



Conseil de recherches en  
sciences humaines du Canada

Social Sciences and Humanities  
Research Council of Canada

Canada

Archives nationales  
du Canada

National Archives  
of Canada



Rapport final

## **Consultation sur les archives nationales de données**

Infrastructure pour l'archivage et  
l'accessibilité des données de recherche  
du Canada

Présenté par le groupe de travail – Consultation  
sur les archives nationales de données (CAND)  
au Conseil de recherches en sciences humaines  
du Canada et à l'archiviste national du Canada

Juin 2002

## Résumé

En octobre 2000, le Conseil de recherches en sciences humaines et l'archiviste national du Canada ont constitué un groupe de travail comprenant des chercheurs et des archivistes et leur ont demandé d'évaluer les besoins en ce qui a trait à un système national d'accessibilité, de conservation et de gestion de données de recherche. Ayant compilé de nombreuses preuves à l'appui d'un tel service de soutien à la création de connaissances par la communauté des chercheurs œuvrant dans le domaine des sciences humaines, le groupe de travail présente maintenant ses recommandations concernant la création d'un nouveau service national d'archivage de données de recherche. Ce service aura trois fonctions de base :

- Conserver les données de recherche compilées par les chercheurs et les organismes gouvernementaux, les entreprises de sondage et autres organismes, qui peuvent être utilisées par les chercheurs pour créer de nouvelles connaissances;
- Gérer les données déposées et en assurer la qualité; sélectionner les données qui seront conservées, élaborer et appliquer des normes relatives aux métadonnées, à l'authentification et à la sécurité et diffuser les données dans l'ensemble des technologies;
- Fournir l'accès aux données de recherche, y compris aux systèmes de sortie Web, aux services de catalogage, aux ententes avec les utilisateurs et les dépositaires afin de protéger la confidentialité et les droits de propriété intellectuelle et aux connexions aux autres dépositaires de données dans le monde entier.

De plus, le groupe de travail recommande qu'un nouveau réseau national d'archives de données de recherche intègre certaines autres fonctions, y compris organiser des cours de formation spécialisée dans les techniques de maniement des données, représenter les intérêts canadiens dans le développement de normes de données internationales, promouvoir l'échange des données comme pratique exemplaire en recherche, effectuer des recherches dans le domaine des sciences de l'information et de l'archivistique, et agir en tant qu'organisme central de coordination pour un réseau de services de données, dans les instituts de recherche canadiens.

L'information numérique compilée aux fins de la recherche joue un rôle de plus en plus important dans l'économie actuelle fondée sur le savoir. À bien des égards, les données sont le moteur de l'innovation et de notre capacité à résoudre des problèmes sociaux et économiques complexes. Bien que des milliards de dollars soient dépensés chaque année pour la collecte de données, le Canada ne possède pas l'infrastructure requise pour en assurer la conservation et l'accès par le public, ce qui limite le rendement du capital investi dans la recherche et mine la bonne gestion publique.

De nombreuses composantes de base requises pour la création des archives nationales de données de recherche existent déjà, notamment, les services de traitement des données des universités, les réseaux de transmission à haute vitesse, les directives juridiques et éthiques et les cadres de travail, les instituts partenaires potentiels, divers dépositaires de données et des initiatives relatives à un portail d'accès ainsi qu'une communauté de

chercheurs qui produisent effectivement des données. L'élément manquant est un service de conservation, de coordination et de gestion des données.

Presque tous les pays développés ont reconnu le besoin d'un service national d'archives de données de recherche et certains pays ont plus d'une génération d'expérience dans l'exploitation de tels services. Le Canada peut bénéficier de cette expérience tout en développant un service d'archives de données de recherche compatible avec son contexte institutionnel et culturel unique. Nous possédons maintenant la capacité technologique et le savoir-faire requis pour créer un « système de confiance » en vue de fournir aux Canadiens et aux Canadiennes un service accessible et étendu qui permettrait aux chercheurs de localiser, de demander, d'extraire et d'utiliser les ressources en données de façon simple, transparente et économique tout en protégeant la vie privée, la confidentialité et les droits de propriété intellectuelle des intéressés. La Fondation canadienne pour l'innovation pourrait financer les coûts liés à l'élaboration de l'infrastructure de ce service. On estime que les coûts de fonctionnement annuels d'une installation et d'un réseau étendu sont d'environ 3 millions de dollars.

Le groupe de travail présente trois options pour ce qui est de la création d'un réseau national d'archives de données de recherche :

- 1) Créer, par le truchement de la législation fédérale, un réseau national d'archives de données de recherche en tant que version modifiée d'un organisme distinct constitué en vertu d'une loi. Cette option serait idéale pour la mise en place d'un service complet, d'un organisme fiable comprenant une installation centrale de conservation et de gestion des données et un ensemble de nœuds de services d'accès localisés dans des instituts de recherche. Le service profite pleinement de l'infrastructure de recherche existante, est stable à long terme et constitue une connexion directe pour les utilisateurs et les producteurs de données de recherche, et est apte à représenter les intérêts du Canada dans le développement des normes internationales de données.
- 2) Créer un réseau national d'archives de données de recherche sous la gouverne du Conseil de recherches en sciences humaines. Cette stratégie englobe les caractéristiques du premier modèle, mais ne requiert pas de dispositions législatives. Elle prend avantage de la connexion directe et immédiate avec les chercheurs, de l'obligation reconnue de rendre compte et des structures de financement.
- 3) Créer un organisme de service spécial au sein des Archives nationales du Canada. Par la création d'une division autonome au sein des Archives nationales, cette option prend avantage de l'infrastructure et de l'expertise existantes en matière d'archivage. Cette stratégie n'a pas été retenue par les autres pays parce que la mission fondamentale d'un service d'archivage national des données de recherche est essentiellement différente. Néanmoins, à titre d'organisme de service spécial, le service pourrait possiblement être stable et en mesure de développer un système fiable d'accès aux données, de gestion et de conservation de celles-ci.

Comme prochaine étape, le groupe de travail recommande que le CRSH et l'archiviste national constituent un comité directeur qui sera chargé de sélectionner la stratégie appropriée pour la mise en place d'un réseau national d'archives de données de recherche ou d'un service d'archivage des données de recherche et de définir plus précisément les

caractéristiques et les exigences en matière de financement d'un tel service et de promouvoir sa mise en place.

## Table des matières

### Sections

1. Introduction
2. En quoi consiste l'archivage des données de recherche?
3. Besoins en matière d'archives nationales de données de recherche
4. Composantes de base des archives nationales de données de recherche
5. Vers un modèle institutionnel : leçons tirées de la scène internationale
6. Principes fondamentaux et hypothèses
7. Options pour le Canada
8. Recommandations concernant la mise en place d'un réseau national d'archives de données de recherche
9. Coûts d'un réseau national d'archives de données de recherche
10. Prochaines étapes

### Annexes

- a) Membres du groupe de travail et du groupe d'étude
- b) Mandat
- c) Rapport sur les modèles internationaux et aperçu d'un service virtuel et volontaire d'archivage des données
- d) Aperçu des options entre plusieurs systèmes institutionnels
- e) Bibliographie

# Consultation sur les archives nationales de données

## Rapport final

### Infrastructure pour l'archivage et l'accessibilité des données de recherche du Canada

#### 1. Introduction

Au Canada, le réseau d'instituts et d'organismes qui constitue l'infrastructure de soutien dont dépend l'économie fondée sur le savoir souffre actuellement de lacunes importantes. Le Canada ne possède aucun organisme national pour conserver et cataloguer les données et fournir un accès systématique, efficace et opportun aux données de recherche. Ces renseignements numériques permettent aux chercheurs de corroborer les connaissances existantes, de reproduire les essais, de vérifier les résultats de la recherche, d'examiner les données et de créer de nouvelles connaissances. L'accès efficace aux données de recherche et leur utilisation peuvent jouer un rôle central pour ce qui est de la capacité d'innovation du Canada. Toutefois, l'infrastructure requise doit être mise en place.

En octobre 2000, le Conseil de recherches en sciences humaines et l'Archiviste national du Canada ont donné à un groupe de travail composé de chercheurs et d'archivistes le mandat de consulter la communauté des chercheurs et archivistes et d'évaluer le besoin d'un service national ou d'une fonction d'archivage des données. Au terme de l'évaluation et de la compilation d'éléments probants confirmant le besoin d'un tel service, le groupe de travail a étudié les archives de données de recherche d'autres pays et examiné les voies possibles en vue de créer une installation de recherche de base au Canada.

Le groupe de travail recommande maintenant la création d'un organisme canadien en vue de combler cette lacune dans l'infrastructure de l'économie canadienne basée sur le savoir – la création et le soutien à long terme d'un système d'archives nationales de données de recherche.

De nos jours, presque toute la recherche est effectuée dans un environnement numérique. Des bases de données statistiques complexes à multicouches, des cartes et images numériques et des textes codés constituent maintenant les outils usuels des chercheurs. Bien que ces ressources aient donné une envergure spectaculaire à la recherche et accru son efficacité, les structures institutionnelles exigées pour la conservation et la gestion des données et l'accès aux données numériques ne sont pas à la hauteur. Cette situation mine la capacité d'innovation des chercheurs canadiens et met en péril des données de recherche précieuses valant des dizaines de millions de dollars.

L'édification d'une société axée sur le savoir, la stimulation de l'innovation et l'examen des problèmes urgents et complexes sur les plans social, politique et économique reposent en grande partie sur la découverte de connaissances par le biais de la recherche. Pour être réceptifs et efficaces, tout en incorporant des perspectives multiples, les chercheurs doivent avoir accès à un large éventail de données de recherche et avoir la possibilité de les partager. Une infrastructure est nécessaire pour atteindre cet objectif. Aujourd'hui, de nombreux éléments sont en place, tels les bibliothèques de recherche et les services informatiques des universités, les conseils de soutien à la recherche, les réseaux de transmission de données à haute vitesse – mais il manque un élément essentiel, soit une installation de stockage, de distribution et de conservation des données de recherche.

La saine gestion publique exige que l'investissement public dans les données de recherche génère un rendement maximal du capital investi. Afin de maximiser le rendement, les données de recherche devraient être utilisées autant de fois que possible dans autant de contextes différents et cela ne peut se produire que si nous mettons en place une infrastructure efficace de données de recherche. Le coût de l'inaction met nos investissements en sciences en péril et mine également une des responsabilités essentielles du gouvernement.

## 2. En quoi consiste l'archivage des données de recherche?

À la différence de nombreuses formes d'archivage traditionnelles, l'archivage des données de recherche consiste non seulement à tenir des registres à des fins juridiques, historiques ou culturelles, mais il doit satisfaire aux besoins des chercheurs qui travaillent dans l'environnement numérique actuel. La mission fondamentale des archives de données de recherche n'est pas de conserver les mémoires enregistrées d'un groupe, d'un organisme ou d'une nation, mais de fournir un service essentiel au milieu de la recherche.

Bien que les nouveaux besoins des instituts en matière de programmes numériques soient de plus en plus nombreux, le groupe de travail n'a examiné que les besoins de la communauté des chercheurs en matière d'accès aux données, de gestion et de conservation de celles-ci, en particulier les données en sciences humaines. À partir de cette perspective, le groupe de travail a défini le processus d'archivage de données comme étant la conservation et la gestion des données de recherche et l'accès du public aux renseignements numériques structurés au moyen de méthodes de recherche dans le but de créer de nouvelles connaissances. Ce processus permet d'utiliser les résultats de la recherche qui sont produits depuis l'information initiale jusqu'à la publication des résultats. Les acquisitions comprendraient les renseignements numériques fournis par les chercheurs et qui présentent de l'intérêt pour eux, sous réserve des limites des ressources financières et des protocoles de conservation élaborés par les milieux de la recherche et des archives de données de recherche eux-mêmes.

Les archives et les organismes nationaux de données de recherche d'autres pays offrent une vaste gamme de services d'accès, de conservation et de gestion à leurs milieux de recherche respectifs, y compris le stockage sur le site et hors site, l'accès à des catalogues

et à des ensembles de données sur Internet, des protocoles de conservation, la création de métadonnées, la migration des données par des logiciels et du matériel informatique, la formation et l'élaboration de normes internationales. En offrant ces services, ces pays jouent un rôle actif et essentiel en tant que diffuseurs de l'information et des connaissances.

---

*Les Archives nationales de données m'apparaissent comme un institut canadien de confiance ayant pour mandat de conserver les données de recherche, de collaborer avec d'autres organismes gouvernementaux et non gouvernementaux pour assurer que leurs pratiques de gestion des données incorporent des normes de conservation, de travailler étroitement avec d'autres instituts canadiens chargés de préserver l'héritage du Canada et de prévenir les lacunes en matière de responsabilités, de représenter le Canada à l'échelle internationale lors des échanges de données de recherche et de l'élaboration de normes connexes, en vue de permettre l'accès à ces données, d'informer les Canadiens sur la manière d'utiliser les données de recherche, de contribuer aux nouvelles recherches en créant de nouvelles données à partir des données archivées, de participer à la protection de la vie privée dans la société canadienne compte tenu du stockage considérable d'informations numériques sur les personnes et de mener des activités de recherche et développement sur tous les aspects de la conservation des données.*

*Charles Humphrey, bibliothécaire de données, University of Alberta,  
Membre du Groupe de travail, Consultation sur les archives nationales de données*

---

### 3. Besoins en matière d'archives nationales de données de recherche

Comme l'a mentionné l'un des membres du groupe de travail, une « tempête de feu » sans précédent est en train de ravager la richesse du Canada en matière de données numériques. Cet énoncé peut sembler exagéré mais, en réalité, il traduit les préoccupations profondes de nombreux archivistes, bibliothécaires et chercheurs dans le monde entier<sup>1</sup>.

Les résultats de recherche sous forme numérique sont extrêmement fragiles, mais des quantités considérables de ces données peuvent être compilées. Aujourd'hui, nous commençons à peine à comprendre comment conserver et gérer efficacement cette information. Bien qu'il n'existe aucun raccourci pour résoudre des problèmes tels l'obsolescence des médias, la « désuétude » du système numérique, le droit d'auteur, la confidentialité, l'élaboration de normes nationales et internationales et les limites des milieux de recherche actuels, éviter ces problèmes ou les ignorer pourrait s'avérer onéreux, à long terme.

---

<sup>1</sup> Nombreux sont les organismes qui se débattent actuellement avec les structures et politiques relatives à l'archivage des données, y compris la *Library of Congress*, l'*Economic and Social Research Council* du Royaume-Uni, les *National Institutes of Health* des États-Unis, les *National Archives and Record Administration*, et le *National Research Council* des États-Unis, l'*International Council for Scientific and Technical Information*, et l'*International Council of Scientific Unions*.



Dans la phase initiale de la Consultation sur les archives nationales de données, le groupe de travail a consulté une vaste gamme d'intervenants qui utilisent, gèrent et produisent des données de recherche relatives aux sciences humaines. L'objet du groupe de travail était d'évaluer le besoin d'un service national ou d'une fonction d'archivage de données de recherche (voir l'annexe E). Cette évaluation a permis de déceler un certain nombre de lacunes structurelles :

- Actuellement, il n'existe aucun institut national ayant le mandat de conserver, de gérer et de produire des données de recherche accessibles au public, qui ait l'envergure requise pour appuyer les efforts de la communauté canadienne des chercheurs. Les Archives nationales du Canada ne disposent pas des ressources requises à cet effet;
- Les services de recherche de données des universités n'ont ni les ressources, ni la responsabilité les habilitant à agir à titre d'archives de données de calibre national. Bien qu'ils s'efforcent de combler les lacunes qu'entraîne l'absence d'un service national d'archivage des données, les services de traitement des données des universités n'ont, en général, que le mandat de fournir aux usagers locaux l'accès aux données facilement disponibles;
- La politique du CRSH sur l'archivage des données, qui incite les chercheurs qu'il soutient à déposer leurs données dans les dépôts d'archives des services de traitement des données des universités, n'a pas atteint ses objectifs. En réalité, en 11 ans, seulement 10 ensembles de données ont été déposés dans les dépôts d'archives de données des universités figurant dans le Guide du CRSH. Bien que certains chercheurs hésitent à partager leurs données, il serait contraire à l'éthique que le CRSH applique cette politique en l'absence d'une installation qui permettrait aux chercheurs de se conformer à la réglementation;
- Le Canada ne dispose d'aucune stratégie coordonnée pour établir des normes internationales de données de recherche dans le cas des projets de métadonnées, telle l'Initiative de documentation des données, ni d'outils pour accéder aux données, tel le projet *Networked Social Science Tools and Resources* (NESSTAR), et de projets collectifs d'infrastructures internationales, comme Frameworks, de l'Union européenne. De même, le Canada n'est pas représenté à l'échelle nationale à l'*International Federation of Data Organizations* et ne participe pas aux initiatives du *Council of European Social Science Data Archives*;
- L'un des problèmes primordiaux auxquels les chercheurs sont confrontés aujourd'hui est la difficulté de localiser les données relatives à la recherche qu'ils effectuent. Il n'existe aucune liste collective ou catalogue des ensembles de données que possèdent les producteurs et distributeurs de données et autres chercheurs. Il s'ensuit que les chercheurs peuvent reproduire inutilement des études coûteuses, se fonder sur de l'évidence anecdotique plutôt que sur l'évidence empirique ou utiliser des données de remplacement provenant d'autres pays. Un service national de données pourrait transférer directement aux ordinateurs personnels des chercheurs l'information concernant les sources des données ainsi que les données elles-mêmes, ce qui entraînerait une économie de temps et d'argent.

Des archives nationales de données de recherche pourraient également répondre aux besoins fondamentaux, de la façon suivante :

- Assurer l'authenticité des données de recherche, ce qui constitue une préoccupation croissante tant pour les producteurs que les utilisateurs de données de recherche. Les procédures d'authentification incorporées au processus de création, de transmission, de réception, d'utilisation, de maintien et de conservation des fichiers de données constituent la façon la plus efficace pour assurer l'authenticité des données au fil du temps. Actuellement, nous ne disposons ni de normes nationales de cette sorte ni d'aucun organisme pour surveiller leur application;
- Reformuler et articuler, à l'échelle nationale, des normes de sécurité pour protéger l'exactitude et l'uniformité des données. Ces normes devraient faire état : (1) de méthodes pour déterminer l'actif des données et de procédures de gestion du risque pour évaluer la vulnérabilité; (2) la détermination des exigences sur les plans juridique, législatif, de la réglementation et des obligations contractuelles et inclure des directives en matière d'éthique et des droits de propriété intellectuelle; (3) un ensemble de principes, de méthodes et de procédures que doivent suivre les organismes pour assurer la fiabilité en ce qui concerne la création, le maintien sûr, l'utilisation confidentielle et la conservation authentique de leurs données.

Si le Canada créait des archives nationales de données de recherche, seraient-elles utilisées? La vaste expérience acquise par d'autres pays démontre une croissance de l'utilisation des données, en fait de fréquence ainsi que de nombre d'utilisateurs (voir l'annexe C).

---

*Parmi les dix ensembles de données les plus demandés par les utilisateurs des archives de données du Royaume-Uni, au cours de l'année financière 2000-2001, quatre des titres provenaient de ministères gouvernementaux, deux étaient subventionnés par des ministères gouvernementaux et quatre étaient parrainés par un important organisme subventionnaire de la recherche.*

*Rapport annuel du R.-U. sur les archives de données*

---

Deux des éléments les plus utilisés pour mesurer les niveaux d'activité des archives de données sont l'envergure de leurs collections et le nombre de clients qu'elles desservent. L'an dernier, par exemple, le ICPSR de la University of Michigan a ajouté 1 835 fichiers de données à sa collection, soit une augmentation de 8 p. 100 par rapport à l'année précédente. D'autre part, le consortium a diffusé, en parallèle, 5 000 gigaoctets de données à ses clients. Durant la même période, les archives de données du R.-U. ont traité plus de 500 acquisitions et desservi 1 000 clients qui avaient passé 2 000 commandes, ce qui, au total, équivalait à près de 9 000 fichiers de données. Il s'agissait d'une augmentation de 2 000 fichiers de données livrés à des utilisateurs au cours d'une période de trois ans.

Plusieurs registres d'archives de données utilisent les statistiques basées sur le trafic Web. L'Oxford Text Archive, par exemple, a fait état de plus de 18 000 téléchargements de textes électroniques en 1999 et 2000. Cette utilisation de textes électroniques est de 39 fois supérieure aux commandes hors ligne de l'Oxford Text Archive. En plus du téléchargement de fichiers, le nombre de contacts des utilisateurs est aussi saisi à partir de statistiques en ligne. Par exemple, le ICPSR a signalé une croissance importante des contacts avec les clients en raison du plus grand nombre d'utilisateurs qui recouraient à Internet pour effectuer de la recherche et pour l'enseignement. Au cours des trois dernières années, alors que les ressources de ICPSR étaient disponibles en ligne, l'organisme a constaté une augmentation de l'accès aux données de plus de mille gigaoctets.

Les archives de données tiennent également des statistiques sur d'autres services. Par exemple, le programme de formation de l'ICPSR soutient, de façon constante, l'inscription annuelle de 500 à 540 participants. Pour citer un autre exemple, le Norwegian Social Science Data Service (NSD) tient des statistiques sur l'utilisation que font les chercheurs de leurs services afin de vérifier si les projets sont conformes à la loi. Le NSD déclare que ce service a connu une croissance allant jusqu'à 65 % dans une année donnée. Les services de référence tiennent habituellement leurs propres statistiques. En 2000-2001, les employés affectés aux archives de données du Royaume-Uni ont reçu 332 demandes de renseignements, suite aux commandes, relativement à des fichiers de données, ce qui ne représente qu'une facette des services de référence. Pendant l'année 2000, le service de données archéologiques a fait état d'un total de 174 demandes de renseignements qui incluaient des questions concernant l'utilisation du catalogue, l'aide technique, et autres renseignements généraux en archéologie. Le History Data Service reçoit environ 480 demandes de références générales pendant la même période. En plus d'offrir des services de référence, l'Oxford Text Archive a également fait l'évaluation technique de 125 demandes de subvention.

Une autre statistique utilisée par certaines archives de données est le volume des permis d'utilisation de logiciels élaborés et distribués par le biais de leurs services. Par exemple, des permis d'utilisation du logiciel NSDstat, mis au point et distribué par le NSD, ont été octroyés à environ 2 000 instituts en Norvège, et à 200 organismes à l'échelle mondiale. Bien que le nombre exact d'utilisateurs par permis soit inconnu, l'expérience démontre qu'un seul exemplaire de permis donne accès aux statistiques du NSD à plusieurs personnes.

Les archives de données plus importantes compilent aussi des statistiques sur leurs activités internationales. Par exemple, le personnel du German Central Archive for Empirical Social Research (ZA) déclare que chaque année, de façon continue, environ 50 chercheurs de calibre international mènent régulièrement, sur place, des recherches au moyen des données, à l'EUROLAB des ZA. Les ZA intègrent également les données et la documentation portant sur un certain nombre de projets internationaux, y compris le International Social Survey Program, pour le compte de 38 pays et sur les eurobaromètres, pour le compte de la Commission européenne.

Dans l'ensemble, l'expérience démontre que les archives de données qui offrent des services étendus (y compris la formation, l'élaboration de logiciels et l'accès en ligne aux fichiers de recherche) sont largement utilisées par les chercheurs d'envergure nationale et internationale, et dans chaque cas, on note une augmentation de l'utilisation.

---

*Nombreuses sont les raisons sous-jacentes au partage des données résultant des études appuyées par le NIH. Le partage des données renforce les enquêtes scientifiques ouvertes, encourage la diversité des analyses et des opinions et les nouvelles recherches, facilite la vérification d'hypothèses et de méthodes d'analyse nouvelles ou de remplacement, appuie les études portant sur la collecte de données et les méthodes de collecte et d'analyse des données, encourage la formation de nouveaux chercheurs, facilite l'examen de thèmes non abordés par les chercheurs initiaux, et permet la création de nouveaux ensembles de données quand les données provenant de sources multiples sont combinées. En évitant de reproduire des activités coûteuses de collecte de données, le NIH est en mesure d'appuyer un plus grand nombre de chercheurs qu'il ne le pourrait si des données semblables devaient être recueillies à nouveau par chaque demandeur.*

*National Institutes of Health (États-Unis),*

*Énoncé de politique sur le partage des données de recherche*

---

#### 4. Les composantes de base des archives de données de recherche

Au cours des dernières années, le gouvernement du Canada a pris des mesures importantes en vue de créer une infrastructure de recherche globale et cohérente et un système de soutien de la recherche au Canada. Des mesures telles que la création de la Fondation canadienne pour l'innovation et de CA\*Net3 ont réussi à combler plusieurs des lacunes existantes. Toutefois, l'un des besoins toujours à combler est celui d'une installation ou d'un institut qui serait responsable d'assurer la conservation des données de recherche et l'accès à celles-ci. Néanmoins, le Canada dispose déjà de nombreuses composantes de base pour la mise en place de cet organisme.

**Services de traitement de données des universités** – Les services existants de traitement de données des universités constituent peut-être la plus importante des composantes de base. Faute de ressources, ces services ne peuvent agir à titre d'organismes multiservices de traitement de données, mais ils possèdent le potentiel pour constituer des nœuds du réseau des archives nationales de données de recherche. Ce fait a été confirmé par l'expérience de l'Initiative de démocratisation des données, alors que les bibliothécaires et les archivistes de données de 66 universités se sont réunis pour former un consortium en vue d'améliorer l'accès aux données de Statistique Canada. Ces professionnels dynamiques gardent contact entre eux, partagent les meilleures pratiques et l'information concernant les sources de données, les méthodes pour améliorer les services aux clients, de même que les progrès récents en matière de normes et de capacités techniques. Si les services de traitement de données des universités disposaient de ressources suffisantes, ils pourraient former un réseau national étendu de points de contact pour les chercheurs qui désirent accéder aux données de recherche compilées par d'autres, déposer les données qu'ils ont compilées eux-même, suivre des cours de

formation afin d'acquérir des compétences supérieures en matière de traitement des statistiques et des données, et obtenir des conseils sur les meilleures pratiques et sur la façon d'agir conformément aux normes sur les données. Toutefois, le plus important serait qu'à long terme, le personnel responsable du réseau des services de traitement de données des universités puisse agir en tant que système de rétroaction pour les utilisateurs, aidant ainsi à structurer et à améliorer les services offerts par les Archives nationales de données de recherche, et, en bout de ligne, les connaissances créées par les chercheurs canadiens.

**Instituts canadiens des archives** – À l'instar des services de traitement de données d'universités, les instituts canadiens des archives, qui ont un mandat de recherche spécifique, offrent d'autres nœuds potentiels dans le cadre d'un réseau national d'archives de données de recherche. Ces instituts existent dans des environnements locaux, régionaux et institutionnels, soit comme entités indépendantes, comme partie d'une institution mère ou au sein des administrations gouvernementales, municipales, provinciales et fédérales. De plus, ils existent dans de nombreuses collectivités qui ne sont pas dotées d'universités. Bien que les archives canadiennes aient, jusqu'à tout récemment, traité principalement des dossiers non numériques, l'infrastructure de leur collectivité, les normes descriptives, les pratiques exemplaires, une expérience exhaustive en matière de protection de la confidentialité et du droit d'auteur, etc. constituent des assises solides qui leur permettront de développer des connaissances et du savoir-faire en vue de participer au réseau national de données de recherche.

**Représentation à l'échelle internationale** – Bien qu'ils n'aient pas d'autorité nationale, les employés qui s'occupent des services de traitement de données de certaines universités assurent actuellement l'une des principales liaisons du Canada avec de nombreux organismes internationaux responsables de la gestion des données de recherche et de l'élaboration de normes internationales en vue de la création de métadonnées, du partage et de la conservation des données. La création de ces normes, les ententes et les pratiques communes sont essentielles dans un monde scientifique qui de plus en plus travaille hors frontières. L'expérience et les compétences de ces personnes peuvent être utilisées pour coordonner les efforts, protéger les intérêts du Canada et représenter ce dernier quand des décisions clés ayant des incidences à long terme sont prises.

**Infrastructure de la transmission de données** – CA\*Net3 relie les services de traitement de données des universités et, sous peu, CA\*Net4, le réseau optique national de transmission de données à très haute vitesse, développé par le Réseau canadien pour l'avancement de la recherche, de l'industrie et de l'enseignement (CANARIE) sera mis en service. CA\*Net3, qui relie maintenant tous les instituts de recherche importants au Canada, fournit le pipeline nécessaire à la transmissions des données de recherche à l'échelle nationale. La capacité considérable de ce réseau permet la transmission rapide, efficace et fiable de très volumineux ensembles de données importants et complexes, ce qui est essentiel dans l'avenir. La dimension et la complexité des ensembles de données de recherche augmentent à un rythme effarant.

**Cadres de gestion** – Les cadres de gestion de l’utilisation des données de recherche sont tout aussi importants que les pipelines numériques et les nœuds d’accès. En raison des renseignements personnels délicats qu’elles contiennent, les données des sciences humaines, en particulier, doivent être gérées au moyen d’un plan d’ensemble et de cadres éthiques ainsi qu’en conformité avec la *Loi sur l’accès à l’information et la protection des renseignements personnels*. Les lignes directrices des trois conseils sur l’éthique de la recherche avec des êtres humains fournissent l’un de ces cadres. Ces directives font état, en termes généraux, des principes qui devraient guider les Archives nationales de recherche dans le traitement des renseignements personnels et de la confidentialité. Ainsi, avec les comités universitaires d’éthique de la recherche, nous possédons les règlements et la capacité institutionnelle pour assurer la protection des renseignements personnels propres à chaque citoyen. Ces comités déterminent les conditions en vertu desquelles des données sensibles peuvent être déposées et diffusées et ainsi incorporent une première étape de présélection aux fins des Archives nationales de données de recherche.

**Recherche et développement** – Dans notre univers numérique qui prend rapidement de l’essor, de nombreuses activités liées au maniement des données de recherche sont encore effectuées sans les connaissances suffisantes. À titre d’exemple, que l’on pense à l’assurance de la qualité, de l’authenticité et de la sécurité des données de recherche. Des Archives nationales de données seront en mesure de tirer parti des nouvelles connaissances résultant de la recherche de pointe dans ce domaine, y compris, par exemple, le projet InterPARES que finance le CRSH.

**Instituts partenaires** – Divers instituts peuvent jouer un rôle important dans le fonctionnement et les services des Archives nationales de données. Tant les Archives nationales du Canada que la Bibliothèque nationale du Canada ont, au fil des ans, développé des compétences appréciables par le truchement de leurs archives respectives ainsi que dans la transition de ces archives sous forme électronique. L’environnement du stockage, les normes descriptives, la migration des formats physiques et logiques et la protection du droit d’auteur ne sont que certains des secteurs où le partage des connaissances et la réalisation de projets conjoints seraient possible.

**Données de recherche** – Les données elles-mêmes constituent les composantes centrales d’un service de données de recherche. Il va de soi que tous les ensembles de données de recherche ne devraient pas nécessairement être conservés. Certains ensembles seront peu utiles après l’achèvement d’un projet aux fins duquel les données ont été compilées; d’autres contiennent des indices personnels qu’il est impossible d’éliminer efficacement ou reproduisent simplement des données compilées ailleurs. Toutefois, la décision de conserver certains ensembles de données et d’en supprimer d’autres est une question préoccupante en archivistique et est cruciale pour assurer un partenariat efficace entre les chercheurs et les archivistes de données.

L’existence d’une grande quantité de données de recherche n’est pas remise en question. Dans la première phase des consultations, le groupe de travail a déterminé que les chercheurs subventionnés par le CRSH produisent annuellement, en moyenne, quelques 400 ensembles de données. Comme le CRSH peut appuyer une partie seulement de la

communauté des chercheurs en sciences humaines, le nombre total des ensembles de données produits chaque année pourrait être de trois à quatre fois supérieur à ce nombre. Ne sont pas inclus les ensembles de données produits par des chercheurs en sciences naturelles, des spécialistes de la santé ou des ingénieurs de la recherche, mais il n'est pas déraisonnable d'estimer que de 4 000 à 5 000 ensembles de données sont produits annuellement, grâce aux fonds publics. Bien qu'il soit impossible de connaître des détails à ce sujet, ces ensembles représentent un investissement public de dizaines millions de dollars par année.

**Données de recherche du gouvernement** – Les recherches du groupe de travail concernant les archives d'autres pays ont révélé qu'en sciences humaines, le gouvernement produit des données de recherche qui sont souvent plus utilisées que les données produites par les chercheurs eux-mêmes. Un nouvel organisme pourrait jouer un rôle important en exploitant une installation de conservation des données, en offrant l'accès au catalogue de données de recherche compilées au gouvernement. Le groupe de travail a souvent entendu des témoignages indiquant que l'accès à ces informations est, dans l'hypothèse la plus optimiste, difficile, et exige beaucoup de temps, et dans l'hypothèse pessimiste, impossible. Toutefois, on estime que les ministères et organismes gouvernementaux tels que Statistique Canada, Développement des ressources humaines Canada (DRHC), Santé Canada, Ressources naturelles Canada, Environnement Canada, Justice Canada et de nombreux autres ministères dépensent plus d'un milliard de dollars annuellement pour la collecte des données. Une administration publique efficace permettrait de donner aux chercheurs des moyens efficaces et efficaces d'utiliser ces données<sup>2</sup>.

**Services de conservation pour d'autres organismes de recherche** – Les technologies actuelles de l'information facilitent considérablement notre accès aux renseignements numériques et notre capacité de les traiter et de les appliquer aux sujets de recherche qui sont d'une importance primordiale pour les Canadiens et les Canadiennes. Toutefois, la conservation à long terme des documents de recherche numériques est un secteur qui n'est pas resté à la hauteur tant du point de vue technologique qu'institutionnel. Actuellement, dans le milieu de la recherche, on met l'accent sur la compilation et l'accessibilité des données sur Internet. Les services de catalogage et de conservation des données interinstitutions sont souvent considérés comme étant d'importance secondaire ou complètement ignorés. L'Institut canadien d'information sur la santé, le Centre canadien de la statistique juridique, le Système canadien d'information pour l'environnement, GéoConnexions, et la Stratégie de données communautaires du Conseil canadien de développement social, qui a été annoncée récemment sont tous dotés d'excellents systèmes d'accès aux données, mais n'ont pas une stratégie bien conçue et fermement appuyée de conservation des données à long terme. Un des rôles les plus importants que puissent jouer les Archives nationales de données de recherche est d'offrir des services de conservation de données et l'expertise pour ces services et de mettre en place de nombreuses autres initiatives d'accès aux données de recherche.

---

<sup>2</sup>Le programme canadien des changements à l'échelle du globe, le Panel sur les systèmes d'information et de données, L'analyse de données et de politiques et les obstacles à l'accès des données au Canada : Sujets de recherche sur le changement mondial, (Société royale du Canada, 1996), p. 7.

---

*Dans le cas des recherches subventionnées par les fonds publics, on devrait exiger que les données produites, les outils de recherche utilisés, les cadres de recherche sur les concepts et l'échantillonnage etc., soient archivés et mis à la disposition des autres chercheurs. Cette mesure serait très importante dans la réalisation d'activités telles que l'incitation à la collaboration, les études longitudinales, les répliques d'études, les études comparatives, la création de concepts de questions « normatives » dans certains secteurs d'examen et d'analyses secondaires. En exigeant l'archivage et l'accessibilité des données, on encouragerait la transparence, l'obligation de rendre compte et la responsabilité. De plus, la prise en compte de telles données devrait devenir l'un des points principaux dans la planification de la « nouvelle » recherche de source directe – cessez de réinventer la roue et réalisez des travaux plus novateurs et créatifs.*

*Répondant au questionnaire*

---

## 5. Vers un modèle institutionnel : leçons tirées de la scène internationale

Le groupe de travail a examiné toutes les archives de données de recherche existantes qui sont axées sur les sciences humaines (voir l'annexe C). Cet examen comprend des entrevues de face à face avec des directeurs d'organismes de traitement de données, l'analyse comparative des politiques et des règlements, l'examen des services, des mandats, des budgets et des structures de gestion. Les principales leçons tirées de l'examen sont les suivantes :

- Nombreux sont les pays qui ont reconnu depuis longtemps le besoin de se doter d'archives de données de recherche pour aider et soutenir les travaux de la communauté des chercheurs. Plusieurs archives de données examinées existent depuis 30 ans ou plus;
- Bien que de nombreux services d'archivage de données, en particulier les services d'accès et de formation, soient diffusés plus efficacement lorsque répartis dans un certain nombre d'installations, pour des raisons d'ordre économique, pratique, d'efficacité et de conservation, la gestion du réseau et des fonctions relatives à l'élaboration de normes est plus efficace lorsqu'elle est effectuée dans une seule installation;
- Les archives de données diffèrent l'une de l'autre. Chacune d'entre elles a été établie dans un contexte national ou disciplinaire spécifique qui reflétait les besoins particuliers du milieu scientifique qu'elle desservait. Il s'agit d'organismes intégrant des disciplines spécifiques, de petite à grande envergure, offrant des services restreints, jusqu'à des instituts nationaux pluridisciplinaires importants axés sur la R-D, qui offrent des services étendus et des réseaux à l'échelle internationale;
- Les archives de recherche florissantes sont directement reliées à l'infrastructure de recherche d'un pays, plutôt qu'au milieu des archives. Elles se caractérisent par l'orientation des services qui met en valeur l'accès aux données de recherche et la conservation des données les plus utiles aux fins de la recherche plutôt que sur la saisie de fichiers de données antérieurs;



- L'archivage des données de recherche est un travail complexe et hautement technique. Les archives de données exemplaires emploient des professionnels dévoués, spécialistes des données et accordent beaucoup d'importance à la formation des gestionnaires de la prochaine génération de données de recherche. La formation de personnel hautement qualifié répond aux besoins tant du milieu de la recherche qu'à ceux d'autres organismes des secteurs public et privé qui doivent traiter de grandes quantités de données;
- Il existe une corrélation directe entre le financement stable des archives de données de recherche et la qualité du soutien qu'elles offrent à la communauté des chercheurs. Par sa nature même, l'archivage est une entreprise à long terme. Les archives de données les plus utiles sont celles dont l'existence est assurée;
- Bien que l'archivage des données de recherche exige des engagements à long terme en matière de financement, les coûts institutionnels ne constituent qu'une très petite fraction des coûts de la collecte de données;
- L'établissement de liens de confiance avec les utilisateurs et les producteurs de données de recherche est essentiel. Si les utilisateurs ne peuvent se fier à la transmission efficace et en temps opportun de données de qualité et si les dépositaires ne sont pas convaincus que leurs droits de propriété intellectuelle seront sauvegardés et que la protection de leurs participants sera assurée, nul ne pourra faire confiance aux services fournis ou les utiliser;
- Les archives de données exemplaires sont souples et indépendantes sur le plan institutionnel. Elles travaillent en étroite collaboration avec de nombreux ministères gouvernementaux et avec des universités mais ne dépendent d'aucun organisme en particulier en matière de stabilité financière ou de prise de décisions. Les archives de données doivent être indépendantes afin d'assurer que les besoins de la communauté des chercheurs en matière d'accès aux données demeurent la principale priorité, plutôt que les besoins relatifs à la tenue des dossiers des ministères gouvernementaux ou des archives traditionnelles de données de recherche sur la culture. La souplesse est importante eu égard à l'adoption de nouvelles technologies et à la capacité de suivre l'évolution des besoins des chercheurs.

Le groupe de travail qui a fait un examen détaillé de 36 instituts a déterminé trois stratégies généralisées visant la conservation et l'accessibilité des données de recherche. Chaque stratégie possède les caractéristiques organisationnelles des services actuels offerts par les services nationaux d'archivage de données :

- **Des archives de données à petite échelle, spécialisées dans les données relatives à des sujets précis**, habituellement hébergées par un département universitaire qui possède une capacité limitée de maniement des données et qui utilise des technologies facilement disponibles sur le marché. La clientèle n'inclut parfois que des chercheurs d'une seule discipline ou d'un petit groupe de domaines de recherche et les budgets de fonctionnement annuels varient de 200 000 \$ à 400 000 \$.
- **Des archives de données de dimension moyenne, intégrées à un organisme**, dont l'organisme parrain est habituellement un institut national de recherche ou un ministère. Le mandat et les activités de collecte de données de ces archives, qui sont souvent situées sur un campus universitaire pour mieux desservir la clientèle de base

composée de chercheurs, sont basés sur celles de l'organisme parent. Ces archives offrent une gamme de services relativement étendue et il arrive que les employés assument parfois des rôles de premier plan dans des organisations nationales ou internationales connexes. Les budgets annuels varient de 500 000 à 1,5 million de dollars.

- **Des archives de données de recherche étendues**, offrant des services à une grande variété de communautés, y compris des chercheurs en milieu universitaire, des analystes de politiques provenant d'organismes non gouvernementaux et gouvernementaux, des organismes publics de services archivistiques, et de simples citoyens. Ces archives de données qui sont souvent établies en vertu d'une loi sont reconnues à titre d'instituts nationaux responsables des principes généraux et des fonctions particulières décrits dans leurs actes fondateurs. Par le truchement d'un ou de plusieurs emplacements pratiques et de l'utilisation exhaustive d'Internet, une gamme étendue de services est offerte. Les services incluent souvent la formation spécialisée et générale, du soutien technique et de la R-D. Les capacités de gestion des données de ces archives sont considérables et souvent développées sur place. Ces organismes ont établi des relations de travail avec d'autres organismes nationaux et ministères, et le personnel est souvent composé de dirigeants d'associations internationales qui participent activement à des échanges internationaux de données. Les budgets annuels de ces organismes varient de 3 à 6 millions de dollars.

---

*Avantages résultant du dépôt et de l'archivage des données :*

- \* *Encouragent les enquêtes scientifiques ouvertes;*
- \* *Encouragent la diversité des analyses et des opinions;*
- \* *Favorisent les nouvelles recherches et permet la vérification des nouvelles méthodes ou des méthodes de remplacement;*
- \* *Améliorent les méthodes de collecte et d'évaluation des données en les soumettant à l'examen d'autres chercheurs;*
- \* *Réduisent les coûts en permettant d'éviter de reproduire les travaux de collecte de données;*
- \* *Constituent une ressource importante aux fins de la formation en recherche;*
- \* *Assurent la sauvegarde des données;*
- \* *Permettent aux propriétaires d'éviter les tâches administratives associées aux utilisateurs externes et à leurs requêtes;*
- \* *Remplissent les obligations voulant que les recherches subventionnées soient mises à la disposition de la communauté des chercheurs;*
- \* *Permettent aux chercheurs de démontrer l'utilisation continue des données après l'achèvement des travaux de recherche initiaux.*

*Site Web du Inter-University Consortium for Political and Social Research, University of Michigan*

---

## 6. Principes fondamentaux et hypothèses

Le groupe de travail a conclu que des archives nationales de données de recherche devraient fonctionner en conformité avec un ensemble de principes fondamentaux. L'objectif global devrait être de créer « un système de confiance » offrant au milieu de la recherche des services accessibles et étendus qui permettent aux utilisateurs de localiser, de demander, d'extraire et d'utiliser les ressources en données de façon simple, transparente et économique. Un tel système devrait être conforme aux principes fondamentaux suivants :

- 1) Des archives nationales de données de recherche devraient encourager la création de connaissances en constituant une partie intégrale du processus de recherche et devraient également faciliter les découvertes et le processus décisionnel au Canada, y compris la formulation de politiques publiques, en conservant les sources des preuves et en facilitant l'accès à ceux-ci;
- 2) Dans la mesure du possible, l'accès aux données devrait être libre et gratuit;
- 3) L'assurance de la confidentialité et la protection de la vie privée et des êtres humains qui participent aux activités de recherche devraient être primordiales;
- 4) Le grand public devrait avoir accès aux données recueillies grâce aux fonds publics sous réserve des conditions relatives à leur utilisation préliminaire équitable par le dépositaire et des dispositions éthiques et juridiques régissant ladite collecte de données.

En maintes occasions, le groupe de travail a appris de sources qui font autorité que l'établissement d'un climat de confiance constitue l'élément clé de la création d'un système exemplaire de conservation et d'accessibilité de données de recherche. Ceci ne peut être accompli que si les utilisateurs et les dépositaires de l'organisme savent que les archives font partie intégrante de leur processus de recherche, qu'elles fourniront des services utiles et ajouteront de la valeur à leurs travaux. De plus, le service de traitement de données doit appuyer et se conformer activement aux règlements et directives établis concernant la protection de la confidentialité, de la vie privée et des droits de propriété intellectuelle. Dans la mise en place d'un système de confiance, les éléments les plus importants pour un nouvel organisme sont la stabilité à long terme, tant sur le plan de la structure institutionnelle que sur celui du financement. Il s'agit de leçons durement apprises par de nombreuses archives de données du monde entier. La source du mandat, de la régie, de l'obligation de rendre compte et un engagement à long terme en matière de ressources financières requises sont des éléments déterminants quant à la réussite ou à l'échec de l'entreprise.

## 7. Options pour le Canada

À la lumière de ces expériences et en consultation avec la communauté des chercheurs, le groupe de travail a conclu qu'il serait à l'avantage du Canada de créer un organisme possédant les caractéristiques générales suivantes :

- Mandat global établi en fonction des besoins d'une grande variété d'intervenants;
- La société et les personnes seraient l'objet central des données cibles;
- Service mettant en valeur la conservation et l'accessibilité des données ;

- Capacité d'assurer la protection de la vie privée et de la confidentialité comme élément fondamental de ses principes de fonctionnement;
- Capacité de traiter des données conformément aux normes internationales, participation aux échanges internationaux de données et représentation des intérêts du Canada lors de négociations internationales;
- Capacité d'effectuer de la recherche de pointe et de faire évoluer les sciences de l'information et de l'archivistique;
- Utilisation des techniques d'information et de communication les plus récentes en vue de maximiser l'accès aux données de recherche tout en réduisant les coûts en temps et en argent pour les chercheurs;
- Capacité d'éduquer et de former tant les producteurs que les utilisateurs des données de recherche et les gestionnaires de la prochaine génération de données;
- Établissement de relations de travail continues avec d'autres organismes nationaux, tels que les Archives nationales et la Bibliothèque nationale ainsi qu'avec des organismes comme CANARIE;
- Adhésions à des organismes et autres ententes officielles relatives aux échanges de données avec d'importantes archives de données à l'étranger, telles que l'ICPSR, aux États-Unis, et le réseau européen *CESSDA*;
- Financement public soutenu à long terme comme source principale de soutien. Ce financement pourrait être complété par la vente de produits à valeur ajoutée dérivés de données et de services de consultation à des organismes à but lucratif, mais ces ventes ne devraient pas constituer le financement de base.

Toutefois, dans le contexte canadien, des Archives nationales de données de recherche doivent également répondre aux critères suivants :

- Service entièrement bilingue;
  - Accès aux données de recherche produites à tous les niveaux du gouvernement tout en respectant les limites de compétences fédérales et provinciales dans les secteurs comme l'éducation et la santé;
  - Respect des lois, règlements et directives sur la propriété intellectuelle canadienne, du droit d'auteur, de la vie privée et de la confidentialité et collaboration à leur développement;
  - Maintien de relations étroites de travail avec les principaux producteurs canadiens de données tels que Statistique Canada et les organismes de statistique provinciaux;
  - Utilisation et soutien des infrastructures de recherche existantes, des services de soutien à la recherche, des programmes de soutien et de financement, y compris les services existants de traitement de données des universités et les bibliothèques de recherche, les conseils de soutien à la recherche, la Fondation canadienne pour l'innovation, le Centre canadien de la statistique juridique, l'Institut canadien d'information sur la santé ;
  - Intérêt dans les données de recherche en sciences humaines et, le cas échéant, en sciences naturelles, en sciences de la santé et en génie.
-

*L'archivage de données comprend un engagement à long terme en matière de ressources, d'expertise et des services publics requis afin d'assurer un accès continu aux fichiers de données, de décrire et de documenter des fichiers et d'offrir un accès à ces fichiers tout en assurant le contrôle intellectuel. L'une des raisons qui expliquerait pourquoi cette idée suscite peu d'enthousiasme chez les chercheurs est qu'il est difficile de savoir quelles données ont été recueillies. Ce serait tout à fait logique d'utiliser des économies d'échelle et de centraliser les ressources requises pour une entreprise de cette envergure.*

Répondant au questionnaire

---

Au Canada, les Archives nationales de données de recherche devraient répondre aux besoins en matière de conservation et d'accessibilité des données et repousser les frontières des sciences de l'information et de l'archivistique. Elles devraient avoir pour assise l'infrastructure de recherche existante tout en intégrant les leçons tirées de l'expérience en archivage de générations de données acquise dans d'autres pays. Il faudrait surtout que les archives soient bien adaptées au contexte social et institutionnel du Canada, tout en répondant aux exigences du public en ce qui concerne l'obligation de rendre compte et la régie efficace.

En étudiant les moyens à utiliser pour créer des Archives nationales de données de recherche, le groupe de travail a examiné les centres de recherche universitaires subventionnés par le gouvernement, a demandé l'avis du Bureau du Conseil privé et utilisé les directives figurant dans le *Cadre d'examen des différents modes d'exécution de programmes* publié par le Conseil du Trésor. Le groupe de travail a pris en compte six options possibles et a examiné chacune d'elles de façon exhaustive. Il a examiné les structures institutionnelles et de régie, les exigences liées au démarrage de l'organisation, à sa stabilité à long terme, ainsi que ses forces et ses points faibles selon la perspective des utilisateurs et des producteurs de données (voir l'annexe D).

En premier lieu, le groupe de travail a étudié l'option consistant à créer une nouvelle division ou un organisme de service spécial chapeauté par un établissement national existant, tel que les Archives nationales ou la Bibliothèque nationale du Canada. Bien que le mandat des Archives nationales soit assez étendu pour inclure les données de recherche non publiées, leur niveau actuel de financement ne leur permettrait pas d'intégrer un tel nouveau secteur de services, alors que l'organisme doit simultanément relever des défis importants, à l'échelle du gouvernement, en ce qui concerne la gestion de l'information dans les secteurs de l'administration électronique, de la transition de ses dossiers à une forme électronique et de la numérisation de ses ressources documentaires existantes. De plus, l'échec d'un essai antérieur en vue de créer une division d'archivage de données au sein des Archives nationales (1973-1986) laisse à penser qu'il existe une disjonction entre le rôle étendu de conservation de la culture qu'assument les Archives nationales et le rôle spécifique de prestation de services qu'assureraient les Archives nationales de données de recherche, au sein de l'infrastructure de recherche du Canada. Ces différences tiennent des stratégies d'acquisition, des compétences du personnel, des programmes descriptifs courants, et des besoins de la clientèle.

La Bibliothèque nationale du Canada collectionne un nombre restreint d'ensembles de données de recherche qui sont conformes à la définition de « publications ». Ces ensembles ne sont, toutefois, que de petits sous-ensembles de données de recherche devant être conservés au Canada. Comme dans le cas des Archives nationales, l'accès aux ressources documentaires actuelles, leur conservation et leur maintien dans les deux organismes ne requièrent pas une connaissance approfondie des méthodes de recherche quantitative, des statistiques et de l'informatique, pour répondre aux besoins des personnes qui utiliseraient un réseau national de données de recherche. Le groupe de travail estime que les exigences particulières du milieu de la recherche relativement aux données requises pourraient marginaliser les activités des archives de données de recherche au sein de ces organismes et ainsi miner la stabilité à long terme requise pour assurer la réussite de l'entreprise.

Une autre option que le groupe de travail a examinée était la création d'un partenariat gouvernemental tel qu'évoqué par le Conseil du Trésor. Il s'agit de l'établissement d'un organisme comme partenariat entre les divers ordres des gouvernements fédéral et provinciaux. Bien que cette voie comporte certains aspects intéressants, elle ne se prête pas à l'établissement de liens directs avec la communauté des chercheurs des milieux universitaire et non gouvernementaux. Comme l'a appris le personnel des archives nationales dans d'autres pays, il s'agit d'un élément essentiel pour la création d'un système de confiance.

Un organisme distinct en vertu d'une loi ou un établissement public possède bon nombre des caractéristiques nécessaires à la création d'un système étendu, fiable et efficace d'archives nationales de données de recherche. Il s'agirait d'un organisme permanent protégé par la loi. Il constituerait un élément de la stratégie d'innovation, axé sur la recherche, le renforcement des capacités, la régie et la compétitivité internationale. Il aurait des directives précises en matière d'autorité et de responsabilité et serait parrainé par un ministère. Le financement serait assuré et stable et proviendrait d'une seule source. À l'instar de Statistique Canada, l'organisme aurait le potentiel et les moyens de se faire reconnaître comme étant un « système de confiance » et pourrait jouir du statut de représentant officiel du Canada sur la scène internationale. Le seul élément d'importance manquant est un lien direct avec les utilisateurs de données de recherche.

La dernière option étudiée, soit un centre situé dans une université, possède, sur place, ce lien direct et immédiat. Une telle installation donnerait un sentiment d'appartenance et le fonctionnement et les politiques seraient entre les mains des membres associés aux universités. L'organisme tire parti des services de données existants, l'expertise et l'infrastructure technologique en place dans les universités; il pourrait aussi utiliser un système hybride centralisé/décentralisé où le Centre s'occuperait de la conservation et du traitement des ensembles de données et où les membres associés offriraient des services locaux d'accès aux données, de dépôt de données, et agiraient à titre de conseillers sur place, d'organiseurs d'activités de formation; l'envergure de l'organisme serait mise à l'échelle et pourrait inclure les domaines scientifiques visés par le Conseil de recherches en sciences naturelles et en génie (CRSNG) et les Instituts de recherche en santé du Canada (IRSC). Enfin, des activités numériques de recherche en matière d'archives

tireraient avantage de la proximité des chercheurs en sciences de l'information qui travaillent à l'université. La principale faiblesse de cette option est que l'organisme ne serait pas stable à long terme. Une deuxième faiblesse est le fait que l'organisme ne serait pas nécessairement habilité à agir à titre d'organisme national sur la scène internationale.

Après avoir examiné et débattu exhaustivement toutes ces options, le groupe de travail a conclu qu'en raison de la nature du système de gouvernement fédéral du Canada, des nouvelles technologies de communication et d'information, des caractéristiques particulières de la communauté des chercheurs et de l'essor des besoins de l'économie axée sur le savoir, une occasion unique d'innovation institutionnelle se présente – la création d'un organisme hybride combinant la stabilité d'un organisme distinct constitué par une loi et les liens avec les utilisateurs d'un centre de recherche situé dans une université.

## 8. Recommandations concernant la mise en place d'un réseau national d'archives de données de recherche

Afin de mettre en place un service national efficace d'archives de données de recherche qui répondrait le mieux aux besoins de l'économie canadienne axée sur le savoir, encouragerait l'innovation, utiliserait les capacités de l'infrastructure existante, assurerait une administration publique efficace et représenterait le Canada sur la scène internationale, le groupe de travail recommande que le gouvernement du Canada prenne les mesures suivantes :

- Crée, par voie législative, un réseau national d'archives de données de recherche en tant que version modifiée d'un organisme distinct constitué en vertu d'une loi;
- Exige que l'organisme relève du Parlement par l'intermédiaire soit du ministre de l'Industrie ou de la ministre du Patrimoine canadien, ou des deux, dans le cadre d'une structure de responsabilisation semblable à celle d'organismes tel le Secrétariat du changement climatique;
- Autorise cet organisme à fonctionner sans liens de dépendance, à l'instar des conseils de soutien de la recherche du gouvernement fédéral;
- Attribue directement des fonds d'exploitation à l'installation centrale et aux nœuds du réseau par des crédits parlementaires annuels dans le cadre du processus budgétaire ordinaire du gouvernement fédéral, ou encore alloue des fonds par l'intermédiaire des conseils fédéraux de soutien de la recherche participants, comme dans le cas des Réseaux des centres d'excellence.

En ce qui concerne les structures et les activités d'un réseau national d'archives de données de recherche, le groupe de travail fait, de plus, les recommandations suivantes :

- Que le nouvel organisme mette en place un réseau étendu de services et une installation centrale responsable de la gestion des données, de l'élaboration et de la conservation des normes et un ensemble de nœuds situés au sein des services de traitement des données de recherche des universités et autres établissements

responsables de l'accessibilité et du dépôt des données et qui sont en mesure d'offrir des services de formation et de consultation aux chercheurs. On suggère que les établissements qui souhaitent devenir des nœuds forment un consortium en vue d'obtenir le financement initial requis auprès de la Fondation canadienne pour l'innovation et des fonds de contrepartie de divers organismes de recherche provinciaux pour l'élaboration de l'infrastructure;

- Que l'organisme crée un Conseil de gestion pour régir le réseau national d'archives de données de recherche. Le Conseil serait composé de représentants des diverses régions du Canada et de divers groupes intéressés qui gèrent, utilisent et produisent des données de recherche;
- Que l'organisme développe, au fil du temps, pour répondre aux besoins déterminés de la communauté des chercheurs, un système d'accès aux données de recherche et des services de gestion et de conservation des données;
- Que l'organisme développe une capacité à approfondir nos connaissances et notre compréhension des sciences de gestion de l'information, des cadres éthique et juridique, des pratiques de gestion de la technologie du savoir, et encourage une culture de partage des données, au sein de la communauté des chercheurs;
- Que l'organisme entretienne des relations formelles de travail collaboratif avec d'autres établissements nationaux tels que les Archives nationales et la Bibliothèque nationale du Canada et qu'il conclue des ententes relatives à la conservation et l'accessibilité des données de recherche avec des producteurs de données importants tels Statistique Canada et les organismes provinciaux de l'analyse statistique;
- Que l'organisme soit habilité à agir au nom du gouvernement du Canada lors de négociations internationales touchant les normes de gestion des données de recherche et des pratiques courantes.

Bien que le groupe de travail soit persuadé que le modèle décrit ci-dessus permettrait au Canada d'être à l'avant-garde dans le domaine de l'archivage de données et des sciences de l'information et qu'il augmenterait considérablement l'avantage concurrentiel du milieu de la recherche canadien, les membres du groupe sont également conscients que la solution idéale n'est pas toujours la plus pratique ou la plus réalisable et proposent deux autres solutions pour créer des archives nationales de données de recherche.

- 1) **Un réseau national d'archives de données de recherche sous la gouverne du CRSH** – Pour donner suite à la stratégie du Economic and Social Research Council, du Royaume-Uni, on créerait, sous la gouverne du CRSH, une installation et un réseau d'archives de données intégrés à une université. L'organisme relèverait du Conseil d'administration du CRSH et serait doté de la même gestion et structure en réseau, et d'une gamme de services semblables à ceux de l'option décrite précédemment. Une loi habilitante ne serait pas nécessaire dans le cas d'un tel organisme, car il serait intégré à la fonction de soutien à la recherche qui fait partie du mandat du CRSH, mais il n'aurait pas la stabilité à long terme que procurent les dispositions législatives. Il tirerait avantage de liens directs avec la communauté des chercheurs ainsi que des relations de travail que maintient le CRSH avec d'importants producteurs de données, tel Statistique Canada. On prévoit qu'il pourrait être mis à l'échelle afin d'inclure tous les secteurs de recherche scientifique et des sciences



humaines et qu'il prendrait avantage de la recherche que finance le CRSH en sciences de l'information et en archivistique.

- 2) **Un organisme de service spécial chapeauté par les Archives nationales du Canada** – Bien qu'en apparence, cette voie puisse sembler la plus logique pour créer des Archives nationales de données de recherche, il est à noter que le groupe de travail n'a reçu que peu de suggestions recommandant l'adoption de cette voie. De plus, l'examen des archives de données de recherche dans d'autres pays a révélé que seules les archives de données de recherche danoises (Danish Data Archives) sont directement rattachées à des archives nationales et des renseignements empiriques laissent croire que cet arrangement a des effets défavorables. Néanmoins, la création d'un organisme de service spécial chapeauté par les Archives nationales serait une solution simple. Il pourrait tirer avantage de l'expérience archivistique du personnel des Archives nationales, utiliser les installations existantes ainsi que l'infrastructure technique et administrative et avoir le statut et l'autorité requis pour agir au nom du Canada dans le cadre du développement de normes et de pratiques internationales. En tant qu'organisme spécial, il jouirait d'une certaine autonomie au sein de la structure de gestion des Archives nationales tout en relevant de l'archiviste national du Canada. Cette voie donnerait une plus grande stabilité à l'organisme que celle de l'ancienne Division des archives ordinolingues. Le plus grand inconvénient de cette approche réside dans le fait que l'organisme n'aurait pas de lien direct et immédiat avec la communauté des chercheurs, soit par le truchement de sa structure de gestion ou par les services de traitement de données des universités. Bien que ces deux éléments puissent être mis en place, l'organisme serait malgré tout intégré à un organisme fédéral dont la mission principale est de préserver la mémoire nationale et les dossiers du gouvernement et non de répondre aux besoins des chercheurs en matière de données.

---

*Les chercheurs conviennent qu'une infrastructure permettant l'échange de données de recherche est depuis longtemps requise au Canada et que nous devons disposer d'une infrastructure cohérente pour la collecte, la documentation, le partage et la conservation des données numériques de recherche. En particulier, il est essentiel de réduire les coûts élevés de la collecte de données et d'assurer que les fichiers sont disponibles à des fins d'analyses secondaires.*

*Présentation de l'University of Calgary,  
Bureau du vice-président (Recherche)*

---

## 9. Coûts d'un réseau national d'archives de données de recherche

Le niveau de financement requis pour l'établissement et le maintien d'un service national de données de recherche dépend de l'envergure de ses activités et de la gamme de services qu'il fournit. Le point crucial est de définir le niveau minimal de financement qui serait requis pour fournir des services adéquats. Un niveau trop faible, en plus de

restreindre la gamme des services qui pourraient être offerts, menacerait également la reconduction du financement. Si le niveau de financement qui permet à l'organisme d'exister est trop bas, ce dernier risque d'être axé sur un nombre limité de disciplines ou de types de services et pourrait susciter du ressentiment de la part des utilisateurs qui ne recevraient pas les services voulus, et ainsi, la reconduction du financement serait compromise.

Le financement d'un réseau national d'archives de données de recherche – à la fois l'installation centrale et ses nœuds – devrait être assumé par le gouvernement du Canada. Bien que des fonds additionnels puissent être obtenus d'autres sources, telles les subventions à la R-D, la vente de produits à valeur ajoutée dérivés de données, et l'imposition de frais pour les services de consultation spécialisés, les activités générales de l'organisme devraient être financées par le gouvernement. C'est le seul moyen efficace d'assurer la stabilité à long terme des archives de données de recherche, qu'elles desservent tous les Canadiens et Canadiennes, dans toutes les régions, et qu'elles répondent aux besoins d'un large éventail de chercheurs et de producteurs de données de recherche en milieu universitaire, d'organismes gouvernementaux et non gouvernementaux et du secteur privé.

L'établissement précis des coûts d'un réseau national d'archives de données de recherche selon le modèle recommandé est impossible parce que le groupe de travail ne dispose pas des ressources requises. Toutefois, l'étude à l'échelle internationale des archives existantes a fourni des points repères importants quant aux niveaux de financement requis pour offrir certains niveaux de service.

En dollars canadiens courants, et une fois que l'organisme sera en service, le budget annuel de fonctionnement d'un organisme offrant des services étendus de traitement de données de recherche est d'environ 3 millions de dollars. Comme le souligne l'étude de faisabilité des archives de données de l'Irlande, environ 40 % des fonds seraient consacrés à l'acquisition, au traitement, au catalogage et à la conservation des données, tandis que les fonds restants (60 %) seraient dépensés pour le développement des procédés requis pour répondre aux besoins des utilisateurs<sup>3</sup>.

Les coûts de l'infrastructure initiale dépendraient d'un certain nombre d'éléments, y compris l'emplacement et la dimension de l'installation centrale de conservation et de traitement des données, le nombre de nœuds du réseau, la répartition des fonctions spécifiques entre les nœuds et l'installation centrale, ainsi que la capacité et la complexité globales du matériel informatique. Si l'organisme était intégré aux instituts de recherche canadiens, plutôt qu'aux Archives nationales ou à un autre ministère du gouvernement fédéral, le financement de l'infrastructure pourrait provenir de la Fondation canadienne pour l'innovation et des fonds de contrepartie de divers organismes provinciaux.

Les services fournis par le réseau, et, par conséquent, ses coûts opérationnels, pourraient être augmentés au fil du temps, au fur et à mesure que croissent le nombre de dépôts et d'utilisateurs. C'est la voie habituelle qu'ont adoptée d'autres pays. Le volume de

<sup>3</sup> The Data Archive, University of Essex, "The Irish Data Archive Feasibility Project", 1997, p. 49.

données conservées n'a pas d'incidences importantes sur les coûts de fonctionnement, car le prix du stockage numérique baisse rapidement. L'expérience des autres pays démontre plutôt que le maniement et la gestion des données, ainsi que les services à valeur ajoutée augmentent en fonction de l'utilisation des services par la communauté des chercheurs et du fait que ces derniers prennent conscience de leurs avantages réels et potentiels.

## 10. Prochaines étapes

Pour lancer le processus de création d'un réseau national d'archives de données de recherche, le groupe de travail fait les recommandations suivantes :

- Le Conseil d'administration du CRSH et l'archiviste national devraient créer un comité directeur pour la mise en œuvre du processus et demander le soutien de l'Association des universités et collèges du Canada (AUCC), de l'Association des bibliothèques de recherche du Canada (ABRC), de la Fédération canadienne des sciences humaines et sociales (FCSHS) et autres organismes connexes, former des parrainages avec des ministères si une loi habilitante est requise, et s'assurer de la participation des intervenants au développement plus poussé des mandats et structures organisationnelles;
- Ce comité devrait établir des contacts appropriés avec les hauts fonctionnaires du ministère de la Justice et du ministère des Finances et avec d'autres représentants de ministères et d'organismes pertinents afin de démarrer le processus de mise en place du nouvel organisme proposé et d'élaborer les détails de son fonctionnement;
- Ce comité directeur devrait être responsable de l'élaboration des critères de sélection pour l'installation centrale et des nœuds du réseau national des archives de données de recherche;
- Le comité devrait également offrir des conseils relatifs aux critères de sélection des membres du Conseil de gestion, au besoin;
- Le groupe de travail recommande fortement que les activités débutent dès que possible afin d'établir le bien-fondé du rôle que peut jouer un réseau national d'archives de données de recherche pour faire progresser davantage la stratégie d'innovation du gouvernement du Canada.

## Annexe A

### Consultation sur les archives nationales de données

### Membres du groupe de travail et du Groupe d'étude

Président – M. John ApSimon, conseiller en sciences auprès du sous-ministre,  
Environnement Canada

M<sup>me</sup> Alexandra Bal, École des arts de l'image, Ryerson University

M. Paul Bernard, département de sociologie, Université de Montréal

M. Gérard Boismenu, département de sciences politiques, Université de Montréal

M. Ernie Boyko, directeur, bibliothèque et centre d'information, Statistique Canada

M<sup>me</sup> Sue Bryant, Secrétariat du Conseil du Trésor, coordonnatrice principale de projet,  
Secrétariat de l'Infrastructure à clé publique

M<sup>me</sup> Joanne Burgess, département d'histoire, Université du Québec à Montréal

M. Joseph Desloges, département de géographie, University of Toronto

M<sup>me</sup> Luciana Duranti, École de bibliothécaires, Études archivistiques et de l'information,  
University of British Columbia

M<sup>me</sup> Fay Hjartarson, directrice, agente de la Gestion des renseignements détenus par le  
gouvernement, Direction générale de la politique stratégique et de la planification,  
Bibliothèque nationale du Canada

M. Douglas Hodges, directeur de projets EPMS, Technologie de l'information,  
Bibliothèque nationale du Canada

M. Charles Humphrey, bibliothécaire de données, University of Alberta

M. José Igartua, département d'histoire, Université du Québec à Montréal

M. Ian Lancashire, département d'anglais, University of Toronto

M. Matthew Mendelsohn, département des études politiques, Queen's University

M. Michael Murphy, directeur du centre Rogers de communications, Ryerson University

M. Frits Pannekoek, directeur, Ressources de l'information, University of Calgary

M. Michael Ridley, bibliothécaire en chef, University of Guelph

M. Geoffrey Rockwell, École des arts, McMaster University

M. Fraser Taylor, département de géographie, Carleton University

M<sup>me</sup> Wendy Watkins, bibliothécaire de données, Carleton University

Gestionnaires – Consultation :

M<sup>me</sup> Yvette Hackett, agente des documents électroniques, Division des archives gouvernementales, Archives nationales du Canada

M. David Moorman, conseiller principal en politiques, Conseil de recherches en sciences humaines du Canada

## Annexe B

### Mandat

#### Contexte

En juin 2001, l'archiviste national et le Conseil d'administration du CRSH ont accepté le Rapport d'évaluation des besoins de la Phase I, du groupe de Consultation sur les archives nationales de données, y compris la recommandation de faire passer le processus de consultation à la Phase II. Le Conseil d'administration du CRSH a également demandé au groupe de travail de mettre en œuvre des activités en vue de sensibiliser davantage le milieu de la recherche aux avantages de l'archivage des données et d'examiner la possibilité d'intégrer de nouveaux organismes partenaires au processus de consultation. De plus, tant les Archives nationales que le Conseil d'administration du CRSH ont demandé que le mandat original de la Phase II soit révisé.

#### Mandat de la Phase II

Dans la Phase II de la consultation, on a demandé au groupe de travail d'examiner les questions suivantes avec l'aide du groupe d'étude, et de faire un compte rendu au Conseil d'administration du CRSH et à l'archiviste national :

- Étant donné qu'un nouveau système national ou une nouvelle fonction d'archivage des données de recherche est recommandé, comment devrait-on procéder pour le mettre en œuvre, quelles fonctions devrait-il assumer et quelle forme institutionnelle devrait-il prendre?
- De quelle façon un système ou une fonction d'accès aux données et de conservation de celles-ci peut-il tirer avantage des nouvelles technologies naissantes de l'information et des communications, pour devenir plus efficace?
- Quelle devrait être l'envergure des ressources documentaires gérées par le nouveau système ou la nouvelle fonction?
- De quelle façon un nouveau système ou une nouvelle fonction devrait-il traiter les questions d'éthique ou juridiques, telles que la confidentialité, la vie privée, le consentement, la propriété intellectuelle, le droit d'auteur, et les préoccupations d'ordre épistémologique, ainsi que les questions liées à l'authenticité et à la sécurité des données?
- Quelles seraient les relations de travail les plus appropriées entre un nouveau système ou une nouvelle fonction et les organismes existants, tels que les Archives nationales du Canada et la Bibliothèque nationale du Canada, en particulier en ce qui a trait aux données produites par le gouvernement? De quelle manière pourrait-on éviter le dédoublement des responsabilités et des services?

Au terme de l'étude, on a demandé au groupe de travail de présenter aux organismes responsables un rapport complet sur ce projet ainsi que des recommandations quant aux mesures que doivent prendre le Conseil, les Archives nationales, les universités et autres organismes connexes.

## Annexe C

### Modèles internationaux de services d'archivage de données

L'un des projets du sous-comité de la recherche du processus de Consultation sur les archives nationales de données était d'examiner les modèles internationaux de services d'archivage de données existants dans le domaine des sciences humaines. Un modèle des fonctions et des services a été élaboré par le sous-comité pour orienter la collecte de renseignements provenant de trente-six services de données et d'archivage représentant les organismes internationaux, chefs de file dans le domaine (voir l'annexe A pour consulter la liste complète de ces organismes). On a recueilli des renseignements détaillés sur les sites Web de chacun de ces organismes en vue de compléter le modèle dans la mesure du possible, à partir de ces sources en ligne. Par suite de cette analyse, on a consulté par courrier électronique chaque organisme afin de confirmer l'exactitude des renseignements recueillis et d'ajouter les éléments qui n'étaient pas disponibles sur les sites Web. Au total, vingt-deux établissements ont participé aux deux phases de la recherche, tandis que dans le cas des quatorze autres, les renseignements recueillis provenaient uniquement des sites Web.

Vingt-huit de ces organismes œuvrent dans le domaine des sciences sociales tandis que les neuf autres sont voués aux sciences humaines. Cinq des services en sciences humaines (ADS, VADS, HDS, PADS, et OTA) offrent des services de traitement de données sur des sujets précis dans le domaine des arts et des sciences humaines, au Royaume-Uni. La plupart des organismes ont une envergure nationale, et parfois, internationale. Les services qui semblaient être axés sur l'échelle régionale n'ont pas été inclus en raison de l'intérêt manifeste que nous portons aux services d'archivage de données à l'échelle nationale.

La typologie des modèles organisationnels basée sur l'examen de trente-six organismes est présentée ci-dessous. La description des modèles est suivie d'un résumé des caractéristiques utilisées pour élaborer cette typologie.

#### **Typologie des organismes d'archivage de données**

On a utilisé les résultats de cette étude pour élaborer la typologie des modèles organisationnels d'archivage de données. À partir des renseignements provenant de trente-six organismes internationaux existants, on a identifié trois modèles généralisés qui regroupent des caractéristiques de ces établissements<sup>4</sup>. En tant que description conceptuelle, ces trois modèles ne décrivent pas nécessairement un organisme en particulier mais représentent plutôt des catégories générales de modèles organisationnels. Comme la moyenne arithmétique, la valeur de la moyenne peut ne représenter aucune des valeurs courantes à partir desquelles elle est calculée. Néanmoins, la moyenne est utilisée

---

<sup>4</sup> Un modèle organisationnel supplémentaire a été déterminé au cours des délibérations du sous-comité de recherche. Bien que ce modèle n'apparaisse pas dans les établissements étudiés et ne soit pas considéré comme approprié par les membres qui ont participé à la Consultation sur un service national d'archivage de données, il est tout de même décrit à l'annexe C, à des fins d'information complémentaire.

comme valeur représentative des lots de nombres à partir desquels elle est obtenue. Dans le même ordre d'idées, les trois modèles ci-dessous représentent, à l'heure actuelle, les caractéristiques organisationnelles des services internationaux d'archivage de données. Même si aucun des trois modèles ne donne une description complète d'un seul organisme, la typologie des organismes d'archivage des données présente un résumé assez complet des divers organismes, tels qu'ils existent actuellement.

### **Modèle A : Les archives de données topiques**

Ce genre d'organisme détient un mandat précis pour desservir un domaine d'étude précis ou quelques disciplines connexes. Le service est situé dans un milieu universitaire et est souvent intégré à un département universitaire individuel. Le mandat de l'organisme est vraisemblablement établi par l'université d'accueil et son soutien financier dépend de l'aide en nature qu'il reçoit de l'organisme d'accueil. Des ressources financières additionnelles sont obtenues par le truchement de travaux à forfait et par l'imposition de frais aux utilisateurs. Le budget total variera de 200 000 \$ à 400 000 \$ CAN.

La structure organisationnelle comprend une unité dotée d'un maximum de cinq employés à temps plein et d'un nombre égal d'étudiants embauchés à temps partiel. Le niveau de traitement des données pour préparer les documents destinés aux archives et pour la diffusion de ceux-ci ne comprend qu'une ou deux étapes. Les formats sous lesquels les documents sont reçus se limitent à quelques pratiques acceptées à grande échelle afin de minimiser les étapes du traitement que doivent subir les produits d'acquisition. L'accès à la collection est fourni par le truchement d'Internet et la stratégie utilisée par le service pour la diffusion de la collection dépend des outils d'extraction des utilisateurs en ligne. La technologie couramment disponible est utilisée pour le soutien des activités de l'unité.

L'unité est la seule du genre dans son pays et est reconnue par d'autres établissements comme emplacement national pour ce service. Au fil du temps, le personnel acquiert du savoir-faire dans le domaine des services spécialisés, et sera éventuellement reconnu à l'échelle internationale pour les compétences qu'il aura développées.

### **Modèle B : Des archives de données intégrées à un organisme**

Ce genre d'organisme reçoit son mandat de l'organisme parrain. Cet organisme parrain est susceptible d'être un organisme national ou un ministère fédéral qui possède un mandat de recherche rigoureux. L'organisme parrain peut choisir de situer les services d'archivage de données sur un campus universitaire en vue de maintenir des liens étroits avec les chercheurs avec lesquels il entretient des relations de travail. Le mandat des archives de données fait partie du mandat plus étendu de l'organisme parrain. Les subventions aux instituts de recherche sont la principale source de revenus pour ce genre de services, bien que ses assises financières puissent être complétées par les droits sur le recouvrement des coûts et les frais imposés aux utilisateurs. Le budget total variera de 500 000 \$ à 1 500 000 \$ CAN. Certains coûts opérationnels seront assumés par l'organisme parrain, tels les frais généraux de l'administration et de la gestion des



dossiers et, par conséquent, ces frais ne seront pas déterminés de façon précise dans le budget des archives de données.

La structure organisationnelle comprend de deux à trois unités ayant un effectif de 10 à 30 employés à temps plein. Le nombre d'employés à temps partiel dépend de la proximité d'une université et de la disponibilité de la main-d'œuvre étudiante. La préparation des données et de la documentation pour l'archivage et la diffusion ultérieure se fait en trois ou quatre étapes de traitement. Les documents sont reçus dans des formats variés et convertis selon les normes des archives. Ces normes sont publiées et disponibles sur le site Web du service. L'accès à la collection est fourni directement en ligne et par l'intermédiaire d'un service de médiation qui est requis en raison des restrictions s'appliquant aux documents pour en assurer la confidentialité. Un employé a pour tâche de négocier et d'assurer la confidentialité des renseignements personnels sur tous les produits. Les services de référence sont offerts en vue d'aider les clients à localiser les données et à répondre aux demandes relatives aux produits.

Il est possible que ces archives ne soient pas uniques dans leur pays, mais elles sont reconnues à titre de service faisant autorité pour ce type de recherche directement associé à la recherche de l'organisme parrain. Le personnel a acquis un niveau d'expertise qui est reconnu à l'échelle nationale et internationale. Il joue un rôle de chef de file auprès des organismes nationaux et internationaux dont les activités sont connexes à son domaine de compétence et aux intérêts de recherche de l'organisme parrain. Les principaux gestionnaires du service participent aux projets et échanges internationaux portant sur les données communes et les intérêts de recherche, y compris le développement de nouvelles technologies et normes.

### **Modèle C : Les archives complètes de données de recherche**

Le mandat de ces archives de données est établi par une variété de groupes qui ont des préoccupations et des intérêts communs. Ces groupes comprennent des chercheurs en milieu universitaire, des analystes des politiques travaillant dans divers organismes gouvernementaux et non gouvernementaux, des fonctionnaires chargés de conserver les renseignements publics, de simples citoyens intéressés à l'accès à l'information et des participants à l'économie basée sur le savoir qui créent et utilisent des données de recherche. Leur mandat, qui vise à servir les intérêts de ces groupes, est établi en vertu de dispositions législatives. Ces archives de données sont reconnues comme un organisme national, responsable des principes généraux énoncés dans les dispositions législatives.

Ce service peut être situé dans plus d'un emplacement. Par exemple, on peut choisir de localiser certaines activités sur un campus universitaire et d'autres à proximité des principaux producteurs de données. Les subventions aux instituts de recherche constituent la principale source de revenus et le service accroît cette source de revenus en effectuant des travaux à forfait pour des organismes et producteurs importants de données, dans les secteurs public et privé. Le budget total du service varie de 3 à 6 millions de dollars canadiens. Certaines dépenses opérationnelles peuvent être

partagées avec d'autres organismes, telle une université où se trouve un service en succursale.

La structure des archives de données consiste en cinq ou six unités principales qui soutiennent l'administration, les archives/collections, le soutien technique, la recherche et développement, les services de référence, l'assistance à l'utilisateur, et le programme de formation. L'exécution des ces fonctions requiert un effectif de 60 à 100 employés, et jusqu'à 20 de ces postes sont, au départ, à temps partiel. L'unité des archives/collections est responsable de la sélection des documents sans égard à la discipline ou au domaine d'étude. L'évaluation de la valeur archivistique du produit est un élément essentiel du processus de sélection. Afin d'assurer la conservation des données et l'accès à celles-ci, jusqu'à six étapes de traitement sont nécessaires pour préparer les données et la documentation.

Les archives de données, qui ont leurs propres normes, reçoivent et traitent des documents présentés sous une vaste gamme de formats. Ces documents sont ultérieurement convertis au format réglementaire des archives. Des précisions concernant le format réglementaire sont accessibles tant sur le site Web du service que lors des ateliers d'information organisés à l'intention des producteurs de données comme partie du programme de formation de l'organisme. Au sein des archives/collections, une sous-unité s'occupe des questions entourant la protection des renseignements personnels et la propriété intellectuelle en vue d'assurer que ces droits sont protégés et que les données de recherche de valeur sont conservées. Cette sous-unité négocie et gère également les conditions d'accès aux données.

Une unité de services de référence/assistance à l'utilisateur facilite aux clients l'accès à la collection. Certaines données seront disponibles en ligne sur Internet, sans intermédiaire. Cette unité soutient également l'accès à d'autres données de diffusion restreinte ou dont l'utilisation est conditionnelle. Des séances de formation seront organisées afin de préparer les utilisateurs à travailler avec les données de la collection et, en règle générale, d'approfondir les connaissances en traitement de données.

Ce service est reconnu comme détenant un mandat national pour l'archivage des données de recherche. Toutefois, il travaille en étroite collaboration avec d'autres organismes nationaux, tels que la Bibliothèque nationale, les Archives nationales et autres archives du pays, à la coordination et à la conservation des données de recherche. Les employés des archives de données qui sont spécialisés dans leur domaine jouissent d'une renommée nationale et internationale et sont des chefs de file au sein des associations nationales et internationales rattachées à leur profession. Les employés utilisent également les archives de données aux fins de projets internationaux et d'échanges de données pour le compte de leur pays. Son unité de recherche et développement a acquis une renommée internationale pour l'adaptation et l'utilisation de la technologie comme moyen de repousser les frontières dans le domaine de la conservation et l'accessibilité des données.

Ceci complète la description des trois modèles généralisés pouvant être utilisés pour les services d'archivage de données qui ont été établis à partir des caractéristiques de trente-

six établissements existants. Dans la section suivante, les résultats qui ont servi à créer les trois modèles généralisés seront examinés.

## Résultats de la recherche

Un résumé des résultats de cette recherche est présenté ci-dessous. Voir l'annexe B pour consulter la liste complète du modèle des fonctions et services utilisés pour ordonner la collecte des renseignements dans le cadre de ce projet et l'annexe D pour obtenir des précisions au sujet de chacun des établissements.

### **Emplacement pratique**

La plupart de ces organismes (au total, 25) sont situés dans une université ou un département universitaire. Quatre des services restants se trouvent dans des instituts de recherche en sciences sociales de plus grande envergure, dont les SDA, la SDB, les SADA, et le SIDOS. Le reste des archives sont localisées dans des organismes indépendants (TARKI, ICSSR, CIS, DDA, SA, ADP, et WISDOM).

### **Mandat imposé par la loi**

Des 36 organismes, seulement six ont déclaré qu'ils existaient en vertu d'un mandat imposé par la loi. Toutefois, chacun de ces services a donné des interprétations différentes de la signification de « mandat imposé par la loi ». À Madrid, le CIS existe en vertu d'un décret royal du gouvernement espagnol. Il s'agit d'un petit organisme qui intègre également les fonctions de centre des enquêtes pour le pays. Un plus grand organisme, tel l'ICPSR, a sa propre constitution et ses propres règlements et dessert directement des membres affiliés rémunérés comprenant des fournisseurs et des utilisateurs de données. Quelques-uns des organismes existent en vertu d'une loi de leur organisme d'accueil. Par exemple, le statut du ESSDA a été déterminé et approuvé par la Tartu University, où les archives sont localisées. Au Royaume-Uni, les fournisseurs de services au Arts and Humanities Data Service développent actuellement des accords sur les niveaux de services offerts par l'organisme-cadre en vue de définir les responsabilités et de déterminer des objectifs précis.

### **Structure du corps administratif**

Vingt-deux établissements donnent à entendre qu'il sont dotés d'un corps administratif sous une forme ou sous une autre. De ces 22 établissements, dix ont un conseil d'administration ou un conseil consultatif composé de spécialistes en sciences sociales, de chercheurs ou de représentants d'établissements (universités, membres, etc.) qui ont des intérêts précis dans les données que fournit leur service. Le nombre de conseillers qui font partie de ces conseils varie de 6 à 35.

Les douze organismes restants déclarent qu'un établissement plus important ou un consortium régit leurs services. Par exemple, les ZA sont régies par le German Social Science Infrastructure Services. Le Research Council of Norway régit le NSD, dont la

structure de régie semble s'articuler autour d'autres responsabilités et organismes. Un autre exemple est celui du ADP slovène qui est administré par le ministère de l'Éducation, des Sciences et du Sport.

### **Origine du mandat de l'organisme**

En ce qui concerne l'origine du mandat, huit organismes déclarent que leur mandat est émis par un institut distinct et onze, par l'université où ils sont localisés. Le mandat de certains instituts est émis tant par l'université d'accueil que par un institut distinct. Par exemple, le mandat du ISSDA est émis conjointement par l'University College de Dublin et par un institut indépendant. La plupart des fournisseurs de services du Arts and Humanities Data Service reçoivent leur mandat directement du AHDS, bien que quelques-uns d'entre eux soient plus anciens que le AHDS (par exemple, les archives de textes d'Oxford). D'autres organismes, tels que l'ICPSR et les SADA, reçoivent leur mandat de consortiums ou d'instituts de la recherche sociale. Deux organismes seulement ont mentionné qu'ils recevaient leur mandat directement du gouvernement fédéral : le ICSSR, en Inde, et le CIS, en Espagne. Le TARKI fonctionne en tant qu'organisme autonome ayant des actionnaires et donc, ne dépend d'aucun organisme gouvernemental. Aux États-Unis, le consortium de données linguistiques (LDC) semble n'être tenu de rendre compte qu'à ses membres. Le mandat de huit établissements n'est pas émis par des organismes.

### **Rapports formels et relations officielles**

Au cours du suivi que nous avons fait auprès des organismes, les définitions de « rapports formels » et de « relations officielles » avaient tendance à être interprétées par les répondants, même si nos assistants de recherche avaient défini les termes. Les rapports formels ont été décrits comme des ententes entre des établissements ou des personnes relativement à l'échange ou au partage des données ou à la prestation de services, tandis que les relations officielles mettaient l'accent sur les associations avec d'autres, sans engagements contractuels, comme l'adhésion à IFDO ou CESSDA. Toutefois, certains organismes ont persisté à définir leur adhésion à IFDO ou CESSDA en tant que rapport formel. Aux fins du présent rapport, l'adhésion à un autre organisme sera considérée comme relation officielle.

La plupart des archives ont fait état de plusieurs rapports formels et relations officielles avec d'autres organismes ou personnes. Nombre d'entre eux ont déclaré entretenir des rapports avec des universités, des départements d'universités, divers niveaux du gouvernement (habituellement, au niveau national), des organismes internationaux d'archivage de données, et d'autres archives nationales de données, y compris l'adhésion à d'autres archives, en tant que membres participants. Seul le CETH a déclaré qu'il n'entretenait aucun rapport formel avec aucun autre organisme.

Les rapports qui ont été mentionnés le plus souvent (au total, 22) étaient les rapports avec d'autres archives de données ou instituts nationaux. Dix autres organismes ont mentionné entretenir des rapports formels avec des départements d'universités. Dix-sept organismes ont déclaré avoir des rapports formels avec des organismes internationaux et six ont fait

mention de rapports avec le gouvernement ou des ministères gouvernementaux. ICPSR, LDC et SSDA-ANU sont les seuls organismes qui maintiennent une forme d'affiliation avec les utilisateurs.

Les organismes ont mentionné un certain nombre de relations officieuses. La plupart de ces organismes (17, au total) semblent être axés sur l'échelle régionale et, de ce groupe, quatre ont précisé avoir des liens avec des établissements d'enseignement. Vingt et un organismes ont mentionné qu'ils sont affiliés à CESSDA; quinze, à IFDO; quatre ont déclaré leur adhésion à l'ICPSR. Comme exemple d'une autre affiliation internationale, neuf établissements ont mentionné qu'ils étaient membres de l'International Social Survey Program.

### **Sommaire budgétaire**

Dix-sept organismes ont confirmé les renseignements relatifs à leurs budgets. Ces établissements tirent des recettes de diverses sources, y compris les subventions, le soutien en nature de la part des établissements-hôtes (habituellement des universités), le travail à forfait, l'imposition de frais aux utilisateurs, et certaines activités de recouvrement des coûts.

Subventions aux instituts de recherche :

Les subventions aux instituts de recherche, définies comme soutien de la part de l'établissement-hôte, sont les principales sources de financement mentionnées par la plupart des organismes (au total, 29). De ceux-ci, douze organismes reçoivent des subventions aux instituts de recherche qui composent au moins la moitié de leur budget. Il s'agit des FSD (100 %), du SSJDA (100 %), des ESSDA (100 %), du SIDOS (90 %), des DDA (90 %), des ZA (80 %), des SA (90 %), du CIS (79 %), du ISDC (70 %) et du SSD (65 %), du HDS (50-60 %), et du UVA (50 %). D'autre part, le VADS, le NSD et l'ICPSR déclarent que les subventions aux instituts de recherche ne constituent que 30 %, 23 %, et 11 %, respectivement, de leurs budgets. Treize autres organismes reçoivent ce genre de soutien, mais n'ont pas mentionné la proportion de l'ensemble de leur budget que représentent les subventions. Trois organismes qui fournissent tous des services au Arts and Humanities Data Service ont mentionné le soutien « en nature » reçu des établissements hôtes.

Organismes subventionnaires :

Vingt-deux organismes reçoivent des subventions pour effectuer leurs travaux, dans des proportions variables. Le budget des PADS est subventionné à 100 %; les budgets d'ADP, de VADS et d'ADP sont subventionnés à 70 %, tandis que celui du NSD l'est à 56 %, et ceux de l'ADS, de l'UVA et du HDS, à 50 %. Trente-six pour cent des recettes de l'ICPSR proviennent de subventions. Dans le cas de quatre autres organismes, 15 % ou moins de leurs budgets provient de subventions. Treize autres organismes n'ont pas confirmé le pourcentage de leurs budgets qui provenait de subventions tandis que neuf autres ont déclaré ne pas recevoir de subventions.

Frais d'utilisation :

Seize établissements imposent des frais aux utilisateurs. Sept d'entre eux n'ont pas précisé quelle proportion de leur budget provenait de frais d'utilisation. Le reste du groupe a mentionné que les frais d'utilisation constituaient une proportion minime de leurs recettes – habituellement moins de 5 %. Dans le cas d'ICPSR (31 %), d'ISDC (20 %), et de NSD (21 %), le pourcentage des recettes provenant des frais d'utilisation était plus élevé.

#### Travail à forfait :

Onze établissements tirent des recettes du travail à forfait. Les recettes de l'ADS provenant du travail à forfait s'élèvent à 50 %, ce qui est le pourcentage le plus élevé déclaré par les organismes pour cette source de recettes, suivi de l'ICPSR, 20 %, et du CIS, 16 %. Le SIDOS, le SSD, les SA et les ADP tirent 5 % ou moins de leurs recettes du travail à forfait. Quatre organismes qui perçoivent de l'argent de cette source n'ont pas indiqué le pourcentage de ces recettes par rapport à leur budget. Seize organismes ne tirent aucunes recettes du travail à forfait.

#### Autres sources de recettes :

Seuls cinq organismes ont déclaré avoir d'autres sources de recettes dont le TARKI, en Hongrie, qui perçoit des recettes de ses investissements et actionnaires. L'ICPSR indique que 20 % de ses recettes proviennent du recouvrement indirect des coûts, tandis que les ADP et le SSD (20 % et 15 %, respectivement) signalent des subventions uniques de leur gouvernement en tant que partie de leur budget pour une année précise. Le UVA reçoit un certain montant d'une fondation. La plupart des organismes (au total, 31) ne reçoivent aucunes recettes d'autres sources.

#### Catégories de dépenses :

Les principales catégories de dépenses mentionnées par les organismes étaient la dotation et les salaires, l'administration et le fonctionnement, les archives, le développement du Web, la publicité, ainsi que la technologie. « Les déplacements et les conférences » ont été mentionnés par un organisme. Certains de ceux qui reçoivent du soutien en nature de l'établissement-hôte (tels que l'ADS, les PADS, l'UVA, les SA, et autres fournisseurs de services de AHDS) n'ont pas déclaré les services d'exploitation comme faisant partie de leurs dépenses. La plupart des répondants n'ont pas donné de précisions au sujet de leurs dépenses et n'ont mentionné que cinq catégories de postes budgétaires. Les UKDA et le SIDOS sont des exceptions, car ils ont fait état de plus de cinq catégories.

#### Taille totale du budget :

Dix-sept établissements ont déclaré le montant de leur budget annuel. Le CIS dispose du plus petit budget d'exploitation qui est légèrement supérieur à 9 000 \$ CAN, suivi par les ADP, de Slovaquie, (42 000 \$ CAN). Les budgets les plus considérables étaient ceux de l'ICPSR, (5 700 000 \$ CAN), des ZA (5,03 millions de dollars CAN), des UKDA, (3,5 millions de dollars CAN), et des DDA (1,397 millions de dollars CAN.). Les budgets des autres organismes varient de 958 000 \$ CAN (SIDOS), 838 000 \$ CAN (FSD), jusqu'à 565 000 \$ CAN (ADS). Le VADS a déclaré que son budget était « confidentiel ». Les SA ont défini leur budget comme étant « partagé » au sein de NIWI, et n'ont pu fournir aucun chiffre.

## Personnel

Le plus grand organisme est l'ICPSR, qui compte plus de 100 employés. Les organismes qui ont de 50 à 100 employés sont les ZA, le NSD, le CIS, le CIDSP et les UKDA. L'ICSSR et le SDB ont un effectif de 20 à 50 employés. Six organismes comptent de 10 à 20 employés (SSD, FSD, DDA, SSDA, SSJDA, et TARKI), et le reste des organismes comptent moins de 10 équivalents temps plein. Dix de ces organismes comptent moins de cinq employés.

On a demandé aux organismes de classer les rôles des employés sous trois catégories : administrative, technique, professionnelle. Le terme *professionnel* a été défini comme comprenant les archivistes, les bibliothécaires de données ou autres spécialistes qui ont des titres semblables. La catégorie *administrative* comprend les gestionnaires, les directeurs ou d'autres employés qui sont axés sur l'exploitation du service. La catégorie *technique* comprend les employés qui s'occupent des aspects technologiques de l'organisme et des ressources documentaires des archives.

La plupart des répondants ont mentionné que moins de dix employés étaient affectés à l'une de ces trois fonctions. Le SA a déclaré que, comme dans le cas de son budget, faire partie de NIWI signifiait qu'il était impossible de déterminer le nombre exact d'employés effectuant soit des tâches administratives ou techniques. Plusieurs organismes n'ont pas fourni de chiffres relatifs à l'une ou deux de ces fonctions.

On a également demandé aux établissements de préciser le nombre de leurs employés qui étaient affectés à la « conservation » et à « l'accès des données ». « L'accès » comprend toute tâche relative à l'assistance à l'utilisateur, la diffusion de l'information aux membres ou d'autres initiatives concernant les utilisateurs et les données. La « conservation » englobe les procédures et activités liées au maintien du contenu des archives de données.

La plupart des organismes ont mentionné que moins de 10 employés sont affectés soit à la conservation ou à l'accès. Dix-huit archives n'ont pas répondu à cette question.

Treize des trente-six établissements embauchent des employés temporaires. De ce groupe, sept embauchent moins de dix employés temporaires annuellement, tandis que six en embauchent plus de dix. Quatre organismes n'en emploient pas du tout et dix-neuf n'ont pas répondu.

## Structure organisationnelle

Plus de la moitié des établissements visés par l'étude ont une structure horizontale, et dix-sept d'entre eux ont mentionné qu'ils étaient désignés comme une simple unité d'exploitation. Les organismes plus importants, tels que les UKDA et l'ICPSR, sont dotés de quinze et cinq unités, respectivement. Le SIDOS, un organisme de taille moyenne, a six différentes unités organisationnelles. L'ADS, le CIS et le NSD maintiennent quatre unités (bien que l'ADS ait mentionné que leurs unités ne sont pas

« définies de façon formelle »). Cinq organismes maintiennent trois unités chacun (SDB, TARKI, HDS, LDC, ZA), tandis que six autres en maintiennent deux (ADPSS, SSJDA, DDA, FSD, SA, CIDSP).

Les titres des divisions dont on a fait mention le plus souvent sont administration, archives/collections, technique, recherche, assistance et services à l'utilisateur, et éducation. Les catégories projets externes et méthodologie, ainsi que eurobaromètre, mentionné par le SIDOS, n'ont reçu qu'une seule mention.

### **Champ d'activités des archives**

Vingt-huit organismes sont axés sur les sciences sociales, cinq sur l'histoire (y compris le SSD et les UKDA, qui rassemblent aussi des données en sciences sociales), cinq sur les beaux-arts, et un, sur la linguistique. Ces organismes reçoivent des données d'une variété de sources, y compris des sources multiples. Nombre des organismes (au total, 26) sont principalement axés sur les chercheurs, les universitaires ou sur d'autres « créateurs » de données. Les ministères ou d'autres organismes du secteur public fournissent des données à seize organismes. Treize de ceux-ci utilisent des ressources commerciales, quatre autres produisent leurs propres données ou les reçoivent de l'établissement dans lequel ils sont situés, et sept reçoivent des données d'autres organismes d'envergure nationale de leur pays. Seulement deux organismes ont mentionné qu'ils compilaient des données par le biais de partenariats d'échange internationaux (ICPSR, ISDC).

Les données recueillies ont habituellement une portée nationale et sont axées sur le pays ou se trouve l'organisme. Une importante exception est LDC, qui rassemble des données linguistiques dans le monde entier et une autre est l'ICPSR et le ISDC, qui participent à des échanges internationaux de données.

Les domaines sont variés. Par exemple, les fournisseurs de services au Arts and Humanities Data Service concentrent leurs efforts dans un ou deux domaines (p. ex. PADS, ADS, VADS). La plupart des établissements de sciences sociales rassemblent des données dans une vaste gamme de domaines comparativement aux archives des sciences humaines qui semblent posséder un mandat de collecte plus restreint. Quelques organismes collectent des données tant dans le domaine des sciences sociales que dans celui des sciences humaines. Par exemple, le SSD de Suède recueille des données dans le domaine des sciences sociales et de l'histoire.

Vingt archives ont des politiques d'acquisition écrites et seize de ces politiques sont accessibles en ligne. Sept archives n'ont pas de politiques formelles.

Dix-huit organismes publient leurs procédures en ligne, dont le degré de précision varie, à l'intention des dépositaires de données. Des dix-sept organismes qui ne publient pas de procédures de dépôt de données, trois (CETH, UVA et CIS) offrent des services dans le domaine des sciences humaines et concentrent leurs efforts sur des projets qui sont élaborés au sein de leur organisme.



Douze organismes abritent des sous-archives topiques qui comprennent tant des archives de données en sciences sociales qu'en sciences humaines. Les sous-archives sont divisées en fonction de divers critères : méthodes de recherche qualitatives et quantitatives (FSD), littérature selon la langue d'origine (UVA), présentation des sources (catégories linguistiques au LDC) et sujet (p.ex., ICPSR, ESSDA, etc.). L'UVA possède le plus grand nombre de sous-archives, soit quatorze, tandis que le reste des organismes en possèdent deux à cinq. Quatorze organismes ont répondu qu'ils ne conservaient pas de sous-archives de ce genre.

### **Dépôt de données**

Quinze organismes ont donné des renseignements concernant le nombre de niveaux de traitement qu'ils utilisent après le dépôt des données. Huit de ces organismes ont expliqué leurs méthodes de façon précise. L'ISDC, en Israël, et l'ICPSR, aux États-Unis, utilisent quatre ou cinq niveaux de traitement. Les NHDA, aux Pays-Bas, en utilisent neuf, les ZA en utilisent huit, et les FSD, jusqu'à trois. Les plus petits organismes, tels que le CETH, acceptent la plupart des formats et leur façon de procéder consiste simplement à verser les ensembles de données. En Suède, le SSD établit une distinction entre le traitement complet et le traitement partiel des données, selon la qualité des ensembles de données déposés. Vingt et un organismes n'ont pas mentionné le nombre de façons de procéder qu'ils utilisent, mais on présume que l'étape d'accès aux ensembles de données est observée.

Les façons de procéder dont on a fait mention le plus souvent sont la vérification des données afin de déceler des erreurs, la vérification de l'anonymat, la création ou la vérification de la documentation, la conversion au format réglementaire, l'uniformisation du marquage en langages XML ou SGML, l'élaboration de métadonnées ou l'obtention d'aide technique pour les ensembles de données, la création de copies de sauvegarde, et enfin, la diffusion des données.

Nombreuses sont les archives qui acceptent une large gamme de formats, en particulier, les archives de données en sciences humaines qui conservent les données que contiennent les documents. Dix-huit archives ont publié leurs formats réglementaires en ligne, à l'intention des dépositaires, tandis que dix ne publient pas ces renseignements.

Vingt-quatre archives ont également publié sur leurs sites ou sur copie papier les formats auxquels les données sont converties. Le CETH, à Rutgers University, ne conserve que les textes en langage XML, dont la présentation est réglementaire. Les formats dont il est fait mention le plus souvent comprennent SPSS (pour les archives de données en sciences humaines, PDF et TIFF (pour les listes de codage ou autre documentation), XML, DDI (Initiative de documentation des données), ASCII et MS Word.

Les autres standards descriptifs le plus souvent mentionnés sont XML/ASCII, PDF, DDI, et RTF. Vingt-deux archives publient en ligne de la documentation sur leurs standards descriptifs à l'intention des dépositaires. Douze organismes ne publient pas ces renseignements en ligne, et l'UVA et le CETH n'utilisent pas de standards descriptifs.

La principale préoccupation des archives en sciences humaines est d'assurer l'anonymat de la source de données, à chaque niveau. Seize archives ont rédigé une politique qui décrit leur façon de procéder en ce qui concerne l'anonymat et les priorités, mais la plupart des archives en sciences humaines n'ont pas mentionné ce sujet comme étant prioritaire (à l'exception du History Data Service et du Archaeology Data Service, pour certains types de données).

Seize archives ont rédigé des politiques concernant le stockage hors site des données, tandis que douze ne l'ont pas fait. La plupart des formats courants des données stockées hors site comprennent le réseau redondant de disques indépendants (RAID), les DC, les disquettes et les bandes de sauvegarde. Deux des cinq fournisseurs de services pour AHDS (ADS et PADS) ont signalé que leurs politiques de stockage des données hors site évoluent et sont actuellement examinées avec AHDS, ce qui signifie que tous les fournisseurs de services de AHDS peuvent coordonner cette fonction avec leur organisation-cadre.

Plusieurs archives dont le CIS, le CETH, les FSD, les NHDA, les ZA, et les SA ont indiqué que leurs procédures de reprise du traitement après sinistre sont liées au stockage hors site. Aucune archive n'a fait mention d'une politique écrite élaborée, bien que l'ICPSR, l'ISDC, l'ADS, et les PADS aient déclaré que l'élaboration de leurs politiques est en cours. Treize archives ont déclaré n'avoir aucune politique en vigueur.

### **Groupes d'utilisateurs et services aux utilisateurs**

Trente-trois organismes ont désigné les organismes universitaires comme sources principales de clients. Dix-sept ont mentionné que la proportion d'universitaires utilisant les services variait de 80 à 100 %, quatre, de 41 à 80 % et deux de 1 à 40 %. Douze établissements ont reconnu que les utilisateurs des universités font partie de la clientèle mais n'ont pas mentionné de pourcentage.

Un nombre moindre d'organismes a conclu que les autres utilisateurs du secteur public faisaient partie de leur clientèle; dix organismes ont estimé que le pourcentage des autres utilisateurs du secteur public variait de 1 à 40 %. Dix autres organismes n'ont pas donné le pourcentage d'utilisateurs, mais ont mentionné des personnes ou des groupes du secteur public comme étant des utilisateurs. Neuf archives ne comptent aucun utilisateur du secteur public.

Vingt organismes ont indiqué que des utilisateurs du secteur privé faisaient partie de leur clientèle, mais la plupart de ces clients ne constituaient que 40 % du nombre total d'utilisateurs. De ces vingt organismes, douze n'ont pas donné de proportion comparativement à d'autres groupes d'utilisateurs, et treize ont déclaré ne desservir aucun utilisateur du secteur privé. Seul le CIS, en Espagne, a signalé que 52 % de ses utilisateurs proviennent du secteur privé.

En ce qui concerne les services, tous les organismes ont un site Web qui diffuse des renseignements dont le degré de précision varie. Une différence observée entre les grandes et les petites archives réside dans le fait que les petits organismes offrent moins

de services et certainement, moins de services en ligne. Douze organismes ont offert des services de référence aux utilisateurs, et de ceux-ci, trois sont des fournisseurs de services au Arts and Humanities Data Service . Six archives n'offrent aucun service de référence. Quatorze offrent des services de référence par le truchement de contacts avec leur personnel, par courrier électronique ou par téléphone. La plupart des organismes exploitent leur service de référence comme un centre d'appel, les services étant offerts par le biais de contacts avec le personnel.

La plupart (vingt-deux) des organismes fournissent certains cours de formation par le truchement d'ateliers individuels ou un programme plus poussé de cours d'été (p. ex., l'ICPSR). Le programme de formation du CIDSP est principalement organisé à l'intention des employés temporaires, qui sont des étudiants au doctorat engagés par l'établissement. Tous les organismes de AHDS organisent des ateliers qui semblent être coordonnés par leur organisation-cadre.

Les publications préparées et diffusées par les organismes comprennent des bulletins d'information, des rapports annuels, des disques compacts de données, des publications sur les données et l'archivage de données, et dans le cas du Arts and Humanities Data Service, des « Guides sur les bonnes pratiques » pour la conservation des données et des archives dans le domaine des sciences humaines. Comme mentionné plus tôt, tous les organismes ont un site Web; et toutes, à l'exception de quatre, préparent des publications.

Quatorze organismes offrent des services d'extraction de données. Certains de ces services, telles les archives de donnée en sciences humaines de CETH et de ADS, considèrent l'extraction des données comme la réalisation de leurs propres projets. D'autres, tels que le CIS, font leurs propres études et extraient des données à partir de ces sources. Treize établissements ont déclaré ne pas extraire de données pour le compte des clients.

Il semble que la plupart des archives nationales offrent un accès direct ou indirect à leurs bases de données. Toutefois, l'accès direct aux fichiers de données semble être une priorité courante pour plusieurs archives nationales. Vingt-cinq archives ont mentionné qu'elles fournissent au moins un service d'accès à leurs collections. Certains de ces services sont réalisés par le truchement d'une coalition ou d'un autre organisme. Douze archives n'offrent aucun service d'accès direct.

L'accès direct aux données semble être le service que veulent offrir la plupart des archives, si elles ne le font pas déjà. Vingt-deux archives négocient l'accès aux données tandis que neuf ne fournissent aucun accès de ce genre. De ces neuf archives, la plupart sont des archives de données en sciences humaines.

Les résumés des services reflètent une vaste gamme de priorités quant aux services qu'offrent ces organismes aux utilisateurs. Outre les services examinés dans les paragraphes précédents, la plupart des archives assument des fonctions de bibliothéconomie et nombreuses sont celles qui publient le catalogue de leur collection en ligne à l'intention des utilisateurs. Certains des services particuliers offerts comprennent

l'analyse des données (NORC, SSD, SA, ICPSR), l'élaboration de logiciels (NSD, LDC, ZA, ICPSR, CETH), la préparation des ensembles de données (ISDC), l'attribution de bourses d'étude (ICSSR), des services de recherche de données dans d'autres archives (VADS, SSDA), et la création de projets indépendants au sein des établissements (CETH). Le CIS réalise ses propres études et en diffuse les résultats et le NORC prépare des questionnaires et élabore d'autres outils de collecte de données. Certains organismes comme l'ICPSR participent à des projets internationaux et à des échanges de données.

Des organismes tels que l'ICPSR et les Archives de données du Royaume-Uni offrent des services très étendus aux utilisateurs. Les archives de moyenne à grande taille sont susceptibles de participer à des projets internationaux et à des échanges de données (LDC, ISDC, SSD, ZA, membres de NESSTAR). Les archives de plus petite taille qui offrent moins de services comprennent ADPSS, en Italie, SSJDA, au Japon, ISSDA, en Irlande, et CIS, en Espagne. Les activités de ces organismes ont tendance à se concentrer à l'échelle nationale.

### **Surveillance du taux d'utilisation**

Deux des paramètres les plus souvent utilisés pour surveiller les niveaux d'activité des archives de données sont la taille de leurs collections et le nombre de clients qu'ils desservent. Par exemple, au cours de l'année financière 2000-2001, l'ICPSR a ajouté 372 titres de données à sa collection, ce qui constituait un taux de croissance de 8 p. 100 par rapport à l'année financière précédente. Ces nouveaux titres contenaient 1 835 fichiers de données, soit une augmentation de 4 p. 100 depuis l'année financière précédente. En ce qui concerne le nombre de clients desservis, l'ICPSR a diffusé cinq mille gigaoctets de données à l'intention de ses clients. Durant la même période de déclaration, les UKDA ont traité plus de 500 acquisitions et desservi 1 000 clients qui avaient passé 2 000 commandes, ce qui au total équivaut approximativement à près de 9 000 fichiers de données. Il s'agissait d'une augmentation de 2 000 fichiers de données diffusés aux utilisateurs au cours d'une période de trois ans.

Plusieurs archives enregistrent l'utilisation des statistiques en fonction du trafic sur le Web. Le nombre de données et de documents téléchargés est propre à ce type de mesure. Les OTA, par exemple, ont signalé plus de 18 000 téléchargements de textes électroniques au cours de l'année financière 1999-2000. Cette utilisation de l'électronique dépasse de 39 les commandes hors ligne des OTA. En plus des téléchargements de fichiers, le nombre de contacts avec les utilisateurs est également saisi à partir des statistiques apparaissant sur le Web. Par exemple, l'ICPSR a fait mention d'une augmentation des contacts avec les clients parce que plus d'utilisateurs effectuent de la recherche et des activités liées à l'enseignement sur Internet. Au cours des trois dernières années financières pendant lesquelles il a facilité l'accès en ligne à plus de ressources, l'ICPSR a déclaré une augmentation de plus d'un millier de gigaoctets de données auxquelles les utilisateurs ont eu accès.

Les archives de données tiennent également des statistiques aux fins d'autres services. Par exemple, le programme de formation de l'ICPSR soutient, sur une base continue,

l'inscription annuelle de 500 à 540 participants dans une variété de disciplines et dans divers pays. Comme autre exemple, les NSD tiennent des statistiques sur l'utilisation que font les chercheurs de leurs services à des fins d'examen des projets de façon à ce qu'ils soient conformes à la loi. Les NSD notent que ce service a connu une augmentation de 65 % au cours d'une année donnée. Les services de référence tiennent habituellement leurs propres statistiques. Pendant l'année financière 2000-2001, le personnel des UKDA a répondu à 332 demandes d'assistance ayant trait aux fichiers de données à la suite de commandes, ce qui ne représente qu'une facette des services de référence. L'ADS a déclaré avoir répondu à 174 demandes d'assistance en 2000, et les questions portaient sur l'utilisation du catalogue, l'assistance technique et des renseignements généraux dans le domaine de l'archéologie. Le HDS a reçu environ 480 demandes de références générales durant la même période. En plus de fournir des services de référence, les OTA font également les évaluations techniques des demandes de subventions. Durant l'année 2000-2001, les OTA ont fourni des services de consultation au sujet de 125 demandes de subventions.

Une autre statistique utilisée par certaines archives de données est le volume de permis d'utilisation de logiciel que leurs services préparent et distribuent. Par exemple, des permis d'utilisation du logiciel NSDstat, qui est mis au point par les NSD, sont délivrés à environ 2 000 établissements en Norvège, et 200 organismes internationaux. Bien que le nombre exact d'utilisateurs individuels par permis soit inconnu, l'expérience démontre que plusieurs personnes ont accès à NSDstat par le truchement d'un seul exemplaire du permis.

Les archives de données de plus grande taille peuvent également tenir des statistiques concernant leurs activités internationales. Par exemple, les ZA soulignent qu'elles reçoivent chaque année, sur une base continue, 50 chercheurs internationaux qui mènent des recherches sur place au moyen de données, au ZA EUROLAB. Les ZA intègrent également les données et de la documentation pour un certain nombre de projets internationaux, y compris l'International Social Survey Program pour le compte de 38 pays et les eurobaromètres, pour le compte de la Commission européenne.

Dans l'ensemble, tout indique que les archives de données qui offrent des services étendus (y compris de la formation, l'élaboration de logiciels et l'accès en ligne aux fichiers de données) sont fort utilisées par les chercheurs de renommée nationale et internationale.

### **Fonctions d'évaluation**

Quinze organismes repèrent l'utilisation de leurs données par le biais des publications des chercheurs, souvent par le truchement des requêtes que reçoivent ces derniers pour tenir les archives au courant des nouveaux travaux qu'ils ont réalisés au moyen des données fournies par l'organisme. Un certain nombre de ces organismes remarquent que, tout en souhaitant que le repérage des données devienne obligatoire, il n'est pas mis en pratique. Douze organismes ne repèrent pas les publications, mais deux autres (PADS et ICPSR) ont affirmé qu'un outil de repérage est en cours d'élaboration.

Encore moins d'organismes repèrent les citations qui font état de leurs collections. Seulement six mentionnent qu'ils effectuent cette fonction. Un des fournisseurs de services de données en sciences humaines pour le compte du Arts and Humanities Data Service, les PADS, remarque que cette fonction est en cours d'élaboration et devrait être en place d'ici deux ans. L'ICPSR publie un guide de citations sur son site Web à l'intention des chercheurs.

Treize établissements repèrent l'utilisation des données qu'ils fournissent directement aux utilisateurs par le truchement du Web. Plusieurs archives signalent qu'elles repèrent le nombre d'accès à leur site Web mais non l'utilisation des données d'une façon systématique. Ces archives comprennent l'ICPSR, les OTA, les PADS, et les NHDA. Aux États-Unis, le Linguistic Data Consortium ne repère l'utilisation des données qu'à la demande du dépositaire. Treize archives affirment qu'elles ne repèrent pas l'utilisation des données.

Le maintien d'une politique sur la propriété intellectuelle est une priorité importante pour vingt-sept organismes et elle est reflétée par les procédures qu'ils ont rédigées afin de protéger les droits des créateurs et des chercheurs. Les politiques en matière de protection des renseignements personnels semblent préoccuper davantage les archives de données en sciences sociales que celles en sciences humaines. En fait, trois archives de données en sciences humaines (en particulier les PADS, le CETH et l'UVA) mentionnent qu'elles ne voient pas l'utilité d'une politique pour protéger leurs données.

## **Technologie**

Les réponses aux questions portant sur la technologie étaient variées. Les organismes n'ont pas tous répondu à cette section du questionnaire. Par conséquent, l'information a souvent été recueillie à partir des sites Web, des rapports annuels ou d'autres publications. Parmi ceux qui ont répondu, la capacité d'entreposage de la collection variait de 4GB (SIDOS) à 83GB (NHDA). Les archives disposant d'un plus grand espace d'entreposage comprenaient le CETH (70GB), l'ICPSR (avec le stockage sur disque à haute vitesse RAID), et le SSD (50GB). Le service de données historiques, en tant que partie des Archives de données du R.-U., partage un stockage avec son organisme parrain et les PADS ont mentionné que leurs systèmes seront bientôt administrés par leur université d'appartenance.

Deux organismes seulement ont indiqué avoir des protocoles de sécurité, lesquels sont protégés par des pare-feu (ceux des SA sont protégés par leur organisation d'attache).

Les plates-formes les plus souvent mentionnées sont Windows 95-2000, UNIX, les serveurs NT4, les serveurs Linux, et les serveurs Sun 4500 Enterprise. La plupart des organismes mentionnent qu'ils utilisent des logiciels à code source ouvert, qui leur permettent d'adapter les programmes en fonction de leurs besoins. Nombreux sont les organismes qui ont mentionné utiliser Microsoft Office pour les ordinateurs des employés. Le stockage RAID a été mentionné par quelques établissements pour la

sauvegarde de copies. Les archives de textes, comme l'UVA et le CETH, utilisent également des logiciels de numérisation tels que Photoshop pour la manipulation des images et OmniPage Pro pour la reconnaissance optique de caractères.

Certaines archives utilisent des bases de données commerciales, y compris Dbase, Data Perfect, Filemaker Pro, FoxPro, Oracle, et Access. Les logiciels de statistique et les logiciels utilitaires qui ont été mentionnés le plus souvent comprennent SPSS, SAS, et StatTransfer. Plusieurs organismes mentionnent l'utilisation de XML (y compris DDI), en tant que convention de codification, en particulier en ce qui concerne les archives en sciences humaines.

## **Services de données et archives en sciences humaines visés par la présente étude**

<b>Acronyme</b>	<b>Nom et pays</b>
ADP	Data Archive for Social Sciences, Slovénie
ADPSS	Data Archive for Social Sciences, Italie
ADS	Archaeology Data Service; partie de Arts and Humanities Data Service, Royaume-Uni
CETH	Center for Electronic Texts in the Humanities, États-Unis
CIDSP	Centre d'informatisation des données socio-politiques, France
CIS	Centre for Social Research, Espagne
DDA	Danish Data Archives, Danemark
ESSDA	Estonian Social Science Data Archives, Estonie
FSD	Finnish Social Science Data Archive, Finlande
HDS	History Data Service; part of the Arts and Humanities Data Service, Royaume-Uni
ICPSR	Inter-University Consortium for Political and Social Research, États-Unis
ICSSR	Indian Council for Social Science Research, Inde
IRSS	Howard W. Odum Institute for Research in Social Science, États-Unis
ISDC	Israel Social Sciences Center, Israël
ISSDA	Irish Social Science Data Archive, Irlande
LDC	Linguistic Data Consortium, États-Unis
NHDA	Netherlands Historical Data Archive, Pays-Bas
NORC	National Opinion Research Center, États-Unis
NSD	Norwegian Social Science Data Services, Norvège
NZSRDA	New Zealand Social Research Data Archive, Nouvelle-Zélande
OTA	Oxford Text Archive; part of the Arts and Humanities Data Service, Royaume-Uni
PADS	Performing Arts Data Archive; part of the Arts and Humanities Data Service, Royaume-Uni
SA	Steinmetz Archive, Netherlands, Pays-Bas
SADA	South African Data Archive, Afrique du Sud
SDA	Sociological Data Archive, Czech Republic, République tchèque
SDB	Social Data Bank, Grèce
SIDOS	Service suisse d'information et d'archivage de données pour les sciences sociales
SSD	Swedish Social Science Data Service, Suède
SSDA-ANU	Social Science Data Archive – Australian National University, Australie
SSJDA	Social Science Japan Data Archive, Japon
TARKI	Social Research Centre Inc., Hongrie
UKDA	UK Data Archive, Royaume-Uni
UVA	Electronic Text Center, États-Unis
VADS	Visual Arts Data Service; part of the Arts and Humanities Data Service, Royaume-Uni
WISDOM	Vienna Institute for Social Science Documentation and Methodology, Autriche
ZA	Central Archive for Empirical Social Research, Allemagne



### Liste alphabétique par nom d'organisme

Data Archive for Social Sciences, Slovénie (ADP)  
 Data Archive for Social Sciences, Italie (ADPSS-Sociodata)  
 Archaeology Data Service; partie de Arts and Humanities Data Service, Royaume-Uni (ADS)  
 Center for Electronic Texts in the Humanities, États-Unis (CETH)  
 Centre d'informatisation des données sociopolitiques, France (CIDSP)  
 Centre for Social Research, Espagne (CIS)  
 Danish Data Archives, Danemark (DDA)  
 Estonian Social Science Data Archives, Estonie (ESSDA)  
 Finnish Social Science Data Archive, Finlande (FSD)  
 History Data Service; part of the Arts and Humanities Data Service, Royaume-Uni (HDS)  
 Inter-University Consortium for Political and Social Research, États-Unis (ICPSR)  
 Indian Council for Social Science Research, Inde (ICSSR)  
 Howard W. Odum Institute for Research in Social Science, États-Unis (IRSS)  
 Israel Social Sciences Center, Israël (ISDC)  
 Irish Social Science Data Archive, Irlande (ISSDA)  
 Linguistic Data Consortium, États-Unis (LDC)  
 Netherlands Historical Data Archive, Pays-Bas (NHDA)  
 National Opinion Research Center, États-Unis (NORC)  
 Norwegian Social Science Data Services, Norvège (NSD)  
 New Zealand Social Research Data Archive, Nouvelle-Zélande (NZSRDA)  
 Oxford Text Archive; part of the Arts and Humanities Data Service, Royaume-Uni (OTA)  
 Performing Arts Data Archive part of the Arts and Humanities Data Service, Royaume-Uni (PADS)  
 Steinmetz Archive, Netherlands, Pays-Bas (SA)  
 South African Data Archive, Afrique du Sud (SADA)  
 Sociological Data Archive, Czech Republic, République tchèque (SDA)  
 Social Data Bank, Grèce (SDB)  
 Service suisse d'information et d'archivage de données pour les sciences sociales (SIDOS)  
 Swedish Social Science Data Service, Suède (SSD)  
 Social Science Data Archive - Australian National University, Australie (SSDA-ANU)  
 Social Science Japan Data Archive, Japon (SSJDA)  
 Social Research Centre Inc., Hongrie (TARKI)  
 UK Data Archive, Royaume-Uni (UKDA)  
 Electronic Text Center, États-Unis (UVA)  
 Visual Arts Data Service; part of the Arts and Humanities Data Service, Royaume-Uni (VADS)  
 Vienna Institute for Social Science Documentation and Methodology, Autriche (WISDOM)  
 Central Archive for Empirical Social Research, Allemagne (ZA)

## Service virtuel d'archives de données offert sur une base de bénévolat

Un quatrième type de modèle d'organisation a été décelé lors des délibérations du sous-comité de recherche. Il s'agit d'une organisation virtuelle (OV)<sup>5</sup> existant sur un réseau diffusé dans une vaste zone et dont le personnel est composé de bénévoles qui partagent les ressources numériques. Bien qu'aucun des trente-six organismes visés par cette étude n'incorpore ce modèle, certains projets réalisés à l'extérieur de cette étude ont été désignés comme exemples d'organisation virtuelle. Par exemple, le projet d'informatisation documentaire Gutenberg reçoit le soutien d'un groupe de bénévoles qui codent dans des fichiers informatiques l'information contenue dans des livres dont le droit d'auteur a expiré. Ces fichiers sont par la suite transmis à un site central sur Internet où la version électronique de ces documents est disponible et peut être diffusée. La communauté à code source ouvert se comporte également comme un organisme virtuel. Les bénévoles élaborent des outils logiciels qui sont partagés et entretenus par cette communauté.

Le groupe de travail et le groupe d'étude n'ont pas appuyé le modèle OV comme modèle institutionnel recommandé aux fins des archives nationales de données de recherche du Canada. En premier lieu, aucun des résultats de la recherche effectuée par le sous-comité n'a donné plus de précisions sur le modèle OV. D'autre part, les délibérations des groupes au sujet de ce modèle ont suscité de graves préoccupations. Toutefois, on a suggéré que l'organisme recommandé mène des recherches sur les utilisations possibles et les fonctions du modèle OV. Cette recherche serait effectuée dans le cadre du programme de recherche et développement de ce nouvel établissement canadien.

Les points ci-dessous résument les avantages et les inconvénients du modèle OV que les groupes ont examiné. Ces renseignements documentent exhaustivement les modèles pris en considération au cours de la consultation.

### 1. Quels sont les inconvénients d'un modèle d'organisation virtuelle?

1.1 Contrôle de la qualité. Une organisation bénévole complique la tâche en matière d'assurance de la qualité des activités essentielles au sein du service, ce qui est particulièrement vrai dans le cas des tâches très complexes. Plus la tâche est complexe, plus il est difficile d'assurer le contrôle de la qualité du travail exécuté sur une base volontaire.

1.2 Service étendu. Une organisation bénévole accomplit bien les tâches qui intéressent les bénévoles et dont ils tirent avantage. Il s'ensuit qu'il est peu probable qu'une telle

---

<sup>5</sup> Pour consulter un exposé sur les organisations virtuelles, voir "The Anatomy of the Grid: enabling scalable virtual organizations," Ian Foster, Carl Kesselman, et Steven Tuecke, *The International Journal of High Performance Computing Applications*, vol. 15(3), 2001.

organisation fournisse des services étendus comprenant non seulement les tâches intéressantes mais également les tâches fastidieuses. On pourrait imaginer les résultats si la conservation des données du recensement était confiée à des généalogistes. Certains documents pourraient être bien conservés tandis que d'autres, qui intéressent peu les bénévoles, disparaîtraient vraisemblablement par négligence.

- 1.3 Niveau d'expertise. Dans une organisation bénévole, les connaissances spécialisées sont susceptibles d'être diffusées et, partant, de demeurer à un niveau amateur, ce qui pourrait être préjudiciable dans le cas de tâches qui exigent une formation spécialisée et des compétences peu communes.
- 1.4 Responsabilité. Une organisation bénévole n'a pas à rendre compte à personne et est difficile à gérer parce que l'on ne peut s'attendre à ce que les bénévoles accomplissent des tâches qui ne les intéressent pas. Généralement, ceux qui travaillent bénévolement ont un contrôle disproportionné sur les tâches exécutées parce qu'ils ne choisissent que celles qui les intéressent. Par conséquent, l'établissement de priorités et le soutien d'orientations stratégiques sont plus difficiles. De plus, les organisations bénévoles courent le risque de devenir restrictives et isolationnistes.
- 1.5 Permanence. La pérennité d'un modèle tient de la solidité de la communauté qui fait don de son temps. Si la communauté de bénévoles s'articule autour d'une personne, elle est susceptible de disparaître dans le cas de la retraite, du décès ou d'un changement de carrière de cette personne.
- 1.6 Sécurité et confidentialité. Une organisation bénévole aura de la difficulté à aborder les questions liées à la sécurité et à la confidentialité parce qu'elle dépend d'un grand nombre de bénévoles relativement peu supervisés qui n'ont peut-être pas les responsabilités professionnelles ou les compétences requises pour se voir confier des tâches en matière de sécurité et de confidentialité.

## 2. Quels sont les avantages d'un modèle d'organisation virtuelle?

- 2.1 Engagement communautaire. Un groupe de bénévoles peut faire appel à la collectivité pour qu'ils comprennent le besoin de la conservation des données et la valeur des services professionnels de conservation. Une organisation bénévole a l'avantage de bénéficier de la participation des personnes mêmes dont le soutien politique est requis.
- 2.2 Utilisation des ressources sous-utilisées. Une organisation bénévole peut aller chercher des ressources sous-utilisées, en particulier pour le traitement et le stockage des ressources du réseau. Par exemple, sur Internet, un nombre considérable d'ordinateurs personnels sont au repos. Un écran de veille conçu adéquatement ou un outil point à point qui exploiterait la capacité superflue de traitement et de stockage pourrait être distribué aux bénévoles sur la plupart des ordinateurs personnels en réseau.

- 2.3 Économie d'argent. L'organisation bénévole a la qualité d'être peu coûteuse. Toute tâche accomplie adéquatement par les bénévoles n'a pas à être rémunérée. Ceci dit, il faut reconnaître qu'il y a des coûts associés à la mise en œuvre d'une organisation bénévole, qui autrement ne seraient pas engagés.
- 2.4 Encouragement des donateurs. Une des questions importantes dans la réussite des archives de données de recherche est de persuader la communauté de contribuer aux archives. Une organisation bénévole constitue une façon d'encourager les donateurs avisés. Ceux qui participent à un projet sont plus enclins à y contribuer de façon appropriée et à passer le mot que les dons sont appréciés. Il faut trouver les moyens qui permettent aux membres de la communauté de contribuer. Par exemple, des outils de migration ou d'émulation devront être élaborés afin de maintenir l'accès aux données de recherche. Ces outils pourraient être élaborés au moyen du code source ouvert, qui a été utilisé avantageusement ailleurs. Par définition, les outils de migration en recherche doivent être ouverts afin que les utilisateurs de données de recherche puissent évaluer les processus qui affectent leurs données.
- 2.5 Élargissement de la portée des services. Une organisation bénévole pourrait s'occuper des dons qui ne sont pas inclus dans le mandat ou la capacité des archives centrales. Dans le cas où des données n'étaient pas clairement définies comme des données de recherche, une organisation bénévole serait peut-être en mesure d'aider à régler les problèmes entourant le contenu problématique. Par exemple, les dons de l'industrie qui ne sont pas sous forme de données de recherche ou les dons qu'on peut considérer comme des publications pourraient être traités gracieusement de façon à ne pas aliéner le donateur. L'organisation bénévole pourrait donner lieu à une organisation complémentaire des archives principales. En acceptant des tâches bien définies qui ne sont pas incluses dans le mandat de base des archives principales, l'organisation bénévole permet aux archives principales de se concentrer sur leur mandat de base. Une organisation bénévole pourrait aussi fournir des services coordonnés en collaboration avec d'autres organismes dont les mandats n'incluent pas les données de recherche.

On estime que certaines fonctions du modèle OV pourraient être incorporées à d'autres modèles d'organisation afin de compléter et d'enrichir les archives nationales de données. Il serait alors nécessaire de trouver un équilibre entre les bénévoles travaillant dans l'organisme d'appui et le personnel des archives nationales de données.

## Annexe D

### Aperçu des choix entre plusieurs systèmes institutionnels pour la création des archives nationales de données de recherche

La présente annexe donne un aperçu des diverses options entre plusieurs systèmes institutionnels pour la création d'un réseau d'archives nationales de données de recherche.

Les modèles incorporent les avis des membres du groupe de travail, les points de vue des directeurs des archives de données du R.-U. et de la Norvège, l'information recueillie de divers intéressés au cours du processus de consultation, et les conseils reçus du Bureau du Conseil privé et du Conseil du Trésor.

En général, dans le monde entier, les archives existantes de données ne sont conformes à aucun modèle spécifique. Toutes ont été constituées en fonction du contexte national particulier qui leur est propre. Toutefois, elles ont en commun l'objectif fondamental d'assurer que la collecte de l'information recueillie par des processus de recherche, ou à des fins de recherche, est conservée et gérée afin que les chercheurs puissent y accéder pour créer de nouvelles connaissances.

Les modèles qui sont décrits dans la présente annexe prennent en considération cet objectif fondamental ainsi que le contexte particulier du milieu canadien de la recherche, les mécanismes de soutien de la recherche, les structures institutionnelles et les systèmes de gestion publique au Canada.

#### 1) Nouvelle division des archives nationales du Canada

##### **Description :**

- La nouvelle division serait créée au sein du cadre de fonctionnement, de gestion et de responsabilité des Archives nationales existantes;
- **Exigences :** reformulation du mandat des Archives nationales; augmentation de base du budget; embauche de nouveaux employés possédant les compétences requises pour exploiter des archives de données de recherche; installations, soutien administratif et infrastructure technologique;

##### **Forces :**

- Prend avantage du cadre de fonctionnement, de gestion et de responsabilité qui existe actuellement et qui est bien établi;
- Prend avantage d'une culture archivistique existante et d'une infrastructure archivistique de premier ordre;
- N'exige pas de disposition législative ou de modification de la structure du gouvernement fédéral;
- La nouvelle division pourrait utiliser les services de soutien existants (p. ex., les ressources humaines, la comptabilité et le service de la paie, le soutien de la technologie de l'information, etc.).

**Faiblesses :**

- Exige que le mandat actuel des archives nationales soit reformulé et élargi;
- Exige une nouvelle orientation des services qui devront être adaptés aux besoins du milieu de la recherche plutôt qu'à ceux du grand public;
- Exige un emplacement unique et central sous l'autorité directe du gouvernement fédéral et, pour ces motifs, serait considéré comme faisant partie de la bureaucratie fédérale;
- N'a pas actuellement de relations bien établies avec l'ensemble de la communauté des chercheurs;
- N'a pas de rapports directs et soutenus avec les services de données des universités;
- N'a pas actuellement dans ses installations, à Ottawa, l'espace voulu pour une nouvelle division;
- Le financement des Archives nationales est sujet à des modifications, de même que l'ensemble des priorités budgétaires du gouvernement.

2) **Organisme de service spécial, intégré aux Archives nationales du Canada**  
(comme le Bureau des passeports)

**Description :**

- Tiré du *Cadre d'examen des différents modes d'exécution des programmes*. Il s'agit d'un « Service opérationnel d'un ministère constitué en organisme de service spécial (OSS) par le Conseil du Trésor, assujéti à un cadre législatif ministériel et aux politiques du Conseil du Trésor, qui doit rendre compte à l'administrateur général. La responsabilisation est établie dans un document cadre et dans un plan d'activités. L'OSS préconise l'application, au sein d'un ministère, d'un mode de fonctionnement qui ressemble davantage à celui du secteur privé. Il dispose de pouvoirs et d'une marge de manœuvre adaptés à ses besoins, délégués par le ministère et le Conseil du Trésor ».
- **Exigences :** autorisation du Conseil du Trésor et soutien de l'archiviste national; reformulation du mandat des Archives nationales et augmentation du budget de base des Archives nationales; embauche de nouveaux employés possédant les compétences requises pour exploiter des archives de données de recherche; des installations, du soutien administratif et une infrastructure technologique.

**Forces :**

- Possède une souplesse de fonctionnement et une certaine autonomie au sein des Archives nationales (doit rendre compte directement à l'archiviste national);
- Prend avantage du cadre de fonctionnement existant bien établi;
- Prend avantage de la culture archivistique existante et d'une infrastructure archivistique de premier ordre;
- Ne requiert aucune disposition législative ni modification de la structure du gouvernement fédéral;

- La nouvelle division serait en mesure d'utiliser les services de soutien existants (p. ex., les ressources humaines, la comptabilité et le service de la paie, le soutien de la technologie de l'information, etc.).

**Faiblesses :**

- Requiert l'élargissement et la reformulation du mandat des Archives nationales;
- Exige un emplacement unique et central sous l'autorité directe du gouvernement fédéral, et, pour ces motifs, sera considéré comme faisant partie de la bureaucratie fédérale;
- Devrait établir des rapports avec l'ensemble de la communauté des chercheurs;
- Devrait établir des rapports directs et soutenus avec les universités;
- Actuellement, les archives ne disposent pas des locaux nécessaires pour une nouvelle division dans ses installations à Ottawa;
- Le financement des Archives nationales est sujet à des modifications, de même que l'ensemble des priorités budgétaires du gouvernement.

3) **Organisme distinct en vertu d'une loi** (comme l'Agence spatiale canadienne, Statistique Canada)

**Description :**

- Tiré du *Cadre d'examen des différents modes d'exécution des programmes*. Les organismes distincts sont « des composantes de la Fonction publique... fonctionnant en vertu d'une loi constituante et relevant d'un ministre responsable ou d'un administrateur général. L'organisme distinct exécute ses fonctions dans le cadre politique établi par le ministre. »
- **Exigences :** Loi du Parlement; crédit annuel du Parlement, dans le cadre du processus budgétaire pour des fonds d'exploitation; des installations, le personnel, le soutien administratif et l'infrastructure technologique;
- L'organisme a aussi besoin d'un ministère d'attache, probablement Industrie Canada, et d'un ministre disposé à promouvoir sa mise en place.

**Forces :**

- Deviendrait un organisme permanent, protégé par une loi;
- Pourrait être considéré comme un élément de la stratégie d'innovation actuelle, axé sur la recherche, le renforcement des capacités, la gestion et la concurrence internationale;
- Autorité et responsabilité clairement définies dans le cadre du gouvernement fédéral;
- Aurait un parrain ministériel, et s'il était intégré à Industrie Canada, un ministre important au sein du Cabinet;
- Étant donné son statut imposé par la loi, l'organisme se verrait vraisemblablement attribuer des crédits budgétaires importants, dès sa création;
- Le financement de l'infrastructure et de son exploitation proviendrait d'une seule source;
- À l'instar de Statistique Canada, l'organisme aurait le potentiel et les moyens de se faire connaître à titre « d'organisme de confiance »;

- Aurait un statut de représentation nationale sur la scène internationale.

**Faiblesses :**

- Sa création nécessiterait des dispositions législatives, et, pour ces motifs, un ministre disposé à le promouvoir au Parlement;
- Le processus pourrait être long et complexe puisqu'il serait assujéti au programme législatif, aux critiques de l'opposition au Parlement, au processus de sélection des priorités du gouvernement, etc.
- Peut ne pas être considéré comme un projet ayant assez d'envergure ou d'importance pour être sanctionné par une loi;
- Serait perçu par les universitaires comme un organisme fédéral et aurait à faire des efforts pour établir des liens avec le milieu de la recherche;
- Constituerait, en quelque sorte, une installation centrale qui pourrait mettre en place des bureaux régionaux dans des universités.

4) Établissement public (comme le Conseil national de recherches du Canada)

**Description :**

- Tiré du *Cadre d'examen des différents modes d'exécution des programmes*. Il s'agit d'une « société constituée par une loi du Parlement... Elle exerce des fonctions administratives et mène des activités de recherche, de supervision et de réglementation de nature gouvernementale. L'établissement public est assujéti aux mêmes exigences et contrôles administratifs que les ministères, mais dispose d'une plus grande latitude dans ses décisions. »
- **Exigences :** Intervention législative du Parlement; attribution de crédits parlementaires annuels dans le cadre du processus budgétaire pour ses fonds d'exploitation; installations, personnel, soutien administratif et infrastructure technologique;
- L'organisme requiert aussi un ministère d'attache, probablement Industrie Canada et un ministre disposé à promouvoir sa mise en place.

**Forces :**

- Essentiellement les mêmes que ceux de l'organisme distinct créé en vertu d'une loi;
- L'organisme aurait, de plus, le même réseau de relations indépendantes que les conseils de recherche, ce qui lui donnerait une liberté d'action considérable;
- Un statut autonome lui permettrait d'entretenir des relations étroites avec le milieu de la recherche;
- Aurait un statut national de représentation officielle sur la scène internationale.

**Faiblesses :**

- Sensiblement les mêmes que ceux d'un organisme distinct en vertu d'une loi, à l'exception du fait que l'envergure de l'organisme lui permettrait d'être considéré comme un établissement public;



- À long terme, son statut est légèrement moins sûr que celui d'un organisme permanent constitué par une loi, mais l'organisme pourrait être considéré comme un organisme du gouvernement (aussi permanent que le CRSH, par exemple);
- Crée une installation centrale mais qui pourrait mettre en place des bureaux régionaux dans des universités.

#### 5) Partenariat public (comme le Centre canadien de statistique juridique)

##### **Description :**

- Tiré du *Cadre d'examen des différents modes d'exécution des programmes*. Il s'agit d'une « relation entre le gouvernement fédéral et au moins un autre ordre du gouvernement ou organisme gouvernemental, en vertu de laquelle les parties s'entendent pour collaborer en vue d'atteindre des objectifs compatibles. Ce partenariat se fonde sur une entente formelle précisant le but et la nature de l'association et les conditions qui la régissent, par ex. le mode de financement, la dotation, les rapports. L'entente de partenariat peut être assujettie à l'autorisation du gouverneur en conseil. Le partenariat public peut inclure des entreprises en coparticipation, c'est-à-dire des sociétés dont les titres de propriété sont détenus par le gouvernement fédéral et d'autres ordres du gouvernement ».
- **Exigences :** Décret pour la mise en place de l'organisme; ministre disposé à promouvoir l'initiative au sein du Cabinet; entente avec les provinces sur une vaste gamme de questions; bureaux, installations, personnel, soutien administratif et infrastructure technologique.

##### **Forces :**

- Même démarche que celle observée pour la création d'organismes semblables à un organisme d'archives de données;
- Un décret assure la une permanence presque autant que des dispositions législatives;
- Élargit vraisemblablement la portée de l'organisme en incluant des rapports directs avec les organismes de l'analyse statistique provinciaux;
- Financement relativement fiable mais provenant de sources multiples.

##### **Faiblesses :**

- Processus exceptionnellement complexe;
- Requiert des ententes relatives aux négociations entre de nombreux intervenants, ententes qui peuvent être longues à conclure;
- Exigerait l'intervention d'un ministre important et engagé au sein du Cabinet;
- Serait assujetti aux politiques et priorités budgétaires d'un nombre de gouvernements pouvant aller jusqu'à 14;
- Ne possède pas l'autonomie d'un établissement public ou l'indépendance d'un organisme constitué par une loi

#### 6) Centre intégré à une université. (comme TRIUMF, University of British Columbia)

**Description :**

- Un centre de soutien à la recherche, établi sur un campus universitaire, exploité comme une entreprise conjointe par des membres associés (universités), avec l'aide d'une contribution financière du CRSH;
- Une demande serait faite à la Fondation canadienne pour l'innovation en vue d'obtenir des fonds pour la mise en place de l'infrastructure;
- Le CRSH disposerait des fonds d'exploitation et les attribuerait au centre. Le centre serait tenu de rendre compte au président et au Conseil d'administration du CRSH;
- Chaque membre associé constituerait un nœud du réseau du centre et siègerait au Conseil de gestion;
- Chaque université canadienne dotée d'un service de traitement des données de recherche pourrait devenir membre associé et prendrait avantage des activités du centre en plus d'en être responsable;
- Le Conseil de gestion serait responsable de la supervision des activités du centre et de l'établissement de l'orientation des politiques générales;
- Un directeur serait chargé de la gestion du centre, de ses activités quotidiennes, des ressources humaines, des transactions financières et de la reddition de compte;
- Le directeur serait secondé par un Comité consultatif composé d'utilisateurs de données et d'archivistes;
- L'emplacement du centre pourrait être choisi par les membres associés au moyen d'un processus d'évaluation par les pairs géré par le CRSH.

**Exigences :** des fonds additionnels obtenus par le biais des crédits budgétaires du CRSH seraient transférés au centre (transfert de fonds); demande de fonds à la Fondation canadienne pour l'innovation aux fins de l'infrastructure, processus de sélection d'un emplacement; création d'une structure de régie et une structure organisationnelle; bureaux, installations, personnel, soutien administratif et infrastructure technologique.

**Forces :**

- Est élaboré d'après un modèle reconnu pour l'établissement sur un campus universitaire d'une installation de recherche financée par le gouvernement fédéral;
- Rapports directs et immédiats, sur place, avec la communauté des chercheurs;
- Le droit de propriété, le fonctionnement et les politiques seraient entre les mains des membres associés des universités;
- Utilise les services de données existants, l'expertise et l'infrastructure technologique dont disposent les universités;
- Pourrait utiliser un système hybride centralisé/décentralisé dans le cas où le centre serait responsable de la conservation et du traitement des ensembles de données et où les membres associés agiraient à titre d'intermédiaires pour l'accès aux données, le dépôt de données et qui offriraient des conseils sur place et organiseraient des activités éducatives;
- Les membres associés pourraient comprendre des organismes non universitaires tels que les archives provinciales;

- La portée de l'organisme peut varier et pourrait inclure des secteurs scientifiques du CRSNG et des IRSC, sous réserve d'ententes avec ces organismes au sujet du financement;
- Les activités d'archivage numérique prendraient avantage de la proximité des chercheurs universitaires en science de l'information.

**Faiblesses :**

- N'a pas la stabilité à long terme qu'assurent les dispositions législatives;
- Il n'existe aucun modèle équivalent en sciences sociales;
- Les fonds proviendraient d'au moins deux sources, soit la Fondation canadienne pour l'innovation et le CRSH;
- Il faudrait que les fonds avancés par la Fondation canadienne pour l'innovation soient complétés à hauteur de 60 % par des fonds de contrepartie (bien qu'il soit possible d'obtenir des fonds de contrepartie des programmes de financement provinciaux);
- La structure de régie exigerait de nombreux partenaires et serait relativement complexe;
- Les membres associés ne disposeraient pas tous d'un niveau semblable de ressources pour assurer les services locaux;
- L'organisme pourrait ne pas avoir le statut officiel de représentant national sur la scène internationale.

## Annexe E Bibliographie spécialisée

### Fonctions

- "Review: Data Historica: A Guide to Electronic Resources for Historians Compiled by the Netherlands Historical Data Archive." *Historical Social Research* [Germany] 23(4), 1998.
- Anderson, B. "Secondary Analysis and Gateway Libraries." *Behavioral and Social Sciences Librarian*. 17:2. 109-110 (1999).
- Austin, Erik W. and Dunn Christopher S. "Protecting Confidentiality in Archival Data Resources." *IASSIST Newsletter*. 22:2. 16-24 (1998).
- Bennett, J. C. "Archiving and Data Management." *Annual Conference – 1996 Sep: York*: National Preservation Office; 1997.
- Bonatti, P., et al. "An Access Control Model for Data Archives." *International Conference; 16th – 2001 June: Paris*: International Federation for Information Processing. Boston; London; Kluwer Academic; c2001.
- Buck, Michael. "Problems in Cataloguing Computer Files." *International Cataloguing and Bibliographic Control* v18 (1989).
- Carley, Michael, and Josefina J Card. "The Social Science Electronic Data Library: Serving the Needs of Data Librarians and Users." *IASSIST-(International-Association-for-Social-Science-Information; 24: No 3 Fall 2000* (2000).
- Church, J. "Getting Our Data Used: A National Statistics Office Perspective." *NATO Advanced Research Workshop – 1997 Aug : Colchester*: IOS Press; Ohmsha; 1998.
- Consultative Committee for Space Data Systems. "Reference Model for an Open Archival Information System (OAIS): Recommendation Concerning Space Data Systems Standards." White Book CCSDS 650.0-W-4.0, September 17, 1998. [http://ssdoo.gsfc.nasa.gov/nost/isoas/ref\\_model.html](http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html)
- Corti, Louise, Janet Foster, and Paul Thompson. "The Need for a Qualitative Data Archival Policy." *Qualitative Health Research* 6, 1, Feb (1996).
- Crothers, A. L. "Institute for Research in Social Science Louis Harris Data Archive (Social and Behavioral Sciences) (Brief Article)." *CHOICE: Current Reviews for Academic Libraries* v38 Special Issue p. 131(2). (2001).
- Dorobek, Christopher J. "Project Aims to Modernize Data Preservation Process." *Government Computer News* v18 ( August 30, 1999).
- Duffy, C. "An Introduction to the Performing Arts Data Service." *Literary & Linguistic Computing : Journal of the Association for Literary and Linguistic Computing*: 1997.
- Economic and Social Research Council. "ESRC Green Paper on Data Policy and Data Archiving" October 2000. <http://www.ilrt.bris.ac.uk/ubris/esrc/p3.html>
- Egret, D., and F. Genova. "The Role of Data Centres in the Era of Electronic Publishing." *Astrophysics and Space Science Library*. 224, (1997) D. Reidel Pub. Co.

- Finke, Roger; et al. "Democratizing Access to Data: the American Religion Data Archive" *IASSIST-(International-Association-for-Social-Science-Information)*; 23: No 1 Spring 1999.
- Fraser, Michael. "From Concordances to Subject Portals: Supporting the Text-Centred Humanities Community." *Computers and the Humanities* v 34 No3 Aug 2000 p. (265-1978).
- Gabrynowicz, Joanne Irene. "The Work of the US National Satellite Land Remote Sensing Data Archive Advisory Committee." *Space Policy* ( Feb2001).
- Green, Ann. Dionne JoAnn. Dennis Martin. "Preserving the Whole: A Two-Track Approach to Rescuing Social Science Data and Metadata." Washington, DC: Digital Library Federation, 1999.
- Jesske Muller, Birgit, and Christoph Reichert. "Social Research on Developing Countries: Reasons for and Aims of a Development Studies Data Collection." *Historical Social Research / Historische Sozialforschung* 15-2 (1990).
- Jones, Margaret. "A Workbook for the Preservation Management of Digital Materials." *New Review of Academic Librarianship* v 6 (2000).
- Jones, Roger. "Australia's Electronic Heritage: Comments on the ACLIS Task Force Recommendations: Augmented Title: Perspectives on the Towards Federation 2001 Conference, Canberra, March 1992." *Australian Library Review* v10 Nov 1993 p. (447-1956).
- Kinzy S. "Data Warehousing Moves GIS Beyond Maps." *GIS-World* 1998; pp. 1958-1961.
- Kuula, Arja. "Making Qualitative Research Material Reusable: Case in Finland." *IASSIST-(International-Association-for-Social-Science-Information)*; 24: No 2 Summer 2000.
- Leadbeter, Deana, and Lorna Guinness. "Development of a Health Data Archive for Bangladesh: an Example of a Cost Effective and Sustainable Approach to Information Sharing in a Developing Country." *IASSIST-(International-Association-for-Social-Science-Information)*; 23: No 2 Summer 1999.
- Lesaoana, Maseka A. "Data Archiving in Africa: The South African Experience." *IASSIST Newsletter*. 21:1. 4-7. (1997).
- Levada, Y. "The Russian Data Infrastructure." *NATO Advanced Research Workshop – 1997 Aug : Colchester*: IOS Press; Ohmsha; 1998.
- Lievesley, D. "Access to European Data." *Conference – 1995 Dec: Dublin*: Dublin; Institute of Public Administration, 1996.
- . "Making Data Available - the Experience of the ESRC Data Archive." *Seminar on the Future History of Our Landscape – 1992 Oct: London*: Royal Society of London; 1994.
- . "The Role and Future Development of the Data Archive." *Conference – 1995 Dec: Dublin*: Dublin; Institute of Public Administration; 1996.
- . "The Value of Data Dissemination to Both Data Users and Producers." *NATO Advanced Research Workshop – 1997 Aug: Colchester*: IOS Press; Ohmsha; 1998.
- Lievesley, D, and I Masser. "Socio-Economic Data: Making the Pieces Fit." *GIS-Europe* 1994.
- McCook, Kathleen-de-la-Pena. "Social Scientific Information Needs for Numeric Data: the Evolution of the International Data Archive Infrastructure." *Collection Management* v 9 Spring 1987. P. 1-53.

- Mochmann, E. "European Co-Operation in Social Science Data Dissemination." *NATO Advanced Research Workshop – 1997 Aug: Colchester*: IOS Press; Ohmsha; 1998.
- National Research Council. "Preserving Scientific Data on Our Physical Universe: A New Strategy for Archiving the Nation's Scientific Information Resources". Washington, DC: National Academy Press, 1995.
- Nielsen, Per. "Merging Cultures: Danish Integration of Academic Data Service into Traditional Archive System." *IASSIST Newsletter*. 19:2. 27-35 (1995).
- Owen, Catherine, et al. "Meeting the Challenge of Film Research in the Electronic Age" *D-Lib-Magazine*; 6: No 3 Mar 2000.
- Rasmussen, Karsten Boye. "Documentation - What We Have and What We Want: Report of an Enquete of Data Archives and Their Staff." *IASSIST Newsletter*. 19:1. 22-35 (1995).
- Richards, Julian D. "Preservation and Re-Use of Digital Data: the Role of the Archaeology Data Service." *Antiquity* v 71 Dec 1997 pp. 1057-1997.
- Ryssevik, Jostein, and Simon Musgrave. "The Social Science Dream Machine: Resource Discovery, Analysis, and Delivery on the Web." *Social Science Computer Review* Vol. 19 Issue 2 ( May2001).
- Scheuch, Erwin K. "From a Data Archive to an Infrastructure for the Social Sciences." *International Social Science Journal* [Great Britain] 1990 42(1): 93-111.
- Schurer, K. "Twenty-Five Years of Preserving and Disseminating Data: the Experience of the ESRC Data Archive. An Abstract." *7th Anglo-Nordic Seminar – 1994 Oct : Windsor*: Nordic Council for Scientific Information and Research Libraries. The British Library; Research and Development Department. NORDINFO; 1995.
- Singh, Jagtar, et al. "Accessing Indian Numeric and Statistical Data: a Critical Study of the Suprastructure and Infrastructure in India." *IASSIST-(International-Association-for-Social-Science-Information)*; 24: No 3 Fall 2000.
- Tanenbaum Eric, and Mochmann Ekkehard. "Towards a European Database for Comparative Social Research." *Historical-Social-Research / Historische-Sozialforschung*; 1996, 21, 2(78), 118-122.
- Treadwell, W. "DDI, the Data Documentation Initiative: An Introduction to the Standard and Its Role in Social Science Data Access." 2000 Jul: *Chicago, IL*: American Library Association. Scarecrow Press; 2002.
- Vries, Repke-Eduard-de. "ILSES: How Library and Data Archive Meet in Active Support of Research in the Social Sciences." *INSPEL* v 31 No4 1997. P. 234-41.
- Whelan, B. J. "An Irish Social Science Data Archive: Prospect and Problems." *Conference – 1995 Dec: Dublin*: Dublin; Institute of Public Administration; 1996.

## Mandate / Principles / Ethical Issues

- "Access to Data From Medical Research." *New England Journal of Medicine* (12/25/86).
- "Case for Free Access to Fundamental Data." *Lancet (North American Edition)* v 353 No9166 May 22 1999. p. 1721
- "Conservatives Challenge Science Community on Data Access." *Issues in Science and Technology* v 15 No4 Summer 1999.
- "Data Access and Privacy Concerns." *Monthly Labor Review* v 117 Jan 1994.
- "Enhancing Access to Primary Sources Internationally: A Joint RLG-CURL Symposium Held at University College London" June 1997. <http://www.rlg.org/psprog.html>
- "Health Privacy Overview: Background and Current Law." *Congressional Digest* v 79 No8/9 Aug/Sept 2000. p. 194-196 .
- History Data Service (HDS). "Collections Development Policy." <http://hds.essex.ac.uk/collpol.asp>
- ICPSR Collections Policy - <http://www.icpsr.umich.edu/ORG/Policies/colldev.html>
- "Law on Access to Research Data Pleases Business, Alarms Science." *Pediatrics* (Oct99 Part 1 of 2).
- "Neuroscientists Debate Data Sharing . Virtual Structure for Canada's Biomedical Research." *Nature*: 2000.
- "News - Scientists Await Impact of Public Data Sharing" *The Scientist*: 2000.
- "Open Exchange of Scientific Data Should Be Protected." *Public Health Reports* v 112 July/Aug 1997. p. 268 .
- "Preserving and Sharing Statistical Material." Working Group on the Preservation and Sharing of Statistical Material: Information for Data Producer/UK Data Archive (UKDA). <http://www.data-archive.ac.uk/home/PreservingSharing.pdf>
- "Protecting Privacy in an Information Sharing Environment." *Fredericton: The Task Force*, 1994.
- "Reinforcing Access to Research Data." *Nature* (1/18/96).
- Adam, John. "Research in Electronic Age Should Be Open." *IEEE Spectrum* (Jun97).
- Alberts, Bruce, and Kenneth Shine. "Scientists and the Integrity of Research." *Science* v 266 Dec 9 1994.
- Aldhous, Peter. "Prospect of Data Sharing Gives Brain Mappers a Headache." *Nature* v 406 No6795 Aug 3 2000. p. 445-446.
- Allen, R S. "An Overview of the Federal Geographic Data Committee, National Spatial Data Infrastructure, National Geospatial Data Clearinghouse, and the Digital Geospatial Metadata Standard: What Will It Mean for Tomorrow's Libraries?" *Bulletin, Special Libraries Association, Geography and Map Division* 180 (1995).
- Altman, Micha, et al. "A Digital Library for the Dissemination and Replication of Quantitative Social Science Research: The Virtual Data Center." *Social Science Computer Review* 19, No. 4 (2001): 458-470.

- Anderson, R. "Undermining Data Privacy in Health Information." *BMJ*. 322(7284):442-3, 2001 Feb 24. .
- Annas, George J. "Rules for Research on Human Genetic Variation- Lessons From Iceland." *The New England Journal of Medicine* v 342 No24 June 15 2000 p. 1830-2000.
- Applegate, David. "AGI Workshop Examines Geoscience Data Preservation." *Geotimes* v 41 (1996).
- Archaeology Data Service (ADS). "ADS Collections Policy: 3<sup>rd</sup> edition."  
<http://ads.ahds.ac.uk/project/collpo13.html>
- Austin, Erik W. and Dunn Christopher S. "Protecting Confidentiality in Archival Data Resources." *IASSIST Newsletter*. 22:2. 16-24 (1998).
- Barber, B., P. Skerman, and R. Saldanha. "The Definition of Data Privacy for Europe." *Conference – 1997 Mar: Harrogate: BJHC Limited; 1997.*
- Baron, James N. "Data Sharing as a Public Good." *American Sociological Review* 1988 53, 1, Feb, vi-viii.
- Beedham, Hilary. "The Royal Statistical Society Working Group on Archiving Data." *IASSIST- (International-Association-for-Social-Science-Information); 23: No 3 Fall 1999 (1999).*
- Blume, Peter ed. "Special Feature: The Citizen in the Information Society; Legal Issues From the Perspective of the Citizen." *Journal of Information, Law and Technology (JILT); Issue 1 F 27 1998.*
- Campbell, Paulette Walker. "Controversial Proposal on Public Access to Research Data Draws 10,000 Comments." *Chronicle of Higher Education (4/16/99).*
- Carew, R. "Institutional Arrangements and Public Agricultural Research in Canada." *Rev-Agric-Econ* 23 (2001).
- Carter, J, and -R. "Perspectives on Sharing Data in Geographic Information Systems." *Photogrammetric-Engineering-and-Remote-Sensing. 1992. 58(11), Pp 1557-1560.*
- Ceci, Stephen J. "Scientists' Attitudes Toward Data Sharing." *Science Technology & Human Values. V13 N1-2 P45-52 Win-Spr 1988 .*
- Cooley, William W. "Confidentiality of Education Data and Data Access." Pennsylvania Educational Policy Studies Number 7.
- Corti, Louise, Janet Foster, and Paul Thompson. "The Need for a Qualitative Data Archival Policy." *Qualitative Health Research* 6, 1, Feb (1996).
- Craig, W. J. "Why We Can't Share Data: Institutional Inertia." *Meeting Entitled "Institutions Sharing Geographic Information" – 1992 Feb: National Center for Geographic Information and Analysis. New Brunswick; Center for Urban Policy Research; 1995.*
- Crissman, James W, James S Bus, and Ron R Miller. "Toxicology: Judge Data or Dollars?" *Environmental Health Perspectives* v 107 No10 Oct 1999.
- Curry, M. R. "Data Protection and Intellectual Property: Information Systems and the Americanization of the New Europe." *Environment-and-Planning-A. 1996. 28/5, 891-908.*
- da Silva, Wilson. "Europe Gets Tough on Data Privacy." *New Scientist* v 154 May 10 1997. P. 10.



- Dalton, Rex. "Young, Worldly and Unhelpful All Miss Out on Data Sharing." *Nature* v 404 No6773 Mar 2 2000.
- Dangel, E. T. 3rd, and B Mishkin. "Collaboration and Data Sharing: Continued." *Science*. 1996 Jan 12;271(5246):129-30.
- Data and Information Systems Panel. "Data Policy and Barriers to Data Access in Canada: Issues for Global Change Research: A Discussion Paper" Canadian Global Change Program [http://www.globalcentres.org/cgcp/english/html\\_documents/publications/data/toc.htm](http://www.globalcentres.org/cgcp/english/html_documents/publications/data/toc.htm)
- Dove, Alan. "Research Data Access Bill Issued." *Nature Medicine* v 6 No1 Jan 2000.
- . "US Public May Gain Access to Research Data." *Nature Medicine* v 5 No1 Jan 1999. P. 8.
- Downing, John D H. "Computers for Political Change: PeaceNet and Public Data Access." *Journal of Communication*. v39 n3 p154-62 Sum 1989.
- Dreier, T. "Development and Perspectives of Copyright: From Gutenberg to Data Highways'." *International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies – 1995 Aug : Vienna: Munich; Vienna; Oldenbourg; Osterreichische Computer Gesellschaft; 1995.*
- Dresser, Rebecca. "Accountability in Science and Government: Is Access the Answer?" *The Hastings Center Report* v 30 No3 May/June 2000.
- Dueker K, et al. "A Geographic Information System Framework for Transportation Data Sharing." *Transportation-Research-Part-C:-Emerging-Technologies* 2000; (1936).
- Duncan, Joseph W. "Confidentiality and the Future of the U.S. Statistical System." *American Statistician* Vol. 30 Issue 1 ( Feb76).
- Dunn, Christopher S. and Erik W. Austin. "Protecting Confidentiality in Archival Data Resources." ICPSR Bulletin, Fall 1998. <http://www.icpsr.umich.edu/ORG/Publications/Bulletin/Fall98/article.html>
- Eberle, Carl Eugen. "Legislation to Protect the Confidentiality of Personal Data: Implications for Social Science Research; Implikationen Des Datenschutzes Fur Die Empirische Sozialforschung." *Zeitschrift-Fur-Soziologie; 1981, 10, 2, Apr, 196-211.*
- Eliot Marshall. "Genome Researchers Take the Pledge: Data Sharing." *Science* 1996 v272 n5261 p477(2).
- Elliott, P, et al. "Intersalt Data - Science Demands Data Sharing." *British Medical Journal* 315 (1997).
- Els, C., T. Verschoor, and H. Oosthuizen. "The Protection of Personal and Medical Data—a Call for Confidentiality." *South African Medical Journal*. 85(8):773-5, 1995 Aug .
- Estabrooks, C. A. Romyn DM. "Data Sharing in Nursing Research: Advantages and Challenges." *Can J Nurs Res*. 1995 Spring;27(1):77-88.
- Farrell, Rita. "Who Owns the Knowledge?" *UNESCO Sources* (May99 Issue 112).
- Fellegi, I. P. "On the Question of Statistical Confidentiality." *Journal of the American Statistical Association* (Mar72).
- Fienberg, S. E. "Statistical Perspectives on Confidentiality and Data Access in Public Health." *Statistics in Medicine*. 20(9-10):1347-56, 2001 May 15-30 .

- Frankel, M. S. "Public Access to Data." *Science*. 1999 Feb 19;283(5405):1114.
- Gibson, J. L. "Cautious Reflections on a Data-Archiving Policy for Political Science." *Political Science & Politics* 28 (1995).
- Goodwin, Candice. "Personal Privacy v Public Health." *New Scientist* v 140 Dec 11 1993.
- Gordon, W. "Commentary on Economical and Ethical Reasons for Protecting Data." *Fordham Annual Conference; 8th – 2000*: Juris Publishing; 2001.
- Greenstein, Daniel I, and Sarah Porter. "Scholars' Information Needs in a Digital Age: Executive Summary." *New Review of Academic Librarianship* v 4 1998.
- Hammersley, Martyn. "Qualitative Data Archiving: Some Reflections on Its Prospects and Problems." *Sociology* Feb 1997 v31 n1 p131(12).
- Haseltine, Florence P. "Unexpected Trouble With the Freedom of Information (FOIA): An Open Letter to the Public." *Journal of Women's Health & Gender-Based Medicine* (Apr99).
- Hilgartner, Stephen. Brandt-Rauf Sherry I. "Data Access, Ownership, and Control: Toward Empirical Studies of Access Practices." *Knowledge: Creation, Diffusion, Utilization*. V15 N4 P355-72 Jun 1994 .
- Howell Robert G. "Database Protection and Canadian Laws, State of Law As of June 15, 1998.". Ottawa: Industry Canada ; Hull, PQ : Canadian Heritage, 1998.
- Imwalle, Christine. "Data Warehousing Enables Sharing." *American City and County* v 113 No8 July 1998.
- Isaacs, Lindsay. "Shared Responsibility, Shared Reward." *American City and County* v 115 No15 Nov 2000. P. 44-52 .
- Jeffords, James. "Confidentiality of Medical Information: Protecting Privacy in an Electronic Age." *Professional Psychology, Research and Practice* v 30 No2 Apr 1999. P. 115-116.
- Johnson, David J. and Michel E. Sabourin. "Universally Accessible Databases in the Advancement of Knowledge From Psychological Research ." *International Journal of Psychology* (Jun2001).
- Jones, Roger. "Comment on 'The Need for the Preservation of Australian Created Electronic Information: a Position Paper'." *Australian Academic and Research Libraries* v 24 Mar 1993 pp. 1947-1993.
- Kaiser, Jocelyn. "Researchers and Lawmakers Clash Over Access to Data." *Science* v 277 July 25 1997 p. 467.
- Karr, Alan F., and Jaeyong Lee. "Disseminating Information but Protecting Confidentiality." *Computer* (Feb2001).
- Kastner, T. "On the Need for Policy Requiring Data-Sharing Among Researchers Publishing in AAMR Journals: Critique of Conroy and Adler (1998)." *Ment Retard*. 2000 Dec;38(6):519-29.
- Key, Sandra W., et al. "Innovations in Research Threatened by Proposals to Restrict Exchange of Data." *AIDS Weekly Plus* (01/05/98).
- Kiernan, Vincent. "Debate Continues on Cost of Global Access to Scientific Data." *Laser Focus World* (Jul97).

- Lawler, Andrew. "Treaty Draft Raises Scientific Hackles." *Science* v 274 Oct 25 1996 p. (494).
- . "U.S., Europe Clash Over Plan to Set Policy on Data Access." *Science* v 268 Apr 28 1995 p. (493).
- Litman, J. "Information Privacy/Information Property." *Stanford Law Review* .
- Longley, P. "Intellectual Property Rights and Digital Data." *Environment and Planning*.
- Love, Denise E., Luis M.C. Paita, and William S. Custer. "Data Sharing and Dissemination Strategies for Fostering Competition in Health Care." *Health Services Research* 36 issue 1 p. 277. (2001).
- Macilwain, Colin. "Business Lobby Set to Take EPA to Court Over Data Access." *Nature* v 403 No6767 Jan 20 2000 p. 236.
- . "Researchers Resist Copyright Laws That Could Endanger Data Access." *Nature* (10/24/96).
- . "Scientists Fight for Right to Withhold Data." *Nature* Vol 397 No 6719 (1999).
- . "US University Fears Allayed Over Access to Research Data." *Nature* (10/21/99).
- Marshall, Eliot. "Agencies, Journals Set Some Rules." *Science* v 248 May 25 1990. P. 954.
- . "Data Sharing: a Declining Ethic?" *Science* v 248 May 25 1990. P. 952+.
- . "Epidemiologists Wary of Opening Up Their Data." *Science* v 290 no 5489 (2000).
- Maurer, S. M. Hugenholtz PB, Onsrud HJ. "Intellectual Property. Europe's Database Experiment." *Science*. 2001 Oct 26;294(5543):789-90..
- Maurer, SM. "Coping With Change: Intellectual Property Rights, New Legislation, and the Human Mutation Database Initiative." *Human Mutation*. 15:22-29, 2000
- Mauthner, Natasha S., Odette Parry, and Kathryn Backett-Milburn. "The Data Are Out There, or Are They?: Implications for Archiving and Revisiting Qualitative Data." *Sociology* Nov 1998 v32 i4 p733(1).
- Mellors, Colin, and David Pollitt. "Policing the Communications Revolution: a Case-Study of Data Protection Legislation." *West-European-Politics*; 9:195-211 O 1986.
- Miller, Heather G, and Wendy H Baldwin. "A Terse Amendment Produces Broad Change in Data Access." *American Journal of Public Health* v 91 No5 May 2001 p. (824-2001).
- Mills, M. E. "Data Privacy and Confidentiality in the Public Arena." *Proceedings / AMIA Annual Fall Symposium*. :42-5, 1997 .
- Mishkin, B. "Urgently Needed: Policies on Access to Data by Erstwhile Collaborators." *Science*. 1995 Nov 10;270(5238):927-8..
- Nobel, J. "Data Confidentiality and Data Access - Practical and Legal Issues in the Netherlands." *International Seminar – 1994 Nov : Luxembourg: Eurostat*.; 1995.
- Normile, Dennis. "Asian Network Seeks Data Sharing." *Science* v 267 Mar 31 1995 p. 1902.
- Ochert, Ayala. "US Fury Over Data Access Legislation." *Times Higher Education Supplement* Issue 1387 (06/04/99).

- Olsen, Florence. "Publishers and Users of Data Bases Disagree Over Intellectual-Property Legislation." *Chronicle of Higher Education* (7/9/99).
- Peterson, M.J.. "Community and Individual Stakes in the Collection, Analysis, and Availability of Data. (Response to Articles by Gary King and Paul S. Herrnson in This Issue, P.444 and P.452, Respectively)." *PS: Political Science & Politics* Sept 1995 v28 n3 p462(3).
- Pickett, R., and D. Quackenbush. "Lock the Door and Throw Away the Key? . If We Cannot Provide Access to Data, Why Collect It?" *Annual Conference – 1995 Nov : New Orleans; LA: CAUSE. CAUSE; 1996.*
- Plater, S. Seeley E, Dixon LA. "Two Routes to Privacy Protection: a Comparison of Health Information Legislation in Canada and the United States." *J Womens Health. 1998 Aug;7(6):665-72.*
- Porter, J, et al. "Circumventing a Dilemma: Historical Approaches to Data Sharing in Ecological Research." *in: Environmental Information Management and Analysis. (Taylor & Francis), 1994, Pp 193-202.*
- Porter, Sarah. "Reports From the Front: Six Perspectives on Scholar's Information Requirements in the Digital Age." *New Review of Academic Librarianship v 4 (1998).*
- Posey, D A, G Dutfield, and K Plenderleith. "Collaborative Research and Intellectual Property-Rights." *Biodiversity and Conservation.*
- Ready, Tinker. "Will Research Be Restricted by New Data Privacy Rules." *Nature Medicine v 7 No2 Feb 2001.*
- Reidpath, Daniel D, and Pascale A Allotey. "Data Sharing in Medical Research: an Empirical Investigation." *Bioethics v 15 no2 (2001).*
- Rennolls, K. "Intersalt Data - Science Demands Data Sharing." *British Medical Journal* 315 (1997).
- Robbin, Alice, and Heather Koball. "Seeking Explanation in Theory: Reflections on the Social Practices of Organizations That Distribute Public Use Microdata Files for Research Purposes." *Journal of the American Society for Information Science and Technology New York* 52 (2001).
- Rockwell, R. C., and R. P. Abeles. "Sharing and Archiving Data Is Fundamental to Scientific Progress." *Journals of Gerontology. Series B, Psychological Sciences & Social Sciences. 53(1):S5-8, 1998 Jan.*
- . "Guest Editorial: Sharing and Archiving Data Is Fundamental to Scientific Progress." *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences v 53B No1 Jan 1998.*
- Rowen, L, G K S Wong, and R P and others Lane. "INTELLECTUAL PROPERTY: Publication Rights in the Era of Open Data Release Policies." *Science: 2000.*
- Rudder, Catherine-E. "APSA Responds to OMB's Draft Regulations." *PS, Political Science and Politics v 32 No2 June 1999. P. 188-189.*
- Samuelson, P. "Copyright, Digital Data, and Fair Use in Digital Networked Environments." *Conference – 1994 May : Montreal; Canada: Kluwer Law International; 1995.*
- Santoro, M. D., and S Gopalakrishnan. "Relationship Dynamics Between University Research Centers and Industrial Firms: Their Impact on Technology Transfer Activities." Chicago, Illinois: *Technology Transfer Society v. 26 (1-2) ( Jan 2001.).*

- Schafer, U. "Public Archives Between Data Access and Data Protection." *DLM-Forum; 2nd – 1999 Oct : Brussels*: Commission of the European Communities. Luxembourg; Office for Official Publications of the European Communities; 2000.
- Schwartz, Ephraim, Heather Harreld, and Tom Sullivan. "Data Sharing's Call to Arms." *InfoWorld Framingham* 23 (2001).
- Shaman, Susan M. Shapiro Daniel. "Data-Sharing Models." *New Directions for Institutional Research. N89 P29-40 Spr 1996* .
- Shelby, R. "Accountability and Transparency: Public Access to Federally Funded Research Data." *Harvard Journal On Legislation*.
- Sheppard, M. "Data Ownership and Copyright." *8th Annual Symposium on Geographic Information Systems in Forestry, Environmental and Natural Resources Management – 1993 Feb : Vancouver; Canada*: Polaris Conferences; 1994.
- Sieber, Joan E ed. "Sharing Social Science Data: Advantages and Challenges." Sage Publications, 1991.
- Siegfried, Susan L. "The Policy Landscape." *The Art Bulletin* v 79 June 1997 p. (209-1997).
- Simpson, Roy L. "How Can We Keep Private Data Private?" *Nursing Management* (May 2001).
- Singer, E., and S. Presser. "Public Attitudes Toward Data Sharing: Findings and Implications." *Annual Meeting – 1997 Aug : Anaheim; CA*: American Statistical Association; Section on Survey Research Methods. ASA; 1997.
- Singer, Eleanor, Schaeffer Nora Cate, and Raghunathan Trivellore. "Public Attitudes Toward Data Sharing by Federal Agencies." *International Journal of Public Opinion Research* Fall 1997 v9 n3 p277(9).
- Skinner, T. "Improving Access to Unidentified Data." *International Seminar – 1994 Nov : Luxembourg*: Eurostat; 1995.
- Smaglik, Paul. "Publication Deal for Celera Sparks Row Over Data Access." *Nature* v 408 No6814 Dec 14 2000 p. (759).
- Stevens, A. R. "Collaboration and Data Sharing." *Science* 270(5243) (1995).
- Stevens, Ann R. "Ownership and Retention of Data." NACUA Publication Series. Washington, DC: National Council of Univ. Research Administrators.
- Sullum, Jacob. "Data on Demand." *Reason* (Feb99).
- Townsend, Sean, Cressida Chappell and Oscar Struijve "Digitising History: A Guide to Creating Digital Resources from Historical Documents." HDS/AHDS.  
[http://hds.essex.ac.uk/g2gp/digitising\\_history/index.asp](http://hds.essex.ac.uk/g2gp/digitising_history/index.asp)
- UK Data Archive "Collections Development Policy" <http://www.data-archive.ac.uk/depositingData/collectionsPolicy.asp>
- Varian, Hal R. "The Information Economy." *Scientific American* v 273 Sept 1995 p. (1995).
- Wheeler, David L. "A Bitter Feud Over Authorship." *Chronicle of Higher Education. V41 N38 PA8,12 Jun 2 1995* .

Whitman, E. D., M. E. Frisse, and M. G. Kahn. "The Impact of Data Sharing on Data Quality." *19th Annual Symposium – 1995 Oct : New Orleans; LA: American Medical Informatics Association. Hanley & Belfus; 1995.*

Williamson, Alan R. "Access to Data and Intellectual Property." *ASM News* 8 (1998).

Young, Douglas L. "A Comment on the Role of Professional Journals in Facilitating Data Access." *Journal-of-Agricultural-Economics-Research; 43:45-7 Spring 1991.*

## Organizational Models

"Access to Cultural Heritage Through an on-Line Multimedia Data Service: Application to the Archive Folders of France's General Inventory of Monuments and Art Treasures." *Conference – 1997 May : Barcelona; Spain: Graphic Communications Association. Graphic Communications Association; [1997].*

Humanities and Arts on the Information Highways: A Profile” Coalition for Networked Information. November 1997. <http://www.cni.org/projects/humartway/humartway-rpt.part1.html>

“Preserving Digital Information: Report of the Task Force on Archiving of Digital Information”. Commission of Preserving Digital Information and the Research Libraries Group, Inc. May 1996. <http://www.rlg.org/ArchTF/tfadi.index.htm>

"The Royal Statistical Society Working Group on Archiving Data. AU: Beedham,-Hilary." *IASSIST- (International-Association-for-Social-Science-Information); 23: No 3 Fall 1999.*

“Trusted Digital Repositories: Attributed and Responsibilities” RLG / OCLC. May 2002 <http://www.rlg.org/longterm/repositories.pdf>

Bainbridge, W. S. "International Network for Integrated Social Science." *Social Science Computer Review* .

Brophy, Peter, et al. “Towards a National Agency for Resource Discovery Scoping Study” British Library Research and Innovation Report 58. 1997. <http://www.ukoln.ac.uk/services/papers/bl/blri058/nard.html>

Corti, Louise, Janet Foster, and Paul Thompson. "The Need for a Qualitative Data Archival Policy." *Qualitative Health Research* 6, 1, Feb (1996).

Eiteljorg, H. "The Archaeological Data Archive Project." *Conference – 1994 Mar : Glasgow: Tempus Reparatum; 1995.*

Fonseca F, et al. "Ontologies and Knowledge Sharing in Urban GIS." *Computers-Environment-and-Urban-Systems* 2000; (251-271).

Gatenby, Pam. “Digital Archiving – Developing Policy and Best Practice Guidelines at the National Library of Australia” International Council For Scientific and Technical Information. Jan 2000. <http://www.icsti.org/2000workshop/gatenby.html>

Garcia, A. "A Conceptual Framework for Clinical Data Warehouses." *A Medical Informatics Odyssey: Visions of the Future and Lessons From the Past: Washington, DC: Hanley and Belfus; 2001.*

Haberkorn-Susan-B, and Traugott-Michael-W. "National Archive of Computer-Readable Data on Aging." *Special-Collections. 1982; Vol. 1 (No. 3-4): P. 73-83.*

- Harris, R. Olby N. "Archives for Earth Observation Data." *Space Policy* .
- . "Earth Observation Data Archiving in the USA and Europe." *Space Policy*.
- Haynes, David, and David Steatfield. "A National Co-Ordinating Body for Digital Archiving?" *Ariadne*. 15. May 1998. <http://www.ariadne.ac.uk/issue15/digital/>
- Hodge, Gail, and Bonnie C Carroll. "Digital Electronic Archiving: The State of the Art and the State of the Practice." International Council for Scientific and Technical Information / CENDI. April 1999. [http://www.icsti.org/99ga/digarch99\\_TOCP.pdf](http://www.icsti.org/99ga/digarch99_TOCP.pdf) ; [http://www.icsti.org/99ga/digarch99\\_ExecP.pdf](http://www.icsti.org/99ga/digarch99_ExecP.pdf) ; [http://www.icsti.org/99ga/digarch99\\_MainP.pdf](http://www.icsti.org/99ga/digarch99_MainP.pdf).
- Leadbeter, Deana, and Lorna Guinness. "Development of a Health Data Archive for Bangladesh: an Example of a Cost Effective and Sustainable Approach to Information Sharing in a Developing Country." *IASSIST-(International-Association-for-Social-Science-Information)*; 23: No 2 Summer 1999.
- Nielsen, Per. "Merging Cultures: Danish Integration of Academic Data Service into Traditional Archive System." *IASSIST Newsletter*. 19:2. 27-35 (1995).
- Pack, Thomas, and Jeff Pemberton. "A Harbinger of Change: The Cutting Edge Library at the Los Alamos National Laboratory." *Online Magazine*, March 1999. <http://www.onlineinc.com/articles/onlinemag/pack993.html>
- Porter, J, et al. "Circumventing a Dilemma: Historical Approaches to Data Sharing in Ecological Research." in: *Environmental Information Management and Analysis*. (Taylor & Francis), 1994, Pp 193-202 .
- Reed, Ken. Blunsdon Betsy. Rimme Malcolm. "The Social and Organizational Life Data Archive (SOLDA)." *Journal of Computing in Higher Education*. V11 N2 P63-74 Spr 2000 .
- Roberts, A. M. Cluckie IDGray LGriffith RJLane AMoore RJPedder MA. "Data Management and Data Archive for the HYREX Programme." *Hydrology and Earth System Sciences*.
- Roghamann, K. J., and R. J. Haggerty. "Mini Data Archives: the Rochester Child Health Studies Masterfiles." *Inquiry*. 9(1):66-71, 1972 Mar.
- Scheuch, Erwin K. "From a Data Archive to an Infrastructure for the Social Sciences." *International-Social-Science-Journal* 42, 1(123), Feb, (1990).
- Schürer, Kevin. "The Irish Data Archive Feasibility Project - A Report" The Data Archive, University of Essex.
- Tanenbaum Eric, and Mochmann Ekkehard. "Towards a European Database for Comparative Social Research." *Historical-Social-Research / Historische-Sozialforschung*; 1996, 21, 2(78), 118-122.
- Thomas, S. P. "Issues in Data Management and Storage." *Journal of Neuroscience Nursing*. 25(4):243-245, 1993 Aug .
- Waters. Donald J. "Toward a System of Digital Archives: Some Technological, Political and Economic Considerations: The Rearranging Effect Presented at the Association of Research Libraries Meeting, October 1998. <http://arl.cni.org/arl/proceedings/131/waters.html>

Zelenock, Tom, and Kaye Marz "Archiving Social Science Data: A Collaborative Process." ICPSR Bulletin, May 1997. Inter-University Consortium for Political and Social Research (ICPSR)  
<http://www.icpsr.umich.edu/ORG/Publications/Bulletin/May97/article.html>

## Technology Usage / Adaptations

- "Defending Secrets, Sharing Data : New Locks and Keys for Electronic Information." *Washington, DC : Congress of the U.S., Office of Technology Assessment, U.S. G.P.O., [1987] Vi, 187.*
- "Selected Papers From TEI 10: Celebrating the Tenth Anniversary of the Text Encoding Initiative." *Computers and the Humanities v 33 No1/2 Apr 1999 p. (206).*
- Badard T, and Richard D. "Using XML for the Exchange of Updating Information Between Geographical Information Systems." *Computers-Environment-and-Urban-Systems 2001; (2001).*
- Balstad Miller, R. "On-Line Access and Analysis of Data: Developing A Social Science Information Management Strategy." *1996 Mar : Arlington; VA: United States; Department of Commerce; Bureau of the Census. Department of Commerce; 1996.*
- Bentley, J. E. "A Data Access Challenge: Build a SAS Frame Application for Access to an Informix Data Warehouse." *Annual International Conference; 23rd – 1998 Mar : Nashville; TN: SAS Institute User Group. SAS; 1998.*
- Bonatti, P., et al. "An Access Control Model for Data Archives." *International Conference; 16th – 2001 Jun : Paris: International Federation for Information Processing. Boston; London; Kluwer Academic; c2001.*
- Brantgarde, Lennart. "Documentation of Academic Research Data and the Hypertext Concept- a Presentation of the Swedish System ELSA." *Tidskrift for Dokumentation v 48 No2 1993 p. 1956-1961.*
- Card, Josefina J. and McKean Elizabeth. "The Development of Software to Facilitate Use of Archived Data Sets." *IASSIST Newsletter. 19:3. 4-11 (1995).*
- Chappell, C. Struijve O. and Anderson S. "The History Data Service – Using Technology to Enhance Access." *IASSIST Newsletter. 22:4. 17-20. (1998).*
- Chen, Su-Shing. "PERSPECTIVES - The Paradox of Digital Preservation - Preserving Digital Information Requires Proven Methods for Maintaining, Accessing, and Authenticating Technology-Generated Data." *Computer: 2001.*
- Chiang, Katherine-S, Jan Olsen, and William-V Garrison. "Beyond the Data Archive: the Creation of an Interactive Numeric File Retrieval System." *Library Hi Tech v 11 No3 1993. P. 57-72 .*
- Cleden, D, Proud R , and Hills N . "New Interface Approach to Complex Data Archive Systems for Remote Sensing and Mineral Exploration Applications." *in: Remote Sensing: an Operational Technology for the Mining and Petroleum Industries. Conference, IMM, London, 1990. (IMM), 1990, Pp 15-25 .*
- Curran, Patrick. "INCORE Metadatabase Server - Technical Aspects and Constraints." *IASSIST Newsletter. 19:3. 12-19 (1995).*
- Ellis, R Darin. And Others. "Gero-Informatics and the Internet: Loading Gerontology Information on the World Wide Web (WWW)." *Gerontologist. V36 N1 P100-05 Feb 1996 .*



- Elson, Lee, Mark Allen, and Jeff Goldsmith. "An Example of a Network-Based Approach to Data Access, Visualization, Interactive Analysis, and Distribution." *Bulletin of the American Meteorological Society* v 81 No3 Mar 2000.
- Emery, W, et al. "On-Line Access to Weather Satellite Imagery and Image Manipulation Software." *Bulletin, -American-Meteorological-Society* 1995.
- Ferguson, James C. "Appendix A: A Computer and Network Resource Guide to Support Data Sharing." *New Directions for Institutional Research. N89 P79-90 Spr 1996* .
- Franzen, H. "Facilitated Access to Information Sources: New Approaches in Organisation, Management and Delivery of Data and." *International Workshop – 1999 Oct : Alghero, Italy:* European Commission. European Commission, Directorate-General for Research; 2001.
- Gardner T, Ruggles S, and Sobek M. "IPUMS Data Extraction System." *Historical-Methods* (1999).
- Gartner H, et al. "OPALIS - An Open Paleocological Information System." *Geo-Information-Systeme* 2000.
- Goffe, William L, and Robert P Parks. "The Future Information Infrastructure in Economics." *Journal of Economic Perspectives* v 11 Summer 1997 p. (1975-1994).
- Grout, Catherine, Phill Purdy, Janine Rymer "Creating Digital Resources for the Visual Arts: Standards and GoodPractice.". Visual Arts Data Service (VADS)/ AHDS.  
[http://vads.ahds.ac.uk/guides/creating\\_guide/contents.html](http://vads.ahds.ac.uk/guides/creating_guide/contents.html)
- Harger, C., et al. "The Genome Sequence DataBase (GSDB): Improving Data Quality and Data Access." *Nucleic Acids Research. 26(1):21-6, 1998 Jan 1* .
- Hawick, K. A. Coddington PD. "Interfacing to Distributed Active Data Archives." *Future Generation Computer Systems* .
- Heikkila, Eric J. "GIS Is Dead; Long Live GIS." *Journal of the American Planning Association* v 64 No3 Summer 1998 p. (350-1960).
- Hinke, T. H., et al. "For Scientific Data Discovery: Why Can't the Archive Be More Like the Web?" *International Conference; 9th – 1997 Aug : Olympia; WA:* IEEE. IEEE; 1997.
- Ide, Nancy M., Sperberg M., and C McQueen. "The TEI: History, Goals, and Future." *Computers and the Humanities* v 29 No1 1995 (1995).
- Jacob, J. "Statistics and Data Protection: A Global View." *International Seminar – 1994 Nov : Luxembourg:* Eurostat; 1995.
- Kerschberg, L. "Knowledge Management in Heterogeneous Data Warehouse Environments." *International Conference; 3rd – 2001 Sep : Munich, Germany:* New York; Springer-Verlag; 2001.
- Kilbride, W. "Piloting an Archive for a New Millennium: Researching into, and Researching With Digital Data." *Conference – 1999 Nov : Stoke on Trent:* Society of Museum Archaeologists. SMA; 2001.
- Kim T. J. "Metadata for Geo-Spatial Data Sharing: A Comparative Analysis." *Annals-of-Regional-Science* 33(2) (1999).
- Kruse, Robin L., Bernard G. Ewigman, and George C. Tremblay. "The Zipper: a Method for Using Personal Identifiers to Link Data While Preserving Confidentiality." *Child Abuse & Neglect* (

- Sep2001).
- Kumar, R. "Ensuring Data Security in Interrelated Tabular Data." *Symposium – 1994 May : Oakland; CA*: IEEE Computer Society Press; 1994.
- Letourneau F, Bedard Y, and Moulin B. "Prospects for the Use of the Data Warehouse Concept for Georeferenced Digital Libraries on the Internet." *Geomatica* 52 (2) (1998).
- Lins, J. E. "On-Line Access and Analysis of Data: Developing A Social Science Information Management Strategy: Floor Discussion." *1996 Mar : Arlington; VA*: United States; Department of Commerce; Bureau of the Census. Department of Commerce; 1996.
- Ludacher, Bertram, Richard Marciano, and Reagan Moore. "Special Section on Advanced XML Data Processing - Preservation of Digital Data With Self-Validating, Self-Instantiating Knowledge-Based Archives." *SIGMOD Record*: 2001.
- McCain, KW. "Sharing Digitized Research-Related Information on the World Wide Web." *Journal of the American Society for Information Science*.
- Miller, P. "The Importance of Metadata to Archaeology: One View From Within the Archaeology Data Service: BAR International Series." *BAR International Series (Supplementary)*.
- Moore, Fred. "Long Term Data Preservation." *Computer Technology Review* Third Quarter (1999).
- . "Profiling the Storage Hierarchy." *Computer Technology Review Los Angeles* 20 (2000).
- Morrison, A. "Delivering Electronic Texts Over the Web: the Current and Planned Practices of the Oxford Text Archive." *Computers and the Humanities* v 33 no 1/2 (1999).
- Morrison, Alan, Michael Popham and Karen Wikander "Creating and Documenting Electronic Texts: A Guide to Good Practice." OTA/AHDS. <http://ota.ahds.ac.uk/documents/creating/>
- Morrison, Christopher. "Coalition Invents Handshakes for Biomed R&D Data Sharing." *IEEE Spectrum New York* 38 (2001).
- Olson, R. J., et al. "Managing Data From Multiple Disciplines, Scales, and Sites to Support Synthesis and Modeling." *Remote-Sensing-of-Environment*. 1999; 70(1): 99-107 .
- Oswald, R. "Shared Vision: Manitoba Partners Build Land Information Data Warehouse." *Geo-Info-Systems* 1996.
- Pirenne, B., P. Benvenuti, and R. Albrecht. "Implementing a New Data Archive Paradigm to Face HST Increased Data Flow." *Annual Conference; 6th – 1996 Sep : Charlottesville; VA*: Astronomical Society of the Pacific; 1997.
- Roddick, J. F. "Data Warehousing and Data Mining - Are We Working on the Right Things?" *Workshop – 1998 Nov : Singapore*: Berlin; Springer; 1999.
- Seligman, Len, and Arnon Rosenthal. "RESEARCH - XML's Impact on Databases and Data Sharing." *Computer*: 2001.
- Shaya, E., et al. "XML at the ADC: Steps to a Next Generation Data Archive." *Alfred Benzon Symposium – 1998 Aug : Copenhagen*: Alfred Benzon Foundation. Munksgaard; 1999.

Sprehe, J Timothy. "The U.S. Census Bureau's Data Access and Dissemination System (DADS)."  
Government Information Quarterly. v14 n1 p91-100 1997.

Vries, Repke Eduard de. "Oracle Eases Data Sharing." *Informationweek*. No. 610, (1996).

SOCIAL SCIENCES AND HUMANITIES RESEARCH COUNCIL OF CANADA

## We build understanding

The Social Sciences and Humanities Research Council of Canada (SSHRC) is an arm's length federal agency that promotes and supports university-based research and training in the social sciences and humanities. Created by an act of Parliament in 1977, SSHRC is governed by a 22-member Council that reports to Parliament through the Minister of Industry.

SSHRC-funded research fuels innovative thinking about real life issues, including the economy, education, health care, the environment, immigration, globalization, language, ethics, peace, security, human rights, law, poverty, mass communication, politics, literature, addiction, pop culture, sexuality, religion, Aboriginal rights, the past, our future.



Social Sciences and Humanities  
Research Council of Canada

Conseil de recherches en  
sciences humaines du Canada

Canada

350 Albert Street  
P.O. Box 1610  
Ottawa, ON K1P 6G4  
Canada

Phone: (613) 992-0691  
Fax: (613) 992-1787  
Internet: [www.sshrc.ca](http://www.sshrc.ca)