

Soutenir la gestion des données de recherche en bibliothèque universitaire : concepts, enjeux et pratiques

Cynthia Lisée, MSI and Édith Robert, MSI

Volume 69, Number 3, July–September 2023

Travailler en milieu documentaire

URI: <https://id.erudit.org/iderudit/1105754ar>

DOI: <https://doi.org/10.7202/1105754ar>

[See table of contents](#)

Publisher(s)

Les Éditions ASTED

ISSN

0315-2340 (print)

2291-8949 (digital)

[Explore this journal](#)

Cite this article

Lisée, C. & Robert, É. (2023). Soutenir la gestion des données de recherche en bibliothèque universitaire : concepts, enjeux et pratiques. *Documentation et bibliothèques*, 69(3), 26–35. <https://doi.org/10.7202/1105754ar>

Article abstract

This article has three objectives. First, we seek to position research data management (RDM) within the broader open science movement. Secondly, we want to demonstrate how the academic librarian could be called upon to be a valuable ally in the scientific ecosystem to better grasp the multiple issues involved in research data management and thus better support faculty and other professionals who gravitate around these practices. Finally, we would also like to highlight some of the particularities induced by MDM for disciplines that make greater use of qualitative methodologies to produce data.

SOUTENIR LA GESTION DES DONNÉES DE RECHERCHE EN BIBLIOTHÈQUE UNIVERSITAIRE: CONCEPTS, ENJEUX ET PRATIQUES

Cynthia LISÉE, MSI

Bibliothécaire en soutien à la recherche à l'UQAM
lisee.cynthia@uqam.ca

Édith ROBERT, MSI

Bibliothécaire à l'École des sciences de la gestion à l'UQAM
robert.edith.2@uqam.ca

RÉSUMÉ | ABSTRACT

Cet article poursuit trois objectifs. Dans un premier temps, nous cherchons à positionner la gestion des données de recherche (GDR) dans le grand mouvement de la science ouverte. Ensuite, nous voulons démontrer comment le bibliothécaire universitaire pourrait être appelé à être un précieux allié dans l'écosystème scientifique afin de mieux saisir les multiples enjeux de la gestion des données de recherche et ainsi mieux soutenir le corps professoral et les autres professionnels qui gravitent autour de ces pratiques. Finalement, nous voulons également mettre en lumière certaines particularités induites par la GDR pour les disciplines qui utilisent plus largement les méthodologies qualitatives afin de produire des données.

This article has three objectives. First, we seek to position research data management (RDM) within the broader open science movement. Secondly, we want to demonstrate how the academic librarian could be called upon to be a valuable ally in the scientific ecosystem to better grasp the multiple issues involved in research data management and thus better support faculty and other professionals who gravitate around these practices. Finally, we would also like to highlight some of the particularities induced by MDM for disciplines that make greater use of qualitative methodologies to produce data.

Introduction

Le mouvement international en faveur de l'ouverture de la science a des retombées dans l'environnement normatif canadien, lesquelles ne sont pas sans conséquences quant à la reconfiguration de nos institutions et dans le changement culturel insufflant de nouvelles pratiques pour la conduite de la recherche. En corollaire, les professionnels des milieux documentaires voient de nouveaux rôles à assumer en participant avec leurs communautés de recherche dans l'appropriation de tous ces nouveaux développements. Investir de nouveaux rôles sans réfléchir aux trames idéologiques qui traversent le discours enthousiaste de la science ouverte et, plus particulièrement dans le secteur des bibliothèques de l'enseignement supérieur, la gestion des données de recherche (GDR), c'est risquer de simplement reproduire des rapports de pouvoir historiquement à l'œuvre. Pour éviter cette instrumentalisation du professionnel de l'information, cet article propose donc de mettre au jour une lecture plus critique de quelques aspects de la

science ouverte pour ensuite mieux positionner le rôle du bibliothécaire en contexte de science ouverte et de GDR

En premier lieu, seront présentés les principes macrosociologiques qui justifient la science ouverte et qui situe le contexte canadien d'émergence de la GDR. La deuxième section sera consacrée à présenter les pratiques actuelles en matière de science ouverte en les contrastant avec ce qui aura été énoncé dans la section précédente sur les vertus de la science ouverte. La troisième section se consacrera à proposer des éléments de réflexion quant au rôle que les bibliothécaires universitaires pourraient être amenés à jouer dans cet écosystème de l'information scientifique pour précisément augmenter le pouvoir d'agir des communautés de recherche qu'ils et elles servent en réponse aux fondements idéologiques traversant leurs interactions avec ces communautés. Puis, nous terminerons en expliquant l'importance de bien connaître certains fondements méthodologiques, notamment qualitatifs, pour être à même de mieux penser la gestion des données de recherche en respectant les

épistémologies disciplinaires et ainsi pouvoir développer des plateformes de dépôts en conséquence. Nous évoquons principalement les préoccupations liées aux médias sociaux et à l'éthique de la recherche, puis nous terminerons en introduisant brièvement les « *Critical data studies* », qui fournissent, à notre avis, un cadre théorique prometteur pour appréhender les enjeux mentionnés.

Écoles de pensée et vertus de la science ouverte

Le discours dominant sur la science continue de s'abreuver à une image noble de la science et le récent mouvement de la science ouverte ne fait pas exception. Par science ouverte, on entend « un savoir transparent et accessible qui est partagée et développé au travers des réseaux de collaboration » (Vicente-Saez et Martinez-Fuentes, 2018, p. 434 [Notre traduction]). Merton (1973) a bien formulé cette conception que la science se pratique selon de hautes valeurs morales que sont l'universalisme, le communalisme, le désintéressement et le scepticisme organisé. Bien des recherches sociologiques et philosophiques ont démontré depuis que les pratiques scientifiques sont loin d'être exemptes des triturations subjectives des individus la finançant, la pratiquant, l'exploitant et la consommant. Il est facile de faire la promotion des caractéristiques de transparence, d'accessibilité et de collaboration de la science ouverte en interpellant à nouveau ces idées nobles que les chercheurs sont prêts à partager de façon désintéressée leurs résultats de recherche, que les problèmes scientifiques sont identifiés collectivement et résolus par la convergence des efforts pour le bien de tous selon les plus hauts standards de rigueur.

Plus spécifiquement, la science ouverte est un ensemble de pratiques couvrant un large spectre d'activités influencées par cinq écoles de pensées identifiées par Fecher et Friesike (2014). Nous nous arrêtons plus particulièrement sur deux d'entre elles qui nous semble avoir le plus d'implication sur la fonction de bibliothécaire. Ainsi, l'école démocratique rejoint les valeurs de la profession bibliothécaire en ce qu'elle présume que l'accès aux connaissances est inégalement distribué et que les pratiques de communication savante devraient donc permettre la libre publication immédiate et gratuite de celles-ci pour le bien de tous. On reconnaît ici le rôle traditionnel de promotion de la publication en libre accès tenu par les bibliothécaires mais on verra plus loin comment ce rôle de promotion pourrait être actuellement en transformation. De nouveaux rôles peuvent aussi être identifiés en lien avec cette école : le soutien au partage de données de données de recherche et l'accompagnement à la production de ressources éducatives libres. Nous ne nous attarderons pas sur ce dernier secteur d'intervention que nous connaissons moins.

Une autre école de pensée qui a particulièrement une incidence sur le rôle du bibliothécaire est l'école pragmatique. Cette école de pensée suppose que la production des connaissances pourrait être plus efficace, notamment avec une meilleure collaboration entre les membres des communautés de recherche. Les pratiques en découlant sont par exemple la publication du protocole de recherche, le recours à un journal de bord électronique (et collaboratif) l'ouverture du processus de révision par les pairs, dernier exemple mais il y en aurait d'autres, le rehaussement de la GDR pour éviter le dédoublement des efforts et favoriser la réutilisation des données. Les bibliothécaires œuvrent donc à la sensibilisation de bonnes pratiques, font connaître à leur communauté de recherche de nouvelles façons de communiquer la science et les familiarisent avec diverses solutions de dépôt.

On comprend qu'on prête aux pratiques de science ouverte de nombreuses vertus que la feuille de route sur la science ouverte du Canada reprend à son compte. Les bibliothécaires sont déjà à pied d'œuvre dans la science ouverte, mais la prépondérance du discours enthousiaste sur la science ouverte n'informe pas nécessairement ce nouveau portfolio de services. Le tableau ci-dessous propose de débiter le questionnement en interrogeant quelques-unes de ces vertus.

TABLEAU 1

Les vertus de la science ouverte interrogées

L'ouverture de la science...	Question
Facilite la reddition de compte	Les politiques sont-elles suffisamment mordantes pour permettre une telle reddition ?
Accroît la reproductibilité des résultats	Est-ce toutes les sciences qui sont reproductibles ?
Augmente la confiance du public à l'égard de la science	Qu'est-ce qu'avoir confiance en la science ?
Réduit le dédoublement des efforts	Est-ce vraiment la majorité des devis de recherche qui produisent des données avec un potentiel de réutilisation souhaitable voire réalisable ?
Accélère l'innovation	Accélère l'innovation pour qui ? Innovation sociale ?
Valorise la diversité des systèmes de connaissance	Qu'est-ce qui qualifie un système de connaissances ?
Favorise la réutilisation de l'information scientifique	Quid du comportement informationnel des chercheurs ?

Ce tableau remet en perspective les avantages généralement présentés lorsqu'il est question de science ouverte. À cette réflexion, s'ajoute également celle de bien saisir l'ensemble des acteurs impliqués dans cet univers, qui ne sont pas toujours sans intérêts. Ce réseau d'influence est composé de divers groupes d'intérêt et d'engagements que le Canada a

pris par sa participation à des instances internationales. Rappelons que l'un des faits saillants est l'importance économique associée aux données de recherche. En effet, la *Recommandation de l'OCDE sur l'accès aux données de la recherche financée sur les fonds publics* est l'une des sources sur laquelle s'appuie les textes canadiens d'orientation politique pour la GDR.

Le réseau d'influences permet aussi de constater que les principes FAIR sont centraux à l'encadrement des nouvelles pratiques de science ouverte canadiennes. La communauté FORCE 11 qui a offert le cadre développement de ces principes a un financement provenant autant de fonds publics que commerciaux (voir leur mode de gouvernance). Finalement, le réseau d'influence du discours canadien se caractérise par une grande concentration de groupes d'intérêt concernés par la publication en libre accès. Notons que les éditeurs commerciaux sont impliqués dans plusieurs d'entre eux. La présence de ces éditeurs commerciaux est un bon prétexte pour rappeler que ces derniers jouent un rôle non négligeable dans l'évolution des pratiques de communication savante. La question des enjeux économiques sous-jacents à une potentielle instrumentalisation au discours de la démocratisation des connaissances devient

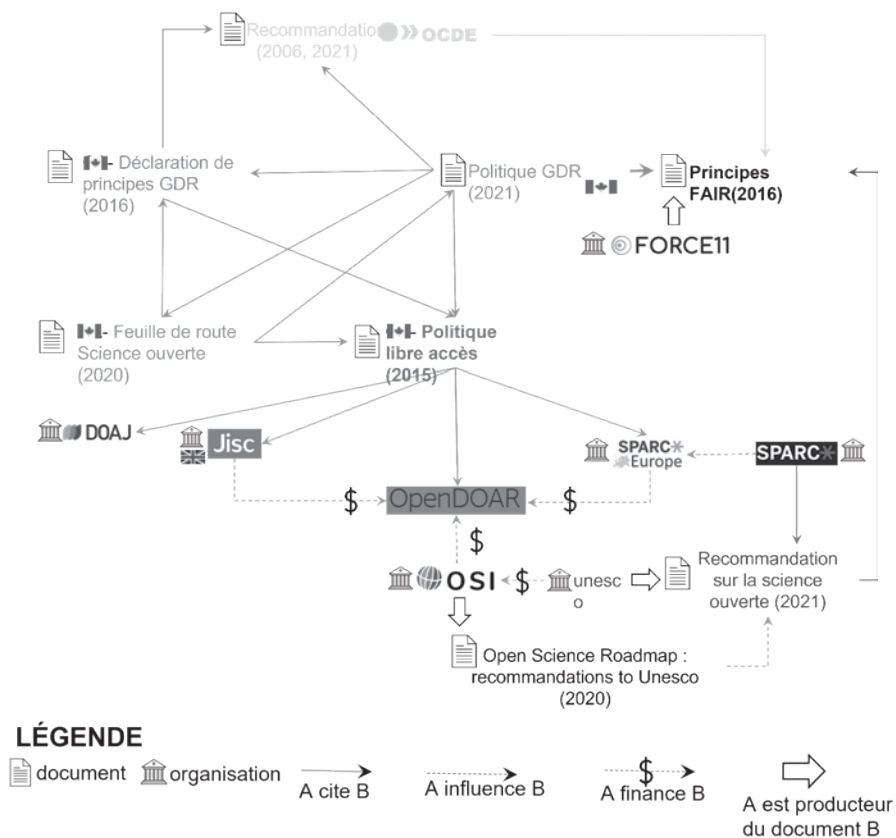
légitime et met la table pour une position plus assumée de la part des bibliothécaires lorsque leur contexte institutionnel les met de plus en plus en position de promotion de toutes ces « bonnes pratiques » d'ouverture.

État des pratiques de sciences ouvertes selon la littérature

Si l'on récapitule, quand on creuse un peu plus pour considérer les « faits de la science » qui soutiennent les grandes affirmations d'accélération de l'innovation, d'augmentation de l'impact de la recherche et d'accessibilité pour les citoyens, on en vient plus à la conclusion que les problèmes contemporains que les pratiques de science ouverte sont censées régler ne sont pas tant si nouveaux ni ne sont pas tant résolus par l'ouverture. Le développement des savoirs scientifiques est une aventure humaine et son passé montre qu'elle est affectée par son contexte sociohistorique qui n'est certes pas imperméable aux idéologies et croyances sculptant ses diverses parties prenantes. Faisons alors le point sur l'impact réel de certaines pratiques de science ouverte. Parmi ces impacts, l'on s'aperçoit que le bibliothécaire universitaire, qui traditionnellement jouait un rôle de

FIGURE 1

Schéma de citations de l'environnement normatif canadien en GDR



soutien, est de plus en plus amené à développer son esprit critique et à agir en tant qu'expert sur certains enjeux de la communication savante et d'offrir un service conseil.

En effet, l'expérience de terrain du bibliothécaire impliqué dans les dossiers relatifs à la science ouverte lui fait vite rencontrer des chercheurs et chercheuses qui expriment moins d'enthousiasme sur les vertus des pratiques de science ouverte. Dans certains cas, le rôle du bibliothécaire est certes un travail de sensibilisation, mais s'il est attentif, ce professionnel saura reconnaître des réticences fondées sur de réelles préoccupations. Miroski (2018) considère ainsi que la science ouverte prendrait racines dans le présent régime néolibéral de la science et que des impératifs du marché seraient surtout à l'œuvre dans la reconfiguration de nos institutions et de la nature des connaissances plutôt que dans la présence de réels nouveaux problèmes requérant une transformation des pratiques de la recherche scientifique. Pour ceux qui sont moins familiers avec ce courant politique, l'article de McKeown (2022) énumère quelques caractéristiques de l'université néolibérale. Voici tout de même quelques exemples rapides pour illustrer cette position de Miroski. Tout d'abord, l'emphase est mise sur la publication en libre accès et, plus récemment, sur le partage des données des recherches subventionnées par les fonds publics. Le Plan S intègre certes tout autant les subventions publiques que privées dans son principe fondateur, mais dans les faits la majorité de la recherche financée par les fonds privés (certainement au Canada) n'est pas couverte par cette obligation d'ouvrir les résultats de la recherche. Et que dire de toute la recherche non subventionnée ? On laisse les chercheurs et chercheuses librement choisir cette option d'ouverture puisque très peu d'institution de recherche possède une politique de libre accès contraignante. Comme Miroski le formule, on est plutôt dans une vision de marché où le consommateur affirme « we paid for it, so we deserve full and free access to the results of research » (p. 178). Miroski mentionne également comment la volonté d'améliorer la qualité de la science par l'ouverture de ses pratiques n'est à ce jour pas encore concluante face au constat que le nombre de rétractation d'articles ne cesse d'augmenter et que les problèmes de biais de publication dénoncés depuis les années 1950 sont encore d'actualité. Ainsi, l'ouverture, peu importe à quelle étape du processus de la recherche elle est implantée, ne garantit pas un devis de recherche bien conçu. Comme dernier exemple, Miroski souligne qu'après presque 20 ans de travail pour le libre accès, nous sommes loin de contrebalancer l'appétit vorace de l'oligopole des éditeurs commerciaux : ceux-ci profitent désormais de la dominance du modèle d'affaires de publication savante en libre accès financé par les frais de traitement d'articles.

Il nous paraît essentiel que les bibliothécaires entrent en intervention auprès des chercheurs avec une éthique de

leur relation de travail rehaussée : le destinataire de l'intervention a peut-être un système de valeurs qui s'entrechoque avec ce contexte néolibéral. La publication en libre accès est une des pratiques de science ouverte où les bibliothécaires ont à mieux articuler leur compréhension car leur rôle, traditionnellement limité à la promotion du libre accès, se tourne désormais vers l'évaluation et la participation à un écosystème de communication savante équitable.

À ce sujet, le Plan S¹ met une forte pression sur les éditeurs de revues commerciales à transformer leur modèle d'affaires vers le libre accès. Les Fonds de recherche du Québec (FRQ) est l'un des premiers organismes de financement nord-américains à avoir rejoint cOAlition S alors qu'en 2018, des organisations de 37 pays réparties sur les cinq continents ont fait une déclaration commune de soutien au Plan S lors de la 14^e conférence de Berlin sur le libre accès (Max Planck Digital Library, 2018). À cause du Plan S, on assiste aujourd'hui à la consolidation du modèle de libre accès financé par les frais de traitement d'article ou FTA (*Article Processing Charge ou APC*) comme modèle dominant. Les revenus des éditeurs en FTA tourneraient autour de 1,5 milliard de dollars (US) en 2021 (Zhang et al., 2022). Ces chercheurs rapportent aussi que les frais de traitement d'article évoluent plutôt à la hausse, davantage que l'indice des prix à la consommation. Leur revue de littérature laisse croire que le modèle de financement de la publication par les frais de traitement d'article s'annonce pour être plus lucratif pour les éditeurs commerciaux que le modèle basé sur les abonnements. À terme, les bibliothèques ne géreront plus des abonnements mais des ententes financières avec les éditeurs commerciaux et veilleront à la bonne administration des fonds publics vers un poste budgétaire dédié aux frais de traitement d'articles. Il y a là tout un rôle d'évaluation qui se pointe à l'horizon. Dans un horizon pas si lointain, on peut présager que toutes les grandes revues commerciales seront en libre accès. Toutefois, lesquelles devrait-elles faire l'objet d'une entente financière avec un éditeur ? Les éditeurs vont-ils permettre aux bibliothèques de sélectionner les revues composant les ententes ? Les revues commerciales qui seront exclues d'ententes financières mais qui sont en libre accès seront certes accessibles en lecture pour tous mais seront-elles accessibles aux chercheurs de l'institution qui n'ont pas de subventions pour financer les FTA afin d'y publier ? On assiste donc à une augmentation des

1. cOAlition S définit ainsi le [Plan S](#) : Le Plan S est une initiative lancée en septembre 2018 soutenant la publication en libre accès. Le plan est soutenu par cOAlition S, un consortium international d'organisations de financement de la recherche d'organisations de recherche. Le Plan S exige que les publications scientifiques qui résultent de recherches financées par des subventions publiques soient publiées en libre accès dans des revues ou plateformes se conformant à certaines exigences.

publications en libre accès mais cela ne se traduit pas nécessairement par une augmentation de l'ouverture de la science et les bibliothécaires sont au cœur de ce nouveau façonnage de l'économie de l'information savante.

À ce rôle d'évaluation des revues s'ajoute un rôle de représentation pour bien faire connaître le contexte politique de la communication auprès de la communauté de recherche. En effet, tous ces changements dans le secteur du libre accès à la publication doivent être accompagnés de mesures de suivis de la conformité des chercheurs et chercheuses aux mandats de libre accès et, éventuellement de partage de données. Les organismes de financement s'attendent à ce que ce soient les institutions qui offrent à leur communauté de recherche un environnement les soutenant vers cette conformité. On devrait voir dans le futur une adoption importante de politiques institutionnelles en lien avec la science ouverte et c'est à ce niveau que les bibliothécaires devraient jouer davantage un rôle de représentation afin de mettre à l'ordre du jour les enjeux de communication savante qui ne peuvent être efficacement pris en charge qu'avec le cadre qu'une politique institutionnelle peut offrir. Un sondage mené auprès d'universités européennes (Morais et al., 2021), constate que la publication en libre accès est la pratique jugée comme ayant la plus haute valeur stratégique parmi toutes les pratiques de science ouverte. Pourtant, seulement 60 % des institutions ayant répondu au sondage ont une politique institutionnelle de libre accès implantée ou en élaboration. Au Canada, on dénombre 18 institutions de recherche ayant déclaré une politique de libre accès sur un potentiel de 96 universités canadiennes (<http://roarmap.eprints.org/view/country/124.html>) et ce chiffre stagne depuis 2015, année où les Trois Conseils ont pourtant adopté leur politique harmonisée de libre accès. Les accords transformatifs, actuellement conclus entre les éditeurs et les bibliothèques et motivés par le Plan S, sont des stratégies légitimes sur le court terme seulement. La présence de politiques institutionnelles sur le libre accès pourrait aider à configurer un écosystème de publication un peu plus équitable et non laissé à la merci des forces du marché. L'intégration de bibliothécaires dotés d'une opinion professionnelle éclairée et critique dans l'exercice de production de ce cadre normatif est certainement un élément de succès pour s'attaquer aux changements structurels permettant l'atteinte de cet avenir plus équitable.

Rôle des intervenants GDR

Nous l'avons évoqué précédemment, les fondements de la science ouverte résonnent particulièrement avec la notion de « communalisme » qui stipule que la science est un bien public et qu'en tant que processus collectif, les résultats de la recherche scientifique doivent circuler librement et ne pas faire l'objet d'appropriation privée (Merton, 1973).

En ce sens, développer des mécanismes afin que des recherches génératrices de données, financées par des fonds publics, soient dans la mesure du possible partageables et réutilisables, témoigne d'un souci de démocratiser la science qui mérite d'être salué.

Rappelons toutefois que le lien naturel entre sciences et démocratie est en soi une posture épistémologique. D'autres courants de pensées, prenant leurs origines chez Auguste Comte, stipulent que la surspécialisation de la science l'éloigne du citoyen et peut être perçue comme antidémocratique et autoritaire. (Gingras, *Sociologie des sciences* p. 21). L'exemple des mouvements antivaccins pendant la COVID-19 en témoignent.

Dans tous les cas, il nous semble important de nous attarder au rôle que peut jouer le bibliothécaire universitaire dans ce contexte et comment il se doit d'être outillé pour accompagner les équipes de recherche.

Dans les universités canadiennes, les bibliothécaires affichent un leadership dans la prise en charge du mandat de soutien des chercheurs et des chercheuses dans le processus de la gestion des données de la recherche. Les bibliothécaires collaborent aussi avec des professionnels d'autres services aussi interpellés par les développements de la GDR. Un référentiel de compétences pertinent pour tous ces intervenants GDR est celui de l'ACRL². La culture des données est une littérature ayant des points communs avec plusieurs piliers de la culture de l'information que le Référentiel de l'ACRL cherche à développer, notamment :

1. *L'information a une valeur*

Les données, comme source d'informations, sont influencées par différents facteurs et leur valeur est bien liée, par exemple, au contexte de leur utilisation.

2. *La recherche savante est une démarche d'investigation*

« La recherche savante est une démarche itérative qui exige de poser des questions d'une complexité croissante ou inexplorée. Les réponses à ces questions génèrent de nouvelles questions ou d'autres pistes d'investigation. » (ACRL, p. 14). L'analyse des données en phase active de projet présente ce même caractère itératif et vient justifier l'importance d'une bonne documentation de ces manipulations.

3. *La production de savoirs résulte d'échanges*

« Au lieu de rechercher des réponses précises à des problèmes complexes, les experts comprennent qu'un problème particulier peut, en situation d'échange, être examiné de plusieurs points de vue divergents » (ACRL, p. 16). On retrouve ici un cadre

2. Association of College and Research Libraries

approprié pour discuter, par exemple, de l'accès restreint à certaines données sensibles.

Cette exploration du Référentiel de l'ACRL gagnerait en profondeur en l'approchant systématiquement sous l'angle de la culture des données pour penser la formation à la GDR. Par culture des données ou *data literacy*, nous entendons la définition de Prado et Marzal (2013, p. 125): « Composante de la culture de l'information (*information literacy*) permettant à un individu « d'accéder, d'interpréter, d'évaluer de façon critique, de gérer, de manipuler et d'utiliser de façon éthique les données » [Notre traduction].

Comme les intervenants GDR font de la sensibilisation et de la formation, aux bonnes pratiques GDR qui s'inscrivent, comme on l'a vu, particulièrement dans le mouvement de la science ouverte, il importe que cet effort de sensibilisation soit réfléchi pour éviter de perpétuer des lieux communs sur les vertus de ces pratiques. Nous proposons de cadrer cette réflexion pédagogique sur les programmes de développement des compétences informationnelles pour que ceux-ci soient mis à jour en y ajoutant des référentiels tenant en compte les compétences nécessaires pour exploiter les données de recherche. À ce sujet, le cadre (framework) développé par Prado et Marzal (2013) est tout à fait excellent et se décline en quatre objectifs :

1. Comprendre les différents types de données, ce qu'elles sont et leur place dans notre société
2. Savoir trouver ou obtenir des données
3. Être mesure de lire, interpréter et évaluer des données
4. Connaître les principes de la gestion des données

Une des forces de ce cadre est son premier objectif qui s'intéresse à la place des données dans la société et permet d'entrée de jeu des lectures critiques sur les usages et discours autour des données. La formation des bibliothécaires devrait non seulement prévoir la compréhension des bonnes pratiques GDR et prévoir le développement de compétences un peu plus techniques pour le traitement des données mais intégrer aussi des contenus sur ce que les théories sociologiques permettent de dégager sur la prégnance des données dans nos sociétés. De plus, le deuxième objectif de ce cadre, avec sa composante « savoir obtenir des données » implique que les bibliothécaires aient davantage de connaissances sur les méthodologies de recherche afin de pouvoir interagir correctement avec leur communauté de recherche. Pour comprendre l'importance de cet enrichissement de la compétence professionnelle du bibliothécaire en matière de culture des données, la section suivante traitera donc de quelques enjeux à la croisée de cette culture des données et des implications épistémologiques que la « datification » de la science fait peser sur la recherche qualitative.

Enjeux méthodologiques et épistémologiques

Jusqu'ici, nous avons établi l'importance de comprendre le contexte et les enjeux économiques et politiques de la gestion des données de recherche. Puis, plus concrètement, nous avons cherché à démontrer comment ces lectures critiques devraient s'ajouter aux nouveaux rôles des bibliothécaires et comment l'adoption d'un référentiel de compétences de culture numérique permettrait d'organiser les activités de sensibilisation et formation en lien avec ces nouveaux rôles. Nous poursuivons maintenant la réflexion en se penchant sur l'impact du discours fort autour des données et le contexte méthodologique de la recherche qualitative. Du point de vue de la pratique scientifique, quel est l'effet de la valorisation des données sur les conditions de production des savoirs ? Il existe une littérature foisonnante qui se préoccupe de ces enjeux, nous allons en exposer quelques-uns, mais d'abord, précisons bien que c'est la question de la donnée numérique qui soulève le plus d'interrogation. En effet, la tendance observée dans les pays membres de l'OCDE, comme le Canada, est de favoriser la libre circulation des données de recherche. Tel qu'énoncé dans les *Principes et les lignes directrice de l'OCDE pour l'accès aux données de la recherche financés sur fonds publics* :

L'accès ouvert aux données de la recherche financée sur fonds publics et leur partage contribuent non seulement à maximiser l'impact des nouvelles technologies et des nouveaux réseaux numériques sur le potentiel de recherche mais permettent aussi un retour plus important sur l'investissement publique dans la recherche (OCDE, 2007, p. 12).

Cet extrait illustre que pour qu'un jeu de données soit déposé et partageable, il doit correspondre à certains critères, dont celui d'être idéalement numériquement exploitable. Cette condition soulève des enjeux éthiques et épistémologiques principalement dans les sciences sociales et les humanités de même que les projets qui comptent utiliser des méthodologies qualitatives. Ils sont également étroitement liés aux questions soulevées par l'utilisation du *Big Data* en recherche.

Données de recherche, médias sociaux et éthique

L'une des préoccupations que l'on retrouve fréquemment dans la littérature concerne le type de recherche que la valorisation des données numériques et, par la bande, les bonnes pratiques de gestion – et de partage – de données de recherche assujetties au discours dominant pourraient favoriser. L'inquiétude se situe notamment sur les risques de surutiliser et de survaloriser les données disponibles sur

les plateformes de réseaux sociaux. Depuis plusieurs années, dans diverses disciplines des sciences sociales, « les réseaux sociaux sont mobilisés comme de nouveaux observatoires des tendances de la société » (Boyadijian, Velcin, 2017, p. 112). En plus de fournir de gros corpus (Hesse et al., 2019), on considère qu'ils permettent de penser les interactions sociales. Bien que maintenant largement utilisées comme sources de données dans le milieu académique, les plateformes de médias sociaux n'ont pas été conçues pour aborder des questions scientifiques (Weinhardt, 2021). À cela s'ajoute également plusieurs considérations éthiques. Rappelons-nous la controverse suscitée par un partenariat entre une équipe de recherche et l'Université de Californie en 2014. Le fil Facebook de plusieurs utilisateurs avait été manipulé dans l'objectif de modifier les informations reçus par les utilisateurs dans le but d'évaluer s'il y avait un lien entre les messages positifs ou négatifs reçus et ceux qui étaient postés par la suite (Metcalf et Crawford, 2016).

Si nous poursuivons avec ces enjeux, l'on pense notamment à l'utilisation des textes issus des médias sociaux qui maintiennent leur caractère privé malgré leur « publication » et qui ne sont donc pas produits dans l'objectif d'être réutilisés pour la recherche à d'autres fins que celles pour lesquelles ils ont été publiés. Comment articuler l'éthique de la recherche dans le respect de l'évolution des lois relatives à la protection de la vie privée qu'on retrouve désormais dans plusieurs pays (ex. : RGPD dans les pays européens, Loi 25 au Québec) ?

La façon d'anonymiser des données qualitatives avec des populations marginalisées est une autre préoccupation importante. Dans le contexte de populations marginalisées, on peut même se demander si l'anonymisation des données est possible : des croisements entre grands jeux de données et l'utilisation de l'intelligence artificielle pourrait permettre de procéder avec un bon niveau de confiance à l'identification des individus, ouvrant la porte à de possibles préjudices. Par ailleurs, même avec les seules métadonnées, il peut y avoir des enjeux de confidentialité qui se présentent. La relation entre éthique de la recherche et numérique se doit d'être mieux comprise afin d'être très bien balisée (Gilbert et al., 2020). C'est en quelque sorte l'éthos des communautés de recherche qualitative qui est ainsi appelé à s'acculturer face aux injonctions à « FAIRiser » ses données : la conduite de la recherche qualitative commence à changer avant que d'importants consensus disciplinaires n'aient émergé.

Une attention particulière devrait être apportée par les responsables de la certification éthique pour que ces derniers s'assurent de prendre en considération le suivi du cycle de vie complet des données récoltées pour des projets de recherche. En plus des détails entourant la collecte, le processus doit s'ouvrir à de nouveaux processus ou

mécanismes d'archivage et de réutilisation des données (le R de FAIR). En développant une vision commune, les différentes parties prenantes (responsables éthiques, curateurs des dépôts, équipes de recherche, populations particulières collaborant avec les chercheurs, etc.) pourraient en amont veiller à ce que l'application des principes FAIR se fassent dans le respect des particularités disciplinaires et de principes reconnus par les communautés participants aux projets de recherche (Karsher, Lafferty-Hesse, 2019).

Des innovations méthodologiques sont appelées à voir le jour dans les prochaines années avec des programmes de financement dédiés spécifiquement à ce type de développement. Par exemple, l'Open Science Fund du Conseil néerlandais de la recherche (NWO) a financé le projet *Innovating Methods for Open Science in Qualitative Management Research* (OPEN-QUAL) dans lequel l'équipe de Hans Berends a exploré le partage d'un protocole d'analyse permettant la réutilisation de données qualitatives issues d'études de cas tout en préservant les ententes de confidentialité. Au lieu de chercher à anonymiser des données sensibles ou confidentielles pour qu'un chercheur lambda puisse les réanalyser, son équipe propose de décentraliser l'interprétation des données en partageant un protocole d'analyse. Chaque ensemble de données qui est sous entente de confidentialité prévoit déjà sélectionner des chercheurs autorisés à les consulter. Ensuite, un chercheur principal d'un nouveau projet X fait alors appel à leur collaboration en convenant avec ceux-ci d'un protocole d'analyse guidant l'extraction de l'information requise pour le projet X. Ce chercheur principal a ainsi accès à plus d'études de cas que s'il travaillait selon l'approche traditionnelle, augmentant ainsi la puissance de son analyse : des études de cas existantes sont réinterprétées à la lumière des besoins d'une autre équipe (Berends, 2022). Il s'agit pour l'instant d'une innovation méthodologique en émergence et qui n'a pas fait ses preuves, mais qui semble prometteuse dans l'optique de partager des données qualitatives sans avoir à enfreindre les ententes de confidentialité des participant.e.s. Nous suivrons cette équipe de près.

Les « Critical data studies »

Finalement, les préoccupations suscitées par ces nouveaux enjeux ont donné naissance à un nouveau domaine d'études que l'on appelle les « *critical data studies* » (CDS). Les chercheurs qui s'identifient à ce courant s'intéressent plus spécifiquement aux rôles des données dans une perspective qui se situe à l'intersection des études en technologies, en sciences sociales et humanités et en droit (Neff et al., 2017). Comme le nom de ce courant l'indique, il s'agit d'adopter un angle critique dans l'étude des données qui s'éloigne de la vision plus traditionnelle ou positiviste qui considère, dans sa forme la plus simple, les données comme des objets

observables (Iwasinki, 2020). Inspirée de la sociologie critique développée par l'école de Frankfort, les CDS rejettent cette vision jugée trop normative et considère les données comme une construction sociale indissociable de son contexte de production (Iwasinki, 2020).

Leurs propos pourraient se résumer par les appels à discussion suivants :

TABLEAU 2

Le cadre de réflexion sur les pratiques autour des données proposé par Neff et al. (2017)

Appel à discussion	Explications
Les données sont en très grandes parties sociales	Il importe de toujours bien communiquer le contexte de production des données, car l'on estime que dans un jeu de données, il y a eu 80 % de nettoyage pour 20 % d'analyse. (Neff et al., 2017)
Le sens qui est donné aux données est toujours un processus collectif	Il y a toujours un processus de négociation dans les équipes de recherche sur le sens qui sera accordées aux données.
Les données ne sont pas une destination	Plusieurs chemins peuvent être suivis dans l'analyse d'un jeu de données. Après avoir élaborer la problématique, elles constituent le point de départ de l'interprétation pour permettre aux résultats de recherche d'être transparents. Elles ne doivent donc pas être considérées comme une fin en soi.
Les données sont une suite d'histoires	Dans une recherche, il y a une histoire qui mène à la collecte de données. Il y a ensuite une histoire qui se déroule durant la production de données et dans l'interprétation. Ce sont l'ensemble de ces histoires qui, une fois intégrées dans un narratif, constitueront la recherche.

Les chercheurs qui se rattachent à ce courant, recommandent que les praticiens des sciences des données travaillent de concert avec ceux qui les critiquent pour qu'ensemble, ils puissent faire en sorte de développer une culture de l'éthique dans le *data management*.

Dans le cas de la GDR, en sciences sociales, les données partagées auront préalablement fait l'objet d'un travail d'interprétation. Un jeu de données, peu importe la discipline est une construction qui ne peut s'abstraire des sujets humains qui la conçoivent. Avant d'être déposé dans une plateforme, il a fait l'objet de délibérations de négociations de prises de décisions d'inclusion et d'exclusion bien ancrée dans des discours dominants, dans des réalités historico-socioéconomiques. Par conséquent, il est impossible de prétendre à la neutralité des données partagées (Neff et al., 2017). La documentation des ensembles de données révèle

particulièrement ici toute son importance tout en laissant ouverte la question des connaissances tacites pouvant échapper à cet effort de documentation, connaissances pourtant nécessaires pour bien comprendre un ensemble de données.

Conclusion

Cet article avait pour objectif de sensibiliser les intervenants GDR en attirant leur attention sur les aspects critiques de la science ouverte. Certains de ces aspects interpellent plus directement la GDR, et soulèvent des enjeux importants pour les recherches en sciences sociales et humanités. La GDR est portée par de belles valeurs de transparence et de démocratisation ainsi que par un grand potentiel de formation, mais il faut être certain de la pratiquer en tenant compte de la complexité des épistémologies disciplinaires et ce, sans ignorer les rapports de pouvoir qui continuent de façonner l'écosystème de l'information savante. À notre avis, le travail de sensibilisation dans une perspective critique devrait continuer à se faire autant auprès des chercheurs et des chercheuses que de l'ensemble des professionnels et professionnelles en soutien à la recherche. Avec un tel apport critique dans le développement et l'appropriation des pratiques GDR, on peut espérer la construction de communautés d'intérêts disciplinaires et méthodologiques gagnant en autonomie d'agir, respectueuses de leurs épistémologies et alors garantes d'un sain développement des systèmes de connaissances bâtissant nos sociétés.

Pour les chercheurs, davantage de travaux permettant de démontrer les avantages de l'ouverture de la science et du partage des données doivent être menées pour que les vertus de la science ouverte soient plus qu'un vœu pieux (voir tableau 1). Pour les bibliothécaires, les différents programmes en sciences de l'information devraient veiller à adapter leur programme de formation pour que les nouvelles générations de bibliothécaires soient outillées en amont à bien saisir ces considérations des pratiques de la science ouverte, tant du point de vue de la compréhension théorique et critique que des enjeux pratiques notamment méthodologique. Le modelage actuel de l'éthos des communautés de recherche sous l'effet des nouvelles pratiques d'ouverture à adopter devrait être central aux discussions disciplinaires à partager avec les chercheurs. C'est en passant par le renforcement de l'autonomie des communautés disciplinaires qu'on pourra assurer le développement d'objets de recherche que ces communautés pourront encore qualifier de scientifiques tout en demeurant pertinents pour les citoyens et citoyennes supposer bénéficier de la grande entreprise scientifique.

SOURCES CONSULTÉES

- Aspesi, C., Allen, N. S., Crow, R., Daugherty, S., Joseph, H., McArthur, J. T., & Shockey, N. (2019, 29 mars). SPARC Landscape Analysis. <https://doi.org/10.31229/osf.io/58yhb>
- Belli, S., Mugnaini, R., Baltà, J., & Abadal, E. (2020). Coronavirus mapping in scientific publications: When science advances rapidly and collectively, is access to this knowledge open to society? *Scientometrics*, 124(3), 2661-2685.
- Berends, Hans. (17 novembre 2022). *Open Science Webinar S1E5: Facilitating the sharing and reuse of qualitative data*. Youtube. <https://www.youtube.com/watch?v=VjznmX5pHYM>
- Bilder G, Lin J, Neylon C (2020), *The Principles of Open Scholarly Infrastructure*. <https://doi.org/10.24343/C34W2H>
- Boyadjian, J. & Velcin, J. (2017). L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives?: La mesure de la popularité des personnalités politiques sur Twitter. *Questions de communication*, 31, 111135. <https://doi.org/10.4000/questionsdecommunication.11078>
- COPIM. (2021). *COPIM statement on the corporate acquisition of OA infrastructure*. <https://copim.pubpub.org/pub/copim-statement-corporate-acquisition-oa-infra/release/1>
- Cuocolo. L. (2019). L'utilizzo dei dati della P.A.: dall'accesso «civico» all'accesso «commerciale». [résumé] *Diritto pubblico comparato ed europeo*. Numero speciale. 163-190. DOI: 10.17394/93046 <https://www.rivisteweb.it/doi/10.17394/93046#>
- Dominik, M., Nzweundji, J.G., Ahmed, N., Carnicelli, S., Mat Jalaluddin, N.S., Fernandez Rivas, D., Narita, V., Enany, S. and Rios Rojas, C., 2022. Open Science – For Whom?. *Data Science Journal*, 21(1) <http://doi.org/10.5334/dsj-2022-001>
- Duan, Y., Luo, J., Zhao, L., Zhang, X., Miao, J., Moher, D., & Bian, Z. (2022). Reporting and data sharing level for COVID-19 vaccine trials: A cross-sectional study. *EBioMedicine*, 78 doi:10.1016/j.ebiom.2022.103962
- Fecher, B. et Friesike, S. (2014). Open Science: One Term, Five Schools of Thought. Dans Bartling S., Friesike S. (dir.) *Opening Science* (p. 17-47). Springer. https://doi.org/10.1007/978-3-319-00026-8_2
- Federer, L. M., Belter, C. W., Joubert, D. J., Livinski, A., Lu, Y. L., Snyders, L. N., & Thompson, H. (2018). Data sharing in PLOS ONE: an analysis of data availability statements. *PloS one*, 13(5), e0194768.
- Fortin, M.-F. et Gagnon, J. (2022). *Fondements et étapes du processus de recherche: méthodes quantitatives et qualitatives (4^e édition)*. Chenelière éducation.
- Gilbert J-P, Ng V, Niu J, Rees EE. Un appel à un cadre éthique lors de l'utilisation des données des médias sociaux pour des applications d'intelligence artificielle dans la recherche en santé publique. *Relevé des maladies transmissibles au Canada* 2020; 46(6):191-6. <https://doi.org/10.14745/ccdr.v46i06a03f>
- Gingras, Y. (2013). *Sociologie des sciences (1^{re} éd, Ser. Que sais-je?, 3950)*. Presses universitaires de France.
- Hesse, A., Glenna, L., Hinrichs, C., Chiles, R., & Sachs, C. (2019). Qualitative Research Ethics in the Big Data Era. *American Behavioral Scientist*, 63(5), 560-583. <https://doi.org/10.1177/0002764218805806>
- Karcher, Sebastian et Lafferty-Hess, Sophia. (2019). An epic journey in sharing: The story of a young researcher's journey to share her data and the information professionals who tried to help. *IASSIST Quarterly*, 43(1), 1-9.
- Iwasiński, Łukasz. (2020). Theoretical Bases of Critical Data Studies. *Teoretyczne podstawy critical data studies.*, 115A(1A), 96-109.
- Max Planck Digital Library. (2018). *Final Conference Statement: 14th Berlin Open Access Conference*. <https://oa2020.org/b14-conference/final-statement/>
- McKeown, M. (2022). The View from Below: How the Neoliberal Academy Is Shaping Contemporary Political Theory. *Society*, 59(2), 99-109.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago and London: Chicago University Press.
- Metcalf, J. et Crawford, K. (2016). Where are human subjects in Big Data research? The emerging ethics divide. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951716650211>
- Mirowski, P. (2018). The future(s) of open science. *Social Studies of Science*, 48(2), 171203.
- Miyakawa, T. (2020). No raw data, no science: another possible source of the reproducibility crisis. *Molecular brain*, 13(1), 1-6.
- Morais, R., Saenen, B., Garbuglia, F., Berghmans, S. et Gaillard, V. (2021). *From Principles to Practices: Open Science at Europe's Universities. 2020-2021 EUA Open Science Survey Results*. European University Association.
- Neff G, Tanweer A, Fiore-Gartland B, Osburn L. (2017). Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science. *Big Data*. 5(2):85-97. doi: 10.1089/big.2016.0050.
- Organisation de coopération et de développement économique. (2007). *Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics*. <https://www.oecd.org/fr/science/inno/38500823.pdf>
- Ortega-Rodríguez C, Licerán-Gutiérrez A, Moreno-Albarracín AL. Transparency as a Key Element in Accountability in Non-Profit Organizations: A Systematic Literature Review. *Sustainability*. 2020; 12(14):5834. <https://doi.org/10.3390/su12145834>
- Read, K. B., Ganshorn, H., Rutley, S., & Scott, D. R. (2021, October 14). Surveying the landscape of CIHR-funded research data sharing practices: An analysis of the published literature. [prépublication] <https://doi.org/10.31219/osf.io/h7byr>
- Schonfeld, R. C. (2017). Who Owns Digital Science? *Scholarly Kitchen*. <https://scholarlykitchen.sspnet.org/2017/10/23/ownership-digital-science/>
- Smale, N, Unsworth, K., Denyer, G. et Magatova, E. (2020), A Review of the History, Advocacy and Efficacy of Data Management Plans. *International Journal of Digital Curation*, 15(1). <https://doi.org/10.2218/ijdc.v15i1.525>
- Wilkins, J. S. (2008). The roles, reasons and restrictions of science blogs. *Trends in ecology & evolution*, 23(8), 411-413.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M. Baak, A., Bomberg, N., Boiten, J.-W., ... Mons, B. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:article no 160018 doi: 10.1038/sdata.2016.18 (2016).
- Weinhardt M. Big Data: Some Ethical Concerns for the Social Sciences. *Social Sciences*. 2021; 10(2):36. <https://doi.org/10.3390/socsci10020036>
- Zhang, L., Wei, Y., Huang, Y., & Sivertsen, G. (2022). Should open access lead to closed research? The trends towards paying to perform research. *Scientometrics*, 1-27.

CRÉDIT DES IMAGES UTILISÉES POUR LA FIGURE 1

Drapeau Canada: domaine public [https://commons.wikimedia.org/wiki/File:Drapeau_du_Canada_\(Pantone\).svg](https://commons.wikimedia.org/wiki/File:Drapeau_du_Canada_(Pantone).svg)

DOAJ: pas suffisamment d'originalité pour un droit d'auteur mais la marque de commerce de l'organisation doit être respectée si vous comptez modifier la figure 1.1. https://commons.wikimedia.org/wiki/File:DOAJ_logo-colour.svg

JISC: pas suffisamment d'originalité pour un droit d'auteur mais la marque de commerce de l'organisation doit être respectée si vous comptez modifier la figure 1.1. https://commons.wikimedia.org/wiki/File:Jisc_logo.png

Drapeau UK: domaine public [https://commons.wikimedia.org/wiki/File:Flag_of_the_United_Kingdom_\(3-5\).svg](https://commons.wikimedia.org/wiki/File:Flag_of_the_United_Kingdom_(3-5).svg)

Force 11 CC BY bas de la page <https://force11.org/info/the-fair-data-principles/>

SPARC Europe CC BY en bas de la page web : <https://sparceurope.org/>

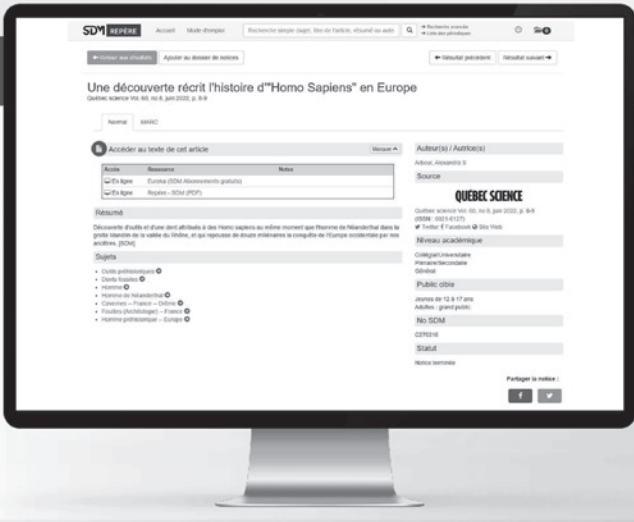
SPARC international: CC BY en bas de la page web : <https://sparcopen.org/>

Open DOAR CC BY-NC-ND 4.0 en bas de la page web <https://v2.sherpa.ac.uk/opensoar/>

Open Scholarship Initiative, CC-BY-NC-SA bas de la page web : <https://osiglobal.org/>

L'accès en ligne à l'information publiée dans les **revues** et **magazines** de langue française.

REPÈRE est bien présent dans les bibliothèques des réseaux d'enseignement et dans les bibliothèques publiques du Québec, ainsi que dans plusieurs autres provinces canadiennes et pays à travers le monde.



REPÈRE C'EST :

- Plus de 650 titres, spécialisés ou d'intérêt général, dont près de 200 titres courants;
- Près de 700 000 notices bibliographiques;
- Près de 13 000 nouveaux articles chaque année;
- Texte intégral : plus de 33 000 articles intégrés à la base de données et plus de 175 000 autres articles en ligne, pouvant généralement être consultés gratuitement;
- Un outil permettant de valoriser vos collections de périodiques;
- Une base de données accessible à distance gratuitement dans les bibliothèques abonnées et ce, sans limites d'utilisateurs simultanés;
- Une base de données conçue au Québec et développée en collaboration avec les utilisateurs afin de couvrir tous les domaines du savoir.

