déterminent les voies de communication par lesquelles l'information circule dans les deux sens. Les lignes vers l'administration sont pointillées dans le seul but de réduire l'encombrement du schéma.

Les *European LTDP Common Guidelines* sont compatibles avec la VADF (voir paragraphe suivant). Les mesures de la VADF sont une méthode d'accréditation et de vérification qui analyse la fiabilité de tous les types de dépôts numériques. Les *LTDP Guidelines* sont, d'autre part, un ensemble de recommandations pratiques portant spécifiquement sur les archives de données d'observation de la Terre et couvrant les mesures de la VADF dans la plupart des cas. Les *European LTDP Common Guidelines* sont un exemple de la façon dont une communauté spécifique productrice de données géospatiales (c.-à-d. les créateurs) a adapté le *modèle de référence du SOAI* à son contexte³. Dans le contexte de l'OT, la préservation est souvent une composante opérationnelle essentielle mise en œuvre dans le cadre de la gestion du cycle de vie d'un document étant donné que la préservation est obligatoire pour la plupart des créateurs de données d'OT. Le créateur dans ce cas est aussi la plupart du temps, le préservateur. L'étude de cas du SGDOT en est un exemple (se reporter à la section 5.1), où le CCT est à la fois le créateur et le préservateur des données. Dans ces directives, comme dans le SOAI, le terme « archives » est utilisé à la place de « système de préservation ». Les huit thèmes sont les suivants :

1. définition et évaluation du contenu des données préservées;
2. fonctionnement et organisation des archives;
3. sécurité des archives;
4. saisie des données;
5. tenue des archives;
6. accès aux données et interopérabilité de celles-ci;
7. exploitation et nouveau traitement des données;
8. prévention de l'élimination des données.

Chacun de ces thèmes est brièvement décrit ci-dessous.

La **définition et évaluation du contenu de données préservées** définit un ensemble cohérent et complet de données pour permettre son utilisation possible actuelle et future. Cela inclut les données, les logiciels de traitement, la documentation de la mission, la description et les métadonnées des archives et à l'intérieur de ceux-ci, des éléments liés aux indicateurs spécifiques de qualités des données d'OT (QA4EO, 2013). Cela comprend également des documents pertinents relatifs à la propriété intellectuelle, aux restrictions d'accès, aux spécifications et aux normes, soit des documents qui feraient partie d'un rapport d'évaluation. Ceci est analogue à une TIS et à une TIA dans le *modèle de référence du SOAI* et aux documents d'archives, tels que définis par le projet InterPARES 2.

Le **fonctionnement des archives** comporte toutes les activités technologiques et administratives quotidiennes qui sont effectuées pour exécuter et contrôler le système de préservation (p. ex., l'exécution et le contrôle des applications, la surveillance du système, les rapports d'anomalie, la récupération d'erreur, et les rapports et statistiques sur les activités). Une des

³ La norme ISO 14721 du SOAI a été utilisée dans la définition de la structure du document des *European LTDP Common Guidelines*.

directives ici est l'adoption du modèle de référence du SOAI. Le système de préservation se situe dans une organisation structurée pour être un centre d'archives, avec des mandats, des politiques, des lois et des ressources durables pour répondre aux objectifs et accomplir les tâches et les processus de préservation à long terme.

La **sécurité des archives** englobe toutes les activités dédiées à la mise en œuvre de mesures de sécurité pour l'accès et l'archivage des données afin de garantir la confidentialité, l'intégrité et la disponibilité des données archivées. Ceci assure que les données sont authentiques, et fournit aux utilisateurs l'information dont ils ont besoin afin de s'assurer que les documents d'archives dans le système sont ce qu'ils prétendent être, que la chaîne de traçabilité est sécurisée, et que les données n'ont pas été altérées. Cela couvre les exigences d'authenticité techniques et non techniques, notamment la sécurité physique, d'information et du personnel.

La **saisie des données** contient les fonctions et services qui, conformément à la norme du SOAI acceptent les trousseaux d'information sur les soumissions (TIS) des créateurs (p. ex., les données et les métadonnées des créateurs), préparent les trousseaux d'information archivée (TIA) aux fins de stockage (c.-à-d. fichiers restructurés, description archivistique, etc.), et veillent à ce que les TIA et leurs données descriptives de soutien (c.-à-d. les métadonnées d'archives) soient stockés dans le système de préservation.

La **tenue des archives** comprend toutes les activités visant à garantir l'intégrité des données archivées. L'intégrité des données garantit que les données archivées sont complètes et inchangées par la perte, l'altération ou la corruption de données. La tenue des archives s'appuie sur le stockage d'équipement et les supports de données dans des locaux sécurisés et à environnement contrôlé et un ensemble d'activités définies à effectuer sur une base régulière (p. ex., la migration vers de nouveaux systèmes et médias en conformité avec la technologie et l'évolution du marché de consommation, le compactage des données et la conversion du format/des trousseaux de données).

L'**accès aux données** correspond aux services et aux fonctions qui rendent les banques d'information d'archives et les services connexes visibles pour les consommateurs. L'**interopérabilité** est liée à la possibilité d'accéder aux données de façon commune et normalisée en dépit des différences intrinsèques entre les ensembles de données et les systèmes utilisés pour y accéder. Cela comprend l'extraction et la livraison des données sous la forme de TIS, et puisque les données d'OT sont hétérogènes, l'interopérabilité et l'harmonisation de l'accès aux données sont très importantes.

PRATIQUE EXEMPLAIRE

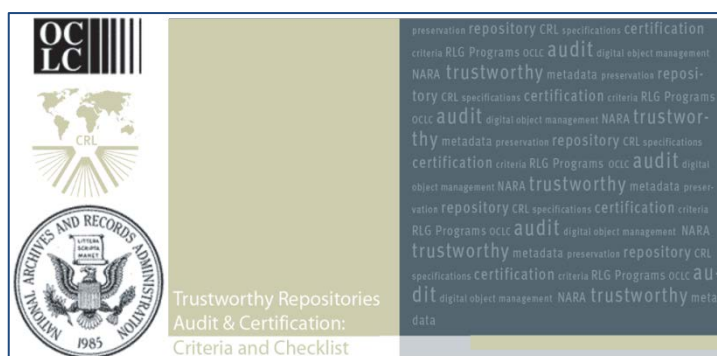
Le document EO Preserved Data Set Content LTDP/PDSC du LTDP (LTDP 2012) est une liste de contrôle sur la préservation utilisée par les créateurs et producteurs de données d'OT pour orienter les décisions de préservation spécifiques de la mission. C'est un excellent exemple pour les producteurs de données géospatiales qui préparent des rapports d'évaluation et des inventaires de données. Le document comprend également une annexe qui cartographie le contenu PDSC au modèle de référence du SOAI.

Le thème sur l'**exploitation et le nouveau traitement des données** contient toutes les activités d'exploitation liées au traitement et au nouveau traitement des données, à la régénération ou à l'amélioration des catalogues (p. ex., avec l'exploration de données), l'intégration de nouveaux services (p. ex., grâce à l'orchestration du déroulement des activités du service) et l'évaluation de la qualité des produits et des services. Il s'agit de garantir la réutilisation de ces données au fil du temps.

Le thème de la **prévention de l'élimination des données** définit un ensemble de procédures à appliquer dans le but de prévenir ou de minimiser la perte de données sur l'observation spatiale de la Terre et s'assurer que les ressources sont appliquées aux activités de préservation et d'accès aux données d'OT au moyen d'un processus d'approbation de la direction. Ce thème revêt une importance particulière lorsque les détenteurs de données d'observation spatiale de la Terre et les propriétaires d'archives ne peuvent plus préserver les données.

4.4 Trustworthy Repositories Audit & Certification: Criteria and Checklist

Le document *Trustworthy Repositories Audit & Certification: Criteria and Checklist* [Critères et liste de contrôle de la vérification et de l'accréditation des dépôts fiables (VADF)] (OCLC et CRL, 2007) fournit aux responsables de dépôts numériques un outil objectif pour mesurer la fiabilité de leur système de préservation. Elle est devenue depuis la norme ISO 16363:2012 *Systèmes de transfert des informations et données spatiales - Audit et certification des référentiels numériques de confiance* (ISO, 2012).



La *Liste de contrôle de la VADF* est divisée en trois sections : l'infrastructure organisationnelle, la gestion des objets numériques (c.-à-d. documents d'archives ou définition du contenu des données préservées), et les technologies, l'infrastructure technique et la sécurité. Chacune d'elles est accompagnée d'une description des attentes générales et est divisée en un certain nombre d'éléments avec les critères et les sous-critères connexes représentés sous la forme d'une série de questions. Les questions portent sur des méthodes pour mesurer l'adéquation (p. ex., fondée sur des indicateurs tels que la preuve documentaire, le degré de transparence, l'atteinte ou non des objectifs déclarés par le service des archives), et la mesurabilité (p. ex., fondée sur des indicateurs tels que le degré de fiabilité). De plus, on s'attend à ce que les systèmes de préservation soient conformes à un certain nombre de normes ISO en matière de qualité, de sécurité et de processus et la norme ISO 14721 sur le *modèle de référence du SOAI*. La *Liste de*

contrôle de la VADF facilite l'évaluation objective de la conformité au SOAI. Voici le contenu des trois sections.

- A. **Infrastructure organisationnelle** – comprend les éléments suivants : la gouvernance, la structure organisationnelle, le mandat ou l'objectif, la portée, les rôles et les responsabilités, le cadre stratégique, le système de financement, les enjeux et les actifs financiers, les marchés, les licences et les passifs, et la transparence. Ces éléments sont évalués en fonction des **critères** suivants : la gouvernance et la viabilité organisationnelle, la structure et la dotation en personnel, le processus de reddition de comptes et le cadre stratégique, la viabilité financière et les marchés, les licences et les passifs.
- B. **Gestion des objets numériques** – comprend les **éléments** organisationnels et techniques liés aux fonctions de dépôt, aux processus et aux procédures nécessaires à la saisie, à la gestion et à l'accès aux objets numériques (c.-à-d. les documents d'archives, le contenu des ensembles de données conservées) à long terme. Les exigences sont les suivantes et elles sont axées sur la fonctionnalité du système de préservation en fonction des entités fonctionnelles du SOAI discutées dans la Section 4.1.
- **Saisie** : l'acquisition d'objets numériques.
 - **Saisie de la création des TIA** : la transformation des objets numériques (c.-à-d. TIS) en TIA.
 - **Planification de la préservation** : stratégies de préservation actuelles, solides et documentées avec des mécanismes pour les tenir à jour.
 - **Archivage et préservation/tenue des TIA** : conditions minimales pour exécuter la préservation à long terme des TIA.
 - **Gestion de l'information** : la capacité du système de préservation à produire et diffuser des versions précises et authentiques des objets numériques.
 - **Gestion de l'accès** : disposition d'accès pour les utilisateurs.
- C. **Technologies, infrastructure technique et sécurité** – ne dicte pas quel matériel et quels logiciels utiliser, mais décrit plutôt les pratiques exemplaires portant sur la gestion et la sécurité des données. Les voici.
- **Infrastructure générale du système**
 - **Technologies appropriées** : bâtir à partir de la configuration système requise et les besoins des communautés désignées.
 - **Sécurité** : comprend les systèmes des TI et la protection physique contre les incendies, les inondations, et les gestes posés par des gens, la gestion des risques de l'infrastructure technique et la gestion des risques de sécurité.

La *Liste de contrôle de la VADF* est particulièrement utile, car elle invite les évaluateurs à chercher des composantes, des documents et des processus concrets qui ne sont pas explicitement mentionnées dans le *modèle de référence du SOAI* ou les *European LTDP Common Guidelines*. La *Liste de contrôle* considère l'environnement des préservateurs.

Le premier dépôt numérique fiable (DNF) accrédité au Canada est le Scholars Portal du Conseil des bibliothèques universitaires de l'Ontario. Le mandat du Conseil des bibliothèques universitaires de l'Ontario est de fournir « un environnement de préservation à long terme solide pour tout le matériel de sa collection, en travaillant à partir de concepts énoncés dans le modèle de référence du Système ouvert d'archivage d'information (SOAI) et codifié dans la *Liste de contrôle de la VADF* et la norme ISO 16363 (*Audit et certification des référentiels numériques de confiance*) ». Le rapport de vérification, les politiques et les plans décrivant le déroulement des activités, et le mandat et les réponses à la liste de contrôle de la vérification sont disponibles sur le site Web du CBUO (CBUO, 2011), où son *Preservation Action Plan – Journals* peut également être localisé. Les données géospatiales ne sont pas actuellement saisies sur le Scholars Portal.

5. Exemples de préservation de données géospatiales

Ce chapitre met en évidence quatre exemples d'initiatives de préservation des données numériques qui sont de bonnes sources de pratiques exemplaires dans le domaine.

5.1 Étude de cas : Système de gestion des données d'observation de la Terre (SGDOT)⁴

5.1.1 Introduction

Le mandat principal du Centre canadien de télédétection (CCT) de Ressources naturelles Canada (RNCCan) est de donner accès à la technologie satellitaire pour surveiller les frontières et les terres du Canada. Le CCT est également le centre d'expertise du GC en matière de télédétection et de géodésie. De plus, le CCT doit se conformer à la *Loi sur les systèmes de télédétection spatiale* (JC, 2007) et au Règlement sur les systèmes de télédétection spatiale (JC, 2007), qui énoncent explicitement que les données brutes seront préservées et que les documents d'archives seront gérés. Grâce à des partenariats avec les intervenants du gouvernement, des liens étroits avec le milieu universitaire et le secteur privé, notamment des collaborations internationales, le CCT garantit que les données satellites sont disponibles pour répondre aux besoins du gouvernement canadien et des Canadiens. Depuis 1971, le CCT a accumulé des archives d'environ 800 téraoctets de données d'OT provenant de différents satellites et de capteurs aéroportés (CCT, 2013).

L'imagerie d'OT du CCT et les produits dérivés appuient les priorités du GC, notamment le développement économique du Nord, la sûreté, la sécurité, la souveraineté et la surveillance environnementale. Les produits sont livrés aux intervenants du gouvernement et aux Canadiens par l'entremise du Catalogue d'observation de la Terre (CEOCat), de la Photothèque nationale de l'air (PNA) et du Cadre national des données d'observation de la Terre (CNDOT). Le CNDOT est un système de préservation pilote d'OT qui rend les données d'OT disponibles en temps opportun aux organisations du gouvernement fédéral. Le CCT saisi, préserve et donne accès aux données de télédétection acquises de deux stations canadiennes de réception satellite existantes, les Offres à commandes principales et nationales (OCPN) pour l'achat d'images satellite commerciales (ISC) et d'autres fournisseurs de données d'OT internationales et du secteur privé.

⁴ Une grande partie du contenu de cette étude de cas a été dérivée d'une entrevue en profondeur avec les principaux membres de la Division de l'acquisition des données du CCT qui participe au projet sur les Services et accès de données d'observation de la Terre du Canada.

D'un point de vue organisationnel, la Division de l'acquisition des données du CCT est responsable de l'acquisition de données d'OT au Canada, de l'entretien du système de préservation des données d'OT, de l'élaboration de systèmes au sol de pointe pour la réception de données efficace, de la préservation et de la distribution des données d'OT, de l'élaboration de systèmes de traitement conviviaux de données d'OT, et de la fourniture de services informatiques et d'ordinateur au CCT. Les services de réception de données, de diffusion et d'accès aux archives des Services de données d'observation de la Terre (SDOT) sont fournis par deux stations de réception satellite du CCT à Prince Albert en Saskatchewan et à Gatineau au Québec et à la Station-relais pour satellites d'Inuvik (SRSI) qui sera bientôt en ligne. L'infrastructure de la composante terrestre des SDOT est capable de recevoir des données directement du territoire nord-américain. Les données mondiales peuvent aussi être reçues par des satellites équipés d'enregistreurs de bord tels que RADARSAT. Les stations au sol reçoivent des données d'OT depuis plusieurs détecteurs à bord de satellite et conservent des données d'archives qui remontent à 1972.

Ressources naturelles Canada a élaboré un plan de relance pour améliorer la capacité de transmission par satellite d'OT du Canada et accéder aux données d'OT, qui a été pris en charge dans le budget fédéral de 2012. En conséquence, le CCT se procure un Earth Observation Data Management System (SGDOT), qui est un système de gestion et de préservation de données et d'accès à ces dernières. Le CCT a aussi obtenu des ressources du programme GéoConnexions afin d'élaborer un projet sur les Services et accès de données d'OT du Canada, qui définit la stratégie de préservation des données d'OT (cadre stratégique et lignes directrices) et des flux d'exploitation de la LTDP pour les activités du SGDOT. Le SGDOT fait partie de la mise en œuvre des Services et accès de données d'OT du Canada et il deviendra une fonction essentielle du CCT et un système de préservation d'OT pour le GC.

5.1.2 Modèle opérationnel en cours d'utilisation – Mise en œuvre à ce jour

En 2010, le CCT a terminé la centralisation des données d'OT dans une nouvelle installation de préservation à Ottawa. Le modèle de système de préservation précédent était un centre d'archives axé sur la mission de capteur et de satellite standard et les utilisateurs étaient principalement des spécialistes d'OT. Le nouveau modèle se dirige vers une architecture axée sur les services (AAS). Une demande de proposition (DP) a été lancée en janvier 2013 pour créer le Système de gestion des données d'observation de la Terre (SGDOT), qui sera le système de préservation pour les archives du CCT. Le SGDOT adhèrera à la norme *ISO 14721:2012 Systèmes de transfert des informations et données spatiales - Système ouvert d'archivage d'information (SOAI) - Modèle de référence* du SOAI et aux *European LTDP Common Guidelines*. Une fois établi, le SGDOT visera la conformité avec les pratiques recommandées de la norme *ISO 16363:2012 Systèmes de transfert des informations et données spatiales – Audit et certification des référentiels numériques de confiance*.

Voici les deux étapes de mise en œuvre des Services et accès de données d'OT du Canada (CCT, 2013).

- Élaborer un cadre stratégique et des lignes directrices.
- Mettre en œuvre le cadre et les lignes directrices, en présentant le système de préservation SGDOT et aussi fournir une série de procédures opérationnelles et de la documentation.

Le CCT sera en mesure de gérer l'augmentation des collectes de données d'OT à l'intérieur du cadre du comité fédéral sur la géomatique et les observations de la Terre (CFGOT) et de ses activités actuelles de collecte de données d'OT, qui sont uniques au Canada. Améliorer l'accès à ces données facilitera la création de connaissances et de produits dérivés de l'OT, ainsi que l'intégration des données d'OT avec d'autres données scientifiques. Le SGDOT sera également en mesure de tenir compte de l'augmentation du volume de données générées à l'avenir, comme de la mission de la Constellation RADARSAT, de la mission de poursuite du programme Landsat et des missions Sentinel.

SGDOT proposé

Le SGDOT remplacera les systèmes de catalogue existants du CCT et améliorera la fonctionnalité de RNCAN, permettra aux utilisateurs autorisés et au public d'accéder aux données d'imagerie (données brutes, produits, produits dérivés d'OT et photographie aérienne (notamment la technologie LiDAR)) des archives de données nationales d'OT multimission du CCT.

Le SGDOT saisira et gèrera les ensembles de données, et inclura la préservation et la mise en catalogue des données, le traitement, la gestion de licences de l'utilisateur final et les trousseaux des produits aux fins de diffusion au moyen de diverses normes de l'Open Geospatial Consortium (OGC). Le système offrira des services et des portails d'accès bilingues qui sont gérés par le biais de profils d'utilisateurs complexes. Il sera modulaire, évolutif et flexible pour répondre aux besoins actuels et futurs et possèdera des API entièrement documentées afin que RNCAN ou des développeurs tiers soient en mesure de personnaliser, d'agrandir, d'améliorer ou d'ajouter des fonctionnalités au système de préservation. Le SGDOT utilisera vraisemblablement des logiciels commerciaux sur étagère (COTS) et libres pour réduire le coût de l'entretien, de la personnalisation et du remplacement du cycle de vie du système de préservation et de ses applications. Il est également possible qu'un logiciel privé soit utilisé en dehors de la solution de base.

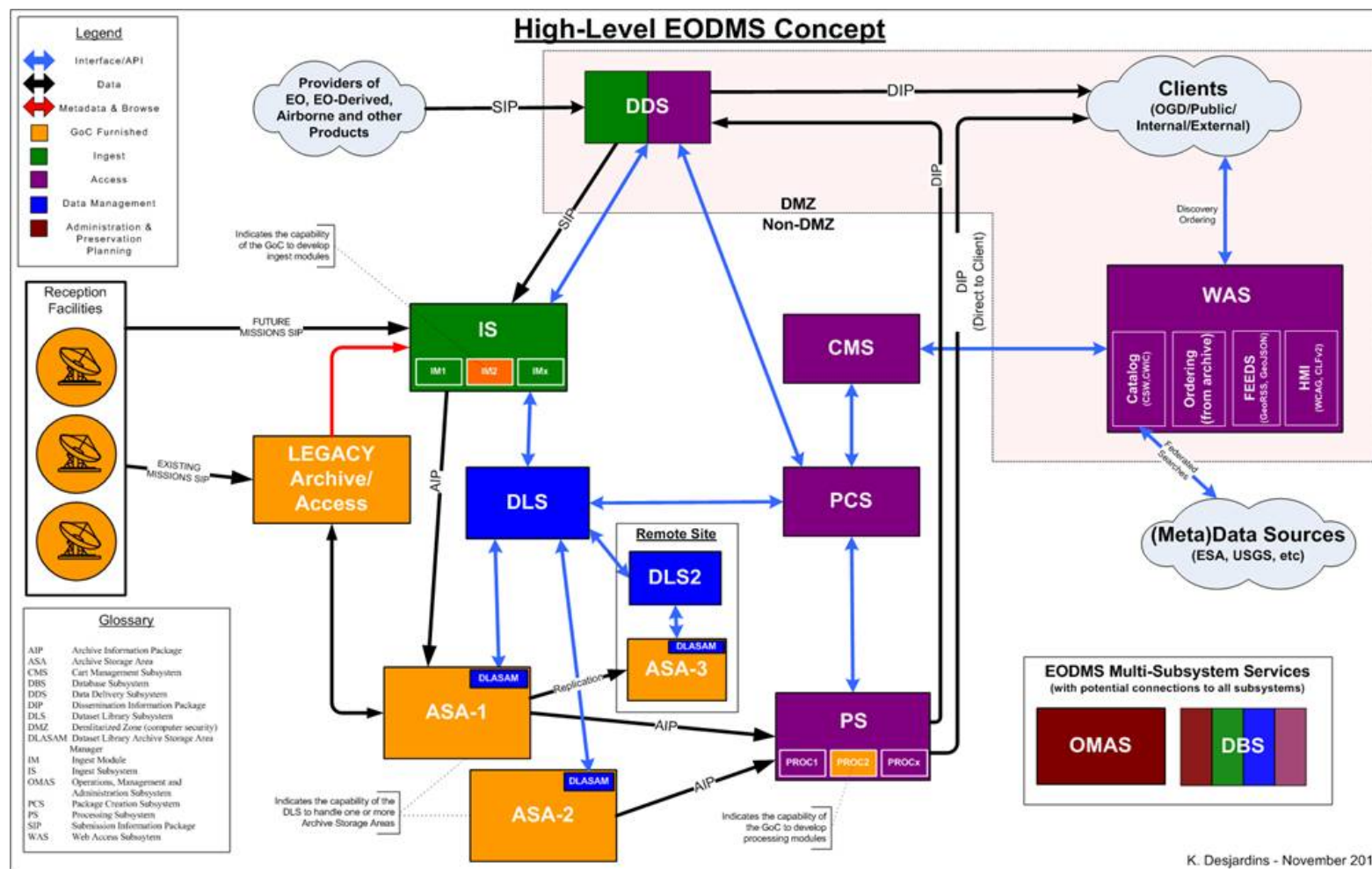
Le SGDOT migrera ou intégrera les technologies de plateforme existantes comme la base de données et les schémas PostgreSQL, le catalogue des comptes de clients et les serveurs FTP, la configuration du serveur FTP et les archives de la rue Booth. De plus, l'initiative comprendra un programme de formation pour les administrateurs, soutenu par un Manuel de formation pour l'administrateur et les opérateurs, des Manuels de formation de l'utilisateur pour les opérateurs et l'administrateur, des Manuels COTS du vendeur et des documents d'aide et de FAQ concernant le SGDOT (solution).

Le SGDOT saisira les types d'ensembles de données qui suivent : données brutes provenant des processeurs ou des installations de réception des données du Canada, produits d'image de fournisseurs commerciaux d'images de satellite, photos, technologie LiDAR, et fournisseurs d'ensembles de données hyperspectrales acquises par avion, archives de donnée satellitaires à long terme dérivées d'OT provenant de sources internes, et autres produits dérivés d'OT de l'interne ou des autres ministères. Celles-ci sont soumises avec des métadonnées, des attributs et des licences comme décrit dans le document LTDP/Preserved Data Set Content. Voici une sélection de quelques-unes des fonctionnalités du SGDOT, qui sont illustrées dans la Figure 3.

- Saisir des TIS de missions multiples et avoir la capacité de visualiser le contenu de la TIS en format ISO 19115, avec des plans de validation et de vérification des données. Les formats d'image qui seront saisis sont : GeoTIFF, NITF, JPEG2000, CEOS, TIFF et JPEG. Le système saisira également des ensembles de données d'OT dérivées de la trame ou du vecteur comme les fichiers de forme ESRI et en format KML/KMZ, avec des métadonnées et un domaine de trame générique avec des métadonnées.
- Voir, gérer, mettre à jour, supprimer et créer des TIA et des TDI grâce à une chaîne de flux de données configurable avec les rapports, les vérifications, les enregistrements chronologiques des données et les défaillances et les alertes de sécurité appropriés.
- Mettre en place différents comptes utilisateurs avec des règles d'accès liées aux attributs des ensembles de données et respecter les règles et les restrictions de licences associées aux ensembles de données.
- Soutenir la découverte d'ensembles de données multimission grâce à un catalogue et une interface de Service de catalogue Web (CSW).
- Assurer l'intégrité et la cohérence des données dans l'ensemble du système et prévenir la corruption, l'abandon ou la perte d'information.
- Soutenir trois environnements de travail
 - *Environnement d'élaboration* – utilisé pour élaborer un nouveau code et de nouvelles interfaces.
 - *Environnement d'analyse* – utilisé pour appliquer des correctifs, des mises à jour et de nouvelles fonctionnalités avant d'être déployés à l'environnement de production. L'environnement d'analyse sera un duplicata de l'environnement de production, car il sera utilisé pour exécuter des simulations de toutes les modifications apportées au système.
 - *Environnement de production* – utilisé pour les activités quotidiennes normales.
- Assurer l'état dans tous les sous-systèmes de la gestion des activités et des entités fonctionnelles de l'administration et générer des rapports présentés au client.
- Gérer les accords de licence d'utilisateur final (ALUF) propres à un ensemble de données individuelles.
- Créer des profils de TDI pour un sous-système de création de trousse, qui définissent des ensembles de données et leurs composantes.
- Signaler et déterminer les interactions des utilisateurs avec le système et effectuer des contrôles de sécurité.

- Le client du sous-système d'accès exportera des centres d'intérêt (CI) à partir d'une interface de cartographie ou d'un profil de client comme un fichier CI (p. ex., fichier de forme ESRI, format KML/KMZ, ou GML) et les utilisateurs seront en mesure de « découper » certaines collections ou sous-collections d'ensembles de données en définissant une zone de découpage ou une étendue géographique à l'intérieur de l'interface de cartographie Web.

Figure 3 : Description du sous-système SGDOT de haut niveau [original demandé par le CCT]



Source : CCT, 2013

Nota – Orange : fournit par le GC; vert : saisie; mauve : accès; bleu : gestion des données; brun : administration et planification de la préservation.

Liens du SGDOT

La stratégie de préservation des Services et accès de données d'OT du Canada s'harmonise avec le mandat du CCT, conformément à la *Loi sur le ministère des Ressources naturelles* visant à « ...participer à l'élaboration et à l'application de normes et de codes en matière de levés techniques, géophysiques et géodésiques... » et « favoriser le développement et l'exploitation de la technologie en matière de télédétection » (1994, ch. 41, art. 6, paragraphes d et g). Aussi, les Services et accès de données d'observation spatiale de la Terre du Canada et le SGDOT contribuent au pilier d'administration des données d'observation spatiale de la Terre du CCT établi par le Directeur général en 2011. Le projet des Services et accès de données d'OT du Canada s'harmonise également avec les politiques suivantes du GC : la gestion de l'information, la norme de sécurité opérationnelle, la Gestion de la sécurité des technologies de l'information (GSTI), la facilité d'emploi des sites Web, l'accessibilité et la normalisation, la norme géospatiale, la norme sur les métadonnées, les technologies de l'information et l'interopérabilité Web. Le SGDOT bénéficiera à l'ICDG puisqu'il élaborera un catalogue à la fine pointe de la technologie et un portail pour l'exploration de données axés sur des normes adoptées par l'ICDG et seront accessibles aux intervenants de l'ICDG. De plus, le CCT doit se conformer au protocole d'entente (PE) propre à la Mission, aux ententes avec les fournisseurs de données du secteur privé, aux Offres à commandes principale et nationale (OCPN), à la loi et au règlement sur la télédétection comme indiqué précédemment, et à d'autres lois générales discutées dans le chapitre 3.

Les Services et accès de données d'OT du Canada soutiendront également le mandat de l'Agence spatiale canadienne (ASC), qui participe à l'élaboration et au fonctionnement de nombreuses missions d'observation spatiale de la Terre et à l'élaboration de produits et services de données d'OT. L'ASC collaborera avec le CCT afin d'atteindre ses objectifs de prestation de services de données d'OT. Le ministère de la Défense nationale (MDN), grâce à son programme Polar Epsilon, a accès aux données RADARSAT-2 qui sont préservées au CCT. Également, le Service de cartographie du MDN consolide une partie de ses services d'accès et vise à interopérer avec les banques de métadonnées du CCT et à rationaliser son processus de commande d'OT. Le CCT collabore également avec le German Aerospace Center (DLR), la National Oceanic and Atmospheric Administration (NOAA) des É.-U. et l'Agence spatiale européenne (ASE) sur l'élaboration du SGDOT. DLR a une relation de longue date avec le CCT, la NOAA a dirigé et financé un projet de catalogue d'interopérabilité et le CCT travaille sur un prototype fonctionnel à cet effet, et l'ASE joue un rôle important dans l'élaboration du SOAI et a offert de partager des outils et de l'expertise en la matière.

Utilisateurs désignés du SGDOT

Les utilisateurs actuels des données d'OT du CCT comprennent 14 ministères et organismes fédéraux (p. ex., Service canadien des glaces, Service canadien des forêts, Parcs Canada, ministère des Pêches et des Océans, Environnement Canada, Agriculture et Agroalimentaire Canada), des universités et des organismes nationaux et internationaux. Certains de leurs

partenaires internationaux comprennent le United States Geological Survey (USGS), la National Aeronautics and Space Administration (NASA), l'ASE, la NOAA, la Brazil Space Agency (INPE) et le China National Satellite Meteorological Center. Le SGDOT donnera accès en fonction des besoins de ces derniers et d'autres utilisateurs. En janvier 2013, le CCT a mené une *enquête sur les exigences des utilisateurs relativement à la préservation à long terme des données d'OT du CCT*. Les résultats étaient destinés à déterminer les communautés désignées du SGDOT et d'informer les configurations du système, en particulier l'entité fonctionnelle d'**accès du modèle de référence du SOAI**. Les utilisateurs auront accès à un portail unique d'ensembles de données d'OT mondiales plutôt que d'avoir à accéder à des portails à missions multiples ou axés sur un thème, alors que le respect des normes d'interopérabilité augmentera la capacité d'intégrer des données plus facilement dans les outils de prise de décision.

5.1.3 Défis éprouvés

Le système de préservation actuel du CCT est limité dans sa capacité à fournir un accès facile aux données d'OT préservées, le volume des données augmente et sa technologie devient obsolète. La transition des archives de missions traditionnelles et le déplacement de sa collection à un nouvel endroit a nécessité la création d'un nouveau système normalisé de préservation numérique des données géospatiales. Il est important de garder à l'esprit que le CCT a su préserver et rendre disponible plus de 800 téraoctets de données de la mission d'OT pendant plus de 40 ans malgré les changements importants dans la technologie, les médias de stockage et l'informatisation au cours de cette période. Les efforts des chercheurs d'OT du CCT, des professionnels et des gestionnaires des TI représentent les principales pratiques de préservation, étant donné qu'aucune autre organisation du GC productrice de données géospatiales ne peut faire de telles allégations. Il n'est donc pas surprenant que le système actuel présente certaines lacunes déterminées suivantes qui doivent être abordées avec le nouveau SGDOT et dans la stratégie de préservation des Services et accès de données d'OT du Canada.

- En raison de la pratique d'utilisation de métadonnées non standards des producteurs de données d'OT, le CCT ne possède pas présentement de fonction de saisie des données normalisée, il n'existe actuellement pas de cadre global de saisie pour les données d'OT, et de nombreux chargeurs ad hoc ont été créés au fil du temps pour saisir des données à partir de différents types de stations de réception satellite commerciales et canadiennes.
- Le catalogue et le système de gestion de données actuels opèrent dans une architecture vieillissante, il y a plusieurs lignes de code provenant de nombreuses organisations dans un système et l'expertise en matière de segmentation de la prestation de services est limitée.
- Il n'y a pas de politique de planification de la préservation ou de plan de préservation des données publiés en place, mais les Services et accès de données d'OT du Canada sont un début pour ce qui est d'une stratégie de préservation.
- Le système de stockage actuel est limité.
- Il y a trop de catalogues et les métadonnées aux fins d'exploration ne sont pas normalisées conformément à la norme du SCT (c.-à-d. ISO 19115).

- Les rôles et les responsabilités ne sont pas clairs et la voie de mise en œuvre de la politique n'est pas définie. Il y a peu de formation du personnel, les procédures et les politiques ne sont pas précisées, et il y a un manque de documentation.
- Il y a peu de ressources financières disponibles pour soutenir la préservation à long terme des données d'OT. Le SGDOT est financé de façon progressive et un modèle de coûts est en cours d'élaboration. Les coûts jusqu'à et y compris 2015 ont été définis.
- Les PE propres à la mission, ainsi que les ALUF, ne sont pas numérisés ou accessibles, car ils sont sur support papier. Ces fichiers sont conservés par plusieurs dirigeants de mission et se trouvent dans différents catalogues propres à la mission. Les exigences et les restrictions sont respectées dans le système, mais elles ne sont pas documentées et ne sont pas facilement accessibles aux utilisateurs. Les modalités, les privilèges d'accès et les attentes de conservation sont énoncés dans ces documents.
- Certains « processeurs » qui transforment les données brutes d'OT en « produits » ne sont pas sous le contrôle du CCT, donc l'accès aux algorithmes et aux logiciels à des fins de préservation est incertain.
- Les licences ne sont pas forcément interoperables, car elles sont liées à des missions spécifiques.
- Les politiques pour le rapatriement des données de la mission de satellites canadiens des stations internationales de réception de satellite doivent être réexaminées, car elles sont incompatibles. Les PE sont généralement les instruments utilisés entre les nations concernant la collecte et le rapatriement des données propres à la mission.
- Les initiateurs des données - souvent des entités du secteur privé - ne sont pas mandatés pour préserver les données dans leurs banques de renseignements ou les données qu'ils traitent, et ne donnent pas accès à des logiciels et des algorithmes de traitement d'image d'OT.
- Les stratégies de conservation sont propres à la mission, mais le CCT a participé activement à la préservation de toutes les données de ses missions.

La plupart de ces problèmes ont été définis dans le plan de préservation des Services et accès de données d'OT du Canada et doivent être traités avec le nouveau SGDOT. Une fois le SGDOT en place, cependant, une analyse des risques et un plan de préparation aux urgences doivent être élaborés. Également absent du plan des Services et accès de données d'OT du Canada, une évaluation d'un tiers du plan mis en place et du système de préservation. De plus, il n'existe actuellement aucune entente explicite entre Services partagés Canada et le SGDOT. Cela est essentiel puisque le système de préservation est un sous-système géré séparément à l'intérieur de SPC et peut nécessiter un environnement administratif différent.

5.1.4 Leçons retenues

Cette étude de cas décrit essentiellement une stratégie de préservation qui est en cours, un système de préservation qui n'est pas encore créé et une compréhension partagée sur un centre d'archives d'OT existant qui sera bientôt remplacé. Indépendamment, voici les leçons importantes retenues de l'expérience du CCT.

- Les besoins des utilisateurs, le volume de données, et les changements dans la technologie, entre autres, ont motivé la transition vers un nouveau système. Ceci a poussé l'organisation à examiner de près son système existant, à définir ses forces et ses faiblesses et d'y remédier en élaborant une stratégie de conservation et des spécifications pour un système de préservation plus formel.
- Le CCT a créé la stratégie des Services et accès de données d'OT du Canada et les spécifications du SGDOT en adoptant les meilleures pratiques internationales, notamment le *modèle de référence du SOAI*, les *European LTDP Common Guidelines* et la *Liste de contrôle de la VADF*.
- Le CCT a participé à la communauté mondiale d'OT et a développé des relations étroites avec un certain nombre de créateurs de données d'OT, et est maintenant en mesure de tirer parti de ces relations dans les collaborations sur les connaissances et le transfert des compétences pour l'aider à mettre ses archives sur pied et d'éclairer les efforts internationaux en cours.

5.2 Profil : Système d'archivage de l'information géographique de l'Ontario (SAIG)

5.2.1 Introduction

La Direction de l'information géographique du ministère des Richesses naturelles (MRN) de l'Ontario a instauré un système d'archivage de l'information géographique (SAIG) en 2009 afin de préserver les données géospatiales obtenues auprès de l'Entrepôt de données sur les terres de l'Ontario (EDTO). L'EDTO du MRN est utilisé pour diffuser des données géospatiales au personnel du MRN, à la fonction publique de l'Ontario (FPO) et aux clients externes. C'est également le dépôt officiel pour les données géospatiales à la fois créées et utilisées par le MRN comme document de référence pour suivre l'évolution des ressources naturelles au fil du temps, et pour appuyer la prise de décisions, la collecte d'informations, les activités du programme, la législation et les politiques, et la recherche scientifique. Les objectifs du SAIG sont les suivants : préserver les données géospatiales essentiels pour se conformer aux exigences des lois et de la politique, soutenir la stratégie de gestion de l'information du MRN, préserver les données de valeur scientifique et historique pour une analyse à long terme, et mieux défendre la prise de décision à long terme en rapport avec la protection des ressources précieuses.

La Direction de l'information géographique a réalisé une analyse environnementale afin de chercher des politiques et des pratiques d'archivage dans d'autres administrations, de miser sur la recherche actuelle et les succès déterminés, et de mieux définir les défis en matière de préservation du MRN. Le personnel du MRN a déterminé le besoin du Ministère en matière de préservation à long terme des données géospatiales retirées de l'EDTO, dont certaines devaient être conservées à perpétuité pour répondre aux besoins du Ministère.

5.2.2 Modèle opérationnel en cours d'utilisation

Le SAIG est une solution de stockage personnalisé, quasi en ligne, dont les données sont stockées sur une série de disques durs. Deux copies de sauvegarde sont créées et les trois copies sont toutes stockées dans des endroits différents. Il existe deux catégories de documents d'archives dans le SAIG – les instantanés de l'EDTO pris de façon annuelle et les documents d'archives retirés, chacun avec une approche de préservation différente. Les instantanés de toutes les données avec leurs métadonnées, les trousseaux du format d'interface normalisé du Système de données intégrées sur la nature et la géographie de l'Ontario (DINGO) (utilisées pour diffuser des ensembles de données aux utilisateurs commandant des données de l'EDTO) et les rapports techniques stockés dans l'EDTO ont été pris chaque année depuis 2009 et entrés dans le SAIG. Les trousseaux du Système de DINGO contiennent des données d'en-tête, qui fournissent des métadonnées sur le contenu de la trousse, notamment la source des données, le fichier avec l'information de projection des données spatiales, la géométrie spatiale des données, et certains éléments des données non spatiales (attributs) de l'ensemble de données qui sont associés aux objets spatiaux dans la trousse du Système de DINGO. Ces données sont extraites de plusieurs tableaux de bases de données différentes et sont fournies sous forme de fichier texte personnalisé dans la trousse du Système de DINGO.

Conformément au calendrier de conservation MRN-4401-01, un instantané annuel est pris de toutes les données géospatiales stockées ou diffusées par le biais de l'EDTO ainsi que la documentation connexe, comme les métadonnées, les modèles de gestion des données et les bulletins techniques (voir l'exemple de la figure 4). La stratégie de conservation adoptée consiste à garder des instantanés annuels pendant 10 ans, après quoi ils seront conservés une année sur cinq pendant 200 ans. Après 200 ans, les documents d'archives sont transférés aux Archives publiques de l'Ontario.

Figure 4 : capture d'écran de la description d'un instantané préservé de l'EDTO

The screenshot displays the 'Ontario' Metadata Management Tool interface. The top navigation bar includes links for Home, Contact LIO, Links, About, and Help. A search bar is present with a 'Search' button and a 'Reset' button. The main content area is titled 'DISCOVERING ONTARIO DATA' and features a 'No preview available' message. The interface is divided into two main sections: 'WHAT?' and 'WHERE?'. The 'WHAT?' section contains a search bar and a list of datasets. The 'WHERE?' section contains a map and a list of datasets. The 'WHAT?' section is further divided into 'Identification info', 'Cited responsible party', 'Abstract', 'Purpose', 'Status', 'Point of contact', 'Maintenance and update frequency', 'Update scope', 'Descriptive keywords', 'Use limitation', 'Access constraints', 'Other constraints', 'Spatial representation type', 'Language', 'Character set', 'Environment description', and 'Extent'. The 'WHERE?' section is further divided into 'Geology Ontario', 'Land Information Ontario (LIO) Warehouse', 'Maps', 'Municipal Planning Provincial Portal (MP3)', 'Ontario Imagery', 'Weather and Water Information Gateway (WWIG)', 'Web Mapping Service (WMS)', 'x - For future use', 'x - For future use', and 'x - For future use'. The 'Identification info' section includes fields for Title, Date, Date type, and Creation. The 'Cited responsible party' section includes fields for Organization name, Position name, Role, Voice, and Electronic mail address. The 'Abstract' section includes a text description. The 'Purpose' section includes a text description. The 'Status' section includes a text description. The 'Point of contact' section includes fields for Organization name, Position name, Role, Voice, and Electronic mail address. The 'Maintenance and update frequency' section includes a text description. The 'Update scope' section includes a text description. The 'Descriptive keywords' section includes a text description. The 'Use limitation' section includes a text description. The 'Access constraints' section includes a text description. The 'Other constraints' section includes a text description. The 'Spatial representation type' section includes a text description. The 'Language' section includes a text description. The 'Character set' section includes a text description. The 'Environment description' section includes a text description. The 'Extent' section includes a text description.

Ontario Powered by **Land Information Ontario**

Metadata Management Tool > Home | Contact LIO | Links | About | Help | English

Username Password Login

WHAT?

WHERE?

- Any - Search Reset Advanced Options

DISCOVERING ONTARIO DATA

No preview available

Identification info

Title: Geographic Data Snapshots
Date: 2010-01-01
Date type: Creation: Date identifies when the resource was brought into existence

Cited responsible party

Organization name: Ministry of Natural Resource
Position name: Information Access Section
Role: Owner: Party that owns the resource
Voice: 1 (705)7551878
Electronic mail address: lio@ontario.ca

Abstract

The Ministry of Natural Resources houses a number of datasets in the Land Information Ontario Warehouse. On an annual basis these are extracted and preserved for historical reference. This practice began in 2010.

Purpose

The preservation of the data is intended to substantiate decision making.

Status

Planned: Fixed date has been established upon or by which the data will be created or updated

Point of contact

Organization name: Ministry of Natural Resources
Position name: Information Access Section
Role: Point of contact: Party who can be contacted for acquiring knowledge about or acquisition of the resource
Voice: 1 (705)7551878
Electronic mail address: lio@ontario.ca

Maintenance and update frequency

Annually: Data is updated every year

Update scope

Dataset: Information applies to the dataset Ministry of Natural Resources (theme).

Descriptive keywords

Each dataset contained in the snapshot is described in an individual metadata record within this catalogue. The metadata will include any and all use and access limitations pertaining to the data.

Use limitation

Restricted: Withheld from general circulation or disclosure

Access constraints

Restricted: Withheld from general circulation or disclosure

Other constraints

Copyright: Exclusive right to the publication, production, or sale of the rights to a literary, dramatic, musical, or artistic work, or to the use of a commercial print or label, granted by law for a specified period of time to an author, composer, artist, distributor

Spatial representation type

The data requires GIS software and is intended for internal MNR staff.

Language

Vector: Vector data is used to represent geographic data

Character set

English

Environment description

UTF8: 8-bit variable size UCS Transfer Format, based on ISO/IEC 10646

Extent

ArcGIS

Geology Ontario

Land Information Ontario (LIO) Warehouse

Maps

Municipal Planning Provincial Portal (MP3)

Ontario Imagery

Weather and Water Information Gateway (WWIG)

Web Mapping Service (WMS)

x - For future use

x - For future use

x - For future use

GeoRSS

Wetland Unit

MOHLTC Long-Term Care Home Location

Census Federal Electoral District 2006

MOHLTC Hospital Location

Ontario Borehole Database

Miscellaneous Line

WRIP Water Well Locations Phase 3

Water Virtual Flow

Ontario In-filled Climate Data

Borehole

5.2.3 Pratiques exemplaires

En collaboration avec la Direction de l'information géographique, la Section de l'accès à l'information du MRN a préparé un *Guideline for Retiring and Retaining Geospatial Data Stored in the Land Information Ontario Warehouse* (MRNO, 2010). Le but de ce document est de veiller à ce que toutes les données géospatiales déplacées de l'EDTO au SAIG appuient le mandat du MRN. De plus, le plan prévoit des directives sur la façon de retirer une catégorie de documents de l'Entrepôt.

Les *Retiring and Retaining Guidelines* sont très exhaustives, articulent clairement les rôles et les responsabilités, et fournissent une liste de questions à considérer pour chaque catégorie de documents d'archives. Dans ce système de préservation, les rôles comprennent les propriétaires d'information, la Section de l'accès à l'information, le Groupe des ressources foncières et les utilisateurs. Par exemple, seul un propriétaire d'information a le droit de retirer des documents d'archives de l'EDTO, et les données ne peuvent pas être retirées sans que le processus de retrait soit terminé et que l'accord du gestionnaire de la Section de l'accès à l'information ait été obtenu. Au moment de retirer un document, le propriétaire d'information doit tenir compte des communications appropriées, des répercussions qu'il pourrait y avoir sur les utilisateurs de

données, et déterminer si le produit de données est officiel et doit être préservé ou transitoire et peut être supprimé. Une série de questions d'évaluation sont prévues pour aider à déterminer la valeur opérationnelle du document (voir l'encadré).

De plus, des ressources de métadonnées et des outils de formation sont fournis avec les modules de formation, les descriptions des normes adoptées (p. ex., ISO 19115) et les guides de mise en œuvre (MRNO, 2011).

PRATIQUE EXEMPLAIRE

Voici les questions qu'utilise le *Retiring and Retaining Guidelines* du MRNO pour aider les utilisateurs à déterminer le type de document.

- Le document de données géospatiales devant être retiré est-il officiel?
- Les données visent-elles ou reflètent-elles un droit légal du gouvernement?
- Les données seront-elles nécessaires pour défendre le gouvernement contre des accusations de fraude de données ou de déclaration fausse?
- Les données pourraient-elles être utiles à d'autres utilisateurs de données géospatiales ou à la communauté de données géospatiales dans son ensemble?
- Les autres utilisateurs pourront-ils avoir accès aux données?
- Les données géospatiales ont-elles été mises à la disposition d'autres utilisateurs par le biais d'un accord de partage des données ou d'un centre d'échanges?
- Les utilisateurs secondaires peuvent-ils comprendre ou interpréter les données sans expertise technique du créateur?
- Les données sont-elles difficiles ou coûteuses à reproduire?
- Y a-t-il des coûts importants ou des conséquences pour le programme si les données sont perdues?
- Les données peuvent-elles être intégrées de façon utile aux nouvelles données résultant de l'amélioration des méthodes de collecte de données?
- La valeur estimée de recherche des données dépasse-t-elle les coûts pour les maintenir aux fins d'une utilisation secondaire par des chercheurs?
- Les données seront-elles utiles pour analyser la répartition géographique au fil du temps?
- Les données appuient-elles l'étude des changements géophysiques au fil du temps?

5.3 Profil : Gestion des données scientifiques intégrées (GDSI)

5.3.1 Introduction

Au sein de Pêches et Océans Canada (MPO), le Service de Gestion des données scientifiques intégrées est responsable de la gestion et de la préservation des données océaniques physiques, chimiques et biologiques recueillies par le MPO, ou acquises par le biais de programmes nationaux et internationaux menés dans des zones océaniques adjacentes au Canada, et de la diffusion des données, des produits de données et des services à la communauté marine (MPO, 2012). Cette information est utilisée en temps réel pour les prévisions météorologiques, le routage des navires, et la prévision des fenêtres de temps pour effectuer des opérations de pêche hauturière sensibles aux conditions météorologiques. Le programme de gestion des données garantit également que l'information puisse être utilisée dans une variété d'applications nécessitant des données sur une période prolongée, comme des modèles de simulation rétrospective de la climatologie des ondes utilisées dans la navigation maritime dans l'océan, l'étude des changements climatiques et l'ingénierie. L'engagement de la tenue des données s'étend parmi toutes sortes de données scientifiques dont le MPO est responsable, comme en témoigne la *Politique de gestion des données scientifiques* du Ministère, qui comporte plusieurs références à l'exigence en matière d'archivage et de préservation des données (MPO, 2001).



5.3.2 Modèle opérationnel en cours d'utilisation

À titre de Centre national responsable des données océanographiques de bouées dérivantes (CNRDO), le service de GDSI gère depuis 1978 un dépôt continuellement mis à jour de données provenant du monde entier (MPO, 2013). Le CNRDO est un centre de données national venant en aide au Centre mondial de données. Afin de préserver les données de bouées dérivantes, il est déposé auprès du service de GDSI le plus tôt possible après la saisie. Ces fichiers initiaux sont alors remplacés par des versions de qualité supérieure des données dès qu'elles sont disponibles. Le service de GDSI préserve toutes les informations liées aux données, notamment tous les indicateurs de qualité des données et l'information sur l'origine (MPO, 2012).

Les données scientifiques du MPO sont gérées dans le cadre d'un système intégré accessible par le biais des centres de données régionaux, zonaux et nationaux. Le service de GDSI agit comme un centre national qui héberge des données ministérielles et partage les fonctions de préservation, au besoin, avec des centres de données régionales existants. Le service de GDSI assure la coordination entre les centres, le cas échéant, pour veiller à ce que toutes les données soient correctement gérées (MPO, 2001).

Toutes les propositions et tous les plans de projets scientifiques du MPO doivent démontrer l'existence d'un plan exhaustif de gestion des données, qui doit inclure des stratégies et des calendriers pour le transfert des données vers le centre de données responsable. Le budget associé au projet doit clairement indiquer la répartition des ressources pour la gestion des données et la façon dont ces ressources seront utilisées. Les gestionnaires de Sciences et Océans du MPO sont responsables de veiller à ce que les collecteurs de données qui relèvent d'eux soumettent leurs données, ainsi que les données recueillies en vertu d'un marché, au centre de données pertinent en temps opportun. Les données couvertes par cette politique comprennent les ensembles de données identifiés ci-dessous, et toute autre donnée scientifique qui peut être créée ou autrement acquise par le MPO.

- Données sur l'habitat marin et d'eau douce
- Données météorologiques
- Données sur les pêches
- Données d'océanographie biologique
- Données hydrologiques (p. ex., débit des cours d'eau)
- Données sur la région des lacs expérimentaux
- Données biologiques d'eau douce
- Données de chimie marine
- Données sur la santé des poissons
- Données biologiques (tirées de l'échantillonnage des prises, des relevés au chalut et acoustiques, des pêches indicatrices et de relevés de l'industrie, de journaux de bord scientifiques, etc.)
- Données recueillies sur le terrain et obtenues en laboratoire à l'appui du processus d'évaluation des stocks
- Données sur les contaminants
- Données océanographiques physiques
- Données recueillies par le Service hydrographique du Canada, sous réserve des accords du SHC et des pratiques opérationnelles.

5.3.3 Pratiques exemplaires

La politique de gestion des données du MPO garantit que les données sont rapidement copiées dans un environnement « géré » où elles sont correctement copiées et sécurisées contre toute perte accidentelle ou circonstancielle, et où les métadonnées connexes sont intégrées aux données afin de préserver l'utilité à long terme de l'ensemble de données préservées.

Chaque mois, un résumé des données de bouées dérivantes reçues en temps réel est publié sur le Web ainsi que des cartes mondiales et régionales du suivi des bouées dérivantes pour le mois précédent.



On s'attaque aux objectifs et aux intentions du SDI-API à l'aide d'un réseau international volontaire de centres de données et de portails. Au Canada, le défi de préserver les données de l'API a été repris par le Réseau canadien d'information polaire (RCIP), qui a été créé en 2010 avec le financement du Bureau canadien du programme de l'Année polaire internationale, d'Affaires autochtones et Développement du Nord Canada (RCIP, 2012). Voici les partenaires du RCIP.

- Institut canadien de l'information scientifique et technique (ICIST)
- Ministère des Pêches et des Océans, Secteur des sciences (MPO)
- Polar Data Catalogue, Université de Waterloo (UW)
- Scholars Portal, Conseil des bibliothèques universitaires de l'Ontario (CBUO)
- Bibliothèques de l'Université de l'Alberta

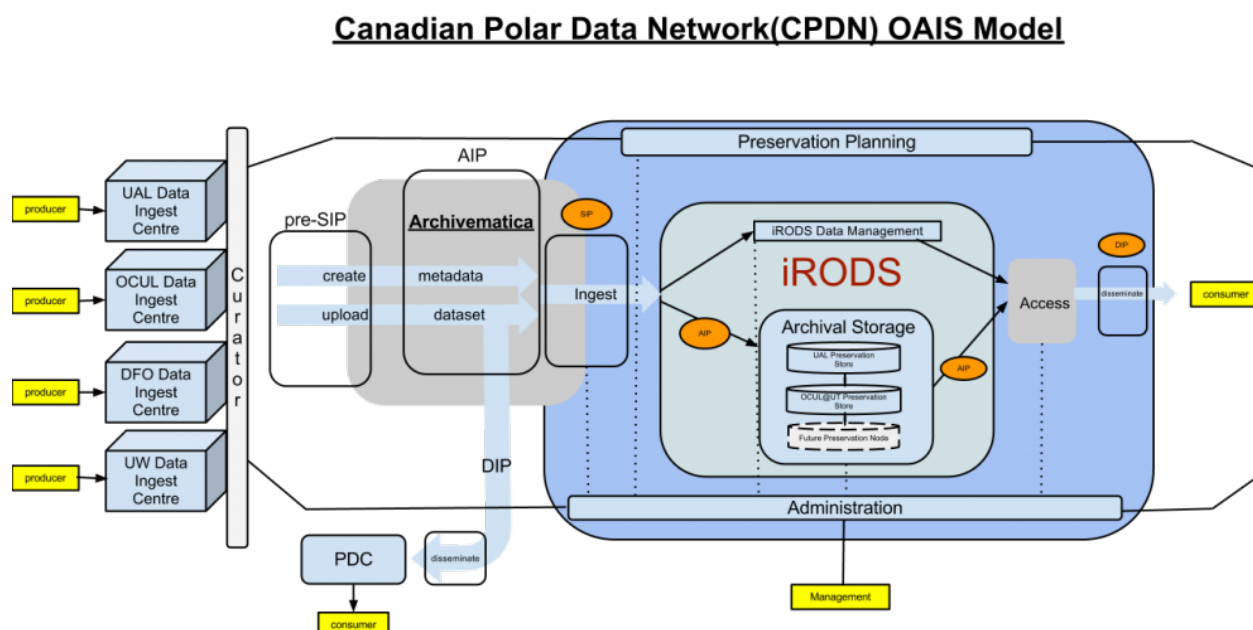
L'objectif de préservation du RCIP est de fournir « un réseau sécurisé hébergeant l'infrastructure nécessaire pour assurer la préservation à long terme des données expérimentales numériques, ... [qui] requiert un système d'archivage d'information qui vérifie en permanence l'intégrité des données et la mise à niveau de nouvelles normes au fil du temps ». Le système d'archivage opérationnel est axé sur le modèle de référence du SOAI, comme illustré à la figure 5.

PRATIQUE EXEMPLAIRE

Politique de gestion des données de l'API
(Parsons et autres, 2011)

...il est essentiel d'assurer la préservation à long terme et l'accès durable aux données de l'API. Toutes les données de l'API doivent être archivées dans leur forme la plus simple et la plus utile et être accompagnées d'une description complète des métadonnées. Un service des données et de l'information de l'API (SDI-API) ([IPYDIS-http://ipydis.org](http://ipydis.org)) doit aider les projets à déterminer les centres de données et les services d'archives appropriés à long terme, mais les projets individuels de l'API sont responsables de faire des arrangements avec des services d'archives à long terme pour garantir la préservation de leurs données. Il faut reconnaître que la préservation des données et l'accès à ces dernières ne doivent pas provenir d'une réflexion après coup et doivent être considérés pendant que les plans de collecte de données sont élaborés.

Figure 5 : Illustration du système d'archives opérationnel du RCIP



Source : Réseau canadien d'information polaire (http://polardatanetwork.ca/?page_id=101)

5.4.3 Défis éprouvés

En 2011, une évaluation a été réalisée sur le rendement de l'API par rapport aux objectifs spécifiques à l'intérieur des éléments de la stratégie de gestion des données (Parsons et autres, 2011). Le tableau 1 résume le rendement lié aux deux objectifs de préservation de la stratégie.

Tableau 1 : Évaluation de rendement par rapport aux objectifs stratégiques de préservation

Objectif stratégique	Rendement
Toutes les données brutes de l'API doivent être préservées et bien surveillées dans des services d'archives à long terme conformément à la norme ISO du modèle de référence du Système ouvert d'archivage d'information.	De nombreuses disciplines n'avaient pas de services d'archives à long terme. Les normes archivistiques à long terme sont encore en évolution et le respect des pratiques exemplaires est très variable selon les projets et les disciplines. Au-delà de l'engagement continu du gouvernement dans certaines disciplines, aucun modèle d'affaires clair et durable n'a émergé pour soutenir l'administration des données à long terme.
Les données doivent être accompagnées d'une documentation complète pour permettre la préservation et l'administration.	La plupart des documents sont <i>ad hoc</i> et largement orientés vers l'exploration. Certaines lignes directrices concernant la documentation ont été élaborées sur une base disciplinaire ou d'un projet, mais certaines questions, telles que la description détaillée et continue de la provenance, n'ont pas été résolues dans la communauté d'archivage en général.

Au moment de l'évaluation, il est apparu que seulement 30 des 124 projets de l'API avaient suffisamment pris en compte la préservation à long terme. Étant donné que plus de 75 % des projets de l'API ont recueilli des données sans plans ou ressources d'archivage clair, il a été difficile de déterminer toutes les données de l'API recueillies, sans parler de s'assurer qu'elles aboutissent dans des archives sécurisées (p. ex., l'obtention des métadonnées demeure un grand défi). Plusieurs chercheurs de l'API n'étaient pas clairs quant à leurs responsabilités en matière de préservation de données ou l'endroit où ils doivent soumettre leurs données. Dans de nombreuses disciplines, des services d'archives à long terme n'existent pas, et il n'y a pas de stratégie détaillée de préservation des données répandue entre les disciplines et les nations. Les causes principales de ces défis sont les suivantes :

- les difficultés que la plupart des scientifiques rencontrent pour trouver le temps de préparer les données pour la préservation;
- le manque de ressources durables pour les centres de données afin de préserver les données de l'API et garantir la coordination entre ces centres.

5.4.4 Leçons retenues

Voici comment les principales leçons retenues concernant l'archivage et la préservation des données de l'API peuvent être résumées (Parsons et autres, 2011).

- Bien que de nombreux gestionnaires de projets de l'API aient supposé que le Centre mondial de données (CMD) du CIUS serait le foyer naturel pour une grande partie de données de l'API, le CMD dans son ensemble n'a pas joué un rôle central ou de premier plan dans la gestion des données de l'API. Il est à espérer que le nouveau Système mondial de données du CIUS⁵ permettra de mieux servir la science polaire à long terme par la croissance d'un véritable réseau de données.
- Les scientifiques ont besoin d'incitations pour échanger et décrire leurs données et afin d'adhérer aux stratégies et aux politiques de données pertinentes. Les bailleurs de fonds doivent permettre la préservation des données comme une dépense acceptable lors de l'octroi des fonds, ainsi que les suggestions d'application ci-dessus.
- La préservation des données à long terme doit être un facteur tout au long du processus scientifique, ce qui nécessite un changement majeur dans certaines des institutions de science. Par exemple, les universités doivent inclure les instructions de gestion des données comme une exigence fondamentale de diplômes d'études supérieures. Ils doivent envisager la publication et l'administration des données à égalité avec leur publication dans la collation des grades, la promotion et la durée des fonctions.
- De nouveaux modèles économiques sont nécessaires afin de pouvoir apporter un soutien durable pour la préservation dynamique et évolutive des données scientifiques. L'expérience de l'API suggère que la préservation des données est plus efficace quand les

⁵ Le SMD s'efforce d'assurer l'administration et la fourniture à long terme de données de qualité évaluées et de services de données à la communauté scientifique internationale.

nations engagent des ressources du programme dans la gestion et la coordination des données et fournissent des dépôts explicites aux fins de la préservation.

- Les nations doivent financer les services d'archives pour combler les lacunes disciplinaires et exiger des services d'archives qu'ils collaborent entre eux sur les normes et l'interopérabilité comme une éventualité de leur financement de la recherche interdisciplinaire.

6. Création d'un système de préservation des données géospatiales

6.1 Introduction

Ce chapitre fournit aux organisations qui envisagent l'élaboration d'un système de préservation numérique des données géospatiales des directives sur la planification et la mise en œuvre d'un tel système, dont les étapes sont illustrées dans la figure 6. Les *Lignes directrices à l'intention des préservateurs* du projet InterPARES 2 ont inspiré la structure logique du chapitre, et certains aspects du *modèle de référence du SOAI*, des *European LTDP Common Guidelines* et des *Lignes directrices à l'intention des créateurs* du projet InterPARES 2 y figurent également. Il est fait référence à la fonctionnalité, aux processus et à la documentation de la *Liste de contrôle de la VADF* qui sont nécessaires pour qu'un système de préservation soit accrédité comme un dépôt numérique fiable (DNF). Bien qu'il y ait chevauchement entre les références, chacun traite l'élaboration de systèmes de préservation d'une perspective différente, et les lecteurs sont invités à choisir les idées qui correspondent le mieux à leurs conditions de préservation. De plus, le lecteur sera dirigé vers des outils et des ressources utiles et des pratiques exemplaires issues de l'étude de cas et des profils. Bien que l'information fournie soit loin d'être exhaustive, elle vise à familiariser le lecteur avec la portée d'ensemble de la création du système de préservation.

Figure 6 : Le processus de mise sur pied d'un système de préservation des données géospatiales



6.2 Déterminer la portée et les objectifs du système

La planification de toute initiative de préservation des documents d'archives commence par la portée. L'exercice de détermination de la portée prend le plus souvent la forme d'un plan d'activités, et comprend ce qui suit : déterminer les facteurs opérationnels pour le système; élaborer les objectifs de préservation programmatique qui s'articulent avec la législation, les règlements, les politiques et les directives; déterminer le niveau minimum et le type de réutilisation que les archives devront maintenir pour la communauté d'utilisateurs; réaliser un inventaire de la technologie et évaluer sa pertinence; définir les normes; établir un budget et créer un processus d'évaluation. Le

Geoarchiving Business Planning Toolkit de GeoMAPP, qui comprend le *Geoarchiving Business Planning Guidebook*, un *Geoarchiving Self-Assessment Template*, et un *System Inventory Template*, sont des outils précieux pour cette partie du processus de planification.

Pratiques exemplaires

Le modèle de référence du SOAI comprend la détermination de la portée comme fonction de gestion et suggère que la portée détermine l'étendue du créateur et des groupes de consommateurs qui sont desservis par les archives.

Le rapport du LTDP *EO Preserved Data Set Content* donne un aperçu des principes de préservation de base et quelques éléments à prendre en considération lors de la détermination de la portée pour la préservation des données de la mission, comme le temps de conservation, les horaires d'accès, ce qui doit être préservé et comment le faire.

Les *European LTDP Common Guidelines* fournissent des recommandations en matière de dotation, mais soulignent que les activités sont régies par une structure organisationnelle qui supervise la planification et les activités.

La *Liste de contrôle de la VADF* fournit des critères pour la responsabilité administrative, la viabilité organisationnelle et la responsabilisation de la procédure. Par exemple, on s'attend à ce que les archives aient ce qui suit : un énoncé de mission qui démontre son engagement à la préservation à long terme; du personnel désigné avec les compétences et la formation nécessaires pour le développement continu; des rôles, des compétences et des descriptions de poste clairement définis; un organigramme; et un plan de développement continu.

Principaux points

- L'élaboration d'un plan d'activités vous aidera à définir la portée.
- Une considération clé est de déterminer la configuration optimale du système.
- Évaluer quelles données doivent être archivées et évaluer les effectifs, les conditions administratives, les logiciels, les réseaux, le stockage et les normes actuels.

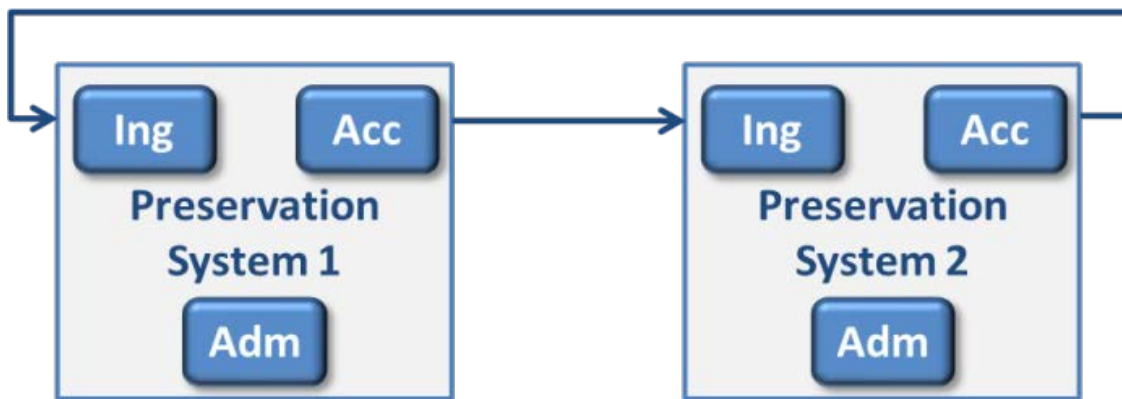
Les objectifs du système de préservation doivent tenir compte des exigences législatives, des politiques et des directives applicables, des besoins opérationnels et des exigences de la communauté désignée. Les objectifs doivent décrire de manière concise le mandat du système, la manière dont il sera atteint, et ce que seront le niveau de service garanti et le mode de participation soutenu par le système de préservation. Un objectif clé sera de veiller à la préservation des métadonnées et des produits de donnée de qualité reconnue à tous les niveaux requis par les utilisateurs, ou la capacité à les générer sur demande grâce au traitement approprié.

L'exercice d'évaluation doit aussi envisager la collaboration entre les organisations, car cela va permettre des arrangements et des processus organisationnels, techniques et opérationnels. Le *modèle de référence du SOAI* décrit les modèles suivants, les trois premiers ayant successivement des degrés supérieurs d'interaction.

Locaux ou thématiques, les **systèmes indépendants de préservation** peuvent être regroupés en un seul endroit ou répartis entre de nombreux sites. Ils peuvent choisir de concevoir des TDI et les mécanismes exploratoires axés sur des normes formelles ou de facto, ce qui pourrait faciliter la coopération volontaire avec d'autres systèmes de préservation qui mettent en œuvre les mêmes normes.

Les **systèmes de préservation coopérants** reposent sur des ententes conclues entre deux services d'archives ou plus. La forme la plus simple de collaboration, c'est quand un système agit en tant que consommateur de matériel d'un autre système; dans un tel cas, il doit accepter les formats des TDI et des TIS du système de production. La collaboration ne se fonde sur aucune norme en matière d'accès commun, de soumission ou de diffusion, mais les groupes qui collaborent doivent soutenir au moins un format commun de TIS ou de TDI pour permettre les requêtes entre les systèmes. La figure 7 illustre le concept de collaboration des systèmes de préservation où les entités de saisie et d'accès partagent les formats TIS et TDI.

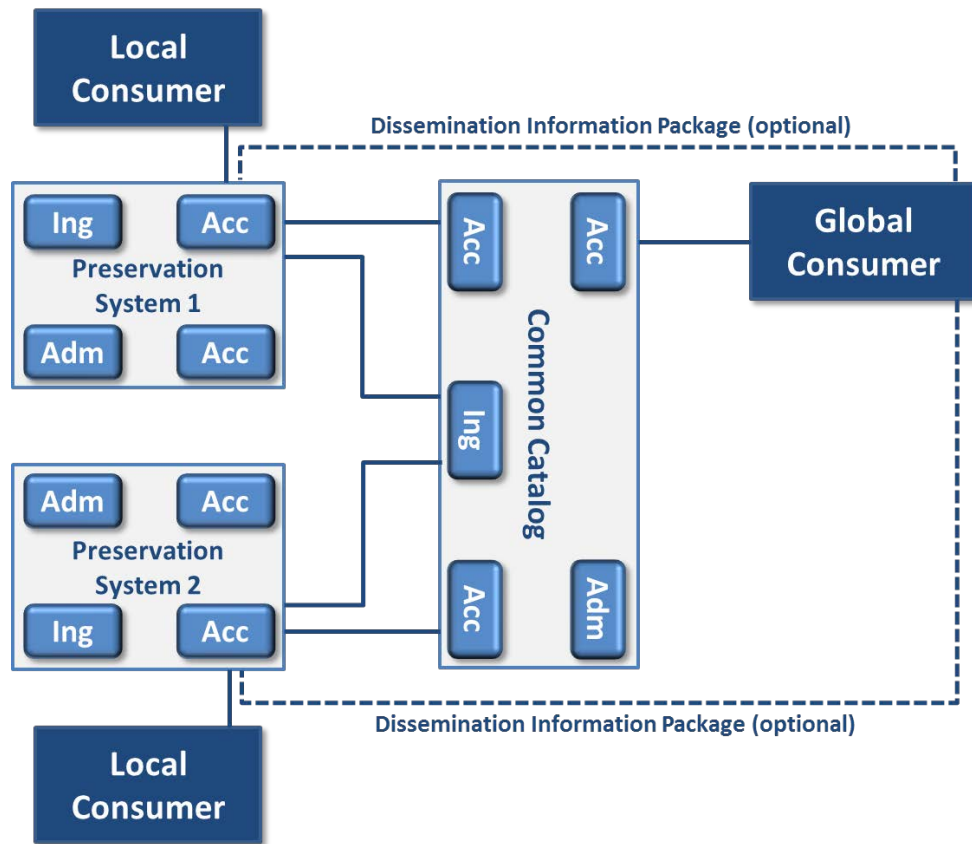
Figure 7 : Systèmes de préservation coopérants avec des ententes de partage mutuel



Source : CCSDS (2012)

Les **systèmes fédérés de préservation** sont conceptuellement axés sur le consommateur, et servent à la fois une communauté locale (c.-à-d., la communauté désignée d'origine) et une communauté mondiale (c.-à-d. une communauté désignée élargie) qui détient des participations dans les banques de plusieurs systèmes de préservation. La communauté mondiale peut influencer les archives à donner accès à leurs banques de renseignements à l'aide d'un ou de plusieurs instruments de recherche de données communes. La figure 8 illustre l'architecture fonctionnelle de deux systèmes de préservation qui ont des communautés désignées similaires et ont décidé de se fédérer pour permettre aux consommateurs de repérer les TIA intéressantes de chacun en une seule session de recherche.

Figure 8 : Systèmes fédérés de préservation utilisant un catalogue commun

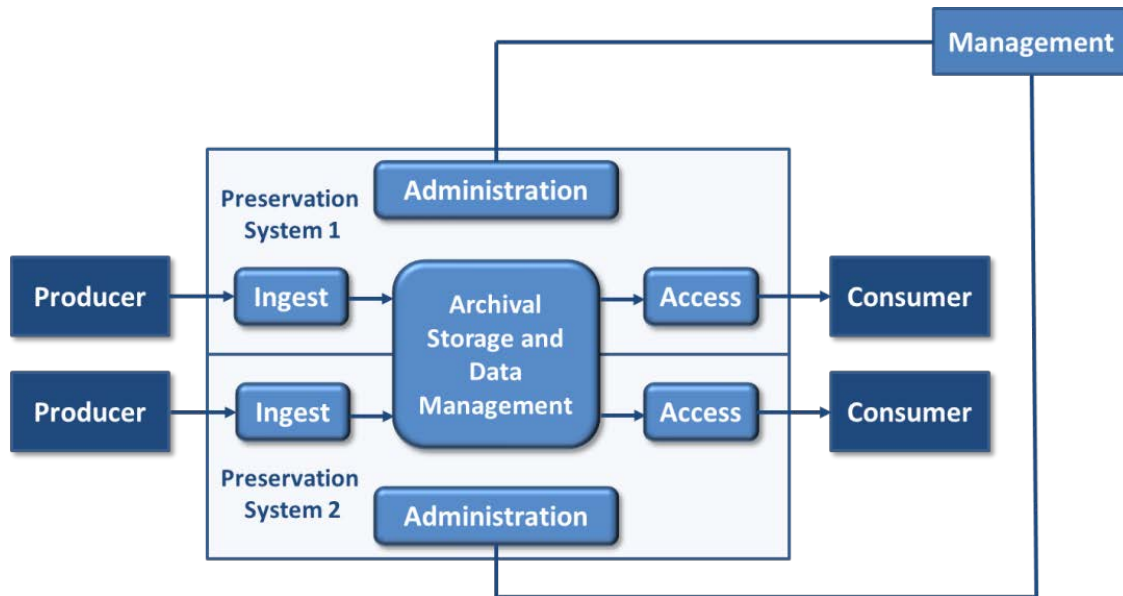


Source : CCSDS (2012)

Le catalogue commun est l'élément externe (mondial) de liaison qui sert de point d'accès commun à l'information dans les deux systèmes. Les TDI contenant les instruments de recherche de chaque système de préservation sont saisies dans le catalogue commun, ce qui peut limiter ses activités en tant qu'instrument de recherche ou inclure la diffusion en commun de produits à partir de l'un ou l'autre ou les deux, comme indiqué par les pointillés sur la figure.

Les **systèmes de préservation à ressources partagées** sont plus intégrants, au moyen desquels la gestion de plusieurs systèmes a conclu des accords afin de partager ou d'intégrer des secteurs fonctionnels. Cette association est fondamentalement différente des exemples précédents, en ce que l'architecture interne du système de préservation doit être prise en compte. La figure 9 illustre le partage d'une fonction de stockage commune, qui consiste en une entité d'archivage et une entité de gestion de données, entre deux systèmes. Chaque système peut servir des communautés totalement indépendantes, mais afin que l'élément de stockage commun fonctionne, des normes sont nécessaires au niveau des interfaces de saisie de stockage et d'accès de stockage internes.

Figure 9 : Systèmes de préservation à ressources partagées



Source : CCSDS (2012)

La situation actuelle de la préservation des données d'OT en Europe, par exemple, est fragmentée avec les systèmes centralisés distribués et indépendants, au moyen de différentes technologies et différentes TIA et TDI. L'un des objectifs des *European LTDP Common Guidelines* (et les documents LTDP/PDSC connexes) est de permettre une plus grande normalisation afin de faciliter la collaboration entre les détenteurs de données d'OT européens et canadiens et leurs systèmes.

6.3 Définir la communauté des utilisateurs du système

Déterminer qui sont les utilisateurs, les consommateurs ou les communautés désignées d'un système de préservation permettra de définir la portée, les objectifs et la configuration du système, du type de données que le système saisira, et du type de fonctionnalité d'accès nécessaire pour répondre aux besoins des utilisateurs. La communauté d'utilisateurs dans son ensemble peut être composée de différentes communautés désignées, qui ont généralement des compétences, des bases de connaissances et de ressources différentes. Les communautés désignées peuvent inclure des scientifiques, des chercheurs, des entreprises, des revendeurs à valeur ajoutée, et le public et peuvent encore être différenciées sur la base des domaines d'application et des zones d'intérêt respectifs (p. ex., l'océan, l'atmosphère, l'infrastructure, la gestion des terres, etc.).

De plus, les nouveaux utilisateurs ayant des objectifs différents pour l'utilisation des données et des compétences et des bases de compétences complètement différentes que celles définies aujourd'hui peuvent vouloir accéder au système de préservation dans le futur. Par conséquent, la définition d'une « communauté désignée » doit être générique et assez vaste afin que le contenu déterminé à être préservé à long terme pour cette communauté permette à d'autres utilisateurs, qui n'ont pas été pris en compte au moment où la préservation a été lancée, d'utiliser les données dans le futur.

C'est une bonne pratique que de définir la communauté des utilisateurs et les objectifs de préservation ensemble pour contenir de façon significative la portée de ce qui doit être préservé à long terme. C'est également une bonne pratique de faire participer les utilisateurs au processus de planification, puisque leur participation permettra d'obtenir l'appui du processus en lui-même et contribuera au plan d'activités qui répond le mieux à leurs besoins. Le programme des Services et accès de données d'OT du Canada du CCT, par exemple, a réalisé une enquête pour obtenir une meilleure compréhension de ses utilisateurs de données d'OT, les résultats éclairant la fonctionnalité d'accès. Le North Carolina Geospatial Data Archiving Project a également réalisé une enquête auprès des organismes de créateurs locaux pour déterminer la fréquence de la saisie des données géospatiales et les pratiques de leur archives (NC CGIA, 2006).

Principaux points

- Les besoins de toutes les communautés d'utilisateurs actuelles désignés doivent être déterminés
- La définition doit être suffisamment large pour répondre aux différents besoins de nouveaux utilisateurs futurs
- La mobilisation des utilisateurs dans la planification du système augmentera les chances de succès.

Dans la *Liste de contrôle de la VADF*, les besoins des utilisateurs définissent comment la documentation est décrite et comprise par cette communauté, et des définitions opérationnelles claires doivent être jointes aux besoins des utilisateurs et à la façon dont les données seront comprises. Alors que les stratégies et les normes de documentation doivent satisfaire à des exigences professionnelles, elles doivent aussi être pertinentes pour les communautés désignées. De plus, les technologies de système de préservation doivent être adaptées aux services fournis aux utilisateurs. Enfin, le thème des *European LTDP Common Guidelines* pour l'accès et l'interopérabilité recommande que les utilisateurs puissent accéder facilement aux données d'OT dans des formats standards et que les logiciels et les services soient interopérables pour garantir un accès homogène aux données d'OT accessibles à partir de systèmes hétérogènes, et ce, à long terme.

6.4 Obtenir et gérer les ressources

D'importantes ressources sont nécessaires pour financer les capacités technologiques et les ressources humaines nécessaires pour faire fonctionner un système de préservation des données géospatiales. Les organisations peuvent acquérir de nouvelles ressources ou réaffecter les ressources existantes ou tirer parti des relations de collaboration avec les autres. Certains des modèles collaboratifs distribués abordés dans la section 6.2 représentent des exemples de mobilisation des ressources. Les ressources doivent également être durables et un plan de communication ferme qui accompagne le plan d'activités est utile afin de convaincre les sources de financement, puisque c'est une stratégie de financement flexible. Le programme des Services et accès de données d'OT du Canada du CCT, par exemple, développe progressivement son infrastructure selon la disponibilité des fonds.

PRATIQUE EXEMPLAIRE

Voici ce que comprend la preuve de pratiques de gestion financière acceptables en vertu de la *Liste de contrôle de la VADF*.

- Des plans opérationnels, des budgets, des rapports financiers et des vérifications sont en place.
- Les plans d'activités sont examinés tous les ans.
- Les pratiques et les procédures financières sont transparentes et vérifiées par des tiers.
- Les risques, les investissements et les dépenses sont régulièrement analysés et présentés.
- Un engagement existe pour combler les déficits de financement.

La viabilité financière est une préoccupation importante dans la *Liste de contrôle de la VADF*, qui recommande que les archives élaborent un plan d'activités viable qui inclut les incidences financières liées à l'élaboration du système de préservation et des activités de production, et indique le niveau de soutien financier des organismes qui contribuent, des souscripteurs et des autres parties. De plus, il doit comprendre comment les coûts futurs de la migration, les améliorations apportées aux immobilisations et les améliorations seront couverts. Des plans d'urgence doivent aussi être élaborés pour faire face aux défaillances majeures et, si le système de préservation est financé par le gouvernement, pour le mettre à l'abri des incertitudes politiques. L'outil

GeoArchiving Comprehensive Cost-Benefit Analysis Guidance élaboré dans le cadre de la *GeoArchiving Business Planning Toolkit* de GeoMAPP peut aider à planifier l'acquisition de processus de ressources (GeoMAPP, 2011).

6.5 Planifier la préservation

La prochaine étape dans l'établissement d'un système de préservation prévoit le processus de préservation des documents d'archives. Les activités de préservation comprennent toujours un mélange de recherche sur les technologies en évolution et l'élaboration de nouvelles stratégies de préservation, ainsi que les activités quotidiennes nécessaires au maintien des banques de renseignements existantes. Chaque cadre divise ces deux types de responsabilités de façon un peu différente.

Pour inspirer confiance, un système de préservation doit reposer sur des politiques, pratiques et procédures fermes et documentées. Dans la *Liste de contrôle de la VADF* (qui comprend un guide pour l'élaboration de la planification et des *stratégies de planification de la préservation*), cette activité consiste à déterminer les mesures les plus appropriées pour garantir la préservation et la facilité d'utilisation future des documents d'archives dans un système de préservation. La planification de la préservation est une entité du *modèle de référence du SOAI* qui fournit les services et les fonctions de surveillance de l'environnement du système de préservation. Le résultat de cette activité sera un document « d'approche et de stratégie de préservation » décrivant toutes les TIA (ou dans le cas des *European LTDP Common Guidelines*, les éléments de l'inventaire du PDSC), les éléments dont ils dépendent ou qui sont nécessaires pour les comprendre et les utiliser, les mesures de préservation connexes déterminées pour chacun d'eux, et l'état de préservation de chaque élément.

Dans le *modèle de référence du SOAI*, les politiques et les procédures sont considérées comme des entrées dans un système de préservation et la planification de la préservation comprend les fonctions suivantes.

- *Surveillance de la communauté désignée* - Cela implique le suivi des changements dans les exigences en matière de service des consommateurs et des créateurs et des technologies de produits disponibles (p. ex., les formats de données, les choix de supports, les préférences pour les progiciels, les nouvelles plateformes informatiques, ou les mécanismes pour communiquer avec les archives). Elle fournit des rapports, des exigences en matière d'alertes et de nouvelles normes pour la fonction *Élaboration de normes et de stratégies de préservation* et envoie des exigences de préservation à la fonction *Élaboration de modèles de trousse et de plans de migration*.
- *Surveillance de la technologie* - Cela inclut le suivi des nouvelles technologies numériques, des normes d'information et des plateformes informatiques pour déterminer les technologies qui pourraient entraîner l'obsolescence dans l'environnement informatique des archives et empêcher l'accès à certaines des banques de renseignements du système. Elle envoie des rapports, des normes de données externes, des résultats sur le prototype et des alertes technologiques à la fonction *Élaboration de normes et de stratégies de préservation* et envoie les résultats du prototype à la fonction *Élaboration de modèles de trousse et de plans de migration*.

- *Élaboration de normes et de stratégies de préservation* - Cela implique l'élaboration et la recommandation de stratégies et de normes, et l'évaluation des risques, pour permettre au système de préservation de faire des choix éclairés. Selon les informations reçues des fonctions *Surveillance de la communauté désignée* et *Surveillance de la technologie* et de l'administration, cette fonction détermine les changements qui exigeraient la migration de nouvelles soumissions ou de certaines banques de renseignements du système de préservation actuelles (p. ex, la mise à jour des TIA avec de l'information représentative supplémentaire ou révisée). Les rapports d'analyse périodiques des risques et les recommandations sur l'évolution du système et les mises à jour des TIA sont envoyés à l'administration.
- *Élaboration de modèles de trousse et de plans de migration* - De nouvelles conceptions de trousse d'information et de plans et prototypes de migration sont élaborés pour mettre en œuvre les politiques et les directives de l'administration, et des directives sont fournies sur leur application aux banques de renseignements et aux soumissions spécifiques des archives. Cette fonction applique les normes sur le format, les métadonnées et la documentation aux exigences de préservation et propose des conceptions de modèles de TIA et de TIS à l'administration. Les objectifs de migration peuvent nécessiter des transformations de l'information sur le contenu pour éviter la perte d'accès en raison de l'obsolescence technologique. Ce travail peut comprendre l'élaboration de nouvelles conceptions de TIA, de logiciels prototypes, de plans d'analyse, de plans d'examen de la communauté et de plans de mise en œuvre pour l'application progressive dans les nouvelles TIA, nécessitant une consultation auprès des autres secteurs fonctionnels du système de préservation et de la communauté désignée. La planification de la préservation élabore, valide et fournit les trousse de migration et les horaires de l'administration et effectue la migration.

La *Liste de contrôle de la VADF* contient les exigences suivantes pour les systèmes de planification de la préservation, qui doivent être articulées dans les politiques et les procédures documentées des archives.

- Prendre des décisions pertinentes sur les formats de fichiers.
- Avoir un déroulement de travail automatisé ou manuel complet pour verser les objets numériques, notamment des protocoles de transfert, des rôles et responsabilités clairs des créateurs et des préservateurs, et une preuve explicite de la conversion de fichiers qui se produit lorsque les TIA sont générées à partir des TIS.
- Prévoir ou appliquer des mesures de préservation relatives aux TIA (p. ex., analyser et appliquer des plans de préservation, créer des registres des mesures, etc.).

Principaux points

- Les plans de préservation doivent garantir la préservation et l'utilité future des documents contenus dans le système.
- Les plans de préservation doivent être bien documentés dans les politiques et les procédures du système.
- Les répercussions du changement technologique sont particulièrement importantes et doivent être surveillées.

- Accéder aux politiques, aux procédures et aux pratiques de stockage du système de préservation afin de garantir l'utilisation efficace de stockage fiable et être adapté à l'évolution technologique.
- Avoir un moyen indépendant de vérifier le contenu du système de préservation basé sur les objets numériques traçables sécurisés (p. ex., un registre d'acquisition vérifiable et un inventaire qui ne peut être modifié)⁶.

Les archives doivent avoir des veilles technologiques formelles en place ainsi que des mécanismes de surveillance et de notification lorsque les métadonnées et les formats approchent l'obsolescence, ou des modifications logicielles et matérielles sont nécessaires, et être en mesure de changer les plans en raison de ses activités de surveillance. Une procédure de veille technologique, des inventaires du matériel et les profils des communautés désignées sont la preuve que ces exigences sont satisfaites. Le centre d'archives doit aussi démontrer que son plan de préservation est efficace et que ceci peut être fait en demandant aux utilisateurs de tester les systèmes au fil du temps.

De plus, un centre d'archives doit démontrer qu'il dispose d'un système d'exploitation bien soutenu et d'autres logiciels d'infrastructure de base, que son système de sauvegarde est suffisant, et que des copies des documents d'archives sont gérées et synchronisées. Aussi, la perte de corruption binaire doit être détectable et toutes les erreurs ou tous les incidents doivent être signalés à l'administration des systèmes. Enfin, les processus de gestion du changement doivent être documentés, les archives doivent avoir une procédure pour analyser les effets des changements et réagir aux mises à jour, et un processus d'analyse des risques doit être mis en place.

Les *Lignes directrices à l'intention des préservateurs* du projet InterPARES 2 comprennent une liste de stratégies de préservation (InterPARES 2, 2007) et la documentation de vérification et d'accréditation disponible sur le site Web Scholars Portal est une autre ressource précieuse.

⁶ La création d'identificateurs d'objets numériques (DOI) pour les ensembles de données est l'une des stratégies pour créer des identificateurs persistants. Le Conseil national de recherches du Canada a lancé *DataCite Canada* en 2012 pour aider les créateurs de données avec ce processus et c'est également l'agent d'allocation de DOI du Canada. Cela fait partie d'une initiative internationale plus vaste des bibliothèques à l'échelle nationale et des organismes de recherche pour rendre les données plus accessibles sur Internet (RNC, 2010).

6.6 Élaborer des politiques et des procédures

Dans le cadre du processus de planification de la préservation, les préservateurs de données doivent élaborer des politiques et des procédures (c.-à-d. des instructions et des règles explicites pour guider les décisions et les actions) pour contrôler le transfert, la maintenance et la reproduction des documents d'archives. Les *Lignes directrices à l'intention des préservateurs* suggèrent que des procédures soient conçues pour satisfaire aux *trois conditions de base* suivantes (InterPARES 2, 2002) pour produire des documents d'archives authentiques et s'assurer de leur identité et de leur intégrité (InterPARES 2, 2007).

- Contrôle sur le transfert, la maintenance et la reproduction de documents, selon lesquels le système et les procédures s'assurent que les transferts entre le créateur et le préservateur disposent de contrôles adéquats pour garantir l'intégrité et l'identité des documents d'archives.
- Documentation de l'activité de reproduction des documents d'archives, notamment : la date de la reproduction; la relation entre les documents d'archives acquis et les copies réalisées par le préservateur; les répercussions du processus de reproduction sur la forme, le contenu, l'accessibilité et l'utilisation; et l'enregistrement et la communication du fait que les éléments d'un document ne sont pas totalement reproductibles.
- Documentation sur les modifications apportées au document au fil du temps.

Les *Lignes directrices à l'intention des préservateurs* recommandent également que les mesures de sécurité soient clairement définies dans les politiques et les procédures et que ces mesures abordent des méthodes pour s'assurer que la chaîne de possession d'un document est maintenue, que les contrôles et les procédures de sécurité soient documentés et respectés, et que les documents d'archives restent inchangés. Les procédures doivent également décrire clairement comment les créateurs transfèrent leurs données au préservateur et comment les préservateurs décrivent l'acquisition (transfert - saisie), la transformation, la maintenance et la disposition d'accès aux documents d'archives de données géospatiales.

Selon le *modèle de référence du SOAI*, la responsabilité d'élaborer des politiques et des procédures concernant les éléments suivants s'inscrit dans un système de préservation :

- négocier et accepter des renseignements adéquats de la part des créateurs;
- obtenir le contrôle de ces renseignements selon le niveau nécessaire à leur préservation;
- déterminer les communautés désignées;
- veiller à ce que les documents d'archives préservés puissent être compris par la communauté désignée;
- protéger les documents d'archives contre les atteintes à la sécurité et les récupérer à la suite de catastrophes;
- veiller à ce que les documents d'archives soient protégés contre toutes les éventualités raisonnables, y compris la dissolution des archives;
- rendre les documents d'archives accessibles aux communautés désignées.

Pour être accrédité comme un DNF, les archives doivent clairement documenter les exigences, les décisions, les développements et les mesures pour assurer un accès à long terme aux documents d'archives qui lui sont confiés. Cette documentation garantit aux utilisateurs, aux créateurs et à la gestion que le système de préservation répond à ses exigences. L'accréditation de DNF est considérée comme l'indicateur le plus clair qu'une bonne pratique fondée sur des normes est facilitée par la responsabilisation de la procédure. La *Liste de contrôle de la VADF* comprend des éléments tels que des politiques écrites, des procédures, des protocoles, des règles, des manuels, des guides et des déroulements de travail, des cycles de révision de la spécification et des mécanismes de mise à jour et d'examen. De plus, le système de préservation doit répondre aux besoins de sa communauté désignée, et maintenir des politiques écrites concernant les

autorisations juridiques pour préserver le contenu numérique au fil du temps (p. ex., pour empêcher le contournement de la GDN). Certains des documents d'archives relatifs à ce qui précède sont des ententes de dépôt, des lois, des politiques et des ententes de service. Les processus de rétroaction de l'utilisateur et la documentation sur l'historique de l'évolution des activités et des procédures sont également nécessaires. Les politiques doivent inclure un engagement à la transparence, par lequel la documentation sur les activités et la gestion des archives est accessible à tous les intervenants.

Pratiques exemplaires

Le *Best Practices for Geospatial Data Transfer for Digital Preservation* de GeoMAPP fournit une approche étape par étape au processus de transfert de données et inclut des modèles et une liste de contrôle (GeoMAPP, 2011).

Le Scholars Portal a rendu disponibles ses politiques et ses plans décrivant son déroulement de travail et son mandat de préservation, qui sont des modèles très utiles pour ceux qui élaborent des systèmes de préservation (CBUO, 2012).

Le *EO Data Set Content de la LTDP* (LTDP, 2012) fournit une liste de contrôle des procédures pour la préservation des données d'OT qui représente une excellente ressource pour les préservateurs de données.

6.7 Évaluer la valeur de préservation des documents d'archives

La dernière étape de planification dans la mise sur pied d'un système de préservation des données géospatiales évalue les documents d'archives qui ont été déterminés pour la préservation potentielle afin de définir s'il y a suffisamment de valeur à leur préservation. L'évaluation est fondée sur des facteurs tels que les objectifs des archives, son mandat législatif, les besoins et les exigences de sa communauté d'utilisateurs, les besoins opérationnels des organisations d'accueil et la valeur intrinsèque des données pour les utilisateurs. Le processus d'évaluation comprend l'examen de la provenance et du contenu d'un document, sa fiabilité et son état et les coûts pour le préserver. Obtenir l'appui des bibliothécaires de l'institution et des archivistes du gouvernement, si possible, peut aider à la création d'un processus d'évaluation. Comme indiqué précédemment, ils devraient également faire partie de la conception du système de préservation. L'*Introduction to Appraisal Mentoring* de GeoMAPP renferme un résumé du processus d'évaluation.

Principaux points

- L'évaluation consiste à évaluer le contenu d'un document, l'authenticité, la provenance et la fiabilité et le coût de sa préservation.
- L'évaluation en collaboration avec les créateurs de documents et les préservateurs est bénéfique.

Les *Lignes directrices à l'intention des préservateurs* du projet InterPARES 2 recommandent l'élaboration de méthodes de transfert et la détermination de stratégies de préservation et de méthodes d'évaluation en collaboration avec les créateurs des documents d'archives. Cela peut inclure de repérer plusieurs propriétaires de données géospatiales, et de comprendre les droits de propriété intellectuelle, les marchés et les ressources. Le processus d'évaluation consiste également à évaluer l'authenticité des documents d'archives, ce qui nécessite un examen minutieux de la chaîne de garde et des pratiques en matière de la gestion de documents d'archives. Un rapport d'évaluation doit documenter les contrôles en place pour garantir l'identité et l'intégrité d'un document d'archives et répondre aux *Conditions de référence requises pour établir la présomption d'authenticité des documents d'archives*

électroniques. Les *Lignes directrices à l'intention des créateurs* fournissent des indications sur la façon de gérer les documents d'archives et évaluer ceux ayant une valeur opérationnelle qui pourraient être transférés aux archives à une date ultérieure.

6.8 Acquérir et saisir les documents d'archives

Une fois que les documents d'archives ont été évalués, la prochaine étape consiste à l'élaboration, par les créateurs et les préservateurs, d'un plan commun pour le transfert des documents d'archives et d'entamer le processus d'acquisition des documents d'archives sélectionnés. Le travail sur le plan peut être commencé tout en évaluant la faisabilité technique pendant le processus d'évaluation. Les procédures d'acquisition doivent être normalisées et appliquées, ce qui, selon les *Lignes directrices à l'intention des préservateurs* du projet InterPARES 2, comprend : l'établissement, le suivi et la mise en œuvre des procédures pour enregistrer le transfert du document; la vérification de l'autorité du transfert; l'assurance que les documents d'archives en cours de transfert correspondent aux documents d'archives désignés pour le transfert; la vérification de l'authenticité des documents d'archives transférés; et l'accès aux documents d'archives.

Dans le *modèle de référence du SOAI*, l'acquisition fait partie de l'entité fonctionnelle de saisie et le transfert de données géospatiales évaluées est couvert par des ententes de dépôt conclues entre le préservateur et le créateur. Les ententes de dépôt déterminent les TIS qui seront soumises, la durée pendant laquelle ces dépôts seront faits au système de préservation, et le modèle de données à utiliser pour les dépôts. Le modèle de données spécifie les composantes logiques de la TIS (p. ex., le contenu des objets de données, les données sur la représentation, l'IDP, les renseignements sur les trousseaux et les données descriptives) et comment (et si) elles sont représentées dans chaque session de dépôt de données.

Voici les objectifs de l'entité fonctionnelle de saisie.

- *Recevoir les trousseaux d'information sur les soumissions (TIS)* – Les TIS numériques peuvent être livrées par les créateurs (ou des éléments internes gérés par l'administration) par virement électronique, chargées à partir des supports soumis au système de préservation, ou simplement placées (p. ex., CD-ROM) sur le système de fichiers du système pour les rendre accessibles. La preuve d'authenticité est fournie par le créateur dans le cadre de l'IDP dans le dépôt, et cette preuve est maintenue, mise à jour, ou majorée par le système de préservation au fil du temps.
- *Effectuer l'assurance de la qualité sur les TIS* – La réussite du transfert de la TIS vers la zone de stockage temporaire est validée. Pour les dépôts numériques, des mécanismes de validation pourraient inclure des contrôles de redondance cycliques (CRC) ou des sommes de contrôle associés à chaque fichier de données, ou l'utilisation de fichiers journaux d'exploitation pour enregistrer et définir tout transfert de fichier ou des erreurs de lecture/d'écriture.
- *Générer des trousseaux d'information archivée (TIA)* – Les TIS sont transformées en une ou plusieurs TIA qui sont conformes aux normes de documentation et aux normes de formatage des données du système de préservation. Cela peut comprendre les conversions de formats de fichiers, la collecte de données sur la représentation adéquate, les conversions des données de représentation ou la réorganisation des renseignements sur le

contenu dans les TIS. Une requête est envoyée à la gestion des données pour obtenir les rapports d'information nécessaires pour produire les informations descriptives qui complètent les TIA.

- *Extraire les données descriptives des TIA* – Les données descriptives sont extraites des TIA et recueillies auprès d'autres sources afin de fournir à la fonction *Mise à jour des coordonnées* et à la gestion des données. Cela inclut des métadonnées pour soutenir la recherche et la récupération des TIA (p. ex., qui, quoi, quand, où et pourquoi), et pourrait aussi inclure des produits d'accès direct spéciaux (p. ex., vignettes, images) qui seront utilisés dans la recherche de données.
- *De la mise à jour des coordonnées au stockage d'archives et à la gestion des données* – Les TIA sont transférées à l'archivage et les données descriptives sont transférées à la gestion des données. Après que le transfert est terminé et vérifié, l'archivage fournit les informations d'identification de stockage pour la TIA, qui sont incorporées dans les données descriptives pour la TIA et transférées à l'entité de gestion des données avec une demande de mise à jour de la base de données.

Principaux points

- Les procédures d'acquisition doivent être élaborées en collaboration par les créateurs et les préservateurs, normalisées et appliquées.
- Les ententes de dépôt sont un bon moyen d'assurer la viabilité à long terme du processus d'acquisition.
- Les processus pour la saisie des documents dans le système de préservation doivent également être bien

Dans la *Liste de contrôle de la VADF*, les éléments saisis liés à la création des TIA sont similaires aux processus du SOAI. Les types d'éléments de preuve liés à la saisie des documents d'archives dans un centre d'archives qui est utilisé dans un processus d'accréditation sont : les énoncés de mission, les ententes de dépôt, le déroulement de travail et les documents de politique et de procédures, les procédures de traitement et la documentation des propriétés qui doivent être préservées. De plus, la preuve nécessite que les mesures techniques appropriées soient en place pour saisir les documents d'archives, que les journaux de procédures et d'authentification soient conservés, et que les registres d'acquisition soient maintenus.

La publication de GeoMAPP, *Best Practices for Archival Processing for Geospatial Datasets*, fournit un excellent cadre au déroulement du travail des organismes d'archivage qui traitent des ensembles de données géospatiales.

6.9 Préserver les documents d'archives

Une fois les données géospatiales acquises et saisies, elles doivent être préservées. Le processus de préservation des documents d'archives versés a été abordé dans les discussions précédentes de la planification de la préservation, des politiques et des procédures, de l'évaluation et de l'acquisition. Cependant, certaines questions spécifiques relatives à l'archivage de documents d'archives restent importantes à aborder. L'entité fonctionnelle du SOAI chargée de l'archivage est conçue de manière à exécuter les processus d'archivage du système de préservation. Dans ce

contexte, le terme « support » est utilisé pour désigner un ou plusieurs mécanismes locaux ou distants pour archiver des informations codées numériquement. Dans ce modèle, les données sont reçues, elles sont gérées dans une hiérarchie d'archivage, les supports sont remplacés (c.-à-d. des TIA peuvent être actualisées), les erreurs sont vérifiées, les mécanismes de gestion des catastrophes sont créés, et les processus sont en place pour fournir les données aux consommateurs.

Voici les objectifs de l'entité fonctionnelle chargée de l'archivage.

- *Recevoir les TIA de la saisie* - Une demande d'archivage et une TIA sont reçues de l'entité fonctionnelle chargée de la saisie et la TIA est déplacée vers un stockage permanent dans le système de préservation. Une indication de la fréquence prévue de l'utilisation des objets de données qui composent la TIA facilitera la sélection des périphériques ou supports d'archivage les plus appropriés. Le type de support est sélectionné, les périphériques ou les volumes préparés, et le transfert physique vers les volumes d'archivage réalisé. À la fin du transfert, cette fonction envoie un message de confirmation d'archivage à l'entité de saisie, notamment l'identification de l'archivage des TIA.
- *Gérer la hiérarchie d'archivage* - Les TIA sont positionnées sur le support approprié, conformes à tous les niveaux spéciaux de service requis pour la TIA ou à toutes les mesures particulières de sécurité, et le niveau approprié de protection de la TIA est assuré. Cette fonction s'assure également que les TIA ne sont pas corrompues durant les transferts, et fournit des statistiques à l'entité fonctionnelle chargée de l'administration sur les supports disponibles, la capacité d'archivage disponible dans les différents niveaux de la hiérarchie d'archivage, et l'utilisation du système de préservation.
- *Actualiser le support de données* - La capacité à reproduire les TIA au fil du temps, sans en altérer les renseignements sur le contenu et l'IDP. Si les attributs dépendants du support (p. ex., la taille des blocs de bande, les informations sur le volume du CD-ROM) ont été inclus dans le cadre des renseignements sur le contenu, un moyen doit être trouvé pour préserver ces renseignements lors de la migration vers un support de plus grande capacité avec différentes architectures d'archivage (voir le *modèle de référence du SOAI* pour les questions sur l'examen détaillé de la migration des données).
- *Réaliser une vérification des erreurs* – Cette fonction fournit une assurance statistiquement acceptable qu'aucune composante de la TIA ne soit corrompue dans l'archivage ou pendant tout transfert interne de données d'archivage. Les procédures doivent être élaborées pour la vérification aléatoire de l'intégrité des objets de données en utilisant les CRC ou un autre mécanisme de vérification des erreurs.
- *Fournir des capacités de reprise après catastrophe* - Des mécanismes sont prévus pour dupliquer les contenus numériques de la collection préservée et, par exemple, archiver le double dans un établissement physiquement séparé, conformément aux politiques de reprise après catastrophe précisées dans l'entité fonctionnelle chargée de l'administration. Cette fonction est normalement réalisée en copiant le contenu des archives sur une certaine forme de support de stockage amovible (p. ex., une bande linéaire numérique, un CD-

ROM), mais peut également être effectuée par le biais du transport de matériel ou des transferts de données du réseau.

- *Fournir des TIA à l'accès* - Cette fonction fournit des TIA sur le type de support requis ou les transfère vers une zone d'archivage temporaire lors de la réception d'une demande de TIA de l'entité fonctionnelle chargée de l'accès. Cette fonction envoie également un avis de transfert de données à l'entité fonctionnelle chargée de l'accès à l'achèvement d'une commande.
- *Gérer la configuration du système* - La fonctionnalité de l'ensemble du système d'archivage est surveillée en permanence et les changements à la configuration sont systématiquement contrôlés. Cette fonction fait la vérification des activités, du rendement et de l'utilisation du système et reçoit des rapports sur les informations du système de l'entité fonctionnelle chargée de la gestion des données et produit des rapports sur les statistiques opérationnelles de l'entité fonctionnelle chargée de l'archivage. Elle résume ces rapports et fournit périodiquement des rapports d'information sur le rendement d'archives et des rapports d'inventaire sur les données d'archives à l'entité fonctionnelle chargée de la planification de la préservation.

6.10 Décrire les métadonnées d'archives

Les données géospatiales sont décrites par les créateurs dans leurs métadonnées. Les métadonnées d'archives décrivent les données saisies et leurs métadonnées (partie de la TIS) et

PRATIQUE EXEMPLAIRE

Le rapport de GeoMAPP, *Archival Metadata Elements for the Preservation of Geospatial Datasets* (GeoMAPP, 2011) fondé sur les métadonnées de description du SOAI est un guide utile. Il comprend un tableau qui fait passer le SOAI au FGDC et aux éléments de métadonnées de Dublin Core. Le tableau a été informé par :

- les éléments de métadonnées de préservation du projet Cedars (Stone & Day, 1999);
- le rapport National Geospatial Digital Archive (NGDA) par Susan Hoebelheinrich (Hoebelheinrich, 2009), qui fournit des recommandations pour les ensembles de données géospatiales complexes;
- le modèle de métadonnées des documents électroniques de données géospatiales (CIESIN, 2005) du Center for International Earth Science Information Network (CIESIN);
- le dictionnaire de données PREMIS pour les métadonnées de préservation (Bibliothèque du Congrès, 2013).

comprennent des informations sur les documents d'archives et leurs contextes (partie de la TIA), qui sont recueillis lors de l'évaluation, du traitement, de la saisie et des phases de préservation par les préservateurs. La *Liste de contrôle de la VADF* comprend une description des données dans son élément de gestion de l'information et la documentation à l'appui qui suit est nécessaire lors de l'examen d'un système de préservation aux fins de l'accréditation : métadonnées descriptives, identificateurs persistants liés à la TIA, ententes avec les déposants, liens des métadonnées à la TIA, journaux documentant l'intégrité référentielle, et documentation du déroulement des opérations.

L'identité des archives dans les *Lignes directrices à l'intention des créateurs* du projet InterPARES est établie dans les éléments de métadonnées qui sont très similaires à la plupart des éléments de données géospatiales. Les

métadonnées d'archives comprennent également des informations sur ce qui suit : des informations sur l'encodage et les enveloppeurs, la date de dépôt d'archivage, la présence ou l'absence d'une signature numérique, les indicateurs d'authentification tels que la corroboration ou les moyens de valider l'ensemble des données, une attestation du document qui est analogue à la signature de l'ensemble de données quand il a été émis, et les numéros des versions d'ébauches. L'information sur l'intégrité d'un ensemble de données est également nécessaire et cela inclut des éléments tels que : les noms des personnes ou du bureau responsables, les noms des autorités compétentes, les indications d'annotations aux données et la présence de modifications techniques, le cas échéant, telles que les enveloppeurs, le codage et les mises à niveau de la version. De plus, les éléments concernant les codes de restriction, les privilèges d'accès, l'importance du document, la disposition prévue (c.-à-d. lorsque les documents d'archives peuvent être transférés aux archives) sont des éléments recommandés.

Les *conditions de base* du projet InterPARES 2 comprennent une exigence de documentation intitulée *Documentation of Reproduction Process and its Effects*, qui recommande notamment ce qui suit :

- la date à laquelle le document a été reproduit et le nom de la personne ou de l'unité d'affaires responsable;
- la relation entre le document obtenu du créateur et les copies créées par le préserveur;
- les effets du processus de reproduction sur la forme, le contenu, l'accessibilité et l'utilisation du document;
- dans le cas où une copie d'un document est connue comme ne reproduisant par intégralement et fidèlement l'élément exprimant son identité et son intégrité, documenter de telles informations et rendre cette documentation facilement accessible aux utilisateurs.

6.11 Gérer et maintenir les documents d'archives

Une fois les documents d'archives saisis et préservés, l'étape suivante consiste à établir des processus pour assurer la gestion et la maintenance adéquats. Les *Lignes directrices à l'intention des préserveurs* du projet InterPARES 2 soulignent la gestion des documents d'archives et les stratégies de préservation, dont plusieurs ont été abordés dans la section de planification de la préservation. Voici les exigences minimales nécessaires pour protéger et maintenir l'accès à des copies authentiques des documents d'archives numériques.

- *Répartition claire des responsabilités* – une personne ou un bureau ayant la responsabilité « sans ambiguïté » de gérer l'archivage et la protection des documents d'archives, avec un personnel qualifié dédié à cette activité.
- *Instauration d'une infrastructure technique adéquate* – ressources matérielles et administratives pour permettre la tenue des documents d'archives tels que les bâtiments, le matériel informatique, les réseaux informatiques, et le personnel auxiliaire.
- *Entretien, soutien et remplacement du système* – un plan de mise en œuvre pour la maintenance, la mise à jour et le remplacement du matériel et des logiciels.
- *Transfert régulier des données sur de nouveaux supports de stockage* – procédures systématiques pour copier les données d'un support de stockage à un autre pour éviter la détérioration du support.
- *Adoption de conditions adéquates pour les supports de stockage* – évitement de la détérioration du support en respectant les conditions environnementales appropriées, telles qu'éviter l'humidité, la poussière et la chaleur.
- *Redondance et emplacement géographique* – la duplication des entités numériques et le stockage des copies multiples sur différents supports de stockage pour protéger contre la défaillance et l'archivage dans des endroits différents pour se protéger contre les catastrophes et les mauvaises conditions environnementales.
- *Sécurité du système* – restriction d'accès aux documents d'archives pour les processus et les utilisateurs autorisés, notamment la restriction de l'accès physique aux ordinateurs et au stockage, et aux documents d'archives sur les ordinateurs.

- *Planification en cas de catastrophe* – plans de reprise des activités après une catastrophe en place avec des procédures détaillées pour restaurer un système endommagé et pour orienter la reprise du système de préservation à la suite d'une catastrophe.

La *Liste de contrôle de la VADF* comprend des stratégies de maintien qui couvrent de nombreux éléments évoqués ci-dessus et comprend également : la fonctionnalité de sauvegarde, la synchronisation de la copie, la détection de la corruption de bits et de la perte de documents d'archives, des évaluations des risques pour contrer les menaces et les attaques entraînant un refus de service, et la mise en œuvre des contrôles pour répondre aux besoins de sécurité. Il y a aussi l'attente que les systèmes de préservation se conformeront à la norme *ISO 17799 Technologies de l'information - Techniques de sécurité - Code de bonne pratique pour la gestion de la sécurité de l'information* (ISO, 2005) et qu'il y aura des plans de reprise après catastrophe, des preuves de copies hors site, des plans de continuité du service, des organigrammes et des journaux des enregistreurs.

Principaux points

- Des stratégies de maintenance et de gestion efficaces sont essentielles pour assurer l'accessibilité continue des copies authentiques des documents numériques
- Le respect de la norme *ISO 17799 Code de bonne pratique pour la gestion de la sécurité de l'information* aidera à garantir que les systèmes de préservation maintiennent correctement leurs documents.

6.12 Fournir un accès aux documents d'archives

Principaux points

- Des politiques et des procédures ouvertes et transparentes concernant l'accès aux documents aideront les utilisateurs à découvrir les données qui correspondent à leurs besoins et à y accéder.
- Une entente de commande, qui peut être un simple formulaire en ligne, sert habituellement de mécanisme aux utilisateurs pour communiquer avec les préservateurs.

La dernière étape dans l'établissement d'un système de préservation des données géospatiales est de s'assurer que les utilisateurs peuvent accéder aux documents d'archives préservés. La manière dont les données préservées sont mises à la disposition des utilisateurs dépend des besoins des utilisateurs et de la mission et des politiques de l'archives. Un portail Web est souvent utilisé pour donner accès aux documents d'archives, qui peuvent être recherchés grâce à des métadonnées descriptives. Des mécanismes doivent être conçus pour rendre plus facile pour les utilisateurs de découvrir des ensembles de données d'intérêt et la façon dont les documents d'archives ont été créés, d'être informés sur les licences, les droits d'utilisation et le coût potentiel pour accéder à ce document, et de passer des commandes. Des mécanismes de transmission des données doivent également être clairement décrits.

Les politiques d'accès doivent être publiques, de sorte que les règles d'autorisation et les exigences d'authentification qui peuvent être incluses dans les ententes de dépôt sont transparentes. Les utilisateurs ont besoin de la fonctionnalité pour faire un lien entre ce à quoi ils ont accédé et l'original. La *Liste de contrôle-de la VADF* renferme un guide utile afin de comprendre les dépôts numériques et la fonction d'accès, et ainsi d'appuyer l'élaboration de la composante d'accès du système de préservation. Les preuves à l'appui des fonctions d'accès acceptables au système de préservation dans la *Liste de contrôle-de la VADF* comprennent : les journaux d'accès, les journaux de demande et de déni, les documents de conception du système, les journaux de production de la TDI, et les processus de revue générale.

Dans le *modèle de référence du SOAI*, l'entité fonctionnelle chargée de l'**accès** offre une interface entre le système de préservation et l'utilisateur. L'utilisateur établit une entente de commande avec le système de préservation, qui identifie les TIA d'intérêt, la façon dont ces TIA doivent être transformées et cartographiées dans les TDI et comment ces TDI seront regroupées dans une séance de diffusion des données. L'entente de commande peut également spécifier d'autres informations nécessaires telles que les détails de livraison (p. ex., le nom, l'adresse postale, etc.), les informations sur les droits (p. ex., les restrictions d'utilisation, les utilisateurs autorisés, ou les droits de licence) et les prix, le cas échéant. L'entente peut être un document officiel ou le processus d'élaboration d'une entente de commande peut n'être qu'un formulaire rempli en ligne pour préciser les TIA d'intérêt.

La fonction *Mise à jour de l'information archivistique* dans l'entité fonctionnelle chargée de l'**administration** soumet également les demandes de diffusion pour obtenir les TDI nécessaires pour s'acquitter de ses fonctions de mise à jour. Cette fonction permet de déterminer si les ressources sont disponibles pour effectuer une demande, s'assure que l'utilisateur est autorisé à accéder aux éléments demandés et de les obtenir, et informe les utilisateurs qu'une demande a été acceptée ou rejetée. Elle transmettra alors la demande à l'entité fonctionnelle chargée de la **gestion des données** ou à la fonction *Générer des TDI* pour exécution.

- *Générer des TDI* – Cette fonction accepte une demande de diffusion, récupère la TIA de l'**archivage**, déplace une copie des données dans une zone d'archivage temporaire pour un traitement ultérieur, et transmet une demande de rapport à l'entité fonctionnelle chargée de la **gestion des données** pour acquérir les données descriptives nécessaires à la TDI. Si un traitement spécial est nécessaire, la fonction accède aux objets de données dans le stockage temporaire et applique les processus demandés (p. ex., l'insertion d'information numérique sur la gestion des droits et le filtrage de données personnelles pour assurer la cohérence avec les droits de l'utilisateur). Elle place la réponse terminée de la TDI dans la zone d'archivage temporaire et informe la fonction *Coordonner les activités d'accès* que la TDI est prête à être fournie.
- *Fournir la réponse* – Pour une remise en ligne, cette fonction accepte une réponse de la fonction *Coordonner les activités d'accès* et la prépare pour la distribution en ligne en temps réel au moyen des liens de communication. Elle détermine le destinataire prévu et la

procédure de transmission demandée, place la réponse dans la zone d'archivage temporaire afin d'être transmise, et prend en charge la transmission en ligne de la réponse. Pour une remise hors ligne, elle récupère la réponse de la fonction *Coordonner les activités d'accès*, et prépare et expédie la réponse. Lorsque la réponse a été expédiée, un avis de commande expédiée est retourné à la fonction *Coordonner les activités d'accès* et les informations de facturation sont soumises à l'entité fonctionnelle chargée de l'**administration**.

Les *European LTDP Common Guidelines* donnent des indications sur la création des mécanismes pour chercher et découvrir des données et de la documentation propre à la mission, maintenir les métadonnées consultables, parcourir un catalogue d'images, et utiliser des outils de visualisation pertinents pour les communautés désignées. Les *lignes directrices* fournissent également une liste de normes d'interface d'accès de l'OGC.

7. Défis et solutions

Il y a de nombreux défis rencontrés au cours de la préservation des données géospatiales. Les *Lignes directrices du projet InterPARES 2*, le *modèle de référence du SOAI*, les *European LTDP Common Guidelines* et la *Liste de contrôle de la VADF* ont été créés pour répondre aux défis organisationnels, d'objet numérique et de préservation technologique les plus communes. Les recommandations de l'*Archivage, gestion et préservation des données géospatiales – Rapport sommaire et recommandations* (GéoConnexions et HAL, 2011) continuent de s'appliquer également. Les outils GeoMAPP et les documents LTDP/Preserved Data Content Definition sont plus spécifiques et adaptés pour aider les préservateurs de données géospatiales, tandis que les directives et les politiques du SCT ont été créés pour orienter les mesures de gestion des données de tous les créateurs de documents d'archives du gouvernement fédéral. Le tableau 2 décrit certaines des défis de préservation les plus courantes, les pratiques exemplaires et les leçons retenues pour répondre à ces défis révélées par les études de cas, les profils et les documents et la littérature de recherche qui ont servi de base à la création de ce guide d'introduction.

Tableau 2 : Défis et solutions des archives fédérées

Défis communs	Solutions
Exigences en fonction du type de données géospatiales numériques	<p>Une évaluation approfondie des types de données qui doivent être saisies est réalisée au cours de la phase d'évaluation. À ce moment, les composantes numériques de chaque ensemble de données sont définies et les éléments nécessaires pour s'assurer qu'elles peuvent être consultées et utilisées sont déterminés. Le registre du format de fichier est un outil utile pour ce processus.</p> <p>Tous ces éléments, y compris les données, doivent être préservés et liés les uns aux autres lors du dépôt et maintenus pendant la préservation et l'accès.</p> <p>Dans le cas de l'ITO, les métadonnées, les trousseaux du format d'interface normalisé du Système de DINGO et les rapports techniques sont chargés dans le SAIG avec les données.</p> <p>Le CCT inclut des éléments d'OT propre à la mission tel que décrit dans le document LTDP/Preserved Data Set Content.</p> <p>[Nota - Dans le cas du CCT, les entités du secteur privé fournissent des produits dérivés d'OT et le CCT n'a pas accès à ces logiciels et algorithmes utilisés pour générer les produits. De nouvelles ententes doivent être en place pour s'assurer que ces derniers sont transférés pour permettre la préservation.]</p>

Défis communs	Solutions
Modification des données géospatiales numériques dynamiques.	<p>Lorsque les données sont mises de côté aux fins de préservation, elles doivent être fixes. La variabilité limitée signifie que le contenu du document (p. ex., les données pour créer la carte et l'algorithme utilisé pour afficher celle-ci sur l'écran) doit être fixe. La forme documentaire du document d'archives (p. ex., spécifications et logiciel utilisé pour créer et visualiser la carte interactive) doit être inaltérable, de sorte que sa présentation reste toujours la même.</p> <p>Des règles fixes doivent être instaurées concernant le choix et la présentation du contenu, permettant ainsi une variabilité stable de la carte interactive ou du modèle.</p> <p>Lorsque les données ne sont pas encore fixes, elles doivent encore être soumises à des contrôles de gestion des données rigoureuses, comme dans le cas des données de bouées dérivantes du MPO.</p>
Détermination de la fréquence d'archivage des données géospatiales numériques.	<p>Les décisions qui animent ce qui précède sont fondées sur les besoins des utilisateurs et les pratiques opérationnelles du créateur et du prescripteur. Les organisations devront déterminer cette fréquence en collaboration avec les utilisateurs, et ce, en suivant les directives des créateurs de données et en examinant la valeur fonctionnelle des documents d'archives en question.</p> <p>Dans le cas du CCT, toutes les données d'OT de la mission sont préservées, et dans certains cas, des distinctions sont faites entre les données essentielles et celles qui ne le sont pas.</p> <p>L'ITO a choisi de saisir un instantané annuel de l'Entrepôt et des documents d'archives retirés qui sont jugés officiels.</p> <p>L'enquête du NGDIA en Caroline du Nord est un outil utile dans ce contexte.</p>
Changements technologiques rapides.	<p>Les cadres, les lignes directrices et les outils abordent cette question dans les phases d'accès, de planification de la préservation et de maintenance. <i>Le modèle de référence du SOAI</i> et la <i>Liste de contrôle de la VADF</i> en particulier abordent le suivi du statut, les plans de transition technologique et la gestion des risques. Les <i>Lignes directrices à l'intention des créateurs et des prescripteurs du projet InterPARES</i> fournissent des recommandations utiles au sujet des manuels et de la tenue de copies des logiciels.</p> <p>Le SGDOT du CCT, par exemple, est conçu de manière à établir un cadre d'élaboration, de mise à l'essai et de production qui fait en sorte que la transition vers de nouvelles technologies ne perturbe pas le système. Les choix de logiciels sont importants et plus ils sont libres et normalisés, plus il sera facile de les gérer. Le financement durable à long terme est également nécessaire.</p> <p>Les <i>Lignes directrices à l'intention des créateurs</i> du projet InterPARES 2 suggèrent également que les créateurs prennent cela en considération.</p>

Défis communs	Solutions
Archivage des données numériques distribuées	La principale recommandation est la détermination de tous les propriétaires de données au moment de la préservation et l'élaboration des stratégies de préservation convenues entre chacun d'eux. Le concept de variabilité limitée s'applique ici
Considérations concernant la protection des renseignements personnels, la confidentialité et les limites d'utilisation	Des lois encadreront la prise de décisions relatives à la protection des renseignements personnels et à la confidentialité, tandis que des PE, des licences et des ententes pourraient exposer les limites d'utilisation. Cela sera déterminé au cours de la phase d'évaluation et cette documentation fait partie de la trousse de soumission. Le système de préservation doit permettre l'exécution de tous les droits d'utilisation et faire en sorte qu'ils soient pris en compte au moment de l'accès. Le CCT est limité par sa mission; à l'ITO, ces limites sont réglées à l'Entrepôt, et pour l'API, associées aux contrats de licence d'utilisation.
Obtention des ressources nécessaires	Des ressources financières durables sont nécessaires pour maintenir un système de préservation, qui comprend des ressources humaines qualifiées, la technologie et l'espace physique. Il est nécessaire d'analyser clairement les coûts et de se doter de plans d'activités; à ce titre, les outils de GeoMAPP sont utiles. La <i>Liste de contrôle de la VADF</i> souligne la révision des plans sur une base annuelle. Le projet de l'API a souffert en raison d'un manque de ressources durables et le projet du CCT est élaboré progressivement et devra assurer le financement dédié à long terme dans le futur.
Pénuries de compétences et d'expérience d'archivage et de préservation	L'éducation et la formation continues du personnel sont une exigence dans le <i>modèle de référence du SOAI</i> et la <i>Liste de contrôle de la VADF</i> . Le transfert de connaissances est également nécessaire. Le CCT mettra en œuvre un programme de formation. L'API a déterminé ceci comme un problème puisque les scientifiques n'avaient pas les compétences requises pour gérer leurs documents d'archives et a recommandé que la gestion des données soit un élément clé de l'éducation pour tous les chercheurs. La GDSI s'appuie sur un réseau distribué de conservateurs sous contrôle centralisé. De plus, de nombreuses institutions gouvernementales ont accès à des bibliothécaires et des archivistes qualifiés, et ils devraient être invités à aider à concevoir les systèmes de préservation et élaborer les procédures et les politiques.

Défis communs	Solutions
Début du processus de préservation	<p>Les activités de préservation commencent avant que les documents d'archives soient créés et continuent tout au long du cycle de vie d'un ensemble de données.</p> <p>Le projet de l'API recommande la création d'un plan de gestion des données combiné avec des mesures incitatives afin que les chercheurs maintiennent correctement leurs documents d'archives et trouvent un centre d'archives approprié pour le dépôt de ces documents.</p> <p>Le SAIG s'inscrit dans le processus de tenue de documents d'archives de l'EDTO, qui se poursuit jusqu'à la préservation.</p> <p>Le CCT compte la préservation parmi ses pratiques courantes pour la collecte de données sur la mission d'OT.</p> <p>Les directives et les politiques du SCT prescrivent aussi que les documents d'archives soient maintenus et gérés tout au long de leur cycle de vie.</p>
Où stocker les données géospatiales	<p>Au moment d'écrire ces lignes, il y a peu d'archives de données numériques.</p> <p>Dès que le SGDOT du CCT sera en ligne, cela fournira un centre d'archives d'OT centralisées du gouvernement du Canada avec le potentiel de croître et de saisir d'autres ensembles de données.</p> <p>La GDSI offre une certaine préservation de documents d'archives et utilise le Centre mondial de données pour certaines de ses données.</p> <p>Au-delà de cela, il y a peu de dépôts numériques qui peuvent saisir des données géospatiales et ceux-ci doivent être mis sur pied si l'on s'attend à ce que les données soient préservées.</p>
Préparation aux situations d'urgence et sécurité	<p>La <i>Liste de contrôle de la VADF</i> la création d'un plan d'intervention d'urgence en cas d'une catastrophe humaine (p. ex., incendies, déversements de produits chimiques) ou naturelle (p. ex., inondations, tremblements de terre) ou des bouleversements politiques. Cela implique d'avoir un stockage hors site, mais aussi d'être en mesure de démontrer que les systèmes d'exploitation et de sauvegarde peuvent reprendre les activités après des événements catastrophiques. Cela signifie également être capable d'accéder aux bâtiments où se trouve le système de préservation, et toute la documentation nécessaire pour commencer à nouveau.</p> <p>Le respect de la norme <i>ISO/IEC 17799:2005 Technologie de l'information – Techniques de sécurité – Code de bonne pratique pour la gestion de la sécurité de l'information</i> est une pratique recommandée.</p>
Rôles et responsabilités	<p>Les politiques et les procédures doivent définir ces éléments de manière claire et nette. Cela signifie avoir une chaîne de commandement claire, une structure de gouvernance et un organigramme qui définit les institutions responsables, ainsi que des rôles et des mandats clairs et explicites.</p>

8. Conclusions

Établir des archives et élaborer les systèmes de préservation connexes est une entreprise complexe, nécessitant une compréhension approfondie des concepts de base d'archivage et de préservation, un plan d'activités solide et l'engagement à long terme de l'hôte des archives et des organisations contributrices de données. Ce guide d'introduction vise à présenter ces concepts aux lecteurs et à fournir un aperçu de haut niveau des étapes de base de la création d'un système de préservation de données géospatiales numériques. Les idées contenues dans ce document proviennent d'un examen de la littérature sur les archives et la préservation des données numériques et de bonnes pratiques et des leçons retenues des premiers à adopter des approches et des méthodes de préservation des données numériques.

Bien que la préservation des données puisse être facultative pour certains créateurs de données, les organisations gouvernementales sont souvent mandatées pour le faire. Les lecteurs sont priés de tenir compte de la législation (et en particulier de la *Loi sur la Bibliothèque et les Archives du Canada*), des règlements, des politiques et des directives du gouvernement du Canada qui touchent la préservation des documents d'archives au moment d'évaluer la pertinence de mettre en place des archives de données géospatiales. Ce guide met en évidence d'importantes obligations, limites et défis liés à la législation et aux politiques clés du GC qui vont influencer sur les décisions de préservation des données.

Le guide introductoire s'appuie fortement sur les directives fournies par cinq documents particulièrement importants : le *modèle de référence du SOAI*, les *European LTDP Common Guidelines*, les *Lignes directrices à l'intention des créateurs*, *Les lignes directrices à l'intention des préservateurs* du projet InterPARES 2, et la *Liste de contrôle de la VADF*. Ces documents suivent souvent les mêmes approches, méthodes et modèles, utilisent des terminologies similaires ou identiques et sont largement acceptés au sein de la communauté archivistique comme les meilleures pratiques internationales. Un autre excellent ensemble de documents d'orientation auquel il est souvent fait référence, généré par le Geospatial Multistate Archive and Preservation Project (GeoMAPP), est particulièrement pertinent pour ceux qui souhaitent préserver des documents d'archives numériques géospatiaux.

Au moment de la publication, le Canada possédait peu d'expérience dans la préservation des données géospatiales numériques. Un certain nombre de facteurs ont contribué à la rareté des exemples des services d'archives numériques géospatiales, notamment : le manque de ressources, la demande limitée des utilisateurs, le manque de connaissance et de compétences en matière de préservation des données numériques, et des incertitudes quant à la fréquence optimale de versement des données dans un système de préservation. Toutefois, quatre exemples d'efforts de préservation des données numériques ont été examinés (le Système de gestion des données d'observation de la Terre prévu du CCT, le Système d'archivage de l'information géographique du MRNO, le système de gestion des sciences intégrées du MPO et les efforts de

préservation des données liés à l'Année polaire internationale), leurs expériences sont décrites et leurs pratiques exemplaires sont mises en évidence.

Plusieurs créateurs de données géospatiales envisagent activement les initiatives de préservation des données, et ce guide introductoire a été élaboré pour fournir une orientation à ces efforts. Les lecteurs sont invités à communiquer leurs pratiques exemplaires de préservation et les leçons retenues avec GéoConnexions, de sorte que ce guide puisse être amélioré, afin d'aider la communauté de l'information géospatiale à accroître son efficacité dans ce domaine complexe d'exercice.

A. Ouvrages de référence

- API. (2010). *About IPY*. Extrait le 3 janvier du site Web de l'Année polaire internationale : <http://ipy.arcticportal.org/about-ipy>
- BAC. (2006). *Bibliothèque et Archives Canada, Gestion de l'information*. Extrait le 18 février 2013 de Gestion du cycle de vie des documents d'archives et de l'information : <http://www.collectionscanada.gc.ca/government/products-services/007002-2012-f.html>
- BAC. (2009). Orientations quant à l'établissement de spécifications en matière de conservation pour les ressources documentaires ayant une valeur opérationnelle. Ottawa, Ontario, Canada : BAC.
- BAC. (2011). Bibliothèque et Archives Canada. (B.A. Canada, producteur) Extrait le 18 février 2013 de Gestion de l'information (GI) – Bibliothèque et Archives Canada : <http://www.collectionscanada.gc.ca/gouvernement/produits-services/007002-6000-f.html>
- BAC. (2012). *Bibliothèque et Archives Canada*. Extrait le 18 février 2013 du Programme de disposition et de tenue de documents d'archives : <http://www.bac-lac.gc.ca/fra/programme-disposition-tenue-documents/Pages/programme-disposition-tenue-documents.aspx>
- Ball, A. (2012). *Review of Data Management Lifecycle Models*. Extrait de OPUS : Boutique de publication en ligne : <http://opus.bath.ac.uk/28587/>
- Bleakly, D. (2002). *Long-Term Spatial Data Preservation and Archiving: What are the issues?* Extrait le 2 février 2013 de <http://prod.sandia.gov/techlib/access-control.cgi/2002/020107.pdf>
- CBUO. (2011). *Dépôt numérique fiable*. Extrait du Conseil des bibliothèques universitaires de l'Ontario : <http://www.ocul.on.ca/node/97>
- CBUO. (2012). *Trusted Digital Repositories Document Checklists*. Extrait du Scholars Portal : <http://spotdocs.scholarsportal.info/display/OAIS/Document+Checklist>
- CCSDS. (2012). *Modèle de référence du Système ouvert d'archivage d'information (SOAI)*. Extrait le 2 janvier 2012 du modèle de référence du Consultative Committee for Space Data Systems <http://public.ccsds.org/publications/archive/650x0m2.pdf>
- CCT. (2013). Services et accès de données d'observation de la Terre du Canada, version 3.0. Ottawa : Ressources naturelles Canada.

- CIESEN. (2005). *Data Model for Managing and Preserving Geospatial Electronic Records*.
Extrait le 11 février 2011 du Centre for International Earth Science Information
Network:Geospatial Electronic Records :
http://www.ciesin.columbia.edu/ger/DataModelV1_20050620.pdf
- Cour suprême du Canada. (2004). CCH Canadienne limitée. c. Barreau du Haut-Canada ([2004] 1
R.C.S. 339). Extrait le 11 mars 2013 du site Web de la Cour suprême du Canada :
<http://scc.lexum.org/decisia-scc-csc/scc-csc/scc-csc/fr/item/2125/index.do>
- DCC. (2013). *Digital Curation Centre Lifecycle Model*. Extrait en 2013 de DCC :
<http://www.dcc.ac.uk/resources/curation-lifecycle-model>
- DPC. (2008). *Preservation Management of Digital Materials: The Handbook*. Extrait le 2 janvier 2013
de la Digital Preservation Coalition :
<http://www.dpconline.org/advice/preservationhandbook/introduction/definitions-and-concepts>
- Erwin, T., & Sweetking-Singer, J. (2009). *The National Geospatial Digital Archive: A
Collaborative Project to Archive Geospatial Data*. Extrait le 11 mars 2013 du Journal of Map
And Geography Libraries : <http://dx.doi.org/10.1080/15420350903432440>
- GéoConnexions. (2008). *La diffusion de données géographiques gouvernementales au Canada –
Guide des pratiques exemplaires*. Extrait le 11 mars 2013 des publications GéoConnexions :
http://geoscan.ess.nrcan.gc.ca/starweb/geoscan/servlet.starweb?path=geoscan/geoscanfastlink_e.web&search1=R=288853
- GéoConnexions et Hickling Arthurs et Low (HAL). (2011). *Archivage, gestion et préservation des
données géospatiales – Rapport sommaire et recommandations*. Ressources naturelles
Canada. Ottawa, Ontario, Canada.
- GéoConnexions et Hickling Arthurs et Low (HAL). (2011). *Document juridique sur la propriété
intellectuelle*. Ressources naturelles Canada. Ottawa, Ontario, Canada.
- GéoConnexions. (2013). Politiques opérationnelles. Extrait le 13 février 2013 des publications
GéoConnexions : <http://geoconnexions.nrcan.gc.ca/19>
- GeoMAPP. (2009). *System Inventory Template*. Extrait le 21 janvier 2013 de Geospatial Multistate
Archive and Preservation Partnership Publications and Tools :
http://www.geomapp.net/docs/GeoMAPP_System_Inventory_Template.pdf
- GeoMAPP. (2010). *Geoarchiving Self-Assessment*. Extrait le 21 janvier 2013 de Geospatial
Multistate Archive and Preservation Partnership Publications and Tools:
http://www.geomapp.net/publications_categories.htm#assess

- GeoMAPP. (2011). *Archival Metadata Elements for the Preservation of Geospatial Datasets*. Extrait le 2 février 2013 de Geospatial Multistate Archive and Preservation Partnership (GeoMapp) Publications and Tools : http://www.geomapp.net/docs/GIS_OAIS_Archival_Metadata_v1.0_final_20110921.pdf
- GeoMAPP. (2011). *Best Practices for Geospatial Data Transfer for Digital Preservation*. Extrait de Geospatial Multistate Archive and Preservation Partnership: Publications and Tools : http://www.geomapp.net/docs/Geo_Data_Transfer_BestPractices_v1.0_final_20111201.pdf
- GeoMAPP. (2011). *Geoarchiving Business Planning Toolkit*. Extrait de GeoMAPP: Publications and Tools: Geospatial Archival Business Planning : www.geomapp.net/docs/00_Geoarchiving_Business_Toolkit_20111231.zip
- GeoMAPP. (2011). *New Partner Mentoring: Introduction to Archival Appraisal*. Extrait de Geospatial Multistate Archive and Preservation Partnership : http://www.geomapp.net/docs/Appraisal_mentoring_presentation_final_20110610.pdf
- GeoMAPP. (2011). *Publications and Tools*. Extrait le 18 février 2013 de Geospatial Multistate Archive and Preservation Partnership : http://www.geomapp.net/publications_categories.htm
- GeoMAPP. (2011a). *Geoarchiving Business Planning Guidebook*. Extrait le 18 janvier 2013 de Geospatial Multistate Archive and Preservation Partnership Publications and Tools: http://www.geomapp.net/publications_categories.htm
- GeoMAPP. (2011b). *Best Practices for Archival Processing for Geospatial Datasets*. Extrait le 18 janvier 2013 de Geospatial Multistate Archive and Preservation Partnership Publications and Tools: http://www.geomapp.net/publications_categories.htm
- GeoMAPP. (2011c). *Geospatial Multistate Archive and Preservation Partnership Publications and Tools*. Extrait le 21 janvier 2013 de http://www.geomapp.net/docs/GIS_Archival_Processing_Process_v1.0_final_20111102.pdf
- GeoMAPP. (2011d). *GeoMAPP Key Findings and Best Practices*. Extrait le 21 janvier 2013 de Geospatial Multistate Archive and Preservation Partnership Project Documents : http://www.geomapp.net/docs/GeoMAPP_ProjectFindings_BestPractices20111231.pdf
- GeoMAPP. (2012). *Geoarchiving Business Cost-Benefit Analysis Guidance*. Extrait le 21 janvier 2013 de Geospatial Multistate Archive and Preservation Partnership Publications and Tools : http://www.geomapp.net/docs/03_Geoarchiving_Cost-Benefit_Analysis_Guidance_20111231.pdf
- GTNO. (1993). *NWT Archives Policy*. Extrait le 11 mars 2013 du gouvernement des Territoires du Nord-Ouest : <http://www.pwnhc.ca/programs/downloads/NWTArchivesPolicy.pdf>

- Hoebelheinrich, N. (2009). *An Investigation Into Metadata for Long-Lived Geospatial Data Formats*. Extrait le 11 février 2013 de Library of Congress: National Geospatial Digital Archive Project : www.digitalpreservation.gov/meetings/documents/ndiipp08/session7_hoebelheinrich_paper.doc
- Hoebelheinrich, N., & Munn, N. K. (2009). *Assessing the Utility of Current Format Registry Efforts for Geospatial Formats*. Extrait le 11 mars 2013 de The National Geospatial Digital Archive, Archiving 2009 : Final Program and Proceedings. The National Geospatial Digital Archive : http://www.ngda.org/docs/Pub_Hoebelheinrich_Arch09_09.pdf
- InterPARES 2. (2002). *Requirements for Assessing and Maintaining the Authenticity of Electronic Records*. Extrait le 18 février 2013 du projet InterPARES 2 : www.interpares.org/book/interpares_book_k_app02.pdf
- InterPARES 2. (2007). *Creator Guidelines : Making and Maintaining Digital Materials: Guidelines for Individuals*. Extrait le 10 février 2013 du projet InterPARES 2 : http://www.interpares.org/public_documents/ip2%28pub%29creator_guidelines_booklet.pdf
- InterPARES 2. (2007). *Preserver Guidelines - Preserving Digital Records: Guidelines for Organizations*. Extrait le 10 février 2013 des produits du projet InterPARES 2 : http://www.interpares.org/ip2/display_file.cfm?doc=ip2%28pub%29preserver_guidelines_booklet.pdf
- InterPARES 2. (2008). *Policy Framework - A framework of principles for the development of policies, strategies and standards for the long-term preservation of digital records*. Extrait le 18 février 2013 du projet InterPARES 2 : [www.interpares.org/ip2/display_file.cfm?doc=ip2\(pub\)policy_framework_document.pdf](http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)policy_framework_document.pdf)
- InterPARES 2. (2013). *Dictionary definitions*. Extrait le 24 janvier 2013 de la base de données terminologique du projet InterPARES 2 : http://www.interpares.org/ip2/ip2_terminology_db.cfm
- ISO. (2003). *ISO 19115:2003 Information géographique -- Métadonnées*. Extrait le 11 février 2013 du site Web de l'Organisation internationale de normalisation à http://www.iso.org/iso/fr/home/store/catalogue_tc/catalogue_detail.htm?csnumber=26020
- ISO. (2005). *ISO/IEC 17799:2005 Technologies de l'information – Techniques de sécurité -- Code de bonne pratique pour la gestion de la sécurité de l'information*. Extrait du site Web de l'Organisation internationale de normalisation à http://www.iso.org/iso/fr/home/store/catalogue_tc/catalogue_detail.htm?csnumber=39612
- ISO. (2005). *ISO 19128:2005 Information géographique – Interface de carte du serveur Web*. Extrait le 10 mars 2013 du site Web de l'Organisation internationale de normalisation à http://www.iso.org/iso/fr/home/store/catalogue_tc/catalogue_detail.htm?csnumber=32546

- ISO. (2012). *ISO 16363:2012, Systèmes de transfert des informations et données spatiales -- Audit et certification des référentiels numériques de confiance*. Extrait le 10 février 2013 du site Web de l'Organisation internationale de normalisation à http://www.iso.org/iso/fr/home/store/catalogue_tc/catalogue_detail.htm?csnumber=56510
- JC. (2007). Ministère de la Justice. *Règlement sur les cartes marines et les publications nautiques (1995) (DORS/95-149)*. Extrait le 18 février 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/reglements/DORS-95-149/>
- JC. (2007). Ministère de la Justice. *Loi sur les systèmes de télédétection spatiale (L.C. 2005, ch. 45)*. Extrait le 18 février 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/R-5.4/TexteComplet.html>
- JC. (2007). Ministère de la Justice. *Règlement sur les systèmes de télédétection spatiale, DORS/2007-66*. Extrait le 18 février 2103 du Site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/reglements/DORS-2007-66/TexteComplet.html>
- JC. (2010). Ministère de la Justice. *Loi sur la prévention de la pollution des eaux arctiques (L.R.C. (1985), ch. A-12)*. Extrait le 18 février 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/A-12/>
- Lauriault, T. P., Craig, B., Pulsifer, P. L., & Taylor, D. R. (2008). *Today's Data are Part of Tomorrow's Research: Tomorrow's Research: Archival Issues in the Sciences*. Archivaria, p. 123-179.
- Library of Congress. (2013). *PREMIS Data Dictionary for Preservation Metadata*. Extrait le 11 février 2013 de Library of Congress : Standards : <http://www.loc.gov/standards/premis/>
- Library of Congress. (2013). *Sustainability of Digital Formats Planning for Library of Congress Collections*. Extrait le 11 mars 2013 de Library of Congress Digital Preservation : <http://www.digitalpreservation.gov/formats/index.shtml>
- LTDP. (2012). *Long Term Data Preservation Earth Observation Preserved Data Set Content LTDP/PDSC*. Extrait de l'introduction du GCSB : http://earth.esa.int/gscb/ltdp/LTDP_PDSC_4.0.pdf
- LTDP Working Group. (2012). *Long Term Preservation of Earth Observation Space Data: European LTDP Common Guidelines*. Extrait le 2 janvier 2013 du Ground Segment Coordination Body (GSCB) : http://earth.esa.int/gscb/ltdp/EuropeanLTDPCommonGuidelines_Issue2.0.pdf
- MacNeil, H. (2000). *Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Term Preservation of Authentic*. Archivaria, p. 52-78.

- Ministère de la Justice. (1985). *Loi sur l'accès à l'information, L.R.C., 1985, c. A-1*. Extrait le 11 mars 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/A-1/TexteComplet.html>
- Ministère de la Justice. (2004). *Loi constituant Bibliothèque et Archives du Canada, modifiant la Loi sur le droit d'auteur et modifiant certaines lois en conséquence (L.C. 2004, ch. 11)*. Extrait le 11 mars 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/L-7.7/TexteComplet.html>
- Ministère de la Justice. (1985). *Loi sur la preuve au Canada (L.R.C. 1985, ch. C-5)*. Extrait le 11 mars 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/C-5/TexteComplet.html>
- Ministère de la Justice. (1985). *Loi sur le droit d'auteur, L.R.C. (1985), ch. C-42*. Extrait le 11 mars 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/C-42/TexteComplet.html>
- Ministère de la Justice. (2000). *Loi sur la protection des renseignements personnels et les documents électroniques, L.C., 2000, ch. 5*. Extrait le 11 mars 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/P-8.6/TexteComplet.html>
- Ministère de la Justice. (1985). *Loi sur la protection des renseignements personnels L.R.C. (1985), ch. P-21*. Extrait le 11 mars 2013 du site Web de la législation (Justice) : <http://laws-lois.justice.gc.ca/fra/lois/P-21/TexteComplet.html>
- MPO. (2001). Politique de gestion des données scientifiques. Extrait le 3 janvier 2013 de Pêches et Océans Canada : <http://www.dfo-mpo.gc.ca/science/data-donnees/policy-politique-fra.htm>
- MPO. (2012). Gestion des données scientifiques intégrées. Extrait le 4 janvier 2013 de Pêches et Océans Canada : <http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/index-fra.html>
- MPO. (2012). *Centre national responsable des données océanographiques*. Extrait le 23 janvier 2013 de Pêches et Océans Canada : <http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/rnodc-cnrdo/index-fra.htm>
- MPO. (2013). Bouées dérivantes. Extrait le 23 janvier 2013 de Pêches et Océans Canada : <http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/drib-bder/index-fra.htm>
- MRNO. (2010). Ministère des Richesses naturelles de l'Ontario. *Guideline for Retiring and Retaining Geospatial Data Stored in the Land Information Ontario Warehouse*. Province de l'Ontario.
- MRNO. (2011). *Les métadonnées*. Extrait le 11 mars 2013 du ministère des Richesses naturelles de l'Ontario, Information sur les terres de l'Ontario : <http://www.mnr.gov.on.ca/fr/Business/LIO/2ColumnSubPage/266895.html>

- MRNO. (2011). *Promouvoir l'excellence dans la conservation des documents*. Extrait le 11 mars 2013 du gouvernement de l'Ontario : <http://www.archives.gov.on.ca/fr/recordkeeping/index.aspx>
- NC CGIA. (2006). *Survey Frequency of Geospatial Data Capture*. Extrait le 11 mars 2013 de Library of Congress North Carolina Geospatial Data Archiving Project: <http://www.digitalpreservation.gov/partners/ncgdap.html>
- NGDA. (2009). *Collection Development Policies and Contracts*. Extrait le 11 mars 2013 du National Geospatial Digital Archive des É.-U. : <http://www.ngda.org/policies.html>
- NRC. (2010). *DataCite Canada*. Extrait le 11 mars 2013 du Conseil national de recherches Canada : <http://cisti-icist.nrc-cnrc.gc.ca/fra/services/icist/datacite-canada/apropos.html>
- NSIDC. (2013). *IPY at NSIDC*. Extrait le 4 janvier 2013 du National Snow & Ice Data Centre : <http://nsidc.org/ipy/>
- OCLC et CRL. (2007). *Trustworthy Repositories Audit & Certification: Criteria and Checklist, Version 1.0*. Extrait le 18 février 2013 du Centre for Research Libraries, Metrics for Repository Assessment : www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf
- Parsons, M., de Bruin, T., Tomlinson, S., Campbell, H., Godoy, O., & LeClert, J. (2011). *Chapter 3,11 The State of Polar Data: The IPY Experience, PART THREE: IPY Observing Systems, Their Legacy and Data Management*. Extrait le 4 janvier 2013 du Conseil international pour la science : <http://www.icsu.org/publications/reports-and-reviews/ipy-summary/ipy-jc-summary-part3.pdf>
- Prince of Wales Northern Heritage College. (2013). Calendriers spécifiques de conservation et d'élimination des documents d'archives. Extrait le 11 mars 2013 du Prince of Wales Northern Heritage College : http://www.pwnhc.ca/programs/archives/record_schedules/Departments/record_schedules.asp
- QA4EO. (2013). *Documentation*. Extrait le 18 février 2013 de Quality Assurance Framework for Earth Observation : <http://qa4eo.org/documentation.html>
- RCIP. (2012). *Purpose du RCIP*. Extrait le 4 janvier 2013 du Réseau canadien d'information polaire : http://polardatanetwork.ca/?page_id=59
- RLG et OCLC. (2002). *Trusted Digital Repositories: Attributes and Responsibilities*. Extrait le 18 février 2013 de Attributes of Trusted Digital Repositories : <http://www.oclc.org/research/activities/trustedrep.html>
- RLG et OCLC. (2002). *Trusted Digital Repositories: Attributes and Responsibilities*. RLG-OCLC.
- RNCan. (2010). Ressources naturelles Canada. Extrait le 18 février 2013 de la Liste des lois et des règlements : <http://www.rncan.gc.ca/lois-reglements/857>

- Roeder, J., Eppard, P., Underwood, B., & Lauriault, T. P. (2008). *Part 3 : Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Government Sectors, Domain 2 Task Force Report*. Extrait le 18 février 2013, du pProjet International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2 : Experiential, Interactive and Dynamic Records Book :
http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf
- SAA. (2013). *Glossary of Archival Records and Terminology*. Extrait le 10 février 2013 de The Society of American Archivists : www2.archivists.org/glossary/terms/a
- SCT. (2012). *Norme sur les données spatiales*. Extrait le 11 février 2013 de la suite des politiques du Conseil du Trésor <http://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=16553§ion=text>
- Sous-comité canadien de la gestion des données de l'API. (2008). *Politique canadienne de gestion des données pour l'Année polaire internationale (API) 2007-2008*. Extrait le 4 janvier 2013 du site Web de l'Année polaire internationale à
http://www.api-ipy.gc.ca/pg_IPYAPI_055-fra.html
- Stone, A., & Day, M. (1999). *Cedars Preservation Metadata Elements: Cedars Project Document AIW02*. Extrait le 11 février 2013 de Cedars Access Issues Working Group :
<http://www.ukoln.ac.uk/metadata/cedars/papers/aiw02/>

B. Glossaire des termes

Sources du glossaire :

- Digital Curation Centre (DCC) Glossary : <http://www.dcc.ac.uk/digital-curation/glossary>
- Digital Preservation Coalition (DPC) :
<http://www.dpconline.org/advice/preservationhandbook/introduction/definitions-and-concepts>
- Geospatial Multistate Archiving and Preservation Project (GeoMAPP)
- Base de données terminologique du projet InterPARES 2 (IP2) :
http://www.interpares.org/ip2/ip2_terminology_db.cfm
- Minnesota State Archives (MSA) :
<http://www.mnhs.org/preserve/records/electronicrecords/erglossary.html#c>
- Système ouvert de l'archivage d'informations (SOAI) :
<http://public.ccsds.org/publications/archive/650x0m2.pdf>
- Quality Assurance Framework for Earth Observation (QA4EO) : <http://qa4eo.org>

Acronyme	Terme	Définition
	Accès	Facilité d'utilisation continue, en cours d'une ressource numérique, tout en conservant toutes les qualités d'authenticité, d'exactitude et de fonctionnalité considérées comme essentielles aux fins pour lesquelles le matériel numérique a été créé et, ou acquis. (DPC)
	Accessibilité	La disponibilité et la facilité d'utilisation de l'information. (IP2)
	Actualiser	Copier le contenu d'information d'un support de stockage au même support de stockage. (DPC)
	Archives	Voir Centre d'archives
	Authentification sans recours à la technologie	L'authentification de documents d'archives axée sur l'utilisation des procédures administratives pour établir une présomption d'authenticité ou, au besoin, la vérification de l'authenticité, notamment par comparaison des données compilées sur l'identité et l'intégrité d'un document et les contrôles procéduraux exercés durant sa création, son utilisation, son maintien ou sa préservation avec les exigences de documents d'archives authentiques. (IP2)
	Base de connaissances	Un ensemble d'information, constitué par une personne ou un système, qui permet à cette personne ou à ce système de comprendre l'information reçue (p. ex., une personne dont la base de connaissances inclut une compréhension de l'anglais sera en mesure de lire et de comprendre un texte en anglais). (DCC)

Acronyme	Terme	Définition
	Centre d'archives	Organisme ou établissement chargé de préserver et de communiquer des documents d'archives sélectionnés préservés en permanence. (IP2)
	Communauté d'utilisateurs désignée	Groupe ciblé de consommateurs potentiels qui devraient être en mesure de comprendre une série de données particulière. La communauté désignée peut être composée de plusieurs communautés d'utilisateurs. (SOAI)
	Composantes numériques	Un objet numérique qui fait partie d'un ou de plusieurs documents d'archives numériques, y compris les métadonnées requises pour organiser, structurer ou présenter son contenu et sa forme, nécessitant une mesure de préservation donnée. (IP2)
	Consommateur	Le rôle joué par les personnes ou les systèmes clients qui interagissent avec les services du SOAI pour trouver des informations préservées d'intérêt et accéder à ces informations en détail. (SOAI)
CRC	Contrôle de redondance cyclique (CRC)	Un code de détection des erreurs le plus utilisé dans les réseaux numériques et les périphériques de stockage afin de détecter toute modification accidentelle des données brutes. Une valeur de contrôle courte est attachée aux données saisies dans ces systèmes, en fonction du reste d'une division polynomiale de leurs contenus; lors de l'extraction le calcul est répété, et des mesures correctives peuvent être prises contre la corruption de données présumée si les valeurs de contrôle ne correspondent pas. (SOAI)
	Conditions de référence en matière d'authenticité	Les exigences qui servent de base à l'évaluation du préservateur de l'authenticité des documents d'archives numériques d'un créateur durant l'évaluation. (IP2)
	Description connexe	L'information décrivant le contenu d'une trousse d'information du point de vue du logiciel ou du document qui permet aux consommateurs de repérer, d'analyser, de commander ou d'extraire des informations des archives. (SOAI)
	Description de la trousse	L'information destinée à être utilisée par l'instrument d'accès. (SOAI)
	Direction	Le rôle joué par ceux qui définissent les politiques d'archivage dans leur ensemble comme un élément dans un domaine stratégique plus vaste, par exemple dans le cadre d'une plus grande organisation. (SOAI)
	Document actif	Un document requis par le créateur dans le but de mener à bien l'action pour laquelle il a été créé ou pour référence fréquente. (IP2)

Acronyme	Terme	Définition
	Données de contrôle	Un décompte du nombre de bits dans une unité de transmission qui est inclus avec l'unité de sorte que le récepteur peut vérifier pour voir si le même nombre de bits est arrivé. Si le décompte correspond, la transmission complète est présumée avoir été reçue. (MSA)
	Données sur la représentation	L'information qui cartographie un objet de données à l'intérieur de concepts plus importants. Un exemple est un logiciel JPEG, qui incarne une compréhension de la norme JPEG et cartographie les bits en pixels qui peuvent ensuite être restitués sous forme d'image aux fins de visualisation. (SOAI)
	Émulation	Un moyen de surmonter l'obsolescence technologique du matériel et des logiciels en développant des techniques pour imiter les systèmes obsolètes sur les générations futures d'ordinateurs. (DPC)
	Entente de commande	Une entente entre les archives et le consommateur dans laquelle les détails physiques de la livraison, tels que le type de support et le format des données, sont précisés. (DCC)
	Entente de dépôt	L'entente conclue entre les archives et le créateur qui spécifie un modèle de données, et toutes les autres dispositions nécessaires au dépôt de données. Ce modèle de données détermine le format/contenu et les constructions logiques utilisés par le créateur et la façon dont ils sont représentés sur chaque support ou dans une session de télécommunication. (SOAI)
	Évaluation	Processus consistant à évaluer la valeur des documents d'archives pour en déterminer la durée et les conditions de conservation. (IP2)
	Exactitude	Le degré auquel les données, l'information ou les documents d'archives sont précis, corrects, honnêtes, libres d'erreurs ou de distorsion, ou pertinents à l'affaire. (IP2)
	Fiabilité	La fiabilité d'un document comme un état de fait. Elle existe quand un document peut soutenir ce qu'il est, et est établie en examinant l'intégralité de la forme du document et le degré de contrôle exercé sur le processus de sa création. (IP2)
	Format logique	Agencement organisé de données sur un support électronique qui permet la reconnaissance et la récupération des structures de contrôle des fichiers et des données par le système d'exploitation de l'ordinateur hôte. Deux formats logiques courants utilisés pour les fichiers et les répertoires sont le format ISO 9660/13490 pour les CD-ROM, et le Universal Disk Format (UDF) pour les DVD. (IP2)
	Forme	Règles de la représentation qui déterminent l'apparence d'une entité et transmettent son sens. (IP2)

Acronyme	Terme	Définition
	Forme fixe	La qualité d'un document qui garantit que son contenu reste complet et inchangé. (IP2)
	Géoarchivage	La détermination des données géospatiales importantes et leur préservation pour une utilisation future. (GeoMAPP)
	Gestion des droits électroniques	Un ensemble de systèmes utilisés pour protéger les droits d'auteur des supports électroniques. (TechTerms.com)
	Gestionnaire de confiance	Préservateur qui est en mesure de démontrer qu'il n'a aucune raison de modifier les documents d'archives préservés ou de laisser quiconque les modifier, et qui est en mesure de faire observer toutes les exigences de la préservation de copies authentiques. (IP2)
IQ	Indicateur de la qualité	Une mesure qui fournit des informations suffisantes pour permettre à tous les utilisateurs d'évaluer facilement la pertinence d'un produit de données pour leur application particulière (c.-à-d. sa « pertinence d'utilisation »). Un IQ doit être fondé sur une évaluation quantifiable de preuves démontrant le niveau de traçabilité aux normes de référence acceptées par la communauté internationale (le système international d'unités (SI), dans la mesure du possible). (QA4EO)
	Information descriptive	L'ensemble des informations, comprenant principalement des descriptions de la trousse, qui est fourni à la gestion des données pour soutenir la recherche, les commandes et la récupération des banques de renseignements du SOAI par les consommateurs. (SOAI)
	Information sur la trousse	L'information qui est utilisée pour lier et déterminer les éléments d'une trousse d'information. Par exemple, cela peut être le volume ISO 9660 et l'information de répertoire utilisée sur un CD-ROM pour fournir le contenu de plusieurs fichiers contenant des renseignements sur le contenu et de l'information sur les descriptions de préservation. (DCC)
IDP	Information sur les descriptions de préservation	L'information qui est nécessaire pour une préservation adéquate des renseignements sur le contenu et qui peut être classée comme de l'information sur la provenance, la référence, la fixité, le contexte et les droits d'accès. (SOAI)
	Instrument de recherche	Un type d'instrument d'accès qui permet à un utilisateur de rechercher et de déterminer des trousseaux d'information archivée d'intérêt. (SOAI)
	Métadonnées	Informations décrivant des aspects importants d'une ressource qui sont nécessaires pour réussir à gérer et à préserver les documents d'archives numériques au fil du temps et qui aideront à garantir que l'information contextuelle, historique et technique essentielle est préservée avec la ressource numérique. (DPC)

Acronyme	Terme	Définition
	Migration	Un moyen de surmonter l'obsolescence technologique en transférant des ressources numériques d'une génération de matériel/logiciel à la génération suivante. (DPC)
	Objet de données	Un objet numérique composé d'un ensemble de séquences de bits. (SOAI)
	Objet d'information	Un objet de données avec ses données de représentation. (SOAI)
	Préservation à court terme	L'accès aux documents d'archives numériques soit pour une période de temps définie tandis que l'utilisation est prévue, mais qui ne s'étend pas au-delà d'un avenir prévisible, ou jusqu'à ce qu'ils deviennent inaccessibles en raison de l'évolution des technologies. (DPC)
	Préservation à long terme	Un accès continu au matériel numérique, ou tout au moins à l'information qu'il contient, de manière indéfinie. (DPC)
	Préservation à moyen terme	Un accès continu au matériel numérique au-delà de l'évolution des technologies pour une période de temps définie, mais pas indéfiniment. (DPC)
	Producteur	Le rôle joué par les personnes ou les systèmes clients qui fournissent l'information qui doit être préservée. (SOAI)
	Reformatage	Copier le contenu d'information d'un support de stockage à un support de stockage différent (reformatage du support) ou convertir d'un format de fichier en un format de fichier différent (reformatage d'un fichier). (DPC)
	Renseignements sur le contenu	Un ensemble d'informations qui est la cible d'origine de préservation ou qui comprend tout ou une partie de ces informations. C'est un objet d'information composée de son objet de contenu en données et ses données sur la représentation. (SOAI)
SOAI	Système ouvert d'archivage d'information	Un centre d'archives, qui consiste en une organisation, qui peut faire partie d'une plus grande organisation, de personnes et de systèmes, qui a accepté la responsabilité de préserver l'information et la rendre disponible à une communauté d'utilisateurs. (SOAI)
	Système de préservation des documents d'archives	Série de règles qui régissent le maintien intellectuel et physique permanent des documents d'archives acquis, et les outils et les mécanismes servant à l'application de ces règles. (IP2)
	Système de préservation fiable	L'ensemble des règles qui régissent la préservation et l'utilisation des documents d'archives du créateur et offrent une probabilité circonstancielle de l'authenticité des documents d'archives, ainsi que les outils et les mécanismes utilisés pour mettre en œuvre ces règles. (IP2)

Acronyme	Terme	Définition
TIA	Trousse d'information archivée	Une trousse d'information composée des renseignements sur le contenu et de l'information sur les descriptions de préservation connexe (IDP), qui est préservée dans un SOAI. (SOAI)
TDI	Trousse de diffusion de l'information	Une trousse d'information, dérivée d'un ou plusieurs TIA, et envoyée par les d'archives au consommateur, en réponse à une demande au SOAI. (SOAI)
	Trousse d'Information	Un conteneur logique composé de renseignements sur le contenu optionnels et d'information descriptive sur la préservation connexe optionnelle. Liée à cette trousse d'information se trouve l'information sur les trousse utilisées pour délimiter et déterminer les renseignements sur le contenu et l'information descriptive sur la trousse utilisée pour faciliter les recherches pour les renseignements sur le contenu. (SOAI)
TIS	Trousse d'information sur les soumissions	Une trousse d'information qui est livrée par le créateur au SOAI pour utilisation dans la mise sur pied ou la mise à jour d'un ou de plusieurs TIA ou l'information descriptive connexe. (SOAI)
	Variabilité limitée	Les modifications apportées à la forme ou au contenu d'un document numérique qui sont limitées et contrôlées par des règles fixes, de sorte que la même requête, demande ou interaction génère toujours le même résultat. (IP2)
	Versement	Action d'accepter la garde légale et physique d'un recueil de documents d'archives et de les inscrire dans un registre. (IP2)

C. Lois et règlements propres à la préservation et à la gestion de l'information du gouvernement

Lois	Obligations	Limitations/Défis
Loi sur la Bibliothèque et les Archives du Canada (2004, ch. 11)	<ul style="list-style-type: none"> ▪ Gestion des données en vue d'un versement possible ▪ Pas de disposition des données sans consentement écrit ▪ Communication des données recueillies dans un but de recherche sur l'opinion publique ▪ Dépôt légal de deux exemplaires de chaque publication ▪ Dépôt des données réputées constituer un patrimoine documentaire 	<ul style="list-style-type: none"> ▪ Comment les ensembles de données modifiés dynamiquement sont traités comme des documents d'archives ▪ Traitement d'ensembles de données avec la contribution de plusieurs créateurs ▪ Traitement des cartes numériques dynamiques/interactives générées au moyen de services Web ▪ Mécanismes d'archivage et contrôles répartis ▪ Préservation de l'information numérique accessible par des portails Web ▪ Soutien institutionnel limité aux archives distribuées
Règlement sur le dépôt légal de publications (DORS/2006-337)	<ul style="list-style-type: none"> ▪ Directive sur la présentation des publications numériques 	<ul style="list-style-type: none"> ▪ Le système n'est pas propice au dépôt de données
Loi sur le droit d'auteur (L.R.C., 1985, ch. C-42)	<ul style="list-style-type: none"> ▪ Les droits et responsabilités en matière de propriété et d'utilisation des données demeurent en vigueur pour les données archivées 	<ul style="list-style-type: none"> ▪ Nécessité de gérer les droits d'utilisation et d'accès
Loi sur l'accès à l'information (L.R.C., 1985, ch. A-1)	<ul style="list-style-type: none"> ▪ Les restrictions d'accès aux données demeurent en vigueur pour les données archivées. 	<ul style="list-style-type: none"> ▪ Nécessité de gérer les droits d'utilisation et d'accès
Loi sur la protection des renseignements personnels (L.R.C., 1985, ch. P-21)	<ul style="list-style-type: none"> ▪ La divulgation des renseignements personnels à BAC est permise dans un but archivistique 	<ul style="list-style-type: none"> ▪ Il peut être nécessaire de regrouper ou de rendre anonymes les données pour en permettre l'accès

Lois	Obligations	Limitations/Défis
Règlement sur la protection des renseignements personnels, DORS/83-508	<ul style="list-style-type: none"> ▪ Directive sur la gestion des renseignements personnels archivés. 	<ul style="list-style-type: none"> ▪ Délais prescrits pour l'accès aux renseignements personnels dans un but de recherche ou un but statistique
Loi sur la protection des renseignements personnels et les documents électroniques [L.C. 2000, ch. 5]	<ul style="list-style-type: none"> ▪ S'applique aux renseignements personnels recueillis, utilisés et communiqués par des organisations du secteur privé 	
Loi sur la preuve au Canada (L.R.C., 1985, ch. C-5)	<ul style="list-style-type: none"> ▪ Établir l'authenticité des données archivées présentées en preuve (métadonnées robustes, modification du suivi au fil du temps, systèmes fonctionnant adéquatement, gestion et entretien des systèmes afin d'assurer l'authenticité des documents d'archives qu'ils contiennent - protocoles de sécurité, protocoles d'accès, etc.) 	

D. Politiques et directives du Secrétariat du Conseil du Trésor (SCT) sur la gestion de l'information

Objectifs	Obligations
Cadre stratégique pour l'information et la technologie (2007)	
<ul style="list-style-type: none"> ▪ Fournit des lignes directrices aux pratiques saines de gestion de l'information et de la technologie au sein du gouvernement ▪ Principes directeurs importants <ul style="list-style-type: none"> ○ Administration – les données doivent être gérées avec rigueur tout au long de son cycle de vie, peu importe le support ou le format 	<ul style="list-style-type: none"> ▪ Les créateurs sont responsables de gérer les documents d'archives ▪ La législation et la réglementation du tableau 1 doivent faire partie du cycle de vie d'un document ▪ La gestion efficace de l'information dans une organisation, de la planification et de l'élaboration des systèmes à l'élimination ou à la préservation à long terme
Politique sur la gestion de l'information (Politique sur la GI de 2007)	
<ul style="list-style-type: none"> ▪ Afin de réaliser une GI efficace et efficiente : <ul style="list-style-type: none"> ○ appuyer la prestation de programmes et de services; ○ assurer la mise en place de processus décisionnels efficaces; ○ faciliter la reddition des comptes, la transparence et la collaboration; ○ préserver l'information et veiller à l'accès à l'information et aux documents d'archives pour le bien de la génération actuelle et des générations à venir. 	<ul style="list-style-type: none"> ▪ Les données géospatiales doivent faire partie d'une stratégie intégrée de gestion de l'information ainsi que d'autres informations ▪ Si les données géospatiales sont utilisées pour soutenir les activités, les politiques et les programmes, il est important de veiller à ce qu'elles soient maintenues dans le cadre d'un ensemble de documents d'archives. Les métadonnées, les méthodes de description de collecte de données, les paramètres sur la qualité des données, l'information contextuelle et tous les autres attributs jugés nécessaires par les créateurs et les responsables doivent aider à la compréhension. Si elles sont utilisées dans le cadre d'une expérience, alors les paramètres de l'expérience doivent également accompagner l'information géospatiale. ▪ Création de mécanismes d'accès aux données et d'explorations de ces dernières. ▪ Veiller à ce que les données soient gérées de façon à respecter les ententes d'utilisation, les conditions applicables aux licences, ou les deux, et à assurer la pertinence, l'authenticité, la qualité et le rapport coût-efficacité de l'information, et cela aussi longtemps que cela est nécessaire pour répondre aux besoins opérationnels et respecter les obligations redditionnelles; ▪ Assure la participation des ministères à la définition des orientations pangouvernementales en matière de gestion de l'information et de tenue de dossiers.

Objectifs	Obligations
Politique sur la gestion des technologies de l'information (2007).	
<ul style="list-style-type: none"> Assurer l'utilisation efficiente et efficace des technologies de l'information pour appuyer les priorités du gouvernement et la prestation de programmes, accroître la productivité et améliorer les services au public. 	<ul style="list-style-type: none"> Cette politique est très limitée Ce qui suit est toutefois recommandé, même si cela ne fait pas partie de la politique : <ul style="list-style-type: none"> l'adoption de spécifications et de normes communes en ce qui a trait à l'accès, l'interopérabilité, l'architecture ouverte, les logiciels au code source libre, les formats de données, les métadonnées et la gestion peut permettre hétérogénéité des TI, tout en assurant l'interopérabilité et, espérons-le, la longévité.
Directive sur les rôles et responsabilités en matière de gestion de l'information (2007)	
<ul style="list-style-type: none"> Déterminer les rôles et les responsabilités de tous les employés des ministères pour appuyer l'administrateur général dans la gestion efficace de l'information (données) au sein de leur ministère. 	<ul style="list-style-type: none"> Veille à ce que les directives, les processus et les outils de gestion appropriés soient en place pour conserver la qualité des données tout au long du cycle de vie de l'information : la planification, la collecte, la création, la réception et la saisie des données, leur organisation, leur utilisation et leur diffusion, leur maintien, leur protection et leur préservation; leur élimination, et leur évaluation. Décisions relatives à la conservation
Directive sur la tenue de documents (2009)	
<ul style="list-style-type: none"> Assurer l'adoption de pratiques efficaces de tenue de documents d'archives qui permettent aux ministères de créer, d'acquérir, de saisir, de gérer et de protéger l'intégrité des ressources d'information (données) ayant une valeur opérationnelle relativement à l'exécution des programmes et à la prestation des services du GC. 	<ul style="list-style-type: none"> Veille à ce que des méthodes, des mécanismes et des outils soient en place pour répondre aux exigences ministérielles en matière de tenue de documents d'archives, tout au long du cycle de vie des données. Déterminent, établissent, mettent en œuvre et maintiennent des dépôts dans lesquels les données ayant une valeur opérationnelle sont archivées ou préservées, que ce soit dans un lieu physique ou dans un environnement électronique. Établissent, utilisent et tiennent à jour des taxonomies ou des structures de classification visant à faciliter l'archivage, la recherche et la récupération des données ayant une valeur opérationnelle, peu importe le format. Établissent, mettent en œuvre et maintiennent des délais de préservation des données ayant une valeur opérationnelle, le cas échéant, en fonction du format. Développent et mettent en œuvre un processus d'élimination documenté pour les données. Exécutent les activités périodiques d'élimination pour toutes les données.

Objectifs	Obligations
Norme pour les systèmes de gestion électronique des documents et des dossiers (SGEDD) du SCT (2010)	
<ul style="list-style-type: none"> ▪ Appuyer la gestion efficace et efficiente de l'information grâce à l'utilisation de normes pour les SGEDD et réduire le coût global grâce à la normalisation et aux économies d'échelle 	<ul style="list-style-type: none"> ▪ Les systèmes automatisés : <ul style="list-style-type: none"> ○ gèrent, protègent et préservent les données depuis leur création jusqu'à leur élimination; ○ maintiennent de l'information contextuelle pertinente (métadonnées); ○ permettent aux organisations d'accéder aux documents d'archives, de les utiliser et d'en disposer (c.-à-d. leur conservation, leur destruction ou leur transfert) d'une manière gérée, systématique et vérifiable.
Norme sur les métadonnées (2010)	
<ul style="list-style-type: none"> ▪ Accroître l'utilisation des métadonnées normalisées et des domaines de valeur à l'appui de la gestion des ressources d'information : <ul style="list-style-type: none"> ○ améliorer l'accessibilité des données, le partage, l'authenticité, la fiabilité et l'intégrité dans tous les ministères; ○ augmenter la capacité des programmes et des services de partager des données de manière efficace et efficiente entre les systèmes et entre les ministères. 	<ul style="list-style-type: none"> ▪ Veille à l'adoption et à la mise en œuvre de normes sur les métadonnées.
Norme sur les données géospatiales (2009)	
<ul style="list-style-type: none"> ▪ Soutenir l'administration et l'interopérabilité des données en veillant à ce que les ministères accèdent aux données géospatiales, les utilisent et les partagent pour appuyer l'exécution des programmes et la prestation des services. 	<ul style="list-style-type: none"> ▪ Veille à l'application de la norme <i>ISO 19115 Information géographique – Métadonnées</i> ▪ Veille à l'application de la norme <i>ISO 19128 – Information géographique – Interface de carte du serveur Web</i>.

E. Faits saillants du Geoarchiving Business Planning Guidebook

Cette directive a été élaborée par le Geospatial Multistate Archive Preservation Partnership (GeoMAPP) pour aider l'État et les autres utilisateurs à évaluer leur situation et leurs besoins actuels, déterminer les coûts et les avantages d'un programme de géoarchivage, et fournir les outils et le cadre opérationnel afin que les organisations réussissent à concevoir, créer et mettre en œuvre un programme de géoarchivage. Le guide préconise une approche selon laquelle les plans sont assemblés grâce à un processus itératif de sensibilisation animé par des discussions de groupe, la recherche, la rédaction et la révision, y compris la participation ou l'orientation des principaux intervenants dans le SIG, l'organisation d'archivage, et le personnel des TI. Il utilise une liste de questions pour faciliter ces discussions et générer du contenu pour chaque section du plan. Le tableau qui suit résume les éléments « hautement recommandés » à l'intérieur de chaque section du guide.

Section du guide	Essence
SECTION PRINCIPALE	
1. Résumé	Quels résultats proposez-vous d'obtenir? Pourquoi avez-vous besoin de le faire?
2. Autoévaluation des géoarchives	Quelles sont les conditions et quels sont les actifs actuels?
3. Clients et intervenants	Qui est visé et qui réalise le cas?
4. Objectifs du programme	Quels sont les « objectifs spécifiques du programme » pour ce plan d'activités? Pour chaque objectif, quels sont les « facteurs de succès » (ou objectifs à l'appui)?
5. Avantages et justification	Quelle est la principale raison faisant en sorte que vous devez faire ce que vous proposez? Quels avantages et quel retour sur investissement seront provisionnés si cela est fait?
6. Exigences et coûts	Quelle est l'approche de votre organisation? Quels sont les coûts totaux estimés de votre proposition?
7. Aperçu de la mise en œuvre	Application progressive et étapes clés Plan budgétaire
8. Mesurer le succès et suggestions de recalibrage	Établir les paramètres des coûts et des avantages et le processus de mise à jour/révision régulière
ANNEXES	
A. Carte et liste de contrôle sur le processus de planification opérationnelle de géoarchivage	Un résumé graphique d'ensemble du processus de planification des activités de géoarchivage et une liste de contrôle qui seront utilisés pendant le processus de planification

Section du guide	Essence
B. Document d'orientation d'analyse coûts-avantages détaillée	Un guide pour l'élaboration d'une analyse des besoins de stockage du projet, les coûts globaux du projet et les calculs coûts-avantages pour un projet de géoarchivage.
C. Trousse d'outils d'analyse coûts-avantages détaillée	Un classeur Excel pour aider l'utilisateur avec les besoins de stockage du projet, les coûts globaux du projet et créer des calculs coûts-avantages pour une période de temps, calculée en années, déterminée par l'utilisateur.
D. Directives et modèles pour soutenir la documentation opérationnelle	Un guide pour l'élaboration d'analyses de rentabilité liées au géoarchivage.
E. Bibliographie	Publications et documents de référence de géoarchivage.