Revue Organisations & territoires

Les enjeux de sécurité du partage des données de recherche dans les projets de recherche mettant en collaboration les établissements de recherche publics et le secteur privé

Marc Bruyère

Volume 30, Number 3, 2021

URI: https://id.erudit.org/iderudit/1086985ar
DOI: https://doi.org/10.1522/revueot.v30n3.1387

See table of contents

Publisher(s)

Université du Québec à Chicoutimi

ISSN

1493-8871 (print)
2564-2189 (digital)

Explore this journal

Cite this article

Bruyère, M. (2021). Les enjeux de sécurité du partage des données de recherche dans les projets de recherche mettant en collaboration les établissements de recherche publics et le secteur privé. *Revue Organisations & territoires*, *30*(3), 141–150. https://doi.org/10.1522/revueot.v30n3.1387

Article abstract

The desire by governments to foster collaboration and innovation in Quebec between public research institutions and private research partners brings to the forefront the challenges of research data security while promoting data sharing. This article explores the issues of scientific data security in the context of open science and the desire for research valorization.

Les enjeux de sécurité du partage des données de recherche dans les projets de recherche mettant en collaboration les établissements de recherche publics et le secteur privé

Marc Bruyère[a]

RÉSUMÉ. La volonté des gouvernements de favoriser les collaborations et l'innovation au Québec entre les établissements de recherche publics et les partenaires privés met à l'avant-plan les défis de la sécurité des données de recherche, tout en favorisant le partage de ces dernières. L'article explore leurs enjeux de sécurité dans un contexte de science ouverte et de désir de valorisation de la recherche.

ABSTRACT. *The desire by governments to foster collaboration and innovation in Quebec between public research institutions and private research partners brings to the forefront the challenges of research data security while promoting data sharing. This article explores the issues of scientific data security in the context of open science and the desire for research valorization.*

La création d'un vaccin pour combattre la COVID-19 a remis de l'avant la question de la collaboration en recherche. En effet, au cours des derniers mois, des entreprises et des universités se sont associées dans le développement et la production d'un vaccin (« COVID has shown the power of science-industry collaboration », 2021; University of Oxford, 2020). Les pays du G7 ont annoncé, en juin 2021, un pacte pour la recherche afin de favoriser la collaboration dans le domaine scientifique et le partage des données de recherche (DR) (G7, 2021). Durant la pandémie, les Fonds de recherche du Québec (FRQ), le Conseil de recherche en sciences naturelles et en génie du Canada (CRSNG) et les Instituts de recherche en santé du Canada (IRSC) ont réitéré l'importance de la science ouverte, c'est-à-dire l'accès pour tous aux publications savantes et aux DR (Collectif des 122 signataires, 2020). Les FRQ (s. d.) ont aussi adhéré au Projet S pour favoriser l'accès libre aux publications savantes. Ces événements rappellent les principes de la science ouverte selon lesquels la mise en commun des savoirs permet l'accélération du développement des connaissances et de l'innovation.

Cette volonté de transformation numérique et de collaboration en recherche, présente bien avant la pandémie, tend à s'accélérer. Récemment, le gouvernement du Québec a créé le Conseil de l'innovation ainsi que le poste d'innovateur en chef afin « de dynamiser le développement de l'innovation au sein des entreprises et de la société québécoise » (Ministère de l'Économie et de l'Innovation du Québec, 2020, s. p.). Il a aussi créé Axelys, une société de développement et de transfert de l'innovation qui favorisera la valorisation de la recherche entre les établissements d'enseignement supérieur (Axelys, s. d.).

Au moment d'écrire ces lignes, un projet de loi avait été adopté pour reconnaitre les données comme un actif gouvernemental afin d'en faire une gestion unifiée (Caire, 2021). Aussi, un poste de gestionnaire de la donnée numérique gouvernementale a été créé

[a] Bibliothécaire professionnel, Université du Québec à Chicoutimi

notamment pour assurer la recension et la qualité des données administratives (Caire, 2021). Un autre projet de loi était en cours pour la mise sur pied d'un ministère de la Cybersécurité et du Numérique, dont l'objectif sera de développer une expertise centralisée et de protéger les données gouvernementales (Caire, 2021). Selon Éric Caire (2021), ministre délégué à la Transformation numérique gouvernementale, ce ministère sera un partenaire extraordinaire et accompagnera la société québécoise dans sa transformation numérique et ces changements législatifs devraient être bénéfiques aux chercheurs. En effet, plusieurs intervenants (Caire, 2021; Malboeuf, 2021; Nadeau, 2021; Quirion, 2016) signalent que l'accès aux données gouvernementales par les chercheurs était très complexe par le passé et demeure un enjeu d'actualité. Les gouvernements ont pris conscience de l'importance des données, notamment en raison de la place de plus en plus grande de l'intelligence artificielle dans l'économie et la recherche (Halin, 2020; Premier ministre du Canada, 2018). « *It has become commonplace to refer to data as the "new oil" of the global economy* » (World Economic Forum, 2019, p. 4).

Ces constats ont amené le gouvernement fédéral à se doter d'une *Feuille de route pour la science ouverte* (Bureau de la Conseillère scientifique en chef du Canada, 2020) pour favoriser un accès plus grand à ses données. Plus encore, il appuie financièrement l'Alliance de recherche numérique du Canada (2021; Innovation, Sciences et Développement économique Canada, 2021c). Récemment créé, cet organisme « coordonnera et financera les activités liées aux volets du calcul informatique de pointe pour la recherche (…), de la gestion des données de recherche (…), et des logiciels de recherche (…) de la stratégie [fédérale] en matière [d'infrastructure de recherche numérique] » (Alliance de recherche numérique du Canada, s. d., para. 3).

En mars 2021, les trois organismes canadiens subventionnaires de la recherche (Conseil de recherches en sciences humaines du Canada [CRSH], CRSNG et IRSC, 2021) ont publié une politique commune sur la GDR qui sera mise en place graduellement. Cette politique instaure l'obligation aux établissements de recherche financés de produire une stratégie publique sur les services et les infrastructures nécessaires à la GDR dans leur organisation. Pour les équipes de recherche, c'est l'obligation de rédiger les actions et les décisions en matière de GDR au sein d'un plan de gestion de données (PGD). Elles devront aussi signaler à tous l'existence des ensembles de DR soutenant leurs publications scientifiques et, dans la mesure du possible et des cadres légaux, publier ces ensembles dans un entrepôt numérique : cette exigence est désignée comme étant le dépôt de données. Cette nouvelle politique commune aura des répercussions sur les projets de recherche en partenariat entre les établissements de recherche publics et l'industrie. En effet, seulement en 2018-2019, 1241 partenaires industriels étaient dénombrés dans les projets subventionnés du CRSNG, tandis qu'en 2020 au CRSH, 99 partenaires industriels et 515 organisations à but non lucratif étaient dénombrés (CRSH, 2021; CRSNG, 2021b).

Les collaborations scientifiques et les nouvelles exigences de la recherche sont un défi à la fois pour les universités, le réseau collégial et les entreprises. Selon Luc Sirois, innovateur en chef du Québec, en parlant de l'innovation en entreprise, « il n'y a pas une transformation numérique à faire au départ, il y a un rattrapage numérique à faire au départ. Ensuite, on pourra peut-être innover à partir des données, du savoir qu'on aura implanté » (Axelys, 2021, 22:06). Pour les universités et les cégeps, l'expertise, les infrastructures et le changement de culture en GDR constituent des enjeux importants (Cooper, Costanzo et collab., 2021; Cooper, Perry et collab., 2020).

Afin de faciliter la tenue de ces recherches en partenariat, il faut s'intéresser au partage de DR et à la GDR entre les acteurs de recherche. En effet, les DR sont à la base des projets de recherche. Si d'abord la question des données peut sembler ne concerner que les projets en intelligence artificielle, elle concerne tous les domaines de recherche, toutes disciplines confondues et de toutes envergures. Qui dit collaboration dit nécessairement partage (minimalement entre collaborateurs). Surgit alors assez rapidement la préoccupation de la sécurité des données, autant pour le respect de la vie privée, pour la confidentialité ou pour la préservation d'avantages concurrentiels.

Cet article décrit les principaux enjeux de cet aspect à considérer dans la gestion des données de recherche et quelques solutions, tout en favorisant le partage des données.

1. Enjeux et défis à relever pour la sécurité des données de recherche

Entre le bien-fondé et les grands principes de ces politiques et grands mouvements de science ouverte, des enjeux et des limites doivent être soulevés. Si de plus en plus l'accent est mis sur la collaboration entre partenaires et acteurs en recherche et en innovation pour favoriser les échanges et l'interdisciplinarité, bien souvent, les enjeux liés à la sécurité peuvent freiner les élans de partage de données de recherche ou sont perçus, à tort, comme insurmontables.

1.1 La protection de la vie privée et la confiance des individus

Comme toute autre information ou tout autre document, les projets de recherche portant sur des êtres humains doivent assurer la protection de la vie privée des participants selon les principes éthiques et les lois sur les renseignements personnels. Dans le cas de projets internationaux, cette superposition de lois et les enjeux de juridiction deviennent rapidement complexes. Les normes et les approbations éthiques varient. Il faut également tenir compte de la possibilité d'identifier des participants anonymes par des croisements de données. En soi, ces couplages de données offrent des pistes de recherche inexploitées et inégalées, encore plus lorsque les données sont partagées. Toutefois, il y a nécessité de considérer les risques et d'établir spécifiquement les données sensibles pour garantir le respect de l'intégrité des personnes impliquées. Par conséquent, le choix ne se résume pas à tout partager ou à ne rien partager des DR, mais à trouver un équilibre entre la protection des individus et le bien commun pour l'avancement de la science.

Certes, une non-conformité au cadre légal peut entraîner des pertes financières, mais, plus largement, une perte de confiance envers l'organisme/entreprise responsable de la recherche et ses partenaires. Les personnes qui participent aux études sont des volontaires et, dans bien des projets de recherche, elles ne reçoivent majoritairement aucune rémunération. Les réticences potentielles des individus en raison des violations de la confidentialité des données ne peuvent rendre que plus ardus le recrutement futur de participants ou l'obtention d'un consentement de leur part pour la réutilisation ultérieure des données de recherche collectées. Dans un contexte d'éloignement géographique, cet enjeu devient d'autant plus important que le bassin de participants est plus limité. Plus largement, une fuite de données peut ternir l'image de l'organisation chapeautant le projet de recherche auprès des bailleurs de fonds, des partenaires de recherche, des gouvernements et du grand public. Plusieurs vols de données de grandes organisations ont reçu une intense couverture médiatique au cours des dernières années telles que Desjardins (Dion, 2020) ou Equifax (Tchandem Kamgang, 2019). Dans certains cas graves, une négligence pourrait mener à une perte de confiance envers la science et ces institutions. D'autres situations peuvent engendrer ce résultat, par exemple une manipulation erronée ou la falsification des données. Conséquemment, ces situations peuvent avoir des impacts sur les conclusions et sur la validité d'une étude. Cela ne peut qu'alimenter la mésinformation, la désinformation, et les fausses nouvelles (*fake news*). Par exemple, des possibilités de traitement contre la COVID-19 par l'hydroxychloroquine ont été largement relayées sur les médias sociaux, malgré l'appel à la prudence de la communauté scientifique et les critiques méthodologiques de ces études (De Rosa, 2021; Ravidsky et Caulfield, 2021). Les médias américains ont relayé l'histoire d'un aîné décédé par automédication avec du phosphate de chloroquine, un traitement antiparasitaire pour les poissons d'aquarium; selon son épouse, gravement malade après avoir pris elle aussi cette substance, elle avait eu l'idée avec son mari après avoir écouté plusieurs conférences de presse du président américain Donald Trump promouvant l'utilisation de la chloroquine et de l'hydroxychloroquine comme traitement contre la COVID-19 (Edwards et Hilliyard, 2020; Waldrop et collab., 2020). Les études les plus médiatisées sur ce sujet sont maintenant accusées de manipulations de résultats par les collaborateurs de ces mêmes études (Ladepeche.fr avec AFP, 2021; Libération, 2021).

Un des piliers de la science est la rigueur méthodologique. Quant à la science ouverte, elle veut favoriser notamment une transparence des travaux de recherche et permettre ultimement une reproduction complète d'une étude par les autres chercheurs ou par les pairs réviseurs. Cela permet aussi de combattre la fraude scientifique puisque certains chercheurs vont jusqu'à créer des DR de toutes pièces en raison de la pression de publier ou d'obtenir des résultats positifs (Malboeuf, 2017).

1.2 La propriété intellectuelle et la concurrence

Comme il a été déjà mentionné, ces objectifs de partage entrent parfois en conflit avec d'autres objectifs de la recherche. Certaines données peuvent ne pas être diffusées ou encore doivent être diffusées à des individus ou à des organisations spécifiques au moment opportun. Pour toute recherche et innovation pouvant mener à la délivrance d'un brevet, toute divulgation préalable de l'invention pourrait anéantir l'obtention d'une telle protection. Les entreprises possédant d'importants ensembles de données peuvent attirer l'attention des organismes de surveillance de la concurrence et soulever des enjeux de monopole (Scassa, 2018). Si l'on ne prend que l'angle des monopoles, cela signifierait qu'un certain partage et qu'une certaine circulation des données permettraient la libre concurrence et l'innovation. D'un autre côté, les analyses de données massives et l'intelligence artificielle nécessitent un grand volume de données de qualité. À moins d'être une très grande entreprise ou un incontournable dans son secteur, les barrières à l'entrée imposées par le non-partage des données sont importantes. Cela devient alors à la fois un enjeu concurrentiel et juridique.

La propriété intellectuelle des données soulève quant à elle des questionnements. Au Canada, les données ne sont pas protégées par le droit d'auteur puisqu'elles constituent des faits, sauf si la base de données contient une sélection créative (Lapointe, 2018; Scassa, 2018). Néanmoins, la définition de la responsabilité des données doit être examinée par l'ensemble des partenaires de recherche : qui peut faire quoi et comment? Pour les chercheurs, la diffusion des DR est parfois nécessaire pour publier un article dans une revue afin de démontrer l'intégrité de leur recherche ou encore pour respecter les conditions des bailleurs de fonds. Sans compter que les ensembles de données peuvent être considérés comme une publication savante à part entière. Pour les entreprises, il y a le besoin de protéger certains avantages de compétitivité : il peut ne pas s'agir du projet de recherche à proprement parler, mais d'éléments analogues sur l'entreprise. Toutefois, il faut mentionner que la science ouverte et le partage des DR peuvent constituer un avantage compétitif. Par exemple, pensons au partage des données sur la COVID-19. Si toutes les DR peuvent ne pas être partagées, plusieurs DR d'un projet peuvent l'être. Par conséquent, il faut aborder clairement les possibilités de partage des DR dès le début du projet de recherche ou de la rédaction du contrat avec les partenaires pour éviter des malentendus ou des déceptions potentiels.

Certains projets de recherche dans des domaines stratégiques sont sujets aux préoccupations de l'espionnage industriel mené par des individus, par des organismes ou par des gouvernements étrangers mal intentionnés voulant dérober les DR à leurs avantages. Le gouvernement fédéral a récemment lancé des travaux en lien avec son récent *Énoncé de politique sur la sécurité de la recherche* (Innovation, Sciences et Développement économique Canada, 2021a) et mis en disposition en ligne un site web de sensibilisation, *Protégez votre recherche* (Innovation, Sciences et Développement économique Canada, 2021b, 2021d). Au-delà de la sécurité informatique, cet énoncé de politique met l'accent sur le choix et la vérification des partenaires de recherche. Cette dernière est même rendue obligatoire pour les demandes de subventions Alliance du CRSNG (2021a). Non seulement la sécurité des données en matière d'accès numérique et de diffusion doit être réfléchie, mais la gestion de l'accès aux environnements physiques doit aussi être considéré dans ces réflexions pour éviter le vol de l'équipement informatique ou les intrusions dans les laboratoires. Les vols de données peuvent être occasionnés par des personnes de l'entourage ou au sein de l'organisation ayant eu accès aux données pour les copier (Schaefer et collab., 2017). Le vol aussi peut entraîner une perte de données causant alors des problèmes opérationnels dans la tenue du projet de recherche, par exemple, causer des délais supplémentaires, fausser les résultats ou, dans les pires cas, obliger l'abandon du projet de recherche. D'autres causes peuvent entraîner une perte de données telles qu'une négligence ou une mauvaise manipulation des DR ou encore un sinistre ou un désastre naturel. Avec la valeur des données et un attrait de plus en plus grand pour la recherche et le développement (R-D), l'intérêt de personnes ou d'organisations mal intentionnées pour ces mêmes données ne pourra que s'accroître.

1.3 La menace climatique

La menace climatique constitue également un enjeu de la GDR. Même immatérielles, les DR consomment des ressources naturelles. En effet, elles

doivent nécessairement être stockées quelque part sur un disque dur, sur un serveur ou dans l'infonuagique (soit au sein de l'entreprise ou d'un partenaire ou encore chez un prestataire de services). Ces serveurs consomment des quantités importantes d'électricité et parfois d'eau pour le refroidissement dans les centres de données (Jones, 2018; Siddik et collab., 2021). Selon la source d'énergie utilisée, ils sont une source d'émission de gaz à effet de serre dans l'atmosphère terrestre (Siddik et collab., 2021). Selon les propos du consultant en informatique Jonathan Koomey, des gains supplémentaires en efficacité énergétique pour les centres de données seront ardus en raison des limites physiques des composants actuels des serveurs (Jones, 2018). Il arrive même que d'anciens serveurs fonctionnent toujours en entreprise sans aucune fin ni aucun objectif utile (Jones, 2018). Également, toute infrastructure informatique nécessite l'utilisation de matériaux miniers pour être produite. Il y a aussi une pénurie de semi-conducteurs et des enjeux éventuels d'approvisionnement pour les minéraux les constituant, ce qui entraîne alors des hausses de coûts (Hanssen, 2021).

Bref, peu importe le support de stockage choisi pour les DR, celui-ci est potentiellement exposé aux désastres naturels. Ces derniers peuvent engendrer des pertes de données causant alors des problèmes opérationnels dans la tenue du projet de recherche, par exemple, causer des délais supplémentaires, fausser les résultats ou, dans les pires cas, obliger l'abandon du projet de recherche. D'autres causes peuvent entraîner une perte de données telles qu'un vol, une négligence ou une mauvaise manipulation des DR. Les catastrophes naturelles n'iront qu'en accentuant au cours des prochaines décennies en raison des changements climatiques (Intergovernmental Panel on Climate Change, 2015). Ainsi, il n'est pas possible de concevoir l'espace de stockage comme étant infini sans considérer les coûts financiers et environnementaux : une sélection structurée des DR doit donc être réalisée.

1.4 L'éloignement géographique et la transformation numérique

L'éloignement géographique représente un défi au recrutement et au développement de l'expertise professionnelle en R-D et en GDR pour les entreprises, les organisations et les établissements de recherche publics. Sans être rattaché directement à la sécurité des données, la disponibilité et l'accès aux infrastructures de recherche peuvent aussi représenter une difficulté en région. Le risque d'une asymétrie des moyens de recherche entre les régions éloignées et les régions métropolitaines est réel. De plus, l'avènement et l'effervescence en intelligence artificielle nécessitent des infrastructures conséquentes.

Parallèlement, la numérisation des DR favorise la démocratisation de l'accès de ces dernières. Par exemple, nul besoin de se déplacer physiquement vers un centre de recherche pour accéder aux données et les partager, comme c'était le cas autrefois. Néanmoins, certaines données sensibles et détaillées demeurent uniquement consultables sur place pour des raisons de sécurité, notamment celles de Statistique Canada et de certains ministères québécois (Malboeuf, 2021; Statistique Canada, 2019b).

Les répercussions du télétravail accentuées par la pandémie de la COVID-19 restent une énigme à ce jour. Autant le télétravail pourrait faciliter le recrutement de personnel spécialisé en élargissant les possibilités de candidats, autant les régions éloignées pourraient devenir en plus grande compétition avec les grands centres urbains pour la recherche de talents.

Plusieurs des enjeux énumérés dans cet article comme la protection de la vie privée, la propriété intellectuelle, l'espionnage industriel et l'éloignement géographique étaient déjà présents par le passé. Ainsi, il n'y a rien de complètement nouveau. Cependant, la facilité de la création numérique provoque une multiplicité et une accélération de la création des DR. Bref, ces enjeux évoluent au fil de la transformation numérique, en raison de l'évolution et l'apparition de nouvelles technologies, et nécessiteront de nouvelles solutions en GDR.

2. Solutions et pistes à envisager

Pour relever les enjeux de sécurité liés aux DR, il est nécessaire d'envisager plusieurs pistes de solutions pour la réalisation d'une GDR efficace et adéquate. Cette dernière ne se limite pas uniquement à la seule dimension de la sécurité des données. Cette GDR permet que le partage de l'information y soit facilité par l'établissement de rôles et de responsabilités ainsi que de systèmes et d'infrastructures définis et adaptés aux besoins du projet. Ainsi, la confusion entre les différentes données aux différentes étapes de la

recherche y est réduite au strict minimum, ce qui permet d'éviter les fausses manipulations, le travail en double ou la perte de données. Conséquemment, une GDR efficace et adéquate fournit un avantage compétitif et concurrentiel dans le déroulement d'un projet de recherche, et facilite le travail collaboratif entre les chercheurs et les partenaires de la recherche.

2.1 Un niveau de sécurité adapté

Il faut rappeler que les solutions employées doivent être proportionnelles à la valeur et à la sensibilité des DR. Toutes les DR ne requièrent pas nécessairement un niveau de sécurité maximal avec des solutions coûteuses pouvant nuire à leur partage. Une évaluation de ces besoins en la matière est alors indispensable.

2.2 Une culture axée sur les pratiques exemplaires

Évidemment, pour assurer la sécurité des DR et, plus largement, une GDR efficace et adéquate, des infrastructures informatiques appropriées font partie des moyens à utiliser. À elles seules, ces infrastructures sont insuffisantes et peut même induire un faux sentiment de sécurité. Il faut avant tout miser sur l'humain dans son éventail de solutions en valorisant le développement d'une culture axée sur les pratiques exemplaires identifiées par son équipe, par son organisation et par ses partenaires. Si cela est plus facile à écrire qu'à faire, c'est toutefois une manière durable d'atteindre ses objectifs de R-D : augmenter sa cybersécurité, augmenter l'efficacité de son équipe de recherche en retrouvant rapidement ses données, diminuer les risques de mauvaises manipulations des données, faciliter la réutilisation des données, assurer la préservation des données, etc. Comme l'affirment Anthony et Cobban (2021) : « *Nothing changes unless people's behavior changes* » (para. 1). Il n'existe pas de solutions toutes faites pour y arriver.

Heureusement, quelques gestes simples peuvent à eux seuls faciliter la collaboration dès la production des DR : se doter d'un endroit commun de stockage de fichiers, d'une appellation normalisée de noms de fichiers pour l'ensemble des collaborateurs, documenter ses données au fur et à mesure du projet de recherche (contexte de collecte, méthode, variables utilisées) et favoriser une sauvegarde régulière pour contrer la perte de données.

2.3 La sensibilisation et la formation

Selon Wilms et ses collaborateurs (2020), il est important de convaincre les chercheurs que la GDR ne requiert pas nécessairement une charge de travail additionnelle. En effet, nombre de chercheurs croient ne pas faire de la GDR dans le cadre de leurs projets de recherche alors qu'au contraire ils effectuent une telle activité (Wilms et collab., 2020). Ces changements ne peuvent pas non plus survenir par la seule présence de politique en GDR et il faut considérer les craintes des chercheurs particulièrement lorsqu'il s'agit de partage des données (Wilms et collab., 2020). Les chercheurs doivent eux-mêmes participer, façonner les changements et décider de ces derniers.

Pour parvenir à ces changements, cela nécessite notamment de la sensibilisation et de la formation. Au sein des universités, plusieurs chercheurs n'ont pas de pratiques exemplaires en GDR (emploi d'espace de stockage inadéquat, documentation insuffisante ou manque de préservation adéquate des DR) (Trimble et collab., 2017).

Concernant les acteurs privés de la recherche, il est encore plus difficile de connaître leurs comportements en GDR dans la réalité. Toutefois, selon l'*Enquête canadienne sur la cybersécurité et le cybercrime de 2019* de Statistique Canada (2019a), au Canada, seulement 42,9 % du secteur privé[1] partageait les pratiques exemplaires et de l'information sur les risques en cybersécurité auprès des employés (en excluant ceux des services de technologies de l'information [STI]). Peu de formations formelles sont offertes par les entreprises canadiennes pour augmenter les compétences en lien avec la cybersécurité, que ce soit pour leurs employés sauf les STI (17,1 %) ou pour les parties prenantes comme les fournisseurs, les clients ou les partenaires (3 %) (Statistique Canada, 2019a). Les statistiques sur la formation et l'information offertes sur la cybersécurité sont beaucoup plus élevées dans les grandes entreprises que dans les petites entreprises. Le manque de compétences en cybersécurité des membres du personnel représente en lui-même une menace existante à la sécurité des DR. De plus, la sécurité des données ne représente qu'une facette de la GDR; il y a donc beaucoup à faire.

2.4 L'expertise des professionnels de l'information

Il apparaît pertinent de souligner l'expertise et le rôle des professionnels de l'information en GDR, soit les bibliothécaires et les archivistes. Ceux-ci peuvent notamment aider dans la formation et la sensibilisation en GDR, dans la rédaction d'un plan de gestion de données par les équipes de recherche, dans la diffusion des DR dans les dépôts de données, dans la préservation et la conservation des DR ainsi que dans le choix de vocabulaires spécialisés pour la description des ensembles de données.

Bien sûr, ils ne sont qu'un des maillons de soutien aux chercheurs dans la chaîne de la GDR. Plusieurs expertises dans les établissements de recherche doivent être mises à contribution, notamment celles du personnel des STI, des comités d'éthique, des administrateurs de la recherche, des juristes, des décideurs et des chercheurs eux-mêmes (liste non exhaustive). D'ailleurs, il existe un grand potentiel d'échange d'expertises entre ces différents acteurs ainsi qu'auprès des partenaires de recherche.

Conclusion

Cet article a mis l'accent sur les défis liés à la sécurité des données, tout en assurant une science ouverte et le partage des DR afin d'éviter le travail en vase clos et de favoriser une innovation ouverte. La multiplication et la valorisation de partenariats des acteurs publics et privés au sein des projets de recherche complexifient la GDR, comme par exemple le besoin d'assurer et de faciliter l'interopérabilité des systèmes, c'est-à-dire comment des systèmes différents d'organismes et de partenaires peuvent communiquer ou se transférer de l'information facilement. Dans les universités canadiennes, pour permettre le partage public des DR après la conclusion d'un projet scientifique, la plateforme Dataverse a été largement adoptée. Les DR sont réunies dans un moteur de recherche appelé Dépôt fédéré de données de recherche (DFDR) pour faciliter le repérage des ensembles de données. L'utilisation et la mise en place de systèmes de partage des DR durant la phase active de la recherche restent une question ouverte et largement inexplorée.

Aux enjeux technologiques s'ajoutent les enjeux de la gestion des différentes réglementations et politiques, des cultures organisationnelles ainsi que des responsabilités des acteurs concernant la GDR, qu'il s'agisse des équipes de recherche, des services de soutien dans les établissements de recherche et des partenaires de recherche. Cet enjeu de la gouvernance des données devient central et de plus en plus urgent. Intérêts communs et divergents, financement des infrastructures et des expertises, lutte de pouvoir et de représentativité ne sont que quelques questions que cet enjeu pose et auxquelles il faudra répondre tôt ou tard.

REMERCIEMENTS

L'auteur désire remercier Marie-Eve Ruest, responsable du développement des collections et de l'accès à l'information à l'Université du Québec à Chicoutimi, pour sa relecture de l'article et pour ses conseils.

NOTE

1 Dans ces statistiques, la définition du secteur privé inclut tous les services éducatifs (universités, cégeps, collèges, etc.).

RÉFÉRENCES

Alliance de recherche numérique du Canada. (s. d.). *Ce qui se passe chez nous*. https://alliancecan.ca/fr/la-noirn/ce-qui-se-passe-chez-nous

Alliance de recherche numérique du Canada. (2021, 31 mars). *Innovation, Sciences et Développement économique Canada renouvelle l'entente de contribution avec la Nouvelle organisation d'infrastructure de recherche numérique* [communiqué de presse]. https://alliancecan.ca/fr/dernier/innovation-sciences-et-développement-économique-canada-renouvelle-lentente-de-contribution-avec-la-nouvelle-organisation-dinfrastructure-de-recherche-numérique

Anthony, S. D. et Cobban, P. (2021, 25 novembre). 3 tactics to accelerate a digital transformation. *Harvard Business Review*. https://hbr.org/2021/11/3-tactics-to-accelerate-a-digital-transformation

Axelys. (2021, 5 octobre). *La chaîne de l'innovation, de l'idée au marché : conférence de MTL 2021* [vidéo]. YouTube. https://www.youtube.com/watch?v=zhQu00BvrQQ

Axelys. (s. d.). *Accueil*. https://www.axelys.ca/fr

Bureau de la Conseillère scientifique en chef du Canada. (2020, février). *Feuille de route pour la science ouverte*. https://www.ic.gc.ca/eic/site/063.nsf/fra/h_97992.html

Caire, É. (2021, 18 novembre). *Allocution de M. Éric Caire, ministre délégué à la Transformation numérique gouvernementale, ministre responsable de l'Accès à l'information et de la Protection des renseignements personnels et leader parlementaire adjoint du gouvernement*. Le grand rendez-vous de l'innovation québécoise, Montréal, QC. https://sync-stream.tv/rv-innovation-diffusion/

Collectif des 122 signataires. (2020, 30 janvier). *Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak* [communiqué de presse]. Wellcome. https://wellcome.org/press-release/sharing-data-and-findings-relevant-novel-coronavirus-ncov-outbreak

Cooper, A., Costanzo, L., Dearborn, D., Perry, C., Szwajcer, A. et Wang, M. (2021). *Sondage sur la capacité des services institutionnels de gestion de données de recherche – Avenir du soutien à la GDR pour les établissements : ressources priorisées, investissements, défis et accélérateurs*. Réseau Portage - Association des bibliothèques de recherche du Canada. https://doi.org/10.5281/ZENODO.4892718

Cooper, A., Perry, C., Szwajcer, A., Wang, M. et Khair, S. (2020). *Sondage sur la capacité des services institutionnels de gestion de données de recherche : sommaire*. Réseau Portage - Association des bibliothèques de recherche du Canada. https://dx.doi.org/10.14288/1.0388723

Conseil de recherches en sciences humaines du Canada (CRSH). (2021, 3 septembre). *Statistiques relatives aux concours : tableau de bord interactif*. https://www.sshrc-crsh.gc.ca/results-resultats/stats-statistiques/index-fra.aspx

Conseil de recherches en sciences humaines du Canada (CRSH), Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) et Instituts de recherche en santé du Canada (IRSC). (2021, 15 mars). *Politique des trois organismes sur la gestion des données de recherche*. https://science.gc.ca/eic/site/063.nsf/fra/h_97610.html

Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG). (2021a, 12 juillet). *Le gouvernement du Canada publie les lignes directrices sur la sécurité nationale pour les partenariats de recherche*. https://www.nserc-crsng.gc.ca/Media-Media/NewsDetail-DetailNouvelles_fra.asp?ID=1280

Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG). (2021b, 2 novembre). *Tableau de bord : investissements du CRSNG au Québec pour 2018-2019*. https://www.nserc-crsng.gc.ca/db-tb/index-fra.asp?year=2019&province=11&category=0

De Rosa, N. (2021, 21 juin). Hydroxychloroquine et COVID-19 : la nouvelle étude de Didier Raoult critiquée. *Ici Radio-Canada*. https://ici.radio-canada.ca/nouvelle/1798258/etude-hydroxychloroquine-azithromycine-didier-raoult-prepublication-hcq-azt-bien-entendu

Dion, M. (2020, 14 décembre). Fuite de données : Desjardins connaissait sa vulnérabilité, mais n'a rien fait. *Ici Radio-Canada*. https://ici.radio-canada.ca/nouvelle/1756980/fuite-donnees-desjardins-enquetes-commission-information-amf

COVID has shown the power of science-industry collaboration [éditorial]. (2021). *Nature*, *594*, 302. https://doi.org/10.1038/d41586-021-01580-0

Edwards, E. et Hilliyard, V. (2020, 23 mars). A man died after ingesting a substance he thought would protect him from coronavirus. *NBC News*. https://www.nbcnews.com/health/health-news/man-dies-after-ingesting-chloroquine-attempt-prevent-coronavirus-n1167166

Fonds de recherche du Québec (FRQ). (s. d.). *Science ouverte*. https://frq.gouv.qc.ca/science-ouverte

Groupe des sept (G7). (2021, 13 juin). *Pacte du G7 pour la recherche*. https://www.international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2021-06-13-research_compact-pacte_recherche.aspx?lang=fra

Halin, F. (2020, 7 décembre). Le gouvernement du Québec injecte 25 millions $ en intelligence artificielle. *Le Journal de Montréal*. https://www.journaldemontreal.com/2020/12/07/quebec-injecte-25m-en-intelligence-artificielle

Hanssen, M. (2021, 27 novembre). Semi-conducteurs : la prochaine crise viendra-t-elle de l'accès aux métaux stratégiques? *La Tribune* (France). https://www.latribune.fr/entreprises-finance/industrie/semi-conducteurs-l-acces-aux-metaux-strategiques-source-de-la-prochaine-crise-897132.html

Innovation, Sciences et Développement économique Canada. (2021a, 24 mars). *Énoncé de politique sur la sécurité de la recherche – printemps 2021 : déclaration*. Gouvernement du Canada. https://www.canada.ca/fr/innovation-sciences-developpement-economique/nouvelles/2021/03/enonce-de-politique-sur-la-securite-de-la-recherche--printemps2021.html

Innovation, Sciences et Développement économique Canada. (2021b, 12 juillet). *Le gouvernement du Canada agit pour protéger la recherche et la propriété intellectuelle au Canada*. Gouvernement du Canada. https://www.canada.ca/fr/innovation-sciences-developpement-economique/nouvelles/2021/07/le-gouvernement-du-canada-agit-pour-proteger-la-recherche-et-la-propriete-intellectuelle-au-canada.html

Innovation, Sciences et Développement économique Canada. (2021c, 16 août). *Infrastructure de recherche numérique : accueil*. https://ised-isde.canada.ca/site/infrastructure-recherche-numerique/fr/infrastructure-recherche-numerique

Innovation, Sciences et Développement économique Canada. (2021d, 12 juillet). *Protégez votre recherche : accueil*. https://www.ic.gc.ca/eic/site/063.nsf/fra/h_97955.html

Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. *Nature*, *561*, 163-166. https://doi.org/10.1038/d41586-018-06610-y

Ladepeche.fr avec AFP. (2021, 20 novembre). Didier Raoult accusé d'avoir falsifié des données : les Hôpitaux de Marseille ouvrent une enquête interne. *Ladepeche.fr*. https://www.ladepeche.fr/2021/11/20/didier-raoult-accuse-davoir-falsifie-des-donnees-lhopital-ouvre-une-enquete-interne-9939580.php

Lapointe, S. (2018, 2 octobre). Affaire MLS : pas le contenu de toutes les bases de données qui s'avère protégé par le régime des droits d'auteur. *Lesjuristes.ca*. https://lesjuristes.ca/fr/affaire-mls-la-jurisprudence-reconfirme-que-ce-nest-pas-le-contenu-de-toutes-les-bases-de-donnees-qui-savere-protege-par-le-regime-des-droits-dauteur

Libération. (2021, 19 novembre). Hydroxychloroquine : des collaborateurs de Didier Raoult dénoncent des falsifications, l'AP-HM ouvre une enquête. https://www.liberation.fr/societe/sante/hydroxychloroquine-des-personnels-de-lihu-de-marseille-denoncent-les-magouilles-de-didier-raoult-20211119_WOY5RXML7NDHRGYAWU43LWCUSM

Malboeuf, M.-C. (2017, 12 septembre). Les tricheurs de la science. *La Presse*. https://www.lapresse.ca/actualites/enquetes/201709/12/01-5132442-les-tricheurs-de-la-science.php

Malboeuf, M.-C. (2021, 18 juillet). Accès aux données de recherche : données sous haute sécurité. *La Presse*. https://www.lapresse.ca/actualites/2021-07-18/acces-aux-donnees-de-recherche/donnees-sous-haute-securite.php

Ministère de l'Économie et de l'Innovation du Québec. (2020, 10 décembre). *Québec se dote d'un Conseil de l'innovation pour accompagner les entreprises*. Gouvernement du Québec. https://www.quebec.ca/nouvelles/actualites/details/quebec-se-dote-dun-conseil-de-linnovation-pour-accompagner-les-entreprises

Nadeau, J.-B. (2021, 8 septembre). Le grand fouillis des données médicales. *L'actualité*. https://lactualite.com/sante-et-science/le-grand-fouillis-des-donnees-medicales

Premier ministre du Canada. (2018, 6 décembre). *Le premier ministre annonce un investissement dans l'intelligence artificielle pour créer plus de 16 000 nouveaux emplois pour les Canadiens*. https://pm.gc.ca/fr/nouvelles/communiques/2018/12/06/premier-ministre-annonce-investissement-lintelligence-artificielle

Quirion, R. (2016, 30 novembre). *Pour un meilleur accès aux données* [diapositives de présentation]. Colloque CIRANO-CIQSS. Faciliter l'accès aux données du Québec : comment et à quelles fins? Montréal, Canada. https://www.ciqss.org/sites/default/files/documents/2016-11-30_Remi-Quirion.pdf

Ravidsky, V. et Caulfield, T. (2021, 18 juin). Désinformation et mythes pendant la pandémie [balado, S1, ép. 8]. Dans *Espaces de courage*. Fondation Pierre Elliott Trudeau. https://www.fondationtrudeau.ca/activites/balados/desinformation-et-mythes-pendant-la-pandemie

Scassa, T. (2018). *Data Ownership*. CIGI Papers nº 187. Centre for International Governance Innovation. https://www.cigionline.org/static/documents/documents/Paper%20no.187_2.pdf

Schaefer, T., Brown, B., Graessle, F. et Salzsieder, L. (2017). Cybersecurity: Common risks. A dynamic set of internal and external threats includes loss of data and revenue, sabotage at the hands of current or former employees, and a PR nightmare. *Strategic Finance*, *99*(5), 54-61.

Siddik, M. A. B., Shehabi, A. et Marston, L. (2021). The environmental footprint of data centers in the United States. *Environmental Research Letters*, *16*(6), 064017. https://doi.org/10.1088/1748-9326/abfba1

Statistique Canada. (2019a). *Enquête canadienne sur la cybersécurité et le cybercrime* [ensemble de données]. SERENE-RISC. https://www.serene-risc.ca/fr/statistique-canada

Statistique Canada. (2019b, 20 décembre). *Centres de données de recherche*. https://www.statcan.gc.ca/fr/microdonnees/centres-donnees

Tchandem Kamgang, A. C. (2019, 23 juillet). Equifax frappé d'une amende de 750 M$: sa fiabilité remise en cause? *Radio Canada International*. https://www.rcinet.ca/fr/2019/07/22/vol-des-donnees-confidentielles-sanction-contre-equifax-equifax-et-protection-des-dossiers-de-clients-desjardins-cybersecurite-federal-trade-commission-desjardins-et-vol-des-informations-privees

Intergovernmental Panel on Climate Change. (2015). *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf

Trimble, L., Dearborn, D., Zaraiskaya, T., Burpee, J., Barsky, E., Sahadath, C., Cheung, M., Mitchell, M. et Gertler, M. (2017). *Across Canada, across disciplines: Research data management practices and needs in the social sciences and humanities* [diapositives de présentation]. IASSIST 2017 Conference, Lawrence, KS. https://doi.org/10.14288/1.0348075

University of Oxford. (2020, 30 avril). *Oxford University announces landmark partnership with AstraZeneca for the development and potential large-scale distribution of COVID-19 vaccine candidate*. https://www.ox.ac.uk/news/2020-04-30-oxford-university-announces-landmark-partnership-astrazeneca-development-and

Waldrop, T., Alsup, D. et McLauglin, E. C. (2020, 25 mars). Fearing coronavirus, Arizona man dies after taking a form of chloroquine used in aquariums. *CNN*. https://www.cnn.com/2020/03/23/health/arizona-coronavirus-chloroquine-death/index.html

Wilms, K. L., Stieglitz, S., Ross, B. et Meske, C. (2020). A value-based perspective on supporting and hindering factors for research data management. *International Journal of Information Management*, *54*, 102174. https://doi.org/10.1016/j.ijinfomgt.2020.102174

World Economic Forum. (2019). *Data science in the new economy: A new race for talent in the fourth industrial revolution.* https://www3.weforum.org/docs/WEF_Data_Science_In_the_New_Economy.pdf