



Septembre 2021

Évaluation de l'état actuel des logiciels de recherche

Rédigé par le groupe de travail de l'Alliance responsable des logiciels de recherche de pointe :

Qian Zhang, Susan Brown, James Colliander, Brian Corrie, Gabor Fichtinger, Scott Henwood, Mark Leggott, Catherine Lovekin, Felipe Pérez-Jvostov, Ghilaine Roquet, and Marc-Étienne Rousseau



**Alliance de recherche
numérique** du Canada

**Digital Research
Alliance** of Canada

Financé par le
gouvernement
du Canada

Canada 

1. SOMMAIRE	5
1.1. LE PAYSAGE CANADIEN DES LR	6
1.2. PRINCIPAUX DÉFIS ET OPPORTUNITÉS	8
1.3. LES PROCHAINES ÉTAPES	12
2. INTRODUCTION	13
3. SURVOL DES LR	15
3.1. QUE SONT LES LR ?	15
3.2. LR, CYCLES DE VIE DES DONNÉES ET DE LA RECHERCHE	16
3.3. CARACTÉRISTIQUES DES LR	21
3.4. TYPES DE LR	23
3.5. LES LR DANS LE MOUVEMENT DE LA SCIENCE OUVERTE	26
3.6. DÉCRIRE ET MESURER LES IMPACTS DES LR	27
3.6.1. <i>Métadonnées de LR</i>	27
3.6.2. <i>Licences pour les LR</i>	28
3.6.3. <i>LR et reproductibilité de la recherche</i>	31
3.6.4. <i>Maintien et pérennité à long terme des LR</i>	43
4. MAIN-D'ŒUVRE POUR LES LR	45
5. LR AU CANADA	47
5.1. LR DANS L'IRN ET L'ALLIANCE	47
5.2. FOURNISSEURS DE SERVICES	48
5.2.1. <i>CANARIE</i>	48
5.2.2. <i>Fédération Calcul Canada (FCC)</i>	52
5.2.3. <i>Hébergement et soutien institutionnel</i>	54
5.2.4. <i>Secteur privé</i>	57
5.2.5. <i>Revue</i>	58
5.3. BAILLEURS DE FONDS POUR LA CRÉATION ET L'UTILISATION DES LR	58
5.3.1. <i>À l'échelle nationale</i>	58
5.3.2. <i>À l'échelle régionale/provinciale</i>	63
5.3.3. <i>Sociétés et fondations (publiques, privées et internationales)</i>	64
6. LR AU-DELÀ DU CANADA (ANNEXE I)	66

7.	TENDANCES ÉMERGENTES D'IRN - ASPECT DES LR.....	67
7.1.	PERSPECTIVE CULTURELLE ET SOCIÉTALE.....	67
7.1.1.	<i>Principes FAIR</i>	68
7.1.2.	<i>Formation et généralisation de l'utilisation des outils</i>	69
7.1.3.	<i>Biais logiciel et algorithmique</i>	69
7.1.4.	<i>Science citoyenne</i>	70
7.1.5.	<i>Les ILR en tant professionnels</i>	70
7.2.	ÉVOLUTION TECHNOLOGIQUE.....	71
7.2.1.	<i>AM et IA</i>	71
7.2.2.	<i>Infonuagique</i>	72
7.2.3.	<i>L'infrastructure en tant que code</i>	73
7.2.4.	<i>Internet des objets (IdO)</i>	74
7.2.5.	<i>Informatique en périphérie</i>	75
7.2.6.	<i>Informatique quantique</i>	76
7.2.7.	<i>Analyse des données</i>	76
7.2.8.	<i>Cybersécurité</i>	77
7.2.9.	<i>Reproductibilité de la recherche</i>	78
7.2.10.	<i>Supercalculateurs exaflopiques et LR</i>	78
8.	ÉVALUATION DE L'ÉTAT ACTUEL.....	79
8.1.	POINTS FORTS.....	79
8.1.1.	<i>PHQ</i>	79
8.1.2.	<i>Financement et partenariats nationaux</i>	80
8.2.	DÉFIS ET OPPORTUNITÉS.....	83
8.2.1.	<i>Manque de financement ciblé et durable</i>	83
8.2.2.	<i>Technique</i>	88
8.2.3.	<i>Compétences, connaissances et formation</i>	90
8.2.4.	<i>Attraction et rétention du PHQ</i>	91
8.2.5.	<i>Juridique/Politique</i>	94
8.2.6.	<i>Sécurité et confidentialité</i>	95
8.2.7.	<i>Collaboration internationale</i>	98

8.2.8.	<i>Mesures incitatives et indicateurs</i>	99
8.2.9.	<i>Culture de la recherche</i>	100
8.2.10.	<i>Politique et stratégie du gouvernement</i>	101
8.2.11.	<i>EDI et représentation</i>	101
9.	PROCHAINES ÉTAPES	102
9.1.	ÉVALUATION DES BESOINS DES CHERCHEURS	103
9.2.	ÉLABORATION D'UN MODÈLE NATIONAL DE PRESTATION DE SERVICES ET D'UN MODÈLE DE FINANCEMENT	103
9.3.	PLAN STRATÉGIQUE DE L'IRN POUR 2022-24	103
	ACRONYMES ET GLOSSAIRE	104
	ANNEXE A : EXEMPLES DE PORTAILS SCIENTIFIQUES AU CANADA.....	107
	ANNEXE B : SOUTIEN ET SERVICES DE LR FOURNIS DANS LE CADRE DES ACTIVITÉS DE RECHERCHE QUOTIDIENNES	108
	ANNEXE C : ARCHÉTYPES D'UTILISATEURS ET DE COMMUNAUTÉS DANS L'ÉCOSYSTÈME RS	112
	ANNEXE D : TYPES DE LR.....	113
	ANNEXE E : SCHÉMAS DE MÉTADONNÉES DE LR REPRÉSENTATIFS	114
	ANNEXE F : BONNES PRATIQUES LIÉES AUX LR.....	115
	F1 : BONNES PRATIQUES DE GESTION DES LR	115
	F2 : BONNES PRATIQUES EN MATIÈRE DE REPRODUCTIBILITÉ INFORMATIQUE (EN TANT QUE CHERCHEURS OU AUTEURS DE DONNÉES/CODES).....	117
	ANNEXE G : LA GESTION DES LR À TRAVERS LE PRISME DE LA COVID-19	119
	ANNEXE H : BAILLEURS DE FONDS POUR LE DÉVELOPPEMENT ET L'UTILISATION DES LR.....	122
	ANNEXE I : PAYSAGE DE LA COMMUNAUTÉ DE LR.....	130

1 Sommaire

En tant que premier exposé de position sur les logiciels de recherche (LR) au Canada, le présent rapport examine et résume les LR en tant que domaine émergent et en tant que secteur de professionnalisation, sur le plan national et international, et documente les forces, les défis et les possibilités au sein de l'écosystème actuel des LR, dans la mesure où ils se rapportent à l'Alliance de recherche numérique du Canada (l'Alliance). Le but de ce document est d'encadrer la compréhension du paysage des LR et d'établir un cadre général de discussion au Canada.

Ce rapport est complémentaire aux documents d'évaluation de l'état actuel de la gestion des données de recherche (GDR) au Canada¹ et d'évaluation de l'état actuel du calcul informatique de pointe (CIP) au Canada² actualisant les exposés de position de 2017 sur la GDR et le CIP soumis à Innovation, Sciences et Développement économique Canada (ISDE) par le Conseil du leadership sur l'infrastructure de recherche numérique (CLIRN). À l'époque, les LR étaient reconnus comme des éléments essentiels du paysage numérique ; toutefois, le domaine n'avait pas la maturité et les communautés structurées permettant la production des deux autres exposés. Ce rapport sur les LR représente donc la première évaluation de l'état actuel des LR au Canada.

Le présent rapport permet à l'Alliance de comprendre l'état actuel et d'en tirer parti, ainsi que de faciliter une stratégie faisant progresser les LR de concert avec d'autres éléments de l'infrastructure de recherche numérique (IRN) afin de soutenir l'excellence en recherche au Canada. Les conclusions et les observations contenues dans ce document ainsi que les publications sur l'évaluation de l'état actuel de la GDR et du CIP ont pour but de fournir des renseignements de base aux analystes et à la direction, au conseil d'administration et au conseil des chercheurs de l'Alliance afin de soutenir l'élaboration du nouveau modèle de prestation de services, la planification stratégique et le modèle de prestation de financement de l'Alliance.

Les LR sont multidimensionnels et internationaux. Dans le cadre de ce document, les LR sont définis d'un point de vue général comme un recoupement émergent et complexe d'outils, de disciplines, de services, de plateformes, de matériel, de ressources et de personnes (utilisateurs, chercheurs, développeurs, intervenants, personnel, communautés, etc.). Les LR sont fondamentaux pour la recherche, étant donné leur rôle pour aider les chercheurs à faire de nouvelles découvertes et à partager leurs résultats avec d'autres. Les tendances technologiques et culturelles de l'IRN entraînent une évolution continue et rapide des LR, ce qui les rend encore plus essentiels pour faire avancer la recherche. La transformation vers une culture de LR axée sur la science ouverte³ dépend de la création d'outils, de plateformes et de services permettant aux chercheurs de mobiliser les connaissances et de rendre les processus de recherche plus

¹ https://engagedri.ca/assets/documents/NIRNO_GDR_PositionPaper.pdf

² Document non publié

³ Dans ce document, la science ouverte est un terme général qui désigne un large éventail d'activités et de formes d'ouverture dans l'environnement universitaire et de recherche qui s'appliquent à toutes les disciplines, notamment le libre accès, les données ouvertes, les sources ouvertes, la recherche reproductible ouverte, la gouvernance ouverte, l'évaluation scientifique ouverte, l'examen par les pairs ouvert, les ressources éducatives ouvertes, la collaboration ouverte, l'éducation ouverte, la science citoyenne, etc.

efficaces, transparents, reproductibles et adaptés aux défis sociétaux. Les éléments particuliers de ce changement sont les suivants : l'augmentation de la collaboration et de l'interaction entre les chercheurs ; le développement d'une infrastructure technique qui favorise l'adoption de pratiques de recherche émergentes ; l'élaboration, la promotion et l'adoption de pratiques scientifiques ouvertes utilisant le code source ouvert. Pour que le Canada puisse relever les défis du 21^e siècle et demeurer concurrentiel sur le plan international, tous ces changements exigent un écosystème agile et réactif, doté d'un solide soutien en personnel hautement qualifié (PHQ) et d'un financement durable.

Les LR sont des composantes d'IRN étroitement liés au CIP et à la GDR, comme on peut le voir dans la stratégie du gouvernement fédéral en matière d'IRN, que l'on analyse les caractéristiques, les types, les fonctions, les besoins ou l'incidence des LR sur la recherche numérique. Dans l'écosystème d'IRN, les éléments fondamentaux des LR, de la GDR, du CIP et de la cybersécurité sont fondamentalement interconnectés et interdépendants. Si l'on considère l'entreprise de recherche à travers le prisme du cycle de vie des LR, des données et de la recherche, on constate des interdépendances et des intersections essentielles à la mise en œuvre d'une approche efficace et durable quant à l'IRN s'appuyant fortement sur les LR.

1.1. Le paysage canadien des LR

Le présent rapport identifie les principaux acteurs du paysage canadien sur le plan local, régional, provincial, national et international, et examine les multiples rôles qu'ils jouent au sein des composantes clés nécessaires pour soutenir les LR sur le plan national.

Avec son programme de financement des LR, CANARIE a joué un rôle essentiel dans l'avancement de l'écosystème canadien des LR. Le programme Logiciels de recherche de CANARIE finance le développement des LR dans les communautés à forte intensité de données et de calcul ainsi que dans les domaines non traditionnels, accélérant la découverte de la recherche en permettant l'accès aux IRN, et favorisant les bonnes pratiques (par exemple, en favorisant la réutilisation des LR pour éviter la réinvention de logiciels [L]). Le portail de logiciels de recherche de CANARIE permet d'accéder aux logiciels de recherche qui ont été financés par CANARIE ainsi qu'aux contributions de la communauté. CANARIE organise chaque année le Colloque canadien sur les logiciels de recherche (CCLR) et participe activement à la communauté internationale des logiciels de recherche. L'initiative Soutien local aux logiciels de recherche (SLRL), lancée récemment par CANARIE, finance le développement et le maintien d'équipes locales de LR afin de faciliter l'accès aux outils et à l'expertise en matière de LR sur le plan institutionnel. Depuis sa création en 2007, le programme Logiciels de recherche a octroyé 50,7 millions de dollars à des équipes de LR au Canada et a mis l'accent sur le développement de LR, l'établissement de pratiques de LR et la création d'une culture et d'une communauté naissante d'ingénieurs en logiciel de recherche (ILR) au Canada⁴. Le plus récent programme Gestion des données de recherche (GDR) de CANARIE investit dans des projets de recherche axés sur la GDR, notamment en finançant le développement de logiciels. CANARIE transférera les programmes de LR et de GDR à l'Alliance d'ici le 31 mars 2022.

⁴ Toutes les valeurs en dollars de ce rapport sont en dollars canadiens, sauf indication contraire.

La Fondation canadienne pour l'innovation (FCI) a également fourni un financement axé sur le développement de LR. Les défis de la FCI en matière de cyberinfrastructure en 2015 et 2017 ont favorisé la conception et le développement de plateformes de LR en incluant le soutien de la Fédération Calcul Canada (FCC) pour améliorer la capacité de recherche impliquant des chercheurs de consortiums de différents établissements, des scientifiques de données et des développeurs de LR. Il s'agissait des premières incursions de la FCI dans le soutien aux plateformes de LR et elles ont été réalisées en réponse à la prise de conscience croissante des défis que pose le maintien de l'infrastructure de logiciel. La FCI a également accordé le Fonds de leaders John R. Evans (FLJE), un programme de subventions de partenariat qui répond aux besoins des chercheurs individuels en matière d'infrastructure de recherche.

Un autre fournisseur national clé de services d'IRN est la FCC, qui fournit aux chercheurs l'infrastructure et les services de CIP, y compris un environnement unifié de LR. Ce cadre national d'IRN permet aux systèmes informatiques indépendants de tous les sites d'installer et de reproduire la même collection de logiciels (également appelée dépôt de logiciels de la FCC) comprenant plus de 3 000 LR d'applications et d'y accéder de manière fiable et efficace. Avec plus de 200 experts basés dans 37 établissements partenaires au pays, la FCC fournit un soutien direct aux chercheurs canadiens en offrant des consultations, un soutien expert et des formations régionales sur de nombreux sujets. Leurs services couvrent un grand éventail de domaines allant de l'allocation des ressources au mouvement des données, et de la visualisation aux sciences humaines numériques. En plus du concours régulier d'accès aux ressources pour le stockage et le calcul et d'allocation de celles-ci, les groupes qui développent des logiciels de recherche à grande échelle peuvent soumettre leur candidature au concours des plateformes et portails de recherche (PPR) de la FCC, qui est davantage axé sur les besoins et les exigences en matière de LR, offrant un soutien direct aux LR et aux infrastructures de LR pour les grandes initiatives scientifiques via le déploiement et le développement de leurs propres plateformes sur les infrastructures de CIP de la FCC. Les ressources distribuées de CIP de la FCC et le PHQ appuient la gamme complète des programmes de recherche, des programmes individuels axés sur les chercheurs principaux (CP) aux nombreuses initiatives scientifiques majeures du Canada.

Tout comme la grande diversité des domaines de recherche, le soutien des établissements aux LR (ainsi que du CIP et de la GDR) est caractérisé par une diversité selon la taille, l'histoire, l'orientation de la recherche, les ressources, le type d'efforts et une foule d'autres facteurs qui forment un paysage complexe. Par exemple, les collèges et les établissements favorisent les partenariats avec l'industrie ; les laboratoires de recherche nationaux mettent en œuvre leur propre technologie de l'information (TI) et leur propre soutien en matière de LR tout en travaillant avec des collaborateurs universitaires ; et les hôpitaux de recherche développent des LR hautement spécialisés combinant la recherche et la pratique clinique, ce qui soulève des questions de sécurité et de confidentialité. Cette diversité a été prise en compte dans les contextes de CIP et de GDR grâce à un réseau de soutien distribué qui s'est développé au cours des deux dernières décennies. Bien qu'il n'existe pas de réseau distribué semblable dans le domaine des LR, une communauté est en train d'émerger grâce aux efforts de CANARIE (p. ex. l'initiative SLRL⁵), de la FCC et de divers groupes de domaines qui contribuent à une approche concertée quant au développement des LR.

⁵ <https://www.canarie.ca/fr/program-news/logiciels-de-recherche/>

Le soutien aux programmes de recherche comportant des exigences de LR au Canada a souvent pris la forme de programmes d'emploi pour les étudiants et le personnel hautement qualifié récemment diplômé en informatique, en génie logiciel et dans d'autres disciplines. Il existe une forte tradition de collaboration entre l'industrie et les établissements par le biais de programmes coopératifs, renforcée par Mitacs⁶ et d'autres incitatifs gouvernementaux. Les activités de CANARIE et des autres organisations ont jeté des bases solides pour la professionnalisation du développement des LR et des ILR afin de soutenir le vaste bassin de talents en développement de LR du pays. Le soutien à la souveraineté des données autochtones dans le budget fédéral de 2021⁷ a des répercussions sur les LR, ce qui est renforcé par l'engagement du Conseil de recherches en sciences humaines⁸ (CRSH) à financer la recherche par les peuples autochtones et avec ceux-ci, ce qui améliore grandement les perspectives de souveraineté des données autochtones.

La plupart des autres organismes nationaux de financement se concentrent sur le financement de la recherche pure ou appliquée, et non pas explicitement sur le développement de logiciels ou d'infrastructures de recherche : le soutien aux LR est indirect et passe généralement par la recherche dans un domaine particulier. De même, aucun financement consacré aux LR n'est disponible auprès des organismes de financement provinciaux ou régionaux. Les bailleurs de fonds privés et les donateurs financent généralement la recherche dans des disciplines précises, plutôt que dans les LR en particulier, bien qu'il y ait des exceptions (p. ex. l'initiative Chan Zuckerberg, le Wellcome Trust). Une augmentation récente des programmes de collaboration offerts par des organismes internationaux et canadiens (p. ex. l'appel conjoint Horizon 2020 de l'UE/IRSC⁹) a permis de financer certains efforts de LR au Canada. Enfin, la récente Initiative des Supergrappes d'innovation pourrait transformer l'innovation régionale et promouvoir la collaboration à grande échelle entre l'industrie et le milieu universitaire.

1.2. Principaux défis et opportunités

Le financement ciblé et durable des LR ne bénéficie pas d'un soutien suffisant.

Le financement traditionnel de la recherche est fondamentalement axé sur l'innovation, laissant le besoin de maintenir les LR à long terme à la capacité des chercheurs principaux de décrire leurs besoins dans un contexte d'innovation. Cela représente une lacune et un défi importants pour les chercheurs canadiens, entraînant une perte d'investissement lorsque les LR ne sont pas maintenus ou généralisés pour être réutilisés, et diminuant le rôle de chef de file du Canada dans le développement des LR. Une nouvelle approche quant au financement des LR doit reconnaître les différents types et phases des LR (p. ex. expérimentaux, émergents mais au niveau de la

⁶ <https://www.mitacs.ca/fr>. MITACS est une organisation nationale à but non lucratif qui s'associe à des établissements canadiens pour mener des recherches et des programmes de formation dans des domaines liés à l'innovation industrielle et sociale.

⁷ <https://www.budget.gc.ca/2021/pdf/budget-2021-fr.pdf>

⁸ <https://www.sshrc-crsh.gc.ca/home-accueil-fra.aspx>

⁹ Horizon 2020 de l'UE désigne le [programme-cadre de l'UE Horizon 2020](#) ; IRSC est l'abréviation d'Instituts de recherche en santé du Canada. Pour en savoir plus : <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/sc1-bhc-05-2018>

production, établis/entreprise), et concevoir des mécanismes d'évaluation, des mesures et des voies de financement appropriés pour chacun.

Par rapport d'autres pays, le Canada accuse un retard dans le soutien d'une communauté de développement de LR efficace, cohésive et concertée.

À quelques exceptions notables près, les LR au Canada sont généralement considérés comme des sous-produits de la recherche dans un domaine particulier, ils ne bénéficient d'un financement dédié et durable et n'ont été reconnus que récemment comme des produits de « première classe » ou une composante essentielle d'une stratégie nationale d'IRN. En revanche, l'Australie et l'Europe ont investi des sommes importantes dans un effort ciblé pour créer des communautés de développement de LR solides dans toutes les disciplines.

L'écosystème des LR est varié et complexe et évolue rapidement.

Tous les logiciels dépendent des autres composants logiciels (ou dépendances) pour leur fonctionnement, et ceux-ci sont divers (par exemple, les systèmes d'exploitation, les systèmes de gestion de bibliothèque, les paquets, les logiciels Web, les EDI, les API). L'écriture de LR pour des plateformes complexes de CIP ou d'infonuagique peut être très difficile, car les programmes doivent être superposés pour exploiter au mieux l'infrastructure sous-jacente. Les services nationaux ont un rôle essentiel à jouer à cet égard, que ce soit en proposant des LR interopérables et réutilisables sur toutes les plateformes ou en soutenant l'éducation et la formation. Les LR sont développés dans le monde entier. Le taux de changement dans le développement des logiciels et la recherche en général est incroyablement rapide : les LR doivent donc évoluer à un rythme semblable en tenant compte des développements internationaux.

Les LR ne sont pas largement diffusés ou partagés et ne sont pas faciles à découvrir, ce qui nuit à la transparence, à la reproductibilité et à la vérification de la recherche.

La plupart des initiatives de LR s'appuient sur des dépôts commerciaux de logiciels de bonne pratique tels que GitHub. Compte tenu de la prolifération des outils et des plateformes de LR, l'augmentation de l'innovation et de l'infrastructure pour répondre aux exigences des diverses communautés de recherche est difficile sans la coordination et le soutien sur le plan national. Il faut donc aider les chercheurs à consigner leur code et leur documentation dans des dépôts de LR accessibles qui assurent le catalogage, l'indexation, la préservation et la curation afin que les LR du Canada puissent devenir une composante novatrice et durable du processus de recherche.

Il faut remédier au manque de compétences en matière de développement de logiciels à tous les stades de l'enseignement et de la recherche.

Il y a une pénurie de formation en LR, avec un besoin particulier de formation en LR dans les domaines qui sont plus récents dans l'utilisation d'IRN à l'appui de leur recherche. La formation communautaire, disciplinaire et régionale au-delà des programmes demeure donc d'une importance capitale. Une formation continue de la sorte est également nécessaire, car les technologies sous-jacentes de LR (et de CIP) sont dynamiques et en constante évolution. Cela ne s'applique pas seulement aux domaines où il n'y a pas eu d'utilisation courante des LR ou d'incorporation du développement des LR, mais aussi à ceux nécessitant une amélioration constante des compétences du PHQ. On pourrait parler d'éducation de manière plus unifiée, en incluant à la fois la formation des ILR et la formation en LR pour tous les chercheurs. Il faut aussi faire plus de sensibilisation sur les bonnes pratiques en matière de pérennité des LR, ce qui peut

être facilité par des organisations agissant comme des points focaux pour l'expertise, non seulement pour partager les connaissances et les compétences, mais aussi pour améliorer la mise en réseau et la collaboration internationales.

La reconnaissance du personnel hautement qualifié (notamment les chercheurs, les étudiants, les postdoctorants et le personnel de soutien à la recherche) dans le contexte des LR est essentielle pour faire évoluer la recherche à l'ère numérique.

« L'ingénieur en logiciel de recherche (ILR)¹⁰ est un titre d'emploi récemment créé qui reconnaît la mesure dans laquelle l'efficacité et les résultats de la recherche sont améliorés par l'intégration de personnel hautement qualifié dans les équipes de recherche. Compte tenu de la demande importante en ILR dans tous les domaines de pratique et du fait que les LR développés dans un domaine peuvent avoir des applications dans d'autres disciplines, il faut élaborer un modèle de financement complet pour les ILR. Bien qu'il existe des communautés d'ILR naissantes dans la plupart des régions, le Canada doit néanmoins déployer des efforts ciblés pour établir à la fois une communauté nationale d'ILR et un parcours de carrière stable pour les ILR au sein de leur établissement d'accueil et ailleurs. Il est nécessaire de mettre en place une solide stratégie de recrutement et de maintien en poste du personnel hautement qualifié (PHQ) en recherche et développement surtout pour les ILR.

Les mesures incitatives liées aux paramètres, au financement, aux récompenses et à la reconnaissance, ainsi qu'à l'avancement professionnel, sont essentielles pour catalyser l'engagement envers les LR.

Tous les intervenants ont la possibilité d'adopter des politiques et des programmes qui reconnaissent le rôle important que jouent les LR dans l'obtention de résultats de recherche dans toutes les disciplines. Des programmes et des mesures d'incitation sont nécessaires pour encourager les chercheurs et les professionnels techniques de la recherche (PTR) à suivre les bonnes pratiques en matière de développement de LR. Les décideurs ont la possibilité de soutenir des initiatives visant à développer des systèmes d'évaluation de la recherche qui récompensent les LR au même titre que les autres résultats de recherche, en réagissant de manière proactive lorsque les bonnes pratiques ne sont pas mises en œuvre. Des mesures appropriées de l'activité de LR faciliteraient le changement de culture nécessaire pour réaliser la promesse de la science ouverte.

Le manque de diversité dans le milieu des LR suggère que nous avons besoin d'une approche plus efficace d'équité, de diversité et d'inclusion (EDI) dans le contexte des LR.

Le manque d'EDI dans ce contexte découle de divers facteurs, notamment le faible taux d'inscription des femmes et des personnes noires, indigènes et d'autres personnes de couleur (BIPOC) dans les programmes axés sur la technologie dans l'enseignement supérieur, ou dans les programmes de développement de logiciels au baccalauréat, à la maîtrise, au doctorat ainsi que le parcours généralement étroit vers l'avancement des femmes et des personnes BIPOC en recherche et en développement. La sensibilisation des intervenants à la promotion de l'EDI dans la profession de LR devrait conduire à la création et à l'amélioration des programmes pour élargir les possibilités et éliminer les obstacles pour les membres des communautés sous-représentées

¹⁰ Même si nous utilisons ce terme communément admis dans ce document, au Canada, les ILR ne sont pas nécessairement des ingénieurs agréés.

qui s'intéressent aux LR. L'établissement de partenariats avec des initiatives de formation technique et de réseautage pour les groupes sous-représentés, ainsi que le soutien de programmes de formation axés sur l'EDI liés à l'utilisation et au développement des LR au sein de communautés de recherche particulières contribueront à élargir le bassin de talents. Il sera donc crucial de créer des politiques, des systèmes et des flux de travail dans les communautés et les équipes de recherche, ainsi que dans les services de ressources humaines (RH) des universités, qui protègent les populations vulnérables et s'attaquent au patriarcat, au colonialisme et au racisme systémiques.

À l'heure actuelle, le Canada ne dispose pas de politiques, de normes et de protocoles bien élaborés pour aider les chercheurs de toutes les disciplines à gérer les LR.

Cette situation nuit à la capacité des chercheurs de tirer parti de l'énorme potentiel des LR et réduit la compétitivité du Canada sur le plan international. Un cadre stratégique national compatible avec les pratiques mondiales est nécessaire pour guider le développement durable des LR au Canada.

Une approche concertée doit être établie pour assurer la sécurité et l'intégrité de la recherche et protéger les intérêts canadiens.

Les équipes de LR sont généralement mal placées pour comprendre la prolifération des menaces à la cybersécurité ou pour y réagir. Par exemple, la conception de solutions de LR pour les passerelles scientifiques¹¹ qui traitent des données sur la santé exige des compétences particulières et une expertise dans les bonnes pratiques de développement de logiciels sécurisés, ainsi qu'une compréhension de la politique de confidentialité des données. Les enjeux de cybersécurité ne cessant de croître, il est essentiel d'élaborer une stratégie nationale solide en matière de cybersécurité qui dépasse le cadre d'une seule équipe de recherche, d'un seul fournisseur de services ou d'une seule université.

Il faut trouver un équilibre entre la gestion de la propriété intellectuelle (PI) et le développement d'une culture dans laquelle le partage de l'information est préconisé.

Les communautés de recherche individuelles sont les mieux placées pour favoriser cette culture, car les normes varient selon la discipline. L'objectif ultime est de protéger adéquatement les résultats de la recherche tout en favorisant une approche ouverte et collaborative de la recherche qui fait progresser la recherche et le développement du Canada à l'échelle mondiale.

Les organismes de recherche fédéraux et provinciaux doivent élaborer une stratégie pour travailler davantage en collaboration avec l'enseignement supérieur et les différents gouvernements.

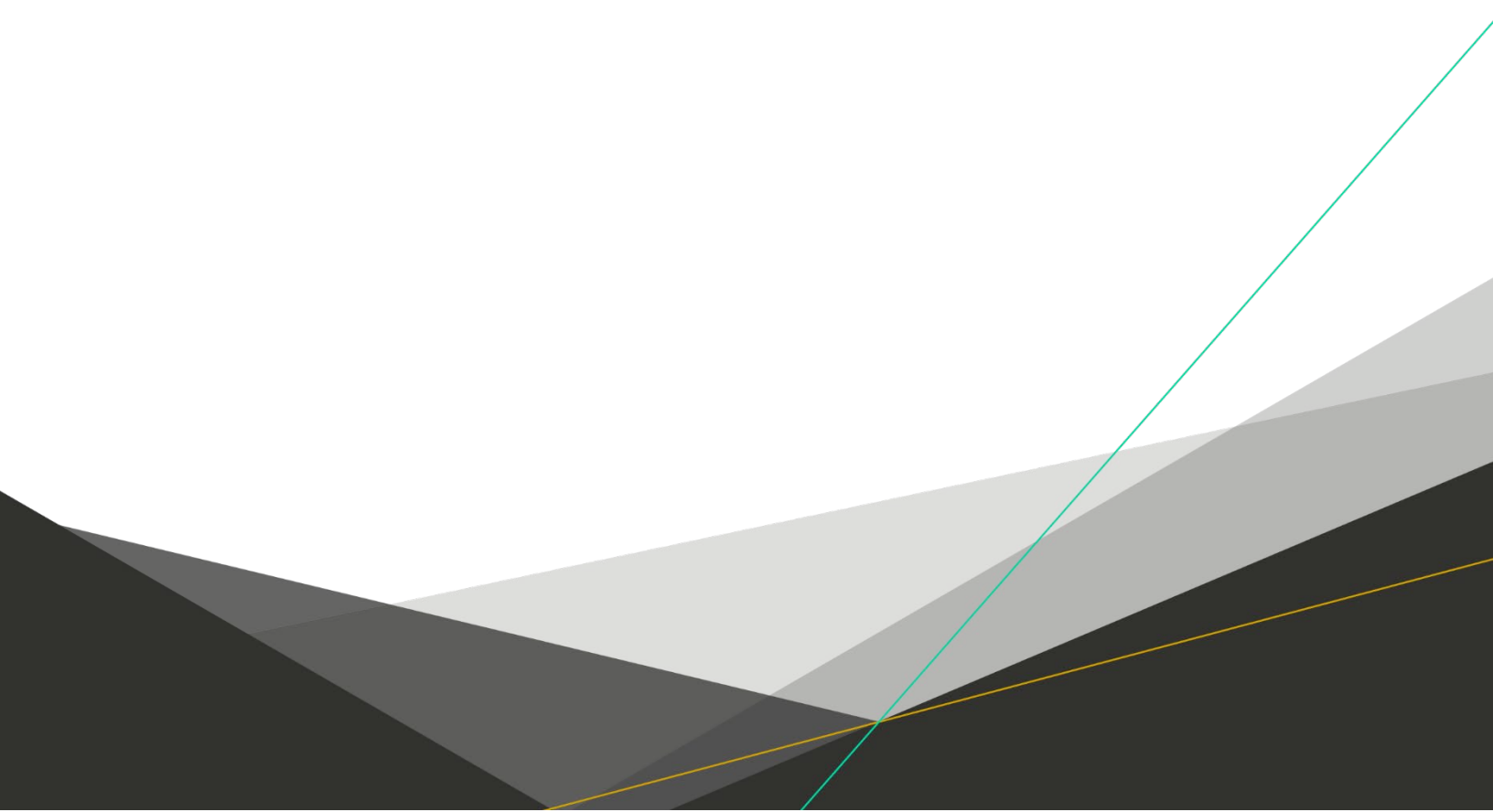
Le grand nombre d'utilisateurs et de communautés, la complexité de leurs rôles et responsabilités (internationaux, nationaux, régionaux et locaux), le nombre de régions impliquées et la diversité des exigences avec lesquelles ils doivent aligner leurs efforts mettent en évidence les défis et la nécessité d'une telle collaboration. Bien que l'on assiste à l'émergence d'un leadership national coordonné en LR, les efforts visant à cristalliser les communautés de LR ont été entravés par le

¹¹ <https://doi.org/10.1016/j.future.2018.12.026>

manque de financement adéquat et de mandat officiel. Sans coordination des investissements dans le contexte des LR, il est difficile de développer les politiques, processus, protocoles, bonnes pratiques et normes partagés qui sont si essentiels.

1.3. Les prochaines étapes

Pour être prêt à relever les défis actuels et à saisir les opportunités potentielles, il est nécessaire de procéder à une analyse approfondie de la situation et du paysage actuels, en accord avec les buts et les objectifs à court et à long terme. Le présent rapport résume l'état actuel du paysage des LR au Canada afin de favoriser une compréhension commune parmi les membres de l'Alliance de l'ampleur et de la complexité de l'engagement des intervenants dans ce domaine ; il sert également de base à partir de laquelle l'Alliance peut établir une voie à suivre pour la stratégie nationale de LR au Canada. Les conclusions et les observations contenues dans ce document ainsi que dans les publications sur l'état actuel de la GDR et du CIP ont pour but d'aider à soutenir le processus d'évaluation des besoins des chercheurs, l'élaboration du nouveau modèle de prestation de services et du modèle de financement de l'Alliance, ainsi que les processus du plan stratégique de l'IRN.



2 Introduction

En tant que premier exposé de position sur les LR au Canada, ce rapport examine et résume les LR en général en tant que domaine émergent et en tant que secteur de professionnalisation, sur le plan national et international, et documente les forces, les défis et les possibilités au sein de l'écosystème actuel des LR. Le but de ce document est d'encadrer la compréhension du paysage des LR au Canada.

L'infrastructure numérique de recherche (INR) est l'ensemble des outils et des services permettant aux chercheurs de transformer les données en percées scientifiques. »¹² En février 2016, la ministre des Sciences, l'honorable Kirsty Duncan, a rencontré les membres du Conseil du leadership sur l'infrastructure de recherche numérique (CLIRN) et leur a demandé d'affirmer leur intention de travailler en collaboration afin de s'assurer que le Canada dispose d'un écosystème IRN solide pour soutenir la recherche essentielle entreprise par la communauté de recherche universitaire du Canada. En mars 2016, le CLIRN a répondu à la demande du ministre Duncan par une lettre qui a été signée par tous les membres du CLIRN, confirmant leur engagement à travailler ensemble.

En novembre 2016, le département fédéral Innovation, Sciences et Développement économique (ISDE) a fourni un financement à la CLIRN pour établir un secrétariat à court terme avec un soutien en personnel à temps plein et a chargé le groupe d'entreprendre ce qui suit :

- Élaboration d'un document de synthèse sur la gestion des données (GD) ;
- Élaboration d'un document de synthèse sur le calcul informatique de pointe (CIP) ;
- Étude et recommandation d'approches potentielles pour la coordination.

À la fin du mois d'août 2017, des groupes de travail dédiés ont élaboré des exposés de position sur la GD¹³ et le CIP¹⁴ en consultant les experts et les intervenants de la communauté. La publication des documents a été suivie d'une nouvelle consultation.

Partiellement en réponse à ces rapports, le gouvernement du Canada investit 572,5 millions de dollars sur 5 ans par le biais du budget 2018¹⁵ pour améliorer l'infrastructure de recherche numérique du Canada. L'objectif du gouvernement est de fournir à tous les scientifiques et universitaires canadiens les outils numériques dont ils ont besoin pour soutenir l'innovation canadienne et occuper des positions de premier plan dans leurs domaines sur le plan international.

¹² Infrastructure numérique de recherche : <https://www.ic.gc.ca/eic/site/136.nsf/eng/home>

¹³ Exposé de position sur la gestion des données pour Innovation, Science et Développement économique Canada : https://engagedrica.sharepoint.com/:b:/s/RS/Eaw-akExLoxCoAWNZJafApoBE75_I2G_XSwCoP1SJ3HLLQ?e=g6TisQ

¹⁴ Exposé de position sur le calcul informatique de pointe (CIP) pour Innovation, Sciences et Développement économique Canada : https://engagedrica.sharepoint.com/:b:/s/RS/ES7QP8wWldIFu6tHJT_NiL8BmewcMHVm7iaNxELNww4hsQ?e=WQ5JVV

¹⁵ <https://www.budget.gc.ca/2018/home-accueil-fr.html>

Le 6 mai 2019, une proposition¹⁶ a été soumise à ISDE et approuvée par celui-ci pour la création d'une nouvelle organisation chargée de coordonner le financement et les orientations stratégiques des activités nationales liées au calcul informatique avancé (CIP), à la gestion des données (GD) et aux logiciels de recherche (LR).

Le 11 mars 2020, l'Alliance de recherche numérique du Canada (l'Alliance) a été officiellement lancée en tant qu'organisme national à but non lucratif et en tant qu'initiative clé de la stratégie nationale en matière d'IRN initiée par ISDE. L'Alliance joue un rôle essentiel en contribuant à l'établissement d'un écosystème d'IRN axé sur les chercheurs, responsable, agile, stratégique et durable pour les chercheurs canadiens en fournissant les outils, les services et l'infrastructure numériques. L'Alliance coordonnera et financera également des activités dans les domaines suivants : CIP, GDR, LR et cybersécurité. L'une des priorités du plan d'entreprise 2020-2021 de l'Alliance¹⁷ était de préparer un exposé de position pour évaluer l'état actuel des LR en complétant l'examen initié avec les exposés de position sur le CIP et la GDR. Fin septembre 2020, un groupe de travail sur les logiciels de recherche (GTLR) a été formé pour élaborer l'exposé de position en question.

Les objectifs ultimes de l'exposé de position sur les LR sont les suivants :

- Renforcer la compréhension de l'Alliance quant à l'état actuel des LR ;
- Contribuer à la définition et au développement de sa relation avec le modèle de prestation de services de l'Alliance ;
- Fournir une base pour le développement du plan stratégique qui guidera les activités de l'Alliance en matière de LR pour les années à venir.

Les membres du GTLR sont :

- Qian Zhang, analyste principal, logiciel de recherche, responsable du GTLR, Alliance;
- Susan Brown, professeure, Université de Guelph ;
- James Colliander, professeur, Université de la Colombie-Britannique ;
- Brian Corrie, responsable technique du projet iReceptor¹⁸, Université Simon Fraser ;
- Gabor Fichtinger, professeur, Université Queen's ;
- Scott Henwood, directeur principal, Programmes, CANARIE¹⁹ ;

¹⁶ Proposition de programme de contribution à l'infrastructure numérique : <https://engagedri.ca/assets/documents/Digital-Research-Infrastructure-Contribution-Program-Proposal-Complete-Submission-May-6-2019.pdf>

¹⁷ Plan organisationnel 2020-2021 : https://engagedri.ca/assets/documents/NOIRN-plan-organizationnel_2021-2022-FINAL.pdf

¹⁸ iReceptor : <http://www.ireceptor.org/>

¹⁹ <https://www.canarie.ca/>

- Mark Leggott, directeur général de Données de recherche Canada ;
- Catherine Lovekin, professeure associée, Université Mount Allison ;
- Felipe Pérez-Jvostov, analyste principal, planification et opérations, Alliance;
- Ghilaine Roquet, vice-présidente de la stratégie et de la planification, Alliance;
- Marc-Étienne Rousseau, directeur des logiciels de recherche, Université McGill.

Le GTLR a sollicité les commentaires de la communauté au sens large et des groupes d'intervenants sur l'exposé de position en matière de LR. Nous tenons à remercier les groupes suivants d'avoir partagé leur expertise : l'équipe d'analystes principaux de l'Alliance, le groupe de travail sur le calcul informatique de pointe et le groupe de travail sur la gestion des données de recherche, le conseil des chercheurs de l'Alliance et les responsables techniques des cinq sites hôtes. Nous remercions tout particulièrement Michelle Barker, directrice de Research Software Alliance (ReSA) et Daniel S. Katz, cofondateur de WSSSSPE, US-RSE et ReSA pour leur révision éditoriale détaillée et leurs commentaires.

3 Survol des LR

3.1. Que sont les LR ?

Les LR sont multidimensionnels et internationaux. Leur définition varie selon les différents contextes (tels que les intervenants, les publics, les fonctions ou les cas d'utilisation). Dans le présent document, nous définissons les LR d'un point de vue général, à l'intersection émergente et complexe d'outils, de disciplines, de services, de plateformes, de matériel, de ressources et de personnes (utilisateurs, chercheurs, développeurs, intervenants, personnel et communautés, etc.

Les LR jouent de multiples rôles dans la recherche²⁰ : ce sont des outils permettant la recherche (par exemple, pour manipuler tous les types de données de recherche tout au long du cycle de vie) ; ils peuvent être un résultat de recherche à part entière, un produit ou une solution primaire d'un projet de recherche ; ils sont aussi des objets de recherche qui sont étudiés ou utilisés pour faciliter la communication savante tout au long du cycle de vie de la recherche. En ce sens, les LR servent également de support pour le partage des idées de recherche et la collaboration. Les LR sont essentiels pour aider les chercheurs à faire de nouvelles découvertes et à partager leurs méthodes, leurs données et leurs résultats avec d'autres. Les LR peuvent être une plateforme de recherche (un ensemble d'outils permettant la recherche) et peuvent être fournis sous forme de code source (pouvant être lu et compilé), d'exécutable (pouvant être exécuté localement) et de service (pouvant être exécuté à distance).

Le terme LR au sens large peut inclure des programmes logiciels, des langages, des bibliothèques, des scripts, des codes de calcul, des modèles, des carnets de laboratoire

²⁰ <https://www.ouvrirlascience.fr/opportunity-note-encouraging-a-wider-usage-of-software-derived-from-research/>

électroniques, des logiciels de dépôt, des outils de gestion du flux de travail et d'autres outils développés pour soutenir la recherche ou faciliter les processus de recherche, pouvant eux-mêmes être mis en œuvre en LR. Les LR peuvent être offerts en ensemble et partagés sous forme de code source, d'exécutables, de conteneurs, de services, etc.

En tant qu'infrastructure de recherche, les LR font souvent référence à des plateformes ou à des services de recherche (aussi appelés laboratoires scientifiques virtuels, environnements de recherche virtuels [ERV] ou passerelles scientifiques^{21,22}) qui instancient une collection de code source de LR, y compris de grands cadres logiciels, des outils propres à une discipline, des services et du code de collage, et qui servent à soutenir les flux de travail de recherche et les communautés de pratique engagées dans la recherche collaborative²³. En règle générale, les capacités d'une plateforme de recherche comprennent l'acquisition et la gestion des données, le traitement et la visualisation, le stockage et la préservation, le partage et la découverte ; les plateformes peuvent fournir la gamme complète ou un sous-ensemble de composants. Les passerelles scientifiques peuvent être propres à une discipline et peuvent soutenir et améliorer la collaboration scientifique et la communication savante en facilitant également l'engagement des citoyens dans la science. Parmi les exemples, citons les portails de données comme le Dépôt fédéré de données de recherche²⁴ (DFDR) qui permet aux chercheurs de découvrir, de stocker ou de transférer des données de recherche, Syzygy²⁵, une plateforme interactive de calcul et d'apprentissage développée par le Pacific Institute for the Mathematical Sciences (PIMS)²⁶, et le Canadian Writing Research Collaboratory (CWRC)²⁷ qui fournit un dépôt, des outils et une plateforme de diffusion pour les études littéraires et culturelles. Voir l'[Annexe A](#) pour un ensemble plus vaste de passerelles scientifiques canadiennes.

3.2. LR, cycles de vie des données et de la recherche

La figure 1 décrit un cycle de vie traditionnel de la recherche, qui commence par l'idéation et la planification du projet et se poursuit par l'acquisition et la collecte de ressources, la gestion active du projet, la rédaction et la publication, la diffusion et l'évaluation.

Si cette vue d'ensemble d'un flux de recherche semble simple, la plupart des étapes ont leurs propres sous-cycles itératifs, représentant des étapes supplémentaires qu'un chercheur pourrait entreprendre. Il convient également de mentionner que le flux de recherche n'est jamais un processus séquentiel ou linéaire, car la connaissance est par nature itérative. C'est pourquoi les chercheurs font régulièrement des allers-retours entre les étapes, passant de l'analyse à la

²¹ <https://doi.org/10.1016/j.future.2018.12.026>

²² Voir la Figure 3 pour savoir dans quelle mesure les LR servent les trois composantes de l'IRN.

²³ Scott Henwood. Livre blanc. Le programme de logiciels de recherche de CANARIE.
<https://www.canarie.ca/wp-content/uploads/2021/03/Livre-blanc-CANARIE-Logiciels-de-recherche.pdf>

²⁴ <https://www.frdr-dfdr.ca/repo/?locale=fr>

²⁵ <https://syzygy.ca/>

²⁶ <http://www.pims.math.ca/>

²⁷ <https://cwrc.ca>

rédaction, et inversement, avant de passer à l'étape suivante de la publication. Les LR sont le plus souvent requis au stade de découverte et d'analyse, ainsi qu'au stade de publication, où l'archivage des ensembles de données et la compilation de paquets de recherche reproductibles par calcul font de plus en plus partie du processus.

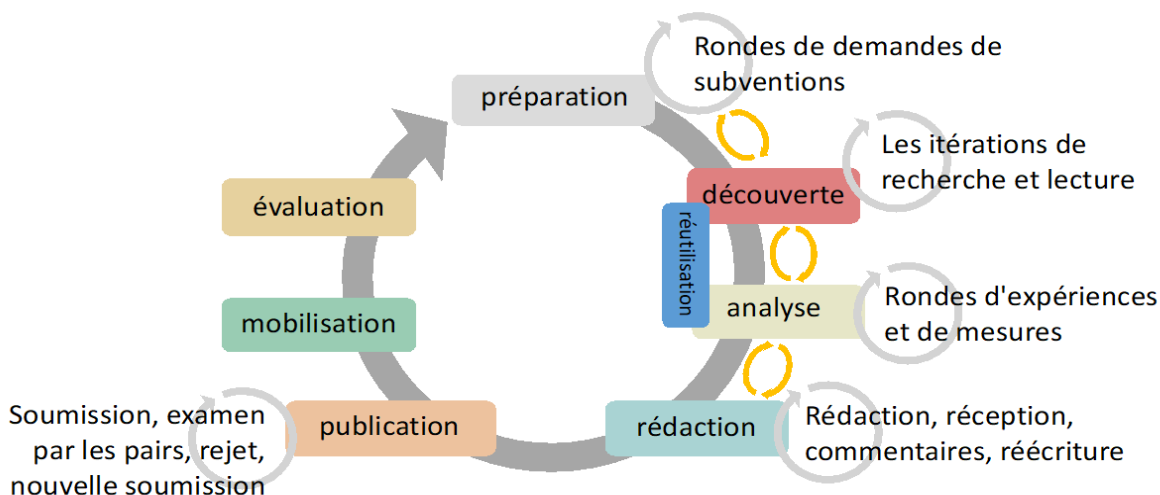


Figure 1 : Un modèle du cycle de vie de la recherche ²⁸

Les données font l'objet de considérations particulières dans le cycle de vie de la recherche, comme la décision de les préserver ou non, car certaines données ne valent pas la peine d'être préservées, ou encore il peut y avoir des considérations particulières (par exemple, des informations personnelles) empêchant la préservation des données au-delà d'une période prescrite. Les données ont généralement une durée de vie plus longue que le projet de recherche de sorte que les chercheurs peuvent continuer à travailler sur les données après l'arrêt du financement, les projets de suivi peuvent réanalyser ou ajouter de la valeur aux données et les données peuvent être réutilisées et réaffectées par d'autres chercheurs ; toutes ces activités reposant sur les LR.

Le cycle de vie des LR présente quelques différences par rapport au cycle de vie standard du développement logiciel²⁹ qui va généralement de la planification à la maintenance, en passant par l'analyse, la conception, la mise en œuvre, les tests et l'intégration. Le cycle de vie des LR est illustré à la figure 2 et comprend les éléments suivants :

- Acquisition et rassemblement des ressources (y compris le financement, l'embauche, les équipes et le développement communautaire).

²⁸ Bosman, Jeroen ; Kramer, Bianca (2016) : 101 Innovations dans la communication savante : re-designing research support services. figshare. Présentation.
<https://doi.org/10.6084/m9.figshare.4054728.v1>

²⁹ <https://online.husson.edu/software-development-cycle/>

- Développement
- Utilisation et réutilisation
- Reconnaissance des contributions et des répercussions
- Maintenance

La phase de développement comprend généralement une ou plusieurs étapes de prototypage. Le prototypage d'un logiciel de recherche peut lui-même, dans des domaines tels que la conception, l'interaction personne-machine et les sciences humaines numériques, constituer un mode de recherche et un résultat de recherche en soi, que les LR soient ou non développés en un système de production^{30,31}. Cela signifie que le soutien au prototypage est une caractéristique importante de l'infrastructure. De plus, l'étape du prototypage peut avoir des répercussions importantes sur le développement ultérieur des LR faisant en sorte que la participation de tous les membres de la communauté de pratique appropriée dès le début des programmes pilotes³² offre la possibilité de s'assurer que l'infrastructure émergente répond aux besoins de la communauté dans divers domaines. De plus, le prototypage basé sur des projets et le développement interne de LR peuvent s'étendre à des infrastructures largement utilisées telles que les carnets Jupyter³³ de plus en plus utilisés dans des disciplines autres que les STIM.

Ce modèle soutient la définition du logiciel pérenne considérée précédemment : la capacité à durer, la disponibilité à long terme et la réutilisabilité.

³⁰ Galey, Alan, et Stan Ruecker. "How a prototype argues". *Informatique littéraire et linguistique* 25.4 (2010) : 405-424.

³¹ Bardzell, Jeffrey, Shaowen Bardzell, et Lone Koefoed Hansen. "Propositions immodestes : Research through design and knowledge". *Actes de la 33e conférence annuelle de l'ACM sur les facteurs humains dans les systèmes informatiques*. 2015.

³² Parmi les exemples actuels de la National Science Foundation figurent l'Open Storage Network (<https://www.openstoragenetwork.org/>), la Global Research Platform (<http://www.theglobalresearchplatform.net/>) et FABRIC (<https://fabric-testbed.net/>).

³³ <https://www.nature.com/articles/d41586-018-07196-1>

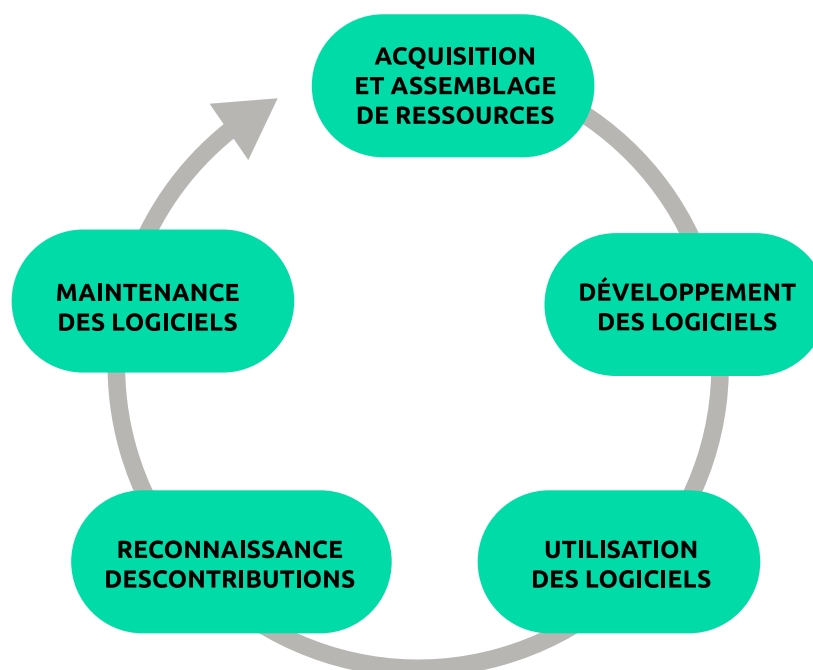


Figure 2 : Un modèle du cycle de vie des LR³⁴

Les deux cycles de vie reflètent un processus itératif dans ce contexte particulier, mais il existe clairement des dépendances et des intersections qui sont essentielles au déploiement d'une approche efficace et durable pour mener des recherches s'appuyant fortement sur les LR. Neil Chue Hong, directeur fondateur et chercheur principal du Software Sustainability Institute, fait remarquer que : « sans données, on peut difficilement valider les résultats, mais sans code, on perd l'occasion de faire progresser la science »³⁵. Les cycles de vie de la recherche et des LR font partie intégrante de l'entreprise de recherche et sont interconnectés ; ainsi nous avons illustré cette relation dans la figure 3 ci-dessous. La figure 3 met en évidence le rôle central que joue les LR : elle se superpose à d'autres domaines de l'infrastructure de recherche numérique dans une relation complexe avec les diverses activités représentées dans la figure 1.

Les activités *planifier-créer-traiter-analyser* de la figure 3 sont en grande partie la responsabilité des chercheurs principaux et des équipes de recherche. Lorsque l'on passe aux phases de *diffusion, de préservation et de réutilisation*, la responsabilité du maintien des LR devient floue, surtout lorsque les ressources sont limitées. Cela peut être dû au manque de financement, à l'absence de personnel hautement qualifié (PHQ) correctement formé ou dédié (ce qui se manifeste dans les sciences sociales et humaines). Le manque de clarté concernant les rôles et les responsabilités pendant les phases d'exploitation et de maintenance du développement peut créer un contexte précaire pour l'épanouissement des LR. La situation s'aggrave encore lorsque

³⁴ <http://wssspe.researchcomputing.org.uk/about-wssspe>

³⁵ <https://twitter.com/npch/status/1258388356431478784>

les personnes (en particulier les développeurs de LR) changent de rôle ou d'établissement (pour les secteurs privés). Un autre problème est la pérennité des LR : il existe actuellement peu d'options pour un soutien continu (par exemple, le transfert de la propriété des LR) une fois que l'effort financé du projet est terminé. Dans de nombreux cas, on demande à l'équipe informatique de l'établissement d'assumer la gestion des LR. Cela peut présenter d'énormes défis, car ces outils de LR peuvent ne pas avoir été développés sur la base de pratiques ou de normes communes (par exemple, la portabilité, l'interopérabilité, l'évolutivité, la bonne documentation) ; l'équipe informatique peut ne pas avoir l'expertise et la capacité de les maintenir pour une utilisation continue. Cela met en évidence deux défis majeurs dans l'écosystème des LR : appliquer les bonnes pratiques dans la conception et le développement des LR et assurer des ressources suffisantes pour maintenir les logiciels.

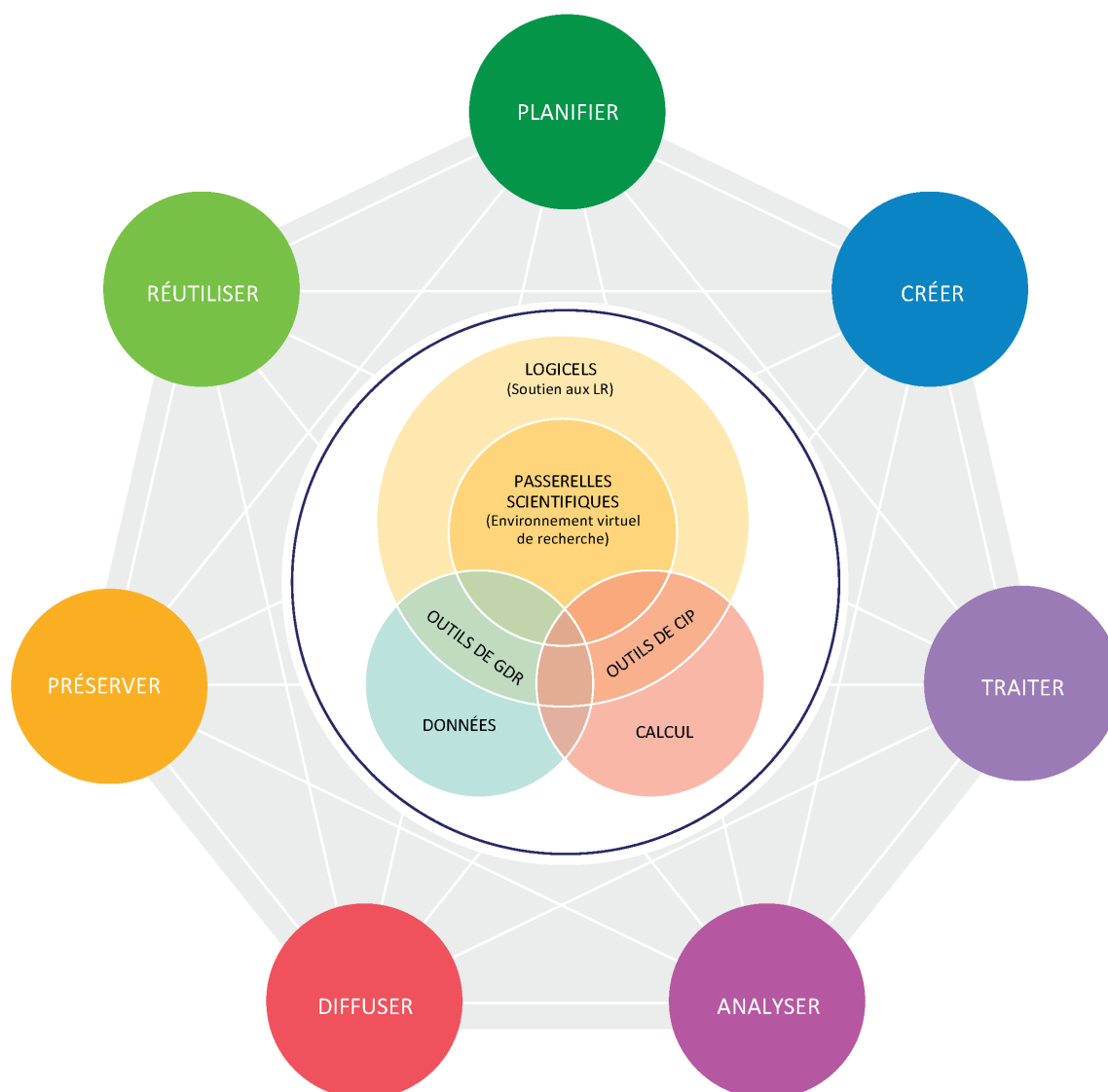


Figure 3 : Le cycle de vie de la recherche et les LR (voir [annexe B](#) et [annexe C](#)).

Interrelations entre les données, l'informatique et les LR (cercles intérieurs) recouvrant les phases du cycle de vie de la recherche (cercles extérieurs). Les utilisateurs et les communautés de recherche, qui sont les agents engagés dans les activités de LR, sont omis dans toutes ces activités. Modifié à partir de l'exposé de position sur la gestion des données (GD) pour Innovation, Sciences et Développement économique Canada. Conseil du leadership sur l'infrastructure de recherche numérique (manuscrit non publié, 31 août 2017) et de l'exposé de position sur le calcul informatique de pointe (CIP) pour Innovation, Sciences et Développement économique Canada. Conseil de leadership pour l'infrastructure de recherche numérique (manuscrit non publié, 31 août 2017).

3.3. Caractéristiques des LR

Les logiciels sont omniprésents dans le paysage de la recherche numérique moderne et font partie intégrante de la création, de l'analyse, de la gestion et de la diffusion des données et des publications.

Les logiciels de recherche (LR) sont fondamentaux pour la recherche et ils doivent être traités comme un produit de recherche *de premier ordre*, à conserver, évaluer et citer, et traités sur un pied d'égalité avec les articles et les données de recherche³⁶. Voici les caractéristiques des LR qui sont fondamentales pour le cycle de vie de la recherche (section 3.2) :

- Métadonnées : administratives, descriptives, techniques, structurelles, etc. ;
- Documentation : Readme, Markdown, manuel d'installation, etc. ;
- Capacité de citation : titre, créateur, identifiant unique, lieu de publication, date de publication, version, type, licence, langue, etc. ;
- Statut de licence : code source ouvert ou fermé, hybride ;
- Statut de publication : publié, non publié ;
- Statut d'instanciation : version, concept ;
- Utilisabilité³⁷ : concept, travail en cours, suspendu, abandonné, actif, inactif, non soutenu, déplacé ;
- Modèle économique : gratuit, commercial ;

³⁶ <https://software.ac.uk/blog/2018-11-28-making-software-first-class-citizen-research>

³⁷ <https://www.repostatus.org/>

- Mécanisme de distribution des logiciels : code source³⁸, exécutable binaire³⁹, paquet, conteneur, image de machine virtuelle, service, etc. ;
- Discipline : générique, propre au domaine ;
- Méthodes d'analyse : quantitative, qualitative ;
- Étapes prises en charge dans le cycle de vie de la recherche : planification, analyse, calcul, visualisation, transfert, stockage, publication, curation et préservation, découverte.

L'octroi de licences pour les logiciels est une caractéristique essentielle qui aide à garantir des conditions d'utilisation claires pour les utilisateurs actuels et futurs, comme l'utilisation, la modification ou la distribution des logiciels originaux ou modifiés. En tant que bien public, les projets à code source ouvert réussissent parce qu'ils ont traditionnellement attiré un grand nombre de bénévoles et de contributeurs qui collaborent à la conception, l'écriture, les tests, le débogage, la distribution, la documentation, le retour d'information, l'innovation, l'inspection et le signalement des bogues du logiciel, ainsi que la promotion du logiciel. Cela met en évidence un changement plus récent dans la nature de la création des logiciels (et dans un sens plus large, des connaissances) : d'une manière ouverte et collaborative, plutôt que dans le style traditionnel « fermé ».

Les logiciels de recherche à code source ouvert offrent également une opportunité sans précédent de communication multidisciplinaire et transdisciplinaire et d'engagement avec un large éventail d'intervenants : une opportunité que la communauté des chercheurs doit mieux comprendre pour l'exploiter plus efficacement. Les logiciels à code source ouvert ont également favorisé la transparence de la recherche en supprimant les restrictions d'accès aux environnements de développement, généralement accompagnées d'un accès libre aux données associées. Les logiciels à code source ouvert ont eu un impact considérable sur la société et l'économie : en plus des avantages généraux tels que la réduction du coût total de possession, le renforcement de la sécurité grâce à un examen plus approfondi du code et l'indépendance vis-à-vis des fournisseurs, si on lui attribue une valeur monétaire à l'impact économique des logiciels à code source ouvert, il pourrait se chiffrer en centaines de milliards de dollars par an^{40,41}. Cette approche a donné naissance à un mouvement social massif dans lequel des développeurs, des contributeurs, des utilisateurs, des chercheurs, des praticiens, des entreprises industrielles, des

³⁸ https://en.wikipedia.org/wiki/Source_code

³⁹ <https://en.wikipedia.org/wiki/Executable>

⁴⁰ <https://www.zdnet.com/article/how-much-are-open-source-developers-really-worth-hundreds-of-billions-of-dollars-say-economists/>

⁴¹ <https://www.oecd.org/going-digital/mdt-roadmap-open-source-software.pdf>

gouvernements et d'autres intervenants collaborent pour créer un bien public dans le paysage numérique⁴².

3.4. Types de LR

Selon la définition plus large des LR [dans la section 3.1](#), les LR peuvent être classées en deux catégories générales ci-dessous :

- 1 Outils, bibliothèques et scripts de logiciel : composants intégrés au niveau du développement qui remplissent généralement une fonction dans le processus ou l'infrastructure de recherche et peuvent également être considérés comme fondamentaux pour des flux de travail, des plateformes de recherche et des passerelles scientifiques plus complexes.
- 2 Plateformes de recherche ou passerelles scientifiques : elles couvrent de nombreux services et présentent généralement à l'utilisateur des sous-catégories de services, de composants et de modules. Si les passerelles scientifiques sont presque toujours propres à un domaine, elles peuvent également être indépendantes du domaine.

En combinant l'ensemble typique de logiciels scientifiques défini par Hinsén⁴³, la taxonomie aplatie de Sochat⁴⁴ et la classification induite par la définition des LR ci-dessus, le groupe de travail sur les LR a élaboré une nouvelle classification à trois niveaux des LR que nous utilisons dans ce document. Dans chacune des catégories, des instances et des exemples spécifiques peuvent être classés dans divers sous-groupes qui démontrent l'étendue des applications de LR de la manière suivante :

- 1 Logiciels dédiés créés ou utilisés pour la recherche

⁴² Un hôte de longue date pour les projets de logiciels à code source ouvert est GitHub (<https://github.com/>), qui en janvier 2020 affirmait avoir plus de 190 millions de dépôts et plus de 40 millions d'utilisateurs. Les logiciels à code source ouvert ont eu un impact sur la concurrence mondiale dans les secteurs des logiciels (et du matériel) informatiques, où les entreprises se faisaient traditionnellement concurrence sur la base de composants protégés « à code source fermé ». Cela a conduit certaines entreprises possédant des produits logiciels protégés à adopter des solutions à code source ouvert dans leurs produits. Cela a également conduit à la création d'un nouveau type d'entreprise, comme Red Hat (<https://www.redhat.com/en>) et MySQL (<https://www.mysql.com/>), qui assemble, distribue et vend des logiciels et des services à code source ouvert. Dans de nombreux pays, les organismes gouvernementaux ont adopté des politiques explicites facilitant le développement et l'utilisation des logiciels à code source ouvert. Par exemple, à l'automne 2015, la division des sciences de la Terre (ESD) de la NASA a publié une politique de logiciel à code source ouvert pour le programme Earth Science Data Systems (ESDS) (https://sites.nationalacademies.org/cs/groups/ssbsite/documents/webpage/ssb_174603.pdf ; <https://earthdata.nasa.gov/earth-science-data-systems-program/policies/esds-open-source-policy>).

⁴³ Konrad Hinsén. Dealing With Software Collapse. Computing in Science and Engineering, Institute of Electrical and Electronics Engineers, 2019, 21 (3), pp.104-108. ff10.1109/MCSE.2019.2900945ff. ffhal02117588f HAL Id: hal-02117588 <https://hal.archives-ouvertes.fr/hal-02117588>

⁴⁴ What is Research Software? <https://vsoch.github.io/2020/what-is-research-software/>

- 1.1. Paquets de logiciels propres au domaine : notamment les bibliothèques, les scripts, les outils, les flux de travail, les codes de modélisation communautaires, etc. (par exemple, IgBLAST⁴⁵, Astropy⁴⁶, NAMD⁴⁷, le modèle climatique CESM⁴⁸, SPM⁴⁹, FSL⁵⁰, AFNI⁵¹, BROCCOLI⁵²).
- 1.2. Passerelles scientifiques propres à un domaine (par exemple, FRACS⁵³, DataStream⁵⁴, Motus⁵⁵, Nunaliit⁵⁶, OBiBa⁵⁷, WARN⁵⁸).
- 1.3. Paquets de logiciels sans domaine particulier : notamment les outils de visualisation, de provenance et de collecte de métadonnées, les gestionnaires de flux de travail, les environnements de développement intégrés (EDI), etc. (par exemple NumPy⁵⁹, scikit-

⁴⁵ <https://www.ncbi.nlm.nih.gov/igblast/>

⁴⁶ <https://www.astropy.org/>

⁴⁷ <https://www.ks.uiuc.edu/Research/namd/>

⁴⁸ <https://www.cesm.ucar.edu/models/>

⁴⁹ https://en.wikipedia.org/wiki/Statistical_parametric_mapping#:~:text=SPM%20est%20logiciel%20écrit%20par,est%20distribué%20comme%20logiciel%20gratuit

⁵⁰ http://ftp.nmr.mgh.harvard.edu/pub/dist/freesurfer/tutorial_packages/OSX/fsl_501/doc/index.html#:~:text=FSL%20est%20une%20bibliothèque%20complète,est%20très%20facile%20à%20installer

⁵¹ <https://afni.nimh.nih.gov/>

⁵² <https://www.nitrc.org/projects/broccoli/>

⁵³ FRACS (FAIR Repository for Annotations, Corpora and Schemas) est un dépôt d'annotations qui facilite la création, le stockage, la récupération, la manipulation et la distribution d'ensembles d'annotations sur des corpus de documents. <https://data.crim.ca/en/>

⁵⁴ DataStream est une puissante plateforme en ligne permettant de partager des informations sur la santé des eaux douces. <https://gordonfoundation.ca/initiatives/datastream/>

⁵⁵ Le Motus Wildlife Tracking System (Motus) est un réseau international de recherche collaborative qui utilise la radiotélémetrie automatisée coordonnée pour faciliter la recherche et l'éducation sur l'écologie et la conservation des animaux migrateurs. <https://motus.org/>

⁵⁶ Le cadre de l'atlas Nunaliit vise à faciliter le récit et la mise en évidence des relations entre de nombreuses formes d'informations provenant de sources diverses, en utilisant des cartes comme moyen central de connexion et d'interaction avec les données. <http://nunaliit.org/>

⁵⁷ OBiBa est un projet international visant à construire un logiciel à code source ouvert pour les études épidémiologiques. <https://www.obiba.org/>

⁵⁸ Le Web-enabled Awareness Research Network (WARN) est une plateforme logicielle de recherche sur la détection des géorisques qui coordonne les informations provenant de dizaines de capteurs afin d'avertir le personnel d'urgence et les systèmes de sauvegarde automatiques d'événements potentiellement catastrophiques. <https://www.canarie.ca/software/platforms/warn/>

⁵⁹ <https://numpy.org/>

learn⁶⁰, YesWorkflow⁶¹, CEDAR Workbench⁶², Zotero, Kepler⁶³, Pegasus⁶⁴, Taverna⁶⁵, Matlab⁶⁶, Jupyter [notebooks]).

- 1.4. Passerelles scientifiques ou dépôts de données sans domaine particulier : (par exemple, OSF, HubZero, DFDR)

AFNI (Analysis of Functional NeuroImages) est une suite de **LR propres à un domaine**, composée de programmes C, Python, R et de scripts shell **permettant de mener directement des recherches**, qui sert à analyser et à afficher de multiples modalités d'IRM : données anatomiques, IRM fonctionnelle (IRMF) et données pondérées par diffusion.

- 3 Logiciels à usage général fréquemment utilisés pour la recherche : notamment les bases de données, les interfaces de programmation d'applications (API), les cadres de travail, les dépôts de données, les outils de données généraux, les moteurs de recherche, les outils d'analyse et de visualisation, etc. (par exemple, PostgreSQL, Neo4j, API REST, API Java, outils de documentation, systèmes de gestion de contenu, Doxygen⁶⁷, Drupal⁶⁸, Excel, Grafana⁶⁹, Elasticsearch⁷⁰, Dataverse, Clowder⁷¹, Invenio⁷²).
- 4 Logiciels utilisés dans le processus de recherche : notamment les systèmes d'exploitation, l'ordonnancement et la gestion des tâches, la gestion de version, les compilateurs, les débogueurs, les profileurs, les systèmes de fichiers parallèles, les ordonnanceurs de tâches,

⁶⁰ <https://scikit-learn.org/>

⁶¹ <https://github.com/yesworkflow-org>

⁶² <https://metadatacenter.org/>

⁶³ <https://kepler-project.org/>

⁶⁴ <https://pegasus.isi.edu/>

⁶⁵ <https://taverna.incubator.apache.org/>

⁶⁶ <https://www.mathworks.com/products/matlab.html>

⁶⁷ <http://www.doxygen.nl/>

⁶⁸ <https://www.drupal.org/>

⁶⁹ <https://grafana.com/>

⁷⁰ <https://www.elastic.co/>

⁷¹ <https://ssa.ncsa.illinois.edu/isda/software/clowder/>

⁷² <https://inveniosoftware.org/>

les outils de surveillance et de performance des systèmes, les éditeurs de fichiers, les outils et plateformes de communication, les applications et services de partage de fichiers ou de stockage infonuagique (par exemple, Linux, Windows, Jira, BBEdit⁷³, Photoshop, Microsoft Word, courrier électronique, Slack, Twitter, Google Drive, Dropbox, git, GitHub).

Pour d'autres façons de classer les LR, voir l'[annexe D](#).

3.5. Les LR dans le mouvement de la science ouverte

Le concept de science ouverte n'a pas la même signification pour tout le monde, mais il englobe en fin de compte : la *philosophie* de la science ouverte, qui consiste à « faire avancer la recherche sur les épaules des géants », un cadre de recherche largement reconnu ; la *politique* de la science ouverte, qui encourage ou impose le partage des ressources numériques produites avec des fonds publics ou selon les exigences de l'éditeur d'une revue ; et la *pratique* de la science ouverte, où chaque discipline définit ses propres bonnes pratiques et outils, notamment s'il est avantageux ou non de partager les résultats de la recherche, et quand, comment et quoi partager.

La science ouverte « est le mouvement visant à rendre la recherche, les données et la diffusion scientifiques accessibles à tous les niveaux d'une société curieuse »⁷⁴. La science ouverte étend les principes d'ouverture à l'ensemble du cycle de vie de la recherche en permettant une connaissance transparente et accessible, en encourageant le partage et la collaboration le plus tôt possible, et en renforçant l'impact de la recherche scientifique, transformant ainsi systématiquement la manière dont la science et la recherche sont menées⁷⁵.

La science ouverte englobe des pratiques telles que la publication de recherches ouvertes, les campagnes en faveur du libre accès, l'encouragement des scientifiques à pratiquer la science du carnet ouvert et, d'une manière générale, la facilitation de la publication et de la communication des connaissances scientifiques.

La science ouverte implique divers mouvements tels que le libre accès (LA), les monographies ouvertes, les données ouvertes (par exemple, les normes de données ouvertes, l'utilisation et la réutilisation des données ouvertes), la recherche ouverte ou reproductible (carnets de notes ouverts, flux de travail scientifiques ouverts, logiciels à code source ouvert), la méthodologie ouverte, la collaboration ouverte, la gouvernance ouverte, les ressources éducatives libres, l'évaluation scientifique ouverte (par exemple, les mesures et les répercussions ouvertes, l'examen par les pairs ouvert), les politiques scientifiques ouvertes (par exemple, les bailleurs de fonds, les éditeurs, les gouvernements, les établissements) et les outils scientifiques ouverts (tels que les dépôts ouverts, les services ouverts, les outils de flux de travail ouverts), les réseaux sociaux scientifiques, la science citoyenne⁷⁶ et le financement participatif de la recherche.

⁷³ <https://www.barebones.com/products/bbedit/>

⁷⁴ <https://www.fosteropenscience.eu/taxonomy/term/7>

⁷⁵ <https://www.fosteropenscience.eu/node/1420>

⁷⁶ https://en.wikipedia.org/wiki/Citizen_science

Avec l'adoption généralisée de la science ouverte et compte tenu du fait que la plupart des recherches scientifiques modernes s'appuient dans une certaine mesure sur les IRN, nous devrions examiner de près ces infrastructures pour nous assurer qu'elles sont elles-mêmes conformes aux principes de l'ouverture et qu'elles sont disponibles pour les chercheurs en tant qu'infrastructure ouverte (voir Invest in Open Infrastructure, IOI⁷⁷). Sans ces qualités, les IRN se positionnent davantage comme un risque pour la science ouverte que comme une opportunité. Si les bailleurs de fonds au Canada d'IRN s'engagent à construire une infrastructure ouverte, nous serons bien placés pour être un chef de file de l'offre d'infrastructure savante mondiale réellement pérenne.

Selon un rapport de la FCI de 2015⁷⁸, bon nombre des défis liés à l'écosystème d'IRN au Canada impliquent un changement de mentalité vers une plus grande ouverture et une plus grande collaboration, comme un changement de mentalité de « mes données » à « nos données », des « priorités de l'individu » aux « priorités du collectif », et d'un écosystème d'IRN « fragmenté » à un écosystème d'IRN « distribué » (FCI 2015). Bien que la stratégie d'IRN du rapport de la FCI ne traite pas explicitement de la science ouverte, elle souligne la nécessité d'améliorer l'infrastructure pour permettre un plus grand partage des données et une plus grande collaboration, et pour améliorer la découverte et l'accessibilité des données de recherche. Autrement dit, pour fournir une base solide pour que la science ouverte puisse s'épanouir. Plus récemment (2020), le bureau de la Conseillère scientifique en chef du Canada a fourni des principes généraux et des recommandations (y compris une politique nationale de science ouverte) pour guider les activités de science ouverte au Canada⁷⁹.

Le développement, le déploiement et la maintenance de LR réutilisables (qu'ils soient de nature informatique ou qu'ils reposent sur une analyse ou une interprétation logicielle) sont de plus en plus reconnus internationalement comme un élément clé pour faciliter les résultats de recherche fiables et reproductibles et la science ouverte. Le partage des LR est une condition nécessaire, mais insuffisante, de la reproductibilité ; le libre accès au code source des LR contribue à améliorer l'impact de la recherche associée. Dans le contexte de la science ouverte, les LR peuvent être produits avec un code source ouvert et une licence permettant la modification, la dérivation et la redistribution (l'approche la plus courante), ou ils peuvent être commercialisés ; les deux approches sont utilisées dans la communauté de la recherche.

3.6. Décrire et mesurer les impacts des LR

3.6.1. Métadonnées de LR

Bien que la recherche moderne s'appuie fortement sur les LR et qu'un nombre croissant de chercheurs s'engagent à développer des LR dans le cadre de leurs recherches, les infrastructures et les bonnes pratiques qui soutiennent l'attribution, la découverte, l'utilisation et la réutilisation et

⁷⁷ <https://investinopen.org/>

⁷⁸ [https://www.ic.gc.ca/eic/site/113.nsf/vwapj/DRI_Strategy_Consultation-fra.pdf/\\$file/DRI_Strategy_Consultation-fra.pdf](https://www.ic.gc.ca/eic/site/113.nsf/vwapj/DRI_Strategy_Consultation-fra.pdf/$file/DRI_Strategy_Consultation-fra.pdf)

⁷⁹ [https://www.ic.gc.ca/eic/site/063.nsf/vwapj/La-Feuille-de-route-pour-la-science-ouverte.pdf/\\$file/La-Feuille-de-route-pour-la-science-ouverte.pdf](https://www.ic.gc.ca/eic/site/063.nsf/vwapj/La-Feuille-de-route-pour-la-science-ouverte.pdf/$file/La-Feuille-de-route-pour-la-science-ouverte.pdf) (en français)

la préservation des LR sont en retard par rapport à celles des données de recherche et des publications. Des métadonnées claires et complètes sont essentielles pour attribuer les résultats des LR à des personnes et des équipes précises. Malheureusement, les dépôts de logiciels, les outils de développement, les langages et les domaines de recherche ne désignent pas ces informations de la même manière, ce qui rend difficile, voire impossible, de travailler sur différentes sources sans perdre des informations précieuses. Bien qu'un défi semblable existe dans le contexte des données de recherche, les détails diffèrent considérablement entre les deux contextes. Il n'y a pas de bonnes pratiques ou de norme unique pour documenter, indexer, partager, archiver et gérer le logiciel ; la citation du logiciel diffère entre les disciplines et les archives.

Il y a diverses initiatives (par exemple CodeMeta⁸⁰, le schéma de métadonnées DataCite⁸¹, DOE Code Metadata Model⁸², lignes directrices pour les métadonnées des logiciels OpenAIRE⁸³) pour faciliter l'attribution des LR pour différents scénarios de préservation, de découverte, de réutilisation et de citation qui sont en cours. Voir l'[annexe E](#) pour plus de détails. Ils sont généralement interopérables actuellement et s'orientent vers une interopérabilité complète basée sur une version de schema.org.

3.6.2. Licences pour les LR

Une licence de logiciel est la base du contrat entre le créateur du logiciel (ou le propriétaire du droit d'auteur) et l'utilisateur final. La licence est un accord ou un instrument fixant des conditions précises qui régissent la façon dont le logiciel peut être utilisé, modifié, dérivé, réutilisé ou distribué par un utilisateur final. Dans la plupart des régions, une œuvre originale est automatiquement protégée par le droit d'auteur, il est donc important pour le titulaire du droit d'auteur (ou le créateur du logiciel) d'inclure un ensemble de permissions formelles ou de « conditions » d'utilisation précises accordées au titulaire de la licence (ou à l'utilisateur final). Sans accord de licence, le logiciel peut être laissé dans un état d'incertitude, car les utilisateurs potentiels peuvent ne pas savoir quelles sont les limites prévues par le créateur. Les créateurs peuvent s'exposer à des actions en justice ou rencontrer des difficultés à contrôler la manière dont leur travail est utilisé. Cela vaut aussi bien pour les logiciels protégés ⁸⁴(généralement des logiciels commerciaux payants) que pour les logiciels libres à code source ouvert⁸⁵. Bien que les utilisateurs finaux puissent rechigner devant des licences de logiciel trop restrictives, l'incertitude causée par l'absence de licence peut également décourager ceux qui souhaitent utiliser l'œuvre. Dans le cas où le logiciel à code source ouvert a été créé par plusieurs contributeurs, la gestion de la propriété et de la licence des droits d'auteur devient encore plus compliquée. Souvent, les politiques institutionnelles locales ou les contrats syndicaux clarifient la question de la propriété.

⁸⁰ <https://codemeta.github.io/>

⁸¹ <https://schema.datacite.org/>

⁸² <https://github.com/doecode/software-metadata/blob/master/XMLSchema/MetadataSchema.xsd>

⁸³ <https://software-guidelines.readthedocs.io/en/latest/>

⁸⁴ https://en.wikipedia.org/wiki/Proprietary_software

⁸⁵ https://en.wikipedia.org/wiki/Free_and_open-source_software

Un autre moyen de gérer et de protéger davantage les droits associés aux contributions, souvent négligés, est l'accord de licence des contributeurs (ALC)⁸⁶ pouvant être personnalisé pour définir, décrire et accorder expressément une licence de droit d'auteur pour la contribution logicielle, les droits et obligations des auteurs contributeurs, le projet à code source ouvert, ou le mainteneur du projet, le cas échéant. À long terme, les ALC permettent non seulement de maintenir l'intégrité du projet, mais aussi de fournir des garanties juridiques à chacun des participants au projet, en facilitant la résolution des litiges concernant la licence ou la propriété des contributeurs de la base de code. Les ALC peuvent inclure des dispositions relatives à la violation des droits d'auteur ou à la résolution alternative des litiges, ou même permettre de changer de licence au fil du temps, ou de distribuer la contribution simultanément sous des licences (exclusives) distinctes sans avoir à demander l'approbation préalable de chaque contributeur. Les ALC ne sont pas normalisées, de sorte que les contributions à différents projets de logiciels à code source ouvert peuvent être régies par plusieurs ALC. Bien que les ALC ne soient pas idéaux pour tous les projets, comme les petits projets à code source ouvert, les projets plus importants peuvent exiger des ALC formels de la part de leurs contributeurs. Les choix de mise en place d'un ALC varieront d'un ALC court et simple à un instrument juridique plus exhaustif⁸⁷.

Dans le milieu universitaire, il convient de noter que les licences peuvent être utilisées pour faciliter l'accès aux LR ou pour le restreindre ; il faut donc choisir la bonne licence de logiciel. Si les licences Creative Commons (notamment CC-BY) sont couramment utilisées dans le contexte universitaire, elles ne sont pas recommandées pour les logiciels⁸⁸, car les licences propres aux logiciels (par exemple Apache-2.0⁸⁹, MIT⁹⁰, GNU0GPLv3⁹¹) fournissent des indications beaucoup plus claires aux utilisateurs potentiels. Il convient également de noter qu'en vertu des politiques (c'est-à-dire des contrats syndicaux) de la plupart des établissements d'enseignement au Canada, c'est généralement le créateur qui détient tous les droits de propriété intellectuelle ⁹²(PI) sur ses œuvres, tandis que dans les collèges et autres types d'établissements de recherche, les droits de PI appartiennent généralement à l'établissement ou au partenaire financier. Il faut également garder à l'esprit que chaque source de financement (tant publique que privée) peut avoir ses propres exigences en matière de PI, et qu'elles peuvent être négociables ou non.

⁸⁶ <https://www.finnegan.com/en/insights/articles/what-you-should-know-about-contributor-license-agreements-in-open-source-projects.html>

⁸⁷ Parmi les exemples de projets à code source ouvert bien connus, citons Apache Software Foundation (<https://www.apache.org/>), Django (<https://www.djangoproject.com/>) et Eclipse Foundation (<http://www.eclipse.org/projects/>). Le projet Islandora est régi par une fondation canadienne et comprend différentes formes d'ALC (<https://www.islandora.ca/contribute>), notamment l'accord de licence des contributeurs individuels (https://islandora.ca/sites/default/files/islandora_cla.pdf) ou l'accord de licence des contributeurs d'entreprise (https://islandora.ca/sites/default/files/islandora_ccla.pdf).

⁸⁸ <https://creativecommons.org/faq/#can-i-apply-a-creative-commons-license-to-software>

⁸⁹ <https://opensource.org/licenses/Apache-2.0>

⁹⁰ <https://opensource.org/licenses/MIT>

⁹¹ <https://choosealicense.com/licenses/gpl-3.0/>

⁹² https://en.wikipedia.org/wiki/Intellectual_property

Les modalités des licences de logiciel varient, mais deux catégories générales sont utilisées dans tout l'éventail des stratégies de licences de logiciel : propriétaire ou gratuit à code source ouvert. Ces dernières années, une troisième catégorie est apparue : la catégorie hybride.

Logiciels propriétaires

Ce type de logiciel est généralement à code source fermé (alias propriétaire), ce qui inclut toute variation qui ne permet pas d'afficher et de réutiliser le code source. Par conséquent, les logiciels sous licence propriétaire sont généralement distribués uniquement sous forme binaire et l'examen du code source ou l'ingénierie inverse de toute partie du code n'est généralement pas autorisée. L'objectif principal d'une licence propriétaire de logiciel est de limiter l'utilisation du logiciel en fonction de la stratégie commerciale du propriétaire. C'est pourquoi les licences propriétaires sont souvent restrictives pour les utilisateurs finaux. Par exemple, elles permettent généralement d'utiliser et d'installer le logiciel sur un seul ordinateur, avec des limites supplémentaires telles que des comptes utilisateurs limités, l'impossibilité de réutiliser le code ou de redistribuer ou modifier le produit. Dans les milieux universitaires, les logiciels propriétaires peuvent occasionnellement diffuser le code source « à des fins d'inspection uniquement » en raison des exigences de publication scientifique ou d'examen par les pairs. La diffusion du code source peut également faire partie des contrats d'achat (par exemple, une clause d'entiercement), dans le cas où une entreprise fait faillite ou cesse de prendre en charge une application particulière. Dans certains cas, les logiciels développés dans le cadre de projets de recherche, même avec un financement public, peuvent devenir propriétaires, car c'est la méthode choisie par les développeurs et le bailleur de fonds pour répondre au besoin d'une source de financement fiable pour le maintenir. Parfois, pour ces cas, il y a un accord pour une distribution gratuite ou à faible coût aux établissements universitaires, tandis que d'autres utilisateurs doivent payer une redevance plus élevée.

Logiciels libres à code source ouvert

Les logiciels à code source ouvert représentent une approche fondamentalement différente de la licence des logiciels, dans laquelle le code source est mis à la disposition des utilisateurs et inclut la possibilité d'utiliser, de visualiser, de distribuer, de modifier et de réutiliser le code. Le principal avantage du logiciel libre à code source ouvert est qu'il facilite l'accès et maximise l'ouverture, réduisant ainsi les obstacles à l'utilisation et à la diffusion des logiciels, ainsi qu'à la dérivation et à l'innovation qui s'ensuivent, ce qui conduit finalement à de meilleurs logiciels. Parmi les autres avantages d'une stratégie de logiciel libre à code source ouvert, il y a notamment : une adoption plus généralisée, une sécurité accrue, un code robuste et pérenne⁹³ avec un repérage et des corrections rapides des bogues, la création d'une communauté, la facilité de collaboration et les contributions des utilisateurs. La disponibilité du code source et l'accès à celui-ci sont de plus en plus importants dans le contexte de la recherche scientifique, où l'examen par les pairs, la reproductibilité et l'exploitation des travaux antérieurs font partie intégrante de l'avancement de la science. De plus, les licences des logiciels libres à code source ouvert peuvent prolonger la durée de vie d'un logiciel (c'est-à-dire sa pérennité) au-delà de l'implication directe des créateurs originaux en simplifiant le développement continu et l'adoption par les successeurs.

⁹³ <https://openresearchsoftware.metajnl.com/article/10.5334/jors.118/>

Il y a une grande variété de licences populaires pour les logiciels à code source ouvert. Des informations sur la manière de choisir une licence de logiciel libre sont disponibles sur [ChooseALicense.com](https://choosealicense.com), SPDX License List⁹⁴, le projet REUSE⁹⁵ et l'Open Source Initiative⁹⁶. Les initiatives nationales pour la science ouverte en Europe (NI4OS Europe) ont publié le License Clearance Tool (LCT)⁹⁷, un portail Web qui offre une approche guidée pour établir la licence de logiciel à code source ouvert appropriée pour la création de nouveaux résultats de recherche (par exemple, ensemble de données, médias, logiciels, etc.) ou pour la réutilisation de contenu existant sans licence. Certaines licences de logiciels à code source ouvert peuvent également être formulées de manière à permettre la dérivation d'alternatives à source fermée, ce qui pourrait être considéré comme une voie vers les licences hybrides.

Logiciels hybrides

Ce modèle de licence est également appelé double licence ou licence multiple⁹⁸. Comme son nom l'indique, l'approche hybride combine une licence de logiciel libre à code source ouvert et une licence propriétaire fermée, facilitant ainsi la pratique de la distribution de logiciels sous deux ou plusieurs ensembles de modalités ou de conditions. Cette option sert quand aucun des deux modèles de licence ne répond aux besoins des développeurs de logiciels. Dans ce schéma de licence hybride, le titulaire des droits peut choisir la licence à appliquer au cas par cas. L'inconvénient de cette approche peut être une charge supplémentaire importante pour le titulaire des droits d'appliquer, d'administrer et de faire respecter des licences multiples, ce qui limitera généralement l'adoption à grande échelle de modèles de licences hybrides aux grandes initiatives de développement de logiciels.

3.6.3. LR et reproductibilité de la recherche

L'objectif principal de la reproductibilité de la recherche est de fournir suffisamment d'informations sur une étude pour que les méthodes de recherche utilisées puissent être utilisées par un autre chercheur (reproductibilité de la méthode) afin de produire les mêmes résultats (reproductibilité des résultats) avec les mêmes matériaux/entrées⁹⁹. La reproductibilité de la recherche est une composante essentielle et continue du processus de communication scientifique, comme l'indiquent les principes FAIR pour rendre les données faciles à trouver, accessibles, interopérables et réutilisables. Les LR, ainsi que les résultats de recherche connexes, favorisent la production, le partage et la mobilisation des connaissances en permettant la reproductibilité. Une approche fondée sur les bonnes pratiques en matière de reproductibilité fait avancer la cause de la science ouverte et de l'intendance numérique par le biais du libre accès, des données ouvertes et, idéalement, du code source ouvert. La gestion des versions de logiciel, un cadre robuste d'essai et de qualité (par exemple, la vérification et la validation), les dépôts de code et

⁹⁴ <https://spdx.org/licenses/>

⁹⁵ <https://reuse.software/>

⁹⁶ <https://opensource.org/licenses/alphabetical>

⁹⁷ <https://lct.ni4os.eu/lct/login>

⁹⁸ <https://en.wikipedia.org/wiki/Multi-licensing>

⁹⁹ What does research reproducibility mean? Science Translational Medicine, 1^{er} juin 2016 : Vol. 8, numéro 341, pp. 341ps12, DOI : [10.1126/scitranslmed.aaf5027](https://doi.org/10.1126/scitranslmed.aaf5027).

la portabilité infonuagique ont tous contribué à l'évolution rapide de la reproductibilité de la recherche.

Depuis l'adoption généralisée des principes FAIR pour les données, de nombreuses initiatives ont vu le jour pour appliquer les principes FAIR à d'autres artefacts de recherche tels que les principes FAIR pour le matériel didactique¹⁰⁰, les principes FAIR pour la sémantique¹⁰¹ et, plus largement, les principes FAIR pour les services¹⁰² et pour les objets numériques¹⁰³. Certains travaux^{104,105,106} ont été étendus à l'application des LR et les principes directeurs qui traduisent cet objectif dans le contexte des LR¹⁰⁷ sont toujours en cours d'élaboration. Plus précisément, des initiatives telles que le groupe de travail FAIR for Research Software (FAIR4RS)¹⁰⁸ et le groupe de travail CURE-FAIR¹⁰⁹ ont travaillé à la coordination d'une série de discussions menées par la communauté sur la manière de définir et d'appliquer efficacement les principes FAIR aux LR.

De plus amples informations sur les bonnes pratiques en matière de gestion des LR et la reproductibilité informatique sont fournies à l'[annexe F](#).

3.6.3.1. Partage de LR

La disponibilité des LR est essentielle à la reproductibilité de la recherche. Les éditeurs de revues et d'autres acteurs de la recherche (tels que des bailleurs de fonds comme le Wellcome Trust¹¹⁰) ont récemment commencé à inclure explicitement les LR dans les politiques relatives à la gestion et au partage des résultats de la recherche universitaire, en exigeant que les LR et les données

¹⁰⁰ <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1007854>

¹⁰¹ <https://www.fairsfair.eu/events/recommendations-fair-semantic>

¹⁰² <https://www.scilit.net/article/b22490fab3437063a08bb94ef6eb9b82>

¹⁰³ <https://datascience.codata.org/articles/10.5334/dsj-2020-015/>

¹⁰⁴ Chue Hong, Neil ; Katz, Daniel S. (2018) : FAIR enough ? Can we (already) benefit from applying the FAIR data principles to software?. figshare. Présentation.
<https://doi.org/10.6084/m9.figshare.7449239.v2>

¹⁰⁵ Christopher Erdmann, Natasha Simons, Reid Otsuji, Stephanie Labou, Ryan Johnson, Guilherme Castelao, ... Tim Dennis. (2019, février). Top 10 FAIR Data & Software Things. Zenodo.
<http://doi.org/10.5281/zenodo.2555498>

¹⁰⁶ Aerts, Dr. (PhD) P.J.C. (DANS) (2017) Sustainable Software Sustainability - Rapport d'atelier. DANS.
<https://doi.org/10.17026/dans-xfe-rn2w>

¹⁰⁷ Lamprecht, Anna-Lena et al. "Towards FAIR Principles for Research Software". 1er janvier 2020 : 37 - 59. <https://doi.org/10.3233/DS-190026>

¹⁰⁸ <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>. La version finale des principes FAIR pour les LR est en phase finale d'examen officiel par la communauté - <https://www.rd-alliance.org/group/fair-research-software-fair4rs-wg/outcomes/fair-principles-research-software-fair4rs>.

¹⁰⁹ <https://www.rd-alliance.org/groups/cure-fair-wg>

¹¹⁰ <https://wellcome.org/grant-funding/guidance/data-software-materials-management-and-sharing-policy>

soient ouvertement disponibles pour examen et reproductibilité¹¹¹. L'absence générale de mandats ou d'exigences formels à cet égard, ainsi que le financement nécessaire pour fournir des résultats de qualité, peuvent faire en sorte de ne pas en tenir compte. Ce qui complique encore les choses, c'est que l'expertise et l'infrastructure en matière de LR dont disposent les chercheurs peuvent être incohérentes et hétérogènes et que de nombreux chercheurs ne sont pas des développeurs de logiciels professionnels étant officiellement formés sur les bonnes pratiques. De nombreux chercheurs construisent et développent des logiciels non pas pour les autres, mais pour leurs recherches personnelles, et peuvent donc avoir tendance à coder rapidement pour résoudre leurs problèmes immédiats, en particulier en sciences humaines et sociales (SHS) où la recherche en solo est encore la norme et où le financement par des équipes plus importantes est rare. Il existe également une pression pour publier ou diffuser un résultat le plus rapidement possible, en avançant dans la course à la publication. Le partage des logiciels peut être considéré comme un avantage pour les concurrents. Une conséquence accidentelle est que le résultat peut ne pas être reproductible ou réutilisable par d'autres, ce qui compromet l'examen par les pairs et l'avancement, et conduit à la reproduction des efforts et au gaspillage des fonds dans de multiples projets.

L'inquiétude que peut susciter l'examen minutieux du processus d'évaluation par les pairs est un autre facteur dissuasif. On admet généralement que le code ouvert permet d'obtenir un meilleur logiciel, à condition qu'une communauté de contributeurs soit disposée à l'améliorer ou à le commenter. GitHub¹¹² est un exemple de plateforme de partage de code qui possède des caractéristiques spécialement conçues pour la communauté scientifique afin de favoriser une collaboration et un partage efficace de code¹¹³. Le partage de code est parfois considéré comme une « arme à double tranchant » par certains chercheurs, car plus il est partagé, plus le risque de voir apparaître des erreurs et des bogues est élevé. Bien qu'en soi cette exposition soit bénéfique pour la science en général, la recherche peut être hautement compétitive, et certains craignent que le processus d'examen ouvert n'entraîne des critiques disproportionnées et sévères jusqu'à compromettre la réputation d'une personne¹¹⁴. Dans un système où les carrières progressent grâce à la publication, on craint qu'une mauvaise interprétation ou qu'une mauvaise utilisation du code, ou que la découverte d'erreurs pourrait compromettre les résultats. Avec une formation suffisante et un accès à des outils tels que les cadres de test des logiciels, la communauté devrait être en mesure d'atténuer ce problème. Avec les licences et l'attribution appropriées, ainsi que l'utilisation régulière des citations de logiciels qui faciliteront le changement de culture nécessaire, les préoccupations relatives au partage des logiciels devraient diminuer. L'émergence de la nomenclature des logiciels (SBOM) est un ajout utile aux projets/codes de LR.¹¹⁵ Grâce à une approche solide de la reproductibilité, les chercheurs, les résultats scientifiques et les ensembles

¹¹¹ Stodden V, Guo P, Ma Z (2013) Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals. PLoS ONE 8(6): e67111.
<https://doi.org/10.1371/journal.pone.0067111>

¹¹² <https://github.com/>

¹¹³ <https://github.blog/2014-05-14-improving-github-for-science/>

¹¹⁴ Cope, J. (2015). Why not to share your data: The “crackpot” argument. eRambler Blog. Récupéré sur <http://erambler.co.uk/blog/crackpot-argument-against-data-sharing/>

¹¹⁵ <https://fossa.com/blog/software-bill-of-materials-formats-use-cases-tools/>

de données bénéficient tous d'une reconnaissance et d'une crédibilité accrues, ce qui renforce le dossier et l'impact de la recherche ainsi que la productivité et la compétitivité d'un individu.

3.6.3.2. Accès aux LR

Le principe d'accessibilité (5) parmi les principes de citation des logiciels de FORCE11 stipule que « les citations de logiciels doivent permettre et faciliter l'accès au logiciel lui-même ainsi qu'aux métadonnées, à la documentation, aux données et aux autres éléments qui lui sont associés et qui sont nécessaires aux humains et aux machines pour utiliser le logiciel référencé en connaissance de cause ». Bien que ce principe n'exige pas que les LR soient librement disponibles, les métadonnées doivent l'être et fournir suffisamment d'informations pour que les LR soient accessibles et utilisés. Pour les logiciels commerciaux, les métadonnées peuvent toujours fournir des informations qui permettent et facilitent l'accès au logiciel particulier, ce qui peut être le numéro de produit d'une société ou un lien vers une page de renvoi contenant des informations permettant d'utiliser le logiciel librement à des fins d'examen universitaire ou de l'acheter. Toutefois, l'accessibilité n'est pas forcément assurée par la simple possibilité de récupération : une description de la manière d'accéder à une version fonctionnelle (par exemple binaire) des LR doit être disponible, ainsi qu'une documentation suffisante et des informations sur les licences. De plus, pour que les LR soient réellement accessibles, il faut fournir des métadonnées sur la façon d'exécuter les LR, y compris les dépendances telles que l'accès aux ensembles de données de recherche, les dépendances des autres logiciels et les environnements d'exécution.

3.6.3.3. Portabilité des LR

Dans le présent document, le concept de portabilité logicielle est défini comme la capacité de déplacer une application de LR entre différents environnements, tels que des plateformes ou des machines, et de sorte que l'application produise les mêmes résultats à partir des mêmes entrées. En raison de la grande diversité des applications logicielles et des dépendances, ce niveau de portabilité peut avoir différentes significations pour différentes personnes¹¹⁶. Dans un scénario, le logiciel n'est portable que si les fichiers exécutables peuvent être exécutés ou réinstallés sur une nouvelle plateforme sans changement : soit la portabilité binaire. Dans un scénario plus courant, le code source est rendu disponible et le programme exécutable peut être construit à partir de ce code pour être réutilisé dans un environnement cible approprié. Il convient de noter que ce scénario implique à la fois l'accès au code et son transport, ainsi que toute conversion qui pourrait être nécessaire (par exemple, la modification du code d'origine pour l'adapter à des architectures de machines, des environnements ou des besoins particuliers d'équipes de recherche locales). Un troisième scénario est celui de la portabilité de niveau intermédiaire, qui implique des aspects des deux scénarios précédents pour atteindre le résultat souhaité. En général, on peut supposer que le coût de transport et d'adaptation du logiciel à un nouvel environnement est inférieur au coût d'un second développement, bien que ce ne soit pas toujours le cas. Une fois qu'une application de LR est portée sur plusieurs environnements, davantage d'utilisateurs en

¹¹⁶ Mooney, J. D. (2004). Developing portable software. Dans *Information Technology* (pp. 55-84). Springer, Boston, MA.

bénéficieront. La portabilité est une question clé pour la réduction des coûts de développement afin que les développeurs puissent consacrer plus de temps à de nouveaux efforts.

La portabilité étant reconnue comme un aspect souhaitable de la qualité des logiciels, une attention accrue est accordée à l'augmentation de la réutilisation et de la reproductibilité de la recherche. On peut également affirmer que la portabilité des logiciels de recherche est elle-même une réponse à la réutilisation et à la reproductibilité. Alors que l'entité principale à porter est généralement un système ou une application logicielle complète, dans certains cas, des composants logiciels auxiliaires tels que des dépendances, des interfaces, des outils, des bibliothèques ou des bases de données peuvent également devoir être portés. La production d'une documentation complète pour le processus de portage doit être considérée comme une activité de reproductibilité qui facilite le développement d'équipes de LR plus qualifiées et expérimentées.

Lorsqu'il s'agit de recherche à forte intensité de calcul, ces environnements informatiques peuvent évoluer très rapidement, avec de nombreuses versions différentes des mêmes systèmes d'exploitation et composants logiciels utilisés simultanément. Dans le cas d'un projet qui dépend d'éléments particuliers de logiciels provenant de sources diverses, il n'est pas rare que les composants changent fréquemment. De nombreux changements (par exemple, des mises à jour de logiciels de niveau inférieur, une modification mineure de la configuration, des changements de systèmes d'exploitation) peuvent entraîner des résultats inattendus et faire échouer¹¹⁷ l'ensemble du flux de calcul ou simplement produire des résultats légèrement ou radicalement différents. En guise de solution, on peut fournir une copie identique du système d'exploitation et des versions exactes de toutes les dépendances logicielles pour que d'autres puissent recréer l'environnement de calcul original. Avec une bonne documentation, les chercheurs peuvent plus facilement créer de nouveaux environnements de recherche et de nouveaux flux de travail et améliorer considérablement la transparence, la portabilité, la réutilisation et la reproductibilité du code. Cependant, l'utilisation d'un grand nombre de ces environnements qui figent le code dans un certain état doit être équilibrée avec des changements souhaitables tels que la correction de bogues et l'ajout de nouvelles fonctionnalités. Parmi les exemples de systèmes actuellement utilisés par les chercheurs pour capturer des environnements de recherche complets figurent les machines virtuelles (par exemple VirtualBox d'Oracle¹¹⁸) et les conteneurs (par exemple Docker¹¹⁹, Podman¹²⁰, Singularity¹²¹, Sarus¹²², NILRC Shifter¹²³, HPC Container Maker

¹¹⁷ Glatard, T., Lewis, L. B., Ferreira da Silva, R., Adalat, R., Beck, N., Lepage, C., et al. (2015). Reproducibility of neuroimaging analyses across operating systems. *Front. Neuroinform.* 9:12. <https://doi.org/10.3389/fninf.2015.00012>

¹¹⁸ <https://www.virtualbox.org/>

¹¹⁹ <https://www.docker.com/>

¹²⁰ <https://podman.io/>

¹²¹ <https://sylabs.io/singularity/>

¹²² <https://user.cscs.ch/tools/containers/sarus/>

¹²³ <https://www.nersc.gov/research-and-development/user-defined-images/>

(HPCCM)¹²⁴). Les Vagrantfiles¹²⁵ et les Dockerfiles¹²⁶ sont des exemples courants d'instructions en texte brut lisible par machine pour créer des machines virtuelles selon une spécification exacte.

3.6.3.4. Réutilisation des LR

La réutilisation des logiciels « consiste à tirer parti des logiciels et des systèmes existants pour créer de nouveaux produits »¹²⁷ et constitue un sujet de recherche depuis plus de 50 ans, depuis sa création lors de la première conférence internationale sur le génie logiciel (ICSE) en 1968¹²⁸. Les chercheurs et les praticiens continuent de développer et d'améliorer les outils, les méthodologies, les processus, les pratiques et les techniques pour accroître la productivité, la qualité et les économies. Une définition semblable considère l'innovation de la recherche elle-même : « La réutilisation des LR consiste à tirer parti de logiciels et de systèmes existants pour d'autres applications ou pour créer de nouvelles *recherches* ». Sous cette définition générale, plusieurs contextes coexistent. Par exemple, il existe deux approches principales du développement des LR : *avec réutilisation* (développement à l'aide de composants préexistants) et *pour réutilisation* (développement de nouveaux composants réutilisables pour la recherche, tels que des bibliothèques logicielles)¹²⁹. La réutilisation des LR peut également être considérée dans un contexte de domaine : la *réutilisation verticale* (réutilisation des LR dans un domaine d'application donné) et la *réutilisation horizontale* (réutilisation des composants de LR dans d'autres applications ou dans plus d'un domaine d'application)¹³⁰. De plus, la réutilisation des LR peut simplement décrire la maintenance et la mise à jour d'une ancienne pièce de LR vers une nouvelle version avancée (de systèmes *utilisés* antérieurement).

En mettant l'accent sur l'accès et la portabilité des LR, la réutilisation des LR améliorera la productivité de la recherche transdisciplinaire de diverses manières. Le premier avantage est l'économie de temps et de ressources, en réduisant le travail redondant nécessaire pour créer de toutes pièces des LR aux fonctionnalités semblables. De plus, grâce à la réutilisation des LR, les chercheurs peuvent consacrer plus de temps à de nouvelles recherches en exploitant les ressources existantes : ils peuvent modifier, mettre à jour, adapter et faire progresser les LR

¹²⁴ <https://github.com/NVIDIA/hpc-container-maker>

¹²⁵ <https://www.vagrantup.com/docs/vagrantfile>

¹²⁶ <https://docs.docker.com/engine/reference/builder/>

¹²⁷ IEEE Standard, 1517-2010 - IEEE standard for information technology-system and software lifecycle processes-reuse processes, 2010.

¹²⁸ McIlroy, M. D., Buxton, J., Naur, P., & Randell, B. (1968, octobre). Mass-produced software components. In Proceedings of the 1st international conference on software engineering, Garmisch-Partenkirchen, Germany (pp. 88-98).

¹²⁹ De Almeida, E. S., Alvaro, A., Lucrédio, D., Garcia, V. C., & de Lemos Meira, S. R. (2005, août). A survey on software reuse processes. Dans *IRI-2005 IEEE International Conference on Information Reuse and Integration, Conf, 2005*. (pp. 66-71). IEEE.

¹³⁰ Martin Griss, Ivar Jacobson, Chris Jette, Bob Kessler, et Doug Lea. 1995. Systematic software reuse (panel) : objects and frameworks are not enough. SIGSOFT Softw. Eng. Notes 20, SI (août 1995), 17-20. DOI : <https://doi.org/10.1145/223427.213969>

existants pour leurs propres besoins de recherche (tout en accordant le crédit approprié aux auteurs du code original). La réutilisation des LR améliorera également la fiabilité, la productivité et la qualité des LR, et favorisera la création d'une communauté en mettant en relation des personnes ayant des intérêts de recherche semblables, créant ainsi un environnement collaboratif propice à l'échange d'idées et à la motivation.

Le degré de réutilisation des LR dépend de divers facteurs, notamment la qualité technique, qui est cruciale pour inciter les autres à réutiliser les LR, ainsi que la qualité de la documentation, les détails de la mise en œuvre, le niveau de portabilité (par exemple, la conteneurisation) et l'accès intégral aux données connexes. Ainsi, les auteurs de code doivent se soucier de la qualité (par exemple, tests, révisions, développement continu, contrôle de version) comme une condition préalable à la réalisation de produits des LR de haute qualité. D'autre part, la réutilisation efficace des LR peut être obtenue par le biais d'une infrastructure de soutien (par exemple, des registres ou des dépôts) pour la publication, la recherche, l'intégration et les références croisées des LR et des actifs connexes. La culture et les facteurs humains (par exemple, l'expérience, les compétences, l'éducation, les incitations et la motivation, etc.

Il convient également de noter que divers langages de flux de travail ont été développés pour aider les chercheurs à documenter, partager, modifier et refaire des expériences, fournissant ainsi un outil essentiel pour une réutilisation efficace des LR. Il existe environ 280 systèmes de gestion de flux de travail — les plus populaires développés pour la science ajoutent la prise en charge du langage de flux de travail commun (CWL) et du CWLProv. Le langage de flux de travail commun est largement adopté dans les environnements de recherche virtuels, en particulier dans les plateformes biomédicales, mais il existe également des applications en hydrologie, en radioastronomie et en physique des hautes énergies. Le langage de flux de travail commun prend en charge les bonnes pratiques, par exemple les fichiers d'entrées et de produits ont tous un PID et des descripteurs riches (métadonnées). Le langage de flux de travail commun prend également en charge Docker et CVMFS, deux systèmes de mise en paquet répandus dans les infrastructures scientifiques.

3.6.3.5. Citation des LR

Les LR sont des éléments essentiels du processus de recherche numérique dans l'ensemble de l'entreprise de recherche, et ils font partie intégrante d'un article de recherche ou d'un ensemble de données pour ce qui est de faciliter la diffusion et la compréhension de la recherche. Il est important que les LR soient correctement cités, car la citation des LR permet aux chercheurs et aux créateurs de LR de publier des LR plutôt que des articles et des données, ou en plus de ceux-ci, afin de participer au système de communication savante et à son système de récompense. Le système de communication savante qui sous-tend la recherche repose sur des mécanismes de citation précis pour garantir l'intégrité de la recherche, accorder le crédit approprié et représenter les relations avec les autres produits de la recherche. Parmi les étapes de la reproductibilité scientifique, les chercheurs doivent être en mesure d'identifier et d'accéder de manière unique à tous les intrants, y compris la version précise des paquets de LR utilisés pour produire les résultats originaux, contribuant ainsi à la transparence et à la traçabilité des résultats de la recherche.

Des systèmes de plus en plus sophistiqués sont construits pour répondre au besoin d'outils de bonnes pratiques pour l'hébergement de code et la collaboration, l'automatisation et la simplification du processus de découverte et de diffusion, la facilitation des recherches dans les documents savants et la garantie de l'interopérabilité. Par exemple, les plateformes de collaboration et de diffusion des LR comprennent [GitHub](#) et [Bitbucket](#), qui sont également destinées au développement collaboratif. [IPOL](#) est une plateforme de publication pour la recherche sur le traitement des images. Pour les versions de distribution de paquets de LR, les plateformes ou registres existants comprennent [PyPI](#) pour Python, [CRAN](#) pour R, [Awesome Research Software Registries](#), [Astrophysics Source Code Library](#) (ASCL) pour l'astrophysique, [MLOSS](#), [DLHub](#) et [OpenML](#) pour les logiciels à code source ouvert d'apprentissage automatique, [Madagascar](#) et [Knowledge Exchange](#). [Software Heritage](#) est une initiative et une infrastructure ouvertes à but non lucratif pour la préservation et l'archivage à long terme des logiciels. Les plateformes et initiatives reproductibles existantes sont notamment [ResearchCompendia](#), [nanoHUB](#), [RunMyCode](#), [Code Ocean](#) et [Whole Tale](#).

L'infrastructure de citation scientifique s'est développée parallèlement à l'édition scientifique de livres et d'articles de revues, de sorte que la reconnaissance et la citation des LR ont bénéficié d'un soutien inégal dans l'écosystème scientifique. Cependant, toutes les disciplines reconnaissent de plus en plus que les LR devraient être cités avec les autres résultats de recherche, bien que les approches varient beaucoup selon la discipline et les éditeurs¹³¹. La citation des LR contribue à l'avancement de la recherche et permet à d'autres chercheurs d'y accéder pour :

- Améliorer la transparence de la recherche et donc la compréhension de ses résultats ;
- Permettre l'examen par les pairs, la vérification et la reproductibilité des résultats de la recherche ;
- Favoriser la collaboration et la réutilisation ;
- Minimiser la duplication des efforts ;
- Encourager l'utilisation à bon escient du travail des autres ;
- Élever les LR au rang d'objet de recherche de « première classe » dans l'écosystème d'IRN.

Le groupe de travail FORCE11 sur la citation des logiciels a élaboré et recommandé un ensemble de six principes de citation pour encourager l'adoption à grande échelle d'une approche cohérente de la citation des logiciels : importance ; crédit et attribution ; identification unique ; pérennité ; accessibilité ; spécificité. Les listes de contrôle Software Citation Checklist for Developers¹³² et Software Citation Checklist for Authors¹³³ fournissent encore plus d'informations pratiques pour ceux qui cherchent à améliorer leur pratique. Le groupe de travail RDA/FORCE11 sur

¹³¹ <https://doi.org/10.7717/peerj-cs.86>

¹³² <https://doi.org/10.5281/zenodo.3482769>

¹³³ <https://doi.org/10.5281/zenodo.3479199>

l'identification du code source des logiciels¹³⁴ a fourni un aperçu de l'état actuel des identifiants pérennes (PID) pour les logiciels. Après avoir analysé les cas d'utilisation et les schémas d'identification de différents domaines universitaires et de l'industrie, le GT a conclu que, plutôt qu'un PID unique, une stratégie de combinaison de plusieurs PID, à la fois extrinsèques (créés par un service externe à l'environnement de développement, comme un DOI) et intrinsèques (comme un code de hachage créé par l'environnement de développement), est nécessaire pour capturer les cas d'utilisation fondamentaux des LR (découvrabilité, accès, pérennité, reproductibilité et réutilisation).

3.6.3.6. Curation des LR

Afin de valider les résultats de la recherche scientifique, les LR doivent rester accessibles et utilisables dans le temps. La curation et la préservation des LR contribueront à l'intégrité et à la transparence des efforts de recherche modernes. Comme d'autres objets numériques, les LR nécessitent une gestion active et continue tout au long de leur cycle de vie, qui peut être défini par deux activités principales : les activités de curation qui permettent la découverte et la récupération des données, le maintien de leur qualité, l'ajout de valeur et la réutilisation dans le temps¹³⁵ ; et les activités de maintenance du logiciel, qui garantissent que le logiciel reste fonctionnel dans le temps. La définition de travail pour la curation des logiciels est la suivante : « *La curation des logiciels englobe les pratiques actives liées à la création, l'acquisition, l'évaluation et la sélection, la description, la transformation, la préservation, le stockage et la diffusion/l'accès/la réutilisation des logiciels sur des périodes courtes et longues* »¹³⁶. En ce sens, la curation et la préservation des LR au fil du temps contribueront à faciliter l'accès et l'utilisation à long terme des LR. La maintenance des logiciels requiert un ensemble de compétences différentes, plus courantes dans la communauté des développeurs et vise à garantir que les technologies et les cadres appropriés (par exemple, les conteneurs avec tout le code associé) sont maintenus, testés et accessibles.

Comme il s'agit d'un ensemble de pratiques émergentes, la curation des LR peut varier. Il est donc crucial d'identifier les besoins et les activités à partir de cas d'utilisation définis. Une exploration plus approfondie des types de LR et la compréhension des multiples intervenants, des utilisations, des fonctionnalités et même de l'emplacement des LR peuvent aider à contextualiser les efforts de curation. Par exemple, les membres du corps enseignant et les étudiants peuvent créer des LR dans le cadre de cours ou de projets, utiliser ou réutiliser des LR dans le cadre de nouvelles recherches : il s'agit de cas d'utilisation très différents qui peuvent nécessiter des approches de curation différentes. Les sites Web sur les projets et les cyberinfrastructures sont de plus en plus courants, et les projets de collaboration repoussent les frontières et les pratiques disciplinaires. Tous ces scénarios reflètent les rôles complexes que les LR peuvent jouer dans l'entreprise scientifique. Comme cas d'utilisation, considérons le processus d'écriture d'un morceau de LR comme outil d'analyse de données. Cet exemple simple

¹³⁴ <https://www.rd-alliance.org/group/rdaforce11-software-source-code-identification-wg/outcomes/use-cases-and-identifier-schemes>

¹³⁵ <http://hdl.handle.net/2142/3493>

¹³⁶ <https://dr maltman.wordpress.com/2017/01/21/guest-post-alex-chassanoff-on-software-curation/>

montre comment les LR peuvent servir à la fois de résultat de recherche à préserver et de moyen méthodologique pour obtenir un résultat d'analyse.

La curation et la préservation du numérique sont un domaine d'étude et de pratique consacré à la garantie d'un accès continu aux artefacts numériques de valeur et d'intérêt durables, mais aussi en font un mélange complexe de politiques, de pratiques et d'infrastructures visant à soutenir le maintien à long terme de l'accès à ces artefacts au-delà de la défaillance des supports ou des changements technologiques. Une approche commune de la curation et de la préservation numériques consiste à caractériser les propriétés significatives ou « l'essence » de l'objet numérique¹³⁷. Quelles sont les composantes essentielles des LR en tant qu'objet savant à organiser et à préserver ? Il est difficile de définir les limites des LR, en partie parce que les LR ont des représentations diverses selon les contextes et les disciplines (voir la section [Que sont les LR](#)). Contrairement aux ensembles de données et aux publications de recherche, les LR sont exécutables, hautement itératifs et souvent dépendants de multiples éléments dynamiques qui sont sensibles aux changements et difficiles à suivre. De plus, les LR sont potentiellement en constante évolution : quelques lignes de code ajoutées peuvent avoir un impact sur de multiples dépendances ; la dépréciation de logiciels et/ou de bibliothèques de logiciels est un phénomène courant dans le développement de logiciels. Les défis de la curation des LR ne se limitent pas à fournir un accès et à garantir la préservation des bits. Permettre une utilisation et une réutilisation adéquates des LR implique des interventions techniques tout au long du cycle de vie pour faciliter différentes stratégies de préservation (par exemple, exécution, migration, émulation, etc.). Le large éventail de cadres et d'approches robustes pour la curation et la préservation des LR en tant qu'objets numériques complexes représente un défi important pour l'accès durable, entravant ainsi la reproductibilité de la recherche.

3.6.3.7. Impact et valeur des LR

Les logiciels sont manifestement essentiels à la recherche, mais dans quelle mesure y recourons-nous ? Une enquête en ligne intitulée « How do scientists develop and use scientific software »¹³⁸, menée en octobre-décembre 2008, a reçu près de 2 000 réponses : 91 % des répondants ont indiqué que l'utilisation de logiciels est importante pour leurs propres recherches ; 84 % des répondants ont indiqué que le développement de logiciels scientifiques est important pour leurs propres recherches. Dans l'enquête UK Research Software Survey de 2014¹³⁹, 70 % des chercheurs ont répondu que leurs recherches seraient impossibles sans logiciels. Trois ans plus tard, l'enquête 2017 de l'US National Postdoctoral Association a donné un résultat similaire (63 %) ¹⁴⁰. De nombreuses personnes de la communauté des chercheurs développent leur propre code, ce qui souligne l'impact des LR dans l'entreprise de recherche.

¹³⁷ Heslop, H., Wilson, A., & Davis, S. (2002). *National archives green paper: an approach to the preservation of digital records*. NAA.

¹³⁸ J.E. Hannay et autres, "How Do Scientists Develop and Use Scientific Software ? *Proc. ICSE Workshop Software Eng. for Computational Science and Eng.* 2009, pp. 1-8.

¹³⁹ Simon Hettrick. (2018, 23 février). *softwaresaved/software_in_research_survey_2014* : Software in research survey (Version 1.0). Zenodo. <http://doi.org/10.5281/zenodo.1183562>

¹⁴⁰ <https://doi.org/10.6084/m9.figshare.5328442.v3>

L'accent mis sur l'impact des revues pour évaluer la valeur et la productivité des chercheurs commence à inclure la reconnaissance du développement de logiciels et d'autres activités dans le cadre des directives de récompense et de promotion universitaires. Par exemple, la Declaration on Research Assessment — DORA (déclaration sur l'évaluation de la recherche) a pour objectif d'améliorer les méthodes d'évaluation des résultats de la recherche universitaire et a été signée par plus de 20 000 personnes et organisations dans 148 pays^{141,142,143}. Comme pour les données et les publications, les facteurs d'impact et les mesures de LR peuvent inclure le nombre de vues, de téléchargements, de citations, d'évaluations par les utilisateurs (p. ex. étoiles GitHub), de contributions à un projet ou à un dépôt (p. ex. forks), de mentions dans les médias sociaux, etc. Les principaux avantages d'une gestion, d'un partage et d'une réutilisation efficaces des LR et des mesures d'évaluation via des métriques précises sont mis en évidence dans les paragraphes suivants.

Les LR facilitent le développement de réseaux durables, comme l'émergence des ingénieurs en logiciels de recherche (ILR). La communauté internationale des ILR a commencé avec UK-RSE¹⁴⁴, qui s'est ensuite élargie avec un certain nombre d'associations internationales de ILR nouvellement formées¹⁴⁵. En tant que PHQ, les ILR peuvent passer d'un projet à l'autre, d'une discipline à l'autre, acquérir une expérience supplémentaire et enrichir la communauté d'ILR dans son ensemble. Les ILR peuvent être exposés à diverses disciplines lorsqu'ils font partie d'une équipe institutionnelle d'ILR ou se concentrer sur une discipline particulière lorsqu'ils font partie de l'équipe d'un chercheur principal. Cet éventail d'expertise en LR et en domaine est un atout d'une communauté qui s'implique de plus en plus dans la recherche interdisciplinaire. Le Colloque canadien sur les logiciels de recherche (CCLR) de CANARIE¹⁴⁶ est un bon exemple d'un événement qui rassemble la communauté de LR de partout au pays pour qu'elle se rencontre et établisse des réseaux favorisant de futures collaborations transdisciplinaires.

Le partage et la réutilisation efficaces des logiciels peuvent potentiellement permettre aux développeurs et aux créateurs de logiciels de gagner en reconnaissance. La citation du logiciel permet de reconnaître l'effort scientifique, facilitant ainsi l'avancement de la carrière et l'augmentation de la réputation professionnelle.

¹⁴¹ <https://sfdora.org/signers/>

¹⁴² <https://www.scienceguide.nl/2019/11/dutch-universities-and-research-funders-move-away-from-the-impact-factor/>

¹⁴³ http://www.cnrs.fr/sites/default/files/press_info/2019-11/Plaqueette_ScienceOuverte.pdf#page=7

¹⁴⁴ <https://society-rse.org/>

¹⁴⁵ Les associations internationales d'ILR formées plus récemment et en cours de formation comprennent de-RSE (<http://de-rse.org/>), nl-RSE (<http://nl-rse.org/>), nordic-RSE (<http://nordic-rse.org/>), us-RSE (<http://us-rse.org/>), RSE-aunz (https://twitter.com/rse_aunz), be-RSE (<https://www.be-rse.org/>), et le groupe d'intérêt spécial DHTech de l'Alliance of Digital Humanities Organizations (<https://dh-tech.github.io/#/>).

¹⁴⁶ <https://www.canarie.ca/software/canadian-research-software-conference/>

Pour illustrer l'impact des LR, prenons l'exemple d'OpenPNM¹⁴⁷, développé par le groupe Porous Materials Engineering and Analysis Lab (PMEAL) de¹⁴⁸ l'Université de Waterloo, un logiciel libre qui fournit aux chercheurs sur les milieux poreux un cadre prêt à l'emploi pour réaliser de nombreuses simulations de réseaux de pores. Le groupe PMEAL s'est engagé dans plusieurs collaborations de recherche depuis son lancement, dont beaucoup n'auraient probablement pas vu le jour autrement. Dans un autre exemple, le Centre McGill de neurosciences intégratives (MCIN) est en train de devenir un chef de file international en développant et en fournissant activement des technologies à forte intensité de calcul (y compris des plateformes et des outils logiciels) qui sont disponibles en libre accès pour faire progresser la recherche collaborative sur le cerveau et les soins cliniques au Canada et chez des partenaires internationaux, ainsi que pour être utilisées par les communautés de recherche canadiennes et internationales¹⁴⁹.

Ces exemples et d'autres au Canada montrent que les LR peuvent avoir un impact considérable sur la réputation dans les communautés de pratique internes et externes, sur le plan national et international.

Les chercheurs s'efforcent d'obtenir des résultats fiables et défendables, ce qui est plus facile à obtenir en utilisant des LR robustes et bien testés. Des LR bien documentés et accessibles facilitent la reproduction et la vérification des résultats par d'autres personnes. Il est dans l'intérêt du Canada de protéger l'important investissement annuel dans la recherche en s'assurant que la communauté des chercheurs possède les compétences nécessaires pour créer des logiciels réutilisables.

Le développement de logiciels peut s'avérer très coûteux, mais il arrive qu'un problème de recherche particulier nécessite un effort de développement « à partir de zéro » ou « d'une friche ». Cependant, dans de nombreux cas, la réutilisation du logiciel peut faire gagner du temps aux chercheurs en exploitant un code qui existe déjà et en résolvant des problèmes qui ont été résolus par d'autres. Le logiciel réutilisable peut être facile à adopter et à étendre, en supposant que la bonne combinaison de compétences est représentée dans une équipe de recherche et que le code réutilisé est bien conçu et mis en œuvre. En s'appuyant sur les logiciels existants, les chercheurs peuvent consacrer plus de temps à la recherche proprement dite et économiser des ressources rares qui peuvent alors être investies dans de nouvelles recherches plutôt que dans le développement de logiciels. Pour garantir un écosystème de recherche robuste et efficace, nous avons besoin à la fois de nouveaux développements de LR et de la réutilisation de LR pour favoriser l'innovation et l'efficacité : la réutilisation consiste à appliquer des techniques connues pour mieux comprendre, tandis que le développement consiste à créer de nouvelles techniques pour mieux comprendre.

L'utilisation d'un ensemble commun de bonnes pratiques de LR, quelle que soit la discipline, peut accélérer l'examen et l'acceptation des résultats de recherche. Cela se manifeste à la fois par la facilité d'accès aux outils de LR utilisés dans des projets de recherche particuliers et par la

¹⁴⁷ <http://openpnm.org/>

¹⁴⁸ <http://www.pmeal.com/>

¹⁴⁹ <https://mcin.ca/technology/>

confiance accrue de la communauté de recherche dans l'exactitude et l'efficacité des LR connexes.

De nombreux types de données de recherche peuvent être dénués de sens sans logiciel pour les lire et les interpréter. Un logiciel bien conçu et durable contribue à garantir l'accès, l'utilisation et la réutilisation à long terme des données de recherche, en tirant parti de la reproductibilité et en assurant le meilleur rendement sur l'investissement réalisé pour produire les données.

Le partage et la réutilisation du code facilitent l'adoption de normes et établissent des communautés de pratique qui peuvent offrir un meilleur soutien par les pairs, la mobilité du personnel entre les projets, etc. Le partage et la réutilisation de codes deviennent l'un des moyens de créer et de maintenir des communautés de recherche. Les LR peuvent également avoir des impacts non scientifiques tels que des avantages économiques et sociaux : nouvelles opportunités d'emploi ; LR en tant que produits finis avec support et services associés ; autres formes de propriété intellectuelle (par exemple, brevets) qui font partie des projets d'innovation. Certains LR propres à un domaine peuvent avoir un impact profond sur l'industrie et l'économie, comme dans le diagnostic médical, les techniques d'ingénierie, les simulations (p. ex. météo, ingénierie, matériaux, épidémiologie, processus chimique), ou la communication et la collaboration. Pour un exemple plus détaillé et actuel à travers le prisme de la COVID-19, voir [l'annexe G](#).

3.6.4. Maintien et pérennité à long terme des LR

La pérennité des logiciels est un élément clé de la science reproductible, car elle fournit un outil essentiel pour l'examen et l'analyse efficaces des résultats publiés, ce qui peut bien sûr conduire à de nouveaux efforts de recherche. Le Software Sustainability Institute¹⁵⁰ définit la pérennité comme suit : « le logiciel utilisé aujourd'hui sera disponible — et continuera d'être amélioré et soutenu — à l'avenir ». Inspirés par le blog de Daniel S. Katz¹⁵¹ et la documentation connexe¹⁵², les auteurs considèrent que le logiciel pérenne est un logiciel facile à maintenir, qui remplit son objectif dans le temps, qui survit à l'incertitude, qui a la capacité de durer et qui soutient les enjeux pertinents (politiques, économiques, sociaux, techniques, juridiques, environnementaux) de sorte qu'il continuera à être disponible à l'avenir, sur de nouvelles plateformes, répondant à de nouveaux besoins. L'importance de la pérennité des logiciels a été soulignée par un certain nombre d'initiatives et d'organisations émergentes telles que le Software Sustainability Institute¹⁵³, Software Carpentry¹⁵⁴, des initiatives de financement de la National Science Foundation aux États-Unis¹⁵⁵ et du Engineering and Physical Sciences Research Council

¹⁵⁰ <https://software.ac.uk/about>

¹⁵¹ <https://danielskatzblog.wordpress.com/2016/09/13/defining-software-sustainability/>

¹⁵² <http://wssspe.researchcomputing.org.uk/wssspe4/>

¹⁵³ <https://software.ac.uk/about>

¹⁵⁴ <https://software-carpentry.org/>

¹⁵⁵ NSF A Vision and Strategy for Software for Science, Engineering, and Education. Disponible à https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf12113

(EPSRC) au Royaume-Uni^{156,157}. De plus, un certain nombre d'ateliers^{158,159,160,161,162,163,164} ont vu le jour, qui se consacrent à l'exploration du thème des logiciels et systèmes pérennes à partir d'une série de perspectives différentes.

La création de logiciels pérennes peut être un défi. Bien qu'il existe de nombreuses bonnes pratiques associées à la pérennité, il n'y a pas d'ensemble unique de bonnes pratiques acceptées. Ce qui est convenu, c'est que les logiciels doivent être maintenus, sinon ils représentent un risque pour le projet et la communauté en général. Dans le cas des passerelles scientifiques, cela va plus loin : elles doivent être activement exploitées, entretenues, surveillées, corrigées et sécurisées, sinon elles ne peuvent pas garantir le soutien de la communauté qu'elles servent et peuvent devenir un risque de cybersécurité. Dans un cas comme dans l'autre, le maintien des LR exige également du personnel hautement qualifié, car cela ne se fait pas tout seul.

Il existe un certain nombre de guides de bonnes pratiques qui fournissent des recommandations pour travailler à toutes les étapes du cycle de vie des LR, par exemple :

- Le guide complet de développement de logiciels du Netherlands eScience Center¹⁶⁵, qui comprend une liste de contrôle utile pour chaque étape du développement.
- Le guide de qualité des logiciels de CLARIAH (Common Lab Research Infrastructure for the Arts and Humanities)¹⁶⁶, qui peut être utilisé à la fois par les développeurs et les utilisateurs de logiciels pour évaluer la qualité et la pérennité des logiciels.
- Les lignes directrices German Aerospace Centre's Software Engineering Guidelines¹⁶⁷ pour soutenir l'ingénierie logicielle pérenne dans la recherche.

¹⁵⁶ <https://webarchive.nationalarchives.gov.uk/20200930204942/https://epsrc.ukri.org/funding/calls/csesoftwareforthefuture/>

¹⁵⁷ <https://webarchive.nationalarchives.gov.uk/20200701162710/https://epsrc.ukri.org/funding/calls/softwareforthefuture/>

¹⁵⁸ <https://sciencegateways.org/engage/focus-week/jumpstart#agenda>

¹⁵⁹ <http://wssspe.researchcomputing.org.uk/welcome/>

¹⁶⁰ <http://birgit.penzenstadler.de/re4susy/>

¹⁶¹ <https://www.knowledge-exchange.info/event/software-sustainability>

¹⁶² https://www.software.ac.uk/attach/Research_Software_Sustainability_Report_on_KE_Workshop_Feb_2016_FINAL.pdf

¹⁶³ <https://software.ac.uk/programmes-and-events/collaborations-workshops>

¹⁶⁴ <https://wosss.org/wosss21-home/>

¹⁶⁵ <https://doi.org/10.5281/zenodo.4020565>

¹⁶⁶ <https://github.com/CLARIAH/software-quality-guidelines>

¹⁶⁷ <https://doi.org/10.5281/zenodo.1344612>

- Les outils de planification de l'amélioration de la productivité et de la pérennité (Productivity and Sustainability Improvement Planning - PSIP)¹⁶⁸ qui comportent une collection de documents soutenant un processus de planification itératif aidant à améliorer la productivité des développeurs de logiciels scientifiques et la pérennité des produits grâce à des pratiques, des processus et des outils logiciels améliorés.

4 Main-d'œuvre pour les LR

Dans le contexte des LR, le personnel hautement qualifié désigne les professionnels qui soutiennent et effectuent des recherches, notamment les chercheurs (notamment les chercheurs en début de carrière, les étudiants et les postdoctorants) et le personnel de soutien à la recherche (par exemple, les administrateurs système, les professionnels techniques de la recherche [PTR], les développeurs de logiciels). Le terme PHQ sert régulièrement à désigner deux types de personnel hautement qualifié : le nouveau PHQ via le développement du PHQ dans le cadre de la recherche financée et le PHQ qui est membre d'une équipe de recherche ou d'une organisation d'infrastructure existante et déjà capable de mener des activités de recherche ou de soutenir la recherche.

Les LR sont développés par le biais d'un réseau hautement distribué impliquant divers acteurs, notamment des chercheurs, des étudiants, des ingénieurs en logiciels de recherche (ILR), des bibliothécaires, des administrateurs d'université, des professionnels d'organisations propres à une discipline et d'organisations à but non lucratif ainsi que des fonctionnaires. Dans la figure 3, les utilisateurs et les communautés de recherche, qui sont les agents engagés dans les activités de LR, sont omis dans toutes ces activités. Avant tout, les utilisateurs (principalement les chercheurs) sont les acteurs essentiels engagés dans les diverses activités tout au long du cycle de vie de la recherche. En ce qui concerne les acteurs, on peut considérer 3 grands groupes d'utilisateurs et de communautés : 1) ceux qui construisent les passerelles scientifiques (par exemple, les développeurs et les équipes de logiciels de recherche) ; 2) ceux qui créent ou déploient les outils logiciels (par exemple, le personnel d'administration des systèmes, les développeurs logiciels) et soutiennent les chercheurs dans leur utilisation ; 3) les chercheurs (par exemple, les non-professionnels ou les scientifiques utilisant des logiciels) qui effectuent des activités de recherche en utilisant les LR en accédant soit aux passerelles par le biais de leurs propres ordinateurs ou ressources locales ou les outils logiciels susmentionnés. Étant donné le caractère relativement nouveau de la communauté de LR en tant que communauté de pratique, ces 3 groupes sont généralement isolés les uns des autres. La communauté qui en résulte est diverse et se prête au développement du PHQ (par exemple, les chercheurs débutants, les stagiaires, etc.). Cette juxtaposition d'outils et d'acteurs facilite également le soutien des communautés, la création d'équipes collaboratives, voire la mise en évidence de synergies entre des groupes multidisciplinaires s'intéressant à de grands défis.

Lorsque nous pensons aux chercheurs qui bénéficient du potentiel offert par les LR, nous avons tendance à imaginer ceux qui travaillent dans des domaines liés aux STIM ; cependant, les LR

¹⁶⁸ <https://betterscientificsoftware.github.io/PSIP-Tools/>

sont de plus en plus utilisés dans les arts, les sciences humaines et les sciences sociales. Comme il est indiqué dans le Future National Digital Research Infrastructure Landscape¹⁶⁹, les LR sont une composante indispensable d'IRN enchevêtrés avec les composantes de la GDR et du CIP dans toutes les disciplines, ce qui entraîne une transformation radicale de l'enseignement supérieur, du paysage universitaire et de la pratique de la recherche, et garantit que les chercheurs canadiens demeurent concurrentiels à l'échelle mondiale et capables de participer à des collaborations internationales. Les chercheurs et leurs équipes de recherche connexes ne se contentent pas d'utiliser les LR ; ils en sont aussi les principaux producteurs. Certaines disciplines, comme la biologie, l'informatique et la statistique, ont donné naissance à des sous-disciplines, comme la bio-informatique et la science des données, qui ont leurs propres revues, leur propre financement et des communautés dynamiques. Dans les sciences humaines, les personnes possédant à la fois une expertise en recherche et des compétences techniques jouent un rôle essentiel en traduisant les enquêtes des sciences humaines en processus auxquels les ordinateurs peuvent contribuer¹⁷⁰. Le développement des LR provient en grande partie de deux groupes : 1) les développeurs de logiciels hautement qualifiés qui travaillent avec des groupes de recherche et sont généralement employés comme ingénieurs ou développeurs de logiciels et sont de plus en plus appelés ILR ; et 2) les chercheurs autodidactes qui ont commencé comme chercheurs postdoctoraux, membres du personnel de recherche ou étudiants des cycles supérieurs.

De plus, le Conseil des dirigeants principaux de l'information des universités canadiennes¹⁷¹(CUCCIO) constitue une autre catégorie d'utilisateurs et de communautés qui sont étroitement engagés dans la prestation de services de LR à l'échelle nationale. Le CUCCIO est une organisation fédérale à but non lucratif, financée par ses membres et constituée en personne morale, qui regroupe les associations provinciales et régionales de dirigeants de TI du secteur de l'enseignement supérieur au Canada. Le CUCCIO aide le milieu universitaire du Canada à progresser grâce à l'utilisation innovatrice et efficace des TI. Les 62 universités membres du CUCCIO tirent parti des occasions de collaborer, de faire progresser et de favoriser les bonnes pratiques, de partager l'information, d'explorer de nouvelles idées et d'établir et de maintenir des relations avec d'autres groupes pour faire avancer des intérêts communs. En tant que seule conférence nationale pour les professionnels des TI dans l'enseignement supérieur, la conférence annuelle de la CUCCIO, CANHEIT (Canadian Higher Education Information Technology)¹⁷², est organisée par l'une de ses universités membres et se concentre sur les questions de TI qui intéressent les universités et les collèges au Canada. Puisque le CUCCIO est membre associé de l'Alliance, l'Alliance a la possibilité de se concerter et de collaborer avec le CUCCIO pour faire progresser le volet des LR.

¹⁶⁹ <https://www.ic.gc.ca/eic/site/136.nsf/eng/home>

¹⁷⁰ Quinn Dombrowski, Tassie Gniady, David Kloster, Megan Meredith-Lobay, Jeffrey Tharsen et Lee Zickel, *Voices from the Server Room : Humanists in HPC, Debates in the Digital Humanities*, à paraître.

¹⁷¹ <https://www.cuccio.net/en/>

¹⁷² <http://wps-canheit-prod.concordia.ca/>

5 LR au Canada

Les LR sont élaborés et offerts par le biais d'un réseau hautement distribué impliquant divers acteurs et intervenants, notamment des personnes (par exemple, des utilisateurs et des communautés), des organisations et des associations, des décideurs (par exemple, des bailleurs de fonds et des éditeurs), des universités, des secteurs privés et des gouvernements, etc. Compte tenu du nombre d'acteurs, l'approche par région peut aider à s'y retrouver : locale, régionale/provinciale, nationale et internationale, où chaque niveau présente un aspect de soutien ou d'habilitation, souvent de manière différente et critique.

Ce chapitre se comporte de deux volets. La section 7.1 donne un aperçu des organisations et associations existantes, des établissements postsecondaires, des hôpitaux de recherche et des instituts affiliés, des secteurs privés et même des revues qui jouent un rôle dans le soutien ou la prestation de services de LR au Canada. La section 7.2 résume les bailleurs de fonds qui soutiennent la création et l'utilisation des LR.

5.1. LR dans l'IRN et l'Alliance

Pour s'aligner sur l'objectif de l'Alliance de soutenir un secteur d'infrastructure de recherche numérique agile et collaboratif au Canada, les activités nationales de LR au sein de l'Alliance, en parallèle avec les composantes de GDR et de CIP, sont consacrées à l'amélioration des capacités des chercheurs afin d'offrir l'accès à des logiciels de recherche appropriés, robustes, évolutifs et pérennes, et intégrés dans un cadre de cybersécurité amélioré.

La recherche moderne à forte intensité de données et de calcul est généralement menée à l'aide d'infrastructure et de systèmes de CIP, où d'énormes quantités de données sont stockées et traitées par les LR. Une part importante des LR est hébergée dans les systèmes de CIP, de sorte que les composants du CIP et des LR peuvent être considérés comme ayant une interaction à dépendance réciproque. Les activités de développement associées au débogage, à l'analyse des données, à la modélisation numérique à grande échelle et à la visualisation sont de plus en plus souvent réalisées directement sur des données également stockées dans cet environnement de CIP. L'accès au CIP transforme la façon dont nous menons nos recherches, fournissons des services, ainsi que des conseils et du soutien ; les capacités de CIP sont donc déterminées par les besoins des processus de LR et de GDR. D'un autre point de vue, la conception, le développement et le déploiement de l'infrastructure de base de CIP reposent eux-mêmes sur le logiciel, et les systèmes de CIP massivement parallèles et dotés d'accélérateurs nécessitent des produits de LR très sophistiqués et hautement évolutifs. Les relations entre les différents éléments des LR, de la GDR et du CIP sont fondamentalement interconnectées et interdépendantes à toutes les couches de l'écosystème. Nous ne pouvons pas avoir l'un sans les autres, et nous ne pouvons plus nous permettre d'avoir chacun d'entre eux fonctionnant dans leurs propres environnements cloisonnés.

La cybersécurité peut être considérée comme une composante intégrale de l'écosystème d'IRN, englobant les mécanismes, les stratégies, les politiques et les pratiques qui empêchent l'accès non autorisé aux actifs numériques, et englobant les solutions logicielles et matérielles pour une vaste gamme de réseaux de communication avec ou sans fil. La cybersécurité est promue conjointement par l'Alliance et CANARIE en concertation avec les établissements locaux et le

réseau mondial de recherche et d'éducation (GREN) par l'exportation d'innovations dans les plateformes, les logiciels et les services aux chercheurs canadiens.

5.2. Fournisseurs de services

Pour que les chercheurs soient en mesure de trouver, de réutiliser et de gérer efficacement les LR et d'avoir accès à ceux-ci, il faut un cadre commun de politiques, de procédures, de protocoles, de normes, de métadonnées et d'IRN qui sont partagés dans un paysage distribué de personnes et d'organisations. Cela est particulièrement vrai dans une confédération comme le Canada, où l'autonomie substantielle des provinces dans les domaines de la recherche et de l'éducation signifie que la coordination des efforts à l'échelle nationale est essentielle, y compris l'accord national sur les cadres et les pratiques. Le milieu canadien de la recherche a reconnu ce besoin, et deux initiatives (voir les [sections 7.1.1](#) et [7.1.2](#)) ont joué un rôle de premier plan en permettant de faciliter et de coordonner les activités liées aux LR au Canada à l'échelle nationale.

5.2.1. CANARIE

Fondé en 1993, CANARIE (anciennement le Réseau canadien pour l'avancement de la recherche, de l'industrie et de l'éducation) est un organisme à but non lucratif qui coordonne le Réseau national de la recherche et de l'éducation (RNRE) du Canada comprenant le réseau à très haute vitesse qui relie les chercheurs, les éducateurs et les innovateurs du Canada entre eux et avec le reste du monde. CANARIE reçoit la majeure partie de son financement du gouvernement du Canada. Poursuivant sa mission de concevoir, de fournir et de favoriser l'adoption d'une infrastructure numérique pour la recherche, l'éducation et l'innovation, CANARIE finance et encourage le développement d'outils logiciels pour la recherche ; soutient l'élaboration de politiques, d'infrastructures et d'outils pour la gestion des données de recherche ; collabore au financement, à la mise en œuvre, au soutien et au renforcement des initiatives en matière de cybersécurité ; fournit des services de gestion de l'accès et de l'identité à la communauté scientifique ; et aide les jeunes entreprises et les PME canadiennes en leur offrant des ressources infonuagiques et une expertise dans les technologies émergentes.

Le programme Logiciels de recherche de CANARIE¹⁷³

Le programme Logiciels de recherche de CANARIE a débuté en 2007 comme mécanisme pour financer le développement d'outils logiciels de recherche qui accélèrent la découverte en offrant l'accès à l'infrastructure numérique, ainsi que pour permettre aux chercheurs de se concentrer sur la recherche plutôt que sur la technologie de soutien. Les logiciels fournis avec l'aide du programme Logiciels de recherche de CANARIE sont conçus de manière à promouvoir les bonnes pratiques, notamment celles qui favorisent la réutilisation des codes afin d'éviter la duplication des efforts de développement de logiciels. Tous les logiciels financés par CANARIE sont accessibles via le [Portail des logiciels de recherche de CANARIE](#). Ce est un endroit central où les chercheurs et les développeurs de logiciels de recherche peuvent découvrir les logiciels fournis par les participants au programme et par les membres de la communauté scientifique. Tous les logiciels enregistrés sur le portail peuvent être utilisés gratuitement par n'importe qui, n'importe où, à des fins de recherche. Le programme Logiciels de recherche de CANARIE investit et soutient à la fois les [plateformes de logiciels de recherche](#) et les [services de logiciels de](#)

¹⁷³ <https://www.canarie.ca/software/>

recherche dans des registres distincts, dans le but d'accélérer la découverte en exploitant les ressources de l'infrastructure numérique. Le Portail des logiciels de recherche est ouvert à tous les logiciels de recherche canadiens, pas seulement aux projets financés par CANARIE. Ainsi, tout développeur de logiciels peut contribuer au registre en y versant des plateformes et des services à l'usage des autres équipes de recherche. Depuis sa création, le programme Logiciels de recherche a accordé 50,7 millions de dollars de financement aux équipes de recherche canadiennes et a facilité le développement de la communauté de LR du Canada.

Le programme Logiciels de recherche de CANARIE est essentiel à l'avancement et au renforcement des capacités de l'écosystème canadien des logiciels de recherche. Le succès de certaines plateformes (p. ex. 3 D Slicer¹⁷⁴, CANFAR¹⁷⁵) que ce programme a financées est non seulement précieux pour les communautés individuelles, mais il est aussi important pour l'écosystème des logiciels de recherche au Canada aujourd'hui. En plus de financer les communautés qui utilisent beaucoup de données et d'ordinateurs, CANARIE joue un rôle crucial dans le financement de domaines non traditionnels (p. ex. le CSÉC¹⁷⁶ et Voyant¹⁷⁷ dans le domaine des sciences humaines numériques) pour lesquels il y avait moins de sources de financement^{178,179,180,181}. L'effet que le Programme a eu en rassemblant des chercheurs de différents domaines pour discuter des pratiques de LR est important : l'accent mis sur la création de LR et l'établissement d'une culture et d'une communauté de LR est la clé qui le différencie des autres programmes.

Afin de promouvoir la réutilisation et les bonnes pratiques en matière de développement de logiciels de recherche, CANARIE organise depuis 2018 un événement annuel, toutes disciplines confondues, l'atelier sur les logiciels de recherche, plus récemment rebaptisé le CCLR¹⁸². Le CCLR rassemble une communauté de professionnels, de développeurs et d'ingénieurs de LR de tout le pays pour échanger des idées et favoriser la collaboration. CANARIE fait la promotion du programme Logiciels de recherche auprès de la communauté par le biais de bulletins d'information périodiques, d'articles de blogue¹⁸³ et de documents.

CANARIE est également un participant actif de la communauté internationale des logiciels de recherche. Par exemple, CANARIE a participé à l'International Research Software Engineering

¹⁷⁴ <https://www.slicer.org/>

¹⁷⁵ <https://www.canfar.net/en/>

¹⁷⁶ <https://cwrc.ca/fr>

¹⁷⁷ <https://voyant-tools.org/>

¹⁷⁸ https://engagedri.ca/assets/documents/whitepapers/CSDH_NDRIO_WhitePaper.pdf

¹⁷⁹ https://engagedri.ca/assets/documents/whitepapers/CADC_NDRIO_White_Paper.pdf

¹⁸⁰ https://engagedri.ca/assets/documents/whitepapers/NDRIO_White_Paper_Medical_Imaging_Consortium.pdf

¹⁸¹ https://engagedri.ca/assets/documents/whitepapers/SBrown-Sustaining_DRI_in_the_Humanities.pdf

¹⁸² <https://www.canarie.ca/event/crsc-2021/>

¹⁸³ <https://www.canarie.ca/category/canarie-blog/>

Leaders Workshop¹⁸⁴ et a été un membre fondateur de la Research Software Alliance¹⁸⁵ (ReSA). En 2017¹⁸⁶ et 2018¹⁸⁷, en collaboration avec le Software Sustainability Institute¹⁸⁸ (SSI) du Royaume-Uni, CANARIE a mené sondage annuel des développeurs de logiciels¹⁸⁹ afin de mieux comprendre les besoins de la communauté canadienne des logiciels de recherche. Les résultats du sondage ont permis de faire évoluer et d'orienter la stratégie du programme Logiciels de recherche de CANARIE, pour faire en sorte que les chercheurs canadiens aient accès à des outils logiciels de calibre mondial pour accélérer leurs recherches. La préparation du sondage sur les logiciels de recherche canadiens de 2021 est en cours.

En 2018, CANARIE a lancé l'initiative Soutien local aux logiciels de recherches (SLRL)^{190,191}. Modelé en partie sur les efforts internationaux, le financement est fourni aux équipes locales de logiciels de recherche pour développer, maintenir, soutenir et faire évoluer les LR afin de doter les chercheurs canadiens d'outils et d'expertise en matière de LR au sein de l'établissement. L'objectif de l'initiative Soutien local aux logiciels de recherches est de promouvoir la valeur d'une équipe de LR solide et d'encourager le soutien à long terme au sein de l'établissement. On espère également qu'une fois que de telles équipes seront en place dans toutes les régions, l'efficacité du développement local de logiciels pourra être étendue à l'échelle nationale et internationale. L'Université McMaster est l'un des trois projets pilotes financés par l'initiative SLRL de CANARIE, qui a été développée en 2018 et suivie d'un projet pilote de deux ans. L'équipe de l'Université McMaster est maintenant autofinancée dans le cadre de l'initiative Research & High Performance Computing — RHPC (Recherche et calcul haute performance), qui a fait évoluer le pilote en finançant des équipes de LR pour travailler directement avec les chercheurs. Les deux autres équipes pilotes de l'Université de Regina et de l'Université Carleton se sont également engagées à maintenir les équipes locales après la fin du financement de CANARIE. De plus, les trois équipes ont ajouté des membres (ILR, personnel de soutien à la recherche dans un domaine, etc.) que CANARIE n'avait pas financés pendant la période pilote, en réponse à la forte demande. Un deuxième appel a été lancé en 2020¹⁹² après la réussite du premier projet pilote.

¹⁸⁴ <https://researchsoftware.org/>

¹⁸⁵ <https://www.researchsoft.org/>

¹⁸⁶ https://github.com/canarie/developer_survey_2017

¹⁸⁷ https://github.com/canarie/developer_survey_2018/blob/master/Results%20Graphs%20-%202018%20Canadian%20Research%20Software%20Developers%20Survey.pdf

¹⁸⁸ <https://www.software.ac.uk/>

¹⁸⁹ https://github.com/canarie/developer_survey_2018

¹⁹⁰ <https://www.canarie.ca/program-news/research-software/>

¹⁹¹ <https://www.canarie.ca/rdm-funding-call-recipients/>

¹⁹² <https://www.canarie.ca/to-equip-canadian-researchers-with-powerful-software-tools-and-expertise-canarie-awards-3-6m-to-six-institutions-to-fund-local-research-software-development-teams/>

Le programme Gestion des données de recherche de CANARIE¹⁹³

Le programme Gestion des données de recherche (GDR) de CANARIE (lancé en 2015) soutient et facilite l'adoption de bonnes pratiques en matière de gestion des données de recherche¹⁹⁴, conformément aux normes internationales, notamment en finançant le développement de logiciels. Le programme de financement de CANARIE est axé sur la GDR, bien que plusieurs des mêmes projets aient été financés par les programmes LR et GDR, ce qui illustre le chevauchement entre les deux. Il y a eu deux appels de propositions pour du financement de GDR de CANARIE, avec des montants de financement de 3,2 millions de dollars et 2 millions de dollars en 2018 et 2019 respectivement. L'appel de propositions de GDR 1¹⁹⁵ était axé sur les composantes et les outils de logiciel qui appuient fortement les principes FAIR et l'encadrement des services de données nationaux¹⁹⁶ (ESDN). Au total, neuf équipes ont été financées¹⁹⁷. L'appel 2¹⁹⁸ sollicitait des propositions qui continueraient à se concentrer sur le respect des principes FAIR et de l'ESDN en ajoutant un soutien pour améliorer l'interopérabilité des dépôts de données et des systèmes, sur le plan national et international. L'appel a permis de financer 4 équipes de recherche¹⁹⁹.

À l'instar de CANARIE qui accueille le CSRC, le programme GDR a organisé des ateliers sur la gestion des données de recherche en 2019²⁰⁰ et 2020²⁰¹ ayant permis de partager et d'échanger des idées, et de promouvoir les possibilités d'accroître l'interopérabilité entre les infrastructures.

CANARIE transférera les programmes Logiciels de recherche et Gestion des données de recherche à l'Alliance d'ici le 31 mars 2022.

¹⁹³ <https://www.canarie.ca/rdm/>

¹⁹⁴ <https://www.force11.org/group/fairgroup/fairprinciples>

¹⁹⁵ <https://www.canarie.ca/rdm/funding/funding-information-rdm-call-1/>

¹⁹⁶ <https://doi.org/10.5281/zenodo.1035843>

¹⁹⁷ <https://www.canarie.ca/rdm-funding-call-recipients/>

¹⁹⁸ L'appel de proposition de GDR 2 est toujours en cours : <https://www.canarie.ca/rdm/funding/call2/>

¹⁹⁹ <https://www.canarie.ca/canarie-awards-2m-to-research-teams-to-extend-the-interopability-of-research-data/>

²⁰⁰ <https://www.canarie.ca/rdm/workshop2020/workshop2019/>

²⁰¹ <https://www.canarie.ca/rdm/workshop2020/>

5.2.2. Fédération Calcul Canada (FCC)

La Fédération Calcul Canada (FCC) est un partenariat entre Calcul Canada²⁰², quatre consortiums régionaux (ACENET²⁰³, Calcul Québec²⁰⁴, Compute Ontario²⁰⁵ et WestGrid²⁰⁶) et des établissements partout au Canada. La FCC fournit la plateforme nationale de CIP du Canada conçue pour accélérer la recherche et l'innovation en coordonnant le déploiement de systèmes de CIP, d'infrastructures numériques, de stockage et de solutions logicielles de pointe. Les services et les ressources de la FCC sont librement accessibles aux établissements de recherche universitaires et aux chercheurs et à leurs collaborateurs du Canada et ont joué un rôle important dans le maintien et le développement du rôle de leader du pays sur la scène internationale de la recherche.

La FCC propose un catalogue des outils, bibliothèques et environnements de LR disponibles^{207,208}. En 2017, la FCC a adopté un dépôt de logiciels commun qui a été mis à disposition par le biais de son groupe de secteurs de CIP et a ainsi fourni un environnement logiciel unifié aux chercheurs. Cet amoncellement logiciel peut être déployé par les utilisateurs²⁰⁹ sur leurs propres ordinateurs ou groupes de secteurs, ce qui leur permet d'utiliser plusieurs versions du même logiciel et d'exécuter leur code simultanément dans plusieurs instances, exactement comme ils le feraient sur les systèmes de la FCC. De cette façon, tous les utilisateurs de la FCC peuvent télécharger le même amoncellement logiciel et l'utiliser selon leurs besoins. Cela garantit une meilleure portabilité et reproductibilité des analyses scientifiques. Le dépôt de logiciels de la FCC s'appuie sur une technologie bien établie développée par la communauté des physiciens du CERN²¹⁰, CernVM-FS²¹¹(CVMFS), pour déployer un système de fichiers distribué sur les groupes de secteurs nationaux. Grâce à ce système distribué de fichiers, des centaines d'outils et de bibliothèques de logiciels peuvent être installés et déployés sur les sites nationaux de manière fiable et efficace, les versions des logiciels peuvent être suivies de manière centralisée et les procédures d'installation peuvent être automatisées. Pour accéder à un logiciel spécifique, il suffit aux utilisateurs d'utiliser une commande pour charger le module approprié, qui charge ensuite le logiciel sur les nœuds de calcul. Grâce à ce dépôt distribué de logiciels, la FCC est en mesure de fournir à la communauté des chercheurs universitaires canadiens un accès facile à un grand nombre d'outils logiciels de recherche génériques et propres à un domaine.

²⁰² <https://www.computecanada.ca/>

²⁰³ <http://www.acceleratediscovery.ca/>

²⁰⁴ <http://calculquebec.ca/en/>

²⁰⁵ <http://computeontario.ca/>

²⁰⁶ <https://www.westgrid.ca/>

²⁰⁷ https://docs.computecanada.ca/wiki/Available_software

²⁰⁸ https://docs.computecanada.ca/wiki/Standard_software_environments

²⁰⁹ https://docs.computecanada.ca/wiki/Accessing_CVMFS#Installation

²¹⁰ <https://home.cern/>

²¹¹ <https://cernvm.cern.ch/fs/>

La FCC s'efforce également de recréer cet environnement dans les nuages publics²¹², en tirant parti d'un projet interne de logiciel libre appelé Magic Castle. Ce cadre prend des ressources de calcul infonuagique (par exemple AWS, Microsoft Azure, Google Cloud, OpenStack et OVH²¹³) et les convertit en un groupe de secteurs virtuel ou une infrastructure de CIP. Après le déploiement, l'utilisateur dispose d'un environnement logiciel de CIP complet qui comprend le dépôt de logiciels de la FCC, l'environnement d'ordonnancement Slurm²¹⁴, un point de terminaison Globus, JupyterHub, LDAP, DNS et plus de 3 000 applications logicielles de recherche. Le personnel de Calcul Canada utilise ce logiciel pour déployer des groupes de secteurs éphémères à des fins de formation ; il devient également populaire auprès des utilisateurs pour le développement, les tests et l'intégration continue.

Dans le but d'aider leurs activités, les experts techniques de la FCC installent, configurent et développent de nombreux systèmes logiciels, allant d'une base de données (base de données de Calcul Canada²¹⁵, appelée CCDB) pour suivre les comptes d'utilisateurs, les ressources d'infrastructure, les allocations de calcul et les concours d'allocation de ressources, à un planificateur de charge de travail sophistiqué, des cadres de surveillance et d'optimisation, des systèmes de stockage et de gestion de groupes de secteurs avancés, des outils et cadres de sécurité, jusqu'à un système de tickets pour l'assistance aux utilisateurs. L'étendue de l'expertise logicielle et des systèmes présents au sein de la FCC est vaste et constitue un écosystème complexe au service de la recherche universitaire canadienne.

Avec plus de 200 experts de 37 établissements partenaires au Canada, la FCC apporte un soutien direct aux chercheurs canadiens en offrant des consultations, un soutien d'experts et des formations régionales²¹⁶ sur de nombreux sujets. Ses services nationaux²¹⁷ couvrent différents domaines allant de l'allocation des ressources au déplacement des données (Globus²¹⁸), et de la visualisation aux sciences humaines numériques. La FCC s'associe également à des chercheurs pour des projets comportant une composante de CIP, en fournissant des ressources en nature (du temps de programmation notamment) pendant une durée déterminée. Par exemple, SHARCNET (l'un des consortiums de Compute Ontario) accorde depuis 2008 un concours de soutien à la programmation dédié²¹⁹ aux chercheurs dans le cadre de projets informatiques. Au total, ce programme populaire a soutenu 43 projets de recherche de 10 établissements différents au cours des 13 dernières années. Chaque projet bénéficie généralement de 4 à 6 mois de soutien, à hauteur de 50 % du temps d'un programmeur. D'autres partenaires de la FCC

²¹² https://github.com/ComputeCanada/magic_castle

²¹³ <https://www.ovh.com/ca/en/>

²¹⁴ <https://slurm.schedmd.com/overview.html>

²¹⁵ <https://ccdb.computecanada.ca/security/login>

²¹⁶ <https://www.computecanada.ca/research-portal/technical-support/training/>

²¹⁷ <https://www.computecanada.ca/research-portal/national-services/>

²¹⁸ <https://docs.computecanada.ca/wiki/Globus>

²¹⁹ <https://www.sharcnet.ca/my/research/programming>

proposent des programmes semblables, en fonction de la demande locale et des ressources disponibles. Outre le fait que tous les chercheurs peuvent accéder à un nombre de ressources par défaut, ceux qui travaillent sur des projets de plateformes logicielles plus importants peuvent demander encore plus de ressources de calcul et de stockage dans le cadre du Concours pour l'allocation de ressources aux plateformes et portails de recherche (PPR) de la FCC²²⁰.

5.2.3. Hébergement et soutien institutionnel

Des établissements postsecondaires, des hôpitaux de recherche et des établissements affiliés de partout au Canada ont commencé à investir dans la prestation de services de LR sur les campus, en offrant un soutien direct à leurs chercheurs, tant sur le plan de l'infrastructure que de l'accès aux outils, aux ressources, aux plateformes de recherche, à l'expertise et à la formation en LR. On s'attend à ce que le nombre d'universités offrant ces services et l'ampleur des investissements nécessaires pour les soutenir augmentent considérablement à mesure que de nouvelles formes d'IRN émergent (collaboration, publication et méthodes de visualisation et d'analyse des données notamment). Il y a un certain nombre de bureaux dans les établissements d'enseignement supérieur qui ont déjà ou auront la responsabilité de soutenir ces efforts à l'échelle locale, notamment :

- Les laboratoires nationaux situés sur les campus (par exemple, les laboratoires du Conseil national de recherches du Canada [CNRC]²²¹) ;
- Les bureaux de recherche (c'est-à-dire les vice-présidents chargés de la recherche (VPR), le personnel des bureaux de recherche, les comités de politique de recherche) ;
- Les départements des technologies de l'information (DTI) ou les services informatiques (c'est-à-dire les directeurs de l'information (DI), le personnel informatique universitaire, les comités institutionnels de gouvernance des données) ;
- Bibliothèques (c.-à-d. bibliothécaires spécialisés/de données, experts en métadonnées, curateurs/archivistes, développeurs et spécialistes du soutien, comités de gestion des données) ;
- Organisations externes (par exemple, organisations de services externes « dérivées » de l'établissement, mais dédiées au soutien de l'établissement, telles que Indoc Research²²² à l'Université de Toronto, Software Carpentry Workshop à l'Université d'Alberta²²³).

²²⁰ <https://www.computecanada.ca/page-daccueil-du-portail-de-recherche/acces-aux-ressources/concours-dallocation-des-ressources/rpp/?lang=fr>

²²¹ <https://nrc.canada.ca/en>

²²² <https://www.indocresearch.org/>

²²³ <http://computecanada.github.io/2016-04-28-ualberta/>

Les chercheurs eux-mêmes participent souvent activement à la création et l'offre de LR à l'intérieur et à l'extérieur de leurs programmes de recherche, notamment dans les laboratoires de recherche ou par le biais de communautés de pratique et de sociétés professionnelles propres à un domaine. Dans de nombreux établissements, des chercheurs spécialisés dans une technologie particulière ou des directeurs d'instituts de recherche peuvent être fortement impliqués dans les efforts locaux de LR.

Cependant, chaque établissement de recherche dispose d'un certain niveau de soutien local en matière de LR, bien que les spécificités soient très variées. Certains établissements disposent d'un ensemble bien développé et étendu de services de LR et, dans certains cas, ces ressources contribuent à la prestation de services à l'échelle régionale et nationale. Ce modèle de réseau distribué de soutien ou de services peut être observé dans l'écosystème canadien d'IRN et se reflète dans la structure de certains des acteurs nationaux, comme la FCC pour le CIP et Portage dans la GDR. Bien qu'il n'y ait pas de réseau distribué semblable dans le paysage des LR, un réseau distribué est en train d'émerger grâce aux efforts de CANARIE, de la FCC et de certaines communautés de domaine.

De nombreux établissements de recherche disposent d'unités informatiques sur différents plans administratifs qui soutiennent les LR de diverses manières. Par exemple, les équipes informatiques des bibliothèques fournissent parfois un soutien et une formation de base, et peuvent également fournir une programmation à petite échelle. La bibliothèque peut également héberger des services de développement de logiciels généraux pour ce qu'elle considère comme des besoins communs à l'ensemble de l'établissement, et comme un élément fondamental de son soutien à la science ouverte. Le service informatique central (ou les services technologiques) peut également fournir des services liés aux LR tels que l'achat de logiciels libres ou à faible coût, l'installation et la formation, etc. Certains départements, en particulier dans les domaines des STIM, disposent de leur propre support informatique local avec un personnel dédié qui travaille de manière intensive avec les logiciels de recherche. Ces équipes peuvent répondre davantage aux besoins particuliers de ce groupe de recherche.

Certaines universités disposent également de centres dédiés à l'informatique de recherche qui facilitent les découvertes multidisciplinaires par le biais du calcul avancé, du génie de logiciels, de l'analyse des données et d'autres méthodes de recherche numérique. Leurs services peuvent inclure le calcul avancé, la science des données, le graphisme et la visualisation, le génie de logiciels et de systèmes, le développement de cyberinfrastructures et la formation.

Les hôpitaux de recherche disposent souvent de leur propre service informatique et d'équipes de LR locales et développent leur propre logiciel de recherche qui combine la recherche et la pratique clinique. Les hôpitaux de recherche sont confrontés à des problèmes uniques et difficiles, car ils doivent traiter des données impliquant des informations sur les patients qui sont soumises à des exigences strictes en matière de confidentialité et de sécurité. Les hôpitaux de recherche ont souvent des équipes de recherche importantes avec des projets de grande envergure ; en raison des contraintes de sécurité et de confidentialité de leurs données, ils sont souvent incapables d'utiliser l'IRN fournie par des organisations comme la Fédération Calcul Canada.

Les collèges et les instituts réalisent généralement le développement des guichets uniques avec des partenaires industriels, parfois avec des partenaires industriels et universitaires. Parmi les exemples de collaboration entre collèges et universités, citons le programme unique de maîtrise en restauration écologique entre le British Columbia Institute of Technology (BCIT) et l'Université

Simon Fraser (SFU), le partenariat entre le Collège George Brown et l'Université de Toronto qui permet aux étudiants d'acquérir des compétences pratiques en fabrication, la collaboration entre le Camosun College Technology Access Centre (CTAC) et le Centre of Aerospace Research (CfAR) de l'Université de Victoria pour le développement de la technologie des drones. Les collaborations entre l'industrie et les établissements d'enseignement supérieur dans les collèges comprennent les programmes d'entrepreneuriat du BCIT, du Nova Scotia Community College, du Northern Alberta Institute of Technology, des cégeps et collèges du Québec et du Collège Centennial ; et les expériences d'enseignement coopératif offertes au Collège Georgian, etc.

Bien qu'elle soit relativement nouvelle au Canada, l'idée d'utiliser des fonds externes pour faciliter le développement d'équipes de soutien locales en LR fait son chemin. Le programme de soutien local de CANARIE en est un exemple²²⁴ : il offre un financement pluriannuel aux établissements pour créer des équipes locales de soutien en LR. De plus, d'autres services semblables voient le jour. Par exemple, le Humanities Computing and Media Centre (HCMC)²²⁵ de l'Université de Victoria existe depuis de nombreuses années en tant que bureau permanent de recherche et développement doté de développeurs de logiciels à temps plein et d'experts en recherche et en conception pédagogique. L'Institut Perimeter n'offre pas seulement une expérience informatique de calibre mondial à ses chercheurs, mais possède également une pratique et une philosophie de génie logiciel de recherche solide, ce qui est inhabituel dans la communauté canadienne des ILR. Au fur et à mesure que l'écosystème canadien et international de LR évolue, les possibilités de développer une expertise locale en LR évolueront en parallèle.

Bien qu'il s'agisse d'une évolution importante des services de LR dans le paysage canadien, les institutions de recherche n'auront pas toutes une demande suffisante pour justifier des équipes à temps plein. Les organisations nationales comme CANARIE qui offrent ce soutien sont importantes et aident à coordonner d'autres modèles de développement de LR, comme le partage d'équipes de LR entre établissements.

Dans un contexte semblable, la plupart des laboratoires de recherche nationaux disposent de leur propre soutien local à l'informatique et aux LR. Bien que les laboratoires gouvernementaux ne soient généralement pas éligibles au financement fédéral, ils sont en mesure de participer à des efforts de recherche collaborative. Malgré ce financement et ce soutien internes, les laboratoires nationaux se tournent vers la FCC et d'autres organismes nationaux pour trouver des occasions de partager des ressources, généralement en travaillant avec des collaborateurs universitaires. Le Canadian Advanced Network for Astronomical Research (CANFAR) consortium²²⁶ est un exemple ; celui-ci est exploité par le Centre canadien de données astronomiques (CCDA) au Herzberg Astronomy & Astrophysics Research Centre, qui est un laboratoire du CNRC, alors que le chercheur principal est de l'Université Western. L'émergence de l'Alliance offre au Canada l'occasion de mettre en place une stratégie plus étroite et plus interopérable pour l'utilisation d'IRN dans l'ensemble de l'entreprise de recherche canadienne.

²²⁴ <https://www.canarie.ca/fr/pour-que-les-chercheurs-canadiens-disposent-de-puissants-logiciels-et-de-expertise-pertinente-canarie-octroie-36-m-a-six-institutions-afin-quelles-mettent-sur-pied-une-equipe-loc/>

²²⁵ <https://www.uvic.ca/humanities/hcmc/index.php>

²²⁶ <https://www.canfar.net/fr/>

Bon nombre d'organisations et de groupes de recherche se consacrent à la recherche auprès des autochtones²²⁷ et contribuent à des sites Web éducatifs²²⁸. Deux de ces plateformes sont le Indigenous Knowledge Social Network²²⁹ (SIKU), qui est une application mobile et une plateforme Web par et pour les Inuits, et l'Initiative d'innovation autochtone²³⁰, hébergée par Grands Défis Canada, qui est une plateforme d'innovation visant à combler les lacunes en matière de ressources, à habiliter les innovateurs et les communautés autochtones et à contribuer au renforcement des capacités d'une nouvelle génération d'innovations autochtones. Le budget fédéral de 2021 prévoit également²³¹ un investissement de plus de 70 millions de dollars pour l'élaboration et la mise en œuvre de la stratégie de gouvernance des données des Premières Nations, ainsi qu'un investissement de 8 millions de dollars pour le soutien de la capacité des données de base des Métis et des Inuits et l'élaboration de stratégies de données fondées sur les distinctions. De plus, le CRSH s'est engagé à financer la recherche menée par et avec les peuples autochtones, notamment les Premières Nations, les Métis et les Inuits²³².

5.2.4. Secteur privé

Le secteur privé fournit couramment des LR aux chercheurs. Par exemple, les établissements de recherche déploient de plus en plus d'analyses à forte intensité de données et de calcul par le biais de services infonuagiques commerciaux de stockage et d'informatique, ainsi que de services comme Watson d'IBM. Les entreprises technologiques et les entreprises en démarrage peuvent créer des LR pour répondre à leurs propres objectifs commerciaux ou utiliser des outils développés dans un environnement universitaire, apportant ainsi leurs propres idées et outils améliorés. Il sera important d'impliquer le secteur privé dans le débat sur l'offre de LR, car il pourrait être un partenaire clé dans l'offre de services durables et rentables. Simultanément, les universités et les chercheurs devront veiller à ce que des garanties appropriées soient mises en

²²⁷ Par exemple, le Réseau canadien des Centres de données de recherche : Indigenous Peoples (<https://crdcn.org/fr/indigenous-peoples>), Centre for First Nations Governance (<http://fngovernance.org/>), CIER : Centre for Indigenous Environmental Resources (<http://www.yourcier.org/>), First Nations Information Governance Centre (<http://fnigc.ca/>), Institute on Governance (<http://iog.ca/>), iPortal : Université de la Saskatchewan (<http://iportal.usask.ca/>), Legacy of Hope Foundation (<http://legacyofhope.ca/>), PCAP (<https://fnigc.ca/ocap-training/>), Ontario Federation of Indigenous Friendship Centres (<http://www.ofifc.org/>), the Scow Institute (<http://scow-archive.libraries.coop/>), Statistique Canada : Peuples autochtones (https://www150.statcan.gc.ca/n1/en/subjects/aboriginal_peoples).

²²⁸ Notamment la carte interactive des profils des Premières Nations (<https://geo.aadnc-aandc.gc.ca/cippn-fnpim/index-eng.html>), iPortal : Indigenous Studies Portal (USask) (<https://login.library.sheridanc.on.ca/login?url=http://iportal.usask.ca>), Four Directions Teachings (<http://fourdirectionsteachings.com/index.html>), Bibliothèque et Archives Canada : Patrimoine autochtone (<http://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/Pages/introduction.aspx>), Algonquian Dictionaries Project (www.atlas-ling.ca), Algonquian Linguistic Atlas (dictionaries.atlas-ling.ca), etc.

²²⁹ <https://siku.org/>

²³⁰ <https://indigenousinnovate.org/>

²³¹ <https://indigenousinnovate.org/>

²³² https://www.sshrc-crsh.gc.ca/society-societe/community-communite/indigenous_research-recherche_autochtone/index-eng.aspx

place pour protéger leur propriété et leur rôle de gestionnaire principal des LR qu'ils génèrent (les ALC notamment).

5.2.5. Revues

Certaines revues (p. ex. Journal of Open Research Software²³³ [JORS], Journal of open-source Software²³⁴ [JOSS], Software Impacts²³⁵, SoftwareX²³⁶) sont explicitement consacrées aux LR. Les éditeurs de revues jouent un rôle important en orientant et en influençant les approches des chercheurs en matière de LR, à la fois par leurs politiques d'édition officielles²³⁷ et par leurs communautés de pratique. Elsevier, qui publie un grand nombre de recherches scientifiques et médicales, a deux revues consacrées à la publication de LR : SoftwareX pour les LR à code source ouvert et Software Impacts pour les LR évalués par les pairs. Parmi les quatre revues, les trois autres^{238,239,240}, à l'exception de SoftwareX, ne sont pas transparentes quant aux critères qu'elles utilisent pour classer les LR, ni quant à la portée des LR qu'elles entendent inclure.

5.3. Bailleurs de fonds pour la création et l'utilisation des LR

Le financement et les bourses de recherche canadiens²⁴¹ se répartissent en deux grandes catégories : 1) le financement de la recherche scientifique ; et 2) le financement de l'innovation et de la recherche commerciale. La liste suivante présente un ensemble représentatif d'organisations qui soutiennent les chercheurs canadiens en les finançant explicitement ou implicitement pour créer et utiliser les LR. Ces organisations varient grandement en ce qui concerne leur portée géographique (nationale, régionale et locale) et leur contenu (universel et disciplinaire).

5.3.1. À l'échelle nationale

5.3.1.1. CANARIE

Depuis sa création en 2007, le programme Logiciels de recherche (LR) de CANARIE s'est fait le champion de la création d'outils LR qui accélèrent la découverte et simplifient l'accès aux IRN en octroyant un total de 50,7 millions de dollars en financement, avec un ratio typique de demandes de financement de 3:1. Parmi les 112 projets financés qui couvrent de nombreux domaines de la recherche universitaire (180 disciplines), 49 projets ont fait appel à des collaborateurs

²³³ <https://openresearchsoftware.metajnl.com/>

²³⁴ <http://joss.theoj.org>

²³⁵ <https://www.sciencedirect.com/journal/software-impacts>

²³⁶ <https://www.sciencedirect.com/journal/softwarex>

²³⁷ <https://www.chorusaccess.org/resources/software-citation-policies-index/>

²³⁸ <https://www.sciencedirect.com/journal/software-impacts/about/aims-and-scope>

²³⁹ <https://joss.readthedocs.io/en/latest/submitting.html#what-we-mean-by-research-software>

²⁴⁰ <https://openresearchsoftware.metajnl.com/about/#q1>

²⁴¹ <https://www.canada.ca/en/services/science/researchfunding.html>

internationaux ; 145 outils, services et plateformes de LR sont à la disposition des chercheurs grâce au financement de LR, avec 617 démonstrations, présentations et documents supplémentaires ; et 1191 équipes de recherche ont bénéficié des logiciels financés (soit 104 équipes de recherche ayant accès à des logiciels avec de millions de dollars de financement). En particulier, les outils de LR qui ont émergé de cette filière de financement donnent au Canada un avantage de premier plan dans l'exploitation d'IRN pour accélérer la découverte et susciter l'innovation. Le programme LR de CANARIE est essentiel non seulement pour faire progresser et renforcer les capacités de l'écosystème canadien des logiciels de recherche, mais aussi pour financer des domaines non traditionnels où les sources de financement étaient moins nombreuses, ainsi que pour bâtir la culture et la communauté des LR au Canada.

Le programme de financement de la GDR de CANARIE, bien qu'il soit axé sur la GDR, présente un chevauchement puisque certains projets ont reçu des fonds à la fois dans le cadre des programmes de LR et de GDR. Pour ce qui est du financement total, le programme de financement de GDR de CANARIE a jusqu'à présent alloué²⁴² 4,5 millions de dollars aux plateformes de GDR et 1,8 million de dollars supplémentaires à Portage. En date du 30 juillet 2021, les deux appels de demandes de financement pour la GDR de CANARIE ont financé un total de 13 équipes ; les projets financés couvrent 26 disciplines, ont rendu disponibles 12 outils, services et plateformes qui profitent à 628 équipes de recherche, et ont donné lieu à 62 démonstrations, présentations et articles. Parmi eux, trois projets ont impliqué des collaborateurs internationaux.

5.3.1.2. Fondation canadienne pour l'innovation (FCI)

Créée par le gouvernement du Canada en 1997, la Fondation canadienne pour l'innovation (FCI) est un organisme indépendant à but non lucratif qui finance des outils et des infrastructures de recherche afin d'aider les chercheurs canadiens à mener des travaux de recherche et de développement technologique de pointe, ainsi qu'à créer et à maintenir un paysage de recherche complet au Canada. Les fonds de la FCI sont accordés en dernier ressort aux établissements canadiens (notamment les universités, les collèges, les hôpitaux de recherche et les organismes de recherche à but non lucratif), et non aux chercheurs individuels. La FCI finance jusqu'à 40 % des coûts d'infrastructure de recherche d'un projet, le reste étant ensuite assuré par des investissements de partenaires des secteurs publics, privés et à but non lucratif. En date du 3 mars 2021, plus de 8,9 milliards de dollars pour 11 618 projets ont été investis dans l'infrastructure des établissements de recherche canadiens depuis sa création en 1997²⁴³.

La FCI offre aux établissements six volets de financement de base qui permettent la création de LR dans le cadre du financement.

- Fonds des leaders John R. Evans²⁴⁴ (FLJE) : ce fonds aide les établissements à obtenir les ressources d'infrastructure (y compris l'infrastructure clé et une partie des coûts de

²⁴² Le terme "alloué" est utilisé ici, car l'appel de demandes de financement pour la GDR 2 est toujours en cours.

²⁴³ <https://www.innovation.ca/fr/projets-resultats/projets-finances>

²⁴⁴ <https://www.innovation.ca/fr/appel-gestion/financement/fonds-leaders-john-r-evans>

fonctionnement et d'entretien) nécessaires pour attirer et retenir les meilleurs chercheurs. Il existe cinq volets de financement différents : le volet non affilié et quatre volets de partenariat (les programmes de chaires de recherche du Canada²⁴⁵ et de chaires d'excellence en recherche du Canada²⁴⁶ du Secrétariat des programmes interorganismes à l'intention des établissements (SPIIE), le programme Subventions Alliance du CRSNG²⁴⁷ et les programmes Subventions Savoir²⁴⁸ et les subventions de partenariat²⁴⁹ du CRSH).

- Fonds des initiatives scientifiques majeures²⁵⁰ (ISM) : Le Fonds des initiatives scientifiques majeures (FISM) soutient les installations de recherche nationales uniques et à grande échelle en leur fournissant les services d'exploitation et d'entretien (E et E) nécessaires pour produire d'excellents développements scientifiques et technologiques de calibre mondial. L'objectif est de faire en sorte que ces installations, qui servent des communautés de chercheurs de tout le pays et de l'étranger, disposent de solides politiques et pratiques de gouvernance et de gestion. Le programme ISM qui finance et soutient actuellement la FCC sera repris par l'Alliance le 1^{er} avril 2022.
- Fonds collège-industrie pour l'innovation²⁵¹ (FCII) : ce fonds aide les collèges et les écoles polytechniques du Canada à renforcer la capacité des infrastructures de recherche appliquée émergentes et l'innovation commerciale en favorisant les partenariats stratégiques avec le secteur privé. Il y a deux volets de financement différents : le volet 1 appuie les demandes d'infrastructure dans le but d'améliorer la capacité existante de recherche appliquée et de développement technologique dans les collèges ; le volet 2 est le programme d'amélioration de l'innovation qui est géré par le CRSNG et appuie l'infrastructure de recherche associée à une demande de subvention de cinq ans du Programme d'innovation dans les collèges et la communauté.
- Fonds d'exploitation des infrastructures²⁵² (FEI) : ce fonds offre un supplément de 30 % aux subventions d'investissement de la FCI qui aide les établissements à couvrir une partie des coûts d'exploitation et d'entretien des infrastructures de recherche financées par la FCI afin d'assurer leur utilisation optimale.

²⁴⁵ <https://www.chairs-chaire.gc.ca/home-accueil-fra.aspx>

²⁴⁶ <https://www.cerc.gc.ca/home-accueil-fra.aspx>

²⁴⁷ https://www.nserc-crsng.gc.ca/innovate-innover/alliance-alliance/index_fra.asp

²⁴⁸ https://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/insight_grants-subventions_savoir-fra.aspx

²⁴⁹ https://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/partnership_grants_stage1-subventions_partenariat_etape1-fra.aspx

²⁵⁰ <https://www.innovation.ca/apply-manage-awards/funding-opportunities/major-science-initiatives-fund>

²⁵¹ <https://www.innovation.ca/fr/appel-gestion/financement/fonds-college-industrie-innovation>

²⁵² <https://www.innovation.ca/fr/appel-gestion/fonds-dexploitation-infrastructures>

- Le Fonds d'innovation²⁵³ : il permet d'investir dans des infrastructures de recherche de pointe dans toutes les disciplines et tous les domaines de recherche.
- Fonds des occasions exceptionnelles²⁵⁴ : ce fonds donne aux établissements la possibilité de demander un financement pour des projets d'infrastructure exceptionnels et urgents qui ne seraient pas pris en compte dans le processus normal d'examen des concours de la FCI.

De plus, les initiatives de cyberinfrastructure de la FCI de 2015 et 2017²⁵⁵ visaient à concevoir et à développer des *infrastructures de données* de recherche par le biais du développement de logiciels et à faire participer des consortiums d'établissements et leurs chercheurs, des scientifiques des données et des développeurs de logiciels, ainsi que Calcul Canada pour mener des recherches de pointe à forte intensité de calcul et de données. Il s'agissait du premier programme de la FCI consacré au soutien des plateformes de LR, qui sont également financées dans le cadre d'autres programmes de la Fondation ; celui-ci reflétait la prise de conscience croissante des défis que pose le maintien des infrastructures.

5.3.1.3. Autres organismes nationaux de financement

La plupart des organismes nationaux de financement sont axés sur le financement de la recherche, mais pas explicitement sur le développement de logiciels ou d'infrastructures de recherche. C'est pourquoi, sur le plan national/fédéral, il y a peu de sources de financement visant à développer des logiciels de recherche comme résultat principal d'un projet de recherche. Les initiatives stratégiques fédérales comme CANARIE et la FCI fournissent des fonds pour le développement d'outils de LR, de passerelles et d'équipes locales et nationales de soutien aux LR. La plupart des organismes et initiatives de financement de la recherche à l'échelle nationale ne financent généralement pas de projets visant à développer explicitement les LR, mais le font plutôt de manière indirecte en finançant des projets qui étudient un problème de recherche propre à un domaine et qui, grâce à cette recherche, développent les LR comme résultat du projet. Par exemple, les trois organismes (les IRSC²⁵⁶, le CRSNG²⁵⁷ et le CRSH²⁵⁸, le Fonds d'urgence pour la continuité de la recherche au Canada²⁵⁹ [FUCR], le Fonds d'excellence en recherche Apogée Canada²⁶⁰ [FERAC], le Fonds Nouvelles frontières en recherche²⁶¹ (FNFR), l'Agence spatiale

²⁵³ <https://www.innovation.ca/fr/appel-gestion/financement/fonds-dinnovation>

²⁵⁴ <https://www.innovation.ca/fr/appel-gestion/financement/fonds-occasions-exceptionnelles>

²⁵⁵ <https://www.innovation.ca/fr/appel-gestion/concours-anterieurs/initiative-cyberinfrastructure-2015-2017>

²⁵⁶ <https://cihr-irsc.gc.ca/f/193.html>

²⁵⁷ https://www.nserc-crsng.gc.ca/index_fra.asp

²⁵⁸ <https://www.sshrc-crsh.gc.ca/home-accueil-fra.aspx>

²⁵⁹ <https://www.sshrc-crsh.gc.ca/funding-financement/crcef-fucrc/index-fra.aspx>

²⁶⁰ <https://www.cfref-apogee.gc.ca/home-accueil-fra.aspx>

²⁶¹ <https://www.sshrc-crsh.gc.ca/funding-financement/nfrf-fnfr/index-fra.aspx>

canadienne²⁶² (ASC) (l'initiative de financement Vols et investigations-terrain en technologies et sciences spatiales ²⁶³[VITES] de l'ASC, le Programme de développement des technologies spatiales²⁶⁴ (PDS) de l'ASC, l'initiative UtiliTerre de l'ASC²⁶⁵, le programme Solutions innovatrices Canada²⁶⁶(ISC) de l'ASC, la coopération internationale de l'ASC avec l'Agence spatiale européenne ²⁶⁷[ESA]), le Programme des chaires de recherche du Canada²⁶⁸ (PCRC) et le Programme des chaires d'excellence en recherche du Canada²⁶⁹ (CERC), l'initiative des Réseaux de centres d'excellence²⁷⁰ (RCE), le financement de MITACS, l'Initiative des Supergrappes d'innovation²⁷¹ et le Programme canadien de l'innovation à l'international²⁷² (PCII). Génome Canada/Québec^{273,274}, bien qu'il ne se concentre pas sur les appels de demandes de financement de LR comme CANARIE, a soutenu le développement de nombreux outils^{275,276,277} et infrastructures²⁷⁸ bio-informatiques pour soutenir la génomique par le biais de projets de recherche financés. Le programme du CRSH sur les images, les textes, le son et la technologie²⁷⁹ constituait une voie de financement spécialisée pour le développement d'outils expérimentaux où certains LR étaient développés, mais le programme a été abandonné et le développement

²⁶² <https://www.asc-csa.gc.ca/fra/default.asp>

²⁶³ <https://www.asc-csa.gc.ca/fra/programmes-financement/programmes/vites/default.asp>

²⁶⁴ <https://www.asc-csa.gc.ca/fra/programmes-financement/programmes/pds/default.asp>

²⁶⁵ <https://www.asc-csa.gc.ca/fra/programmes-financement/programmes/utiliterre/default.asp>

²⁶⁶ <https://www.asc-csa.gc.ca/fra/programmes-financement/programmes/sic/default.asp>

²⁶⁷ <https://www.asc-csa.gc.ca/fra/programmes-financement/opportunités-financement/default.asp>

²⁶⁸ <https://www.chairs-chaire.gc.ca/home-accueil-fra.aspx>

²⁶⁹ <https://www.cerc.gc.ca/home-accueil-fra.aspx>

²⁷⁰ https://www.nce-rce.gc.ca/IndeX_fra.asp

²⁷¹ <https://www.ic.gc.ca/eic/site/093.nsf/fra/accueil>

²⁷² https://www.deleguescommerciaux.gc.ca/funding-financement/ciip-pcii/index.aspx?lang=fra&_ga=2.99399966.185868441.1637103103-513118306.1637103103

²⁷³ <https://www.genomecanada.ca/fr>

²⁷⁴ <https://www.genomequebec.com/>

²⁷⁵ <https://www.genomecanada.ca/fr/logiciel-visant-identifier-et-quantifier-les-peptides-dans-les-megadonnees-de-la-spectrometrie-de>

²⁷⁶ <https://www.genomecanada.ca/fr/prediction-rapide-de-la-resistance-antimicrobienne-partir-dechantillons-metagenomiques-donnees>

²⁷⁷ <https://www.genomecanada.ca/fr/ontarget-logiciel-fonde-sur-les-megadonnees-pour-lillustration-des-regions-cis-regulatrices>

²⁷⁸ <https://www.genomecanada.ca/fr/de-nouveaux-outils-logiciels-pour-les-donnees-volumineuses-lappui-de-la-recherche-de-pointe-sur-le>

²⁷⁹ https://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/itst/research_grants-subsventions_recherche-fra.aspx

d'« outils » est maintenant intégré au programme Connexion²⁸⁰ que la plupart des chercheurs associent au soutien de conférences ou d'ateliers. Par conséquent, il y a moins de travaux sur les LR dans les sciences sociales et humaines qu'auparavant. Dans le contexte de la Vérité et la Réconciliation, le soutien à la recherche autochtone et à la souveraineté des données dans le budget fédéral de 2021 a des répercussions sur les LR.

5.3.2. À l'échelle régionale/provinciale

Les organismes provinciaux/régionaux traditionnels de financement de la recherche ne financent pas explicitement les LR : par exemple, la Michael Smith Foundation for Health Research²⁸¹(MSFHR), en tant qu'organisme provincial de financement de la recherche en santé, vise à favoriser le développement des compétences dans la recherche en santé en Colombie-Britannique et à optimiser la recherche en santé et l'application des connaissances dans cette province ; le Fonds pour la recherche en Ontario — Excellence en recherche (FRO-ER)²⁸², les bourses de jeunes chercheurs²⁸³ et le récent Fonds ontarien de recherche pour l'intervention rapide contre la COVID-19²⁸⁴ offrent pour l'enseignement supérieur un financement de recherche et de développement afin de soutenir la recherche et l'innovation qui aide à couvrir le coût des activités de recherche (salaires, fournitures, etc.) seulement. Cependant, les organismes de financement provinciaux/régionaux financent parfois les LR, comme le Fonds pour la recherche en Ontario — Infrastructure de recherche²⁸⁵. Mais il existe plusieurs organisations provinciales et régionales qui ont un impact sur le développement et la prestation des services de LR. Les organisations régionales peuvent soutenir le développement des LR soit directement (le FIA de l'APECA^{286,287} et les programmes régionaux connexes en sont un exemple) soit indirectement par le biais d'un versement égal de financement pour d'autres programmes. On observe également une tendance récente à reconcevoir les organismes provinciaux de financement, essentiellement axés sur la santé, pour qu'ils soutiennent la recherche interdisciplinaire en plus de celle uniquement axée sur la santé. L'Alliance canadienne des organismes provinciaux de financement de la recherche en santé (NAPHRO)²⁸⁸ est un exemple d'organisation qui orchestre une approche collaborative au pays.

²⁸⁰ https://www.sshrc-crsh.gc.ca/funding-financement/umbrella_programs-programme_cadre/connection-connexion-fra.aspx

²⁸¹ <https://www.msfhr.org/>

²⁸² <https://www.ontario.ca/fr/page/fonds-pour-la-recherche-en-ontario-excellence-en-recherche>

²⁸³ <https://www.ontario.ca/fr/page/bourses-de-nouveaux-chercheurs>

²⁸⁴ <https://www.ontario.ca/fr/page/fonds-ontarien-de-recherche-pour-lintervention-rapide-contre-la-covid-19>

²⁸⁵ <https://www.ontario.ca/fr/page/fonds-pour-la-recherche-en-ontario-infrastructure-de-recherche>

²⁸⁶ <https://www.canada.ca/fr/promotion-economique-canada-atlantique/services/fondsdinnovationdelatlantique.html>

²⁸⁷ <https://www.canada.ca/fr/promotion-economique-canada-atlantique.html>

²⁸⁸ <https://www.naphro.ca/>

Les gouvernements provinciaux sont également des acteurs clés pour aider l'innovation en matière de guichet unique dans les établissements d'enseignement supérieur. Ce soutien varie d'une province à l'autre et d'un territoire à l'autre, mais peut inclure un financement direct, un financement de contrepartie et une gamme de programmes de soutien. Citons, par exemple, la Bourse aux jeunes chercheurs de l'Ontario, Innovation PEI²⁸⁹ et Alberta Innovates²⁹⁰. Dans bon nombre de ces programmes provinciaux, le développement économique régional constitue une stratégie d'investissement clé, souvent motivée par les intérêts et les capacités des établissements. Dans le secteur collégial, le secteur privé peut être un important facilitateur de l'innovation en matière de guichet unique ; il compte beaucoup sur les partenaires industriels pour le financement et le développement.

5.3.3. Sociétés et fondations (publiques, privées et internationales)

Les partenariats de recherche avec les entreprises sont également un point central de l'innovation et de l'esprit d'entreprise, offrant de vastes possibilités au-delà du campus et faisant le lien entre le monde universitaire et l'industrie, développant de nouvelles technologies²⁹¹ et découvrant des talents. De nombreuses universités ont créé des bureaux de transfert de technologie pour faciliter les partenariats avec le secteur privé et les activités dérivées de la recherche, mais la plupart d'entre eux se concentrent sur les licences et les approches de la propriété intellectuelle. Les Open-Source Program Offices — OSPO (bureaux de programmes à code source ouvert)²⁹² ont fait leur apparition sur le plan international, mais ils n'ont pas encore été adoptés au Canada. En offrant une aide à la commercialisation, on fait progresser les idées issues de la recherche, on tire parti des investissements commerciaux dans la recherche et on donne aux étudiants la possibilité d'explorer des carrières et d'acquérir une expérience professionnelle tout en obtenant un diplôme universitaire. Plus récemment, le secteur privé a commencé à créer des bureaux de programmes à code source ouvert²⁹³, créés pour tirer parti de la valeur croissante des cadres à code source ouvert. Bien qu'il n'y ait actuellement aucun OSPO dans les établissements d'enseignement supérieur canadiens, on en voit apparaître ailleurs²⁹⁴. Certains de ces partenariats de recherche avec des entreprises peuvent également découler de programmes financés par le gouvernement ou financés conjointement, notamment par le CRSNG, Mitacs²⁹⁵, les Centres d'excellence de l'Ontario (CEO), le CRSH, etc.

La plupart des fondations, des organisations caritatives et des entreprises (publiques, privées ou internationales), financent généralement la recherche dans une discipline précise, mais pas en

²⁸⁹ <https://www.princeedwardisland.ca/fr/information/innovation-i-p-e/information-and-communications-technology>

²⁹⁰ <https://albertainnovates.ca/>

²⁹¹ <https://uwaterloo.ca/news/news/rogers-and-university-waterloo-partner-build-made-canada-5g>

²⁹² <https://www.linuxfoundation.org/tools/creating-an-open-source-program/>

²⁹³ <https://demand.fossa.com/what-is-an-ospo?source=Website>

²⁹⁴ <https://blogs.library.jhu.edu/2020/12/hopkins-ospo-featured-on-impactful-open-source-podcast/>

²⁹⁵ <https://www.mitacs.ca/fr>

LR en particulier. Par exemple, la fondation McLean²⁹⁶ accorde des subventions dans de nombreux domaines, dont le mieux-être social, l'éducation, la conservation de l'environnement, la santé et les arts ; la fondation Graham Boeckh²⁹⁷ (GBF) vise à transformer les soins de santé mentale ; la fondation Jarislowsky²⁹⁸ se concentre sur la promotion, le soutien et l'avancement de l'excellence en éducation, en médecine et dans les arts ; Grands Défis Canada²⁹⁹ (GCC) soutient l'intégration des sciences et de la technologie, l'innovation sociale et commerciale en santé mondiale³⁰⁰, les grands défis humanitaires³⁰¹, et l'Initiative d'innovation autochtone³⁰². Certains programmes financent la recherche en collaboration avec d'autres organismes : le programme de bourses de découverte de la fondation de recherche Banting³⁰³ investit dans la recherche innovante en santé et en biomédecine et dans la science des données, avec un cofinancement supplémentaire en partenariat avec d'autres organisations (notamment l'Institut canadien des sciences statistiques³⁰⁴ [INCASS], la Fondation de recherche médicale sur la dystonie³⁰⁵, Mitacs, le Centre McLaughlin de l'Université de Toronto³⁰⁶). Il ne fait aucun doute que les LR jouent un rôle indispensable pour faciliter l'ensemble des projets de recherche financés.

Certains bailleurs de fonds et fondations internationales disposent d'importants programmes de financement direct des LR : l'initiative Chan Zuckerberg a un programme de financement de la science ouverte qui vise à promouvoir le partage des résultats de la recherche, y compris les logiciels de recherche à code source ouvert³⁰⁷. Wellcome Trust³⁰⁸ est un autre exemple, qui non seulement investit dans la recherche sur la santé et le réchauffement climatique, mais qui a récemment financé directement les LR par le biais de ses subventions de développement technologique³⁰⁹. Wellcome Trust joue également un rôle important dans le soutien au libre

²⁹⁶ <http://www.mcleanfoundation.ca/>

²⁹⁷ <https://grahamboeckhfoundation.org/fr/>

²⁹⁸ <https://www.charitydata.ca/charity/the-jarislowsky-foundationla-fondation-jarislowsky/894744036RR0001/>

²⁹⁹ <https://www.grandchallenges.ca/>

³⁰⁰ <https://www.grandchallenges.ca/fr-ca/programs/sante-mondiale/>

³⁰¹ <https://humanitariangrandchallenge.org/fr/>

³⁰² <https://www.grandchallenges.ca/fr-ca/programs/initiative-dinnovation-autochtone/>

³⁰³ <https://www.bantingresearchfoundation.ca/discovery-award-program/>

³⁰⁴ <https://canssionario.utoronto.ca/>

³⁰⁵ <https://dystoniacanada.org/fr/>

³⁰⁶ <http://www.mclaughlin.utoronto.ca/>

³⁰⁷ <https://chanzuckerberg.com/rfa/essential-open-source-software-for-science/>

³⁰⁸ <https://wellcome.org/>

³⁰⁹ <https://wellcome.org/grant-funding/schemes/technology-development-grants>

accès^{310,311} et aux données ouvertes. En 2016, Wellcome Trust a lancé la plateforme de publication en libre accès Wellcome Open Research³¹² et Open Research Fund³¹³ qui encourage à rendre les résultats des recherches (co)financées par Wellcome conformes aux principes FAIR. Plus précisément, l'appel de demandes de financement pour la recherche ouverte de 2021 vise à encourager le partage et l'ouverture dans la recherche en santé en couvrant les coûts de développement (construction d'un nouvel outil logiciel ou d'une nouvelle application) et les coûts de licence et de calcul (par exemple, l'achat de stockage infonuagique). En ce qui concerne l'évaluation de la recherche, Wellcome Trust fait partie des premiers utilisateurs qui ont recours au DORA pour améliorer les méthodes d'évaluation des résultats de la recherche universitaire³¹⁴.

Pour une liste plus exhaustive des sources de financement pour le développement et l'utilisation des LR, voir l'[annexe H](#).

6 LR au-delà du Canada ([Annexe I](#))

L'[annexe I](#) présente une analyse environnementale globale du paysage de la communauté de LR : initiatives, communautés, entreprises, revues et éditeurs, entités à but non lucratif, plateformes, projets, registres, index, agrégateurs, groupes, infrastructures et services. Ce travail s'appuie sur : 1) les résultats du groupe de travail de la Research Software Alliance (ReSA) qui a identifié et analysé le paysage des LR^{315,316}; et 2) les Scholarly Infrastructures for Research Software³¹⁷ (SIRS) du groupe de travail sur l'architecture de l'EOSC qui a fourni un grand survol de 9 infrastructures européennes représentatives des LR. En plus de leurs activités communautaires, ces organisations pourraient servir à diffuser les dernières nouvelles et mises à jour, les appels de la communauté³¹⁸, les opportunités de financement³¹⁹, etc.

Pour que les chercheurs canadiens soient compétitifs dans l'espace de recherche et qu'ils puissent tirer parti des LR créées par leurs homologues ailleurs, les pratiques de LR canadiennes doivent être coordonnées avec celles du monde entier, idéalement avec un élément de concertation et de collaboration des bailleurs de fonds. Parmi les initiatives énumérées à l'[annexe](#)

³¹⁰ <https://wellcome.org/grant-funding/guidance/open-access-guidance/open-access-policy>

³¹¹ <https://wellcome.org/grant-funding/guidance/data-software-materials-management-and-sharing-policy>

³¹² <https://wellcomeopenresearch.org/>

³¹³ <https://wellcome.org/grant-funding/schemes/open-research-fund>

³¹⁴ <https://wellcome.org/grant-funding/guidance/data-software-materials-management-and-sharing-policy>

³¹⁵ <https://wellcome.org/grant-funding/guidance/data-software-materials-management-and-sharing-policy>

³¹⁶ <https://doi.org/10.5281/zenodo.3699950>

³¹⁷ <https://www.softwareheritage.org/2020/10/23/towards-an-open-architecture-of-scholarly-infrastructures-for-research-software/?lang=fr>

³¹⁸ <https://us-rse.org/2021-02-19-community-call-education/>

³¹⁹ <https://www.fordfoundation.org/campaigns/critical-digital-infrastructure-research/>

L, il y a plusieurs organisations et événements internationaux auxquels les organisations, associations et communautés de recherche canadiennes en LR ont participé activement à l'échelle internationale. Il s'agit notamment de :

- UK Society of Research Software Engineering ³²⁰
- US Research Software Engineer Association ³²¹(US-RSE)
- Research Software Engineers ³²²
- International Research Software Alliance ³²³ (ReSA)
- Groupe de travail FAIR4RS de la Research Data Alliance (RDA) ³²⁴
- Du point de vue de la communauté de LR, CANARIE a largement joué le rôle de représentant des LR canadiens sur le plan international (p. ex. le rôle de CANARIE au sein de la Society of Research Software Engineering, de la ReSA et de UK SSI ; du groupe de travail FAIR4RS de la RDA).

Malgré la participation minimale du Canada aux communautés ou organisations mondiales de LR, des détails de mise en œuvre sur le remodelage de l'ensemble de l'IRN du Canada pour l'adapter à un contexte international émergeront de nombreuses façons, notamment dans le cadre du rapport d'évaluation des besoins des chercheurs, du nouveau modèle national de prestation de services et du modèle de financement, ou du plan stratégique d'IRN pour 2022-2024.

7 Tendances émergentes d'IRN — Aspect des LR

7.1. Perspective culturelle et sociétale

Un nombre croissant d'acteurs de la recherche et de la société civile plaident pour une adoption généralisée de la science ouverte, ce qui est abordé dans la section 5. La promotion de la science ouverte exige en fin de compte un changement culturel dans la manière dont la recherche est effectuée et reconnue. L'efficacité de la science ouverte dépend de la création de plateformes, de services, d'outils et d'applications librement accessibles qui permettent aux chercheurs de

³²⁰ <https://society-rse.org/>

³²¹ <https://us-rse.org/usrse-map/>

³²² <https://researchsoftware.org/2021/01/27/introducing-the-international-council-of-RSE-associations.html>

³²³ <https://www.researchsoft.org/>

³²⁴ <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>

mobiliser les connaissances. En ce sens, les éléments spécifiques de ce changement comprennent l'augmentation de la collaboration et de l'interaction entre les scientifiques et le développement d'une infrastructure technique (par exemple, les passerelles scientifiques ou la conteneurisation) qui favorise le développement de pratiques de recherche émergentes en plus des réseaux informatiques conventionnels, ainsi que le développement de pratiques de « science ouverte ».

7.1.1. Principes FAIR

Les LR sont différents des données et des communications savantes, mais l'objectif général des principes FAIR pour les données³²⁵ s'applique également aux LR. Les principes FAIR sont conçus pour assurer et améliorer l'utilisation optimale, la transparence, la reproductibilité et la réutilisation des données de recherche et autres objets de recherche. La manière dont les principes généraux FAIR devraient être appliqués aux LR est devenue un sujet de discussion au cours des dernières années, et bien qu'il soit généralement admis que les quatre principes fondamentaux³²⁶ s'appliquent aux LR, les caractéristiques particulières des logiciels (par exemple, le caractère exécutable, la nature composite, l'évolution continue, la gestion des versions) nécessitent la révision et l'extension des principes directeurs originaux. Les moteurs, les intervenants et les incitations, malgré leur recoupement, ne sont pas identiques pour les données et les LR, et la variété des LR et de leurs canaux de distribution pose un défi lors de l'adaptation des principes actuels FAIR pour les données. Plusieurs groupes travaillent à l'élaboration d'un ensemble de principes directeurs FAIR pour les LR, notamment le groupe de travail sur les principes FAIR pour les logiciels de recherche³²⁷ (FAIR4RS WG), codirigé par RDA, FORCE11³²⁸ et ReSA³²⁹, et les efforts du Software Sustainability Institute (SSI)³³⁰. Le document intitulé *Towards FAIR Principles for Research Software*, publié en 2020³³¹, est un exemple de travail dans ce domaine. Le récent rapport intitulé *Implementing FAIR for Research Software: Attitudes, Advantages and Challenges*, publié en 2021³³², aborde différents sujets autour des LR, notamment les attitudes, les avantages, les motivations et les obstacles à l'adoption des principes FAIR pour les LR, les pratiques actuelles d'adoption des LR dans les organisations, l'EDI et la pérennité des logiciels.

³²⁵ <https://www.force11.org/fairprinciples>

³²⁶ M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg et al., The FAIR guiding principles for scientific data management and stewardship, *Scientific Data* 3 (2016), 160018. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).

³²⁷ <https://www.rd-alliance.org/groups/fair-4-research-software-fair4rs-wg>

³²⁸ <https://www.force11.org/>

³²⁹ <https://www.researchsoft.org/resa-taskforces-join-us/>

³³⁰ <https://software.ac.uk/>

³³¹ Lamprecht, A.L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., Dominguez Del Angel, V., van de Sandt, S., Ison, J., Martinez, P.A. et McQuilton, P., 2020. Towards FAIR principles for research software. *Data Science*, 3(1), pp. 37-59. <https://content.iospress.com/articles/data-science/ds190026>

³³² [https://ssi-cw.figshare.com/articles/conference_contribution/Implementing FAIR for research software attitudes advantages and challenges/14453031](https://ssi-cw.figshare.com/articles/conference_contribution/Implementing_FAIR_for_research_software_attitudes_advantages_and_challenges/14453031)

Le premier programme de financement pour la GDR de CANARIE avait comme exigence fondamentale la conformité aux principes FAIR.

7.1.2. Formation et généralisation de l'utilisation des outils

Aujourd'hui, la plupart des recherches ont des fondements numériques, et les exigences en matière d'IRN de domaines qui n'étaient pas auparavant actifs dans la communauté de CIP traditionnellement axée sur les STIM augmentent de façon spectaculaire. Par exemple, des défis particuliers sont posés par les nombreuses méthodes qui relèvent du terme générique de sciences humaines numériques (SHS). Les SHS associent des outils et des méthodes numériques à des ensembles de données, petits et grands, pour faire progresser la recherche dans de nombreuses disciplines. Elles impliquent souvent une recherche collaborative, multidisciplinaire ou interdisciplinaire, étroitement liée à l'enseignement et à la diffusion (par exemple, d'ensembles de données, de sources historiques remédiées, de collections de recherche thématiques et d'autres études numériques). Les chercheurs en SHS adaptent à la fois des outils et des algorithmes provenant d'autres domaines, par exemple l'IA, l'apprentissage machine (AM) et le séquençage génétique, pour des activités telles que l'exploration de textes, le traitement d'images et l'analyse de données, et développent eux-mêmes des outils. Les sciences humaines englobent de nombreuses disciplines avec de multiples méthodes d'enquête, des types de données et des besoins divers en matière de logiciels de recherche. Il n'existe donc pas de solution unique pour ces domaines mal desservis lorsqu'il s'agit de fournir un accès équitable au CIP ; certains travaux peuvent ne pas ressembler à du CIP d'un point de vue traditionnel. Étant donné que les méthodes informatiques et l'analyse des données ne sont pas encore intégrées à la plupart des programmes d'études en sciences humaines, il existe un besoin pressant pour d'autres formes de formation aux méthodes appropriées. En ce qui concerne les LR et la formation quant à leur utilisation, les initiatives communautaires existantes et émergentes (Digital Humanities Summer Institute³³³ [DHSI] et d'autres initiatives de formation canadiennes, Alan Turing Institute's Digital Humanities & Research Software Engineering summer school³³⁴, CODATA-RDA School of Research Data Science³³⁵ notamment) et les communautés de pratique peuvent servir de partenaires pour la formation avec l'Alliance afin de répondre aux besoins de la communauté de recherche.

7.1.3. Biais logiciel et algorithmique

Comme les algorithmes exploitent une variété de collections de macro-données et de micro-données pour construire des outils d'AM et d'AI, une attention accrue est accordée à l'étude du

³³³ <https://dhsi.org/>

³³⁴ <https://www.eventsforce.net/turingevents/frontend/reg/thome.csp?pageID=23222&eventID=72&traceRedir=2>

³³⁵ <https://codata.org/initiatives/data-skills/research-data-science-summer-schools/>

biais algorithmique³³⁶. Le biais peut être dû à de nombreux facteurs, notamment la conception de l'algorithme lui-même, l'utilisation involontaire de données d'entraînement incomplètes ou non représentatives, ou la manière dont les données sont codées, collectées, sélectionnées ou utilisées par un algorithme. Les biais algorithmiques se retrouvent dans toute la gamme des outils, y compris dans la plupart des plateformes basées sur l'IA (par exemple, les résultats des moteurs de recherche, la reconnaissance faciale, les plateformes de médias sociaux, les algorithmes de justice pénale), et peuvent avoir des conséquences allant de violations involontaires de la vie privée au renforcement des préjugés sociaux (fondés notamment sur la race, le sexe, la sexualité et l'origine ethnique) et soulever des problèmes éthiques³³⁷. L'étude des biais algorithmiques peut être partiellement abordée dans des cadres juridiques, tels que le Règlement général sur la protection des données³³⁸ (RGPD) de l'Union européenne, mais une réglementation plus complète, de nouvelles approches pédagogiques pour enseigner une science des données responsable³³⁹ et des stratégies de détection des biais sont nécessaires à mesure que les technologies émergentes deviennent de plus en plus avancées et opaques. La communauté des LR peuvent clairement bénéficier de l'utilisation de logiciels d'AM et d'IA, mais l'approche pour l'intégration des algorithmes et des données nécessite une compréhension et une attention particulières.

7.1.4. Science citoyenne

Alors que la science ouverte ouvre les portes du monde universitaire, la science citoyenne³⁴⁰ invite la participation de tout le monde. Dans le cadre de la science ouverte, l'interaction entre les citoyens et les scientifiques réduit le fossé entre les deux groupes, éclaire l'approche traditionnellement isolée de la recherche et peut améliorer le profil de la science dans la société. De cette façon, les experts et les profanes travaillent ensemble sur des problèmes d'intérêt commun, selon la rigueur de la méthode scientifique. Les LR facilitent la science citoyenne en permettant à des volontaires de participer à des activités de collecte de données, de faciliter l'analyse des données et, dans le cas des logiciels à code source ouvert, de co-développer les outils qui sous-tendent les processus de recherche.

7.1.5. Les ILR en tant professionnels

Le concept des ingénieurs en logiciel de recherche (ILR) en tant que professionnels ne remonte que d'une décennie. La naissance du terme ILR peut être attribuée au mouvement d'ILR au Royaume-Uni, avec une série d'ateliers organisés par la SSI qui ont soulevé que les développeurs de logiciels manquaient de reconnaissance, entres autres. Les ILR sont en pleine croissance, avec un nombre croissant d'offres d'emploi et de rôles dans ce domaine. La France, par exemple, a codifié des catégories de professions de soutien à la recherche universitaire (par exemple, des

³³⁶ https://fr.wikipedia.org/wiki/Biais_algorithmique

³³⁷ <https://www.jstor.org/stable/j.ctt1pwt9w5>

³³⁸ <https://gdpr.eu/>

³³⁹ <https://link.springer.com/article/10.1007/s40593-021-00241-7>

³⁴⁰ https://fr.wikipedia.org/wiki/Sciences_participatives

postes d'« ingénieurs de recherche »³⁴¹ associés au traitement des données dans toutes les disciplines³⁴²). Bien que le monde entier utilise le terme « ILR », le Canada n'autorise malheureusement pas légalement l'utilisation du terme « ingénieurs », mais utilise plutôt le terme « développeurs de logiciels de recherche ». L'importance des ILR adéquatement formés devient de plus en plus évidente, à la fois au sein des établissements et des organismes de financement (par exemple, le programme Software for the Future de l'EPSRC³⁴³, le CSSI de la NSF³⁴⁴ [anciennement le financement³⁴⁵ du de la NSF SI²]), grâce à des initiatives comme les programmes de CANARIE, SSI, Society of Research Software Engineering, et le Hidden REF³⁴⁶, mais aussi dans le grand public, étant donné le rôle des ILR dans la mobilisation pour l'intervention rapide quant à la pandémie COVID-19³⁴⁷.

7.2. Évolution technologique

En plus les perspectives culturelles et sociétales mentionnées à la section 8.1, un certain nombre de technologies et de tendances numériques plus récentes et en pleine évolution gagnent en dynamisme et en popularité et ont un impact sur les LR.

7.2.1. AM et IA

De nos jours, l'apprentissage machine (AM) et l'intelligence artificielle (IA) sont devenus l'un des domaines les plus intrigants de la recherche en informatique ; la technologie (par exemple, le contrôle des machines, l'apprentissage automatique, l'apprentissage profond (AP) et les réseaux neuronaux) évolue rapidement, surtout au cours des cinq dernières années, le matériel informatique et les machines ayant évolué au même rythme que les logiciels. Du point de vue de la GDR, l'AM, l'IA et d'autres méthodes statistiques modernes offrent de nouvelles possibilités d'opérationnaliser des sources de données auparavant inexploitées et en croissance rapide. Les solutions logicielles d'IA et d'AM ne sont plus seulement une tendance, mais une force dominante dans des domaines aussi divers que les soins de santé, la finance, l'éducation, les transports, ainsi que dans le développement de LR, comme les tests automatisés et la génération de code automatisée. Les occasions de développer des applications concrètes de la technologie ne manquent pas, et les possibilités de percées dans ce domaine sont immenses. En 2017, le gouvernement du Canada a mandaté le CIFAR de mettre au point la Stratégie pancanadienne en matière d'intelligence artificielle avec un budget de 125 millions de dollars — la première stratégie nationale d'IA au monde³⁴⁸. Le CIFAR travaille en étroite collaboration avec les trois

³⁴¹ <https://www.enseignementsup-recherche.gouv.fr/cid23194/ingenieur-de-recherche.html>

³⁴² <https://metiersit.dsi.cnrs.fr/index.php?page=cartofamille&codeBAP=D&codeFamille=A>

³⁴³ <https://epsrc.ukri.org/research/ourportfolio/themes/researchinfrastructure/subthemes/einfrastructure/software/>

³⁴⁴ https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505505

³⁴⁵ <https://www.nsf.gov/pubs/2013/nsf13511/nsf13511.htm>

³⁴⁶ <https://hidden-ref.org/>

³⁴⁷ <https://society-rse.org/rse-contributions-to-covid-19-research/>

³⁴⁸ <https://cifar.ca/fr/ia/>

instituts nationaux d'IA du Canada (Amii³⁴⁹ à Edmonton, Mila³⁵⁰ à Montréal et l'Institut Vector³⁵¹ à Toronto), ainsi qu'avec des universités, des hôpitaux et des organisations de tout le pays. Les groupes interdisciplinaires qui poursuivent des projets de recherche et d'impact impliquant l'AM et l'IA auraient également intérêt à aborder explicitement une série de questions concernant la transparence, la reproductibilité, l'équité, la responsabilité, l'éthique et l'efficacité (TREE³⁵², FAT³⁵³), ainsi que les principes FAIR mentionnés précédemment.

En réponse à l'appui du programme national d'activités de la Stratégie pancanadienne en matière d'IA, afin de soutenir la collaboration croisée entre les meilleurs scientifiques en IA du Canada et d'inspirer des domaines de recherche à haut risque et à haute récompense : neuf projets ont été financés par le programme de subventions Catalyseur en IA du CIFAR.

7.2.2. Infonuagique

L'infonuagique est devenue une force dominante dans les systèmes de prestation de ressources. Il s'agit d'une approche sur demande par Internet pour accéder aux ressources informatiques (applications, serveurs physiques/virtuels, stockage de données, bases de données, logiciels, outils d'analyse et de développement, infrastructure, capacités de réseautage, intelligence notamment). Les organisations mettent de plus en plus en œuvre des services d'infonuagique pour leurs chercheurs (par exemple, le nuage de Calcul Canada³⁵⁴), en tirant parti des avantages des différents modèles de coût, de la haute performance et de la productivité, de la fiabilité, de la facilité de gestion, de l'évolution élastique, de la centralisation des données, de la sécurité appropriée, etc. La plupart des services d'infonuagique se répartissent en quatre grandes catégories : l'infrastructure en tant que service (IaaS), la plateforme en tant que service (PaaS), les services sans serveur (également appelés fonctions en tant que service) et les logiciels en tant que service (SaaS), parfois appelés amoncellement infonuagique parce qu'ils peuvent s'appuyer les uns sur les autres. Bien que l'infonuagique existe depuis un certain temps, son utilisation du point de vue d'activités de recherche et des LR est encore émergente et n'est pas largement adoptée. Les thèmes liés aux LR dans le domaine de l'infonuagique comprennent les affaires (par exemple, l'adoption de l'infonuagique, la vie privée, les questions juridiques et éthiques, la confiance, la cryptographie et la sécurité de l'infonuagique pour la protection contre les attaques et les violations étrangères), la technologie (par exemple, l'analyse infonuagique, les plateformes infonuagiques, les services et l'architecture, le développement de logiciels, l'équilibrage de la charge infonuagique, la conteneurisation, l'informatique en périphérie, la

³⁴⁹ <https://www.amii.ca/>

³⁵⁰ <https://mila.quebec/>

³⁵¹ <https://vectorinstitute.ai/>

³⁵² <https://www.bmj.com/content/368/bmj.l6927>

³⁵³ <https://www.fatml.org/>

³⁵⁴ <https://www.computecanada.ca/page-daccueil-du-portail-de-recherche/services-nationale/service-dinfonuagique-de-calcul-canada/?lang=fr>

gestion des centres de données, la sécurité) et les domaines d'application infonuagique (par exemple, dans l'éducation, la science en ligne, la santé, la gestion des connaissances, les logiciels à code source ouvert, l'infonuagique mobile et l'infonuagique verte axée sur la création de centres de données et de serveurs virtuels).

L'infonuagique est particulièrement pertinente pour les sciences humaines numériques pour des raisons allant de la nature souvent itérative et collaborative des activités de recherche à l'interactivité des outils et des données, comme les jeux, et aux exigences de certains types de données, comme les données ouvertes liées, qui doivent être accessibles 24 heures par jour et 7 jours par semaine. Netlytic³⁵⁵ fournit des outils en ligne pour la collecte et l'analyse des médias sociaux. Le projet Archives Unleashed³⁵⁶, y compris le nuage Archives Unleashed, répond à la nécessité de rendre de grands volumes de données Web obsolètes utilisables par les chercheurs en sciences humaines, peu importe s'ils savent programmer. Il convient également de noter que, même s'ils ne répondent pas aux définitions conventionnelles des LR, les sites de renforcement de la communauté et de diffusion de la recherche jouent un rôle important dans l'écosystème de l'infrastructure de SHS en faisant circuler la littérature grise et en favorisant le dialogue et les échanges. Ces sites peuvent être axés sur un domaine précis, comme la Nouvelle initiative Canadienne en histoire de l'environnement³⁵⁷ (NiCHE), ou plus largement ciblés, comme le Canadian HSS Commons, actuellement en version bêta, qui ajoute une fonctionnalité de gestion de projet à sa plateforme de publication et de réseautage basée sur les principes FAIR. Les sciences humaines numériques sont donc de grands utilisateurs de l'infonuagique par opposition à l'informatique par lots ; la plateforme la plus utilisée sur le nuage de Calcul Canada est la suite Voyant Tools pour l'analyse et la visualisation de textes.

7.2.3. L'infrastructure en tant que code

La maintenance d'une infrastructure infonuagique peut être une tâche difficile, car des configurations compliquées peuvent être fragilisées à une échelle supérieure sans une architecture appropriée. Malgré cette complexité, les avantages offerts par l'infonuagique et la technologie de virtualisation sont nombreux, et l'un des plus importants est peut-être l'amélioration de la gestion de l'infrastructure de manière cohérente et reproductible. L'un des moyens d'y parvenir est l'Infrastructure comme Code³⁵⁸ (IaC), qui crée une couche d'abstraction permettant de conditionner, d'organiser et de construire des environnements d'infonuagique à la demande par le biais de scripts automatisés, sur la base d'un modèle existant plutôt que de devoir reconstruire à chaque fois. Pour ceux qui travaillent avec des constructions complexes d'infonuagique, l'IaC simplifie la gestion à long terme. En tant que technologie infonuagique, l'IaC permet également de faire évoluer et de redistribuer facilement l'infrastructure à la demande. Cet aspect est essentiel pour la reproductibilité de la recherche, car il permet de définir et de modéliser l'infrastructure de manière formelle par le biais d'un logiciel, de stocker le code de l'infrastructure dans des systèmes de gestion de version, d'écrire le code de configuration sous forme de documentation, de réutiliser des parties de l'architecture dans de futurs projets ou dans n'importe

³⁵⁵ <https://netlytic.org/>

³⁵⁶ <https://archivesunleashed.org/>

³⁵⁷ <http://niche-canada.org/>

³⁵⁸ https://fr.wikipedia.org/wiki/Infrastructure_as_code

quel nuage souhaité, et de faciliter le partage. Du point de vue de l'automatisation, l'IaC peut contribuer à garantir l'efficacité du déploiement des ressources tout en offrant une configuration et une flexibilité rapides et cohérentes, tout en réduisant le coût global de la gestion de l'infrastructure. Les ressources peuvent être distribuées via le nuage ou sur site selon les besoins, ce qui peut s'avérer utile lors de crises majeures — comme le début de la pandémie de COVID-19, lorsque certaines organisations ou entreprises se sont retrouvées soudainement dans l'obligation d'évoluer.

7.2.4. Internet des objets (IdO)

L'Internet des objets (IdO) est un paradigme émergent, utilisant des dispositifs intelligents dans un contexte de connexion à l'Internet pour fournir des solutions innovantes à divers défis et problèmes liés aux entreprises, aux gouvernements et aux industries publiques/privées partout au monde. Des vêtements aux appareils de cuisine connectés, les dispositifs d'IdO sont omniprésents. Le marché mondial de l'IdO devrait dépasser les 4 milliards de dollars américains d'ici 2025³⁵⁹.

L'IdO peut être bénéfique pour de nombreuses disciplines telles que la finance, l'environnement, l'industrie, l'éducation, les transports, mais surtout pour les sciences de la santé (également appelées soins de santé intelligents ou santé en ligne). Les soins de santé en ligne basés sur l'IdO (par exemple, la télésurveillance de la santé, les notifications d'urgence, l'imagerie numérique, les applications mobiles pour l'édition et l'affichage des dossiers médicaux, la chirurgie robotique) peuvent contribuer à créer un système de soins de santé numérique (par exemple, des infrastructures et des plateformes) pour collecter des données et connecter les ressources et les services de soins de santé disponibles. Le potentiel de l'IdO peut être observé dans d'autres domaines, tels que la gestion intelligente de la circulation, la sécurité et la surveillance, l'agriculture et les fermes intelligentes, la fabrication intelligente, les dispositifs intelligents (par exemple, les vêtements, les véhicules, les robots), les villes et les maisons intelligentes, le commerce de détail, l'énergie et le contrôle des réseaux, etc. L'IdO, combiné à d'autres innovations telles que l'analyse des données et l'IA/AM/AP, est important pour la société.

Bien que l'IdO fasse l'objet d'une multitude d'études de recherche cruciales, il reste encore de nombreux défis et problèmes à résoudre pour en exploiter tout le potentiel. L'un des plus importants est la sécurité et la confidentialité des dispositifs IoT en raison des menaces, des cyberattaques, des risques et des vulnérabilités. L'interopérabilité pose également un défi en raison de la prolifération rapide et de la nature hétérogène des différentes technologies d'IdO. On peut considérer quatre niveaux d'interopérabilité en ce qui concerne l'IdO : technique, sémantique, syntaxique et organisationnel³⁶⁰. L'IdO soulève d'autres problèmes, notamment l'évolutivité, la disponibilité et la fiabilité (qualité de service [QoS] de l'informatique, des dispositifs, des communications, des réseaux et des applications).

L'IdO pose aussi des défis de programmation. La programmation dans le paysage de l'IdO est très difficile, car elle doit s'adapter à de multiples dispositifs, chacun étant différent en termes de

³⁵⁹ <https://www.mckinsey.com/featured-insights/Internet-of-things/how-we-help-clients>

³⁶⁰ Van-der-Veer H, Wiles A. Achieving technical, interoperability-the ETSI approach, ETSI White Paper No. 3. 2008.
<http://www.etsi.org/images/files/ETSIWhitePapers/IOP%20whitepaper%20Edition%203%20final.pdf>

normes, d'API et de formats de métadonnées. De plus, la programmation des appareils portables est très différente de celle des serveurs et des ordinateurs de bureau, car elle nécessite des langages de niveau inférieur (C notamment) et des compétences logicielles particulières (programmation en temps réel).

7.2.5. Informatique en périphérie

Dans un avenir riche en données, avec des milliards d'appareils connectés à l'Internet, un traitement plus rapide et plus fiable des données deviendra crucial. L'informatique en périphérie (Edge Computing) est également un domaine nouveau et émergent : il s'agit d'un système informatique distribué doté d'une puissance de traitement décentralisée, qui optimise l'utilisation des ressources dans l'infonuagique (Cloud Computing), les données étant traitées plus près de la source — à la périphérie du réseau — plutôt que sur une plateforme gérée de manière centralisée ou dans des entrepôts de données. De plus, elle permet d'économiser de la bande passante et d'améliorer la sécurité des données en raison de leur proximité avec la source et de l'élimination des retards de réponse et de la latence, sans affecter la productivité. Selon Grand View Research, le marché mondial de l'informatique en périphérie valait 3,5 milliards de dollars US en 2019 — et ce chiffre devrait atteindre 43,4 milliards de dollars US d'ici 2027³⁶¹.

L'informatique en périphérie complète l'infonuagique dans un environnement informatique hybride. Alors que l'infonuagique s'appuie sur des centres de données centralisés, l'informatique en périphérie s'appuie sur des microcentres de données distribués à la périphérie du réseau, où les données sont utilisées plus près de la source. L'informatique en périphérie est nécessaire pour remédier aux lacunes des applications et services basés sur l'infonuagique en matière de performances et d'exigences réglementaires. C'est un enjeu, car la tendance à la numérisation pour améliorer l'efficacité et les performances alimente la demande d'applications qui nécessitent des performances de pointe, notamment les applications d'IdO. Les applications d'IdO nécessitent souvent une bande passante importante, une faible latence et des performances fiables, tout en respectant les mandats réglementaires et de conformité, ce qui en fait des candidats classiques pour l'informatique en périphérie. Le principe des applications de l'IdO et de l'informatique en périphérie est le même, quelle que soit la mise en œuvre exacte : des dispositifs ou des capteurs à une extrémité envoient des données à un centre de données en périphérie pour le traitement et éventuellement l'analyse, puis à une application plus centralisée (souvent dans le nuage) qui fournit le service. Les solutions logicielles de virtualisation, telles que les machines virtuelles (MV), les conteneurs et leurs outils de gestion ou de migration, ont été des catalyseurs actifs des systèmes informatiques en périphérie. Plus récemment, les réseaux gérés par logiciel ont permis d'explorer davantage le potentiel de l'informatique en périphérie en facilitant les technologies appropriées (par exemple, le réseau défini par logiciel (SDN), la virtualisation de la fonction réseau (NFV) et le réseau superposé).

Malgré des recherches et des améliorations constantes, la technologie informatique en périphérie est encore relativement jeune et des défis concernant les capacités des appareils — notamment la capacité de développer des logiciels et du matériel capables de gérer le déchargement des calculs à partir du nuage — sont susceptibles de se poser. La capacité des machines à basculer entre un calcul qui peut être effectué à la périphérie et un calcul qui nécessite le nuage est également un défi. La commercialisation éventuelle du réseau de communication mobile de

³⁶¹ <https://www.grandviewresearch.com/industry-analysis/edge-computing-market>

cinquième génération (5G) offre de nouvelles possibilités pour le développement de l'informatique en périphérie, notamment dans les environnements en temps réel. Les défis prometteurs pour la recherche et les technologies futures dans le domaine de l'informatique en périphérie comprennent la gestion de la mobilité, la sécurité des données et la protection de la vie privée, l'hétérogénéité, la fiabilité, les modèles de données de stockage et les environnements de simulation. Bien que certains cas d'utilisation puissent prouver la valeur de l'informatique en périphérie plus clairement que d'autres, l'impact potentiel sur l'écosystème croissant et connecté pourrait être substantiel.

7.2.6. Informatique quantique

La puissance de la recherche quantique exploite les lois quantiques de la nature pour développer de nouvelles technologies puissantes. La recherche quantique interdisciplinaire englobe la théorie et l'expérience, encourage les collaborations au-delà des frontières scientifiques et se concentre sur des domaines de recherche fondamentaux tels que l'informatique quantique³⁶², la communication quantique, la détection quantique et les matériaux quantiques. L'informatique quantique, en exploitant le comportement quantique des atomes, des molécules et des circuits nanoélectroniques pour une méthode de calcul radicalement différente et fondamentalement plus puissante, pourrait stimuler le développement de nouvelles percées dans les domaines de la science, des matériaux et de la médecine. Les applications et impacts potentiels de l'informatique quantique incluent la sécurité (par exemple, la cryptographie quantique), l'optimisation (par exemple, le recuit quantique³⁶³ et l'optimisation quantique adiabatique³⁶⁴, les problèmes de recherche non structurés), la suprématie quantique³⁶⁵, la simulation de processus physiques, chimiques et biologiques, et la conception de médicaments et de matériaux. Bien que cette technologie soit véritablement émergente, elle n'en est encore qu'à ses débuts et comporte de nombreux éléments inconnus. Par exemple, la manière d'écrire des codes traditionnels et d'avoir des parties qui tirent réellement parti de l'informatique quantique et qui le font d'une manière significative et efficace est un énorme défi, car il n'y a pas beaucoup de cadres permettant une programmation facile dans ce domaine. De plus, la plupart des chercheurs ne savent pas comment aborder un ordinateur quantique ni ce qu'il faut en faire. Les répercussions pourraient donc nécessiter une toute nouvelle méthode de programmation et de tout nouveaux langages.

7.2.7. Analyse des données

Depuis les années 1990, l'analyse des données a permis une révolution dans la production de connaissances au sein de la science et au-delà de celle-ci, en ouvrant la voie à de nouvelles méthodes très efficaces pour planifier, mener, diffuser et évaluer la recherche, ainsi que la création de nouvelles méthodes pour produire, stocker et analyser les données. L'analyse des données a facilité l'émergence du domaine de la science des données, qui rassemble des techniques informatiques, de codage, algorithmiques, statistiques et mathématiques, ainsi qu'une expertise du domaine pour extrapoler les connaissances issues de l'analyse des données. De la

³⁶² https://fr.wikipedia.org/wiki/Calculateur_quantique

³⁶³ https://en.wikipedia.org/wiki/Quantum_annealing

³⁶⁴ https://fr.wikipedia.org/wiki/Calcul_quantique_adiabatique

³⁶⁵ https://fr.wikipedia.org/wiki/Supr%C3%A9matie_quantique

finance aux soins de santé, l'analyse des données est omniprésente et, parallèlement à la promotion de la science ouverte et des données ouvertes, elle a encouragé le partage et l'interconnexion d'artefacts de recherche hétérogènes à l'aide d'IRN. La disponibilité de grandes quantités de données dans des formats exploitables par les machines incite à créer des procédures efficaces pour collecter, organiser, visualiser et modéliser ces données. Ces infrastructures axées sur les données servent à leur tour de plateformes pour le développement d'approches d'IA dans le but d'accroître la fiabilité, la rapidité et la transparence des processus de création de connaissances. Les chercheurs de toutes les disciplines peuvent bénéficier de la capacité de relier des données provenant de sources diverses, ce qui améliore la précision et le pouvoir prédictif des résultats scientifiques et aide à définir les futures orientations de la recherche. Comme l'illustre l'augmentation des financements, des programmes de formation et des lieux de publication dédiés, l'analyse des données est largement considérée comme l'avènement d'une nouvelle façon de mener des recherches et de remettre en question les conceptions actuelles de ce qui constitue la connaissance scientifique.

L'adoption des technologies d'analyse des données (par exemple Apache Hadoop, Spark, R, Python, lacs de données, bases de données NoSQL, analyse prédictive, analyse en continu, informatique en périphérie, blockchain, calcul en mémoire, sécurité des données, IA, gouvernance des données, analyse prescriptive) ne devrait pas ralentir de sitôt. Par exemple, la simulation à grande échelle est l'un des grands bénéficiaires de l'analyse des données massives et d'une puissance de calcul sans précédent. L'apprentissage machine est en train de fusionner avec l'analyse et les réponses vocales, tandis que l'IA permet d'obtenir des informations plus approfondies et une automatisation de plus en plus sophistiquée. Les applications à code source ouvert ont fini par dominer l'espace d'analyse des données, et cette tendance devrait se poursuivre. Les solutions d'analyse des données peuvent apporter une contribution décisive au débat sur le changement climatique et les événements approximatifs, du Big Bang à l'époque actuelle de l'univers.

7.2.8. Cybersécurité

Aujourd'hui, toutes les infrastructures de recherche à grande échelle dépendent des ressources des technologies de l'information et de la communication (TIC), offrant de nouvelles possibilités de collaboration et de partage géographiquement distribuées. Cette dépendance à l'égard des TIC a accru la nécessité de trouver des synergies et de développer des moyens de relever les défis liés aux TIC à un niveau fondamental, en fournissant des services efficaces et rentables qui peuvent être d'un usage général. Les IRN pour la mise en réseau, le CIP, la GDR et l'infonuagique évoluent rapidement ; la consolidation et l'intégration facilitent ainsi la prestation de services aux projets de recherche répartis à l'échelle internationale. Au-delà des e-infrastructures et des dispositifs numériques pour le réseau, le calcul et la GDR, il est nécessaire de développer des infrastructures de recherche spécifiques dans le domaine des sciences informatiques, en soutenant l'expérimentation de systèmes déstabilisateurs, y compris les infrastructures électroniques, le matériel informatique, les intergiciels et les logiciels, les protocoles, le calcul et les questions de cybersécurité. Les développeurs de LR n'ont généralement pas tenu compte de l'importance de la cybersécurité dans la conception de leurs logiciels ou de leurs infrastructures électroniques. En raison du nombre croissant d'intrusions et de violations d'informations confidentielles et de données de santé en particulier, les ingénieurs en logiciels de recherche doivent presque faire passer cet enjeu en tête de liste des priorités et s'assurer qu'ils disposent de mécanismes robustes qui instaurent une bonne sécurité. Dans le monde de la recherche

numérique hyperconnecté d'aujourd'hui, la cybersécurité n'est plus facultative, ce qui en fait un autre volet croissant de la recherche en informatique. Il s'agit d'un effort de co-conception entre tous les intervenants, où les nouveaux besoins des chercheurs ouvrent la voie aux efforts d'innovation.

7.2.9. Reproductibilité de la recherche

Il faut que les LR produisent des résultats reproductibles, surtout en cas de crise comme la pandémie actuelle. Le Saint Graal de la reproductibilité totale des LR a fait l'objet d'une grande attention dans de nombreuses disciplines, avec des outils développés allant des environnements de virtualisation ou de conteneurisation, aux carnets de notes. Cela pourrait inclure l'enregistrement automatique de toutes les valeurs des paramètres (y compris le réglage des graines aléatoires à des valeurs prédéterminées) ainsi que l'établissement des exigences dans l'environnement (par exemple, les dépendances). Parmi les exemples de systèmes actuellement utilisés par les chercheurs pour capturer et reproduire des environnements de recherche, il y a notamment les machines virtuelles et les conteneurs, l'infonuagique (7.2.2) et l'infrastructure en tant que code (7.2.3). Ces systèmes font maintenant partie du nouveau modèle de livraison de la pérennité (préservation). En plus des systèmes utilisés dans le modèle de livraison décrit ci-dessus, des pipelines de LR et des flux de travail scientifiques/computationnels³⁶⁶ sont construits pour automatiser la gestion des tâches dans une analyse complexe, permettant la reproductibilité de calcul. Pour plus d'exemples, voir l'[annexe F](#).

Les LR soutiennent et aident également la GDR (par exemple, la diffusion et le partage des données) — un élément crucial de la reproductibilité. L'adoption croissante de programmes comme Jupyter et R markdown, qui produisent des documents informatiques dynamiques contenant du code vivant et du texte descriptif, permet également à ceux qui réutilisent les données d'interagir directement avec les données ouvertes. Le support de ces outils logiciels peut de plus en plus être fourni par des dépôts pour aider la reproductibilité informatique (par exemple, l'intégration Git — Zenodo - Binder³⁶⁷).

7.2.10. Supercalculateurs exaflopiques et LR

Les supercalculateurs exaflopiques désignent des machines capables d'effectuer individuellement 10^{18} opérations en virgule flottante par seconde (un exaflop). Les architectures de ces derniers sont généralement des architectures hybrides constituées d'une hiérarchie de composants parallèles (nœuds, mécanismes d'échange, cœurs), dont les mécanismes d'échange sont souvent constitués de composants hétérogènes comprenant des coprocesseurs hautement parallèles tels que les GPU³⁶⁸. Chacun de ces composants possède généralement une mémoire embarquée, ce qui donne lieu à des hiérarchies de mémoire complexes avec différents niveaux de performance d'accès à la mémoire. Ces composants sont ensuite reliés par des interconnexions dont la bande passante et la latence de communication varient.

³⁶⁶ Par exemple, Kurator (<https://github.com/kurator-org>), Nextflow (<https://www.nextflow.io/>), nf-core (<https://nf-co.re/>), Vistrails (https://www.vistrails.org/index.php/Main_Page), Galaxy (<https://galaxyproject.org/>).

³⁶⁷ <https://blog.jupyter.org/binder-with-zenodo-af68ed6648a6?gi=87913247c7e4>

³⁶⁸ <http://www.intertwine-project.eu/introduction-interopability>

Ces architectures sont extrêmement difficiles à programmer efficacement, avec les techniques et les bibliothèques de programmation parallèle traditionnelles (OpenMP, MPI, OpenCL, CUDA, etc.) optimisées pour un seul niveau de cette architecture hiérarchique (par exemple, MPI pour le parallélisme entre les nœuds/cœurs et CUDA pour le parallélisme entre les GPU multicœurs). L'écriture de LR pour les systèmes de classe exaflopique requiert un ensemble de compétences uniques qui n'impliquent pas seulement des connaissances sur l'utilisation de bibliothèques parallèles telles que MPI et CUDA, mais aussi des connaissances et des compétences sur la combinaison de ces bibliothèques et l'optimisation des codes de sorte que l'application utilise les interconnexions entre les composants de la hiérarchie. Pour obtenir une efficacité élevée sur les systèmes de classe exaflopique, il faut éviter les goulots d'étranglement à tous les niveaux de cette architecture hétérogène et hiérarchique. Bien que la communauté des chercheurs travaille sur de nouveaux modèles de programmation pour ces architectures hétérogènes complexes, aucune solution de ce type n'est actuellement largement acceptée et utilisée. L'informatique exaflopique est abordée plus en détail dans l'exposé de position de l'Alliance sur le CIP.

8 Évaluation de l'état actuel

8.1. Points forts

8.1.1. PHQ

Au Canada, nous bénéficions à la fois d'un bon soutien aux programmes de recherche qui incluent des efforts de développement de logiciels et de solides programmes d'études en informatique sur le plan national. Cela comprend les programmes universitaires en génie logiciel³⁶⁹, les centres et instituts de recherche en logiciel³⁷⁰, des groupes de recherche³⁷¹, des programmes coopératifs³⁷² et des collaborations entre industrie et établissements dans les

³⁶⁹ Par exemple, l'Université Ontario Tech (<https://ontariotechu.ca/index.php>), le programme de génie logiciel de l'Université de Victoria (<https://ontariotechu.ca/index.php>), la maîtrise en leadership d'ingénierie (MEL) en systèmes logiciels fiables de l'Université de la Colombie-Britannique (<https://apscpp.ubc.ca/programs/mel/dependable-software-systems/>).

³⁷⁰ Par exemple, le Software Engineering Research Centre de l'Université Concordia (<https://www.concordia.ca/research/software-engineering-research-centre-serc.html>), l'Institut canadien des sciences statistiques (CANSSI) (<http://www.canssi.ca/>), PIMS (<https://www.pims.math.ca/>), le Fields Institute for Research in Mathematical Sciences (<http://www.fields.utoronto.ca/>), le Centre de recherches mathématiques (CRM) (<http://www.crm.umontreal.ca/en/index.shtml>), l'Atlantic Association for Research in the Mathematical Sciences (AARMS) (<https://aarms.math.ca/>).

³⁷¹ Par exemple, Indoc (<https://indocresearch.org/aboutus.html>), High Energy Physics group at SFU (<http://hep.phys.sfu.ca/>), l'équipe de développement de logiciels de la bibliothèque Robertson à l'UPEI (<https://library.upei.ca/vre>), Research Commons à l'UBC (<https://researchcommons.library.ubc.ca/>).

³⁷² Par exemple, l'Université de Waterloo (<https://uwaterloo.ca/co-operative-education/>), le programme d'enseignement coopératif de l'Université de Victoria (<https://www.uvic.ca/engineering/software/co-op/index.php>).

collèges³⁷³. Les activités de CANARIE, notamment le Colloque sur les logiciels de recherche³⁷⁴, la participation à l'International Research Software Engineering Leaders Workshop³⁷⁵, le rôle de membre fondateur de la Research Software Alliance³⁷⁶ (ReSA) et le lancement de l'initiative de soutien local aux logiciels de recherche, constituent également une base solide pour le développement des LR et des ILR. À titre d'exemple d'initiative axée sur les LR, le Perimeter Institute offre non seulement une expérience informatique de calibre mondial à ses chercheurs, mais il a également une pratique et une philosophie de génie logiciel de recherche bien établies, ce qui est inhabituel dans la communauté canadienne des LR. Au fur et à mesure que d'autres groupes de recherche universitaires développent et améliorent leurs équipes d'ILR, il sera possible de rassembler ces équipes au sein d'une organisation nationale d'ILR.

Un important bassin de talents en matière de développement de LR se dessine, en particulier au sein de l'industrie, ce qui attire des personnes plus qualifiées vers un écosystème prospère et dynamique. Par exemple, Mitacs a établi des partenariats avec 70 universités canadiennes, 6 000 entreprises et le gouvernement du Canada et les gouvernements des provinces afin d'exploiter des programmes de recherche et de formation à l'appui de l'innovation industrielle et sociale. Financé conjointement par ces paliers de gouvernement, ainsi que par des partenaires universitaires et d'innovation³⁷⁷, Mitacs soutient plus de 20 000 projets de recherche, plus de 3 600 collaborations de recherche internationales et a formé plus de 33 000 étudiants et postdoctorants au cours des 20 dernières années. L'un des objectifs de Mitacs est de combler l'écart de compétences entre l'industrie et le milieu universitaire et de faciliter le développement d'un solide bassin de talents. Pendant de nombreuses années, le CRSNG a offert des subventions d'engagement, qui ont également facilité les partenariats entre l'industrie et les universités, y compris de nombreux projets dans le cadre desquels des développeurs universitaires ont créé des codes pour des entreprises en fonction de leurs besoins particuliers. Les subventions Engage ne sont plus offertes aux universités, mais elles sont encore offertes aux collèges et elles sont aussi disponibles par l'entremise du CRSH.

L'existence d'un bassin de talents et d'une industrie logicielle dynamique peut être à la fois un atout et un défi, car la recherche doit rivaliser avec l'industrie pour attirer le personnel hautement qualifié, qui est généralement mieux rémunéré dans l'industrie. Malgré ce désavantage concurrentiel, les universités et les collèges sont en mesure d'offrir d'autres avantages (par exemple, la qualité de vie, un travail qui vise à avoir un impact social plutôt que financier) qui peuvent attirer de bons développeurs qui ont tendance à rester longtemps.

8.1.2. Financement et partenariats nationaux

Au Canada, il existe deux programmes de LR à l'échelle nationale : CANARIE et la FCI. Le programme Logiciels de recherche de CANARIE, qui finance le développement et le soutien des LR, sensibilise les gens aux LR et aux ILR en faisant la promotion des bonnes pratiques, ainsi

³⁷³ Par exemple, Mohawk (<https://www.mohawkcollege.ca/>)

³⁷⁴ <https://www.canarie.ca/fr/event/cclr-2021/>

³⁷⁵ <https://researchsoftware.org/>

³⁷⁶ <https://www.researchsoft.org/>

³⁷⁷ <https://www.mitacs.ca/en/about/partners>

qu'en tirant parti des ressources de l'infrastructure numérique de recherche. Les programmes de financement de CANARIE en matière de LR ³⁷⁸ et de GDR³⁷⁹ sont en place depuis plus de 12 et 7 ans respectivement et sont essentiels à de nombreuses plateformes de recherche canadiennes. Ce financement vise à la fois à encourager la réutilisation des LR existants et à promouvoir la sensibilisation et le suivi de la disponibilité et de la pérennité des outils. La FCC ne fournit pas de financement, mais offre un soutien décentralisé et régional aux utilisateurs ainsi que des plateformes pour mener des recherches sur les groupes de secteurs de CIP et les nuages de recherche.

Le défi sur la cyberinfrastructure lancé par FCI comprenait un concours en 2015 et un en 2017, visant à concevoir et à développer des infrastructures de recherche impliquant des chercheurs de consortiums, des scientifiques des données et des développeurs de logiciels de différents établissements. Ce financement comprenait un soutien à la FCC pour améliorer la capacité des chercheurs canadiens à mener des recherches de pointe à forte intensité de calcul et de données. Il s'agissait des premières incursions de la FCI dans le domaine du soutien des logiciels au niveau de la plateforme (par opposition au soutien du matériel), en réponse à la prise de conscience croissante des défis que pose le maintien de l'infrastructure des logiciels. Une autre bourse de la FCI est le Fonds des leaders John R. Evans (FLJE)³⁸⁰, un programme de subvention en partenariat³⁸¹ avec le CRSH, le CRSNG ou le CRC/CERC qui répond aux besoins des chercheurs individuels en matière d'infrastructure de recherche. Il est également possible de soumettre des demandes parallèles à une évaluation conjointe en même temps ou de faire une demande subséquente d'une bourse du FLJE.

Des organismes non canadiens ont financé certains efforts de LR au Canada, notamment dans le domaine des SHS. Plusieurs plateformes ont réussi à obtenir un financement de l'UE dans le cadre du programme H2020, notamment iReceptor³⁸² et CanDIG³⁸³ par le biais de l'appel conjoint de l'UE Horizon 2020 et des IRSC³⁸⁴. Il devient assez courant que les IRSC s'associent au Wellcome Trust et à d'autres organismes internationaux pour financer les LR (p. ex. le projet FAIRsFAIR³⁸⁵). La Mellon Foundation finance un nombre important de projets d'infrastructure de

³⁷⁸ <https://www.canarie.ca/software/funding/lrss-call1/>

³⁷⁹ <https://www.canarie.ca/rdm/funding/call2/>

³⁸⁰ <https://www.innovation.ca/fr/appel-gestion/financement/fonds-leaders-john-r-evans>

³⁸¹ https://www.innovation.ca/sites/default/files/Funds/JELF/nov2018/guidelines_for_completing_a_proposal_-_partnerships.pdf

³⁸² <https://gateway.ireceptor.org/login>

³⁸³ <https://www.distributedgenomics.ca/>

³⁸⁴ <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/sc1-bhc-05-2018>

³⁸⁵ <https://www.fairsfair.eu/news/rori-and-fairsfair-partner-fairware-initiative-funders-across-uk-europe-and-canada-join-forces>

recherche numérique et de développement de logiciels dans le monde entier^{386,387}. Depuis 2012, avec le soutien de la Fondation Sloan³⁸⁸, le programme de bourses postdoctorales du CLIR (The Council on Library Resources)³⁸⁹ a apporté un soutien à la fois à la curation des données de recherche et à la curation des logiciels en plaçant 49 scientifiques et spécialistes des sciences sociales dans 38 établissements d'accueil aux États-Unis et au Canada. Ces boursiers ont aidé à cultiver une compréhension plus sophistiquée de la curation des données et des logiciels qui contribue à un environnement numérique pérenne pour la recherche.

L'Initiative des Supergrappes d'innovation³⁹⁰ est une stratégie récente soutenue par le gouvernement du Canada qui transformerait les efforts régionaux d'innovation et de création d'emplois, en mettant au défi les organisations du secteur privé canadien de toutes tailles de collaborer avec les institutions de recherche. En 2018, l'Initiative des Supergrappes d'innovation a annoncé cinq Supergrappes au Canada (Technologie numérique³⁹¹, IPC³⁹², NGen³⁹³, Scale AI³⁹⁴, et Économie océanique³⁹⁵), avec 950 millions \$ sur 5 ans pour promouvoir la collaboration à grande échelle entre les leaders de l'industrie, les PME et les universitaires pour développer et étendre des technologies à fort potentiel au Canada. La Supergrappe Technologie numérique dispose d'enveloppes de financement pour des composantes particulières des LR, comme les dépôts³⁹⁶.

Le programme de bourses de nouveaux chercheurs (BNC) de l'Ontario est un programme discrétionnaire, non admissible, qui encourage les nouveaux chercheurs travaillant dans des établissements de recherche financés par les fonds publics de l'Ontario à constituer une équipe de recherche, quelle que soit leur discipline. Chaque bourse accordée à un chercheur principal fournit jusqu'à 100 000 \$, avec un montant supplémentaire de 50 000 \$ devant être égalé par

³⁸⁶ https://mellon.org/grants/grants-database/advanced-search/?amount-low=&amount-high=&year-start=&year-end=&city=&state=&country=Canada&q=infrastructure&per_page=25

³⁸⁷ https://mellon.org/grants/grants-database/advanced-search/?amount-low=&amount-high=&year-start=&year-end=&city=&state=&country=Canada&q=software&per_page=25

³⁸⁸ <https://sloan.org/>

³⁸⁹ <https://www.clir.org/fellowships/postdoc/>

³⁹⁰ <https://www.ic.gc.ca/eic/site/093.nsf/fra/accueil>

³⁹¹ <https://www.ic.gc.ca/eic/site/093.nsf/fra/00011.html>

³⁹² <https://www.ic.gc.ca/eic/site/093.nsf/fra/00012.html>

³⁹³ <https://www.ic.gc.ca/eic/site/093.nsf/fra/00010.html>

³⁹⁴ <https://www.ic.gc.ca/eic/site/093.nsf/fra/00009.html>

³⁹⁵ <https://www.ic.gc.ca/eic/site/093.nsf/fra/00013.html>

³⁹⁶ <https://www.ic.gc.ca/eic/site/093.nsf/fra/00018.html> : parmi les exemples de projets, il y a l'analyse environnementale par satellite, les données communes sur l'eau douce, l'agriculture de précision pour améliorer la santé des cultures, etc.

l'établissement du chercheur ou un organisme partenaire. Ces fonds peuvent être utilisés pour embaucher du personnel hautement qualifié, pouvant servir au développement de logiciels³⁹⁷.

8.2. Défis et opportunités

8.2.1. Manque de financement ciblé et durable

Le Canada distribue des fonds importants pour la recherche et le développement (R-D), l'innovation et les projets de commercialisation³⁹⁸, mais il y a peu de financement ciblé, spécifique ou même à long terme pour les LR. En comparaison avec d'autres pays ([annexe I](#)), le Canada accuse un retard dans le soutien d'une communauté de développement des LR efficace, cohésive et coordonnée, à l'exception des programmes de CANARIE. En revanche, l'Australie et l'Europe sont des exemples de régions qui ont investi des sommes importantes dans un effort ciblé visant à créer des infrastructures de recherche solides pour tous les chercheurs. La raison fondamentale en est que le développement des LR et des plateformes n'a pas été une priorité stratégique des organismes de financement, à l'exception de CANARIE et de l'initiative de la FCI mentionnée précédemment. En général, la FCI finance les projets d'infrastructure informatique centralisée nationale qui proposent le développement des logiciels nécessaires au fonctionnement de l'infrastructure, mais tout ce qui a trait au développement des LR après la période d'attribution n'est pas admissible. Cela peut laisser les chercheurs dans l'embarras pour trouver d'autres programmes afin de financer la maintenance et le développement ultérieur, ce qui est particulièrement problématique lorsqu'une communauté de pratique considérable dépend de cette infrastructure logicielle. Cela s'explique en partie par le fait que les LR sont généralement développés comme un sous-produit de la recherche propre à un domaine et que, jusqu'à récemment, ils n'ont pas été considérés comme un produit de « première classe » ou un « constituant d'IRN ». Cela représente une lacune et un défi importants pour les chercheurs canadiens et diminue le rôle du Canada comme chef de file dans le développement de l'infrastructure de recherche. Les programmes de financement qui existent (section 6.1.1.2) sont soit propres à un domaine (p. ex. les programmes nationaux et régionaux de Génome Canada et de Genome British Columbia³⁹⁹ [Genome BC]), soit ils financent indirectement le développement continu et durable des LR, par exemple lorsqu'un chercheur principal développe lui-même des LR ou utilise des fonds de recherche à cette fin. Il s'agit d'une approche provisoire pour la mise en place d'une infrastructure de recherche au Canada ; il y a de grandes possibilités de croissance de la communauté des LR par le biais d'intervenants travaillant ensemble pour coordonner le financement dans ce domaine.

La FCC a créé des logiciels d'infrastructure de CIP dans de nombreux domaines, notamment la planification des tâches, la distribution des logiciels, l'automatisation du déploiement infonuagique et la base de données de Calcul Canada (CCDB)⁴⁰⁰, mais il s'agit généralement d'un processus organique qui doit trouver un équilibre entre la nécessité, les ressources disponibles et l'expertise du personnel existant au lieu d'avoir une vision stratégique à long terme. Plus important encore,

³⁹⁷ <https://www.ontario.ca/fr/page/bourses-de-nouveaux-chercheurs>

³⁹⁸ <https://www.canada.ca/fr/services/science/innovation/financement.html>

³⁹⁹ <https://www.genomebc.ca/>

⁴⁰⁰ Base de données de Calcul Canada (CCDB) : <https://ccdb.computecanada.ca/security/login>

l'équivalent des investissements internationaux destinés aux logiciels d'infrastructure de CIP (par exemple, ceux des États-Unis ou de l'UE) n'a pas été réalisé au Canada. Cette lacune pourrait être comblée en reconnaissant son importance, ainsi que le financement et le soutien connexes pour les plateformes de type Portail des sciences et les intergiciels sous-jacents qui pourraient être transmis aux fournisseurs d'infrastructure comme la FCC qui n'ont pas le mandat stratégique de soutenir de tels projets.

Un autre défi connexe auquel se heurtent certains organismes de financement est la manière d'équilibrer le soutien entre les infrastructures génériques et les infrastructures axées sur les disciplines. La plupart des infrastructures de recherche dans tous les domaines comprennent désormais une infrastructure numérique importante dans leurs éléments de base, et il peut être difficile de différencier l'infrastructure numérique fournie par ces IRN et les IRN génériques. Comme le note le rapport RISCAPÉ⁴⁰¹, un point de différenciation est que les infrastructures de recherche dans des domaines particuliers ont tendance à utiliser principalement les services de base d'IRN : « Il est difficile de trouver des infrastructures de recherche qui utilisent des services de niveau supérieur développés dans le domaine de l'infrastructure électronique ; elles développent plutôt leurs propres services dans leurs domaines respectifs ». Il s'agit d'un défi que CANARIE a tenté de relever par le biais d'un financement ciblé et en encourageant les plateformes existantes à intégrer les communautés de recherche en dehors de leurs disciplines.

Bien que les chercheurs comprennent généralement la valeur des LR en tant qu'élément fondamental du processus de recherche et en tant que produit ayant un impact, les organismes de financement ont été plus lents à reconnaître et à soutenir ce travail universitaire. Par exemple, les trois organismes subventionnaires ne permettent généralement pas l'achat ou le développement de LR comme dépense admissible⁴⁰². Ainsi, les étudiants des cycles supérieurs du laboratoire développent les LR plutôt que d'embaucher des ingénieurs logiciels de recherche qualifiés pour les projets. Le CRSH disposait auparavant d'un volet de financement dédié au développement d'outils expérimentaux, y compris les LR, mais ce volet a été supprimé. S'il est possible d'obtenir des fonds pour développer un logiciel de recherche en tant qu'élément de soutien d'un projet de recherche, il y a très peu de sources de financement conçues pour développer un logiciel de recherche en tant que résultat principal d'un projet de recherche. Cette lacune dans le financement est une occasion pour que d'autres organismes de financement soutiennent le développement des LR d'une manière délibérée et complémentaire aux trois organismes.

Il existe d'autres obstacles administratifs et opérationnels auxquels les chercheurs peuvent être confrontés lorsqu'ils travaillent à la création d'importants outils de LR. Comme nous l'avons mentionné à la section 9.1.2, malgré la possibilité de présenter des demandes parallèles avec les bourses du FLJE, celles-ci sont déduites des enveloppes de la FCI, ce qui signifie que la possibilité de présenter une demande est soumise à des processus de priorisation internes. Il y a aussi une exigence de fonds égalé, ce qui, dans certains cas, peut signifier que les chercheurs

⁴⁰¹ <https://riscape.eu/riscape-report/>

⁴⁰² https://www.unb.ca/research/ors/resources/tri-agency_eligible_research_grant_expenses_and_supporting_evidence_requirements_6mar2017.pdf

principaux ayant obtenu une subvention de partenariat⁴⁰³ ne peuvent pas demander de bourse du FLJE pour l'infrastructure d'accompagnement, malgré la possibilité théorique de le faire, et donc que les projets qui ont été évalués par des pairs et qui ont été retenus dans le cadre du programme de financement de la recherche en sciences humaines le plus prestigieux du pays ne peuvent pas demander d'infrastructure pour soutenir cette recherche. Il serait avantageux d'avoir un financement sans enveloppe (comme c'était le cas avec le programme de cyberinfrastructure de la FCI) pour promouvoir la collaboration entre établissements dans le développement de l'infrastructure. Un autre défi a été le modèle de financement 40/40/20 de la FCI, qui peut s'avérer particulièrement difficile avec les subventions de collaboration qui couvrent plusieurs provinces, et la gamme d'approches de l'appariement provincial. En pratique, le prototypage et le développement de LR sont financés par diverses voies de financement du CRSH, notamment les subventions Développement de partenariats et Savoir⁴⁰⁴, mais historiquement, l'absence de lignes directrices claires sur l'admissibilité au financement du développement de LR par le CRSH a fait en sorte que les résultats dépendaient fortement de l'expertise des comités d'évaluation.

La relation différente entre la production, la diffusion et la réutilisation des données, dans laquelle la recherche, les LR, la GDR et la publication se chevauchent, complique la pérennité des LR dans le contexte du CRSH. Possiblement plus que dans d'autres domaines, bon nombre d'artefacts numériques, c'est-à-dire beaucoup de données, dans les sciences humaines demeurent pertinents au-delà du cycle de recherche initial au cours duquel ils ont été créés ; ils font l'objet d'une évaluation, d'une interprétation, d'une actualisation, d'une amélioration et d'une réutilisation continues par d'autres chercheurs. Les éditions savantes de sources primaires, les archives numérisées et les collections de recherche thématiques de matériaux épars, notamment, constituent une forme de publication qui fait progresser la recherche des chercheurs qui les produisent⁴⁰⁵ et qui est souvent essentielle au travail des autres chercheurs et étudiants, et précieuse pour le grand public, dans le domaine et ailleurs, pendant une génération ou plus. Ainsi, le fait de rendre interopérables et calculables de grandes collections numérisées ou des sous-ensembles de celles-ci produits par les chercheurs, comme les Data Capsules du Hathi Trust Research Centre qui permettent le traitement informatique d'ensembles de contenus culturels protégés par le droit d'auteur, est en soi un aspect important des LR pour de nombreux domaines⁴⁰⁶. Cela signifie également que le maintien d'un accès significatif à de tels ensembles de données de recherche en sciences humaines sur le Web, en plus de la préservation à long terme, est un élément d'infrastructure essentiel pour soutenir la recherche en sciences humaines⁴⁰⁷. Lorsque les LR des SHS ne sont pas maintenus, il y a un risque important de perte

⁴⁰³ https://www.sshrc-crsh.gc.ca/about-au_sujet/partnerships-partenariats/partnership_grants-bourses_partenariats-fra.aspx

⁴⁰⁴ https://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/insight_grants-subventions_savoir-fra.aspx

⁴⁰⁵ Muñoz, Trevor, 2013. Data curation as publishing for the digital humanities. *Journal of Digital Humanities*, 2(3), pp.14-22.

⁴⁰⁶ Zeng, Jiaan, Guangchen Ruan, Alexander Crowell, Atul Prakash, et Beth Plale. "Cloud computing data capsules for non-consumptive use of texts". Dans les Actes du 5th ACM workshop on Scientific cloud computing, pp. 9-16. 2014.

⁴⁰⁷ Terras, Melissa, James Baker, James Hetherington, David Beavan, Martin Zaltz Austwick, Anne Welsh, Helen O'Neill, Will Finley, Oliver Duke-Williams et Adam Farquhar. "Enabling complex analysis of

de données ou de perte d'accès^{408,409}, par exemple le site menacé du Métis Archival Project (MAP) ou la défunte plateforme ArtMob ; il y a aussi le risque de répétition d'un travail important pour sauver les données en vue d'une utilisation active, c'est-à-dire l'accessibilité en ligne, si elles sont restaurées, comme dans le cas des archives numériques Fred Wah. Le code lui-même est un artefact culturel qui nécessite des efforts considérables pour le préserver et le rendre accessible, comme en témoignent des initiatives telles que les archives du patrimoine culturel soutenues par l'UNESCO⁴¹⁰. Les artefacts de code interactifs complexes tels que la littérature électronique et les jeux vidéo posent des défis uniques qu'il est essentiel de relever pour les rendre susceptibles d'être analysés socialement et culturellement à l'avenir.

L'infonuagique et la gestion efficace des LR dans le nuage et les solutions pour garder le contenu accessible sur le Web sont donc un aspect essentiel de l'IRN des SHS. Le projet Endings⁴¹¹ promeut les bonnes pratiques pour permettre la « dégradation gracieuse » (Nowviskie et Porter) des projets Web terminés, par exemple par l'utilisation de sites Web statiques (par exemple en utilisant github.io). Ces méthodes propres au SHS doivent être promues et l'offre de versions gérées de LR libres, courantes et basées sur des normes, comme Omeka ou Mukurtu, les rendrait plus attrayantes et accessibles aux chercheurs et présenterait des avantages encore plus grands en termes de cohérence des métadonnées et du format des données pour soutenir le partage et la réutilisation des données. Les méthodes et les outils utilisés dans les disciplines du CRSH sont si diversifiés qu'aucun campus ne peut les prendre tous en charge. Un bassin commun d'expertise sur le plan national permettrait donc aux chercheurs de faire des choix plus éclairés parmi un plus grand nombre d'options. De nombreux chercheurs adoptent actuellement des approches à faible barrière, comme les sites WordPress, qui sont utilisés pour présenter à la fois des données utiles à d'autres et des résultats de recherche importants (p. ex. Yellow90s⁴¹²). Le soutien de ces deux stratégies permettrait de créer des liens dès les premiers stades avec les chercheurs de manière à faire progresser les bonnes pratiques en matière de méthodes de recherche et de GDR, idéalement par le biais d'équipes locales de soutien liées à des réseaux plus vastes donnant accès à une expertise pour les LR en question.

La pérennité est la capacité de durer. Pour les LR, la pérennité signifie que les LR continueront à être disponibles et fonctionnels à l'avenir, tant que le besoin existe ou jusqu'à ce qu'ils soient remplacés, ce qui nécessite un effort humain pendant toute cette période. Puisque la pérennité des LR accorde un avantage global pour la communauté des chercheurs et une meilleure utilisation des fonds de recherche, les bailleurs de fonds doivent s'y intéresser davantage. CANARIE a reconnu ce besoin et a choisi de passer du financement de nouveaux LR à celui de

large-scale digital collections: humanities research, high-performance computing, and transforming access to British Library digital collections." *Digital Scholarship in the Humanities* 33, no 2 (2018) : 456-466.

⁴⁰⁸ Edmond, Jennifer, et Francesca Morselli. " Sustainability of digital humanities projects as a publication and documentation challenge". *Journal of Documentation* (2020).

⁴⁰⁹ https://robincamille.com/presentations/death_of_digital_scholarship/

⁴¹⁰ <https://www.softwareheritage.org/>

⁴¹¹ <https://endings.uvic.ca/>

⁴¹² <http://www.1890s.ca/>

la pérennité des LR en développant et en élargissant les projets déjà financés et en encourageant la réutilisation des LR. Bien entendu, cela a entraîné une diminution du financement des nouveaux projets de LR par CANARIE. Cela démontre les défis auxquels font face les bailleurs de fonds lorsqu'il s'agit de trouver un équilibre entre le financement de nouveaux projets de LR et la pérennité des projets existants qui sont encore utiles. L'établissement d'un tel équilibre financier est important pour la santé de l'écosystème des LR.

Un autre défi concernant le financement des projets de LR concerne l'administration et la gestion du financement. Deux des principaux programmes canadiens de financement des LR, le programme Logiciels de recherche de CANARIE et le programme Défi 1 de la cyberinfrastructure de la FCI, ne font pas partie du modèle courant de subvention des trois organismes. Par conséquent, la demande et la gestion de projets dans le cadre de ces programmes ne font pas partie de l'expérience habituelle des chercheurs en matière de subventions. Les deux programmes ont des processus de demande (p. ex. les subventions de la FCI sont généralement destinées à de grandes infrastructures physiques plutôt qu'à des logiciels), des implications administratives, des ressources humaines et juridiques (p. ex. le programme de CANARIE est un modèle de contrat basé sur les réclamations) et des exigences en matière de rapports (p. ex. CANARIE et la FCI ont tous deux des exigences strictes en matière de rapports) très différents de ceux des projets financés par les organismes des trois organismes. Par conséquent, la demande et l'administration de ces projets entraînent de lourds frais administratifs qui dissuadent même les candidats les plus qualifiés et que même les grands bureaux de recherche ont du mal à assumer. Il convient de noter que ces frais généraux sont inhérents aux exigences des bailleurs de fonds envers le gouvernement fédéral (p. ex. l'accord de contribution de CANARIE) et que les bailleurs de fonds n'ont d'autre choix que de les imposer. Il existe donc des possibilités de rationaliser les processus de demande et de gestion des subventions de LR.

Une autre déconnexion clé dans les filières traditionnelles de financement de la recherche est que les bailleurs de fonds, et donc les évaluateurs, recherchent la nouveauté directe et immédiate comme l'un des critères d'examen des propositions sollicitées, ce qui, à un certain point du cycle de vie des LR importants, est contraire aux besoins de la pérennité des LR, où la maintenance conduit à l'utilisation qui conduit à la nouveauté ultérieure par ces utilisateurs, plutôt que la nouveauté immédiate, est la clé. Cette exigence de nouveauté incite les chercheurs à développer de nouvelles LR plutôt que de réutiliser des composants existants ou d'ajouter des améliorations itératives, ce qui constitue un gaspillage et est contre-productif. Elle conduit également à des propositions compétitives de la part de chercheurs principaux qui seraient mieux servis en travaillant dans un environnement plus collaboratif où les bonnes pratiques de développement de LR sont la norme plutôt que l'exception.

Les LR sont souvent financés en tant qu'infrastructure dans le cadre d'un cycle de subventions de trois à cinq ans, alors que le projet typique de LR devrait durer beaucoup plus longtemps, surtout lorsqu'il dessert une grande communauté. Le maintien, l'adaptation et l'expansion des LR sont des exigences fondamentales de la pérennité des LR. Si les bailleurs de fonds prévoient de favoriser la pérennité des LR, ils doivent fournir des fonds à ces fins. Il a été noté que l'EPSRC et le Biotechnology and Biological Sciences Research Council (BBSRC) du Royaume-Uni ont tous deux lancé des appels de demandes de financement pour la maintenance de logiciels, à

savoir Software For The Future II⁴¹³ et Bioinformatics and Biological Resources (BBR) Fund, respectivement⁴¹⁴.

Même les bons LR ne seront pas utilisés si personne n'est au courant de leur existence. Le fait de permettre la découverte et la disponibilité des LR réduira les efforts de duplication et favorisera la réutilisation des LR, ce qui permettra de concentrer le financement et les efforts des chercheurs sur la recherche plutôt que sur le développement des LR. Si les développeurs sont en mesure d'articuler des cas d'utilisation pour leur LR, il y a plus de chances qu'ils parviennent à obtenir un financement pour maintenir ce LR. Il est tout aussi important de concevoir le logiciel en tenant compte de la réutilisation en dehors du groupe d'origine, en tirant parti de la flexibilité, des architectures de plugiciel, etc. qui est peut-être le plus grand obstacle à la réutilisation par d'autres. Dans les logiciels commerciaux, un plus grand nombre d'utilisateurs signifie généralement plus de revenus, ainsi que la possibilité de soutenir ces utilisateurs, d'ajouter de nouvelles fonctionnalités, etc. Cependant, dans le domaine des LR, en l'absence de financement pour embaucher du personnel hautement qualifié (PHQ) pour le soutien aux utilisateurs, un plus grand nombre d'utilisateurs signifie plus de temps nécessaire pour soutenir ces utilisateurs et moins de temps pour ajouter des fonctionnalités, entreprendre la recherche réelle, etc. Ces facteurs dissuasifs entravent le développement durable des LR. Une nouvelle approche quant au financement des LR doit reconnaître différents types et phases de LR (p. ex. expérimentaux, émergents, mais au niveau de la production, établis/entreprise), et concevoir des mécanismes d'évaluation, des mesures et des flux de financement appropriés pour chacun.

8.2.2. Technique

L'écosystème des LR est diversifié et complexe. Les LR incluent de nombreux logiciels, des systèmes aux applications, en passant par la modélisation, les passerelles, l'analyse, les algorithmes, les intergiciels et les bibliothèques ; des progiciels hautement développés avec des algorithmes et des cadres de physiques multiples utilisés par d'importantes bases d'utilisateurs aux modèles de programmation et aux abstractions pour les données, le matériel et la science, ou encore des scripts, des carnets de notes et des programmes très courts et simples, écrits par les chercheurs pour leur propre usage. En ce qui concerne les plateformes de LR, le défi du développement d'une infrastructure propre à un domaine par rapport à une infrastructure générique est également compliqué. Dans le cas de certains programmes de financement de CANARIE, si le projet proposé n'était pas en mesure de prendre des outils ou des plateformes de LR axés sur un domaine et les rendre plus généraux, il n'était pas admissible à ce programme de financement.

Au cours des 15 à 20 dernières années, d'énormes efforts ont été déployés pour faire de la sensibilisation sur l'importance des données pour la recherche, où l'intérêt et les investissements dans la GDR et les bonnes pratiques connexes ont considérablement amélioré la façon dont les données sont traitées et conservées. Cependant, les logiciels ne sont pas des données. Par rapport aux données qui sont généralement statiques, les logiciels sont des entités évolutives qui doivent s'adapter aux changements constants de leur environnement : s'ils n'évoluent pas, les

⁴¹³ <https://epsrc.ukri.org/funding/calls/softwareforthefuture/>

⁴¹⁴ <https://webarchive.nationalarchives.gov.uk/ukgwa/20200930155721/https://bbsrc.ukri.org/funding/filter/2019-bioinformatics-biological-resources-fund/>

logiciels se dégradent. Bien que les données de recherche puissent aussi se dégrader, le contexte est très différent pour les LR.

Les logiciels dépendent des autres logiciels pour leur fonctionnement normal : ces composants sont connus sous le nom de dépendances, et ils sont variés : du système d'exploitation, des bibliothèques du système, à d'autres paquets nécessaires (par exemple un navigateur, l'environnement d'exécution Java [JRE]). L'une des façons dont les logiciels se dégradent est par le biais d'une modification de l'un des composants logiciels dépendants qui peut affecter le fonctionnement des logiciels ; les grandes plateformes de LR augmentent le risque que cela se produise. Si l'on veut que les LR restent utiles, ils doivent être soutenus pendant leur cycle de vie actif, au lieu d'être simplement préservés.

Le développement de logiciels a connu une croissance exponentielle au cours du dernier demi-siècle, et les outils et les plateformes qui les prennent en charge ont évolué à un rythme rapide. Le nombre d'éléments de code source publics originaux double tous les 22 mois et les engagements originaux doublent tous les 30 mois⁴¹⁵. Les logiciels fonctionnent à l'échelle mondiale ; l'accélération du développement de l'innovation, de la recherche et de la technologie est incroyablement rapide. Pour rester pertinents, les LR doivent évoluer à un rythme comparable.

Un autre impact sur le développement des LR est le fait que la recherche elle-même évolue à un rythme rapide, ce qui met encore plus de pression sur les acteurs de l'écosystème des LR pour réagir en conséquence. La réponse à la pandémie de COVID-19 en est un bon exemple : de nombreux programmes de recherche au Canada et ailleurs ont rapidement pivoté pour répondre aux multiples défis présentés par la crise. Par exemple, les plateformes iReceptor⁴¹⁶, QUÉBEC COVID-19 BIOBANK⁴¹⁷ (BQC19), CanCoGen⁴¹⁸ et CanDig⁴¹⁹ ont réagi rapidement en apportant des changements à leurs logiciels afin d'accommoder la recherche axée sur la COVID-19.

Le partage des LR est encouragé par certaines communautés et promu par les bailleurs de fonds, les éditeurs et les défenseurs des LR, mais n'est généralement pas obligatoire. Le manque de diffusion et de partage des LR entrave la transparence, la reproductibilité et la vérification de la recherche. Mis à part les dépôts de LR suivant des bonnes pratiques comme GitHub, il n'y a aucun autre dépôt, catalogue, registre ou service de LR au Canada, à l'exception du portail de LR de CANARIE⁴²⁰ et du cadre de la FCC. De plus, la FCC offre divers logiciels (de 400 à 800 selon les types de plateformes matérielles) sur des sites hôtes pour leurs utilisateurs. Le nombre

⁴¹⁵ Rousseau, G., Di Cosmo, R., & Stefano Zacchiroli. (2020). Software provenance tracking at the scale of public source code. *Empirical Software Engineering*, 1–30. <https://doi.org/10.1007/s10664-020-09828-5>

⁴¹⁶ <http://www.ireceptor.org/> Scott et al. The adaptive immune receptor repertoire community as a model for FAIR stewardship of big immunology data, *Current Opinion in Systems Biology* Volume 24, décembre 2020, pages 71-77, DOI : 10.1016/j.coisb.2020.10.001

⁴¹⁷ <https://bqc19.ca/fr>

⁴¹⁸ <https://virusseq-dataportal.ca/>

⁴¹⁹ CanDIG : Secure Federated Genomic Queries and Analyses Across Jurisdictions: <https://www.biorxiv.org/content/biorxiv/early/2021/03/31/2021.03.30.434101.full.pdf>

⁴²⁰ <https://science.canarie.ca/researchsoftware/home/main.html>

réel d'outils et de plateformes de LR est difficile à quantifier, mais il est plusieurs fois supérieur à ce qui est facile à répertorier à partir des services nationaux. L'élargissement de l'innovation et de l'infrastructure pour répondre aux différentes exigences des communautés de pratique est un défi sans la coordination ni le soutien à l'échelle nationale. Pour assurer la pérennité des efforts de LR du Canada, il faut aider les chercheurs à consigner leurs logiciels dans des dépôts de LR accessibles qui catalogueront, indexeront, préserveront, organiseront et fourniront des identifiants pérennes. Le sous-groupe sur les bonnes pratiques pour les registres de logiciels du groupe de travail sur la mise en œuvre de la citation de logiciels de FORCE1⁴²¹ a présenté un ensemble de neuf bonnes pratiques pouvant aider à définir la portée, les pratiques et les règles qui régissent les registres et les dépôts de LR individuels⁴²². L'élaboration de ces bonnes pratiques ne sera bénéfique que si elle s'accompagne d'une stratégie d'investissement durable pour le développement d'outils et de plateformes de LR.

Il est encore plus difficile d'écrire des LR pour une plateforme complexe de CIP. Si l'on prend l'exemple de la programmation basée sur les groupes de secteurs, les logiciels et le matériel vont de pair lorsqu'il s'agit de réaliser du CIP sur un groupe de secteurs. Les programmes doivent être codés pour exploiter le matériel sous-jacent, tandis que les programmes traditionnels à architecture non parallèle doivent être réécrits pour être performants sur un groupe de secteurs. De plus, il faut tenir compte des différentes approches de la programmation parallèle (approche distribuée ou à mémoire partagée), de la gestion des intrants et des extrants, des bibliothèques qui pourraient ne pas être prises en charge à l'avenir, etc. En dépit de ces défis, lorsque les LR pour les groupes de secteurs sont développés et gérés de manière durable, ils peuvent être maintenus pendant des décennies, améliorant finalement le rapport coût-efficacité ainsi que les avancées de la recherche. Les services nationaux ont un rôle essentiel à jouer à cet égard, qu'il s'agisse d'offrir des LR interopérables et réutilisables sur plusieurs plateformes ou de soutenir l'éducation et la formation.

Un défi connexe est de savoir comment intégrer et coordonner les trois composantes du CIP, de la GDR et des LR dans un écosystème aussi dynamique et évolutif. Grâce à une infrastructure intégrée et inclusive (p. ex., un « guichet unique »), les chercheurs peuvent bénéficier de la production, de la gestion et du partage des résultats de leur recherche tout en permettant la réutilisation et la reproductibilité de la recherche, ainsi que de collaborations et d'engagements facilités entre les domaines et les établissements.

8.2.3. Compétences, connaissances et formation

Comme le souligne l'analyse de l'environnement de l'IRN au Canada préparée dans le cadre du sommet de 2014, « il existe un important besoin non satisfait de mise à niveau des compétences, de formation et de mentorat dans l'utilisation de l'informatique en périphérie, en particulier dans

⁴²¹ <https://github.com/force11/force11-sciwg>

⁴²² <https://arxiv.org/abs/2012.13117v1>

les disciplines qui n'ont pas eu un engagement important dans la recherche intensive en données jusqu'à récemment »⁴²³.

Certains chercheurs savent coder, mais peu d'entre eux comprennent l'ensemble des compétences nécessaires pour développer des logiciels fiables, durables, reproductibles et réutilisables ; les compétences pour celles-ci relèvent généralement du génie logiciel. D'un autre côté, certains chercheurs ne devraient pas être obligés de devenir des ILR. Au contraire, les chercheurs devraient se perfectionner en fonction de leur degré de participation dans les logiciels. Des organisations comme Software Carpentry s'attaquent à ce problème en formant les chercheurs sur le plan des compétences informatiques de la recherche. Cette formation est également l'une des priorités de la FCC qui propose une grande variété de formations - programmation, MPI, analyse de données, AM, CUDA, écriture de script, etc. — dispensées sous de multiples formes (cours d'été universitaires, webinaires, ateliers, etc.) à des milliers de participants. Ce type de formation contribue à accroître le niveau de compétences général de la communauté des chercheurs, mais si nous voulons améliorer les compétences de manière significative, il faut veiller à ce que les compétences en matière de logiciels soient acquises au tout début d'une carrière de chercheur, afin de garantir que ces compétences soient utilisées tout au long de cette carrière. Il faut remédier au manque de compétences en matière de développement de logiciels dans les filières d'enseignement et de recherche et les intégrer dans les programmes de formation doctorale. De plus, l'absence d'une instance organisée d'ILR rend plus difficile l'offre de formation à ce groupe. Il est également important de parvenir à une plus grande sensibilisation aux bonnes pratiques en matière de développement durable des guichets uniques, ce qui peut être facilité par des organisations (centralisées ou distribuées) qui agissent comme des points focaux pour l'expertise, non seulement pour partager les connaissances et les compétences, mais aussi pour améliorer la mise en réseau et la collaboration à l'échelle internationale.

8.2.4. Attraction et rétention du PHQ

Dans le contexte des LR, le personnel hautement qualifié (PHQ) désigne les professionnels qui soutiennent et effectuent des recherches, y compris les chercheurs (notamment les chercheurs en début de carrière, les étudiants et les postdoctorants) et le personnel de soutien à la recherche (notamment les administrateurs système, les professionnels techniques de la recherche (PTR), les développeurs de logiciels). Le PHQ prend généralement en compte 2 types de personnes : les nouveaux membres du PHQ qui se perfectionnent au cours de la recherche financée ; les membres du PHQ qui font partie d'une équipe de recherche ou d'une organisation d'infrastructure existante et qui sont déjà capables de mener des activités de recherche ou de soutenir la recherche.

Parmi les PTR, le concept d'un ILR est assez nouveau, n'ayant été développé qu'au cours de la dernière décennie. La naissance du terme ILR peut être attribuée au mouvement d'ILR au Royaume-Uni, combiné à une série d'ateliers organisés par SSI qui ont mis en évidence la nécessité pour les développeurs de logiciels d'avoir quelque chose de plus fondamental qu'une simple reconnaissance de rôle. L'ILR n'est pas traditionnellement considéré comme une discipline

⁴²³ The Leadership Council. (2014). *Canadian DI Environmental Scan: A Supplement to the Background Précis Document Provided to DI Summit 2012*, 7. Consulté sur le site <http://digitalleadership.ca/wp-content/uploads/2014/01/Canadian-DI-Environmental-Scan.pdf>.

formelle, mais selon la définition britannique d'ILR⁴²⁴, c'est un profession unique en raison d'un ensemble de compétences distinctives qui associe des compétences techniques approfondies en matière de développement de logiciels à une compréhension des cultures et des environnements de recherche spécifiques.

L'efficacité et les résultats de la recherche pourraient être améliorés si les ILR étaient intégrés dans les équipes de recherche. Déléguer la responsabilité des LR à un ILR réduit également la pression exercée sur les chercheurs pour qu'ils maîtrisent un autre ensemble de compétences, mais il est difficile d'attirer et de retenir ce personnel hautement qualifié dans le milieu universitaire. Une pratique courante dans le milieu universitaire est que les chercheurs principaux encouragent les étudiants des cycles supérieurs à entreprendre le développement des LR, ce qui met en évidence l'un des parcours les plus courants pour le développement des ILR. Cette approche est facilitée par un modèle de financement qui encourage l'embauche de postdoctorants tout en décourageant le développement d'une ressource d'ILR plus professionnalisée. Deuxièmement, l'absence d'un plan de carrière officiel et durable pour les ILR rend difficile le recrutement et la rétention de ce personnel précieux. On peut faire valoir que les universités canadiennes offrent un environnement de travail positif et flexible par rapport à celui de l'industrie, mais le cheminement de carrière dans le milieu universitaire peut être moins clair étant donné que les postes d'ILR dépendent généralement du financement. Cela peut entraîner un taux de roulement élevé, et même lorsqu'ils restent, les ILR sont souvent dans des postes précaires. Prenons l'exemple d'un projet de recherche financé sur deux ans : un ILR pourrait être embauché rapidement à ce poste, mais se retrouver dans une impasse sans possibilité de reconnaissance, de récompense ou d'avancement de carrière, car la structure universitaire typique et la culture académique récompensent les publications et non les ILR. À l'inverse, l'industrie offre davantage de récompenses financières, d'esprit d'équipe, de perfectionnement des compétences et de possibilités de promotion. Par conséquent, il est difficile d'embaucher et de retenir les ILR dans le milieu universitaire, ce qui nuit à la pérennité ainsi qu'au développement proactif de la communauté des ILR. Les difficultés de recrutement des ILR dans le milieu universitaire peuvent être exacerbées par les besoins disparates des départements locaux des RH et des finances (par exemple, les contrats d'emploi institutionnels locaux, les échelles de salaire et les avantages sociaux, la longévité et les plans de carrière), ainsi que par la préférence des universités et des bailleurs de fonds pour les chercheurs « traditionnels » par rapport aux ILR, et la perte de personnel hautement qualifié au profit de l'industrie. Il faut donc mettre en place une stratégie solide de recrutement et de maintien en poste du personnel hautement qualifié, et en particulier des ILR.

Les programmes de LR de CANARIE ont marqué le début des efforts visant à développer une culture de LR au Canada en organisant des événements axés sur les LR (p. ex. le Colloque canadien sur les logiciels de recherche) et en finançant des équipes de soutien local aux logiciels de recherche (SLRL). Dans le cadre du programme SLRL, après le financement de démarrage de deux ans de CANARIE, l'équipe de l'Université McMaster⁴²⁵ a obtenu un financement directement de l'établissement ; l'Université McMaster poursuit sa collaboration avec deux autres équipes pilotes, l'une de l'Université de Regina et l'autre de l'Université Carleton. En réponse à

⁴²⁴ <https://society-rse.org/about/>

⁴²⁵ <https://research.mcmaster.ca/research-resources/local-it/rhpcs/researchsoftware/>

la forte demande, les trois équipes ont ajouté des membres⁴²⁶ (notamment des ILR, du personnel de soutien à la recherche propre au domaine, etc.) que CANARIE n'avait pas financés pendant la période de financement.

Le développement de la communauté de LR progresse lentement au Canada et la communauté de LR reste informelle. Un modèle international de rétention du personnel hautement qualifié en LR est l'University College London, un pionnier dans la constitution d'équipes d'ILR avec un soutien direct de l'établissement. Il s'agit d'un nouveau modèle d'organisation de l'expertise en LR, où les équipes de recherche emploient des ILR de façon permanente. Leur modèle consiste à financer une capacité de base et une équipe de direction, en réservant une partie des frais généraux du financement de la recherche, avec des ILR supplémentaires financés par des services rémunérés. Ce modèle permet également au groupe d'ILR de se perfectionner et de croître au fil du temps, bien que d'autres modèles bénéficiant d'un soutien national voient le jour. Par exemple, aux Pays-Bas, une approche impliquant le partage de l'expertise, du financement et de la responsabilité entre les centres DANS⁴²⁷ et SURFsara⁴²⁸ est envisagée, agissant comme un « groupe transversal qui partage l'expertise entre les organisations nationales et fait campagne pour la pérennité des logiciels à l'échelle européenne »^{429,430}.

L'importance des ILR bien qualifiés devient de plus en plus évidente tant au sein des établissements que des bailleurs de fonds (p. ex. le Software for the Future de l'EPSRC⁴³¹, le financement⁴³² de la NSF SI²), et grâce à des initiatives comme celles de CANARIE, de la SSI, de la Society of Research Software Engineering, du Hidden REF⁴³³, etc. Alors que l'IRN se dirige vers le calculateur exoflopique, nous arrivons à un point charnière dans l'utilisation efficace de la prochaine génération de technologies. Les chercheurs ont accès à de grandes quantités de données, mais ils ont besoin de l'aide et des compétences des ILR pour leur permettre d'utiliser toutes ces données. Il y aura une augmentation de la demande de compétences des ILR pour les établissements et les communautés de recherche, tant à l'échelle nationale qu'internationale.

Reconnaître les ILR comme un rôle officiel est une chose, mais financer et soutenir les parcours professionnels dans ce domaine en est une autre⁴³⁴. Les mises en œuvre nationales sont définies par de nombreux facteurs différents : la façon de structurer le financement, les organisations clés et le développement de stratégies pour le financement des ILR. Compte tenu de l'importante demande d'ILR dans tous les domaines de pratique et du fait que les ILR spécialisés dans un

⁴²⁶ <https://research.mcmaster.ca/research-resources/local-it/rhpcs/rhpcs-staff-directory/>

⁴²⁷ <https://dans.knaw.nl/en/>

⁴²⁸ <https://www.surf.nl/en/research-it>

⁴²⁹ Hettrick, S. (2016). Research software sustainability: Report on a Knowledge Exchange workshop.

⁴³⁰ <http://doi.org/10.1109/SE4Science.2019.00009>

⁴³¹ <https://epsrc.ukri.org/research/ourportfolio/themes/researchinfrastructure/subthemes/einfrastructure/software/>

⁴³² <https://www.nsf.gov/pubs/2013/nsf13511/nsf13511.htm>

⁴³³ <https://hidden-ref.org/>

⁴³⁴ <https://codeforthought.buzzsprout.com/1326658/7123159-funding-research-software-engineering>

domaine pourraient servir dans d'autres disciplines, il est nécessaire d'élaborer un modèle de financement des ILR complet qui soit inclusif pour les PTR et les ILR dans toutes les disciplines. Bien que le Canada ait une communauté d'ILR naissante dans la plupart des régions du pays, il faut mettre davantage l'accent sur l'établissement d'une communauté nationale d'ILR et d'un cheminement de carrière stable pour les ILR à l'intérieur et à l'extérieur de leur établissement.

Il s'agit d'un défi pour l'ensemble de la communauté des IRN soulignant la nécessité de travailler ensemble. Il faut que les bailleurs de fonds fassent comprendre à tous les acteurs de l'écosystème que les contributions des ILR et des PTR sont encouragées et valorisées. L'Europe a reconnu ces nouvelles « rencontres entre le secteur technique et l'université traditionnelle », et ce développement des LR bénéficie d'un soutien important. Par exemple, la France dispose d'une catégorie codifiée de professions de soutien à la recherche universitaire (notamment les « ingénieurs de recherche »⁴³⁵ liés à l'analyse de données dans toutes les disciplines⁴³⁶). Toutefois, ces postes de service public intégrés dans des instituts ou des projets de recherche universitaire sont récemment devenus des postes plus précaires en raison du financement de la recherche par projet. Bien qu'historiquement, les équipes⁴³⁷ informatiques du CNRS et les laboratoires du Portail des métiers aient eu du personnel de carrière, ces postes sont aussi devenus moins stables. Au Royaume-Uni, les ILR et les programmes de bourses connexes sont en augmentation⁴³⁸. Une entité comme le Turing Institute⁴³⁹ dispose de postes tels que DevOps et Research software, tout en tissant des liens avec d'autres partenaires de recherche. Les signataires du concordat britannique visant à soutenir le développement de la carrière des chercheurs⁴⁴⁰ constituent un autre bon exemple d'une approche destinée à améliorer l'emploi et le soutien d'un riche profil de carrières de recherche, y compris les ILR. À la suite de ces exemples internationaux, le Canada peut passer du niveau conceptuel actuel de soutien et de reconnaissance des ILR à une approche plus systématique et stratégique visant à aligner les principes et les politiques sur les mesures incitatives à la carrière des ILR.

8.2.5. Juridique/Politique

Les logiciels de recherche sont internationaux. L'harmonisation des politiques (p. ex. la gestion des licences), des normes (p. ex. les principes FAIR pour les LR⁴⁴¹) et des protocoles (p. ex. la réutilisation des LR) est essentielle à l'excellence dans le domaine des LR, tant sur le plan national qu'international. Contrairement aux bonnes pratiques en matière de GDR, où les efforts de Portage, de DRC et des communautés propres à un domaine sont déployés depuis des décennies, il reste encore beaucoup de travail fondamental à faire dans le contexte des LR. À l'heure actuelle, le Canada ne dispose pas de politiques, de normes et de protocoles bien

⁴³⁵ <https://www.enseignementsup-recherche.gouv.fr/cid23194/ingenieur-de-recherche.html>

⁴³⁶ <https://metiersit.dsi.cnrs.fr/index.php?page=cartofamille&codeBAP=D&codeFamille=A>

⁴³⁷ <https://metiersit.dsi.cnrs.fr/>

⁴³⁸ <https://www.software.ac.uk/programmes-and-events/fellowship-programme>

⁴³⁹ <https://www.turing.ac.uk/>

⁴⁴⁰ <https://www.vitae.ac.uk/policy/concordat-to-support-the-career-development-of-researchers>

⁴⁴¹ <https://sorse.github.io/programme/workshops/event-016/>

élaborés pour aider les chercheurs de toutes les disciplines à gérer les LR. Cela nuit à notre capacité de tirer parti de l'énorme potentiel des LR, mais aussi à notre compétitivité dans le cadre de nos collaborations internationales. Il faut créer un cadre stratégique national qui guidera le développement durable des LR au Canada, tout en s'assurant que le cadre canadien est compatible avec les pratiques mondiales.

D'autres régions, comme l'UE, l'Australie et les États-Unis, ont investi dans ce domaine ; le Canada peut s'inspirer de la communauté internationale des LR. Il y a quelques exemples d'efforts du Canada dans ce domaine comme la participation à l'International Research Software Engineering Leaders Workshop⁴⁴² et le fait d'être un membre fondateur de l'International Research Software Alliance⁴⁴³. Au cours des dernières années, plusieurs initiatives internationales de la communauté de LR ont vu le jour, comme le Software Source Code IG (SSC IG) de la RDA⁴⁴⁴, le Software Source Code Identification (SCID) WG⁴⁴⁵, le FAIR4RS WG⁴⁴⁶ et le Software Citation Implementation Working Group (SCIWG) de FORCE11⁴⁴⁷. En plus de donner accès à des réseaux de pairs et de contribuer à façonner la pratique internationale en LR, ces initiatives et communautés internationales offrent aussi d'importantes possibilités en termes de modèles de réussite : le Canada pourrait bénéficier des outils et des ressources qui ont déjà été développés.

8.2.6. Sécurité et confidentialité

Le succès de la recherche et de l'innovation au Canada repose avant tout sur la grande intégrité de la recherche et une entreprise de recherche ouverte et collaborative, entre les types d'organismes et les juridictions, qui favorisent la découverte et repoussent les limites de la recherche. En même temps, ce fondement d'innovation ouverte doit être équilibré par des mécanismes qui protègent les biens intellectuels, découragent le détournement des artefacts et des résultats de la recherche et assurent une gestion responsable des fonds publics. Les comportements qui violent ces principes et valeurs compromettent l'intégrité de la recherche et peuvent même présenter des risques pour la sécurité de la recherche⁴⁴⁸. De ce point de vue, la sécurité et l'intégrité de la recherche sont interdépendantes. Au cours de la dernière décennie, des individus (à l'échelle nationale et internationale) et des gouvernements ont déployé des efforts de plus en plus sophistiqués pour exploiter, menacer et compromettre la sécurité et l'intégrité de la recherche, indépendamment de la situation géographique, de la taille et des disciplines. Ces comportements incluent des défauts de divulgation (par exemple, financement, emploi, affiliations et nominations, laboratoires parallèles, conflits d'intérêts financiers), des comportements inappropriés ou de l'exploitation (par exemple, détournement de la propriété

⁴⁴² <https://researchsoftware.org/>

⁴⁴³ <https://www.researchsoft.org/>

⁴⁴⁴ <https://www.rd-alliance.org/groups/software-source-code-ig>

⁴⁴⁵ <https://www.rd-alliance.org/groups/rdaforce11-software-source-code-identification-wg>

⁴⁴⁶ <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>

⁴⁴⁷ <https://www.force11.org/group/software-citation-implementation-working-group>

⁴⁴⁸ https://www.ic.gc.ca/eic/site/063.nsf/fra/h_97955.html

intellectuelle ou d'autres droits légaux, rupture de contrat et de confidentialité), des activités contraires à l'éthique et parfois criminelles (par exemple, violation des normes éthiques et des politiques des organismes fédéraux, fraude de subventions, cyberattaques, vol de données). Ces comportements peuvent avoir des répercussions négatives importantes tant sur les chercheurs individuels que sur les organismes de recherche, et menacer l'intégrité de l'entreprise de recherche, notamment en faussant les décisions relatives à l'utilisation appropriée des fonds de recherche, en entraînant la perte de financement, en détournant des informations confidentielles ou exclusives et des données de prépublication vers des entités étrangères, en portant atteinte à la réputation, à la carrière et aux finances, et en perdant la confiance du public. Afin de relever efficacement les défis liés à la sécurité et à l'intégrité de la recherche, il faut établir une approche concertée, équilibrée, fondée sur les risques et les preuves, pour protéger les intérêts canadiens et maintenir l'écosystème d'innovation qui a contribué à soutenir le leadership mondial du Canada en matière de sciences et de technologie.

Alors que les préoccupations liées à la cybersécurité ne cessent de croître, la cybersécurité devient un élément important et amplificateur de la gestion globale des risques d'une organisation. Il est donc de plus en plus important d'élaborer une stratégie nationale solide en matière de cybersécurité qui dépasse le cadre d'un seul fournisseur de services ou d'une seule université de recherche. Les menaces de cybersécurité peuvent concerner les systèmes, les données, les personnes, les actifs et les capacités, et mettre en péril la sécurité, l'économie, la réputation, la sécurité publique et la santé. Parmi les risques identifiés figurent la protection de la propriété intellectuelle, des données sensibles et des informations personnelles, l'élaboration et la mise en œuvre de normes et de politiques en matière de cybersécurité, la formation à la sensibilisation et la protection de la vie privée. Pour contrer efficacement les menaces à la sécurité, il faut s'efforcer de comprendre et de gérer les risques liés à la cybersécurité, d'identifier et de mettre en œuvre des mesures visant à améliorer la sécurité des données, de prévenir les violations internes, d'établir des processus de réponse aux incidents et de maintenir la conformité aux exigences pertinentes. C'est pourquoi la cybersécurité est en train de devenir une considération majeure pour de nombreux développeurs de LR, qu'il s'agisse de la protection des services essentiels par le biais de garanties appropriées (par exemple, l'écriture d'un code de grande qualité conforme aux normes et aux bonnes pratiques, la gestion des identités et des accès, le renforcement des normes et de la validation cryptographiques, la promotion de plateformes fiables), de la sécurisation des technologies émergentes (par exemple, la cybersécurité de l'IdO⁴⁴⁹), de la détection et de l'identification d'événements de cybersécurité, de la réponse rapide aux incidents de cybersécurité, de la restauration et du rétablissement des capacités ou des services altérés, ou du maintien de plans de résilience.

Plus que dans d'autres domaines du développement de logiciels, le rythme rapide des nouveaux types de menaces liées à la cybersécurité signifie que la plupart des équipes de LR ne sont pas bien placées pour comprendre ces menaces ou pour y répondre. Cela est particulièrement important pour les systèmes d'infrastructure critiques comme les passerelles scientifiques, qui peuvent être particulièrement sensibles aux menaces de cybersécurité parce qu'ils sont conçus pour interagir avec des utilisateurs externes. Bien que ce problème puisse être moins important pour un chercheur écrivant des LR sur un groupe de secteurs de CIP, étant donné que

⁴⁴⁹ <https://www.nist.gov/blogs/cybersecurity-insights/rounding-your-iot-security-requirements-draft-nist-guidance-federal>

l'environnement et les aspects de sécurité sont dans de nombreux cas sous le contrôle de l'opérateur du système, et non du développeur ou de l'utilisateur de la plupart des logiciels d'application typiques, la cybersécurité doit tout de même être prise en considération. Les ILR qualifiés jouent un rôle clé à cet égard, car ils sont plus susceptibles d'être formés à l'utilisation des bonnes pratiques de sécurité dans le cadre du développement de leurs logiciels.

En réponse au besoin identifié par le milieu canadien de la recherche et de l'éducation pour une coordination et d'une harmonisation nationales des efforts en matière de cybersécurité, CANARIE a lancé en 2020 un nouveau programme national, le programme Initiatives en cybersécurité (PIC)⁴⁵⁰, pour financer des initiatives en cybersécurité qui renforceront le secteur grâce à des technologies de pointe, des processus améliorés et une expertise élargie dans l'ensemble du secteur canadien de la recherche et de l'éducation. De plus, la récente publication des pratiques recommandées pour renforcer la sécurité et l'intégrité de l'entreprise de recherche scientifique et technologique américaine⁴⁵¹ par le Science and Technology Research Enterprise (NSTC) des États-Unis⁴⁵² a donné des recommandations aux organismes de recherche et complète le protocole présidentiel sur la sécurité nationale (National Security Presidential Memorandum NSPM-33⁴⁵³) qui demandait aux départements et organismes fédéraux d'agir pour protéger la recherche financée par le gouvernement fédéral. Le cadre de cybersécurité du NIST constitue⁴⁵⁴ également une ressource utile pour aider les organismes de recherche à établir et à maintenir des mesures de cybersécurité efficaces. Le NIST élabore actuellement un cadre pour la GDR basé sur le cadre de cybersécurité.

Le contexte des données de santé — pour lequel il faut protéger la vie privée des patients en fonction de l'utilisation des données — pose des problèmes de sécurité particuliers et complexes. Les données réservées sont une occasion manquée de faire avancer la recherche dans l'intérêt de tous, mais une divulgation inappropriée des données peut compromettre les intérêts de l'individu. Les plateformes de recherche telles que Science Gateways qui traitent des données de santé nécessitent des compétences et des connaissances particulières qui incluent les bonnes pratiques de développement de logiciels sécurisés, la compréhension de la politique de confidentialité des données et la capacité de créer une solution de LR en tenant compte de ces enjeux. Par exemple, le cadre de protection de la vie privée du NIST⁴⁵⁵ fournit un ensemble d'outils⁴⁵⁶ et de procédures pour évaluer les contrôles de sécurité et de protection de la vie privée servant à protéger les informations sensibles, mais non classifiées par le gouvernement,

⁴⁵⁰ <https://www.canarie.ca/fr/cybersecurite/pic/>

⁴⁵¹ <https://trumpwhitehouse.archives.gov/wp-content/uploads/2021/01/NSTC-Research-Security-Best-Practices-Jan2021.pdf>

⁴⁵² <https://www.whitehouse.gov/ostp/nstc/>

⁴⁵³ <https://trumpwhitehouse.archives.gov/presidential-actions/presidential-memorandum-united-states-government-supported-research-development-national-security-policy/>

⁴⁵⁴ <https://www.nist.gov/cyberframework>

⁴⁵⁵ <https://www.nist.gov/privacy-framework>

⁴⁵⁶ <https://csrc.nist.gov/publications/detail/sp/800-172/final>

l'industrie et les universités à l'appui de divers programmes fédéraux. Les ILR doivent être en mesure d'intégrer de tels cadres stratégiques dans leurs solutions de LR.

Les établissements transforment la manière dont ils répondent aux risques de sécurité, car le gouvernement insiste pour que des mesures de sécurité solides soient prises, en particulier dans la recherche sur les êtres humains. Une approche traditionnelle, souvent intégrée dans les cadres politiques institutionnels, consiste à contrôler ce que les chercheurs peuvent faire en matière de logiciels, par le biais d'une politique d'achat ou de l'embauche de développeurs de logiciels. Cependant, ce type d'approche n'aura pas toujours le résultat attendu en ce qui concerne la réponse à la politique connexe. Cette approche compromet la sécurité institutionnelle et nationale, ce qui fait qu'il incombe aux différents acteurs de trouver des moyens de répondre aux besoins des chercheurs de manière proactive et efficace.

La sécurité est trop souvent érigée en un mur qui empêche les utilisateurs d'accéder aux résultats de la recherche, ce qui entraîne une tension dans l'entreprise de recherche qui doit être gérée avec soin. Grâce à la mise en œuvre des bonnes pratiques, à la sensibilisation, à la formation et à l'éducation, ainsi qu'à la gestion des risques, les résultats de la recherche peuvent être protégés de manière appropriée, tout en préservant l'approche ouverte et collaborative de la recherche, essentielle au positionnement mondial du Canada en matière de recherche et de développement.

8.2.7. Collaboration internationale

Les collaborations internationales en matière de LR sont actuellement assez difficiles. À quelques exceptions près, les fonds de collaboration provenant d'organismes de financement nationaux ne peuvent être distribués qu'aux établissements du pays en question. Par exemple, les établissements non canadiens ne sont généralement pas en mesure de recevoir directement des fonds des programmes des trois organismes, ou le pourcentage des fonds disponibles pour le PHQ international doit être justifié et peut être plafonné à un seuil très bas (p. ex. les programmes de CANARIE). Bien qu'il soit possible de financer des chercheurs internationaux à partir des fonds des trois organismes, il est difficile d'établir des collaborations officielles par le biais de ces programmes de financement. Il en va de même pour d'autres programmes internationaux. Par exemple, le programme de financement Horizon Europe ne permet généralement pas aux établissements canadiens de recevoir directement des fonds, et le financement de chercheurs à l'extérieur des pays admissibles de l'UE n'est généralement pas autorisé, sauf dans de rares circonstances. Ainsi, les collaborations internationales en matière de LR sont souvent faiblement jumelées, impliquant généralement le financement national de composantes nationales, la collaboration internationale se faisant par le biais d'ententes informelles entre les collaborateurs de recherche.

Il y a plusieurs exceptions notables à cette règle. En 2018, les agences Horizon 2020 de l'UE et IRSC ont lancé un cycle de financement international axé sur les projets de données sur la santé^{457,458}. Ce programme de financement a permis de soutenir six grands projets internationaux de LR avec des contributions substantielles d'institutions de tout le Canada⁴⁵⁹. De plus, certains

⁴⁵⁷ <https://cihr-irsc.gc.ca/f/50691.html>

⁴⁵⁸ https://cordis.europa.eu/programme/id/H2020_SC1-BHC-05-2018

⁴⁵⁹ [Projets H2020 SC1-BHC-05-2018](#)

programmes de financement nationaux permettent aux collaborateurs internationaux de recevoir un financement direct (par exemple, les NIH américains permettent aux universités canadiennes de recevoir un financement) et il y a certains programmes de financement internationaux pour les LR (par exemple, Chan-Zuckerberg Open Source for Science⁴⁶⁰).

Enfin, certains domaines sont très internationaux dans leur orientation quant aux LR, avec des collaborations internationales de longue date sur les LR. Par exemple, la communauté de la physique des particules a une histoire de collaboration internationale autour du développement des LR, les physiciens des particules canadiens étant à la fois de grands utilisateurs et des contributeurs importants à ce développement des LR.

Bien qu'il soit actuellement difficile au Canada d'établir des collaborations internationales sur le développement des LR (par le biais du financement de la recherche), il y a quelques possibilités de travailler avec les intervenants pour résoudre les problèmes et faciliter l'utilisation et l'expansion de certaines des réussites énumérées ci-dessus.

8.2.8. Mesures incitatives et indicateurs

Les mesures incitatives liées aux paramètres, au financement, aux récompenses et à la reconnaissance, à la progression de carrière et à l'évolution des normes scientifiques sont essentielles pour catalyser l'engagement dans les LR. De manière générale, les chercheurs des disciplines et des domaines de recherche allant des STIM aux domaines des sciences humaines qui reposent sur l'informatique et les données développent et réutilisent régulièrement les LR et les considèrent comme des résultats de recherche valables. Les chercheurs des autres disciplines peuvent ne pas avoir la même vision des LR en tant que résultat primaire. Tous les intervenants ont la possibilité d'adopter des politiques et des programmes qui reconnaissent le rôle important que jouent les LR dans l'atteinte des résultats de la recherche, dans toutes les disciplines. De plus, des programmes et des mesures incitatives sont nécessaires pour encourager les chercheurs et les professionnels de soutien à la recherche (p. ex. les ILR, les PTR et les gestionnaires de données) à adhérer aux bonnes pratiques de développement de logiciels pour accroître l'expertise de base en LR et pour former les professionnels de soutien à la recherche.

La saisie d'indicateurs sur la qualité, le niveau de développement et l'effort alloué à un projet spécifique de LR est un autre défi, y compris les différences entre les communautés de pratique. Dans les sciences humaines, par exemple, on hésite à citer des versions numériques d'artefacts, car on considère qu'il est plus authentique de citer des objets physiques. Ces considérations, ainsi que d'autres, suggèrent que les décideurs politiques ont la possibilité de soutenir des initiatives visant à développer des systèmes d'évaluation de la recherche (par exemple, les principes DORA utilisés par des organismes de recherche tels que le Wellcome Trust⁴⁶¹) qui récompensent les résultats des logiciels au même titre que les publications, les données et les autres résultats de la recherche, tout en veillant à ce que des réponses proactives soient apportées lorsque ces systèmes ne sont pas mis en œuvre. Les mesures traditionnelles de publication, telles que les mesures basées sur les citations, devraient être considérées comme

⁴⁶⁰ <https://chanzuckerberg.com/rfa/essential-open-source-software-for-science/>

⁴⁶¹ <https://wellcome.org/grant-funding/guidance/data-software-materials-management-and-sharing-policy>

complémentaires aux mesures émergentes (par exemple, Altmetrics⁴⁶², GitHub forks, commit, couverture médiatique) afin de mesurer l'impact des LR de manière appropriée pour tous les acteurs de l'écosystème de la recherche. La promotion et la création de mesures appropriées pour l'activité de LR faciliteraient le changement de culture nécessaire pour réaliser la promesse de la science ouverte.

8.2.9. Culture de la recherche

Un effort intellectuel et financier considérable a été investi dans le développement d'infrastructures, d'outils et de bonnes pratiques qui permettent la science ouverte. Malgré ces investissements, ces efforts ne font toujours pas partie de la routine de la plupart des chercheurs. L'incidence des pratiques de science ouverte sur les logiciels et la reproductibilité exige des connaissances et des compétences substantielles de la part de l'équipe de développement des logiciels. Par exemple, la mise à disposition de LR nécessite un ensemble de compétences particulières, différentes du contexte de l'ensemble de données. Il s'agit là d'une occasion de collaboration entre les organismes de financement et d'autres intervenants pour créer de nouvelles mesures d'incitation et d'évaluation permettant de trouver un équilibre entre la science ouverte et la culture traditionnelle de publication pour ne pas périr.

Dans le contexte de la recherche, la PI (données et produits de recherche tels que les publications savantes ou les œuvres créatives) est la monnaie de la réussite universitaire de même que le soutien à sa production tel que les subventions ; toutefois, la PI prend des formes très différentes selon les disciplines. Traditionnellement, les données ne sont pas largement partagées, car elles ne sont pas protégées par le droit d'auteur, et le contrôle de leur distribution a permis aux chercheurs de tirer parti de la valeur de ces actifs. Il y a également des cas, par exemple les données du recensement de Canadian Peoples⁴⁶³ pour lequel un ensemble de données ne peut pas être entièrement ouvert, mais peut être partagé si le soutien personnel adéquat est disponible. Inversement, l'abandon involontaire par les auteurs universitaires des droits d'auteur aux éditeurs entrave le partage et la réutilisation des publications de recherche en tant que données ouvertes. Il faut donc développer une culture dans laquelle le partage de l'information est explicitement discuté et encouragé ; ce processus est mieux encouragé par les communautés de recherche individuelles, puisque les normes varient selon la discipline. Les LR ont un rôle à jouer pour soutenir ce changement de culture.

Pour exploiter véritablement les nouvelles IRN, il faut accorder une importance égale aux facteurs technologiques et culturels afin de faire progresser la recherche qui est essentielle à l'établissement et au maintien de la prospérité économique et sociale du Canada. Pour atteindre cet objectif, il faut élaborer et maintenir un ensemble de bonnes pratiques menant à un accès libre et sans obstacle aux LR. Le passage à une culture des LR axée sur la science ouverte rend les processus scientifiques plus efficaces, plus transparents et plus réactifs aux défis sociétaux, et rend les connaissances scientifiques plus facilement accessibles. Les chercheurs négligent souvent la contribution fondamentale que les logiciels apportent à la fiabilité et à la reproductibilité de leurs résultats scientifiques ; les éditeurs négligent la nécessité d'identifier les logiciels comme un élément essentiel du processus de publication ; les bailleurs de fonds négligent la nécessité

⁴⁶² <https://www.altmetric.com/>

⁴⁶³ <https://thecanadianpeoples.com/>

de dégager des fonds pour maintenir les logiciels de manière durable et négligent le besoin croissant de faire appel à des experts en logiciels pour les projets de recherche. Les établissements et les organisations de recherche négligent la nécessité de développer l'expertise en logiciels dans les équipes de recherche ; les décideurs politiques négligent l'importance des logiciels pour la recherche en omettant l'application de mesures d'impact appropriées. Une campagne soutenue de sensibilisation est nécessaire auprès de tous les chercheurs et tous les acteurs afin de convaincre les intervenants de la recherche que les logiciels sont un produit de recherche précieux en adéquation avec l'investissement qu'ils reçoivent et la recherche qu'ils facilitent.

Enfin, du point de vue culturel, les chercheurs actifs peuvent ne pas connaître d'autres personnes, dans leur discipline ou ailleurs, qui entreprennent des efforts similaires ou complémentaires dans le contexte des LR. De nombreuses raisons peuvent expliquer cette situation, mais le manque de reconnaissance des LR en tant que travail savant, l'écart de compréhension entre le développement des LR et les besoins en recherche, ou le manque de financement pour les ILR et le personnel connexe, et le manque d'initiatives nationales de la communauté des LR sont des exemples qui peuvent avoir un impact important. Des initiatives comme International Research Software Engineers⁴⁶⁴ et le colloque sur les LR de CANARIE, ainsi que leur programme SLRL, offrent des possibilités de croissance et d'expansion de la communauté d'ILR au Canada.

8.2.10. Politique et stratégie du gouvernement

Les organismes de recherche du gouvernement fédéral et des provinces (p. ex., le CNR et NAPHRO) ont l'occasion d'élaborer une stratégie pour travailler davantage en collaboration avec leurs partenaires universitaires et les autres gouvernements. Le grand nombre d'utilisateurs et de communautés, la complexité entourant la clarté de leurs rôles et responsabilités (à l'échelle internationale, nationale, régionale et locale), le nombre de régions impliquées et la diversité des exigences avec lesquelles ils doivent aligner leurs efforts, soulignent les défis de cette tâche. Bien que l'on assiste à l'émergence d'un leadership national concerté en matière de LR, les efforts visant à cristalliser les communautés de LR n'ont pas été financés de manière adéquate ou mandatés de manière formelle. Ces deux facteurs ont rendu difficile la planification stratégique et concertée. Si l'on ajoute à cela le manque de coordination des investissements dans le contexte des LR, il est difficile d'élaborer les politiques, processus, protocoles, bonnes pratiques et normes partagés qui sont si essentiels. Il y a des modèles qui peuvent servir de base à cette coordination ; le « EU Open Science Cloud » est un exemple de gouvernance et d'ingénierie permettant d'intégrer les LR et les données de plusieurs pays dans un cadre cohérent⁴⁶⁵.

8.2.11. EDI et représentation

Comme nous l'avons vu précédemment, le développement d'une approche efficace quant à l'EDI peut être problématique dans le contexte des logiciels de recherche. Malgré une base solide de programmes d'informatique au Canada, il n'y a pas de programmes de formation spécifiques en génie de LR au Canada, ce qui conduit à un manque de compétences en développement de LR dans les activités de recherche, mais aussi à un manque d'opportunités de promotion pour les

⁴⁶⁴ <https://researchsoftware.org/>

⁴⁶⁵ <https://eosc-portal.eu/>

ILR. Le manque d'EDI dans le contexte des logiciels de recherche découle de divers facteurs, y compris une représentation significativement plus faible des femmes et des personnes appartenant à un groupe racialisé dans les programmes de sciences informatiques et dans les carrières liées aux logiciels, ainsi que le rétrécissement général de cheminement vers la titularisation pour les chercheuses⁴⁶⁶. Cela indique que nous devons créer une main-d'œuvre diversifiée en sensibilisant les divers intervenants (individus, organismes de recherche et décideurs) aux défis et aux possibilités liés à l'EDI au profit de la profession de LR et en créant ou en faisant progresser des programmes visant à élargir les possibilités et à éliminer les obstacles pour les communautés sous-représentées intéressées par les LR⁴⁶⁷. Certains efforts ont été entrepris dans ce domaine, mais ils ne suffisent pas : le groupe de travail sur la diversité, l'équité et l'inclusion (DEI-WG) de l'US-RSE a été formé et travaille activement à l'organisation d'initiatives DEI et de sujets connexes, à l'engagement avec des groupes technologiques DEI tels que Women in HPC⁴⁶⁸ et Code 2040⁴⁶⁹, ainsi qu'à l'élaboration de conseils sur la façon de rédiger des descriptions de poste d'ILR inclusives afin de s'assurer qu'elles attirent un public diversifié.

Un enjeu plus récent, mais connexe est de savoir si la science ouverte et la science reproductible sont suffisamment inclusives pour certains membres de la communauté qui pourraient être lésés par le contrôle ou la protection des chaînes d'outils⁴⁷⁰. Cet aspect de la sécurité aborde la question du contrôle des données et du code en tant que facteur de différenciation du pouvoir. Ainsi, il faudrait remettre en question certains aspects de la culture de la recherche et du risque, qui peuvent être partiellement résolus en créant des systèmes et des flux de travail dans les communautés ou les équipes de recherche qui protègent les populations vulnérables et s'attaquent au patriarcat, au colonialisme et au racisme systémiques, ce qui permet d'intégrer les dimensions d'équité, de diversité et d'inclusion (EDI) dans le contexte des LR. L'orientation sur les bonnes pratiques de reproductibilité et de science ouverte n'est qu'un moyen de faciliter les progrès et l'EDI et les avantages pour les stagiaires et les jeunes chercheurs.

9 Prochaines étapes

Pour être prêt à relever les défis actuels et à saisir les opportunités potentielles, il faut commencer par une analyse approfondie de la situation et du paysage actuels en fonction des objectifs à court et à long terme. Comme ce rapport résume l'état actuel du paysage des LR au Canada afin de soutenir le développement d'une compréhension commune parmi les membres de l'Alliance de l'ampleur et de la complexité de l'engagement des intervenants dans ce domaine, il sert également de base à partir de laquelle l'Alliance peut établir une voie à suivre pour la stratégie nationale des LR au Canada. Les conclusions et les observations contenues dans ce document, ainsi que dans les publications sur la situation de la GDR et du CIP, ont pour but d'aider à soutenir

⁴⁶⁶ <https://www.nplusonemag.com/issue-34/essays/sexism-in-the-academy/#fn1-11214>

⁴⁶⁷ <https://developerrelations.com/dev-rel/mentored-sprints-for-diverse-beginners>

⁴⁶⁸ <https://womeninhpc.org/>

⁴⁶⁹ <http://www.code2040.org/>

⁴⁷⁰ <https://en.wikipedia.org/wiki/Toolchain>

le processus d'évaluation des besoins des chercheurs, le nouveau modèle de prestation de services de l'Alliance et le développement du modèle de financement, ainsi que les processus du plan stratégique pour l'IRN.

9.1. Évaluation des besoins des chercheurs

L'Alliance a consulté la communauté de recherche au Canada afin d'évaluer les outils, les services et le soutien actuels en matière d'IRN ; d'identifier et d'aborder l'état futur idéal de l'IRN au Canada ; et de déterminer comment l'Alliance pourrait atteindre cet état du point de vue du CIP, des LR et de la GDR, respectivement. Le processus de consultation complet comprend quatre étapes : Étape 1 : Appel de livres blancs⁴⁷¹ sur l'écosystème futur d'IRN au Canada (terminé) ; Étape 2 : Appel de documentation actuelle (terminé) ; Étape 3 : Sondage en ligne (terminé) ; Étape 4 : Discussions en assemblée générale virtuelle (terminé). Le Conseil des chercheurs⁴⁷² supervise le processus d'évaluation des besoins des chercheurs qui fera appel à de nombreuses communautés de chercheurs et d'associations disciplinaires et fournira des conseils indépendants à l'Alliance sur des questions liées à la prestation de services et de programmes pour la communauté des chercheurs.

À la suite du processus d'évaluation des besoins, les résultats seront intégrés aux conclusions des rapports sur l'état actuel du CIP, des LR et du GDR qui fourniront des recommandations pour soutenir l'effort de planification stratégique de l'Alliance.

9.2. Élaboration d'un modèle national de prestation de services et d'un modèle de financement

En collaboration et en consultation avec Deloitte et les partenaires d'IRN, l'Alliance affinera un nouveau modèle de prestation de services qui définira les services nationaux, régionaux et locaux pour la GDR, le CIP et les LR, y compris les niveaux de service attendus, les nouveaux modèles de financement, les feuilles de route de transition, ainsi que les rôles et responsabilités des intervenants.

9.3. Plan stratégique de l'IRN pour 2022-24

L'Alliance présentera une stratégie et une vision nationales pour le CIP, les LR et la GDR en intégrant les résultats de ses activités d'évaluation et de sensibilisation. Le plan stratégique comprendra une feuille de route pour transformer l'écosystème actuel dans lequel le CIP, les LR et la GDR sont traités comme des entités distinctes, en un état futur souhaité plus intégré où la recherche est soutenue par des IRN solides tout au long de son cycle de vie.

⁴⁷¹ <https://engagedri.ca/fr/%C3%A9valuation-canadienne-des-besoins-en-mati%C3%A8re-de-recherche-num%C3%A9rique/soumissions-expos%C3%A9s-de-position>

⁴⁷² <https://engagedri.ca/fr/conseil-des-chercheurs>

Acronymes et glossaire

Full name	Acronym	Full name	Acronym
Leadership Council for Digital Research Infrastructure	LCDRI	Research Software Engineers	RSEs
digital research infrastructure	DRI	Canadian Research Software Conference	CRSC
data management	DM	Porous Materials Engineering and Analysis Lab	PMEAL
advanced research computing	ARC	McGill Centre for Integrative Neuroscience	MCIN
Research Software	RS	Engineering and Physical Sciences Research Council	EPSRC
New Digital Research Infrastructure Organization	NDRIO	Common Lab Research Infrastructure for the Arts and Humanities	CLARIAH
Working Group on Research Software	RSWG	Vice-Presidents Research	VPR
Pacific Institute for the Mathematical Sciences	PIMS	Information technology departments	ITD
Software	SW	Chief Information Officers	CIO
Earth Science Division	ESD	National Alliance of Research Health Funding Organizations	NAPHRO
Earth Science Data Systems	ESDS	Compute Canada Federation	CCF
Interactive development environments	IDEs	Canada Foundation for Innovation	CFI
Department of Energy	DOE	Canadian Institutes of Health Research	CIHR
Office of Scientific and Technical Information	OSTI	Natural Sciences and Engineering Research Council	NSERC
free and open-source software	FOSS	Social Sciences and Humanities Research Council	SSHRC
intellectual property	IP	open-source Program Offices	OSPOs
National Initiatives for Open Science in Europe	NI4OS Europe	Ontario Centres of Excellence	OCE
Contribution License Agreement	CLA	International Research Software Alliance	ReSA
HPC Container Maker	HPCCM	Research Data Alliance	RDA
Astrophysics Source Code Library	ASCL	Journal of Open Research Software	JORS
International Conference on Software Engineering	ICSE	Journal of open-source Software	JOSS
Highly Qualified Personnel	HQP	open access	OA

National Research and Education Network	NREN	Canada Research Continuity Emergency Fund	CRCEF
Software Sustainability Institute	SSI	Canada Emergency Wage Subsidy	CEWS
Research Data Management	RDM	Canadian International Innovation Program	CIIP
National Data Services Framework	NDSF	National Research Council of Canada Industrial Research Assistance Program	NRC IRAP
Canadian University Council of Chief Information Officers	CUCCIO	Partnership Development Activities	PDA's
Canadian Higher Education Information Technology	CANHEIT	Requests for Proposals	RFP's
Research & High Performance Computing	RHPC	small and medium-sized enterprises	SME's
Canadian Advanced Network for Astronomical Research	CANFAR	Canadian Statistical Sciences Institute	CANSSI
National Research Council Canada	NRC	Smart Prosperity Institute	SPI
Indigenous Knowledge Social Network	SIKU	Michael Smith Foundation for Health Research	MSFHR
New Frontiers in Research Fund	NFRF	Canadian Partnership Against Cancer	CPAC
Canada Research Coordinating Committee	CRCC	Dystonia Medical Research Foundation Canada	DMRFC
Tri-agency Institutional Programs Secretariat	TIPS	Graham Boeckh Foundation	GBF
Canadian Space Agency	CSA	Integrated Youth Services	IYS
Flights and Fieldwork for the Advancement of Science and Technology	FAST	International Alliance of Mental Health Research Funders	IAMHRF
Space Technology Development Program	STDP	Global Affairs Canada	GAC
announcement of opportunity	AO	Grand Challenges Canada	GCC
research and development	R&D	Bill and Melinda Gates Foundation	BMGF
Innovation, Science and Economic Development Canada	ISED	Declaration on Research Assessment	DORA
Innovative Solutions Canada	ISC	Ministry of Research, Innovation and Science	MRIS
European Space Agency	ESA	Ontario Research Fund – Research Excellence	ORF-RE
Canada First Research Excellence Fund	CFREF	Ontario Institute for Cancer Research	OICR
Canada Research Chairs Program	CRCP	Scholarly Infrastructures for Research Software	SIRS
Canada Excellence Research Chairs	CREC	FAIR For Research Software Working Group	FAIR4RS WG
Networks of Centres of Excellence	NCE	General Data Protection Regulation	GDPR

transparency, reproducibility, ethics, and effectiveness	TREE	Partnership Grant	PG
infrastructure as a service	IaaS	Biotechnology and Biological Sciences Research Council	BBSRC
platform as a service	PaaS	Bioinformatics and Biological Resources	BBR
software as a service	SaaS	Java Runtime Environment	JRE
Infrastructure as Code	IaC	research technical professionals	RTPs
Internet of Things	IoT	local research software support	LRSS
Quality of Service	QoS	RDA's Software Source Code IG	RDA SSC IG
Edge Computing	EC	Software Source Code Identification	SCID
Virtual Machine	VM	FORCE11 Software Citation Implementation Working Group	FORCE11 SCIWG
Software-Defined Networking	SDN	Cybersecurity Initiatives Program	CIP
Network Function Virtualization	NFV	US National Science and Technology Council	NSTC
Quantum computing	QC	National Security Presidential Memorandum 33	NSPM-33
information and communication technology	ICT	Equity, diversity and inclusivity	EDI
Master of Engineering Leadership	MEL	Diversity, Equity, and Inclusion Working Group	DEI-WG
Centre de recherches mathématiques	CRM		
Atlantic Association for Research in the Mathematical Sciences	AARMS		
John R. Evans Leaders Fund	JELF		
humanities and social sciences	HSS		
The Council on Library Resources	CLIR		
Protein Industries	PIC		
CCF central database	CCDB		

Ce tableau d'acronymes et de glossaire est destiné à aider les lecteurs à comprendre la signification de certains termes utilisés dans le présent document.

Annexe A : Exemples de portails scientifiques au Canada

Portail scientifique	Portée nationale ou internationale	Contenu général ou spécifique à une discipline	Discipline scientifique	catalogue de services
<u>2i2c</u>	internationale	général	interdisciplinaire	outils informatiques; traitement et analyse; formation et soutien
<u>ATLAS-Canada</u>	internationale	spécifique à une discipline		calcul; gestion des données; traitement et analyse; stockage
<u>CANFAR</u>	nationale	spécifique à une discipline	sciences naturelles /sciences physiques/ astronomie	
<u>CBRAIN</u>	nationale	spécifique à une discipline	activités de soutien/ développement de l'infrastructure; sciences médicales et de la santé/informatique biomédicale; interdisciplinaire	gestion des données, calcul, sécurité et opérations, traitement et analyse, stockage, formation et soutien
<u>CWRC</u>	internationale	spécifique à une discipline	sciences sociales et humaines	traitement et analyse; formation et soutien
<u>Érudit</u>	internationale	spécifique à une discipline	activités de soutien/ archivage numérique	gestion des données, sécurité et opérations, partage et découverte, stockage
<u>GENAP</u>	nationale	spécifique à une discipline	sciences médicales et de la santé; Informatique biomédicale; génomique	gestion des données; réseautique; traitement et analyse; sécurité et opérations; partage et découverte; stockage; formation et soutien
<u>iReceptor</u>	internationale	spécifique à une discipline	sciences de la santé; Immunologie	calcul; gestion des données; traitement et analyse; partage et découverte; stockage; formation et soutien
<u>MERIDIAN</u>	internationale	spécifique à une discipline	sciences naturelles/ acoustique sous-marine	gestion des données, calcul, sécurité et opérations, traitement et analyse, partage et découverte, stockage, formation et soutien
<u>Oceans 2.0</u>	internationale	spécifique à une discipline, océanographie	sciences naturelles / sciences de la Terre et de l'environnement; gestion des océans; prévention des catastrophes et protection de l'environnement; océanographies physique, chimique, biologique et géologique	Gestion des données; traitement et analyse; sécurité et opérations; partage et découverte; stockage

PanGeo	internationale	spécifique à une discipline	sciences de la Terre/ géoscience	outils informatiques; traitement et analyse; stockage; formation et soutien
PAWS	internationale	général	interdisciplinaire	outils informatiques; traitement et analyse
SNOLab	nationale et internationale	spécifique à une discipline	sciences naturelles/ sciences physiques/ astroparticule	traitement et analyse
Syzygy	nationale	général	interdisciplinaire	calcul, traitement et analyse, partage et découverte, formation et soutien
Voyant Tools	internationale	général	interdisciplinaire, analyse de texte	partage et découverte; traitement et analyse; formation et soutien

Le [tableau A](#) présente plusieurs exemples de passerelles scientifiques. D'autres exemples se trouvent dans le Registre des plateformes logicielles de recherche de CANARIE⁴⁷³.

Annexe B : Soutien et services de LR fournis dans le cadre des activités de recherche quotidiennes

⁴⁷³ <https://science.canarie.ca/researchsoftware/platforms/list/main.html>

Cycle de vie de la recherche							Données	Calcul	Logiciels	
Planifier	Créer	Traiter	Analyser	Diffuser	Conserver	Réutiliser	Outils de plan de gestion des données (PGD)	Performance du système et outils de surveillance	Outils autres que CIP/la	Outils de portail scientifique
découvrir et générer une idée	concevoir, créer et développer le LR			Accorder la permission aux contributeurs de lancer le LR			outils de plan de gestion des données	langages	transfert	intégration
conceptualiser et planifier la gestion du LR (à long terme)	lancer/utiliser/personnaliser le LR			considérer (de nouveau) de choisir une licence de LR		Acquérir, recueillir, recevoir et assembler les ressources : LR et dépendances, données connexes, publication et documentation, flux des travaux de recherche (computational), etc.	générer/produire/créer / simuler les données et les métadonnées connexes		SE	intergiciels
définir, acquérir, recueillir, recevoir les ressources (fonds et personnes) et les assembler en équipes et en communautés	décrire/documenter le LR (p.ex. fichier <i>Lisez-moi</i>)			évaluer et sélectionner le LR qui sera lancé		Installer l'environnement d'exécution du LR (p.ex., conteneurs, carnets, environnement de reproductibilité infonuagique)	collecter les données	bibliothèques	analyse	chaînes de compilation
	contrôler les versions			préparer les métadonnées, la documentation et les instructions pour le LR (utilisation)		Reproduire, vérifier et valider les résultats afin de protéger l'intégrité scientifique	nettoyer/traiter/analyser (modélisation, simulation, analyse/visualisation) / manipuler/transférer les données	scripts	virtualisation de bureau	technologies Web
	mettre à l'essai, déboguer et intégrer le LR			organiser et stocker / sauvegarder le LR		Intégrer le LR dans de nouvelles recherches	querelles et visualisation de données	pipelines	Visualisation	interfaces API
	organiser et stocker/sauvegarder le LR			contrôler les versions			assurance/contrôle de la qualité	cadres/progiciels d'envergure de modélisation/simulation numérique (modèle + données)	prétraitement	

Activités, soutien et services de LR

Cycle de vie de la recherche							Données	Calcul	Logiciels	
Plan	Create	Process	Analyze	Disseminate	Preserve	Reuse	RDM Tools	System performance and monitoring tools	non-ARC/non-RDM Tools	Science Gateways Tools
	distinguer les apports au logiciel et les apports de ce dernier			choisir une plateforme de publication pour le LR / enregistrer le LR dans un registre	choisir une plateforme fiable de longue durée pour le LR / enregistrer le LR dans un registre		décrire, documenter et interpréter les données	cadres	post-traitement	
	Respecter les licences de logiciel en vigueur / choisir une licence de LR pour le logiciel nouvellement développé			Ingérer/déposer/partager/libérer le LR	ingérer/déposer le LR pour conservation		intégrer les données	transfert de données		
	assurer la maintenance et la mise à l'essai en continu/déboguer le LR						organiser les données	organiser le stockage		
							stocker et sauvegarder les données (outils et plateformes)	équipement		
							contrôler les versions	gestion, traitement, calcul et analyse des données avancés et en continu		
							partager/publier/consulter les données dans des dépôts/registres/plateformes	prétraitement/post-traitement		
							données FAIR	Visualisation		
							réutiliser les données	SE		
							reproduire des recherches	systèmes de fichiers		
							accéder aux données	nuages (de base)		
							découvrir/rechercher (de nouveau) les données (outils et plateformes)	programmeur		

Activités, soutien et services de LR

Cycle de vie de la recherche										Logiciels		
Plan	Create	Process	Analyze	Disseminate	Preserve	Reuse	Données	Calcul	Logiciels	non-ARC/non-RDM Tools	Science Gateways Tools	
							RDM Tools	System performance and monitoring tools				
							protection/conservation / archivage des données (sélection, évaluation, conversion, etc. des données)	infrastructure de bureau virtuel (VDI)				
							anonymiser les données sensibles	intégrer un intergiciel				
							sécuriser et protéger les données (outils et plateformes)	outils de performance et de surveillance de système				
							systèmes de fichiers, SE	outils de test de performance et d'extensibilité				
							transférer les données	outils de débogage				
							mesurer l'incidence de la recherche (consultations, téléchargements, citations, altmètrie, etc.)	outils de gestion de systèmes en grappes				
							faire migrer/transformer des données originales dans des formats différents au moyen de supports appropriés	dépôts de logiciels				
							numériser					
							portails scientifiques					
							automatiser : flux des travaux réutilisables (p. ex. YesWorkflow, Taverna, Kepler)					
Activités, soutien et services de LR												

Annexe C : Archétypes d'utilisateurs et de communautés dans l'écosystème RS

Profil type/rôle professionnel	Description	Types connexes, rôles ou exemples
chercheur	personne qui fait de la recherche	Professeurs, étudiants, postdoctorants, chercheurs non universitaires (p. e. chargés de cours, chercheurs scientifiques et en laboratoire, associés / moniteurs / documentalistes/personnel de recherche) dans le secteur universitaire; chercheurs en R&D dans le secteur non universitaire
bibliothécaire	personne qui travaille dans une bibliothèque, donnant accès à l'information, aux instructions et à la formation documentaire	Experts en métadonnées, conservateurs/archivistes, bibliothécaires spécialisés, catalogueraphes, soutien/AQ, bibliothécaires de copyright
administrateur système	personne chargée de planifier, exécuter, superviser, surveiller, soutenir, de faire tourner, de maintenir et de configurer les systèmes informatiques, les ordinateurs simples ou multiples ou les serveurs.	opérateurs de portails, soutien logiciels/applications, personnel de fonctions, ingénieurs de systèmes, soutien de la sécurité informatique, personnel des licences/droits d'auteur
développeur	personne qui rédige le code, les logiciels ou les programmes informatiques	développeurs professionnels (p. ex. ingénieurs en logiciels de recherche), analystes, développeurs autodidactes/formés (p. ex. chercheurs, postdoctorants, étudiants diplômés), communautés de code source libre (p.ex. Numpy, Pandas, Jupyter); développeurs de logiciels (ingénieurs en logiciels, analystes, etc.) pour la recherche du secteur non universitaire
groupe de recherche	groupe de chercheurs et de chercheuses qui collaborent sur différents aspects d'un projet (planification, exécution, supervision, éducation, etc.)	gestionnaires de projet, formation/soutien des utilisateurs, gestionnaires de bourses
fournisseur commercial partenaire	entité ou entreprise commerciale membre d'une alliance fournissant des services ou des biens	développeurs commerciaux, en R&D; fournisseurs de données; internes, universitaires affiliés, participants ONG
soutien de la communauté	aide offerte ou obtenue d'une communauté spécialisée	listes de diffusion, forums, communautés de code source libre (Numpy, Pandas, Jupyter, ...)
recherche universitaire et non universitaire	alliance entre des entités universitaires et non universitaires	stagiaires, universitaires affiliés, participants ONG, scientifiques/chercheurs citoyens
 bailleur de fonds	organismes, sociétés et fondations de financement	bailleurs de fonds des secteurs public et privé (p.ex. Wellcome Trust)
revue et éditeur	personne, périodique, entreprise ou organisation qui prépare, distribue et publie des recherches universitaires	Elsevier, Journal of Open Research Software (JORS), Journal of open-source Software (JOSS), Software Impacts, SoftwareX
autre partie prenante	profil type autre que ceux qui sont listés	gouvernement, milieu de l'enseignement, industrie et commerce

Annexe D : Types de LR

D'autres façons de catégoriser les LR sont basées sur ses caractéristiques et ses particularités. Par exemple, les LR peuvent être classés selon les catégories suivantes :

- Par fonctions : par exemple, logiciel d'application, logiciel système, outils de programmation, service logiciel, plateforme de recherche.
- Par couches⁴⁷⁴ : par exemple, plateforme et infrastructure, outils et bibliothèques
- Par statut de licence : en libre accès, protégé, hybride
- Par statut de publication : publié, non publié
- Par statut d'instanciation : preuve de concept, produit opérationnel fonctionnel réel
- Par coût : par exemple, gratuit, commercial
- Par le mécanisme de distribution du logiciel : par exemple, code source, exécutable binaire, paquet, conteneur, image de machine virtuelle, service
- Par disciplines : général, spécifique au domaine
- Par les méthodes de recherche : quantitative, qualitative
- Par les objectifs des processus du cycle de vie de la recherche : planification, analyse, calcul, visualisation, transfert, stockage, publication, curation et préservation, découverte.

En pratique, un élément de LR peut ne pas correspondre à un seul type. Dans ce cas, les LR ont des types hybrides. Par exemple, les LR peuvent avoir à la fois de composants à code source ouvert et à code source fermé, ou même d'éléments publiés et non publiés. Un autre exemple est celui où il y a un code source et autre chose, par exemple un service ou un exécutable qui est construit à partir de ce code source, mais pas par l'utilisateur. Dans un cas comme dans l'autre, il est important de pouvoir identifier chaque fichier ou composant de manière unique et indépendante afin de s'assurer que les métadonnées et le type de logiciel s'appliquent à chaque composant : licence, version, créateurs, etc. Il pourrait également y avoir un hybride de LR et d'autres objets de recherche, comme un carnet de notes Jupyter ou un flux de travail ou un script de calcul qui sont publiés dans des objets de type mixte connus sous le nom d'objets de recherche⁴⁷⁵ ou de compendiums de recherche⁴⁷⁶ qui représentent des pipelines de calcul avec des informations de provenance associées.

⁴⁷⁴ <https://hal.archives-ouvertes.fr/hal-02117588/document>

⁴⁷⁵ <https://zenodo.org/communities/ro/search?page=1&size=20>

⁴⁷⁶ <https://biostats.bepress.com/bioconductor/paper2/>

Annexe E : Schémas de métadonnées de LR représentatifs

Format de fichier des citations (FFC)

Le format de fichier de citation ⁴⁷⁷(FFC) est un format de fichier en texte brut lisible par l'homme et la machine (CITATION.cff) en YAML⁴⁷⁸ qui fournit des métadonnées de citation pour les logiciels. Les développeurs de code peuvent inclure le fichier de FFC dans leurs dépôts pour indiquer aux autres comment citer correctement leurs logiciels. Le fichier FFC est actuellement pris en charge par GitHub⁴⁷⁹, Zenodo⁴⁸⁰, et Zotero⁴⁸¹, etc. Le fichier de FFC est un format simple avec lequel les auteurs de code peuvent enregistrer des métadonnées qui peuvent ensuite être traduites dans d'autres formats ou schémas.

Initiative CodeMeta

L'objectif du projet CodeMeta est de créer un vocabulaire de concepts, c'est-à-dire un schéma minimal de métadonnées pour les logiciels et les codes universitaires, qui peuvent servir à normaliser l'échange de métadonnées des logiciels entre les plateformes (par exemple, les dépôts) et les organisations. CodeMeta a utilisé la passerelle⁴⁸² pour générer un ensemble de concepts de métadonnées de logiciels (CodeMeta-2.0⁴⁸³ et CodeMeta-1.0⁴⁸⁴), organisés respectivement en JSON-LD et XML. CodeMeta travaille actuellement à la mise à jour de schema.org afin que tous les termes de CodeMeta y soient inclus. CodeMeta pourrait devenir une communauté qui soutient les logiciels de recherche au sein de schema.org, plutôt qu'un schéma indépendant.

Schéma des métadonnées de DataCite

Le schéma de métadonnées DataCite est une liste de propriétés de métadonnées de base choisies pour l'identification complète et cohérente d'une ressource à des fins de citation et de

⁴⁷⁷ <https://github.com/citation-file-format/citation-file-format>

⁴⁷⁸ <https://yaml.org/spec/1.2.2/>

⁴⁷⁹ <https://docs.github.com/en/repositories/managing-your-repositorys-settings-and-features/customizing-your-repository/about-citation-files>

⁴⁸⁰ https://twitter.com/ZENODO_ORG/status/1420357001490706442

⁴⁸¹ <https://twitter.com/zotero/status/1420515377390530560>

⁴⁸² <https://github.com/codemeta/codemeta/blob/master/crosswalk.csv>

⁴⁸³ <https://doi.org/10.5063/schema/codemeta-2.0>

⁴⁸⁴ <https://doi.org/10.5063/schema/codemeta-1.0>

récupération, ainsi que des instructions d'utilisation recommandées. La dernière version est Metadata Schema 4.3^{485,486} publiée en 2019.

Initiative DOE CODE

L'Office of Scientific and Technical Information (OSTI) du département de l'énergie (DOE) a travaillé sur un projet appelé DOE Code Metadata Model ⁴⁸⁷ qui définissait les métadonnées utilisées pour les logiciels.

Directives OpenAIRE pour les métadonnées de LR

Le projet OpenAIRE a créé des lignes directrices sur les logiciels pour définir et mettre en œuvre des politiques locales de gestion des logiciels en exposant les métadonnées des logiciels et en les rendant citables. OpenAIRE a créé des passerelles et des cartographies⁴⁸⁸ à partir d'autres initiatives pour rendre les dépôts OpenAIRE conformes à tous les autres en termes de découverte et de citation des logiciels.

Annexe F : Bonnes pratiques liées aux LR

L'utilisation du terme « bonnes pratiques » a des interprétations différentes selon les communautés. Un article récent intitulé « Good Enough Practices in Scientific Computing » présente un ensemble de bonnes pratiques informatiques que chaque chercheur doit adopter, quel que soit son niveau de compétence en informatique⁴⁸⁹.

F1 : Bonnes pratiques de gestion des LR

Les principes de base pour la gestion des LR sont les principes FAIR (FAIR4RS) qui ont été dérivés des principes FAIR pour les données et qui sont adaptés et adoptés pour les objets de recherche autres que les données de recherche. Dans le cadre de FAIR4RS, les considérations portent sur la manière dont les principes originaux pourraient être adaptés au cas des LR, sur la manière de promouvoir et de mesurer l'adoption de FAIR4RS, et sur la manière de soutenir la mise en œuvre.

⁴⁸⁵ <https://schema.datacite.org/meta/kernel-4.3/metadata.xsd>

⁴⁸⁶ <https://doi.org/10.14454/f2wp-s162>

⁴⁸⁷ <https://github.com/doecode/software-metadata/blob/master/XMLSchema/MetadataSchema.xsd>

⁴⁸⁸ https://docs.google.com/spreadsheets/d/1mKs-Pg_JuLcpqEkQqISCs2gGC7nEEbxdTbloGcU6NI/edit#gid=0

⁴⁸⁹ <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005510>

Tout comme les DMP, les plans de gestion de logiciel⁴⁹⁰ (PGL) sont fortement recommandés par certains organismes de financement^{491,492}. Ils définissent et formalisent un ensemble de structures et d'objectifs permettant de comprendre les LR pour ce qui est des modalités de rôles et de responsabilités tout au long du cycle de vie des LR, de la phase de développement à la prise en compte de la pérennité. Certaines lignes directrices et certains modèles^{493,494,495} ont été élaborés pour fournir les éléments de base des PGL, garantissant que les LR soient accessibles et réutilisables à court, moyen et long terme.

Au cours de la gestion active des LR, lors du développement ou de la modification des LR, les bonnes pratiques^{496,497,498} incluent notamment l'utilisation d'analyses en scripts (par exemple, R et Python) par le biais d'une interface de type pointer-cliquer avec des logiciels commerciaux (par exemple, Excel de Microsoft, SPSS d'IBM), suivant les lignes directrices de style (par exemple, PEP8⁴⁹⁹ pour Python), les commentaires et la documentation de code (par exemple, le fichier README, les outils de documentation automatisés tels que Sphinx⁵⁰⁰ et Doxygen⁵⁰¹), les tests (par exemple, pytest⁵⁰²), les versions publiées (par exemple, git), l'intégration continue (par exemple, Travis CI⁵⁰³), l'inclusion de fichiers de métadonnées dans le code source (par exemple, CodeMeta, CFF) et l'ajout d'une licence enregistrée (par exemple, SPDX, projet REUSE).

De nos jours, les chercheurs et les développeurs de logiciels de recherche s'appuient sur des composants et des infrastructures de LR actuels pour la gestion des LR. Ces infrastructures de LR comprennent des environnements de développement intégrés (par exemple, Jupyter notebook), des systèmes de gestion de flux de travail (par exemple kepler, pegasus), des

⁴⁹⁰ Les PGL sont aussi désignés par d'autres termes, par exemple les plans de développement logiciel, les plans de gestion des logiciels de projet, les déclarations de travail par CANARIE et le département de l'énergie (DOE) aux États-Unis, etc.

⁴⁹¹ <https://webarchive.nationalarchives.gov.uk/ukgwa/20201001001808/https://epsrc.ukri.org/files/funding/calls/2012/computational-science-and-engineering-software-for-the-future/>

⁴⁹² <https://www.nsf.gov/pubs/2016/nsf16532/nsf16532.htm>

⁴⁹³ [10.5281/zenodo.1422656](https://zenodo.org/record/1422656)

⁴⁹⁴ <https://cs.uwaterloo.ca/~apidduck/se362/Assignments/A2/spmp.pdf>

⁴⁹⁵ https://sceweb.uhcl.edu/helm/RationalUnifiedProcess/webtmpl/templates/mgmt/rup_sdpln_sp.htm

⁴⁹⁶ Best Practices for Scientific Computing: <https://doi.org/10.1371/journal.pbio.1001745>

⁴⁹⁷ Good enough practices in scientific computing: <https://doi.org/10.1371/journal.pcbi.1005510>

⁴⁹⁸ Software carpentry: <https://software-carpentry.org/>

⁴⁹⁹ <https://www.python.org/dev/peps/pep-0008/>

⁵⁰⁰ [Sphinx-doc.org](https://sphinx-doc.org)

⁵⁰¹ <https://www.doxygen.nl/index.html>

⁵⁰² <https://docs.pytest.org/en/6.2.x/>

⁵⁰³ <https://travis-ci.org/>

archives (par exemple, Software Heritage⁵⁰⁴, Zenodo, des dépôts institutionnels), des catalogues ou des registres (par exemple swMath, ASCL⁵⁰⁵, OpenAire, ScanR, CRAN, PyPi, RunMyCode⁵⁰⁶) et des éditeurs (par exemple IPOL, eLife, JOSS), qui constituent également de bonnes solutions et de bonnes pratiques⁵⁰⁷ pour améliorer la conformité des logiciels aux principes FAIR, ainsi que pour communiquer et soutenir la diffusion des logiciels, en offrant de meilleures possibilités de curation et de partage des LR.

F2 : Bonnes pratiques en matière de reproductibilité informatique (en tant que chercheurs ou auteurs de données/codes)

La reproductibilité informatique définit particulièrement cette forme de reproductibilité comme « l'obtention de résultats cohérents en utilisant les mêmes données d'entrée, les mêmes méthodes de calcul et les mêmes conditions d'analyse »⁵⁰⁸. Un nouveau guide a été publié pour tenter de répondre à la question « Dans quelle mesure les logiciels de recherche doivent-ils être reproductibles ? »⁵⁰⁹, qui définit quatre niveaux de reproductibilité, suggère des critères pour aider à déterminer le degré idéal de reproductibilité de votre LR et recommande des pratiques pour atteindre ce degré de reproductibilité.

Un principe de général de reproductibilité informatique consiste à fournir une description claire, précise et complète de la manière dont un résultat rapporté a été obtenu, bien que différents domaines d'étude ou types d'enquête puissent exiger différents types d'informations. Un dossier de recherche reproductible par calcul peut inclure intégralement ou partiellement des éléments suivants :

- Données primaires (et documentation) collectées et utilisées dans l'analyse ;
- Données secondaires (et documentation) collectées et utilisées dans l'analyse ;
- Résultats de la production des données primaires (et documentation) par l'analyse ;
- Résultats de la production de données secondaires (et documentation) par l'analyse ;
- Programmes logiciels (et documentation) pour le calcul des résultats publiés ;

⁵⁰⁴ <https://www.softwareheritage.org/howto-archive-and-reference-your-code/>

⁵⁰⁵ Astrophysics Source Code Library (ASCL): <https://ascl.net/>

⁵⁰⁶ <http://www.runmycode.org/>

⁵⁰⁷ Au cours des années 2019-2020, le [groupe de travail sur les bonnes pratiques pour les registres de logiciels](#) du groupe de travail sur la mise en œuvre de la citation des logiciels FORCE11 a travaillé à la création de [neuf bonnes pratiques pour les registres et dépôts de logiciels scientifiques](#).

⁵⁰⁸ National Academies of Sciences, Engineering, and Medicine. (2019). Reproducibility and replicability in science. National Academies Press.

⁵⁰⁹ <https://zenodo.org/record/4761867#.YKfD4ZNKg-Q>

- Programmes logiciels de recherche pour la reproduction des résultats publiés ;
 - Documenter le fonctionnement des LR et rendre cette documentation disponible ;
 - Gestion des versions ;
 - Capture de l'environnement d'exécution et de calcul :
 - Notamment les caractéristiques du matériel (par exemple, le nombre de cœurs dans toutes les unités centrales) et les caractéristiques des logiciels (par exemple, le système d'exploitation, les langages de programmation, les paquets de soutien, les autres logiciels installés, ainsi que leurs versions et configurations).
 - Documentation sur les dépendances du logiciel, y compris :
 - Logiciels installés, bibliothèques intégrées et versions associées ;
 - Métadonnées et spécifications ;
 - Structure du code source ;
 - Composants individuels qui supportent la fonctionnalité ;
 - Paramètres ;
 - Environnement de construction et d'exécution.
- Documentation sur les LR et détails de mise en œuvre de haute qualité ;
- GDR et accès intégral aux données utilisées dans la recherche ;
- Enregistrement des informations sur le flux de travail et la provenance des recherches informatiques ;
- Article de revue publié.

Il y a de nombreuses ressources visant à fournir aux chercheurs les outils, les connaissances, la formation et la confiance nécessaires pour suivre les bonnes pratiques de la science ouverte reproductible. En voici quelques exemples :

- [Méthode Turing](#)
- [Guide de reproductibilité de la science ouverte R](#)
- Outils de reproductibilité (infonuagiques) : [Conteneur Docker](#), [Binder](#), [Code Ocean](#), [Whole Tale](#), [eLife Reproducible Document Stack](#), [Galaxy](#), [Gigantum](#), [Manuscripts](#), [o2r](#), [REANA](#), [ReproZip](#), [Sweave](#), [Stencila](#), [Renku](#), [binder](#), [EaaS](#), [ReScience](#), Jupyter notebook, etc.

En plus des plateformes et des ressources, des pratiques sociales telles que les communautés fondamentales (par exemple, le UK Reproducibility Network⁵¹⁰) ont vu le jour.

Annexe G : La gestion des LR à travers le prisme de la COVID-19

La nouvelle maladie à coronavirus (COVID-19) a été officiellement déclarée pandémie mondiale par l'Organisation mondiale de la santé (OMS) le 11 mars 2020⁵¹¹. Au moment de la rédaction du présent rapport, des ressources massives ont permis aux chercheurs du monde entier de s'attaquer à cette crise sous de multiples angles, notamment la recherche médicale, les normes sociétales, la participation communautaire, les données, la puissance de calcul et les logiciels, ce qui a donné lieu à des résultats de recherche extraordinaires grâce à une plus grande adoption des pratiques de la science ouverte. Tirant pleinement parti de ce paysage, la réponse des communautés de recherche nationales et internationales pour aider à résoudre cette crise fournit un excellent exemple de l'impact de pratiques de gestion et de soutien solides en matière de LR. En voici quelques exemples :

- Les agences internationales de financement ont annoncé une intervention rapide pour financer la recherche liée à la COVID-19, dont la plupart comportaient des exigences pour que les résultats de la recherche (données, codes et publications) soient partagés ouvertement^{512,513,514}. Au Canada, cela inclut les programmes de financement de réponse rapide des trois organismes^{515,516}.
- Les établissements et les entreprises mondiales ont également augmenté leurs investissements dans le calcul informatique de pointe grâce à la COVID-19. Par exemple, afin de maximiser l'impact de la recherche, AMD a lancé un groupe de travail pour les bénéficiaires du fonds HPC COVID-19 et les ingénieurs d'AMD afin de discuter conjointement des domaines de recherche et des résultats ainsi que des optimisations matérielles et logicielles qui peuvent accélérer leur travail collectif⁵¹⁷. Google a également

⁵¹⁰ <https://www.ukrn.org/>

⁵¹¹ <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>

⁵¹² <https://wellcome.org/coronavirus-covid-19/open-data>

⁵¹³ <https://coronavirus.frontiersin.org/covid-19-research-funding-monitor>

⁵¹⁴ <https://nebigdatahub.org/covid19/>

⁵¹⁵ <https://www.idrc.ca/en/news/funding-opportunity-rapid-research-response-novel-coronavirus-outbreak>

⁵¹⁶ <https://cihr-irsc.gc.ca/e/51890.html>

⁵¹⁷ <https://www.amd.com/en/press-releases/2020-09-14-amd-covid-19-hpc-fund-adds-18-institutions-and-five-petaflops>

annoncé les subventions de recherche sur la COVID-19 de Google Cloud^{518,519}. Le consortium COVID-19 Performance Computing (HPC) est un autre exemple qui a réuni le gouvernement fédéral, l'industrie et les leaders universitaires pour fournir un accès aux ressources de calcul haute performance les plus puissantes au monde afin de soutenir la recherche sur la COVID-19⁵²⁰.

- Les communautés de recherche collaborent pour développer les bonnes pratiques pour le partage et la réutilisation des résultats de la recherche sur la COVID-19. Sur le plan international, le groupe de travail sur la COVID-19 de la Research Data Alliance (RDA) a réuni des experts multidisciplinaires du monde entier pour élaborer des lignes directrices et des recommandations sur le partage des données dans le contexte actuel de la COVID-19⁵²¹. Les recommandations et lignes directrices sur la COVID-19 de la RDA comprennent une section soulignant les bonnes pratiques pour les logiciels de recherche. Ces recommandations se concentrent sur le partage et le rôle clé qu'ils jouent dans l'analyse des données et l'avancement de la recherche.
- Les plateformes de partage de la recherche (telles que les dépôts) et les services de curation ont uni leurs forces pour soutenir la recherche liée à la COVID-19. L'un des plus importants est le dépôt de données Zenodo COVID-19 Community, qui est gratuit et ouvert aux chercheurs du monde entier pour partager de nombreux résultats de recherche (publications, données, logiciels, protocoles, etc.) liés à l'épidémie de COVID-19 ou au virus SRAS-CoV-2⁵²². La curation de ces ensembles de données est soutenue par le programme européen OpenAIRE, qui a lancé une version bêta en ligne de COVID-19 Open Research Gateway⁵²³ pour agréger, relier et fournir un accès aux produits scientifiques sur la COVID-19 à travers de multiples infrastructures et sources de communication savante. Données de recherche Canada (DRC) avait l'intention de rassembler les outils et les plateformes logicielles sur la COVID-19 dans sa représentation du réseau canadien de partage des données sur la COVID-19⁵²⁴.

⁵¹⁸ https://edu.google.com/programs/credits/research/?modal_active=none

⁵¹⁹ <https://uwaterloo.ca/research/google-apply-google-cloud-research-credits>

⁵²⁰ <https://covid19-hpc-consortium.org/>

⁵²¹ [RDA COVID-19 Recommandations et lignes directrices pour le partage des données](#)

⁵²² <https://zenodo.org/communities/covid-19/?page=1&size=20>

⁵²³ <https://beta.covid-19.openaire.eu/>

⁵²⁴ <https://share.mindmanager.com/#publish/Zv6lhDk-aGEptiZFkhjDBPr7HTzvX6vvj8bEpcx>

- Les milieux de la recherche ont également réagi à la crise. La plateforme iReceptor⁵²⁵ (basée à l'Université Simon Fraser) et le portail VirusSeq de CanCoGen⁵²⁶ (qui fait partie de l'initiative de recherche CanCOGeN COVID-19 de Génome Canada) sont des exemples de la contribution du Canada axée sur la génomique humaine et virale autour de la COVID-19 et du SRAS-CoV-2. iReceptor fournit une passerelle scientifique⁵²⁷ pour explorer The AIRR Data Commons, un réseau distribué de dépôts de données fondés sur des normes qui stockent les données de séquençage de la réponse immunitaire humaine anticorps/cellules B et T (données AIRR-seq). À l'aide d'un cadre qui suit les principes fondamentaux FAIR pour le développement de logiciels, l'équipe iReceptor a pu étendre la plateforme pour conserver, stocker et partager les données AIRR-seq provenant d'une série d'études internationales sur la COVID-19⁵²⁸. Depuis le 28 mai 2021, plus d'un milliard de séquences annotées provenant de 15 études sur la COVID-19 sont disponibles dans AIRR Data Commons⁵²⁹. Le portail VirusSeq de CanCoGen assure le suivi et l'analyse génomique de l'évolution des caractéristiques du virus du SRAS-CoV-2 au Canada, dans le but d'identifier et de suivre les tendances de la transmission à l'échelle régionale, provinciale, nationale et internationale, de faciliter la détection de nouveaux groupes de cas et d'éclousions et de découvrir l'évolution des caractéristiques virales. En date du 27 avril 2021, le portail VirusSeq contient 111 génomes viraux provenant de deux études réalisées au Canada.

La COVID-19 a démontré ce qui est possible dans le monde de la recherche lorsqu'il y a un besoin impérieux. En regardant les progrès qui ont été réalisés en termes de recherche scientifique sur la COVID-19, la prépublication avec une approche de science ouverte très itérative a joué un rôle important. Les instances de financement ont été en mesure d'agir de manière très agile (en priorisant les ressources pour les travaux sur la COVID-19 en quelques semaines au lieu de quelque mois) pour soutenir les recherches importantes et urgentes sur la COVID-19. Cela a été possible grâce aux partenariats préexistants entre les instances de financement et les communautés de recherche. Un autre facteur clé a été la possibilité d'utiliser les technologies numériques à distance.

De plus, il a été fascinant de constater la force de cohésion de toutes les intervenants (y compris les communautés de chercheurs, les bailleurs de fonds) et la rapidité avec laquelle ils ont pu se

⁵²⁵ Corrie et al. iReceptor : A platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories, *Immunological Review*, Volume 284, Issue 1.
<https://doi.org/10.1111/imr.12666>

⁵²⁶ <https://virusseq-dataportal.ca/> Le portail canadien de données VirusSeq (CanCOGeN VirusSeq) constitue un lien essentiel entre les unités de santé publique du Canada et les chercheurs qui suivent l'évolution du virus et des variantes préoccupantes.

⁵²⁷ <https://gateway.ireceptor.org/login>

⁵²⁸ Scott et al. The adaptive immune receptor repertoire community as a model for FAIR stewardship of big immunology data, *Current Opinion in Systems Biology*, Volume 24, December 2020
<https://doi.org/10.1016/j.coisb.2020.10.001>.

⁵²⁹ <http://www.ireceptor.org/covid19>

réunir, faire avancer les choses, obtenir l'accès aux installations informatiques et la vitesse à laquelle ceux-ci ont pu déployer les LR et les plateformes de LR.

Bien qu'il soit trop tôt pour évaluer l'impact total des initiatives et des programmes connexes sur la pandémie elle-même, on observe néanmoins un changement de mentalité évident dans l'écosystème d'IRN vers une plus grande ouverture et collaboration, ainsi que l'adoption des bonnes pratiques du mouvement de la science ouverte à l'échelle nationale et internationale. En ce qui concerne le paysage des LR au Canada, ces développements ne se produiraient pas sans l'investissement continu du Canada dans les bonnes pratiques de développement des LR et son engagement à faire du Canada un chef de file dans la communauté des LR. La question probablement la plus importante est de savoir comment le degré de collaboration et de partage de données autour de la COVID-19 peut être encouragé pour qu'il se poursuive après la pandémie et comment nous pouvons profiter de cette occasion pour aider à coordonner cette collaboration et la faire progresser.

Annexe H : Bailleurs de fonds pour le développement et l'utilisation des LR

Nom	site Web	Région/type	Notes
Agriculture et Agroalimentaire Canada : Initiatives Agri-risques : Microsubventions	https://agriculture.canada.ca/fr/programmes-services-agricoles/initiatives-agri-risques-microsubventions	fédéral/sources de financement gouvernementales	
Alberta Cancer Foundation	https://www.albertacancer.ca/	gouvernements provincial/régional/organismes de développement (AB)	
Alberta Centre for Child, Family and communauté Research (ACCFR)	https://policywise.com/	gouvernements provincial/régional/organismes de développement (AB)	
Alberta Innovates (AB Innov)	https://albertainnovates.ca/	gouvernements provincial/régional/organismes de développement (AB)	
Société Alzheimer du Canada	https://alzheimer.ca/fr	fondations et sociétés (secteur public)	
Amazon Research Awards	https://www.amazon.science/research-awards	investissement de société privée	
Andrew W. Mellon Foundation	https://mellon.org/	fondations et sociétés (secteur public)	subventions versées par une fondation américaine pour financer le développement des sciences humaines et sociales et le développement d'un LR
Arthritis Society	https://arthritis.ca/	fondations et sociétés (secteur public)	
Agence de promotion économique du Canada	https://www.canada.ca/fr/promotion-economique-canada-	gouvernements provincial/régional/organismes de	

atlantique (APCA)	atlantique.html	développement (Atlantique)	
Banting Research Foundation	https://www.bantingresearchfoundation.ca/	fondations et sociétés (secteur public)	
Bill and Melinda Gates Foundation (BMGF)	https://www.gatesfoundation.org/	financement international (secteur privé)	établi aux États-Unis
Brain Canada	http://www.caro-acro.ca/	fondations et sociétés (secteur public)	
Fondation canadiennes des tumeurs cérébrales	https://www.braintumour.ca/fr/	fondations et sociétés (secteur public)	
C17 Council (C17)	http://www.c17.ca/	fondations et sociétés (secteur public)	
Chaires d'excellence en recherche du Canada (CERC)	https://www.cerc.gc.ca/home-accueil-fra.aspx	fédéral/financement gouvernemental d'initiatives stratégiques	
Fonds d'excellence en recherche Apogée Canada (FERAC)	https://www.cfref-apogee.gc.ca/home-accueil-eng.aspx#	fédéral/financement gouvernemental d'initiatives stratégiques	p. ex. monde Water Futures - université de la Saskatchewan
Chaires d'excellence en recherche du Canada (CERC)	https://www.cerc.gc.ca/home-accueil-fra.aspx	fédéral/financement gouvernemental d'initiatives stratégiques	
Fonds d'urgence pour la continuité de la recherche au Canada (FUCRA)	https://www.sshrc-crsh.gc.ca/funding-financement/crcef-fucrc/index-fra.aspx	fédéral/financement gouvernemental d'initiatives stratégiques	
Le Réseau national de la recherche et de l'éducation du Canada (RNRE)	https://www.canarie.ca/fr/rnre/	consortium	
Association des bibliothèques de recherche du Canada (ABRC)	https://www.carl-abrc.ca/fr/	consortium	
Société canadienne du sang	https://www.blood.ca/fr	fondations et sociétés (secteur public)	
Société canadienne du cancer (SCC)	https://cancer.ca/fr/	fondations et sociétés (secteur public)	
Fondation canadienne pour l'innovation (FCI)	https://www.innovation.ca/fr	fédéral/financement gouvernemental d'initiatives stratégiques	
Société canadienne de l'hémophilie (SCH)	https://www.hemophilia.ca/fr/	fondations et sociétés (secteur public)	
Instituts de recherche en santé du Canada (IRSC)	https://cihr-irsc.gc.ca/f/193.html	fédéral/sources de financement gouvernementales	
programmeme canadien de l'innovation à l'international (PCII))	https://www.delequescommercial.gc.ca/funding-financement/ciip-pcii/index.aspx?lang=fra&qa=2.175023846.574400765.1642427776-809489988.1642427776	fédéral/financement gouvernemental d'initiatives stratégiques	
Agence canadienne de développement économique du Nord (CanNor)	https://www.cannor.gc.ca/fra/1351104567432/1351104589057	fédéral/sources de financement gouvernementales /communautés/organismes autochtones	
Fondation canadienne	https://whenithurtstomove.org/fr	fondations et sociétés (secteur public)	

d'orthopédie		public)	
partenariat canadien contre le cancer	https://www.partnershipagainstcancer.ca/fr/	fondations et sociétés (secteur public)	
Canadian Prostate Cancer Research Initiative (CPCRI)	https://webapps.cihr-irsc.gc.ca/cris/agency_info_E?p_agency_cd=CPCRI&p_popup=T	fondations et sociétés (secteur public)	
Agence spatiale canadienne (ASC)	https://www.asc-csa.gc.ca/fra/default.asp	fédéral/sources de financement gouvernementales	
Initiative canadienne de recherche sur le contrôle du tabac	http://www.ohpe.ca/node/4717	fondations et sociétés (secteur public)	
CANARIE	https://www.canarie.ca/fr/rnre/	fédéral/financement gouvernemental d'initiatives stratégiques	
Oncologie Nouvelle-Écosse	Nova Scotia Health Authority (nshealth.ca)	gouvernements provincial/régional/organismes de développement (N.-É.)	
Oncologie Ontario	https://www.cancercareontario.ca/en	gouvernements provincial/régional/organismes de développement (ON)	
Société de recherche sur le cancer	https://www.societederecherchesurlecancer.ca/fr	fondations et sociétés (secteur public)	
ActionCancer Manitoba	https://www.cancercare.mb.ca/Patient-Family/cancercare-team	gouvernements provincial/régional/organismes de développement (MB)	
Children's Oncology Group (COG)	https://childrensoncologygroup.org/	fondations et sociétés (secteur public)	Le groupe d'oncologie pédiatrique (COG) est financé en grande partie par l'institut national du cancer et reçoit des fonds additionnels d'autres organismes subventionnaires et de sources philanthropiques (thecogfoundation.org)
Colleges et Instituts Canada	https://www.collegesinstitutes.ca/fr/	fédéral/sources de financement gouvernementales/organismes/communautés autochtones	
Collèges et instituts Canada - Stages en technologies propres	NSCC lie le talent et les possibilités avec des stages en technologies propres (collegesinstitutes.ca)	fédéral/sources de financement gouvernementales/organismes/communautés autochtones	
Communitech DATA.BASE programme	https://www.communitech.ca/	fondations et sociétés (secteurs public et privé)	
Consortium de recherche et d'innovation en aérospatiale au Québec (CRIAQ)	https://www.criaq.aero/	consortium	
Consortium de RECHERCHE et d'INNOVATION en transformation métallique (CRITM)	https://www.critm.ca/	consortium	
Consortium de recherche et innovations en	http://www.cribiq.gc.ca/	consortium	

bioprocédés industriels au Québec (CRIBIQ)			
Consortium québécois sur la découverte de médicaments (CQDM)	https://cqdm.org/fr/	consortium	
Council on Library and Information Resources (CLIR)	https://www.clir.org/	fondations et sociétés (secteur public)	établi aux États-Unis, le programme de bourses de recherche postdoctorale du CLIR finance la conservation des logiciels (p. ex. pour les applications de réalité virtuelle et d'image en 3D)
Crohn et Colite Canada	https://crohnetcolite.ca/	fondations et sociétés (secteur public)	
Relations Couronne-Autochtones et Affaires du Nord Canada - Stages en technologies propres	Relations Couronne-Autochtones et Affaires du Nord Canada - Canada.ca	fédéral/sources de financement gouvernementales /organismes/communautés autochtones	
Relations Couronne-Autochtones et Affaires du Nord Canada - Stages en technologies propres - programmeme de lutte contre les contaminants dans le Nord	https://science.gc.ca/eic/site/063.nsf/fra/h_7A463DBA.html	fédéral/sources de financement gouvernementales /organismes/communautés autochtones	
Patrimoine canadien	https://www.canada.ca/fr/patrimoine-canadien.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Patrimoine canadien - Fond pour l'histoire du Canada	https://www.canada.ca/fr/patrimoine-canadien/services/financement/fonds-histoire-canada/renseignements.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Défense nationale	https://www.canada.ca/fr/ministere-defense-nationale.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Fondation de recherche médicale sur la dystonie Canada	https://dystoniacanada.org/fr	fondations et sociétés (secteur privé)	
Bourses de nouveaux chercheurs	https://www.ontario.ca/fr/page/bourses-de-nouveaux-chercheurs	gouvernements provincial/régional/organismes de développement (ON)	
Emploi et Développement social Canada	https://www.canada.ca/fr/emploi-developpement-social.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Environnement et Changement climatique Canada	https://www.canada.ca/fr/environnement-changement-climatique.html	fédéral/financement gouvernemental d'initiatives stratégiques	
European Commission "Horizon 2020"	https://ec.europa.eu/programmes/horizon2020/en	financement international (secteur public)	
Fédération canadienne des municipalités - Fonds municipal vert	https://www.fcm.ca/fr/programmes/fonds-municipal-vert	fondations et sociétés (secteur privé)	
FRQ - Nature et Technologies (FRQNT)	https://frq.gouv.qc.ca/nature-et-technologies/	gouvernements provincial/régional/organismes de développement (QC)	

FRQ - Santé (FRQS)	https://frq.gouv.qc.ca/sante/	gouvernements provincial/régional/organismes de développement (QC)	
FRQ - Société et Culture (FRQSC)	https://frq.gouv.qc.ca/societe-et-culture/	gouvernements provincial/régional/organismes de développement (QC)	
<u>Genome BC</u>	https://www.genomebc.ca/	gouvernements provincial/régional/organismes de développement (BC)	
Genome Canada	https://www.genomecanada.ca/fr	fédéral/financement gouvernemental d'initiatives stratégiques	
Génome Québec	https://www.genomequebec.com/	gouvernements provincial/régional/organismes de développement (QC)	
Affaires mondiales Canada	https://www.international.gc.ca/monde-affaires-affaires-mondiales/home-accueil.aspx?lang=fra	financement international (secteur public)	
Google Cloud research credits	https://edu.google.com/intl/en_ca/programmes/credits/research/?modal_active=none	investissement de société privée	
Gordon and Betty Moore Foundation	https://www.moore.org/	financement international (secteur privé)	établi aux États-Unis
Gordon Foundation	https://gordonfoundation.ca/	fondations et sociétés (secteur privé)	
Gouvernement du Canada	https://www.canada.ca/fr/emploi-developpement-social/services/financement.html	fédéral/sources de financement gouvernementales	
Fondation Graham Boeckh	https://grahamboeckhfoundation.org/fr/	fondations et sociétés (secteur privé)	
<u>Grand Challenges</u>	https://grandchallenges.org/subvention-opportunities	financement international (secteur public)	
<u>Grand Challenges Canada (GCC)</u>	https://grandchallenges.org/subvention-opportunities	financement international (secteur public)	
subventions.gov	subventions.gov	financement international (secteur public)	programmes de subvention américains aux États-Unis, p.ex. ministère de la défense (DOD), instituts nationaux de la santé (NIH), fondation scientifique nationale (NSF), ministère de l'énergie (DOE)
Santé Canada	https://www.canada.ca/fr/sante-canada.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Heart And Stroke Foundation	https://www.heartandstroke.com/	fondations et sociétés (secteur public)	
<u>Helmsley Charitable Trusthelmsleytrust.org</u>	https://helmsleytrust.org/	financement international (secteur privé)	établi aux États-Unis
Huawei Innovation Research programme (HIRP)	https://innovationresearch.huawei.com/IPD/hipr/portal/index.html	investissement de société privée	
Impact Canada	https://impact.canada.ca/fr	fédéral/financement gouvernemental d'initiatives stratégiques	
Initiative des	https://www.ic.qc.ca/eic/site/093	fédéral/financement gouvernemental	

supergrappes d'innovation	nsf/fra/accueil	d'initiatives stratégiques	
Innovation, Sciences et Développement économique Canada	https://www.ic.gc.ca/eic/site/icgc.nsf/fra/accueil	fédéral/financement gouvernemental d'initiatives stratégiques	
InnovÉÉ	https://innov-ee.ca/	consortium	
Centre de recherches pour le développement international	https://www.idrc.ca/fr	financement international (secteur public)	
Jarislowsky Foundation	https://www.charitydata.ca/charity/the-jarislowsky-foundation-la-fondation-jarislowsky/894744036RR0001/	fondations et sociétés (secteur public)	
McLean Foundation	http://www.mcleanfoundation.ca/	fondations et sociétés (secteur privé)	
Recherche en santé mentale Canada	https://www.mhrc-rsmc.ca/	fondations et sociétés (secteur public)	
Michael J. Fox Foundation (MJFF)	https://www.michaeljfox.org/	financement international (secteur public)	établi aux États-Unis
Michael Smith Foundation for Health Research (MSFHR)	https://www.msfhr.org/	fondations et sociétés (secteur public); gouvernements provincial/régional/organismes de développement (BC)	
Ministère de l'Économie et de l'Innovation (MEI) (anciennement le MESI)	https://www.economie.gouv.qc.ca/	gouvernements provincial/régional/organismes de développement (QC)	
Ministère des relations internationales et de la Francophonie (MRIF)	http://www.mrif.gouv.qc.ca/fr/	gouvernements provincial/régional/organismes de développement (QC)	
Ministère des Collèges et des Universités	https://www.ontario.ca/fr/page/ministere-des-colleges-et-universites	gouvernements provincial/régional/organismes de développement (ON)	
Ministry of Research, Innovation and Science (MRIS)	https://www.ontario.ca/page/ministry-economic-development-job-creation-trade	gouvernements provincial/régional/organismes de développement (ON)	
MITACS	https://www.mcgill.ca/research/research/funding/Fédéral/Government/mitacs	fédéral/financement gouvernemental d'initiatives stratégiques	
National Research Council Canada (NRC)	https://nrc.canada.ca/en	fédéral/financement gouvernemental d'initiatives stratégiques	
Sciences naturelles Resources Canada (NRCan)	https://www.nrcan.gc.ca/home	fédéral/financement gouvernemental d'initiatives stratégiques	
Sciences naturelles Sciences and Engineering Research Council of Canada (NSERC)	http://www.nserc-crsng.gc.ca/index_eng.asp	fédéral/sources de financement gouvernementales	
Networks of Centres of Excellence (NCE)	https://www.nce-rce.gc.ca/Index_eng.asp	fédéral/financement gouvernemental d'initiatives stratégiques	
New Brunswick Health Research Foundation (NBHRF)	https://www.nbhrf.com/	gouvernements provincial/régional/organismes de développement (NB)	

New Frontiers in Research Fund (NFRF)	https://www.sshrc-crsh.gc.ca/funding-financement/nfrf-fnfr/index-eng.aspx	fédéral/sources de financement gouvernementales	
Nova Scotia Health Research Foundation (NSRF)	http://www.cdha.nshealth.ca/discovers-innovation/research-fund	gouvernements provincial/régional/organismes de développement (N.-É.)	
Ocean Protection Funding Waitt Foundation	https://www.waittfoundation.org/	fondations et sociétés (secteur privé)	siège social à Washington, DC
Institut ontarien du cerveau	https://braininstitute.ca/fr/	gouvernements provincial/régional/organismes de développement (ON)	
Centres de l'excellence de l'Ontario	https://www.oce-ontario.org/	gouvernements provincial/régional/organismes de développement (ON)	
Fonds ontarien de recherche pour l'intervention rapide contre la COVID-19	https://www.ontario.ca/fr/page/fonds-ontarien-de-recherche-pour-lintervention-rapide-contre-la-covid-19	gouvernements provincial/régional/organismes de développement (ON)	
Ontario Genomics	https://www.ontariogenomics.ca/	gouvernements provincial/régional/organismes de développement (ON)	
Institut ontarien de recherche sur le cancer	https://oicr.on.ca/fr/a-propos/	gouvernements provincial/régional/organismes de développement (ON)	
Ministère de l'Agriculture, de l'Alimentation et des Affaires rurales de l'Ontario	http://www.omafra.gov.on.ca/french/index.html	gouvernements provincial/régional/organismes de développement (ON)	
Ministère de l'Agriculture, de l'Alimentation et des Affaires rurales de l'Ontario	http://www.omafra.gov.on.ca/french/index.html	gouvernements provincial/régional/organismes de développement (ON)	
Ministère de la Santé de l'Ontario	Ministère de la Santé Ontario.ca	gouvernements provincial/régional/organismes de développement (ON)	
Ministère des transports de l'Ontario	https://www.ontario.ca/fr/page/ministere-des-transport	gouvernements provincial/régional/organismes de développement (ON)	
Fonds pour la recherche en Ontario – Excellence en recherche	https://www.ontario.ca/fr/page/fonds-pour-la-recherche-en-ontario-excellence-en-recherche	gouvernements provincial/régional/organismes de développement (ON)	
Fonds pour la recherche en Ontario – Infrastructure de recherche	https://www.ontario.ca/fr/page/fonds-pour-la-recherche-en-ontario-infrastructure-de-recherche	gouvernements provincial/régional/organismes de développement (ON)	
Cancer de l'ovaire Canada	https://ovairecanada.org/Home-1?lang=fr-CA&_ga=2.174196258.1089352693.1642432680-823893308.1642432680	fondations et sociétés (secteur public)	
Pancreatic Cancer Canada Foundation (PCCF)	https://pancreaticcancercanada.ca/	fondations et sociétés (secteur public)	

Paper Presentation subventions	https://www.mcgill.ca/research/research/funding/internal	financement interne	
Partnership for Clean Competition (PCC)	https://cleancompetition.org/	fondations et sociétés (secteur public)	établi aux États-Unis
Savoir polaire Canada	https://www.canada.ca/fr/savoir-polaire.html	fédéral/sources de financement gouvernementales/organismes/communautés autochtones	
PRIMA Québec (anciennement NanoQuébec)	https://www.prima.ca/	consortium	
PROCURE	https://www.procure.ca/	fondations et sociétés (secteur privé)	
PROMPT (anciennement CINQ)	https://promptinnov.com/	consortium	
Cancer de la prostate	https://cancer.ca/fr/about-us/prostate-cancer	fondations et sociétés (secteur public)	
Agence de la santé publique du Canada	https://www.canada.ca/fr/sante-publique.html	fédéral/sources de financement gouvernementales	
Sécurité publique Canada	https://www.securitepublique.gc.ca/index-fr.aspx	fédéral/financement gouvernemental d'initiatives stratégiques	
Fondation cancer du sein du Québec	https://www.charityintelligence.ca/charity-details/612-quebec-breast-cancer-foundation	fondations et sociétés (secteur public)	
Research Manitoba	https://researchmanitoba.ca/	gouvernements provincial/régional/organismes de développement (MB)	
La Société royale du Canada	https://rsc-src.ca/fr	fondations et sociétés (secteur public)	
Sabbatical Leave Research subventions	https://www.mcgill.ca/research/research/funding/internal	financement interne	
Saskatchewan Cancer Agency	http://www.saskcancer.ca/	gouvernements provincial/régional/organismes de développement (SK)	
Saskatchewan Health Research Foundation (SHRF)	https://www.shrf.ca/	gouvernements provincial/régional/organismes de développement (SK)	
Institut indo-canadien Shastri	https://www.shastriinstitute.org/	fondations et sociétés (secteur public)	organisation binationale
Simons Foundation	https://www.simonsfoundation.org/	fondations et sociétés (secteur privé)	
Sloan Foundation	https://sloan.org/	financement international (secteur privé)	établi aux États-Unis
Conseil de recherches en sciences humaines (CRSH)	https://www.sshrc-crsh.gc.ca/home-accueil-fra.aspx	fédéral/sources de financement gouvernementales	
SSHRC General Research Fund (GRF) et NSERC GRF	https://uwaterloo.ca/research/find-and-manage-funding/manage-funding/closing-research-accounts/general-research-fund-guidelines-nserc-and-sshrc-only	financement interne	
Condition féminine Canada	https://cfc-swc.gc.ca/index-fr.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Technology Academy Finland (TAF)	https://taf.fi/	financement international (secteur public)	

Agence d'évaluation d'impact du Canada	https://www.canada.ca/fr/agence-evaluation-impact.html	fédéral/financement gouvernemental d'initiatives stratégiques	
Institut de prévention des sinistres catastrophiques	https://www.iclr.org/	fondations et sociétés (secteur public)	
La Fondation Terry Fox	https://terryfox.org/fr/	fondations et sociétés (secteur public)	
Transports Canada	https://tc.canada.ca/fr	fédéral/financement gouvernemental d'initiatives stratégiques	
U.S. Air Force - Research Interests of the Air Force Office of Scientific Research	https://www.Fédéral/Government/subventions.com/Research-Interests-of-the-Air-Force-Office-of-Scientific-Research-68274.html	financement international (secteur public)	fédéral/gouvernement américain - subventions
UK Research and Innovation (UKRI)	https://www.UK.ri.org/	financement international (secteur public)	établi au R.-U.
Wellcome Trust	https://wellcome.org/	financement international (secteur privé)	établi aux États-Unis
Weston Family Foundation	https://www.westonfoundation.org/	fondations et sociétés (secteur privé)	
Workplace Safety and Insurance Board Ontario (WSIB)	https://www.wsib.ca/fr/programme-recherche-et-subvention	gouvernements provincial/régional/organismes de développement (ON)	
NumFOCUS	https://numfocus.org/	États-Unis	développement de logiciels

Annexe I : Paysage de la communauté de LR

Nom	Type	Site Web	Région	Discipline	Objet du logiciel	Notes
(International) Society of Research Software Engineering/UK-RSE	initiative communautaire, événements	https://society-rse.org/	monde	toutes	général	
ACM Badging schema	projets	https://www.acm.org/publications/policies/artifact-review-and-badging-current	monde	toutes	évaluation des artefacts, politique	ACM. (2016). Artifact review and badging. https://www.acm.org/publications/policies/artifact-review-and-badging-current ⇒ Examen des artefacts et badges version 1.1 - 24 août 2020
Alliance of German Science Organisations: Working group on		https://www.allianzinitiative.de/fields-of-action-projets/digital-	Allemagne	toutes	général	logiciels de recherche - organisations, initiatives,

Software and Services		tools-software-and-services/?lang=en				programmes, projets, subventions, communautés de pratique, réseaux, revues, conférences
Allianz-Initiative	initiative communautaire/gouvernementale	https://www.allianzinitiative.de/fields-of-action-projets/digital-tools-software-and-services/?lang=en	Allemagne	toutes	général	
Apache Software Foundation	initiative communautaire	https://www.apache.org/	monde	toutes	science ouverte	
Astrophysics Source Code Library	registre	https://ascl.net/	monde	astronomie	indexage, citation	
Australian Research Data Commons	initiative gouvernementale	https://ardc.edu.au/collaborations/strategic-activities/software-and-platforms/	Australie	toutes	développement et pérennité de logiciels	
Awesome Research Software Registries	registre, catalogue, index, dépôt, plateforme	https://github.com/NLeSC/awesome-research-software-registries	monde	toutes	registre	
Better Scientific Software (BSSw)	initiative communautaire	http://bssw.io	É.-U., international	toutes	durabilité, productivité	
Colloque canadien sur les logiciels de recherche	événements	https://www.canarie.ca/fr/event/cclr-2021/	Canada	toutes	général, renforcement d'une communauté, durabilité	
CANARIE (Canada)	initiative gouvernementale	https://www.canarie.ca/fr/event/cclr-2021/	Canada	toutes	durabilité	
Code.gov	plateforme, programme, registre	https://code.gov/	monde	toutes	pratiques exemplaires, code source libre, partage, collaboration, réutilisation,	
CodeMeta initiative	initiative communautaire, groupes	https://codemeta.github.io/	monde	toutes	métadonnées, citation	
communauté Surface Dynamics Modeling System (CSDMS)	initiative communautaire	https://csdms.colorado.edu/	monde	sciences de la Terre	général	

CRAN	registre	https://cran.r-projet.org/	monde	code (langue) R	gestionnaires de paquets	
Dagstuhl	revue, projets	https://www.dagstuhl.de/en/publications/oasics	Allemagne	Informatique	édition	
Dagstuhl Artifacts Series (DARTS)	projets	https://www.dagstuhl.de/en/publications/darts/	Allemagne	toutes	évaluation des artefacts, politique	Wagner, M. (2017). Hitting the Bull's eye with darts: Artifact evaluation in computer science. https://doi.org/10.5281/ZENODO.583007
DARPA Open catalogue	catalogue	http://www.darpa.mil/opencatalogue/	É.-U.	toutes	conservation, registre,	
de-RSE	groupe RSE (research software engineers)	https://www.de-rse.org/en/	Allemagne	toutes	général	
de-RSE - Society for Research Software Engineers	initiative communautaire	https://de-rse.org/de/index.html	Allemagne	toutes	général	
Debian Astro	initiative communautaire	https://blends.debian.org/astro	monde	astrophysique	productivité	
DEVLOG réseau	initiative communautaire, groupes	http://devlog.cnrs.fr/reunion	France	toutes	renforcement de la communauté, général	
Digital préservation Coalition	initiative communautaire	https://www.dpconline.org/	monde	toutes	durabilité	quelques événements liés aux logiciels : https://www.dpconline.org/knowledge-base/pr%C3%A9servation-lifecycle/software
EaaS	entité à but non lucratif, initiative communautaire, programme, projet, subvention, communauté de pratique, réseau	https://www.softwarepreservationnetwork.org/eaas/	É.-U., international	toutes	conservation	
eLife	entité à but non lucratif, initiative communautaire, revue,	https://elifesciences.org/	monde	sciences biomédicales	code source libre, partage, culture, pratiques, collaboration,	

	projet, communauté de pratique			biochimie et sciences biologiques	reproductibilité computationnelle, EDI, productivité	
ELIXIR	initiative gouvernementale internationale	https://elixir-europe.org/	monde	sciences biologiques	registre d'outils, formation	
EOSC Architecture WG - Scholarly Infrastructures for Research Software (SIRS)	initiative communautaire, groupes	https://www.eoscsecretariat.eu/working-groupes/architecture-working-group	Europe	toutes	infrastructure, durabilité, interopérabilité, identification des paramètres	Vers une architecture ouverte d'infrastructures savantes pour les logiciels de recherche
EOSC-Context: OpenAireAdvanced, EOSChub, EOSCpilot	initiative gouvernementale internationale	https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud	Europe	toutes	interopérabilité, registre d'outils	
ePLAN	initiative communautaire	https://escience-plateforme.nl/	Pays-Bas	toutes	général	
European Environment for Scientific Software Installations (EESSI)		https://www.eessi-hpc.org/	Europe			
F1000Research	groupes, éditeur, plateforme	https://f1000research.com/	monde	toutes	général	nouveau groupe sur l'examen des codes. Plan du groupe de travail sur l'examen du code F1000
Five recommendations for FAIR software	projet	https://fair-software.nl/	monde	indifférent au domaine	général	
FORCE11 Software Citation Implementation Working Group	initiative communautaire	https://www.force11.org/group/software-citation-implementation-working-group	monde	toutes	citation	
German Research Foundation	initiative gouvernementale	https://www.dfg.de/en/index.jsp	Allemagne	toutes	durabilité	
Helmholtz science ouverte	entité à but non lucratif	https://os.helmholtz.de/open-science-in-der-helmholtz-gemeinschaft/forschungsoftware/	Allemagne	sciences de la nature, sciences biologiques	science ouverte	

				ques, génie		
Hidden REF	initiative communautaire	https://hidden-ref.org/	R.-U.	toutes	citation	
Hidden Ref	initiative communautaire	https://hidden-ref.org/	R.-U.	toutes	impact	
HUBzero	plateforme	https://hubzero.org/	monde	toutes	développement de sites Web, code source libre, édition, partage, collaboration scientifique, renforcement de la communauté, outils de calcul et d'analyse	
IDEAS Software Productivity	projet	https://ideas-productivity.org/	É.-U.	toutes	durabilité, productivité	
Image Processing On Line (IPOL)	revue, projets	https://www.ipol.im/	monde	traitement d'images, étude d'images, mathématiques	reproductibilité, science ouverte,	
INCF/OCNS Infrastructure/Software/Tools SIG	entité à but non lucratif, initiative communautaire	https://neurostars.org/t/ocns-infrastructure-software-tools-sig-meet-and-greet-initial-discussions/15560/17 , https://github.com/orgs/OCNS/teams/software-sig	monde	neurosciences	général	
Journal of Open Research Software (JORS)	revue	https://openresearchsoftware.metajnl.com/	monde	toutes	citation, préservation	
Journal of code source libreSoftware (JOSS)	revue	http://joss.theoj.org	monde	toutes	citation, préservation	
Knowledge Exchange (cooperation of Jisc, SURF, CSC, DFG, DEFF, CNRS)	collaboration, initiative communautaire	http://www.knowledge-exchange.info/	Europe	toutes	libre accès, recherche ouverte, partage	

Knowledge Exchange (cooperation of Jisc, SURF, CSC, DFG, DEFF, CNRS)		http://www.knowledge-exchange.info/	monde	toutes	durabilité	
Mass Open Cloud	projet	https://massopen.cloud	É.-U.	toutes	reproductibilité, partage	
Mentored Sprints	initiative communautaire, groupes	https://mentored-sprints.netlify.app/	É.-U.	toutes	pratiques exemplaires, formation	
Molecular Science Software Institute	projet	https://molssi.org	monde	sciences moléculaires		
Mozilla	société	https://www.mozilla.org/en-US/	monde	toutes	science ouverte	
NASA code source libreSoftware	catalogue	https://code.nasa.gov/	É.-U.	toutes	partage, registre	
Pays-Bas eScience Center	initiative gouvernementale	https://www.esciencecenter.nl/	Pays-Bas	toutes	général	
réseau for Computational Modeling in Social and Ecological Sciences (CoMSES Net)	initiative communautaire, plateforme, dépôt	https://www.comses.net/	monde	sciences sociales et écologie	pratiques exemplaires, citation, conservation, reproductibilité, réutilisation	
Nouvelle-Zélande eScience Infrastructure (NeSI)	initiative gouvernementale	https://www.nesi.org.nz/	Nouvelle-Zélande	toutes	général	
NIH Data Commons	initiative gouvernementale	https://commonfund.nih.gov/commons	É.-U.	toutes	interopérabilité	
NL-RSE	groupe RSE (research software engineers)	https://nl-rse.org/	Pays-Bas	toutes	général	
Nordic-RSE Group	initiative communautaire	https://nordic-rse.org/	pays nordiques	toutes	général	
NumFOCUS	entité à but non lucratif	https://numfocus.org/	monde	toutes	science ouverte	
NumFOCUS	entités à but non lucratif	https://numfocus.org/	monde	toutes	pratiques exemplaires, renforcement de la communauté	
Open BioInformatique Foundation	initiative communautaire	https://www.open-bio.org/	monde	sciences	général	

				biologiques		
Open Force Field Initiative	initiative communautaire	https://openforcefield.org/	monde	sciences biologiques, chimie, matériaux, physique	développement et interopérabilité de logiciels	
OpenAIRE Explore	plateforme, infrastructure, agrégateur	https://explore.openaire.eu/	Europe	toutes	interopérabilité, registre d'outils, science ouverte, partage, découvrabilité, transparence, pratiques exemplaires	
OpenForum Europe (OFE)	entités à but non lucratif	https://openforumeurope.org/	Europe	toutes	interopérabilité, code source libre, politique	
PLAN-E	initiative communautaire	https://plan-europe.eu/	Union européenne	toutes	durabilité	
PyPI	registre	https://pypi.org/	monde	code Python	gestionnaires de paquets	
RDA CURE-FAIR WG	initiative communautaire, groupes	https://www.rd-alliance.org/groupe/cure-fair-wg	monde	toutes	conservation, reproductibilité, durabilité, FAIR	
RDA FAIR for Research Software (FAIR4RS) WG	initiative communautaire, groupes	https://www.rd-alliance.org/groupe/fair-research-software-fair4rs-wg	monde	toutes	FAIR, général, durabilité	
RDA Software Source Code Identification (SCID) WG	initiative communautaire, groupes	https://www.rd-alliance.org/groupe/rd-force11-software-source-code-identification-wg	monde	toutes	identification, général	Identification du code source de logiciel Cas d'utilisation et schémas d'identificateur pour l'identification persistante du code source de logiciel
RDA Software Source Code IG (SSC IG)	initiative communautaire	https://www.rd-alliance.org/groupe/software-source-code-ig	monde	toutes	général	

RDA-COVID19-Software	initiative communautaire, groupes	https://www.rd-alliance.org/groupe/rda-covid19-software	monde	toutes	partage, durabilité, pratiques exemplaires	Recommandations et lignes directrices de la RDA sur la COVID-19 pour le partage des données
Red Hat Collaboratory	initiative communautaire, sociétés	https://www.bu.edu/rhcollab/	É.-U.	toutes	renforcement de la communauté, code source libre	
ReSA (Research Software Alliance)	initiative communautaire	https://www.researchsoft.org/	monde	toutes	général	Groupe de travail de la Research Software Alliance (ReSA)
Research Data Alliance	initiative communautaire	https://www.rd-alliance.org/	monde	toutes	général	
Portail des logiciels de recherche @CANARIE	registre	https://science.canarie.ca/researchsoftware/home/main.html	Canada	toutes	registre, indexage, durabilité	
rOpenSci	initiative communautaire	https://ropensci.org	monde	toutes	général	
RSE Assoc of Aust/NZ (RSE-aunz)	initiative communautaire	https://rse-aunz.github.io/	Australie/ Nouvelle-Zélande	toutes	général	
RSE Assoc of Belgium (be-RSE)	initiative communautaire	https://www.be-rse.org/	Belgique	toutes	général	
ScanR	agrégateur	https://scanr.enseignementsup-recherche.gouv.fr/	France	toutes	agrégateur, indexage, général	
SciCodes	initiative communautaire	https://scicodes.net/	monde	toutes	registres, dépôts, pratiques exemplaires, reproductibilité, préservation, découvrabilité, citation	
Science Code Manifesto	initiative communautaire, projets	http://sciencecodemano.org/ , https://github.com/codeissscience/manifesto	monde	toutes	citation, conservation, licences, pratiques exemplaires	Barnes, N. (2010). Science code manifesto. http://sciencecode manifesto.org
Science Gateways catalogue	registre, catalogue	https://catalogue.sciencegateways.org/	É.-U.	toutes	portails scientifiques, général, réutilisation	

Science Gateways communauté Institute (SGCI)	initiative communautaire / projet	http://sciencegateways.org	É.-U.	toutes (portails de sciences)	général	
	initiative communautaire	https://asclnet.github.io/SWRegistreWorkshop/Products/Products.html	monde	toutes	registres, dépôts, normes de métadonnées, citation	
SE4Science	initiative communautaire	https://se4science.org/workshops/	monde	toutes	science ouverte, durabilité	
Software Heritage	plateforme	https://www.softwareheritage.org	monde	toutes	conservation, citation	
Software Impacts	revue	https://www.sciencedirect.com/journal/software-impacts	monde	toutes	citation, conservation	
Software préservation réseau (SPN)	initiative communautaire	https://www.softwarepreservationnetwork.org/	monde	toutes	conservation	
Software durabilité Institute (SSI)	initiative communautaire / projet	https://www.software.ac.uk/	R.-U.	toutes	général	
SoftwareX	revue	https://www.sciencedirect.com/journal/softwarex	monde	toutes	examen des pairs, reproductibilité, citation	libre accès
SORSE (Series of Online Research Software événements)	événements	https://sorse.github.io	monde	toutes	renforcement de la communauté, collaboration	
swMATH	registre, initiative communautaire	https://swmath.org/	monde	mathématique	registre, libre accès, découvrabilité, documentation, publication connexe	
The Carpentries	entité à but non lucratif	https://carpentries.org/	monde	toutes	développement de logiciels	
The EinsteinPy projet	projet	https://einsteinpy.org	Inde, Europe	physique de la gravité, relativité générale	durabilité, productivité, conservation	
UKRI (conseils de recherche du R.-U.) groupe d'experts en	initiative gouvernementale	https://www.ukri.org/	R.-U.	toutes	durabilité, productivité	groupe d'experts en infrastructure électronique,

infrastructure électronique, boursiers RSE, etc						boursiers en génie logiciel de recherche ("RSE"), etc.
URSSI	initiative communautaire/gouvernementale	http://urssi.us/	É.-U.	toutes	durabilité	
US-RSE Association	initiative communautaire	https://us-rse.org/	É.-U.	toutes	général	
WSSSPE (Working towards Sustainable Software for Science: Practice and Experiences)	initiative communautaire, événements	http://wssspe.researchcomputing.org.uk	monde	toutes	général	
Neuroimaging Tools & Resources Collaboratory (NITRC)	communauté, plateforme, initiative, infrastructure, registre	https://www.nitrc.org/	É.-U.	neuroinformatique	renforcement de la communauté, collaboration, découverte, libre accès, pratiques exemplaires, dépôts, gestionnaires de paquets, reproductibilité	
Whole Tale	projet, plateforme, infrastructure	https://wholetale.org/	É.-U.	toutes	code source libre, reproductibilité, collaboration, découverte, édition, partage, réutilisation	