

En favorisant la génération de volumes considérables de données, à la fois manipulables et exploitables, l'évolution actuelle des technologies a permis l'émergence d'un quatrième paradigme : celui de la science guidée par les données. S'y ajoute l'effacement des barrières géographiques et disciplinaires au profit d'une science mondialisée et transdisciplinaire. Tant technologiques que scientifiques, ces évolutions s'inscrivent enfin dans la mouvance *open data*, incitation à l'ouverture des données publiques, renforcée à l'échelle européenne par le programme cadre de recherche « Horizon 2020 ».

L'Inist-CNRS face au défi des données de la recherche

Les pratiques et usages des chercheurs ont évolué en conséquence, favorisant l'émergence de nouveaux besoins, qui témoignent du passage d'une logique de services centrés sur les publications à une nouvelle gamme d'objets numériques, parmi lesquels les données de la recherche figurent en bonne place.

Ce contexte global a conduit l'Institut de l'information scientifique et technique (Inist), unité propre de service du CNRS – reconnue pour ses activités traditionnelles de fourniture de documents et de production de bases de données bibliographiques – à entamer une vaste réflexion prospective sur l'extension de son périmètre opérationnel dans le but d'y inclure des services de gestion et de valorisation des données de la recherche.

MISE EN PLACE D'UNE ORGANISATION DÉDIÉE

Créée en septembre 2012, une cellule de coordination « Données de la recherche » composée de deux ingénieurs, vise à assurer la gestion et l'harmonisation des activités de l'Inist relatives aux données de la recherche. Ces deux coordinateurs animent un groupe de huit correspondants-documentalistes issus des services scientifiques en sciences, techniques et médecine (STM) et sciences humaines et sociales (SHS) dont le rôle est de diffuser l'état d'avancement des travaux de la cellule au sein de leurs services respectifs.

Ils gèrent parallèlement les relations entre les différents pôles métiers de l'Institut et interviennent sur le plan national et international avec les partenaires externes dans les activités et projets liés aux données de la recherche.

RENFORCEMENT DES COMPÉTENCES : UNE APPROCHE CONSTRUCTIVISTE

À l'instar de nombreux pays, la recherche française souffre de fortes réductions budgétaires qui contraignent au redéploiement du personnel en place plutôt qu'à des recrutements supplémentaires. Bien qu'en

France aucune offre de formation continue ne soit encore disponible, l'expérience d'autres pays plus avancés en la matière comme le Royaume-Uni ou les États-Unis s'avère en revanche riche d'enseignement, exploitable notamment dans des méthodes d'apprentissage constructiviste ou de type projet. Une étude de faisabilité a déterminé les nouveaux besoins des chercheurs en matière de gestion de données de la recherche, ainsi que les connaissances et compétences nécessaires pour y répondre. Elle a permis aux coordinateurs, qui bénéficiaient d'une expérience acquise sur de précédents projets, d'initier une approche de formation continue au sein de la cellule.

Les correspondants continuent pour leur part à être impliqués dans l'indexation des bases de données bibliographiques Pascal et Francis, tout comme dans les activités liées à la terminologie, en consacrant désormais près d'un tiers de leur temps de travail à l'apprentissage de nouvelles compétences. Ces spécialistes de l'information développent progressivement une culture des données dans leurs domaines d'expertise scientifique, destinée à court terme à guider les chercheurs en matière de bonnes pratiques (entrepôts de données, standards de métadonnées, outils et terminologie...).

En parallèle, certains d'entre eux participent actuellement à des projets pilotes de gestion de données de recherche au cours desquels ils bénéficient d'une période d'immersion dans les laboratoires et acquièrent par là même une meilleure compréhension du processus de recherche.

Cette approche méthodologique visant à accompagner l'évolution du métier de documentaliste a été présentée à Munich lors de la 42^e conférence annuelle¹ de la Ligue des bibliothèques européennes de recherche (Liber).

PARTENARIATS NATIONAUX ET INTERNATIONAUX

L'Inist a récemment établi des coopérations sur le plan local avec deux unités mixtes de recherche

[1] www.liber2013.de

[2] www.sim4rdm.eu

[3] <http://rbdd.cnrs.fr>

[4]

www.donneesdelarecherche.fr

[5] www.ortolang.fr
(ANR-11-EQPX-0032)

[6] www.datacite.org

[7] Contact : datasets@inist.fr



© Phot. Georges Fessy

CNRS-Université de Lorraine :

- une application de gestion des matériels d'étude propres à un laboratoire de biologie moléculaire, qui permettra de faciliter les échanges et la réutilisation des données ;
- une étude de cas réalisée dans le cadre d'un projet européen *Support Infrastructure Models for Research Data Management (SIM4RDM)*² amenant le laboratoire à adopter les bonnes pratiques de gestion de données.

Au niveau national, l'Inist collabore avec le Réseau Bases de données (R-BDD)³, dont la mission consiste à valoriser les bases de données produites par les communautés de recherche en les rendant interopérables sur les plans technique et sémantique. L'Institut héberge et alimente également une plateforme d'information et de veille⁴ sur les données de la recherche, créée en 2012 à l'initiative du ministère de l'Enseignement supérieur et de la Recherche. Par ailleurs, depuis 2011, l'Inist est membre de l'Équipex Ortolang (*Open Resources and TOols for LANGUAGE*)⁵, validé dans le cadre des Investissements d'avenir. Son but est de proposer une infrastructure en réseau offrant un réservoir de données (corpus, lexiques, dictionnaires, etc.) et d'outils sur la langue et son traitement disponibles et documentés. Cette infrastructure poursuit un double objectif :

- la mutualisation d'acquis et le partage de connaissances sur le français et les langues de France pour les communautés de linguistes et de traitement automatique du langage ;
- l'usage et le transfert des ressources et outils mis en place au sein des laboratoires publics vers les partenaires industriels.

L'Inist assure le rôle d'opérateur technique (infrastructure) pour l'hébergement et l'exploitation de la plateforme. De plus, il participe à la définition de la structure des métadonnées attendues, préparant ainsi sa fonction de maintenance des corpus déposés dans le temps.

Au niveau international, l'Inist est membre du consortium DataCite⁶, qui a comme objectif de valoriser les données de la recherche et les informations non textuelles dans la communication scientifique. L'Institut est à ce titre habilité à attribuer des DOI (*Digital Object Identifier*)⁷, identifiant pérenne qui facilite la citabilité contribuant ainsi à une meilleure visibilité et accessibilité des données.

DE NOUVELLES ORIENTATIONS STRATÉGIQUES

Soucieux de relever le défi des évolutions technologiques et scientifiques, l'Inist a mis en place une organisation fonctionnelle spécifiquement dédiée. Accompagner l'évolution du métier d'ingénieur-documentaliste vers la notion anglo-saxonne de « *data librarian* », développer de nouvelles compétences destinées à apporter de nouveaux services : ainsi s'affiche sa volonté de concevoir et développer une véritable infrastructure sociale de soutien aux chercheurs. D'ores et déjà impliqué dans de nombreux projets nationaux et internationaux, l'Institut fait montre de réactivité afin de relever les nouveaux enjeux de l'*Open Science*.

CATHERINE POUPON-CZYSZ

Responsable du Pôle Structuration & valorisation des connaissances en IST, Inist-CNRS
Catherine.Poupon-Czys@inist.fr



Le bâtiment de l'Inist à Vandœuvre-lès-Nancy.



POUR EN SAVOIR PLUS

CONTACT :

DonnéesRecherche@inist.fr