

## LE PARTAGE DES DONNÉES SCIENTIFIQUES À L'ÈRE DE L'E-SCIENCE : L'INSTRUMENTATION DES PRATIQUES AU SEIN D'UN COLLECTIF MULTIDISCIPLINAIRE

[Florence Millerand](#)

ENS Paris-Saclay | « Terrains & travaux »

2011/1 n° 18 | pages 215 à 237

ISSN 1627-9506

Article disponible en ligne à l'adresse :

-----  
<https://www.cairn.info/revue-terrains-et-travaux-2011-1-page-215.htm>  
-----

Distribution électronique Cairn.info pour ENS Paris-Saclay.

© ENS Paris-Saclay. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Florence Millerand

# Le partage des données scientifiques à l'ère de l'*e-science* : l'instrumentation des pratiques au sein d'un collectif multidisciplinaire

**L**A MISE en réseau des chercheurs, des laboratoires et des institutions scientifiques, à travers le développement d'un vaste ensemble d'infrastructures numériques, constitue l'une des grandes mutations techniques contemporaines<sup>1</sup>. Parmi ces développements technologiques de grande envergure, de grandes bases de données se mettent en place pour favoriser la promotion de la collaboration scientifique à grande échelle via la publication et le partage de données de recherche. À l'instar des premiers projets de banques de données en sciences de la vie (*GenBank* ou la *Protein Data Bank*), un mouvement similaire se développe en sciences de la nature, où les défis scientifiques actuels (le changement climatique ou la biodiversité par exemple) requièrent d'une part la mise en commun d'expertises multidisciplinaires, et de l'autre le partage de ressources (instruments, données, etc.) Parallèlement, le mouvement actuel en faveur de la publication des données scientifiques à travers des archives ouvertes, qui vise à rendre publiquement accessibles les données de recherche financées à même les fonds publics encourage le développement de grandes bases de données partagées (Arzberger *et al.*, 2004a, 2004b).

---

1. D'importants programmes de financement de projets d'infrastructures basées sur les réseaux numériques ont ainsi été mis en place dans la dernière décennie, tels que l'*Initiative Cyberinfrastructure* de la *National Science Foundation* aux États-Unis en 2000 ou le programme *E-Science* du Conseil de recherche britannique initié en 1999.

Les hypothèses sont nombreuses sur les changements associés à ces récents développements. Prenant pour exemple la biologie moléculaire, où le dépôt des séquences d'ADN, protéines et autres, dans les grandes bases de données du domaine est désormais incontournable (Brown, 2003), on évoque l'émergence d'un nouveau régime de production de connaissances parallèlement au modèle canonique de la publication d'articles (Hilgartner, 1995). Certains vont même jusqu'à évoquer un changement de paradigme (Gilbert, 1991). Beaucoup plus rares sont les études qui cherchent à appréhender ces mutations à l'échelle des collectifs à partir d'études empiriques. Hine (2006) envisage les bases de données comme des instruments de recherche – et pas seulement comme un médium de publication – susceptibles de faire émerger de nouvelles formes d'organisation du travail scientifique. Bowker (2000) insiste sur les dimensions politiques et éthiques des bases de données en montrant les traitements différenciés qui sont réservés aux données de biodiversité en fonction de la formation académique des chercheurs. Plus largement, les travaux récents sur les infrastructures d'information ("*Infrastructure Studies*") fortement inspirés des travaux de Star sur les infrastructures matérielles (Star et Ruhleder, 1996 ; Star et Bowker, 2002) invitent à situer l'analyse sur les pratiques de conception et de développement de ces innovations sociotechniques pour saisir dans l'action, et non pas lorsqu'elles sont stabilisées, les modalités par lesquelles elles naissent, évoluent, et éventuellement contraignent les contenus et les usages (Bowker et Star, 1999 ; Edwards *et al.*, 2007 ; Karasti *et al.*, 2010 ; Millerand et Bowker, 2008, 2009 ; Ribes et Finholt, 2009).

Cet article vise à rendre compte des résultats préliminaires d'une recherche empirique ethnographique réalisée au sein d'une communauté scientifique canadienne, le réseau *ArcticNet*, qui développe depuis 2007 une infrastructure numérique pour la publication et le partage de données de recherche : le *Polar Data Catalogue*, une grande base de données sur la recherche polaire. La démarche ethnographique mobilisée dans le cadre de cette recherche vise à saisir les modalités de développement de l'infrastructure et la manière dont ses usages modifient les pratiques de recherche<sup>3</sup>. Les techniques de l'observation participante, l'analyse de

---

2. Le label « *Infrastructure Studies* » regroupe les travaux récents qui se sont développés sur les infrastructures d'information à partir d'une réflexion sur les modalités de leur émergence. Voir en particulier Bowker *et al.* (2010) et Edwards *et al.* (2009), qui posent les jalons d'un possible programme de recherche sur les infrastructures d'information scientifiques.

3. La recherche ethnographique, initiée en 2008, est financée pour une durée de 3 ans par une subvention du Fonds Québécois de Recherche sur la Science et la Culture (FQRSC).

documents (comptes rendus de réunions, documents de présentation de l'infrastructure...), l'analyse des prototypes (incluant la documentation technique) et la conduite d'entretiens semi-directifs auprès des différents groupes d'acteurs concernés (développeurs, gestionnaires de données, chercheurs, étudiants, personnel de recherche...) ont été utilisées. Les éléments présentés dans cet article résultent d'une première analyse conduite sur des comptes rendus d'observations, des documents, des prototypes ainsi que sur les verbatims de 24 entretiens.

Ce qui nous intéresse plus particulièrement ici, c'est la question de l'*instrumentation* des activités de publication et de partage de données par l'entremise d'infrastructures techniques numériques comme les grandes banques de données partagées. Nous cherchons à caractériser le rôle structurant des infrastructures matérielles dans les activités de production de connaissances. Notre questionnement est triple : Dans quelle mesure peut-on observer de nouvelles pratiques de publication et de partage de données au sein du collectif de recherche ? Quels seraient les nouveaux rôles et formes de division du travail entre les acteurs, en lien avec les pratiques développées ? Dans quelle mesure cette infrastructure technologique – et les nouvelles pratiques qu'elle suscite – façonnent-elles et modifient-elles les modalités de représentation et d'échange des connaissances, par exemple en donnant à voir de nouveaux rapports aux savoirs ?

Nous faisons l'hypothèse que l'infrastructure de partage de données développée au sein d'*ArcticNet* participe à la structuration du collectif de recherche à partir d'une mise en visibilité différenciée des données scientifiques qui y sont produites d'une part, et par le développement d'une routine organisationnelle qui vient renforcer les formes de division du travail en place d'autre part. Le développement de l'infrastructure s'accompagne d'un processus d'instrumentation des activités de publication et de partage de données qui donne à observer l'existence de cultures épistémiques fortement différenciées au sein du collectif, et qui interviennent dans les rapports des chercheurs à leurs données de recherche. Plus généralement, les nouvelles pratiques associées à l'usage de l'infrastructure posent la question de la possibilité d'une réorientation du regard des scientifiques sur leurs données de recherche, désormais représentées sous la forme de descriptions codifiées (standardisées) et dépourvues de leur statut de propriété individuelle au profit de celui de bien commun.

Nous commençons par exposer le cas à l'étude, à savoir la communauté scientifique *ArcticNet* et son projet de grande base de données. Nous présen-

tons par la suite trois éléments d'analyse du processus d'instrumentation des activités de publication et de partage des données scientifiques qui participent de la structuration du collectif : la prégnance des cultures épistémiques – révélées par la grande disparité des pratiques développées ; le renforcement des rôles et des formes de division des tâches au sein des collectifs ; et la possibilité d'une réorientation du regard des scientifiques sur leurs données de recherche. Nous concluons en proposant des pistes de recherche pour l'étude du rôle des grandes bases de données dans les pratiques scientifiques.

## ■ L'instrumentation du partage de données au sein du collectif *ArcticNet*

Créé en 2002, *ArcticNet* est un réseau de recherche interdisciplinaire canadien centré sur l'étude des conséquences du changement climatique sur les écosystèmes arctiques. Il s'agit d'un réseau de centres d'excellence du Canada, une innovation en matière de politique scientifique canadienne qui entend favoriser la collaboration scientifique entre les disciplines et entre les milieux académiques, gouvernementaux et industriels (Fisher *et al.*, 2001 ; Atkinson-Grosjean, 2006)<sup>4</sup>. Collectif hétérogène s'il en est, la communauté *ArcticNet* rassemble un peu plus de 150 chercheurs, provenant des sciences naturelles, des sciences de la santé et des sciences sociales, répartis dans trente universités canadiennes et huit ministères canadiens fédéraux, auxquels s'ajoutent une centaine de partenaires provenant des organisations inuites, des communautés nordiques, du secteur privé ainsi que d'équipes scientifiques internationales<sup>5</sup>. Les disciplines représentées sont majoritairement les sciences de la nature (80% des chercheurs) et dans une moindre mesure les sciences de la santé (10%) et les sciences sociales (10%).

En 2007, à l'initiative d'une poignée de chercheurs du collectif et en réponse aux recommandations croissantes des organismes de financement canadiens, un projet d'infrastructure de publication et de partage de données est mis en place. Le projet prend forme parallèlement à la

---

4. La création des Réseaux de centres d'excellence du Canada date de 1988. Le premier réseau a vu le jour en 1990.

5. Au total, le réseau compte un peu plus de 400 personnes en incluant les étudiants, les chercheurs postdoctoraux, les techniciens, le personnel de recherche et les partenaires (site Web d'*ArcticNet* : <http://www.arcticnet.ulaval.ca/>).

définition d'un plan de gestion des données au sein du collectif, qui donnera lieu à la publication d'une charte en matière de partage de données<sup>6</sup>. L'infrastructure vise à constituer, à l'origine, la plus grande base de données sur l'Arctique pour s'étendre, deux ans plus tard, à toute la recherche polaire (incluant donc l'Arctique et l'Antarctique). Concrètement, le *Polar Data Catalogue* est une base de données en ligne librement accessible sur le Web, qui vise à centraliser l'ensemble des données de recherche sur les régions polaires incluant, au premier chef, celles collectées par le collectif *ArcticNet*. Précisons que l'initiative a émergé au sein même du collectif (le principal acteur du projet étant le responsable scientifique d'un des quatre thèmes de recherche du collectif, auquel se sont joints deux chercheurs partenaires) et il est, par ailleurs, fortement appuyé par la direction scientifique d'*ArcticNet*.

Le *Polar Data Catalogue* vise, à terme, à favoriser la production de nouvelles connaissances en offrant, d'une part l'accès à des données jusque là inaccessibles et d'autre part, la possibilité de mener des recherches croisées sur des jeux de données provenant de disciplines différentes ou portant sur des couvertures spatiales et temporelles jusque-là jamais comparées. À titre d'exemple, l'infrastructure permettra aux chercheurs intéressés d'analyser des données sur l'évolution du permafrost dans une région précise (données géophysiques) en lien avec les changements dans la disponibilité de certaines ressources alimentaires (données biologiques) qui pourraient affecter l'état de santé (données médicales) et le mode de vie (données sociologiques) de la population autochtone locale.

Trois ans après son lancement, un peu plus de 50% des membres du collectif ont contribué à la base de données en y publiant leurs métadonnées, c'est-à-dire des informations qui décrivent leurs données. Les données de recherche en tant que telles représentent 25% du contenu, contre 75% pour les métadonnées<sup>7</sup>. Le projet vise d'abord la publication des métadonnées pour ultérieurement aboutir à une base de données complète, et cela dans le cadre d'une stratégie de diffusion progressive de la base au sein du collectif – la publication des métadonnées requérant a priori moins de travail de la part des membres du collectif. À la manière d'un catalogue de références de bibliothécaire, qui permet de repérer une liste de titres à partir de mots-clés puis de les consulter grâce à leur cote, un catalogue de

6. *ArcticNet Data Policy* (18 janvier 2008) : <http://www.polardata.ca/pdc/public/data-policy.pdf>.

7. Document interne.

métadonnées permet de repérer des jeux de données existants et de trouver l'information permettant de se les procurer (par exemple en fournissant les noms et coordonnées des chercheurs à contacter, des informations sur la couverture spatiale ou temporelle des données, etc.)

Les chercheurs du collectif sont vivement encouragés à alimenter la base de données. La direction scientifique d'*ArcticNet* multiplie les présentations et conférences à ce sujet (par exemple lors du dernier congrès annuel d'*ArcticNet*, où l'un des responsables du projet a présenté le *Polar Data Catalogue* en séance plénière à plus de 400 personnes<sup>8</sup>) et met en place des mesures incitatives auprès des chercheurs (par exemple, l'obligation d'indiquer les métadonnées et données publiées dans les rapports annuels et demandes de renouvellement de fonds<sup>9</sup>). En outre, le collectif s'est doté depuis 2008 d'un cadre réglementaire en matière de gestion de ses données de recherche, qui recommande – sans toutefois l'exiger – leur publication dans un délai de trois ans après leur production<sup>10</sup>.

Dans le cadre de notre étude ethnographique, nous nous sommes intéressées à l'activité pratique de publication des métadonnées par les chercheurs. En pratique, ces derniers doivent remplir un formulaire en ligne dans lequel ils sont invités à fournir une série d'informations : nom et adresse, résumé de la recherche, mots-clés, dates de collecte, zone géographique couverte (latitude, longitude), unités de mesure utilisées, méthodes d'échantillonnage, instruments, etc. Cette tâche, en apparence technique et fastidieuse, s'est avérée problématique pour la plupart des chercheurs et a cristallisé un certain nombre de tensions. Contrairement à ce que nous aurions pu penser, les enjeux liés au respect de la confidentialité des sources ou les craintes associées au risque de récupération frauduleuse des données de recherche n'ont pas été les sources principales de résistance des chercheurs. En revanche, le risque que les données publiées soient mal réutilisées parce que mal interprétées (et donc mal documentées), de même que la lourdeur de la tâche en termes de temps et d'expertise requis pour décrire adéquatement les jeux de données ont été les principaux handicaps à l'usage de la base. Décrire ses données de recherche dans un langage compréhensible par d'autres collègues que ceux appartenant à son

---

8. Notes d'observation, 8-11 décembre 2009, *ArcticNet All Scientists Meeting*, Victoria, Canada.

9. Document interne.

10. Ce délai de trois ans correspond à ceux exigés par les organismes de financement canadiens et américains en sciences de la nature. À l'instar des politiques de ces mêmes organismes, il s'agit pour le moment d'une recommandation plus que d'une exigence, d'où l'absence de mesures punitives en cas de non respect du délai.

champ disciplinaire, préciser les méthodes utilisées de façon à limiter les risques d'interprétation erronée ou encore, plus simplement, accepter de consacrer du temps à cette activité sont quelques unes des difficultés majeures que nous avons repérées chez les chercheurs interviewés. En premier lieu, ces difficultés sont apparues être directement liées à l'existence de cultures épistémiques différentes au sein du collectif.

## ■ La confrontation des cultures épistémiques

L'usage du *Polar Data Catalogue* au sein du collectif *ArcticNet* s'est heurté aux différentes cultures épistémiques qui le composent. Il existe, en matière de gestion et de partage des données scientifiques, des traditions (ou des « différences culturelles » pour reprendre les mots des scientifiques interviewés) qui agissent comme des normes et conventions et qui gouvernent en grande partie les représentations et des comportements des chercheurs. Les usages de la nouvelle base sont apparus directement dépendants de la nature des données de recherche et des rapports différenciés que les chercheurs entretiennent vis-à-vis de leurs données de recherche, deux éléments intrinsèquement liés aux traditions épistémiques dans lesquelles s'inscrivent les membres du collectif *ArcticNet*.

Les travaux en histoire et sociologie des sciences montrent comment la façon dont on « connaît », c'est-à-dire les manières d'appréhender les phénomènes et les objets de connaissance, sont fortement liées à l'existence de traditions épistémologiques propres. Malgré la prégnance du mythe selon lequel il existerait une méthode scientifique unique, les sciences apparaissent en réalité beaucoup plus diverses dans leurs approches et méthodologies (Galison et Stump, 1996). La façon dont on connaît est intimement liée à la nature des instruments utilisés (Clarke et Fujimura, 1992), au sens et à l'importance réservés aux données de recherche, à la façon dont les communautés se structurent, aux moyens de communication qu'elles utilisent, aux savoirs implicites qu'elles partagent, etc., bref aux « cultures épistémiques » qui les caractérisent (Knorr-Cetina, 1999). Les manières dont on archive ou partage des données de recherche peuvent varier considérablement. Ainsi, les champs disciplinaires où la collaboration scientifique à grande échelle autour d'instruments complexes a une longue histoire ont des cultures de partage de données bien en place. C'est le cas de la physique des hautes énergies étudiée par Knorr-Cetina (1999), où les équipes de recherche peuvent rassembler plusieurs dizaines de chercheurs



autour de gigantesques instruments (les accélérateurs de particules) et où les données « appartiennent » à tous. À l’opposé, les disciplines où les pratiques sont plus dispersées et artisanales – comme l’écologie – se caractérisent généralement par des pratiques spécifiques marquées par les façons de faire, normes et conventions propres aux équipes et aux laboratoires, voire propres aux styles personnels des chercheurs (Borgman *et al.*, 2007).

### ***Des rapports différenciés aux données de recherche***

D’emblée, les questions de confidentialité surviennent pour les données en sciences sociales et en sciences de la santé, alors qu’elles sont généralement de moindre importance en sciences de la nature (hormis pour les projets qui portent sur des populations animales ou végétales protégées ou à risque). Comme l’explique cette chercheuse en nutrition, il est délicat de publier des données sur les populations humaines dès lors qu’elles incluent des informations personnelles :

*« Nous, nos données ce sont des données sur la population inuite, qui concernent la santé, des indicateurs socioéconomiques (...) on a aussi des données personnelles (...) sur les questions de violence conjugale, consommation de drogues, abus sexuels – vous savez que c’est une préoccupation centrale pour les communautés nordiques n’est-ce pas – donc nous devons prendre plusieurs précautions par rapport à ces données avant de pouvoir les publier. » (20-Elizabeth-09-12-2009)<sup>11</sup>*

Comment assurer le respect de la confidentialité des sources tout en publiant des données suffisamment riches et précises pour qu’elles puissent faire l’objet de nouvelles recherches et analyses ? Les défis sont de taille et les solutions restent à venir. À titre d’exemple, dans le cas de données sur des populations inuites, le fait d’avoir supprimé toute indication géographique pour préserver l’anonymat des participants à l’enquête a rendu impossible toute nouvelle analyse comparative en fonction de la localisation spatiale<sup>12</sup>. À l’opposé, la publication de la plupart des données en sciences de la nature se fait sans problème, en particulier lorsqu’il s’agit de données physiques. Il s’agit alors de s’assurer de leur qualité en ôtant les artefacts, anomalies, erreurs de mesure, etc. Le travail de préparation des données, préalable à leur publication en ligne, n’exigera pas le même

11. Afin de respecter la confidentialité des propos des personnes interviewées, les noms indiqués sont fictifs et sont utilisés uniquement à des fins de repérage des extraits cités dans le corpus de verbatim.

12. Notes d’observation (décembre 2009) et entretien n°20.

effort ni ne soulèvera les mêmes enjeux, par comparaison aux données en sciences sociales ou de la santé.

Si la nature même des données peut faciliter ou, au contraire, limiter leur publication, l'importance de l'implication personnelle du chercheur ou de la chercheuse dans leur production apparaît être un élément encore plus déterminant. L'engagement personnel des anthropologues dans les enquêtes ethnographiques est généralement associé à un fort sentiment de propriété, voire d'identification, vis-à-vis du matériel produit. Le problème de l'anthropologue n'est-il pas, comme le dit Ricœur (1969), la compréhension de soi-même par le détour de la compréhension de l'autre ? Cependant, la proximité du chercheur par rapport à ses données de recherche et la réticence à leur publication ne sont pas l'apanage des chercheurs en sciences humaines et sociales. Il est fréquent que les chercheurs en sciences de la nature souhaitent garder le droit de propriété sur leurs données dès lors que leur recueil est laborieux ou exige un certain niveau de savoir-faire (Borgman *et al.*, 2007), du moins pendant le temps nécessaire à leur analyse, comme l'explique ce géographe :

*« Je ne pense pas que ce soit juste en fait que quelqu'un soit obligé de rendre publiques ses données après, disons, 12 mois ou peut-être même après 24 mois. Je pense que, à un moment donné, peut-être que ça doit être 3 ans, peut-être 5, ok il pourrait y avoir une forme d'obligation de publication des données si on n'a pas pu avancer dessus. Mais il faut prendre en compte le fait que la personne a fait l'effort d'écrire le projet pour la demande de subvention (...) c'est du travail ! Et quoi, après, devoir renoncer à toutes ces données, ce n'est pas correct selon moi. »*  
(24-Rob-10-12-2009)

Autre exemple : certains chercheurs en biologie végétale au sein du collectif travaillent sur de très petites quantités de données, généralement sur des zones géographiques très spécifiques, ce qui implique de longues heures de présence sur le terrain et un travail substantiel de préparation pour la mise au point du processus de collecte de données et pour le repérage des sites de collecte. Compte tenu de l'ampleur de la tâche et de l'importance de l'implication personnelle du chercheur dans le processus, on comprend que ceux-ci hésitent à publier leurs données dans une méga banque de données anonyme. À l'opposé, en océanographie physique par exemple, il est commun de travailler sur des données recueillies par d'autres, dans la mesure où la collecte, fortement instrumentée, peut s'étendre sur plusieurs années et être distribuée entre plusieurs équipes

de collaborateurs. Aussi les étudiants travaillent-ils généralement sur des données recueillies par d'autres :

« Il y a une grosse, grosse différence entre la physique pis les autres disciplines, c'est que, un étudiant en biologie, son directeur va l'envoyer à bord du bateau, il va prendre ses données pis il va faire sa thèse sur ses données. Moi, neuf fois sur dix, mes étudiants en physique vont faire leur thèse sur des données qui ont été prises un an avant ou deux ans avant, parce que je veux pas qu'ils passent deux ans à attendre leurs données. Tu sais, si quelqu'un est en maîtrise pis qu'il attend deux ans avant que ses données soient prêtes, ça n'a pas de sens (...). Donc en général, ceux qui travaillent sur des données, à quelques exceptions près, comme en turbulence, ne travaillent pas sur des données qu'ils ont prises eux-mêmes. » (5-Pat-22-10-2009)

Dans la mesure où le « script » (Akrich, 1992) du *Polar Data Catalogue* propose un cadre d'usages et de pratiques unifié, on comprend qu'il vienne heurter les différentes traditions épistémiques au sein du collectif, et donc que le développement de nouvelles pratiques instrumentées de partage de données devienne un objet de conflit. Le succès de la nouvelle base repose sur l'adoption, par l'ensemble des chercheurs du collectif, de ce que Fujimura appelle un « *package* standardisé » (*standardized packages*) (Fujimura, 1988, 1992), ici matérialisé, en premier lieu, dans le fameux formulaire de métadonnées. Si l'on admet que le *Polar Data Catalogue*, en tant qu'infrastructure de partage, s'apparente à un « objet-frontière » (Star et Griesemer, 1989 ; Star, 2010), au sens où il constitue un instrument par lequel différents groupes s'alignent et se coordonnent, le formulaire de métadonnées – qui prescrit un ensemble de pratiques (en matière de description de données) ainsi qu'une certaine vision du partage de données (en accès libre et ouvert) – promeut l'adoption de certaines manières de faire et de penser en matière de publication de données, non seulement au sein du collectif *ArcticNet*, mais aussi pour l'ensemble des chercheurs qui travaillent dans le domaine de la recherche polaire. Le développement de « *packages* standardisés » autour d'objets-frontière aide à délimiter et à réduire l'éventail des actions possibles (Fujimura, 1992). Une fois développés et incorporés dans des routines, ces ensembles de méthodes et de techniques peuvent soutenir l'action collective. Dans le cas qui nous concerne, le formulaire de métadonnées vise précisément à standardiser les formats de publication des données et, par là, à en faciliter le partage au sein de la communauté scientifique.

L'analyse des versions successives du formulaire a révélé l'existence de manières différentes de « parler » des données au sein du collectif. À titre d'exemple, dans la version 1, il était demandé d'indiquer la « couverture temporelle » des données, ce qui avait suscité la réaction d'un anthropologue en particulier qui avait signalé que « le critère n'était pas adéquat » pour décrire ses données d'entretiens : s'agit-il de la période durant laquelle les entretiens ont été réalisés (en l'occurrence, deux ans) ou de la période temporelle couverte par les données elles-mêmes (en l'occurrence, plusieurs générations, dans la mesure où le contenu des entretiens portait sur l'héritage culturel autochtone) ?<sup>13</sup>

*« Nous, on fait pas la différence entre les deux. Non, parce qu'en sciences naturelles, la date de collecte, c'est aussi la période temporelle couverte par les données. (...) On a résolu le problème en disant : Mettez x comme date de début mais dans le résumé, spécifiez que les données, les entrevues ont été faites en août 2009 par exemple. » (19-Georgia-15-02-2010)*

Comme l'explique cette membre de l'équipe de développement et de maintenance du *Polar Data Catalogue*, la distinction entre la date de collecte et la période temporelle couverte n'a pas lieu d'être en sciences naturelles (en l'occurrence le domaine de provenance des développeurs de la base de données). Le formulaire a donc dû être corrigé pour accommoder l'ensemble des chercheurs.

### ***La configuration organisationnelle du collectif***

Au-delà des cultures épistémiques, les contextes organisationnels des collectifs sont à prendre en compte dans la problématique du partage de données. Dans le cas d'*ArcticNet*, d'une part la collaboration interdisciplinaire est à la base du programme de recherche, d'autre part une grande partie du travail se fait en mer, dans le cadre de missions scientifiques où, littéralement, tous les chercheurs sont sur le même bateau.

*« Je suis un océanographe physicien (...) je fais partie d'un projet de biologie sur les points chauds, les hot spots qu'ils appellent, où on regarde comment la physique va impacter la biologie. (...) [Sur le bateau] les biologistes [sont] derrière nos épaules pis ils nous disent "Oh, moi j'aimerais que tu me ramènes un échantillon de 100 mètres ou de 200 mètres ou de 400 mètres" (...) Donc on ramène des échantillons d'eau*

13. Notes d'observation (septembre 2009) et entretien n°19.

à partir desquels les biologistes et les chimistes font leurs analyses. »  
(5-Pat-22-10-2009)

En pratique, comme l'illustre l'extrait ci-dessus, le recueil des données requiert une grande collaboration entre les chercheurs, qui doivent s'entendre sur les protocoles de collecte, partager des instruments, etc., tout en poursuivant leurs propres objectifs de recherche (qu'il s'agisse de répertorier une espèce, mesurer les courants marins, prélever une carotte de glace, ou encore interviewer des représentants des populations inuites sur les changements dans leurs pratiques de chasse ou dans leurs habitudes alimentaires.) Les océanologues physiciens, en particulier, jouent un rôle central dans le processus de collecte des données, dans la mesure où ils sont responsables de la production de certaines données clefs, nécessaires aux autres équipes de recherche.

« [Ceux qui réalisent] des recherches sur les ours polaires ou les phoques, eux, ils ont besoin de la physique, ils ont besoin des propriétés physiques de l'eau de mer, ils ont besoin des courants, ils ont besoin de tout ça donc (...) même si mes étudiants travaillent sur des projets en physique, ils sont un peu les océanographes de service, physiciens de service, pour fournir des données aux autres disciplines qui en ont absolument besoin. (...) les données CTD il suffit que tu nous demandes et on les donne, les données de physique. » (5-Pat-22-10-2009)

Les physiciens se démarquent par une culture de partage déjà bien en place, mais leur statut de fournisseurs de données est, dans le cas d'*ArcticNet*, institutionnalisé, dans la mesure où ils sont financés par le programme pour collecter les données « noyau »<sup>14</sup> et les rendre disponibles à tous.

Si le développement de la base de données est venu heurter les différentes cultures épistémiques du collectif, il s'est aussi accompagné par un renforcement des rôles et des formes de division du travail au sein des équipes de recherche.

---

14. Il s'agit de données de base sur l'environnement physique dans lequel sont effectués l'ensemble des prélèvements. Appelées données « noyau » (*core data*), ce sont des données sur les propriétés physiques de l'eau (température, salinité, etc.) qui sont généralement collectées en continu, grâce à un instrument dédié. Il s'agit, en quelque sorte, de données contextuelles, dont l'équivalent en sciences sociales seraient, par exemple, les données socioéconomiques d'une population (statistiques sur les moyennes d'âge, de revenu, etc.) au sein de laquelle une enquête sociologique serait conduite.

## ■ Le renforcement des formes de division du travail

Dans son ethnographie d'une base de données sur le séquençage du génome de la souris, Hine (2006) montre comment le développement de la ressource participe de la structuration du laboratoire en y faisant émerger un nouvel ordre social, où les nouvelles responsabilités des chercheurs, en l'occurrence le dépôt systématique des échantillons d'ADN, suscitent certaines tensions sociales. Dans le cas du *Polar Data Catalogue*, comment ces nouvelles pratiques de publication et de partage de données se traduisent-elles sur le plan des rôles et des formes de division du travail au sein des équipes de recherche ?

Nos observations nous conduisent à penser que les nouvelles pratiques associées à l'usage du *Polar Data Catalogue* s'inscrivent en continuité, voire viennent renforcer, les formes de division du travail préexistantes, et cela, quelle que soit la discipline. En l'occurrence, ces pratiques sont prises en charge par les étudiants, techniciens et personnels de recherche, autrement dit par ceux qui assurent généralement la plus grande partie du travail de manipulation des données. D'une part la nature des tâches impliquées est perçue comme étant purement technique et de second ordre – et donc logiquement confiée aux personnes qui occupent les fonctions et statuts les moins importants (Shapin, 1989) ; d'autre part l'ensemble des activités liées au traitement des données, depuis leur collecte jusqu'à leur analyse, est généralement considéré comme une tâche technique qui relève de la vaste catégorie du travail « invisible » – au sens où il a été défini dans les travaux sur l'organisation sociale du travail (Star et Strauss, 1999 ; Orr, 1996 ; Suchman, 1995) et dans les travaux sur le travail technique dans les sciences (Shapin, 1989 ; Cambrosio et Keating, 1988 ; Lynch, 1982, 1985 ; Collins, 1974 ; Barley et Bechky, 1994). En sciences expérimentales, le travail de manipulation des données incombe généralement au personnel de recherche ou aux étudiants. C'est d'ailleurs une étape fondamentale de la formation des étudiants en écologie (Roth et Bowen, 2001). C'est l'occasion, pour ces derniers, de « se faire la main » en se frottant aux aspects les plus tangibles (et souvent les plus contraignants) du processus de recherche. C'est aussi, dans certains cas, de longues tâches techniques, laborieuses, complexes ou encore répétitives, pour lesquelles un certain savoir-faire est indispensable – d'où leur prise en charge par des personnels spécialisés. La publication des métadonnées et données revient ainsi presque naturellement aux techniciens et aux étudiants :

*« En fait, ceux qui entrent les métadonnées, c'est toujours les mêmes, c'est les étudiants. En fait, c'est soit un étudiant soit le technicien responsable du laboratoire (...) Ce sont les étudiants qui manipulent les données ou les techniciens, plus que les chercheurs. Les chercheurs, c'est rare qu'ils manipulent les données brutes, ils supervisent l'écriture, ils supervisent les étudiants. » (19-Georgia-15-02-2010)*

C'est sur le tas et par essais et erreurs que les techniciens ou les étudiants découvrent comment décrire adéquatement les jeux de données. Même si tous s'accordent pour dire que la tâche reste relativement simple, elle exige d'y consacrer du temps et, surtout, de « se poser les bonnes questions » :

*« On n'a pas l'automatisme, et puis à chaque fois, il faut réfléchir à la façon dont on fait ça. (...) c'est pas si compliqué, ça prend pas une semaine, mais... » (2-Edouard-21-10-2009)*

En pratique, la plupart des entrées de données doivent être revues et, le plus souvent, corrigées – signe que la tâche est en réalité loin d'être aussi triviale qu'elle n'en a l'air :

*« Par exemple, y'en a une, récemment, je lui ai renvoyé ses métadonnées, j'ai dit : "Écoute, c'est parce qu'on comprend pas les..." (je ne sais plus). Parce qu'elle avait mis tous les sites en abrégés pis tout ça alors je lui ai dit : "On peut pas vraiment connaître tous tes sites, on veut savoir c'est quoi : qu'est-ce que t'as échantillonné comme données ?" Pis là, elle l'a modifié par rapport à ça. Donc faut que je regarde... » (19-Georgia-15-02-2010)*

Publier des données n'est pas une opération banale, mais une activité qui requiert du temps et une certaine expertise. Il faut, d'une part s'assurer qu'elles soient de bonne qualité et d'autre part, les décrire avec précision de façon à ce qu'elles soient interprétées correctement, et donc réutilisées de manière appropriée. La qualité de la documentation (c'est-à-dire des métadonnées) est cruciale dans le processus. En outre, la responsabilité du chercheur ou du laboratoire, par exemple en cas de réutilisation erronée des données due à une documentation déficiente ou incomplète, est en jeu. L'histoire du *Polar Data Catalogue* est encore trop récente (et le nombre de jeux de données disponibles trop faible) pour que de tels cas se soient déjà présentés. Cependant, les exemples de réutilisation inappropriée ne manquent pas ; c'est d'ailleurs l'une des principales sources de résistance à

la publication des données de recherche dans les sciences (Costello, 2009). Malgré l'importance de cette préoccupation, la question de la formation des chercheurs ou du personnel de recherche (incluant les techniciens) à la production des documentations des données n'est pas à l'ordre du jour. La publication des données s'ajoute à la liste des multiples tâches confiées au personnel de recherche, et rares sont les communautés de recherche qui ont les moyens de dédier des ressources spécifiques à cette fonction (par exemple en finançant des postes de gestionnaires de données<sup>15</sup>). Plus largement, on pointe du doigt l'absence de mécanismes incitatifs pour le partage de données.

« Ils [les chercheurs] sont pas contre l'idée. C'est juste que pour eux, c'est une charge de travail supplémentaire qui s'ajoute aux demandes de permis, aux évaluations environnementales, aux rapports, etc. Je pense que c'est un peu ça qui bloque. Ça nous amène à nous dire "comment est-ce qu'on récompense la publication des données ?" »  
(19-Georgia-15-02-2010)

L'absence de mécanismes de rétribution pour le partage de données en sciences de la nature en général, et dans le cas du *Polar Data Catalogue* en particulier, constitue sans aucun doute le principal obstacle au développement de bases de données partagées. Les bénéfices à court terme ne sont pas évidents pour tous les chercheurs, et la question de savoir quels seraient les meilleurs incitations (citation, rémunération, etc.) est actuellement en discussion.

## ■ Une réorientation du regard des scientifiques

La généralisation de l'usage du *Polar Data Catalogue* au sein du collectif implique l'inscription d'une nouvelle routine organisationnelle dans le quotidien du travail scientifique. Weick *et al.* (2005) envisagent la mise en place de nouvelles routines comme des occasions de construction de sens. En d'autres termes, l'adoption d'une nouvelle routine, en tant qu'ensemble appris et répété d'arrangements pratiques et matériels, peut participer de la construction d'une nouvelle manière d'appréhender les choses. Ribes et Bowker (2009) ont observé un changement de perspective chez des chercheurs en sciences de la terre parallèlement au développement d'un projet

15. Sur la fonction de gestionnaire de données, voir Karasti et Baker (2004).



de cyberinfrastructure, en l'occurrence un mode d'appréhension des données et des connaissances à travers le prisme des sciences de l'information, voire des sciences informatiques, et non plus seulement à travers celui de leurs disciplines respectives. Dans quelle mesure de nouvelles pratiques de publication et de partage de données peuvent-elles participer de l'évolution du regard des scientifiques sur les données de recherche ? À la lumière de nos premières observations, nous discutons la possibilité d'une réorientation du regard des chercheurs sur les données dans deux directions. Premièrement, l'usage de la base passant par la création de représentations codifiées des données, les chercheurs sont amenés à les considérer non plus seulement comme des données « scientifiques » mais aussi comme des données « informatiques », c'est-à-dire présentant un contenu informationnel qu'il s'agit de décrire et de structurer. Deuxièmement, ils sont amenés à considérer leurs données de recherche en tant que « bien public », et non plus en tant que propriété individuelle, du laboratoire ou de l'institution.

### ***Des représentations codifiées : données, métadonnées...***

Publier ses données et métadonnées dans une base peut paraître trivial à première vue. Après tout, il s'agit simplement de remplir un formulaire en ligne. Mais ce serait sous estimer les dimensions politiques des artefacts (Winner, 1985) et des systèmes de classification (Bowker et Star, 1999) d'une part, et la force contraignante des représentations d'autre part. Habités au modèle canonique de la publication d'articles, les scientifiques sont tentés de reproduire le format textuel de l'article lorsqu'il s'agit de décrire leurs données :

*« Le problème souvent, c'est qu'ils voient les métadonnées comme un article scientifique. Souvent, faut que je réécrive pour dire : "C'est quoi les unités, c'est quoi les variables échantillonnées. C'est pas le résultat de ta recherche qui m'intéresse, c'est de savoir c'est quoi que t'as comme données..." Y'a une étudiante entre autres, à qui j'ai écrit, je lui a dit : "Écoute, je vais te renvoyer ta métadonnée pis j'aimerais ça que tu me mettes les spécifications de ton échosondeur" »  
(19-Georgia-15-02-2010)*

Savoir comment décrire adéquatement des données de recherche n'est pas fréquent chez les chercheurs du collectif. Ceux-ci doivent apprendre à les formater d'une certaine manière, en fonction d'une certaine norme et, en premier lieu, à distinguer les métadonnées des données, ce qui exige un

certain apprentissage.

« – *Et, donc vous seriez plutôt favorable à publier vos métadonnées ?*  
– *Hum, mes... mes métadonnées ?? qu'est-ce que ça pourrait bien être ?*  
*Mes sites d'échantillonnage, je suppose ?* » (24-Rob-10-12-2009)

À l'instar de ce chercheur qui s'interroge sur ce que pourraient être ses métadonnées, les chercheurs du collectif sont amenés à se poser de nouvelles questions sur leurs données. Ce faisant, ils sont conduits à les considérer non plus seulement comme des données scientifiques mais aussi, comme des « informations » (au sens de la théorie de l'information), c'est-à-dire comme des ensembles de données présentant un contenu informationnel qu'il s'agit de décrire et de structurer sous forme de représentations codifiées (en l'occurrence sous la forme d'une liste de métadonnées). Les chercheurs doivent donc agir et réfléchir à leurs données de recherche du point de vue de leurs caractéristiques à la fois scientifiques et informationnelles et, par là, envisager le domaine scientifique du collectif du point de vue des disciplines, catégories, terminologies, nomenclatures, etc., qui y sont représentées. Le simple fait de remplir le formulaire conduit à se poser des questions du type : « *Cette unité de mesure est-elle "standard" ?* », « *Puis-je dire que j'ai utilisé cette méthode d'échantillonnage même si c'est un raccourci propre au labo, donc pas tout à fait comme on devrait le faire mais qui "marche" quand même ?* »<sup>16</sup>. L'usage de la base conduit les chercheurs à s'interroger sur ce qu'ils prennent généralement pour acquis et qui reste invisible dans les publications, c'est-à-dire les propriétés des données de recherche.

### ***Les données scientifiques comme « bien public »***

Le changement de perspective sans doute le plus important chez les scientifiques a trait au changement du statut des données, considérés jusque là comme la propriété individuelle du chercheur (la donnée étant du matériel de recherche propre à chaque chercheur voire à chaque laboratoire), vers un statut de ressource commune pour la communauté ou le domaine scientifique, voire le grand public. Même si cette vision des données comme un « bien commun » (*data commons*) fait l'objet d'un consensus grandissant au sein de la communauté scientifique – rappelons que le partage de données est considéré dans la tradition mertonienne comme le sceau de la science moderne (Merton, 1968), – l'absence de mesures de

16. Notes d'observation (décembre 2009).

récompense et d'encadrement institutionnel des pratiques de partage ne peut que conforter les réticences individuelles des chercheurs. Cela étant dit, le projet du *Polar Data Catalogue* prend appui sur la structure particulière du collectif, une structure en réseau, fondée sur la collaboration scientifique pluridisciplinaire. En cela, il refléterait l'esprit même du réseau :

*« Le Polar Data Catalogue fait partie de l'esprit d'ArcticNet. Donc je pense que c'est les réseaux comme ArcticNet qui vont changer la façon de faire la recherche. Et le Polar Data Catalogue fait partie de cette nouvelle mentalité là je dirais, d'être multidisciplinaire pis de partager ces informations là pour être justement capable de résoudre des problèmes concrets. »* (19-Georgia-15-02-2010)

En ce sens, les pratiques de partage de données sont fort probablement appelées à se développer, que ce soit au sein du réseau à travers le projet *Polar Data Catalogue* ou à travers d'autres grandes bases de données nationales ou internationales. Plus largement, l'actualité de la question du partage de données est directement liée à la mise en place de projets de recherche à grande échelle, qui mobilisent de larges collectifs et qui traversent les frontières disciplinaires.

*« Le partage de données c'est important, ces très grands projets, vous savez, bon par exemple l'Admunsten [le navire scientifique du réseau ArcticNet] c'est un gros navire, c'est beaucoup d'argent, et il ne s'agit pas d'une personne qui travaille sur une question scientifique, c'est un grand nombre de scientifiques qui travaillent. Et je pense que de rendre ces données disponibles c'est une bonne décision. Il y a trop d'argent en jeu pour une seule personne. Je pense que c'est une question d'échelle, oui. »* (24-Rob-10-12-2009)

Comme l'explique ce chercheur, la portée et l'envergure du programme scientifique d'*ArcticNet* renvoient à des enjeux scientifiques et financiers qui militent pour la mise en commun des données de recherche. Encore faut-il rappeler que la raison d'être du collectif repose précisément sur l'interdisciplinarité et sur la collaboration entre les milieux académiques, gouvernementaux et industriels pour s'attaquer au défi scientifique que représente l'étude du changement climatique.

## ■ Conclusion

La question du partage des données scientifiques, telle qu'elle est formulée dans le contexte actuel de projets de grandes bases de données, conduit les chercheurs à se poser de nouvelles questions sur la nature de leurs données et sur les possibilités de leur réutilisation par d'autres. Ce faisant, on observe le début d'un processus de réorientation du regard des chercheurs sur leurs données de recherche, qu'ils appréhendent non plus seulement comme du matériel de recherche propre à chacun, mais aussi comme un « bien commun » qu'il s'agit de partager ouvertement au sein de la communauté scientifique. Ce changement de perspective est en cours. Il se manifeste, certes à travers l'expression de résistances individuelles chez les scientifiques, mais aussi et surtout à travers la confrontation des cultures épistémiques que les nouvelles pratiques viennent heurter.

Le partage de données reste un objet de recherche relativement peu étudié par les sociologues des sciences. Pourtant, les développements actuels en matière d'infrastructures de collaboration scientifique viennent jeter un nouvel éclairage sur ces enjeux et, plus largement, sur le rôle des grandes bases de données dans la production de la recherche. Une des pistes qui nous apparaît parmi les plus intéressantes est la question du rôle « performatif » des grandes bases de données scientifiques. Dans la lignée des réflexions de Star et Bowker (Bowker, 2000 ; Bowker et Star, 1999), il s'agit d'interroger les dimensions éthiques et politiques des dispositifs et d'aborder les modalités pratiques des processus de représentation (Latour, 1987). Rendre les données mobiles n'est pas seulement une question technique (choisir entre une base de données ou une autre, entre un formulaire de métadonnées ou un autre), c'est aussi et surtout une question scientifique et politique, qui implique des négociations sur les catégories et terminologies à privilégier dans les formats d'inscription et de représentation des données (Millerand et Bowker, 2008, 2009). Quelles sont les données, et donc les perspectives disciplinaires, qui sont effectivement mises en réseau dans ces grandes bases de données ? Quelles sont celles et ceux qui restent, au contraire, isolées ou peu visibles ?

## ■■■ références

**Akrich M.**, 1992. The De-scription of Technical Objects, in W. E. Bijker, J. Law (eds.), *Shaping technology/Building Society. Studies in Sociotechnical Change*, Cambridge MA, MIT Press, 259-263.

**Arzberger P. W., Schroeder P., Beaulieu A., Bowker G. C., Casey K., Laaksonen L., Moorman D., Uhlir P., Wouters P.**, 2004a. Promoting Access to Public Research Data for Scientific, Economic, and Social Development, *Data Science Journal*, 3, 135-152.

**Arzberger P. W., Schroeder P., Beaulieu A., Bowker G. C., Casey K., Laaksonen L., Moorman D., Uhlir P., Wouters P.**, 2004b. An International Framework to Promote Access to Data, *Science*, 303, 1777-1778.

**Atkinson-Grosjean J.**, 2006. *Public Science, Private Interests: Culture and Commerce in Canada's Networks of Centres of Excellence*, Toronto, University of Toronto Press.

**Barley S. R., Bechky B. A.**, 1994. In the Backrooms of Science: The Work of Technicians in Science Labs, *Work and Occupations: An International Sociological Journal*, 21(1), 85-126.

**Borgman C. L., Wallis J. C., Enyedy N.**, 2007. Little Science Confronts the Data Deluge: Habitat Ecology, Embedded Sensor Networks, and Digital Libraries, *International Journal on Digital Libraries*, 7, 17-30.

**Bowker G. C., Star S. L.**, 1999. *Sorting Things Out: Classification and its Consequences*, Cambridge MA, MIT Press.

**Bowker G. C., Baker K. S., Millerand F., Ribes D.**, 2010. Towards Information Infrastructure Studies: Ways of Knowing in a Networked Environment, in **J. Hunsinger, L. Klastrup, J. M. Allen** (eds.), *International Handbook of Internet Research*, Londres, Springer, 97-118.

**Brown C.**, 2003. The Changing Face of Scientific Discourse: Analysis of Genomic and Proteomic Database Usage and Acceptance, *Journal of the American Society for Information Science and Technology*, 54(10), 926-938.

**Cambrosio A., Keating P.**, 1988. "Going Monoclonal": Art, Science, and Magic in the Day-to-Day Use of Hybridoma Technology, *Social Problems*, 35(3), 244-260.

**Clarke A. E., Fujimura J. H.** (eds.), 1992. *The Right Tools For the Jobs: At Work in Twentieth-Century Life Science*, Princeton, Princeton University Press.

**Collins H. M.**, 1974. The TEA Set: Tacit Knowledge and Scientific Networks, *Science Studies*, 4, 165-186.

**Costello M. J.**, 2009. Motivating Online Publication of Data, *BioScience*, 59(5), 418-427.

- Edwards P. N., Jackson, S. J., Bowker G. C., Williams R.**, 2009. Introduction: An Agenda for Infrastructure Studies, *Journal of the Association for Information Systems*, 10(5), 364-374.
- Edwards P. N., Jackson S. J., Bowker G. C., Knobel C. P.**, 2007. *Understanding Infrastructure: Dynamics, Tensions, and Design*, National Science Foundation, Office of Cyberinfrastructure.
- Fisher D., Atkinson-Grosjean J., House D.**, 2001. Changes in Academy/Industry/State Relations in Canada: The Creation and Development of the Networks of Centres of Excellence, *Minerva*, 39, 299-325.
- Fujimura J. H.**, 1988. The Molecular Biological Bandwagon in Cancer Research: Where Social Worlds Meet, *Social Problems*, 35(3), 261-283.
- Fujimura J. H.**, 1992. Crafting Science: Standardized Packages, Boundary Objects, and 'Translation', in A. Pickering (ed.), *Science as Practice and Culture*, Chicago, University of Chicago Press, 168-211.
- Galison P., Stump D. J.**, 1996. *The Disunity of Science: Boundaries, Contexts, and Power*, Stanford, Stanford University Press.
- Gilbert W.**, 1991. Towards a Paradigm Shift in Biology, *Nature*, 349(6305), 99.
- Hilgartner S.**, 1995. Biomolecular Databases: New Communication Regimes for Biology? *Science Communication*, 17(2), 240-263.
- Hine C.**, 2006. Databases as Scientific Instruments and their Role in the Ordering of Scientific Work, *Social Studies of Science*, 36(2), 269-298.
- Karasti H., Baker K. S., Millerand F.**, 2010. Infrastructure Time: Long-term Matters in Collaborative Development, *Computer Supported Collaborative Work*, 19(3-4), 377-415.
- Karasti H., Baker K. S.**, 2004. *The Long-Term Information Management Trajectory: Working to Support Data, Science and Technology*, Scripps Institution of Oceanography Technical Report, University of California San Diego.
- Latour B.**, 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*, Cambridge MA, Harvard University Press.
- Lynch M.**, 1985. *Art and Artifact in Laboratory Science : A Study of Shop Work and Shop Talk in a Research Laboratory*, Londres, Routledge & Kegan Paul.
- Lynch M.**, 1982. Technical Work and Critical Inquiry: Investigations in a Scientific Laboratory, *Social Studies of Science*, 12(4), 499-534.
- Merton R. K.**, 1968. *Social Theory and Social Structure*, New York, Free Press.
- Millerand F., Bowker G. C.**, 2009. Metadata Standards. Trajectories and Enactment in the Life of an Ontology, in M. Lampland, S. L. Star (eds.),

- Standards and Their Stories*, New York, Cornell University Press, 149-165.
- Millerand F., Bowker G. C.**, 2008. Metadata, trajectoires et « énonction », in C. Rosental (dir.), *La Cognition au prisme des sciences sociales*, Paris, Éditions des Archives Contemporaines, 277-303.
- Orr J. E.**, 1996. *Talking about Machines: An Ethnography of a Modern Job*, New York, Cornell University Press.
- Ribes D., Finholt T. A.**, 2009. The Long Now of Technology Infrastructure: Articulating Tensions in Development, *Journal of the Association for Information Systems*, 10(5), 375-398.
- Ribes D., Bowker G. C.**, 2009. Between Meaning and Machine: Learning to Represent the Knowledge of Communities, *Information and Organization*, 19(4), 199-217.
- Ricoeur P.**, 1969. *Le Conflit des interprétations. Essais d'herméneutique*, Paris, Le Seuil.
- Roth W.-M., Bowen, G. M.**, 2001. Of Disciplined Minds and Disciplined Bodies: On Becoming an Ecologist, *Qualitative Sociology*, 24(4), 459-481.
- Shapin S.**, 1989. The Invisible Technician, *American Scientist*, 77, 554-563.
- Star S. L., Bowker G. C.**, 2002. How to Infrastructure, in L. A. Lievrouw, S. Livingstone (eds.), *Handbook of New Media. Social Shaping and Consequences of ICTs*, Londres, Sage Publications, 151-162.
- Star S. L., Griesemer J.**, 1989. Institutional Ecology, 'Translations,' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939, *Social Studies of Science*, 19, 387-420.
- Star S. L., Ruhleder K.**, 1996. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces, *Informations Systems Research*, 7(1), 111-134.
- Star S. L., Strauss A.**, 1999. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work, *Computer Supported Cooperative Work*, 8(1-2), 9-30.
- Star S. L.**, 2010. Ceci n'est pas un objet-frontière. Réflexions sur l'origine d'un concept, *Revue d'anthropologie des connaissances*, 4(1), 18-35.
- Suchman L. A.**, 1995. Making Work Visible, *Communications of the ACM*, 38(9), 56-64.
- Weick K. E., Sutcliffe K. M., Obstfeld D.**, 2005. Organizing and the Process of Sensemaking, *Organization Science*, 16(4), 409-421.
- Winner L.**, 1985. Do Artifacts Have Politics? in D. MacKenzie, J. Wajcman (eds.), *The Social Shaping of Technology*, Philadelphia, Open University Press.



**Florence MILLERAND** (millerand.florence@uqam.ca)  
Professeure agrégée  
Département de communication sociale et publique  
Centre interuniversitaire de recherche  
sur la science et la technologie (CIRST)  
Université du Québec à Montréal