

Tirer profit des données massives : vers un nouveau cadre stratégique pour l'évolution de la recherche numérique au Canada

Document de consultation (16 octobre 2013)

CRSH  SSHRC

Conseil de recherches en sciences humaines (CRSH)



Instituts de recherche en santé du Canada (IRSC)



Conseil de recherches en sciences naturelles et en génie (CRSNG)



Fondation canadienne pour l'innovation (FCI)

Rétroaction : Envoyez vos commentaires à l'adresse digitalscholarship@sshrc-crsh.gc.ca.

Table des matières

Résumé	1
Introduction	3
Évolution de la recherche numérique	3
L'infrastructure numérique au Canada	7
Vers l'élaboration d'une politique et l'établissement de conditions générales efficaces	8
Annexe 1 : éléments du contexte de l'infrastructure numérique.....	11
Annexe 2 : carte du domaine et des intervenants.....	12
Annexe 3 : bibliographie	21
Annexe 4 : définitions	28

Tirer profit des données massives : vers un nouveau cadre stratégique pour l'évolution de la recherche numérique au Canada

Document de consultation
(16 octobre 2013)

Résumé

Grâce à divers investissements et projets de collaboration auxquels participent des établissements d'enseignement postsecondaire, des organismes sans but lucratif et à peu près tous les ordres de gouvernement, le Canada a mis en œuvre un grand nombre des éléments permettant une infrastructure numérique efficace en matière de recherche et d'innovation. Cela a offert une plateforme de base à une grande partie de la communauté des chercheurs. De plus, il semble évident que le potentiel de la recherche axée sur une grande quantité de données surpasse progressivement et rapidement la capacité de gérer et de faire progresser l'environnement numérique permettant de répondre aux besoins du 21^e siècle.

En collaboration avec Génome Canada, les organismes subventionnaires du Canada, à savoir le Conseil de recherches en sciences humaines (CRSH), le Conseil de recherches en sciences naturelles et en génie (CRSNG), les Instituts de recherche en santé du Canada (IRSC) et la Fondation canadienne pour l'innovation (FCI) - désignés sous l'appellation collective « groupe 3+ » -, se sont réunis afin de relever ce défi. Dans le document de consultation ci-joint, ils proposent une nouvelle harmonisation collective de leurs politiques de financement en ce qui concerne la gestion de données produites dans le cadre de projets menés grâce au financement des organismes subventionnaires. Plus particulièrement, ils proposent les initiatives suivantes :

1. **établir une culture de gestion** : l'avènement du « déluge de données » menace de dépasser l'évolution d'une culture de gestion de données au sein de la communauté des chercheurs. Pour y remédier, les organismes subventionnaires, les établissements de recherche et les associations scientifiques professionnelles - en ce qui concerne les données existantes et les pratiques exemplaires - devraient collaborer à l'élaboration de politiques et de lignes directrices claires afin de faire ressortir cet important aspect de la culture de recherche ainsi que de promouvoir la création de capacités et de systèmes de gestion de données efficaces. À cet effet, après avoir consulté la communauté des chercheurs et d'autres intervenants, on propose que le groupe 3+ établisse les fondements d'un plan interorganismes ciblé quant à la gestion de données;
2. **coordonner l'engagement des intervenants** : le milieu de la recherche canadienne se caractérise par un grand engagement envers la collaboration. En effet, des personnes et des organisations travaillent ensemble afin d'assurer une planification à long terme et de prendre des mesures ascendantes visant à orienter l'évolution de l'environnement numérique. Pour appuyer ces importantes mesures, qui offrent la richesse de l'engagement de la communauté, on pourrait élaborer un mécanisme de coordination permettant d'assurer une plus grande harmonisation, ce qui contribuerait à rentabiliser le temps investi par tous les intervenants. Afin de favoriser une contribution maximale d'importants intervenants, le groupe 3+ pourrait collaborer avec d'autres organisations et des groupes de travail afin d'assurer une consultation et une coordination de tous les intervenants - notamment les provinces - quant à la création d'une infrastructure numérique nationale en matière de recherche;

Tirer profit des données massives : vers un nouveau cadre stratégique pour l'évolution de la recherche numérique au Canada

Document de consultation
(16 octobre 2013)

3. **créer une capacité et de futurs paramètres de financement** : il faudrait réexaminer les paramètres du financement d'une infrastructure numérique nationale coordonnée afin de créer un milieu de recherche numérique étant orienté vers l'avenir. Plus particulièrement, il faudrait revoir l'équilibre entre les rôles et les responsabilités des intervenants nationaux, provinciaux et institutionnels afin d'en assurer l'efficacité et la rentabilité. Ainsi, le groupe 3+ propose une vaste collaboration quant à l'élaboration d'un plan coordonné visant à favoriser la création et le fonctionnement continu de centres de calibre mondial se spécialisant dans la gestion de données et exerçant des fonctions interreliées, notamment les suivantes : l'administration, les activités, les politiques et l'accès; des réseaux et une infrastructure plus efficaces; le développement de compétences; la formation de diplômés et de chercheurs.

Ces mesures comportent trois objectifs : favoriser la collecte et surtout le partage de données entre chercheurs conformément à l'initiative « Gouvernement ouvert » (www.data.gc.ca) du gouvernement du Canada; favoriser la création d'un milieu de recherche numérique canadien qui permet une plus efficace coordination des intervenants; contribuer à s'assurer de maximiser les futurs investissements liés à l'infrastructure numérique et à la formation au profit des Canadiens.

On fait actuellement parvenir le document aux intervenants et à la communauté des chercheurs dans le cadre d'un vaste processus de consultation visant à appuyer la future création du cadre stratégique proposé et les étapes connexes qui suivront.

Introduction

Le présent document a été préparé conjointement par les trois organismes subventionnaires fédéraux (le Conseil de recherches en sciences humaines, le Conseil de recherches en sciences naturelles et en génie et les Instituts de recherche en santé du Canada) ainsi que la Fondation canadienne pour l'innovation (FCI), collectivement appelés le « groupe 3+ », en collaboration avec Génome Canada. Il propose des changements à leurs cadres stratégiques de financement, qui favorisent d'excellentes pratiques de gestion des données, assurant ainsi l'évolution de la recherche numérique et de l'infrastructure numérique dans son ensemble au profit des Canadiens. Les propositions spécifiques qu'il contient (p. 11), sont motivées par des changements importants associés à l'évolution rapide de la recherche numérique dans le monde et à l'impact qui en résulte sur la recherche au Canada. Dans le cadre d'un processus de consultation national visant à en appuyer le contenu définitif, on les présente ici à la communauté des chercheurs et aux intervenants canadiens du milieu de la recherche numérique à des fins de discussion et de rétroaction.

Évolution de la recherche numérique

Le Canada se trouve actuellement au cœur d'une révolution des communications et de l'information dont les incidences sociales, économiques et culturelles se traduiront par de profondes transformations conceptuelles qui ont été grandement favorisées, soutenues et influencées par de nouvelles technologies dynamiques, comme les technologies de l'information et des communications (TIC). Cette révolution reflète l'évolution des technologies elles-mêmes, ainsi que les prodigieuses quantités d'informations - ou données - que ces outils peuvent maintenant enregistrer et traiter, et ce, qu'il s'agisse du comportement des consommateurs, de l'observation astronomique ou de la santé humaine. De fait, la forte intensification de la circulation de données venant de diverses sources constitue elle-même une source appréciable d'innovation et de « perturbations ».

« Nous en faisons tous l'expérience : une marée montante d'information, qui balaie les professions, les familles, le monde entier. Nous la créons, la diffusons, l'entreposons, la recevons, la consommons - puis, souvent, nous la traitons de nouveau et le cycle recommence. Cela nous donne un pouvoir - sans précédent dans l'histoire de l'humanité - de comprendre et de contrôler le monde. Toutefois, cela bouscule aussi nos institutions, bouleverse nos habitudes de travail et impose du stress imprévisible sur nos vies et nos sociétés. » [Traduction libre.]

Commission européenne, 2010¹

À mesure que cette transformation s'accélère et s'amplifie, l'analyse des données se concentre rapidement non plus sur le simple progrès technologique, mais aussi sur de nouvelles perceptions de l'expression et des comportements sociaux, économiques et culturels. Les nouvelles TIC favorisent la transformation du tissu social à mesure que les données deviennent la nouvelle monnaie d'échange dans les domaines de la recherche, de l'enseignement, de l'administration publique et du commerce. Les données deviennent rapidement un torrent qui envahit toutes les sphères de l'activité économique,

¹ Commission européenne. *Riding the wave. How Europe can gain from the rising tide of scientific data*, rapport final du groupe d'experts à haut niveau sur les données scientifiques, Commission européenne, octobre 2010.

sociale et culturelle mondiale². Selon des études récentes, alimentés par d'énormes quantités de données, on verra peut-être déferler une immense vague d'exploration, d'innovation, de productivité et de croissance à mesure que les particuliers, les établissements postsecondaires, les entreprises, les gouvernements et les organisations de tous ordres commenceront à en exploiter tout le potentiel. En outre, l'aptitude à saisir, à analyser et à diffuser de grandes quantités de données débouchera sur de nouvelles méthodes et pratiques de recherche - notamment la bio-informatique et la biologie computationnelle - pour créer ce que Tony Hey, de Microsoft Research, a appelé le quatrième paradigme de l'exploration scientifique³. Ce nouveau paradigme redéfinit également les relations entre les catégories traditionnelles d'enseignement postsecondaire, de recherche et de service. De fait, les données servent maintenant tout à la fois à la recherche, à l'apprentissage et à l'innovation. De plus, cette transformation entraîne de profondes préoccupations, notamment quant à la formation, à la capacité d'analyse, à l'accessibilité, à l'éthique, à la propriété intellectuelle et à la vie privée.

À l'instar des autres volets de la révolution de l'information et des communications, la mesure dans laquelle le progrès technologique et les données sont mises à profit pour favoriser l'économie et améliorer la société canadienne est intimement liée à un processus décisionnel humain. La façon dont le Canada gère ses données influera directement sur l'aptitude des Canadiens à lancer de futurs travaux de recherche-développement de pointe et à en tirer parti. Les données et l'information n'ont que faire des frontières - autant celles des établissements d'enseignement que les autres types de frontières. Pour gérer les données, il faut bien plus qu'appuyer l'excellence de la recherche. Les données numériques sont la matière brute de l'économie du savoir et deviennent de plus en plus importantes pour tous les secteurs de la société, y compris l'industrie. Dans son rapport de 2011 sur les données massives, le McKinsey Global Institute rappelle qu'à l'instar d'autres facteurs de production essentiels, comme les biens matériels et le capital humain, on constate de plus en plus souvent qu'une grande partie de l'activité économique actuelle, de l'innovation et de la croissance n'existeraient tout simplement pas sans les données. On pourrait dire la même chose de la capacité de les enregistrer, de les gérer et de les conserver ou de la nécessaire formation d'un personnel qui puisse fonctionner efficacement dans ce contexte.

Le Canada est maintenant en concurrence directe avec de nombreux pays, entre autres les États-Unis, les pays de l'Union européenne, l'Australie et d'autres pays techniquement avancés, dans la course à l'élaboration d'une stratégie efficace pour maîtriser la vague numérique. Avec une approche bien coordonnée d'élaboration de politiques, de formation, de création d'outils d'infrastructure et analytiques ainsi que de gestion et de maintien en poste du personnel et d'accès aux données de recherche, les Canadiens pourraient tirer des avantages socioéconomiques accrus et à plus long terme des grands investissements publics dans la recherche. Cette approche constituerait un prolongement logique de l'Initiative du gouvernement du Canada pour un gouvernement ouvert (www.data.gc.ca), qui vise déjà à accroître l'accès aux données du gouvernement.

² McKinsey & Company. *Big Data: The next frontier for innovation, competition, and productivity*, mai 2011.

³ Tony Hey. *The Fourth Paradigm: Data Intensive Scientific Exploration and Discovery*, Microsoft.

« Même si seule une fraction des données est destinée à la publication dans des revues, l'usage potentiel des données de façon indépendante et déconnectée est immense. Toute donnée produite par un chercheur peut constituer le morceau manquant du casse-tête d'un autre chercheur - même si leurs travaux n'ont, à première vue, aucun lien. La capacité d'exploiter de telles données et de les appliquer de nouvelles façons et dans de nouvelles directions promet d'accélérer considérablement la recherche et l'innovation. » [Traduction libre.]

Rapport du Thomson Reuters Industry Forum, 2013⁴

Une importante question de politique publique se pose alors : comment utiliser des modes de découverte fondés sur le traitement de données massives et sur une grande capacité de calcul au sein des secteurs public et privé pour appuyer la création d'emplois, la productivité et l'essor économique, stimuler l'innovation et trouver de nouvelles solutions à des défis d'envergure mondiale, tels que les changements climatiques, la prestation de soins de santé, la réduction de la pauvreté, la sécurité énergétique ou les diagnostics médicaux et la prévention des maladies. On doit aussi trouver des façons de répondre à des questions d'ordre éthique, juridique et social complexes découlant de nouvelles formes de collecte, d'exploration et de partage des données. Cela est d'autant plus important que les téléphones intelligents, les réseaux sociaux et les bases de données massives sur les consommateurs et les bénéficiaires de services de santé permettent maintenant de suivre et d'analyser des comportements individuels⁵.

Bien que la société axée sur les données, ainsi appelée par Alex Pentland, du Media Lab du MIT, semble poser des défis de taille, le Canada occupe une place de choix pour tirer pleinement parti du potentiel social et économique du déluge de données et pour découvrir de nouvelles façons de protéger les données personnelles, la vie privée et l'intérêt public. Par exemple, selon certains rapports récents, la recherche menée dans le domaine des TIC constitue l'un des principaux secteurs d'activité des secteurs public et privé du pays du point de vue des résultats, et 44 p. 100 des brevets canadiens sont rattachés au domaine des TIC. Une infrastructure numérique évoluée durable et de classe mondiale est définitivement un facteur déterminant de la réussite économique de ce secteur et possiblement de bon nombre d'autres secteurs.

Le Canada possède actuellement tous les éléments nécessaires permettant d'appuyer la création d'une telle infrastructure, entre autres des ordinateurs de grande capacité, des réseaux très efficaces, des dépôts de données, des effectifs hautement qualifiés et tous les outils et services connexes. Toutefois, comme les organisations et les fournisseurs de services ne peuvent pas encore compter sur une vision cohérente et unificatrice, l'établissement d'une meilleure coordination et d'une plus grande cohérence ne pourra que favoriser la mise en place d'un système de soutien plus efficace et plus efficient pour la recherche et le développement technologiques axés sur les données. Le Canada pourra également réduire l'écart de productivité face à ses concurrents à l'échelle mondiale, accélérer le rythme de

⁴ Thomson Reuters. *Unlocking the Value of Research Data*, rapport, juillet 2013.

⁵ Steve Lohr. « The Promise and the Peril of the Data Driven Society ». *The New York Times*, février 2013.

l'innovation et stimuler l'essor économique et la création de richesse dans tous les secteurs de l'économie.

Besoins liés aux systèmes permettant d'établir une infrastructure numérique efficace

L'un des grands défis actuels d'une nouvelle infrastructure numérique touche à la fois l'infrastructure et la gestion de grandes quantités de données - leur encodage pour en faciliter la conservation et la consultation, leur circulation, leur stockage, ainsi que leur accès et leur transformation en connaissances. Souvent, on n'est plus limité par la disponibilité des données, mais bien par leur interprétation. De fait, lors de la conférence sur la cyberscience tenue à São Paulo en 2013, on a souligné que la transformation des données en connaissances est maintenant le principal défi de l'ère numérique. À cette fin, on a besoin d'une infrastructure numérique évoluée qui favorise l'accès direct, l'utilisation, la réutilisation et l'intégrité des données. Il faut aussi mettre l'accent sur l'établissement et l'utilisation des processus nécessaires pour recueillir, gérer, analyser, interpréter, communiquer et archiver des données volumineuses. L'infrastructure numérique doit intégrer des moyens permettant aux chercheurs de tous les secteurs de bien utiliser les outils technologiques, étant donné que l'infrastructure humaine est aussi importante que l'infrastructure technologique. Les divers éléments de cette dernière doivent aussi être cohérents, bien coordonnés et harmonisés.

Il existe de nombreuses façons de décrire une infrastructure numérique évoluée robuste. Les principes suivants (adaptés du rapport du Sommet de l'infrastructure numérique de 2012) constituent un bon point de départ :

- **Intégration** - Tous les éléments indispensables font partie d'un système national qui permet d'éviter les doublons inutiles, la fragmentation et le chevauchement des services, ainsi que la concurrence inutile.
- **Inclusivité** - Traiter avec tous les intervenants fédéraux, provinciaux, territoriaux, régionaux et institutionnels dans tous les secteurs et toutes les communautés de chercheurs.
- **Durabilité** - La conception et le soutien favorisent l'évolution, l'adaptabilité, l'agilité et la réactivité aux utilisateurs et aux parties prenantes, ainsi que l'évolution de la technologie.
- **Exhaustivité** - Gamme complète de services portant sur tous les éléments de l'infrastructure numérique évolutive (réseaux, installations de calcul, outils de travail et services, gestion des données, conditions générales et ressources humaines).
- **Accessibilité** - Optimiser l'accès, quel que soit l'endroit, le domaine ou le niveau d'expertise. Du point de vue des utilisateurs, le guichet unique est préférable à l'obligation de négocier des ententes avec de nombreux fournisseurs de services.
- **Appréciation** - Les parties prenantes appuient le système national. La priorité des services aux utilisateurs est déterminée par l'intérêt pour les parties prenantes et tient compte des occasions offertes par l'évolution de l'infrastructure.
- **Intendance** - Le système national fait l'objet d'une surveillance efficace, d'un engagement ferme des parties prenantes et d'une harmonisation efficace de ses composantes.
- **Agilité** - Réceptivité aux besoins des utilisateurs et concentration sur des services rapides et efficaces.
- **Éthique** - Appuyer et promouvoir l'utilisation éthique des données et permettre la liberté de recherche.

L'infrastructure numérique au Canada

Des initiatives de coopération et des investissements conjoints des gouvernements fédéral, provinciaux et territoriaux et des institutions - et souvent de l'industrie - ont permis au Canada de mettre en place de nombreux éléments d'une infrastructure numérique évoluée et efficace de recherche et d'innovation et de se doter ainsi de la plateforme de base d'un élément indispensable à la communauté des chercheurs. Les Canadiens peuvent effectivement être fiers du réseau à grande largeur de bande et à faible latence de classe mondiale établi au Canada (et géré par CANARIE et les réseaux optiques régionaux évolués [RORE]) ainsi que des installations et services de calcul à grand rendement offerts partout au Canada par Calcul Canada. De plus, les bailleurs de fonds, les institutions, les chercheurs et les bibliothécaires s'intéressent de plus en plus à la gestion des données et à la formation du personnel. À ce jour cependant, le potentiel de la recherche à forte utilisation de données dépasse la capacité du Canada de gérer la vaste infrastructure numérique nécessaire pour satisfaire aux besoins du 21^e siècle et d'en assurer l'essor.

À l'heure actuelle, le paysage numérique est habité par un grand nombre d'acteurs et de parties prenantes qui doivent tous participer au processus de développement de l'infrastructure. À ce propos, mentionnons :

Les bailleurs de fonds

Les trois organismes subventionnaires fédéraux jouent un rôle primordial dans la définition de la structure et le financement de l'infrastructure numérique et de la formation avancée du Canada. Le partage des responsabilités a évolué depuis la formation de la FCI, en 1997, et de Génome Canada, en 2000. Depuis 15 ans, la FCI en particulier est peu à peu devenue le principal bailleur de fonds des ressources en matière de calcul qui vont au-delà de ce qu'il est possible d'obtenir avec des subventions destinées au financement de projets. Elle a avant tout pris l'initiative de rationaliser les services de calcul de haute performance.

Les provinces (et potentiellement les territoires) sont également d'importants bailleurs de fonds, par l'entremise parfois de leurs propres organismes de financement et parfois de divers ministères, sous forme de soutien fonctionnel des établissements d'enseignement postsecondaire, des systèmes électroniques de soins de santé (notamment les hôpitaux) et des RORE. Les fonds de contrepartie provenant d'établissements d'enseignement postsecondaire et du secteur privé représentent une autre source de soutien en matière d'infrastructure.

Les fournisseurs de services

Les principaux fournisseurs de services canadiens sont CANARIE, les RORE et Calcul Canada. Les rôles de CANARIE dans la fourniture d'un réseau à large bande et à faible latence et de Calcul Canada dans le développement de ressources de calcul sont bien connus et compris. CANARIE assume depuis peu la gestion de la Fédération canadienne d'accès, un environnement d'accès de confiance (par code d'accès unique) pour les chercheurs et les établissements d'enseignement supérieur du Canada.

On trouve également des fournisseurs de services de contenu, entre autres i) le Réseau canadien de documentation pour la recherche; ii) le Conseil national de recherches du Canada, qui gère le service canadien d'enregistrement de données DataCite; iii) le Consortium pour l'avancement des standards d'administration de l'information en recherche, qui s'occupe des données sur l'administration de la recherche et qui devrait adopter des normes relatives aux métadonnées.

Mentionnons également les infrastructures nationales spécialisées dans la gestion des données de recherche dans des domaines précis, comme i) le Réseau canadien des Centres de données de recherche, qui fournit des services de conservation et un accès protégé aux données de Statistiques Canada et qui souhaite ajouter d'autres ensembles de données confidentielles; ii) le Canadian Polar Data Network, issu d'une activité florissante de gestion des données conçue pour appuyer l'Année polaire internationale; iii) le Centre canadien de données en astronomie, un dépôt de données de confiance lié à un réseau international de données en astronomie recueillies dans le cadre de projets auxquels participent des chercheurs canadiens.

Ajoutons enfin les infrastructures de gestion de données de recherche liées à des projets, comme Ocean Networks Canada et l'Étude longitudinale canadienne sur le vieillissement.

Les bénéficiaires et les collaborateurs

Les établissements de recherche universitaire et leurs communautés de chercheurs sont autant des collaborateurs que les principaux bénéficiaires d'une infrastructure numérique évoluée robuste. Les intérêts de ces établissements dans ce domaine sont défendus par :

- les recteurs et les vice-recteurs à la recherche et d'autres administrateurs supérieurs d'établissements postsecondaires et d'hôpitaux;
- des services de technologie de l'information, des dirigeants principaux de l'information et leur organisme national, le regroupement des dirigeants des technologies de l'information des universités canadiennes;
- les bibliothèques universitaires, les bibliothécaires en chef et leur organisme national, l'Association des bibliothèques de recherche du Canada.

Les systèmes d'innovation régionaux sont également des bénéficiaires par l'intermédiaire de l'utilisation de CANARIE par le secteur privé et des RORE d'envergure locale ou régionale, ainsi que par l'accès à des chercheurs et à des travaux de recherche soutenus par les trois organismes subventionnaires, la FCI et Génome Canada.

Les groupes de collaboration

Les préoccupations propres à certaines parties prenantes à l'égard de la durabilité et du futur développement de l'infrastructure numérique avancée du Canada ont donné lieu à la formation de deux groupes. Le premier, le Groupe de travail sur la stratégie des données de recherche, est devenu Données de recherche Canada en 2012, une tribune où les parties prenantes travaillent ensemble à l'amélioration de l'intendance des données de recherche. Le second, le Conseil de direction de l'infrastructure numérique, met l'accent sur l'élaboration d'une stratégie nationale visant à renouveler et à renforcer l'infrastructure numérique évoluée du Canada, en accordant la priorité aux questions de gestion des données, de financement durable et de planification intégrée.

Vers l'élaboration d'une politique et l'établissement de conditions générales efficaces

Compte tenu du très grand nombre de parties prenantes et de la complexité du système, il reste encore une grande marge pour amorcer une vaste réflexion stratégique nationale et prendre des mesures à l'égard d'enjeux vitaux pour l'avenir de l'infrastructure numérique évoluée du Canada. De fait, un nombre relativement restreint d'initiatives permettrait de favoriser un environnement permettant de mieux orienter la recherche de pointe qui est utile aux secteurs public et privé et d'assurer son développement continu.

Ainsi, dans ce document, le groupe 3+ ne propose pas simplement qu'un autre organisme ou une autre organisation surveille le vaste ensemble de ressources et de services qui font déjà partie du domaine de l'infrastructure numérique. Il préconise plutôt de procéder à une nouvelle harmonisation des politiques de financement des organismes subventionnaires en mettant l'accent sur une plus grande clarté et la coordination globale du rôle joué par les trois organismes subventionnaires fédéraux et la FCI dans le domaine de la recherche numérique dans tous les secteurs de recherche. À terme, cela devrait marquer une première étape vitale pour l'établissement, au Canada, d'un milieu de recherche numérique et de formation qui assurera plus de clarté et une coordination efficace des actions prises par les diverses organisations de parties prenantes, notamment les autres bailleurs de fonds, les fournisseurs de services et les chercheurs, afin que leurs objectifs et leurs efforts contribuent de façon optimale à l'atteinte de buts et d'objectifs précis touchant l'infrastructure numérique du Canada.

Dans la recherche de cet objectif et pour bien lancer la mise en place d'une infrastructure numérique robuste et durable au Canada, les mesures qui suivent, proposées par les trois organismes subventionnaires et la FCI visent à amorcer un débat chez les parties prenantes et l'ensemble des chercheurs, et à obtenir leurs rétroactions.

1. **Établir une culture de saine gestion** - Le début d'un « déluge de données » menace de prendre de court l'évolution d'une culture d'intendance des données au sein de la communauté de recherche. À cette fin, les organismes de financement de la recherche, les établissements de recherche et les associations scientifiques professionnelles, se référant aux données existantes et aux pratiques exemplaires dans le monde, devraient collaborer à l'élaboration de politiques claires et de lignes directrices visant explicitement à mettre en valeur ce volet important de la culture de recherche et à promouvoir la mise en place de systèmes et de moyens appropriés pour la gestion des données. Le CRSH propose donc que le groupe 3+ définisse les éléments de base d'un plan de gestion de données interorganismes qui est ciblé. Ces éléments pourraient comprendre ce qui suit :
 - a. une exigence relative à l'inclusion, dans toutes les demandes de subvention, de plans particuliers de gestion des données incluant les coûts de la collecte et de l'analyse des données et la conservation des résultats et des ensembles de données connexes;
 - b. une définition des éléments précis de ces plans de données dont les examinateurs tiendront compte pour évaluer les demandes de financement;
 - c. des lignes directrices précisant les données à conserver ainsi que le mode de conservation;
 - d. les politiques et les lignes directrices relatives au libre-accès aux données (conjointement au travail déjà entrepris le groupe 3+);
 - e. des lignes directrices permettant aux chercheurs de choisir des dépôts de données acceptables;
 - f. la reconnaissance des dépôts de données canadiens qui satisfont les normes mondiales pertinentes;
 - g. des lignes directrices permettant d'assurer un consentement informé quant à l'utilisation de données ainsi qu'à la protection de renseignements personnels et de données confidentielles.
2. **Coordonner la participation des parties prenantes** - Le milieu canadien de la recherche se caractérise par un degré d'engagement élevé en matière de collaboration et de coopération de particuliers et d'organisations qui unissent leurs forces dans des activités de planification à long terme et des actions émergent de la base afin de définir l'évolution du paysage

numérique. Ces actions importantes qui enrichissent l'engagement collectif pourraient tirer profit de l'établissement d'un mécanisme de coordination destiné à assurer une meilleure harmonisation afin de rentabiliser davantage le temps investi par toutes les parties. Afin d'optimiser la contribution des acteurs clés, le groupe 3+ collaborerait avec d'autres organisations et groupes de travail pour assurer des consultations permanentes et la coordination avec toutes les parties prenantes, y compris les gouvernements provinciaux, pour établir l'infrastructure numérique de recherche du Canada.

3. **Accroître les ressources et définir les paramètres de financement futurs** - Afin de favoriser l'établissement d'un environnement de recherche numérique tourné vers l'avenir, il conviendrait de réexaminer les paramètres du financement d'une infrastructure coordonnée à l'échelle nationale, et en particulier l'équilibre des rôles et des responsabilités entre les parties prenantes nationales, provinciales et institutionnelles, afin d'assurer un soutien efficace et de faire des gains d'efficacité. Le groupe 3+ contribuerait à élaborer un plan coordonné visant à favoriser la mise sur pied ou l'amélioration et le fonctionnement continu de divers centres de classe mondiale spécialisés dans la gestion de données et pouvant appuyer :
 - a. les pratiques exemplaires concernant l'administration, le fonctionnement, les politiques et l'accès;
 - b. des réseaux et une infrastructure évolués, ainsi que le soutien de recherches de pointe visant, par exemple, à :
 - i. déceler de nouveaux écarts et définir des solutions pour le développement de la technologie, des services et de l'infrastructure de recherche numérique du Canada;
 - ii. définir les meilleures pratiques à l'échelle mondiale en matière de gestion des données;
 - iii. harmoniser et prévoir l'intégration d'autres partenaires possibles dans les domaines de la recherche numérique et de la gestion de données, en misant notamment sur la participation du secteur privé;
 - iv. établir les meilleures méthodes de collaboration internationale en gestion des données;
 - c. le développement de compétences ainsi que la formation de diplômés et de chercheurs.

Ces mesures proposées par le groupe 3+ visent à accélérer la mise en œuvre d'une approche coordonnée et efficace grâce à laquelle une compréhension commune favorise des actions particulières de la part des divers acteurs du paysage numérique du Canada. Cela permettra de produire des résultats concrets qui sont précis, notamment l'accroissement de la quantité de données de recherche dont peuvent profiter les secteurs public et privé ainsi que le développement de compétences contribuant à une meilleure utilisation de cet environnement. Dans l'ensemble, nous espérons qu'une définition conjointe des principales caractéristiques d'une infrastructure numérique robuste et durable et une coordination explicite des nombreuses étapes nécessaires pour implanter cet écosystème permettront au Canada de vite devenir un chef de file mondial dans l'établissement des sociétés prospères, résilientes et justes de l'ère numérique.

Annexe 1 : éléments du contexte de l'infrastructure numérique

- **Les conditions générales** - Les politiques et le cadre juridique qui s'appliquent à la recherche numérique; les mécanismes de coordination et d'harmonisation des divers éléments de l'environnement de recherche numérique; la pertinence des systèmes de financement et de récompense de la recherche numérique; la capacité du Canada de traiter avec des acteurs internationaux dans le domaine de la recherche numérique.
- **Expertise et compétences (intrants)** - L'exhaustivité et la qualité du personnel qualifié dans des domaines généraux et spécialisés pour assurer l'utilisation efficace des éléments de l'infrastructure numérique.
- **Outils et services** - Les logiciels, les applications et les services de soutien personnel qui permettent aux chercheurs de valoriser leurs données et d'optimiser l'utilisation des instruments et des systèmes de recherche.
- **Les données de recherche et l'infrastructure de gestion des données de recherche** - Les données, du point de vue à la fois de l'infrastructure et des systèmes de gestion de données, c'est-à-dire la collecte, l'organisation, la normalisation, l'archivage, la conservation et la diffusion des données et les caractéristiques des systèmes, c'est-à-dire la souplesse, la sécurité, l'accessibilité, l'interopérabilité, le caractère abordable, le libreaccès et le rendement élevé.
- **Les installations, outils et services de calcul** - Le matériel et les ressources logicielles connexes qui facilitent la recherche à forte intensité de calcul et la recherche à forte intensité de données, ainsi que les services et les outils qui permettent de valoriser les installations, ce qui comprend l'infonuagique et le calcul dans des réseaux maillés.
- **Les réseaux et les outils et services** - Moyens permettant de relier les chercheurs à des sources de données et de transporter des données entre divers endroits.
- **L'infrastructure et les outils de collaboration** - Moyens de relier des chercheurs à d'autres chercheurs et à des partenaires géographiquement dispersés ou qui utilisent des ensembles de données et des outils de travail communs, dans le cadre de projets de recherche multisectoriels.

Annexe 2 : carte du domaine et des intervenants

Courte description d'organismes et de groupes de travail qui font la promotion de la recherche numérique, qui l'appuient et qui constituent la trame de l'infrastructure numérique canadienne.

Données de recherche Canada (DRC)

- **Nature** - Organisme national dirigé et soutenu par des parties prenantes qui vise à promouvoir la vision des données de recherche au Canada.
- **Qui?** - Diverses organisations de parties prenantes qui désirent toutes la mise en œuvre d'une infrastructure, de processus et de mécanismes de soutien axés sur la concrétisation de la vision des données de recherche au Canada, et même y participer, entre autres CUCCIO, Calcul Canada, CANARIE, l'ABRC, la FCI, les IRSC, le CRSNG, le CRSH, CASRAI, le CNRC, le CT, l'API, BAC et le CODATA.
- **Mandat** - Élaborer une stratégie, faciliter les communications et les partenariats entre diverses initiatives en matière de données, promouvoir l'éducation et la formation dans les compétences connexes, évaluer les progrès dans la mise en œuvre de la vision, attirer l'attention sur les lacunes et servir de point de convergence canadien d'initiatives internationales en matière de données (mandat défini à l'interne).
- **Fonctions** - Les activités se concentrent sur cinq domaines : les politiques, l'infrastructure, les normes et l'interopérabilité, l'éducation et la formation, ainsi que les relations internationales. DRC ne prévoit pas posséder ou exploiter une infrastructure.
- **Convergences/rerelations** - Légèrement différentes de celles du Conseil de direction de l'infrastructure numérique, car elles portent sur des questions relatives aux données de recherche et à la gestion du cycle de vie des données. Il semble y avoir un certain chevauchement, toutefois, entre les deux organisations en ce qui concerne les domaines d'intérêt et la participation.
- **Actions récentes** - La réunion inaugurale de DRC a eu lieu en janvier 2013. Les priorités de 2013 sont les suivantes :
 - **lancer Données de recherche Canada** à titre d'organisation formée de nombreuses parties prenantes bénévoles qui lancera de nouvelles actions visant à tirer pleinement partie de la valeur des données de recherche canadiennes;
 - **encourager une vaste participation** afin de bien refléter la diversité des parties prenantes qui s'intéressent aux données de recherche;
 - **faire progresser le travail des comités de DRC** : l'infrastructure, l'éducation et la formation, les politiques, les normes et l'interopérabilité ainsi que les relations internationales;
 - **tenir une série de séminaires Web** sur la gestion de données (orateurs canadiens et étrangers, divers sujets et publics);
 - **cogérer la circulation de données dans le cadre de la conférence du CASRAI sur les vastes ensembles de données** qui aura lieu en octobre 2013 et coordonner la tenue d'un atelier sur la conférence préliminaire des centres de données;
 - **poursuivre les relations internationales** établies avec la Research Data Alliance, la Global Research Data Infrastructure, la DataCite Federation et d'autres initiatives;
 - **établir un conseil consultatif national** réunissant des représentants principaux de l'industrie, des milieux universitaires, des laboratoires de recherche gouvernementaux, des organismes de financement et des décideurs, afin de conseiller Données de recherche Canada;

- lancer une consultation nationale en ligne afin de communiquer les résultats du Sommet de 2011 sur les données de recherche canadiennes à un plus grand nombre de parties prenantes partout au pays.
- **Historique :**
 - DRC a pris le relais du Groupe de travail sur la stratégie des données de recherche (GTSDR) en 2012, lui-même formé afin de déceler et de définir les défis et les enjeux relatifs à l'accès aux données issues de recherches canadiennes et à leur conservation.
 - Le GTSDR a organisé le [Sommet de septembre 2011 sur les données de recherche canadiennes : Portrait de la situation des données](#).
 - Le rapport final du Sommet proposait d'établir une stratégie nationale relative aux données de recherche canadiennes comprenant entre autres une vision, des objectifs généraux et un cadre d'action assorti d'échéances souples, ainsi que la répartition des tâches entre les principaux groupes de parties prenantes.
 - Une analyse des écarts au Canada a également été faite : http://publications.gc.ca/collections/collection_2009/cnrc-nrc/NR16-123-2008F.pdf
- Le GTSDR a cédé ses activités à Données de recherche Canada en décembre 2012.

Conseil de direction de l'infrastructure numérique

- **Nature** - Initiative nationale dirigée et soutenue par des parties prenantes pour élaborer une stratégie nationale visant à renouveler et à renforcer l'infrastructure numérique évoluée actuelle destinée à la recherche, à l'innovation et à l'éducation au Canada.
- **Qui?** - Groupe intersectoriel coprésidé par Jay Black, de l'Université Simon-Fraser et membre de CUCCIO, et par Steven Liss, vice-principal à la recherche, Université Queen's, et qui regroupe le RCDR, CUCCIO, Calcul Canada, CANARIE, l'ABRC, la FCI, le CNRC, Industrie Canada, les ICRS, le CRSNG, le CRSH et CASRAI.
- **Mandat** - Tirer parti des travaux accomplis lors du Sommet, élaborer les plans des premières initiatives et établir le mécanisme qui assurera la participation continue de nombreuses parties prenantes.
- **Fonctions et priorités** - Faire une analyse des écarts, établir une feuille de route pour combler les écarts et organiser un suivi du Sommet de 2012.
- **Convergences/reliations** - Le Conseil se distingue de DRC par le fait qu'il vise avant tout à offrir une perspective générale ainsi que les mécanismes permettant d'appuyer une approche intégrée et durable dans le cadre de l'infrastructure numérique de pointe du Canada.
- **Actions récentes** - Découle du Sommet de 2012 sur l'infrastructure numérique organisé par CUCCIO à Saskatoon.

Association des bibliothèques de recherche du Canada (ABRC)

- **Nature** - Organisme directeur des bibliothèques de recherche canadiennes.
- **Qui?** - Les 31 principales bibliothèques de recherche du Canada.
- **Mandat** - Faire preuve de leadership au nom des bibliothèques de recherche canadiennes et accroître leur capacité de faire progresser la recherche et l'enseignement supérieur; viser l'efficacité et la durabilité des communications de recherche et des politiques publiques, ce qui permet un grand accès aux renseignements liés à la recherche.
- **Fonctions et priorités** - L'orientation stratégique de 2013-2016 comprend les éléments suivants :
 - favoriser les initiatives de collaboration visant à partager et à préserver les collections de recherche canadiennes;

- coordonner des initiatives de gestion de données de recherche;
- promouvoir le libre accès et de nouvelles formes de communication de recherche.
- **Actions récentes dans le domaine de la recherche numérique**
 - L'ABRC a dirigé l'élaboration d'un concept de réseau distribué et dépôts de données, qui visait à appuyer une proposition de financement de la FCI présentée en 2012. En fin de compte, aucune proposition n'a été à la FCI, car le Fonds de l'avant-garde et des initiatives nouvelles ne constituait pas un programme de la FCI qui était pertinent à cet effet.
 - Avec le RCDR, en octobre 2012, l'ABRC a publié un rapport sur les rôles pouvant être joués par les bibliothèques universitaires quant à la mise en œuvre d'une politique de libre accès aux publications de recherche.
 - En janvier 2013, l'ABRC a offert un cours de quatre jours intitulé « Introduction à la gestion des données de recherche » à des bibliothécaires, et environ 60 personnes provenant de 30 universités y ont participé. Ces participants ont poursuivi la discussion sous la forme d'une communauté de pratique en ligne.
 - L'ABRC has worked to facilitate the development of open access digital repositories at (by now) all member libraries; Most CARL member libraries are developing local research data management services for researchers.
 - L'ABRC has advocated for federal government support for national research data management infrastructure; it has developed web content, conference presentations, workshops, and articles on both open access and research data management.
 - L'ABRC is working with CASRAI and Research Data Canada to produce CASRAI's 2013 "Big Data" conference.
 - L'ABRC has created a Data Management Subcommittee headed by its Vice-President/President-Elect (Martha Whitehead, Queen's) that is proposing the formation of a national collaborative network of local/regional/other initiatives for collecting, preserving and providing access to research data produced in Canada.

Intersects/Interfaces - Both RCDR and Canadiana.org (Canada's premier organization for the digitization of Canadian historical documentation and the exposure of Canadian digital documentary collections) had their origins as CARL initiatives; CARL collaborates closely with them. L'ABRC is a supporting member of Research Data Canada and the Leadership Council for Digital Infrastructure. CARL is the major Canadian association member of SPARC (the Scholarly Publishing and Academic Resources Coalition) and COAR (the Coalition of Open Access Repositories), both of which promote and support open access and repositories internationally. Locally, CARL member library directors work in consultation with CUCCIO member CIOs in the context of various digital services.

Conseil des bibliothèques universitaires de l'Ontario (CBUO)

- **Nature** - Consortium de bibliothèques.
- **Qui?** - Les 21 bibliothèques universitaires de l'Ontario.
- **Mandat** - Améliorer les services d'information en Ontario et ailleurs par des achats collectifs et le partage de l'infrastructure d'information numérique, une planification concertée, des activités de promotion, des évaluations, des recherches, des partenariats, les communications et le perfectionnement professionnel.
- **Fonctions** - Assurer l'accès à diverses ressources d'apprentissage et de recherche et veiller à leur conservation dans le cadre de pratiques de gérance durables et responsables; diriger la formation de partenariats afin d'appuyer l'expansion de l'infrastructure de recherche numérique canadienne; exploiter le Scholars Portal; encourager l'amélioration de l'accès aux

données numériques, entre autres à celles qui résultent de l'Initiative de démocratisation des données; accroître l'accès aux cartes, aux données géospatiales et à d'autres ressources cartographiques sous forme numérique et imprimée.

- **Activités récentes** - Le Scholars Portal est le premier dépôt numérique fiable certifié au Canada. Cette certification, qui est la seule certification générale reconnue pour des archives numériques, a été décernée par le [Center for Research Libraries](#).

Regroupement des dirigeants des technologies de l'information de l'enseignement supérieur au Canada (CUCCIO)

- **Nature** - CUCCIO est une association sans but lucratif financée par ses membres. Des dirigeants de la technologie de l'information (TI) du secteur de l'enseignement supérieur travaillent ensemble pour aider les universités canadiennes à exceller dans l'utilisation novatrice et efficace de la TI.
- **Qui?** - Dirigeants principaux de l'information de plus de 50 universités canadiennes.
- **Priorités stratégiques et mandat** (selon le plan stratégique)
 - Favoriser l'application de pratiques exemplaires dans la gestion des technologies de l'information par les universités canadiennes.
 - Définir, développer et commanditer des services sectoriels axés sur la collaboration.
 - Concevoir et mettre en œuvre des programmes et des services destinés à appuyer le perfectionnement professionnel du personnel des services de TI.
 - Établir et cultiver des relations avec les gouvernements, des organismes publics, des entreprises et d'autres groupes d'intérêt du domaine de l'enseignement supérieur afin de faire valoir les intérêts communs des universités canadiennes.
- **Activités récentes en matière de recherche numérique**
 - Organisation du *Sommet de 2012 sur l'infrastructure numérique*, pour :
 - élaborer une vision d'une infrastructure numérique exhaustive, intégrée et durable destinée à appuyer la recherche, l'enseignement et l'innovation au Canada;
 - élaborer un plan d'action particulier assorti d'étapes précises;
 - obtenir des engagements de la part des parties prenantes à l'égard de la mise en œuvre de ce plan d'action.
 - Le sommet a conduit à la formation d'un groupe intersectoriel, le *Conseil de direction de l'infrastructure numérique*, chargé de poursuivre les travaux lancés lors du Sommet (voir ci-dessus).
 - On prévoit actuellement donner suite au sommet de 2012.
- **Convergences/rerelations** - Chevauchements partiels avec les domaines d'intérêt de Calcul Canada, de CANARIE et de Données de recherche Canada concernant le tableau d'ensemble de l'infrastructure numérique.

Calcul Canada

- **Nature** - Organisme sans but lucratif constitué en société qui fournit la plateforme nationale canadienne de ressources de calcul de haute performance (CHP). Un président et un conseil d'administration ont été nommés à la fin de 2012.
- **Qui?** - L'organisme est formé de 29 universités canadiennes. Chaque université ou collègue du Canada dont au moins un chercheur utilise un système de calcul évolué peut adhérer à Calcul Canada et avoir accès à ses ressources. Calcul Canada offre des services à l'échelle du Canada par l'intermédiaire de nœuds régionaux :
 - Compute West – WestGRid

- Compute Ontario - HPCL, SciNet, SHARCNet
- Calcul Québec - auparavant le Réseau québécois de calcul de haute performance et le Consortium Laval, Université du Québec à Montréal, McGill et l'est du Québec
- Compute Atlantic - ACENet
- **Fonctions** - Calcul Canada offre ses services au moyen de systèmes de CHP gérés par des consortiums établis dans diverses régions du Canada et par l'intermédiaire du réseau à large bande CANARIE.
- **Mandat** - Promouvoir et appuyer l'utilisation partagée de ressources de calcul évoluées afin d'assurer la compétitivité de la recherche et de l'innovation au Canada. Calcul Canada n'est pas un organisme de normalisation et ne définit pas de politiques.
- **Actifs** - La plateforme de Calcul Canada comprend des installations de calcul, de stockage en ligne et à long terme, des connexions avec le réseau CANARIE, ainsi que des services de soutien des utilisateurs. Elle s'adresse principalement aux grands systèmes de calcul qui servent à faire des simulations et des recherches exigeant une grande capacité de calcul.
- **Financement** - La FCI finance une partie des installations. Diverses institutions participent aussi au financement. Calcul Canada reçoit aussi une aide importante pour son exploitation et la maintenance de la part du Fonds des initiatives scientifiques majeurs de la FCI.
- **Actions récentes et en cours**
 - Proposition récente infructueuse à la FCI (concours du Fonds de l'avant-garde et des initiatives nouvelles) pour combler des besoins en médecine et en sciences humaines relatifs à :
 - des systèmes de « données volumineuses », c'est-à-dire des systèmes ayant des rapports élevés en matière de besoin de stockage et de traitement de données (qui dépassent largement l'offre actuelle de Calcul Canada) et un débit élevé;
 - des environnements de traitement de données protégés (par exemple pour traiter des données confidentielles).
 - Plus axé sur la capacité que sur l'orientation de la recherche (ce qui est jugé acceptable).
 - Élaboration du plan stratégique organisationnel de 2013.
- **Convergences/reliations** - Provinces, CANARIE, DRC, Conseil de direction de l'infrastructure numérique, secteur privé.

CANARIE

- **Nature** - L'un des réseaux les plus grands et les plus rapides au monde consacré à la recherche de pointe, à l'enseignement et à l'innovation.
- **Qui?** - Société sans but lucratif financée par des utilisateurs du gouvernement fédéral et du secteur privé.
- **Mandat** - Concevoir, mettre en œuvre et promouvoir l'infrastructure numérique destinée aux chercheurs et aux enseignants canadiens.
- **Fonctions** - Gérer un réseau ultrarapide qui relie des universités, des centres de recherche, des laboratoires gouvernementaux, des hôpitaux et d'autres établissements scientifiques partout au Canada. Utilisé de concert avec les RORE, ce réseau permet aux chercheurs de communiquer, de partager et d'analyser des données.
- **Objectifs** - Pour la période du 1^{er} avril 2012 au 31 mars 2015 :
 - **Exploitation du réseau** - Exploiter le réseau CANARIE à titre d'infrastructure optique de recherche indispensable, c'est-à-dire agrandir les sections du réseau qui

- « appartiennent » à l'organisme pour offrir plus de souplesse et réduire les coûts de prestation d'une largeur de bande supérieure, notamment les services réseau offerts à :
- [la Fédération canadienne d'accès](#),
 - [le Service de diffusion du contenu](#), et
 - poursuivre la mise en œuvre du Programme patrimonial d'extension des infrastructures (auparavant le Programme d'extension des infrastructures) et les Programmes Alliance réseaux.
- **Innovation technologique** - Concevoir, présenter et mettre en œuvre les technologies de nouvelle génération qui feront du réseau CANARIE un réseau de recherche de pointe, à l'aide de nouveaux outils logiciels comprenant une trousse de services réutilisables. Deux programmes appuient cet objectif :
- Les interfaces de plateformes de recherche (IPR) - Tirer parti des services de l'ancien Programme de plateformes sur réseau (PPR) par la création d'un ensemble de services de plateformes (IPR) fondés sur les plateformes de réseau existantes que pourront utiliser plusieurs plateformes de recherche.
 - Le Programme de plateformes sur réseau (PPR) - Conception de plateformes logicielles évoluées qui facilitent la collaboration entre les chercheurs et l'accès aux outils et aux données de recherche.
- **Innovation dans le secteur privé** - Tirer parti du réseau CANARIE pour aider des entreprises établies au Canada et les universités canadiennes à innover et à commercialiser des produits et des services qui stimuleront l'innovation technologique au Canada.

Réseaux optiques régionaux évolués

- **Nature** - Réseaux optiques régionaux évolués (RORE)
- **Qui?** - Douze réseaux régionaux :
 - Est - ACORN-NL, ACORN-NS, New Brunswick Advanced Network, Prince Edward Island Advanced Network
 - Québec - Réseau d'informations scientifiques du Québec (RISQ)
 - Ontario - ORION
 - Ouest - BCNet, Cybera, SRnet, MRnet
 - Territoires - Aurora College, Yukon College
- **Mandat** - Appuyer l'exploitation et l'expansion de réseaux et de services régionaux évolués destinés à soutenir la recherche et l'innovation.
- **Convergences/reliations** - À l'échelle nationale. Les RORE servent à rapprocher les chercheurs des ressources opérationnelles de réseaux à grande vitesse situés à proximité. Pour sa part, CANARIE réunit les réseaux de communications régionaux au sein d'un service de recherche national.

Fédération canadienne d'accès

- **Nature** - Environnement d'accès de confiance (par code d'accès unique) pour les chercheurs et les établissements d'enseignement supérieur du Canada.
- **Qui?** - Créé initialement par CUCCIO et transféré à CANARIE en 2012.
- **Mandat** - Faire en sorte que le partage de ressources protégées soit plus facile, plus sûr et plus évolutif à l'ère des ressources et des services numériques, c'est-à-dire :

- permettre aux employés, aux étudiants et aux enseignants qui visitent d'autres organisations d'avoir accès à des réseaux sans fil et à des ressources Internet à partir de l'accréditation donnée par leur organisation d'origine;
- offrir aux participants une méthode d'accès économique et sécurisé aux outils de gestion;
- favoriser la protection des renseignements personnels en évitant aux chercheurs, aux étudiants et aux enseignants l'obligation d'avoir plusieurs comptes protégés par des mots de passe;
- aider les organisations à mieux gérer l'accès à leurs ressources, en fonction du statut et des privilèges reconnus par l'organisation d'origine de l'utilisateur.

Réseau canadien de documentation pour la recherche (RCDR)

- **Nature** - Partenariat d'universités canadiennes axé sur l'expansion des contenus numériques destinés aux projets de recherche universitaire au Canada.
- **Qui?** - Organisation sans but lucratif constituée en société qui réunit 75 universités canadiennes. Les bibliothèques universitaires sont au cœur des initiatives du RCDR. Elles jouent un rôle primordial en mettant leur savoir-faire et leurs ressources à la disposition des chercheurs universitaires canadiens.
- **Mandat** - Entreprendre de grands projets d'acquisition de contenus et d'utilisation de contenus sous licence afin de doter les universités canadiennes d'une infrastructure du savoir et d'une capacité de recherche et d'offrir à toutes les universités canadiennes un accès équitable et économique à des données de recherche.
- **Activités récentes** - Avec l'ABRC, publication, en octobre 2012, du rapport *Mise en œuvre du libre accès : Rapport du groupe de travail ABRC-RCDR sur le libre accès*.
- **Convergences/rerelations** - Participation de nombreux bibliothécaires membres de l'ABRC.

Conseil national de recherches du Canada - Institut canadien de l'information scientifique et technique (CNRC - ICIST)

- **Nature** - Bibliothèque scientifique nationale du Canada.
- **Qui?** - L'ICIST est une division du CNRC.
- **Mandat** - L'ICIST a entrepris sa transformation d'éditeur et de bibliothèque de publications scientifiques en un fournisseur de services de bibliothèque électroniques.
- **Activités récentes**
 - L'ICIST est en train d'établir un site miroir de PubMed Central des National Institutes of Health des États-Unis. Financé par les IRSC, ce site offrira aux chercheurs canadiens l'accès aux publications américaines enregistrées dans PubMed Central et servira de portail pour le dépôt d'articles de revues canadiennes. Le portail permettra plus tard aux IRSC d'établir une politique de libreaccès entièrement obligatoire.
- **Convergences/rerelations** - L'ICIST est également l'hôte et l'administrateur de divers groupes de travail qui s'intéressent aux données et au libreaccès, entre autres le nouvel organisme Données de recherche Canada. Ses activités l'amènent à collaborer notamment avec l'ABRC, CASRAI et le RCCDR.

CASRAI

- **Nature** - Organisme de normalisation sans but lucratif qui se concentre sur les données relatives à l'administration de la recherche et dont le conseil d'administration est formé de nombreuses parties prenantes.

- **Qui?** - Organisations de recherche (bailleurs de fonds, gouvernements fédéral et provinciaux, établissements, chercheurs) qui collaborent à l'évolution de la terminologie normalisée de la recherche et au progrès de la plateforme de normalisation de l'interopérabilité de la recherche. Les représentants des organisations membres font partie de divers comités, cercles d'examen et conseil consultatifs. CASRAI dirige l'établissement de miroirs internationaux de CASRAI Canada.
- **Mandat** - Offrir une tribune et les mécanismes nécessaires pour normaliser les données que les chercheurs, leurs établissements et leurs bailleurs de fonds doivent produire, stocker, échanger et traiter tout au long des activités de recherche.
- **Priorités** - Faire avancer les normes sémantiques qui faciliteront l'efficacité de l'exploitation dans un cadre numérique commercial.
- **Convergences/rerelations** - Participation au Conseil de direction de l'infrastructure numérique et à DRC.

Réseau canadien des Centres de données de recherche (RCCDR)

- **Nature** - Le Réseau sert de tribune et de structure pancanadiennes pour offrir aux chercheurs canadiens l'accès à des statistiques dans le domaine des sciences sociales et sur la santé de la population. Il diffuse de l'information destinée à la planification et à l'élaboration de politiques publiques efficaces. Le RCCDR surveille un ensemble national de centres de données de recherche. Un centre de données de recherche est un laboratoire de recherche universitaire qui emploie un analyste de Statistique Canada et qui offre aux chercheurs des services sur place :
 - *Accès protégé à des microdonnées confidentielles* - Recensements et enquêtes de Statistique Canada, et de nombreuses bases de données administratives; bases de données d'autres ministères fédéraux.
 - *Tout ce qu'il faut pour analyser les données* - postes de travail bien équipés, logiciels d'analyse statistique et soutien technique.
- **Qui?** - Quelque 45 établissements universitaires et Statistique Canada forment le noyau du RCCDR.
- **Mandat**
 - *Améliorer l'accès aux données* en donnant aux chercheurs canadiens un accès gratuit à des microdonnées détaillées venant d'un nombre croissant de répertoires de données administratives et d'enquêtes et de recensements.
 - *Accroître le nombre de spécialistes de la recherche quantitative* au Canada et former la prochaine génération de chercheurs.
 - *Accroître la pertinence de la recherche* en améliorant les communications entre les chercheurs en sciences sociales et les utilisateurs potentiels du savoir qu'ils produisent.
- **Convergences/rerelations**
 - Statistique Canada, CANARIE, les RORE, l'ICIST, divers ministères et organismes fédéraux œuvrant dans le domaine de la politique en matière de santé.

Canadian Polar Data Network (CPDN)

- **Nature** - Réseau canadien et organisme chargé de l'application de normes issu du centre de données établi dans le cadre de l'Année polaire internationale.
- **Qui?** - Partenariat formé par University of Alberta Libraries, University of Waterloo Canadian Cryospheric Information Network, le Scholars Portal du CBUO, Pêches et Océans Canada (gestion intégrée des données scientifiques) et l'ICIST.

- **Mandat** - Offrir une infrastructure durable de données de recherche qui englobe la conservation et l'accès pour la recherche scientifique sur les pôles (Arctique et Antarctique) et les activités de surveillance émanant du Canada ou faites au Canada.
- **Convergences/rerelations** - Ministères et organismes désignés dans la législation canadienne (comme la *Loi sur les océans*) afin de recueillir des données permettant de comprendre l'environnement, les ressources et les écosystèmes.
- **Priorités** - Le CPDN traitera en priorité des données qui ne font pas partie des mandats des divisions d'archivage de données gouvernementales.

Centre canadien de données en astronomie (CCDA)

- **Nature** - L'une des principales installations mondiales d'archivage et d'exploration de données en astronomie.
- **Qui?** - Division du CNRC.
- **Mandat** - Gestion, conservation, préservation et consultation de toutes les données de projets auxquels participent des astronomes canadiens (surtout des universitaires).
- **Convergences** - Les astronomes étrangers; l'Agence spatiale canadienne.

Annexe 3 : bibliographie

Royaume-Uni

Digital Curation Centre (DCC) UK. *A series of guides designed to build capacity and skills in research data management*. 2011 et 2012.

How to Appraise and Select Research Data
How to Cite Datasets and Link to Publications
How to Develop a Data Management and Sharing Plan
How to License Research Data
How to Write a Lay Summary

<http://www.dcc.ac.uk/resources/how-guides>

Green, Ann, Stuart Macdonald et Robin Rice. Policy-making for Research Data in Repositories: A Guide
Ann Green. 2009. <http://www.coar-repositories.org/files/guide.pdf>

JISC. e-Science Curation Report. Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision. 2003. http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal.pdf

JISC. Data centres: their use, value and impact. Septembre 2011.
<http://www.jisc.ac.uk/media/documents/publications/general/2011/datacentres.pdf>

JISC. Digitisation in the UK: The Case for UK Framework. Novembre 2005.
www.jisc.ac.uk/publications/programmerelated/2005/pub_digi_uk.aspx

Hey, Tony, éditeur. *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Microsoft Research, 2009.

Hey, Tony et Anne E. Trefethen. *The UK e-Science Core Programme and the Grid*
<http://users.ecs.soton.ac.uk/ajgh/FGCSPaper.pdf>

Hey, Tony et Anne Trefethen. *The Data Deluge: An e-Science Perspective 2003*.
http://eprints.soton.ac.uk/257648/1/The_Data_Deluge.pdf

Hey, Tony et Jessie Hey. « e-Science and Its Implications for the Library Community »
Library Hi Tech 24, n° 4 (2006): 515-28.
<http://www.emeraldinsight.com/Insight/ViewContentServlet?Filename=Published/EmeraldFullTextArticle/Articles/2380240404.html>

Maron, Nancy L, Jason Yun et Sarah Pickle. *Sustaining our digital future: Institutional Strategies for Digital Content*. 29 janvier 2013. JISC, Strategic Content Alliance.
<http://www.sr.ithaka.org/research-publications>

Maron, Nancy L, et Matthew Loy. *Funding for Sustainability: How Funders' Practices Influence the Future of Digital Resources*. Juin 2011. www.sr.ithaka.org/research-publications/funding-sustainability-how-funders-practices-influence-future-digital

STFC. Building an Open Data Infrastructure for Science: Turning Policy into Practice. Novembre 2012.
www.ambafrance-uk.org/IMG/pptx/5_AngloFrenchBicarregui.pptx

STFC. *e-Science Department Strategy 2011*. <http://www.stfc.ac.uk/e-science/resources/pdf/stfce-sciencedepartmentstrategy2011.pdf>

STFC *e-Science Annual Review* Daresbury Laboratory and Rutherford Appleton Laboratory 2009-10.
http://www.stfc.ac.uk/e-Science/resources/PDF/e-science_AnnRep_2010.pdf

Whyte, Angus et Jonathan Tedds. *Making the Case for Research Data Management*. DCC Briefing Paper
Septembre 2011. <http://www.dcc.ac.uk/resources/briefing-papers/making-case-rdm>

The Royal Society. *Science as an open enterprise*. Juin 2012.
<http://royalsociety.org/policy/projects/science-public-enterprise/report/>

États-Unis

ACLS. *Our Cultural Commonwealth, The report of the American Council of Learned Societies
Commission on Cyberinfrastructure for the Humanities and Social Sciences*. 2006.
<http://www.acls.org/cyberinfrastructure/ourculturalcommonwealth.pdf>

Edwards, Paul N., Steven J. Jackson, Geoffrey C. Bowker et Cory P. Knobel. Understanding
Infrastructure: Dynamics Tension and Design. *Report of a Workshop on "History & Theory of
Infrastructure: Lessons for New Scientific Cyberinfrastructures"*. Février 2007.
[http://deepblue.lib.umich.edu/bitstream/handle/2027.42/49353/UnderstandingInfrastructure2007.pdf
;jsessionid=2EB84DB8B71BC66EAFBD60D0CB5E2C8?sequence=3](http://deepblue.lib.umich.edu/bitstream/handle/2027.42/49353/UnderstandingInfrastructure2007.pdf;jsessionid=2EB84DB8B71BC66EAFBD60D0CB5E2C8?sequence=3)

Kahn, Scott D. *On the Future of Genomic Data* *Science* 331, 728 (2011); pp 728-729
DOI: 10.1126/science.1197891.

National Academies Press. *Ensuring the Integrity, Accessibility, and Stewardship of Research Data in
the Digital Age*. 2009. http://www.nap.edu/catalog.php?record_id=12615

National Academies Press. *Steps Toward Large-Scale Data Integration in the Sciences: Summary of a
Workshop*. 2010. http://www.nap.edu/catalog.php?record_id=12916

National Academies Press. *The Future of Scientific Knowledge Discovery in Open Networked
Environments: Summary of a Workshop*. 2012. http://www.nap.edu/catalog.php?record_id=18258

National Institutes of Health. *Data and Informatics Working Group Draft Report to
The Advisory Committee to the Director*. 15 juin 2012.
<http://acd.od.nih.gov/Data%20and%20Informatics%20Working%20Group%20Report.pdf>

National Science Foundation. *Revolutionizing Science and Engineering Through Cyberinfrastructure:
Report of the National Science Foundation Blue Ribbon Advisory Panel on Cyberinfrastructure*.
Janvier 2003.
<http://www.nsf.gov/od/oci/reports/atkins.pdf>

National Science Foundation. *Cyberinfrastructure Framework For 21st Century Science And Engineering*.
Mai 2012. <http://www.nsf.gov/od/oci/cif21/CIF21Vision2012current.pdf>

National Science Foundation. *Cyberinfrastructure for 21st Century Science and Engineering. Advanced
Computing Infrastructure Vision and Strategic Plan*. Février 2012.
<http://www.nsf.gov/pubs/2012/nsf12051/nsf12051.pdf>

National Science Foundation. *Cyberinfrastructure Framework for the 21st Century. A vision and
strategy for software for science, engineering and education*. 2012.
<http://www.nsf.gov/pubs/2012/nsf12113/nsf12113.pdf>

NITRD. *Five-year strategic plan for the Federal Networking and Information Technology Research and Development (NITRD) Program*. Juillet 2012.
<http://www.nitrd.gov/Publications/index.aspx>

NITRD. *Harnessing the Power of Digital Data for Science and Society*. Janvier 2009.
http://www.nitrd.gov/About/Harnessing_Power_Web.pdf

OSTP Policy Memorandum. *Increasing Access to the Results of Federally Funded Scientific Research*. 22 février 2013. <http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research>

President's Council of Advisors on Science and Technology (PCAST). *Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology*. 2010. Report to the President and Congress. <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd-report-2010.pdf>

President's Council of Advisors on Science and Technology (PCAST). *Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology*. 2010. Report to the President and Congress Janvier 2013.
<http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd2013.pdf>

Understanding infrastructure: Report of a Workshop on "History & Theory of Infrastructure: Dynamics, Tensions and Design. Lessons for New Scientific Cyberinfrastructures". 2007.
<http://deepblue.lib.umich.edu/bitstream/handle/2027.42/49353/UnderstandingInfrastructure2007.pdf;jsessionid=61ABF07F9DE8AFB603E22CBCB9F29F73?sequence=3>

Weller, Martin. *The Digital Scholar: How Technology Is Transforming Scholarly Practice*. 2011.
http://www.bloomsburyacademic.com/view/DigitalScholar_9781849666275/book-ba-9781849666275.xml

Australie

Australian Code for the Responsible Conduct of Research. 2007.
http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/r39.pdf

Gouvernement de l'Australie. *An Australian e-Research Strategy and Implementation Framework*. Final Report of the e-Research Coordinating Committee. Avril 2006.
<http://ncris.innovation.gov.au/Documents/eRCCReport.pdf>

Gouvernement de l'Australie. *Platforms for Collaboration. Summary Investment Plan*.
<http://ncris.innovation.gov.au/Documents/PfCInvPlansum.pdf>

Gouvernement de l'Australie. *Platforms for Collaboration Final Investment Plan*.
<http://ncris.innovation.gov.au/Documents/PfCInvPlanFinal.pdf>

Australian National Data Service (ANDS). *Diverse guides and policies*.
<http://ands.org.au/guides/index.html>

O'Brien, Linda. *The Changing Scholarly Information Landscape: Reinventing Information Services to Increase Research Impact*. Juin 2010. Proceedings ELPUB2010 - Conference on Electronic Publishing.
http://elpub.scix.net/cgi-bin/works/Show?111_elpub2010

Wolski, Malcolm, Joanna Richardson, Mark Fallu, Robyn Rebollo, Joanne Morris. 2011. *Developing the Discovery Layer in the University Research e-Infrastructure*.

Wolski, Malcolm, Joanna Richardson et Robyn Rebollo. Building an Institutional Discovery Layer for Virtual Research Collections. D-Lib Magazine, Mai-juin 2011, Volume 17, Numéro 5/6.
<http://www.dlib.org/dlib/may11/wolski/05wolski.html>

Europe

ASPIRE - *A Study on the Prospects of the Internet for Research and Education. Middleware and Managing Data and Knowledge in a Data-rich World*. Funded by FP7.
<http://www.terena.org/activities/aspire/docs/ASPIRE-data.pdf>

CERN. *Open Federated Identity Management for Research Collaborations* Avril 2012. CEWRN 2012-006.
<https://cdsweb.cern.ch/record/1442597/files/CERN-OPEN-2012-006.pdf>

Dallmeier-Tiessen S, Darby R, Gitmans K, Lambert S, Suhonen J, Wilson M (2012). *Opportunities for Data Exchange. Compilation of results on drivers and barriers and new opportunities*. 9 juillet 2012, pour FP7.
<http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/08/ODE-CompilationResultsDriversBarriersNewOpportunities1.pdf>

Deutsche Forschungsgemeinschaft (DFG). Taking Digital Transformation to the Next Level: The Contribution of the DFG to an Innovative Information Infrastructure for Research. Juillet 2012.
http://www.dfg.de/download/pdf/foerderung/programme/lis/strategy_paper_digital_transformation.pdf

DFG Committee on Scientific Library Services and Information Systems; Subcommittee on Information Management Recommendations for Secure Storage and Availability of Digital Primary Research Data. État de situation au 26 juin 2008.
http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf

Commission européenne. *Infrastructures TIC pour la science en ligne*. Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions. 2009.
<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0108:FIN:FR:PDF>

Commission européenne. *Open Infrastructures for Open Science*. Horizon 2020. Rapport de consultation.
<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/open-infrastructure-for-open-science.pdf>

Commission européenne. *Riding the wave. How Europe can gain from the rising tide of scientific data*. Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission. Octobre 2010.
<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

Commission européenne. *Advancing Technologies and Federating Communities A Study on Authentication and Authorisation Platforms For Scientific Resources in Europe*. 2012.
<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/aaa-study-final-report.pdf>

Commission européenne. *Development of impact measures for e-Infrastructures*. 2012.
<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/impact-study-final-report.pdf>

Commission européenne. *Developing World-class Research Infrastructures for the European Research Area (ERA). Report of the ERA Expert Group*.
http://ec.europa.eu/research/infrastructures/pdf/ri_era-expert-group-0308_en.pdf

Commission européenne. *G8+05 Global Research Sub-group on Data*. Draft Report, 28 octobre 2011. <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/g8.pdf>

Commission européenne. *Knowledge without Borders*. GÉANT 2020 as the European Communications Commons. Octobre 2011. <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/geg-report.pdf>

Commission européenne. *Global Virtual Research Communities*. ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/e-infrastructure/dge01224j-brochure5-global-virtual-research-com-low_en.pdf

Commission européenne. *Financing a Software Infrastructure for Highly Parallelised Codes* – IDC Final Report for the DG Information Society of the European Commission. Juillet 2011. <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/fisofi4hpc.pdf>

Commission européenne. *Digital Object Identifiers and Uniquer Author Identifiers to Enable Services for Data Quality Assessment, Provenance and Access*. Draft Final Report. Octobre 2011. <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/digoiduna.pdf>

Commission européenne. E-Infrastructure Reflection Group. *e-IRG Roadmap*. 2012. http://www.e-irg.eu/images/stories/publ/e-irg_roadmap_2012-final.pdf

Commission européenne. E-Infrastructure Reflection Group. *e-IRG Cloud Computing for research and science: a holistic overview, policy, and recommendations*. Final Version. 30 octobre 2012. http://www.e-irg.eu/images/stories/dissemination/e-irg_cloud_computing_paper_v.final.pdf

Commission européenne. Franco Accordino. *Digital Science and its impact on Scientific Societies*. 2010. www.cs.uu.nl/groups/AD/panel-DigitalScience.pdf

Commission européenne, Celina Ramjoué. *Towards a European Policy on Open Access*. Novembre 2012. http://www.unica-network.eu/sites/default/files/2012-11_27_UNICA-Celina.pdf

Commission européenne. EGI. *Linking digital resources across Eastern Europe for European science and Innovation*. 2012. <http://www.terena.org/activities/development-support/epe2012/presentations/11.pdf>

Commission européenne. *Strategic Plan for e-infrastructure. (European Grid Infrastructure EGI) EGI-InSPIRE*. 2010. <https://documents.egi.eu/public/RetrieveFile?docid=1098&version=3&filename=EGI-1098-D230-final.pdf>

Commission européenne. *Report on Integration of Data and Publications*. 17 octobre 2011. Susan Reilly, Wouter Schallier, Sabine Schrimpf, Eefke Smit, Max Wilkinson. <http://www.libereurope.eu/sites/default/files/ODE-ReportOnIntegrationOfDataAndPublication.pdf>

EU and UK Department for Culture, Media and Sport. *Dynamic Action Plan for the EU co-ordination of digitisation of cultural and scientific content*. 2005. <http://www.minervaeurope.org/publications/dap/dap.pdf>

GRDI2020. *Final Roadmap Report, Global Research Data Infrastructures: The Big Data Challenges*. Rapport présenté à l'Union européenne, février 2012. <http://www.grdi2020.eu/Repository/FileScaricati/e2b03611-e58f-4242-946a-5b21f17d2947.pdf>

Siena. *Roadmap on Distributed Computing Infrastructure for e-Science*. Mai 2012. <http://www.sienainitiative.eu/>

Autres documents étrangers

McKinsey & Company. *Big Data: The next frontier for innovation, competition, and productivity*. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers. Mai 2011.
http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation

OCDE. *Principles and Guidelines for Access to Research Data from Public Funding*. OCDE, 2007.
www.oecd.org/dataoecd/9/61/38500813.pdf

Canada

Backgrounder 2011 *Canadian Research Data Summit 2011 Canadian Research Data Summit. Mapping the Data Landscape*.
http://rds-sdr.cisti-icist.nrc-cnrc.gc.ca/docs/Summit_Backgrounder.pdf

Canadian Polar Data Network (CPDN) Governance Charter. 2013.
http://polardatanetwork.ca/wp-content/uploads/CPDN_Governance.pdf

ABRC. The Canadian National Collaborative Data Infrastructure Project. Final Report. January 2012. Prepared by Lynn Copeland, CNCDI Project Coordinator, and Martha Whitehead, Chair, CARL Data Management Subcommittee.
http://carl-abrc.ca/uploads/pdfs/carl_cncdi_final_report.pdf

Un environnement numérique canadien pour la recherche, l'innovation et l'éducation. Mémoire présenté dans le cadre de la Consultation sur l'économie numérique par Canadian Digital Media Network, Réseau canadien de la documentation pour la recherche, Conseil des DPI des universités canadiennes, CANARIE inc. et Calcul Canada. 28 juin 2010.
http://www.canarie.ca/templates/about/publications/docs/DES_Submission_F.pdf

Association canadienne des utilisateurs de données publiques (ACUDP). *Consultation the Future Role of the National Archives of Canada and the National Library of Canada*. Décembre 1998.
<http://datalib.library.ualberta.ca/data/capdu-english-consultation-submissions.pdf>

Document d'information sur le libre accès aux publications et aux données de recherche à l'intention des organismes subventionnaires canadiens.
<http://www.science.gc.ca/default.asp?lang=fr&n=2360F10C-1>

Digital Infrastructure Discussions: A précis of the discussions leading up to the Digital Infrastructure Summit, June 13-14, 2012 at the University of Saskatchewan, Saskatoon. 2012.
English, John. *The Role of the National Archives of Canada and the National Library of Canada*. A Report submitted to the Honourable Sheila Copps (1999-2000?).
<http://datalib.library.ualberta.ca/data/englishreport1999.pdf>

Génome Canada. Rapport de l'atelier sur la bio-informatique et la génématique. Toronto, Ontario, Canada. 5 et 6 décembre 2011. <http://www.genomecanada.ca/medias/pdf/fr/bioinformatics-meeting-report.pdf>

Humphrey, Chuck. *Preserving Research Data in Canada*. Blogue. Décembre 2012.
<http://preservingresearchdatainCanada.net/>

Bibliothèque et Archives Canada. *Stratégie canadienne sur l'information numérique; Version préliminaire à commenter*. Octobre 2007.
<http://www.lac-bac.gc.ca/obj/012033/f2/012033-1000-f.pdf>

Bibliothèque et Archives Canada. *Canadian Digital Information Strategy (CDIS): Final Report of consultations with stakeholder communities 2005 to 2008*. Publié en février 2010.
http://datalib.library.ualberta.ca/data/CDIS_FinalReport_eng_REVISED_Final.pdf

Mapping the Data Landscape: Report of the 2011 Canadian Research Data Summit. Décembre 2011. Produit par le Groupe de travail sur la stratégie des données de recherche.
http://rds-sdr.cisti-icist.nrc-cnrc.gc.ca/docs/data_summit-sommet_donnees/Data_Summit_Report.pdf

Site Web de Données de recherche Canada. <http://rds-sdr.cisti-icist.nrc-cnrc.gc.ca/fra/index.html>

Société royale du Canada. *Data Policy and Barriers to Data Access in Canada: Issues for Global Change Research. A Discussion Paper by the Data and Information Systems Panel of the Canadian Global Change Program*. Disponible à la Société royale du Canada. 1996.

Shearer, Kathleen et Diego Argáez. *Addressing the Research Data Gap: A Review of Novel Services for Libraries*. Report to CARL. 2010.
http://carl-abrc.ca/uploads/pdfs/library_roles-final.pdf

Mapping the Data Landscape: Report of the 2011 Canadian Research Data Summit.
http://rds-sdr.cisti-icist.nrc-cnrc.gc.ca/eng/events/data_summit_2011/index.html

Pearce, Nick, Martin Weller, Eileen Scanlon, Sam Kinsley. *Digital Scholarship Considered: How New Technologies Could Transform Academic Work* Technology & Social Media (Special Issue, Part 2)], in *education*, 2010, 16(1).
<http://www.ineducation.ca/article/digital-scholarship-considered-how-new-technologies-could-transform-academic-work>

Conseil de recherches en sciences humaines du Canada. *Consultation nationale sur les archives de résultats de recherche; Phase un : Rapport d'évaluation des besoins*. Mai 2001.
http://www.sshrc-crsh.gc.ca/about-au_sujet/publications/da_phase1_f.pdf

Conseil de recherches en sciences humaines du 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.
http://www.sshrc-crsh.gc.ca/about-au_sujet/publications/da_finalreport_f.pdf

Gérance des données de recherche au Canada : une analyse des écarts. 2009.
http://publications.gc.ca/collections/collection_2009/cnrc-nrc/NR16-123-2008F.pdf

Strong, David F. et Peter B. Leach. *Consultation nationale sur l'accès aux données de la recherche scientifique (CNADRS). Rapport final*. 31 janvier 2005.
http://publications.gc.ca/collections/collection_2012/crsng-nserc/NR16-84-2005-fra.pdf

Annexe 4 : définitions

<i>Terme ou expression</i>	<i>Définition</i>	<i>Source</i>
Données	Valeurs qualitatives ou quantitatives ou nombres qui sont considérés exacts (ou présumés l'être). Les données peuvent être brutes, ou primaires (découler directement de mesures), ou dérivées des données primaires, mais sans être le produit d'une analyse ou d'une interprétation autre qu'un calcul.	Science as an Open Enterprise
Données volumineuses	Ensembles de données dont la taille dépasse les capacités de saisie, de stockage, de gestion et d'analyse des outils logiciels de base de données ordinaires. Cette définition laisse entendre que la taille des ensembles de données assimilées à des « données volumineuses » augmentera au fil des progrès techniques. La définition peut également varier par secteur, selon la nature des logiciels disponibles et de la taille habituelle des ensembles de données dans une industrie particulière. Cela dit, dans de nombreux secteurs, les données volumineuses vont de quelques douzaines de téraoctets à un grand nombre de pétaoctets (milliers de téraoctets).	McKinsey Global Institute - Big data: The next frontier for innovation, competition, and productivity
Données de recherche	Dossiers factuels utilisés comme principales sources de la recherche, et couramment acceptés dans le milieu des chercheurs comme nécessaires pour valider les résultats de recherche.	Sommet 2011 sur les données de recherche canadiennes
Métadonnées	Littéralement, des « données sur les données » qui contiennent de l'information sur un ensemble de données, par exemple la raison et la méthode de collecte, ainsi que leur auteur et la date de création. Elles peuvent également être techniques et décrire la structure de l'ensemble de données, les conditions d'utilisation et les normes respectées.	Science as an Open Enterprise
Données ouvertes	Données qui respectent le critère de transparence intelligente. Elles doivent être accessibles, utilisables, évaluables et intelligibles.	Science as an Open Enterprise
Données sémantiques	Données auxquelles sont rattachées des métadonnées particulières. Les métadonnées peuvent servir à établir des liens entre des données.	Science as an Open Enterprise

Conservation	Activité qui consiste à gérer et à promouvoir l'utilisation des données dès leur création, afin qu'elles correspondent aux fins visées et qu'elles soient disponibles à des fins de découverte et de réutilisation. En ce qui concerne les données dynamiques, cette activité peut également comprendre leur valorisation ou leur mise à jour pour qu'elles servent encore leurs fins principales. Au niveau supérieur, la conservation comprend aussi l'établissement de liens annotés et avec d'autres publications.	JISC e-Science Curation Report
Archivage	Activité de conservation qui consiste à choisir des données et à les stocker afin d'y avoir accès et à assurer leur intégrité logique et matérielle à long terme, ce qui comprend la sécurité et l'authenticité.	JISC e-Science Curation Report
Préservation	Activité d'archivage destinée à assurer la gestion à long terme de certains éléments de données afin que l'on puisse encore les consulter et les comprendre, malgré les changements technologiques.	JISC e-Science Curation Report
Gestion des données	Ensemble des activités de traitement des données, de l'administration aux éléments techniques.	Sommet de 2011 : Portrait de la situation des données
Gérance des données	Plan organisationnel des rôles et responsabilités des personnes chargées de surveiller la gestion des données à tous les stades de leur cycle de vie, y compris leur préservation. Un grand projet de recherche peut faire intervenir plusieurs gérants de données aux divers stades du cycle de vie des données. Pour un petit projet, la gérance des données peut relever au premier chef du chercheur principal et de l'organisation chargée de préserver les données.	Sommet de 2011 : Portrait de la situation des données
Infrastructure de gestion des données de recherche (IGDR)	Configuration des effectifs, des services et des outils réunis pour appuyer la gestion des données tout au long d'une recherche et en particulier pour assurer une couverture complète des diverses étapes du cycle de vie des données. Elle peut être structurée à l'échelle locale ou mondiale, ou les deux, pour appuyer les activités relatives aux données de recherche tout au long de la recherche.	Blogue de Chuck Humphrey
Plan de gestion des données	Document officiel décrivant de quelle façon un chercheur ou un projet de recherche traite ses données pendant et après un projet de recherche.	Sommet de 2011 : Portrait de la situation des données

Politique de gestion des données	Principes généraux qui définissent un cadre de direction pour la gestion des données. Elle peut porter sur des éléments stratégiques, comme l'accès aux données, sur des enjeux juridiques pertinents, sur des questions de gestion des données et sur la garde des données, leur acquisition et d'autres enjeux.	Sommet de 2011 : Portrait de la situation des données
Interopérabilité des données	Façon de structurer les données de manière à pouvoir intégrer les divers ensembles de données.	Sommet de 2011 : Portrait de la situation des données
Ensemble de données	Groupe structuré de données.	Sommet de 2011 : Portrait de la situation des données
Métadonnée (1)	Information descriptive ou contextuelle qui désigne un autre objet ou une ressource ou qui y est rattachée. Il s'agit habituellement d'un ensemble d'éléments structuré qui décrit les ressources d'information et en facilite l'identification, la recherche et l'extraction par les utilisateurs, ainsi que la gestion de l'accès et du contenu.	Site Web du Data Curation Centre, Royaume-Uni
Métadonnée (2)	Littéralement, des « données sur les données ». L'information comprend habituellement des balises sémantiques associées à des données et qui servent à décrire les données. Elles peuvent être schématisées et utilisées à titre d'attributs dans une base de données.	Hey et Trefethen. The Data Deluge
Métadonnée descriptive	Permet aux utilisateurs d'identifier, de rechercher et d'extraire des ressources d'information et comprend souvent l'utilisation de vocabulaires contrôlés pour le classement et l'indexation des données, ainsi que des liens vers des ressources connexes.	Site Web du Data Curation Centre, Royaume-Uni
Métadonnée technique	Description des processus techniques employés pour produire un objet numérique ou exigés pour l'utiliser.	Site Web du Data Curation Centre, Royaume-Uni
Métadonnée administrative	Sert à gérer les éléments administratifs d'un objet numérique, comme les droits de propriété intellectuelle et l'acquisition de ces droits. Les métadonnées administratives décrivent également de l'information sur la création, la modification et le contrôle des versions des métadonnées elles-mêmes. On parle alors de méta-métadonnée!	Site Web du Data Curation Centre, Royaume-Uni
Métadonnée d'utilisation	Pour gérer l'accès par les utilisateurs et faire le suivi des utilisateurs et comprend des données sur les versions multiples.	Site Web du Data Curation Centre, Royaume-Uni
Métadonnée de préservation	Description des mesures prises pour préserver une ressource numérique, comme les migrations et les calculs des totaux de contrôle.	Site Web du Data Curation Centre, Royaume-Uni

Réseau maillé	<p>Infrastructure répartie qui est fédérée pour réunir des ressources venant de nombreuses organisations gérées par des domaines administratifs différents.</p> <p>Le réseau maillé vise à coordonner la mise en commun des ressources dans un cadre institutionnel dynamique et multi-institutionnel afin d'ajouter des fonctions supplémentaires aux éléments constitutifs : courtage, coordination des processus de travail, intégration des installations de calcul et de stockage. À cette fin, l'interopérabilité et les normes doivent être définies à divers niveaux : accès aux ressources, coordination et outils opérationnels, stockage et gestion des données, accès au réseau, etc.</p>	<p>Commission européenne, Advancing Technologies and Federating Communities</p> <p>Ibidem</p>
Cyberscience (1)	<p>Terme employé pour décrire les relations de collaboration croissantes à l'échelle mondiale entre les gens et les ressources partagées et qui seront nécessaires pour résoudre de nouveaux problèmes scientifiques et techniques.</p>	<p>Hey et Trefethen. The Data Deluge.</p>
Cyberscience (2)	<p>Science axée sur le calcul informatique et exécutée dans des environnements de réseaux fortement répartis ou science qui utilise d'immenses ensembles de données qui exigent des calculs en réseau. Ce terme englobe parfois les technologies qui permettent la collaboration répartie, comme le réseau d'accès.</p>	<p>Wikipédia</p>
Cyberscience (3)	<p>La cyberscience n'est pas une discipline scientifique à proprement parler. Elle désigne l'ensemble des outils et des technologies nécessaires pour appuyer la science collaborative en réseau. L'infrastructure de cyberscience vise à offrir aux scientifiques des méthodes de recherche plus rapides, de meilleure qualité et différentes.</p>	<p>Hey, Tony et Jessie Hey. "e-Science and Its Implications for the Library Community." <i>Library Hi Tech</i> 24, n° 4 (2006): 515-28.</p>
Recherche numérique	<p>La recherche numérique comprend :</p> <ul style="list-style-type: none"> • la constitution d'une collection numérique d'information à étudier et à analyser; • la création d'outils permettant de constituer des collections; • la création d'outils permettant d'étudier et d'analyser des collections numériques; • l'utilisation de collections numériques pour produire de nouveaux produits intellectuels; • la création de systèmes-auteurs adaptés aux nouveaux produits intellectuels, sous forme soit numérique soit traditionnelle. 	<p>Our Cultural Commonwealth</p>

Cyberinfrastructure	<ul style="list-style-type: none"> • Réseaux maillés de centres de calcul. • Bibliothèques exhaustives d'objets numériques. • Collections de données scientifiques bien conservées. • Instruments en ligne et vastes réseaux de capteurs. • Trousses logicielles commodées. 	In Our Cultural Commonwealth, à propos du rapport Aitkins (Royaume-Uni, 2003)
Infrastructure de recherche numérique	Matériel, installations et services de TIC qui appuient la recherche institutionnelle et dans des systèmes d'innovation nationaux et qui permettent aux chercheurs d'entreprendre des recherches de grande qualité et de produire des résultats axés sur l'innovation.	Présentation de Rys Francis
Cyberinfrastructure	Couches intercalaires entre la technologie sous-jacente (domaine de l'informatique) et une discipline scientifique. La priorité est accordée aux systèmes et services à valeur ajoutée qui peuvent être largement partagés entre des domaines scientifiques et qui appuient et favorisent de grandes avancées dans des sciences interdisciplinaires et multidisciplinaires tout en réduisant la répétition des efforts et le chevauchement des ressources, par exemple en tenant compte du matériel, des logiciels, du personnel, des services et des organisations.	Rapport Aitkins
Infonuagique	<p>Paradigme de calcul distribué à grande échelle axé sur les économies d'échelle et dans lequel une puissance de calcul regroupée, abstraite, virtualisée, évolutive sur demande et gérée, des installations de stockage, des plateformes de traitement et des services sont fournis sur demande à des clients externes par Internet. Ses éléments clés sont :</p> <ul style="list-style-type: none"> • un paradigme de calcul distribué spécialisé; • une évolutivité massive; • la possibilité de former une entité abstraite qui offre divers niveaux de service à des clients situés à l'extérieur du nuage d'information; • la recherche des économies d'échelle; • des services qui peuvent être configurés de façon dynamique (par la virtualisation ou d'autres méthodes) et fournis sur demande. <p>L'infonuagique est un moyen efficace d'offrir l'accès à de puissantes ressources de calcul sans devoir investir dans une nouvelle infrastructure, former de nouveaux employés ou acquérir des licences d'exploitation de logiciels.</p>	GRDI 2020

Infrastructure de données scientifiques	Ce dont les chercheurs ont besoin pour créer, stocker et mettre en commun les données issues de leurs recherches et pour trouver, consulter et traiter les données dont ils ont besoin.	Advancing technologies and federating communities
	<p>Données brutes recueillies ou produites pendant des expériences, l'étude ou l'observation de divers phénomènes (conformément à un modèle de recherche initial); les données sont ensuite analysées, et les résultats, publiés. Un processus de préservation est nécessaire à toutes ces étapes.</p> <ul style="list-style-type: none"> · Données structurées et ensembles de données produits par la sélection et le traitement des données (et qui appuient un modèle formel particulier). · Données publiées organisées de manière à appuyer une théorie scientifique ou des résultats de recherche. · Publication de données pour appuyer le regroupement, l'intégration et la transparence des données. 	
Flux de travaux	Description précise d'une procédure scientifique, c'est-à-dire les processus en plusieurs étapes qui permettent de coordonner des tâches multiples et de reproduire un scénario complexe. Chaque tâche représente l'exécution d'un processus de calcul. Les données produites par une tâche sont utilisées par des tâches ultérieures selon une topologie graphique préétablie qui permet d'orchestrer l'enchaînement des tâches.	Hey - The Fourth Paradigm. Goble et de Roure p. 138
Formats de fichiers de données	Les formats préférés sont conçus par un dépôt de données qui peut garantir que les données peuvent être converties en formats qui resteront lisibles et utilisables. Il s'agit habituellement des normes concrètes utilisées par un groupe particulier.	Extrait de Policy-making for Research Data in Repositories: A Guide 2009