

Enquête sur les initiatives canadiennes et internationales en matière de la gestion des données

par Diego Argáez et Kathleen Shearer

pour le compte du Groupe de travail sur la gestion des données de l'ABRC

(Document de travail)

28 avril 2008



Introduction

Les travaux de recherche sont maintenant de plus en plus axés sur les données et les chercheurs doivent utiliser des ensembles de données vastes et complexes. Grâce à l'accès sans précédent aux données et aux outils numériques, la recherche devient de plus en plus efficace et les découvertes sont facilitées. Faire en sorte que les données de recherche soient facilement accessibles afin qu'elles puissent être utilisées le plus souvent et le plus largement possible, c'est une question de bonne gestion des ressources publiques. (OCDE 2004)

En effet, l'efficacité de notre activité de recherche dépend de la façon dont ces données sont gérées. Dans le passé, l'archivage des données était géré par chacune des disciplines, par chacun des groupes ou chacun des chercheurs. Toutefois, étant donné le coût élevé de la création de collections de données et la complexité de leur gestion et de leur conservation, cette façon de procéder n'est plus jugée satisfaisante.

Les bibliothèques de recherche ont un rôle à jouer dans ce nouveau contexte où les données ont une importance primordiale. Dans un sondage qu'elle a réalisé en 2007, l'ABRC a constaté que la plupart de ses membres s'intéressent à la gestion des données de recherche, mais que peu ont une politique officielle d'archivage des données. L'ABRC a formé un groupe de travail sur la gestion des données de recherche afin d'aider les membres à recueillir, à organiser et à préserver les données de recherche, et à y donner accès, ainsi que de formuler une approche concertée pour l'ABRC.

Le présent rapport vise à donner une vue d'ensemble des types d'activités de gestion des données qui sont entreprises au Canada et à l'étranger. Nous présentons les diverses options offertes aux bibliothèques et nous préparons le terrain pour une enquête plus approfondie par le groupe de travail sur les rôles possibles des bibliothèques.

Nature de l'examen

Dans le rapport de la National Science Foundation intitulé *Long-lived Digital Data Collections* (<http://www.nsf.gov/pubs/2005/nsb0540/>), trois types de collections de données sont recensés (NSF, p. 20-21) :

[TRADUCTION]

Les collections de données de recherche sont les produits d'un ou plusieurs projets de recherche définies et elles contiennent généralement des données qui font l'objet d'un traitement limité. Elles peuvent ou non être conformes à des normes collectives, comme les normes pour les formats de fichier, la structure des métadonnées et les politiques d'accès au contenu. Ces collections bénéficient de budgets relativement modestes, souvent de subventions de recherche pour un projet donné.

Les collections de données sur les ressources ou propres à un groupe sont utilisées par un seul groupe de scientifiques ou d'ingénieurs. Ces collections numériques établissent souvent des normes propres au groupe soit par l'adoption de normes déjà existantes, soit par l'élaboration de nouvelles normes à la suite d'une intervention de l'ensemble du groupe, lorsqu'il n'y en a pas ou qu'elles sont insuffisantes. Les budgets pour les collections de données sur les ressources ou de données propres à un groupe sont moyens et proviennent généralement de subventions directes fournies par des organismes.

Les collections de données de référence visent à servir de vastes segments des milieux scientifiques et de l'éducation. Parmi les caractéristiques de cette catégorie de collections numériques, mentionnons un vaste éventail et un ensemble divers de groupes d'utilisateurs, c'est-à-dire des scientifiques, des étudiants et des éducateurs provenant de

disciplines, d'établissements et de lieux très différents. Dans ces circonstances, l'adoption de normes solides, bien établies et complètes est essentielle, et la sélection de normes par les collections de référence a souvent pour effet de créer une norme universelle.

Le présent examen visait à recenser les pratiques exemplaires et les modèles possibles pour l'archivage à long terme et l'accès aux données dans les bibliothèques. Nous avons donc cherché surtout à trouver des modèles différents qui font partie des deux dernières catégories. En outre, pour la plus grande partie, les centres de données ou les services qui n'archivent pas des données, mais offrent exclusivement l'accès aux données, ont été omis de cette étude (à l'exception de ODESI).

Il y a des milliers de collections de données dans le monde. Nous avons voulu dans la présente étude non pas dresser une liste exhaustive des initiatives d'archivage de données, mais plutôt recenser les divers modèles pour la gestion des données de recherche actuellement utilisés. Comme il s'agit d'un domaine qui évolue rapidement, nous avons aussi énuméré dans l'annexe du rapport les projets de démonstration ou en voie d'élaboration que nous avons relevés pendant l'étude.

Méthodologie

Nous avons fait le recensement des initiatives à la fois au Canada et à l'étranger que nous avons pu trouver dans la littérature et nos recherches sur Internet. Pour l'examen des initiatives canadiennes, nous avons aussi envoyé un questionnaire aux directeurs de l'ABRC afin de recenser d'autres initiatives utiles.

Le plus possible, nous avons recueilli les renseignements ci-après sur chaque initiative :

- Nom de l'initiative
- Organisme et pays
- Types de données recueillies
- Étendue des données recueillies
- Discipline
- Métadonnées
- Technologies
- Personnel
- Fonds
- Archivage et conservation
- Acquisition de données
- Accès
- Services
- Modèles d'organisation

Résultats

Les données numériques existent sous de nombreuses formes et se présentent sur un continuum d'analyse (depuis les données brutes jusqu'aux données amplement traitées). Les activités de gestion des données diffèrent beaucoup selon le type de données et le mandat de l'organisme en question. Il faut tenir compte de ces différences lorsqu'on examine la diversité des activités de gestion des données en cours.

Types de données : Tous les types de données ont été retenus dans l'étude, c'est-à-dire les chiffres, les images, les transmissions vidéo et audio, les logiciels et l'information sur les versions de logiciels, les algorithmes, les équations, les animations et les modèles ou simulations. Les types de données recueillies par chacune des archives de données sont très propres à une discipline. Il y a peu d'archives qui recueillent une grande variété de types de données dans un dépôt unique (DANS aux Pays-Bas, Arts and Humanities Data Service au Royaume-Uni), mais, pour la plus grande partie, les archives de données ont été spécifiquement conçues pour recueillir et gérer un nombre restreint de types de données dans une discipline définie.

Étendue des données : L'étendue des archives de données diffère selon le mandat et la discipline des archives en question. Généralement parlant, les archives nationales dans les sciences sociales recueillent des données relatives au pays en question, et les archives scientifiques recueillent généralement des données provenant de l'extérieur du pays. Des archives recueillent des données brutes, d'autres privilégient les données après analyse seulement, et certaines recueillent les deux.

Disciplines : Nous avons relevé des initiatives d'archivage des données dans quatre grandes catégories de disciplines : (1) arts et humanités; (2) sciences sociales, (3) sciences de la santé et (4) sciences naturelles et génie.

Métadonnées : La plupart des dépôts respectent des normes de métadonnées propres à chacune des disciplines pour les métadonnées descriptives. Dans les sciences sociales, les centres de données se conforment habituellement à la norme Data Documentation Initiative (DDI). Il n'existe pas de modèle général pour la représentation des métadonnées des études scientifiques. Toutefois, il y a actuellement un mouvement nouveau visant l'élaboration d'un ensemble commun de métadonnées de sorte que les ensembles de données de diverses disciplines scientifiques soient interopérables.

Technologies : Les technologies utilisées varient. De nombreuses archives utilisent des technologies qu'elles ont développées elles-mêmes, surtout pour l'analyse des données. Il existe aussi de grandes entreprises technologiques qui ont participé étroitement au développement de dépôts et d'outils (p. ex., Microsoft).

Traitement et conservation : Beaucoup d'archives de données visent principalement à traiter les données en vue d'en faciliter une analyse plus approfondie. Dans de nombreux cas, on ne sait pas exactement quelles activités de conservation sont entreprises par les archives. Les pratiques les plus souvent mentionnées sont le contrôle de la qualité, le stockage et la sauvegarde des données, ainsi que les métadonnées descriptives. Pour ce qui est de la conservation, beaucoup d'archives se conforment aux pratiques exemplaires, comme le modèle de référence Open Archival Information System, mais on ne sait pas exactement jusqu'à quel point ces types de normes rigoureuses sont respectées dans chaque dépôt. Les coûts et la reproductibilité des données sont deux facteurs à considérer pour la conservation.

Personnel : Le personnel des archives examinées s'échelonne de 5 employés à plein temps à plus de 50 employés à plein temps, selon la taille.

Acquisition des données : Les données sont acquises de diverses façons :

- Dans les sciences naturelles, les données sont habituellement acquises au moyen d'instruments scientifiques, comme les télescopes, mais certaines archives de plus grande importance accueillent aussi les données déposées par les chercheurs.
- Au R.-U., les centres de données administrés par le gouvernement exigent le dépôt par les chercheurs.
- Dans les arts et les sciences sociales, il s'agit habituellement à la fois de dépôt par les chercheurs et d'acquisition (parfois moyennant paiement) d'ensembles de données extérieures provenant d'autres organismes.
- Dans les sciences de la santé, où il y a une tradition selon laquelle les revues scientifiques exigent le dépôt des données avant la publication d'articles connexes et

imposent souvent aux chercheurs l'obligation de déposer les données, il y a beaucoup plus de dépôts par les chercheurs.

Fonds : Pour la plus grande partie, les archives de données que nous avons examinées sont subventionnées de l'une ou plusieurs des façons suivantes :

- Un ou plusieurs organismes subventionnaires
- Un organisme gouvernemental
- Un département universitaire
- Des revenus tirés de licences

Comme les collections de données de longue durée ont une envergure internationale et qu'elles sont conservées au-delà de la durée d'un projet de recherche, il reste encore difficile d'obtenir des fonds durables pour la conservation à long terme.

Accès : De nombreuses initiatives offrent l'accès gratuit à au moins une partie des ensembles de données. Voici les divers modèles utilisés :

- Adhésion par l'établissement ou obligation d'acquiescer à une licence OU accès payant à chaque fois pour les non-membres.
- Accès gratuit à des fins de recherche ou d'éducation ET accès payant à des fins commerciales.
- Accès gratuit à certaines données; et paiement au prix coûtant pour d'autres données OU interdiction d'accès à d'autres données à l'extérieur du groupe.
- Accès aux membres du groupe seulement. Pas d'accès à l'extérieur.

La plupart des archives de données imposent aux utilisateurs la signature d'accords de licence, qui régissent des éléments comme la confidentialité, la mention des auteurs et la redistribution des données. Pour ce qui est des données personnelles, toutes les archives qui offrent ce type de données ont établi des mécanismes assurant l'anonymat des données avant de les offrir à d'autres. Dans certains cas, les données pour lesquelles il est impossible d'assurer l'anonymat ne sont tout simplement pas disponibles. Un certain nombre d'archives permettent également aux chercheurs de restreindre l'accès à leurs données pour une période définie après le dépôt, afin qu'ils puissent publier leurs constatations avant la diffusion des données.

Les **services** offerts par les archives varient. On peut les grouper en quatre catégories :

Les services de dépôt qui offrent aux chercheurs des moyens de structurer, de marquer et de déposer les données

Les services d'interrogation, c'est-à-dire de recherche et d'extraction de données. Ces services sont axés surtout sur les technologies.

Les services d'analyse des données – comme la visualisation, le groupement. Ils sont également axés surtout sur les technologies et ces types de services sont très utilisés dans le domaine scientifique.

Les services-conseils par des experts et les services de soutien – il s'agit d'aider les créateurs de données à élaborer des plans de gestion des données, à préparer les données et à faire le marquage des métadonnées.

Modèles d'organisation : Quatre modèles d'organisation sont représentés :

- Centre de données indépendant
- Centre de ressources centralisées et partie d'un réseau distribué
- Série de bancs d'essai appuyés par un centre de coordination
- Réseaux distribués de centres de données interopérables

Autres tendances à signaler

- Les grands acteurs commerciaux commencent à s'intéresser à la collecte et au stockage des données de recherche (Microsoft, Google)
- Pour ce qui est des modèles d'organisation, il y a de moins en moins de centres indépendants centralisés et plus de réseaux distribués de centres de données interopérables
- Les nouvelles fonctionnalités exigeront l'interopérabilité entre les diverses disciplines et les divers types de données
- Le R.-U., les Pays-Bas et l'UE examinent actuellement le rôle des dépôts institutionnels pour la collecte et la gestion des données
- Les métadonnées sont reconnues comme la clé de l'accès aux données et de leur convivialité dans l'avenir (ce qui comprend la documentation du processus de recherche, et non seulement des données comme telles)
- Le nombre d'outils d'analyse des données augmente et ils deviennent plus complexes (p. ex., l'utilisation du SIG dans de nombreux domaines).
- On privilégie le traitement des données pour leur réutilisation, et non pour leur conservation à long terme.
- Les créateurs de données ont besoin d'aide pour la structuration et le marquage de leurs données afin qu'elles soient comprises par d'autres.
- Le libre accès aux données devient plus courant.

Rôle des bibliothèques

Il est manifeste que, si les bibliothèques doivent travailler étroitement avec l'ensemble des chercheurs, tous les services qu'elles offrent doivent correspondre aux besoins de l'ensemble des chercheurs qu'elles veulent aider. Dans la présente étude, nous avons relevé un certain nombre d'aspects préliminaires pour lesquels les bibliothèques de recherche pourraient contribuer à la gestion des données.

1. Dépôts de données – Établir et gérer des dépôts de données institutionnels, en vue de réaliser éventuellement un système transparent de bibliothèques et de services de données des bibliothèques qui sont en réseau et qui contribuent à l'étude et au traitement des données.

2. Élaboration des métadonnées – Les métadonnées sont un volet essentiel des données de recherche. Les bibliothèques de recherche peuvent être les chefs de file de l'élaboration de métadonnées automatisées uniformes et ontologiquement riches pour ces ensembles de données. L'élaboration et la gestion des métadonnées sont déjà des tâches reconnues dans le milieu des bibliothèques, mais les pratiques actuelles ne suffisent pas étant donné l'ampleur des travaux envisagés.

3. Interopérabilité – Les bibliothèques ont toujours collaboré ensemble. L'accès aux données et l'utilisation multidisciplinaire des collections de données distribuées exigeront l'adoption de méthodes uniformes de description lorsque les données seront créées.

4. Soutien et formation – Soutien pour la gestion de l'information personnelle, à mesure que les ensembles de données et l'information connexe deviennent de plus en plus transférables.

5. Préservation – Les pratiques actuelles dans les dépôts de données portent surtout sur l'utilisation immédiate des données, et non pas à long terme. Il est important de se donner les moyens pour ce qui est de la préservation.

Un certain nombre de rapports expliquant le rôle des bibliothèques en matière de gestion des données ont été publiés récemment, en particulier un rapport de l'ARL en 2006, intitulé *Agenda for Developing E-Science in Research Libraries*

(http://www.arl.org/bm~doc/ARL_EScience_final.pdf) et le rapport du R.-U., *Dealing with Data: Roles, Rights, Responsibilities and Relationships* (<http://connect.educause.edu/Library/Abstract/DealingwithdataRolesright/44533>). Ces rapports donnent un contexte important pour l'examen plus approfondi des constatations dont il est fait état ici. Dans les mois à venir, le Groupe de travail sur la gestion des données de l'ABRC définira davantage les rôles recensés ici et présentera des recommandations aux membres de l'ABRC pour la participation des bibliothèques dans ce domaine.

Constations sommaires pour 25 initiatives actuelles de gestion des données au **Canada** en matière d'**accès**, de **services** offerts, d'**acquisition de données**, de **modèles d'organisation**, de **préservation** et de **financement**.

Accès	N ^{bre} d'initiatives en matière de données	Services	N ^{bre} d'initiatives en matière de données	Acquisition de données	N ^{bre} d'initiatives en matière de données
Modèle de consortium, accès pour le personnel et les étudiants des établissements membres participants	2	Interrogation et extraction	11	La plupart des ensembles de données appartiennent déjà aux principaux établissements et les autres ont été acquis auprès de sources extérieures	1
Réservé au personnel et aux étudiants de l'université (un seul établissement)	2	Outils d'analyse en ligne	6	Entrevues téléphoniques assistées par ordinateur	2
Gratuit, fichiers de microdonnées et données agrégées à la disposition du public.	1	Outils d'analyse en ligne avec le logiciel Nesstar	3	Questionnaires d'enquête	6
Les ensembles de données sont retenus jusqu'à ce que leur anonymat soit assuré. Réserve aux chercheurs qui font une demande et sont autorisés	5	Services-conseils, services de formation sur les pratiques de gestion des données, personnel disponible pour donner de l'aide, etc.	5	Données acquises auprès de sources extérieures (p. ex., Statistique Canada, une entreprise de sondage, etc.)	7
Gratuit	12	Dépôt de données	4	Données d'enquête extraites de fichiers administratifs	4
Une partie du contenu est gratuit, une autre partie est payante	2	Préparation des données, anonymisation, formatage, etc.	4	Dépôt pour les chercheurs	3
Les ensembles de données demandés sont payables « au prix coûtant »	1	Fichiers d'aide en ligne et didacticiels	1	Instruments scientifiques	7
Deux niveaux d'accès : gratuit pour le personnel et les étudiants dans les établissements participants qui versent une cotisation; les chercheurs qui présentent une demande et sont autorisés.	1	Divers rapports et publications sur les données ou documents explicatifs du contexte des ensembles de données	3	Inconnu	1
		Rapport de données personnalisées (payant)	2	Documents du recensement national	5
		Accès à des données et à des images en temps réel	2		

Modèles d'organisation	N^{bre} d'initiatives en matière de données
2 établissements principaux, appuyés par le Groupe consultatif externe et le Groupe de gestion opérationnelle des partenaires	1
Projet indépendant soutenu par des universités partenaires, le gouvernement fédéral et un autre établissement de recherche	1
Centre/projet de données indépendant	7
Centre/projet de données indépendant avec participation de l'établissement parent ou d'un autre organisme de gestion des données	1
Organisme ayant un mandat fédéral regroupant divers partenaires à l'échelle nationale	1
Réseau distribué	6
Projet de démonstration	1
Centre/projet de données indépendant faisant aussi partie d'un réseau distribué	8

Préservation	N^{bre} d'initiatives en matière de données
Contrôle de la qualité de l'annotation structurée (métadonnées)	11
Méthode de traitement inconnue	10
Une certaine forme d'archivage	3
Système d'archivage de gestion des données (DMAS) pour les données à la fois courantes et archivées	2

Financement	N^{bre} d'initiatives en matière de données
Financement par un organisme subventionnaire	8
Source de financement inconnue	3
Financement partiel par les établissements parents/membres	2
Financement par les pouvoirs publics	17
Financement offert par plus d'un organisme subventionnaire	5
Financement par les cotisations des établissements participants	2
Subvention par le secteur privé	8
Financement supplémentaire par divers organismes de recherche	4

Sommaire des constatations pour **17** initiatives actuelles de gestion des données **internationales** en matière d'**accès**, de **services** offerts, d'**acquisition de données**, de **modèles d'organisation**, de **préservation** et de **financement** :

Accès

Dans le cas de **12** projets, les données sont offerts gratuitement à tout le monde. Dans un cas, le chercheur doit présenter une demande avant d'obtenir les données sans frais (Henry A. Murray Research Archive); dans le cas de l'un des projets, les données sont offertes gratuitement mais seulement aux chercheurs des établissements participants (Australian Social Science Data Archive) et les chercheurs non participants paient l'accès « au coût » pour chaque ensemble de données. Dans le cas d'un autre projet, le libre accès est accordé conformément à certaines modalités de licence et à la condition que la source des données utilisées soit mentionnée.

Dans le cas de deux projets de gestion des données à l'étranger, l'accès est assorti de droits : Australian Social Science Data Archive « au prix coûtant » par ensemble de données demandées pour les chercheurs non participants, et les centres de données du Natural Environment Research Council (NERC) exigent des droits selon l'utilisation prévue des données demandées.

Dans le cas d'un projet, le modèle d'accès aux données est fondé sur l'accès aux membres qui versent des droits au consortium : Inter-University Consortium for Political and Social Research.

Dans le cas d'un projet, l'accès « au prix coûtant » est accordé à certains ensembles de données et l'accès est gratuit en général pour les membres des milieux de l'éducation supérieure ou permanente, et on offre aussi beaucoup de documentation gratuite au grand public : Arts and Humanities Data Service.

Deux des initiatives de gestion des données offrent l'accès par la voie d'un modèle d'abonnement : Biology Image Library (taux pour les particuliers, les établissements et les sociétés) et Cambridge Structural Database.

Dans les sciences de la santé, on offre de plus en plus l'accès gratuit aux données; **5** des **6** initiatives de gestion des données en sciences de la santé à l'étranger que nous avons recensées offrent l'accès gratuit à leurs ensembles de données archivées.

Accès gratuit aux ensembles de données dans d'autres disciplines :

Arts et humanités : **1**

Sciences sociales : **3**

Sciences naturelles et génie : **3**

Initiatives **internationales** (suite)

Services	N ^{bre} d'initiatives en matière de données
Service de formation aux pratiques de gestion des données	5
Aide (p. ex., fichiers d'aide en ligne ou assistance d'experts)	3
Aide au traitement (p. ex., normes reconnues de documentation des données, contrôle de la qualité pour les ensembles de données déposées, etc.)	6
Préservation de longue durée	3
Dépôt de données	6
Outils d'interrogation et d'extraction en ligne	10
Outils d'analyse en ligne (p. ex., analyse des données sans téléchargement de fichiers, visualisation, présentations graphiques, etc.)	8
Conversion des données	4

Acquisition de données	N ^{bre} d'initiatives en matière de données
Dépôt pour les chercheurs	6
Sources externes (p. ex., universités, organismes gouvernementaux, entreprises d'études de marché, etc.) et dépôt pour les chercheurs	6
Sources externes (p. ex., universités, organismes gouvernementaux, entreprises d'études de marché, etc.)	1
Inconnu	2
Dépôt pour les chercheurs et pour les revues	1
Instruments scientifiques	1

Modèles d'organisation	N ^{bre} d'initiatives en matière de données
Réseau distribué	6
Centre de données indépendant (certains avec des unités exécutant certaines opérations – p. ex., dans le cas de l'ICPSR, développement de collections, organisation de collections et ressources éducatives)	5
Centre centralisé de ressources et partie d'un réseau distribué	5
Série de bancs d'essai avec l'aide d'un centre de coordination	1

Préservation	N^{bre} d'initiatives en matière de données
Une certaine forme d'archivage	5
Sauvegarde quotidienne, archivage et planification de la migration	2
Inconnu	6
Contrôle de la qualité de l'annotation structurée (métadonnées)	6

Financement	N^{bre} d'initiatives en matière de données
Organismes subventionnaires et l'établissement d'accueil	2
Organismes subventionnaires, l'établissement d'accueil et soutien en nature d'un autre organisme (p. ex., un organisme d'informatique avancée)	1
Divers établissements et organismes subventionnaires	1
Plus d'un établissement ou organisme subventionnaire	4
Cotisations à un consortium ou abonnement	3
Ensemble d'autofinancement (à partir d'un <i>budget principal</i>) et acquisition de licences de produits de données	1
Fondation	1
Un seul organisme subventionnaire	3
Inconnu	1
Cotisations volontaires internationales	1

Initiatives **internationales** (suite)

Sur les **19** projets de démonstration et projets de données en développement à l'étranger, nous avons retenu les observations ci-après. Remarque : elles sont incomplètes comparativement au résumé des autres initiatives de gestion des données documentées parce qu'elles en sont à diverses étapes de développement et qu'il n'y avait pas beaucoup d'information à leur sujet sur Internet.

- Dans le cas de **trois** initiatives, les services d'interrogation et d'extraction sont mentionnés. P. ex., les outils d'interrogation et d'extraction en ligne – cinq grilles de données fédérées permettant aux chercheurs de n'importe quel site du World Universities Network (WUN) de rassembler en toute transparence des collections représentant diverses ressources de nombreux établissements
- L'interopérabilité est un élément commun dans la planification de certains des projets. P. ex., l'élaboration d'intergiciels pour des projets communs interdisciplinaires. **Cinq** des projets de démonstration ou initiatives en voie d'élaboration prévoient l'interopérabilité.
- Dans le cas d'au moins **trois** des initiatives internationales, les dépôts institutionnels seront un élément important de l'infrastructure. Dans un cas, l'un des services sera l'établissement d'un lien entre les publications et les ensembles de données auxquelles les publications renvoient.
- Dans au moins **cinq** cas, on a l'intention d'archiver les données déposées pour l'accès à long terme. P. ex., il est question dans un cas de l'archivage dans les DI, dans un autre cas, il s'agit d'offrir l'archivage durable des données numériques et, dans un autre cas, d'offrir des outils propres au contexte et des métadonnées pour le traitement afin de faciliter la réutilisation ultérieure de l'information déposée.
- Au moins **trois** offriront des services de dépôt de données, dont l'un un modèle spécialisé pour le dépôt d'ensemble de données « orphelines »
- Il est question dans **un** cas de services prévus pour l'extraction des données en ligne
- Dans le cas de **l'un** des projets de démonstration, on prendra des dispositions pour offrir des images brutes.
- Dans au moins **trois** cas, il y aura des activités d'éducation et de sensibilisation pour les modèles opérationnels. P. ex., « On cherchera à sensibiliser davantage les scientifiques aux possibilités des archives numériques. »