



2024-01-18

Rapport du sondage sur les logiciels de recherche de l'Alliance



Digital Research
Alliance of Canada

Alliance de recherche
numérique du Canada



Table des matières

Collaboratrices et collaborateurs	2
Principales conclusions.....	3
Renseignements généraux et contexte	5
Les répondantes et répondants	5
Les logiciels de recherche en tant que produit de la recherche	8
L'utilisation de logiciels de recherche	9
Soutien à l'utilisation des logiciels de recherche	13
Financement de l'utilisation de logiciels de recherche	16
Le développement de logiciels de recherche.....	17
Types de logiciels de recherche.....	23
Participation au développement de logiciels de recherche	24
Soutien au développement de logiciels de recherche.....	25
Financement du développement de logiciels de recherche	26
La pérennité des logiciels de recherche	30
Conclusion	31

A decorative background on the left side of the page, featuring a dark grey diagonal shape, a yellow horizontal band, and a teal triangular shape at the bottom. A woman wearing a VR headset is visible in profile, looking to the right.

Cliquez ou appuyez
ici pour saisir du
texte.

Cliquez ou appuyez ici pour saisir du texte.



Collaboratrices et collaborateurs

Le sondage sur les logiciels de recherche a été élaboré par le groupe de travail (GT) sur la stratégie pour les logiciels de recherche (LR) de l'Alliance :

Girma Bitsuamlak (Université Western)
Susan Brown (Université de Guelph)
Liseanne Cadieux (Alliance)
Brian Corrie (Université Simon-Fraser)
Carolyn Côté-Lussier (Institut national de la recherche scientifique)
Fares Dhane (Alliance)
Félix-Antoine Fortin (Université Laval)
Kimberley Hartley (Alliance)
Mark Leggott (Alliance)
Tamanna Moharana (Alliance)
Pierre-Olivier Quirion (McGill)
Y.G. Rancourt (Alliance)
Seppo Sahrakorpi (Alliance)
Abdel Yousif (Université de Calgary)
Qian Zhang (Alliance)

Auteurs du rapport :

Fares Dhane (Alliance)
Brian Corrie (Université Simon-Fraser)
Felipe Pérez-Jvostov (Alliance)



Principales conclusions

Développement de logiciels de recherche

Souvent, les équipes chargées du développement de logiciels de recherche regroupent différents rôles et les personnes étudiantes au doctorat sont les principales développeuses, notamment en sciences et génie (SG) et en recherche en santé (RS).

La taille et la composition des équipes varient, les répondantes et répondants en sciences sociales et humaines (SSH) et en RS ayant généralement moins d'expérience dans le développement de logiciels que leurs homologues en SG.

La maîtrise de la programmation et du codage varie également d'un domaine à l'autre, les répondantes et répondants en SG ayant généralement une expertise approfondie.

Les répondantes et répondants en SSH consacrent moins de temps à développer des logiciels que les chercheuses et chercheurs en SG et en RS.

Les répondantes et répondants en SG allouent plus de temps au développement de logiciels de recherche que leurs homologues d'autres domaines.

Si le contrôle des versions est monnaie courante, on constate une variation importante dans le respect des normes et des pratiques exemplaires. Une documentation informatisée appropriée demeure une préoccupation.

Partage et publication de logiciels de recherche

Les répondantes et répondants partagent des logiciels principalement sur des plateformes comme GitHub, mais les pratiques de documentation ne sont pas uniformes.

Plusieurs répondantes et répondants disposent de publications faisant référence à leur logiciel de recherche, mais la connaissance des principes de développement de logiciels de recherche est très variable.

La majorité des répondantes et répondants ont reçu une formation sur les logiciels, principalement par autoapprentissage et en ligne.

Il y a un intérêt pour la formation, mais moins de motivation pour les compétences liées à la pérennité; la participation aux associations de génie logiciel est faible, surtout en SSH.

Soutien et financement pour le développement de logiciels

L'accès au soutien pour le développement de logiciels a été considéré comme important dans tous les domaines, mais beaucoup de répondantes et répondants ne l'ont pas.



Le financement est devenu un obstacle important à la pérennité des logiciels de recherche dans tous les domaines. L'appui des établissements est essentiel pour le développement de logiciels de recherche, mais le financement réservé à cette fin est limité.

La majorité des répondantes et répondants s'entendent qu'il est important de faciliter la réutilisation, la reproductibilité et la science ouverte, mais il n'y a pas d'entente claire sur les rôles et les responsabilités en matière de pérennité des logiciels.

Préservation des logiciels et défis

La majorité des répondantes et répondants connaissent des dépôts courants comme GitHub, mais ce n'est pas aussi uniforme entre les domaines pour d'autres dépôts comme Zenodo ou DockerHub.

Les difficultés d'accès et d'utilisation des logiciels préservés comprennent des contraintes financières et de temps, un manque d'expertise et la méconnaissance de certains dépôts.

L'examen du code par les pairs et les services de soutien pour la documentation et la pérennité des logiciels ont été jugés utiles.



Renseignements généraux et contexte

Les logiciels de recherche jouent un rôle crucial dans la recherche universitaire moderne axée sur les données. Il n'est pas surprenant qu'ils soient un pilier fondamental de l'infrastructure de recherche numérique (IRN), car ils constituent une condition sine qua non à l'analyse et à la manipulation des données et permettent d'optimiser l'utilisation des ressources informatiques par les chercheuses et chercheurs. Les logiciels de recherche sont également de plus en plus reconnus comme des produits de la recherche, dans la mesure où, à l'instar des données, ils peuvent être partagés pour accroître leur réutilisation, leur reproductibilité et le développement du savoir. Pourtant, même avec un examen global de l'IRN visant à comprendre comment les chercheuses et chercheurs du Canada développent, utilisent et financent leurs logiciels de recherche, il n'a pas été possible de mesurer l'évolution rapide des pratiques ni le degré d'adoption dans différentes communautés de recherche. Pour combler ce manque, l'Alliance de recherche numérique du Canada (l'Alliance) a mené un sondage exhaustif auprès de la communauté de la recherche du Canada afin de mieux comprendre les lacunes et les forces existantes dans le contexte canadien.

Le sondage était accessible du 10 mai au 8 juin 2020. Le présent rapport en décrit les principaux résultats et se divise ainsi : les répondantes et répondants, les logiciels de recherche en tant que produit sous-estimé de la recherche, l'utilisation des logiciels de recherche, le développement des logiciels de recherche et la pérennité des logiciels de recherche. Pour faciliter la lecture et le recoupement avec le sondage, les références à des questions précises sont entre parenthèses et en italique (p. ex. [B3]).

Les répondantes et répondants

Au total, 548 chercheuses et chercheurs ont répondu au sondage. Dans ce contexte, il s'agit de personnes qui ont participé au sondage en répondant à au moins une question, qu'elles aient rempli ou non l'entièreté du sondage (152 chercheuses et chercheurs ont répondu sondage complet). Chaque question a reçu en moyenne 181 réponses. Le corps professoral universitaire représentait 38 % de tous les répondantes et répondants (corps professoral – professeur(e), corps professoral – autre), suivis du personnel de recherche (12 %), de l'administration (12 %) ainsi que des ingénieures et ingénieures/spécialistes en logiciels de recherche (8 %) [B3]. Le pourcentage des autres répondantes et répondants et de leurs rôles est indiqué dans la figure 1 [B3]. Des répondantes et répondants, 71 % ont déclaré occuper un poste permanent à temps plein, tandis que 22 % ont un contrat à durée déterminée, 3 %, un poste à temps partiel et 3 %, un autre type de poste [B5].

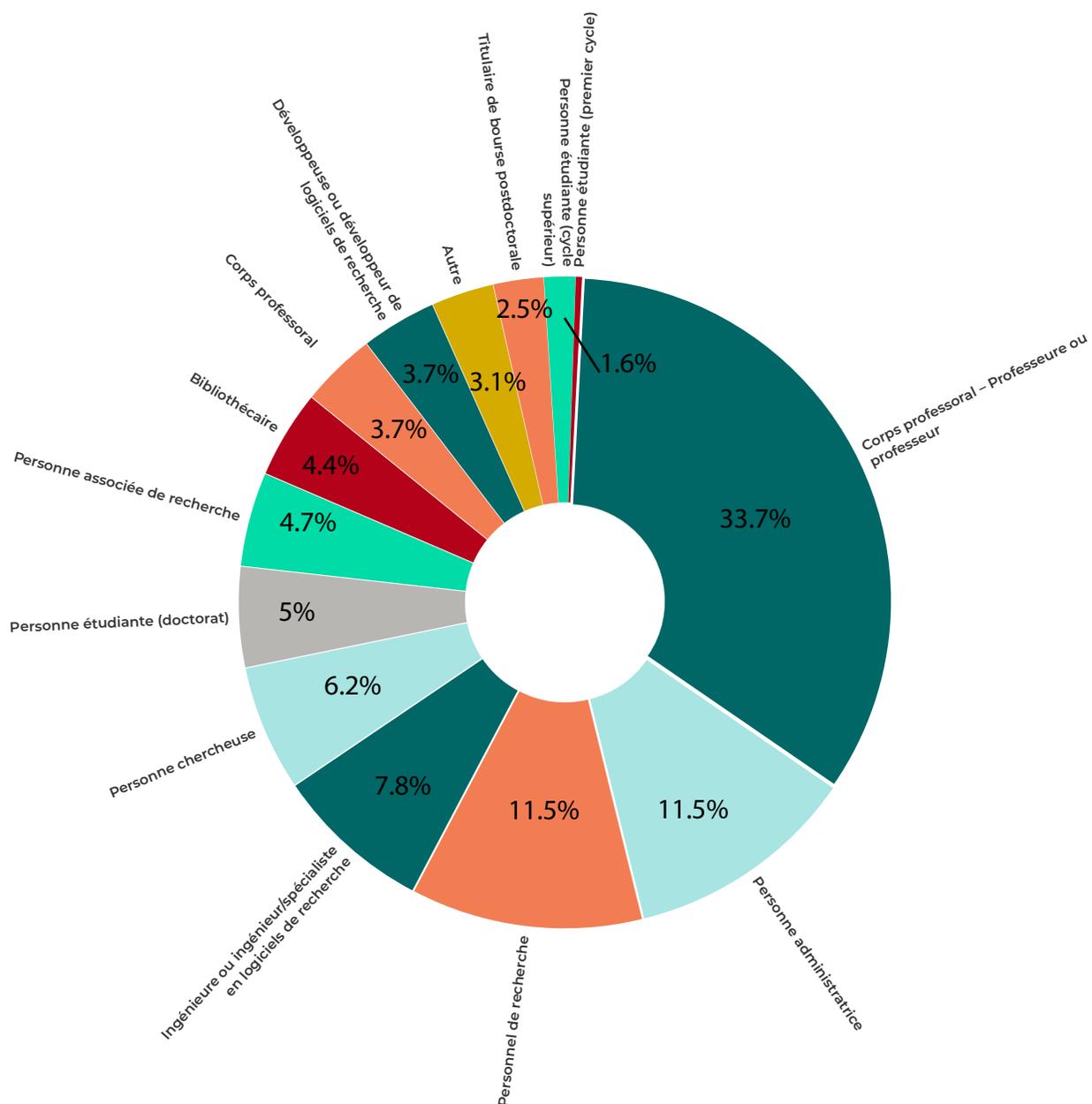


Figure 1. Question [B3]. Répartition des répondantes et répondants par rôle.

On a invité les participantes et participants à s'identifier sur la base de la [Classification canadienne de la recherche et développement \(version 1, 2020\)](#) [B2]. On note une diversité de la participation dans les divisions : 30 % en sciences naturelles, 24 % en sciences médicales, de la santé et de la vie, 17 % en génie et technologies, 14 % en sciences humaines et arts, 12 % en sciences sociales et 3 % en sciences agricoles et vétérinaires (figure 2). Pour faciliter l'interprétation et la présentation du présent rapport, ces six catégories ont été regroupées en trois grands groupes. Cela devrait permettre au lectorat de se pencher sur les tendances générales plutôt que sur les différences intrinsèques entre les divisions. Les trois grandes



catégories sont les suivantes : *sciences et génie* (SE), qui comprennent les sciences naturelles, le génie et les technologies, et les sciences agricoles et vétérinaires; *sciences sociales et humaines* (SSH), qui comprennent les sciences humaines et les arts, et les sciences sociales; et *recherche en santé* (RS), qui comprend les sciences médicales, de la santé et de la vie. La répartition des répondantes et répondants dans les trois catégories consolidées est présentée dans la figure 3.

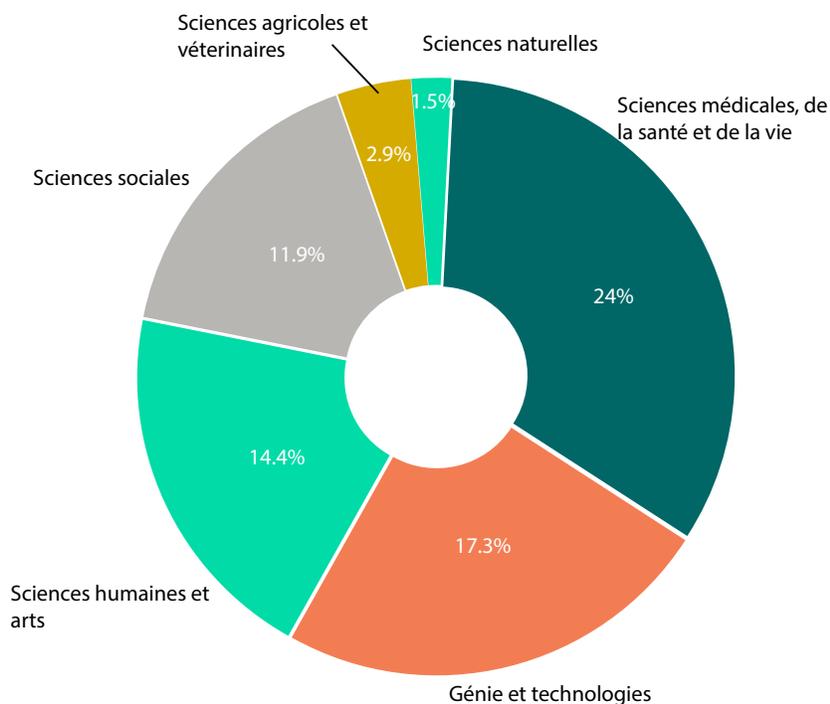


Figure 2. Répartition des répondantes et répondants selon la Classification canadienne de la recherche et développement.

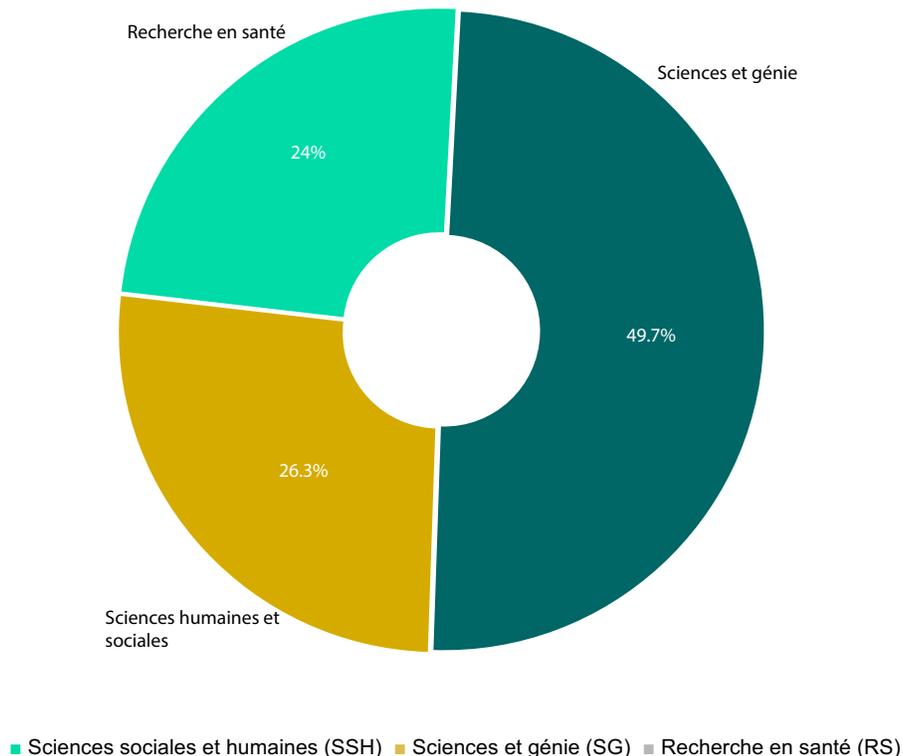


Figure 3. Répartition des répondantes et répondants dans les trois domaines de recherche regroupés.

Les logiciels de recherche en tant que produit de la recherche

L'une des premières questions du sondage portait sur le concept de logiciels de recherche et visait à déterminer si la définition actuellement utilisée par l'Alliance concorde avec la compréhension que les chercheuses et chercheurs ont des logiciels de recherche [C1]. La définition fournie était la suivante : « *Les logiciels de recherche sont définis dans une perspective large comme une intersection émergente et complexe d'outils, de disciplines, de services, de plateformes, de matériel, de ressources et de personnes (utilisatrices et utilisateurs, chercheuses et chercheurs, développeuses et développeurs, parties prenantes, personnel et communautés, etc.) qui les utilisent et y contribuent.* » La majorité des personnes répondantes (65 %) étaient d'accord avec la définition fournie, mais beaucoup ne l'étaient pas (35 %), soit parce qu'elles la trouvaient trop vague, soit parce qu'elles étaient elles-mêmes incertaines. Pour certaines, les logiciels de recherche englobent tout outil logiciel utilisé dans leurs activités universitaires, tandis que d'autres les associent principalement au code développé pour leur recherche. Il importe donc de tenir compte de ces divergences d'interprétation à la lecture des conclusions générales tirées des résultats et présentées dans le présent rapport.



La majorité des répondantes et répondants (91 %) considèrent l'utilisation de logiciels de recherche comme essentielle à leur recherche [C2]. Cinquante pour cent des répondantes et répondants considéraient le développement de logiciels de recherche comme un produit principal de leur recherche [C3], avec des différences marquées entre les domaines (SSH – 36 %; SG – 59 %; RS – 48 %). Parmi les répondantes et répondants qui considèrent ces logiciels comme un produit principal de la recherche, 43,4 % étaient du corps professoral – professeur(e), suivis des ingénieures et ingénieurs/spécialistes en logiciels de recherche (10 %). Ces chiffres concordent avec des résultats internationaux similaires : 92 % des chercheuses et chercheurs au Royaume-Uni et 95 % aux États-Unis dépendent des logiciels de recherche¹. Le sondage mené au Royaume-Uni a également révélé que 56 % des chercheuses et chercheurs au pays développent leur propre logiciel de recherche². Cela souligne l'importance de ces logiciels dans la recherche moderne et indique qu'ils jouent un rôle semblable au Canada et dans d'autres pays.

L'utilisation de logiciels de recherche

Beaucoup de répondantes et répondants (85 %) ont indiqué avoir un accès facile à un logiciel de recherche ou savoir déjà quel logiciel répond à leurs besoins [E2]. Ainsi, seule une petite proportion de leur temps (10 % ou moins) est consacrée à la recherche de logiciels. Cependant, la manière dont les chercheuses et chercheurs trouvent les logiciels qu'elles et ils utilisent varie d'une discipline à l'autre. Par exemple, les chercheuses et chercheurs en SG et en RS s'appuient principalement sur des publications scientifiques et leurs compléments, tandis que les chercheuses et chercheurs en SSH s'appuient davantage sur leur communauté de spécialistes, les ateliers et les cours ainsi que les conférences [E1] (figure 4). Les chercheuses et chercheurs en SG et RS utilisent également des plateformes d'hébergement de code source telles que GitHub pour trouver des logiciels, avec respectivement 37 % et 32 % des répondantes et répondants de ces disciplines qui les utilisent. Cependant, seulement 14 % des chercheuses et chercheurs en sciences sociales et humaines ont dit utiliser GitHub.

¹ <https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers>, <https://zenodo.org/record/1183562>; Nangia et al., « Track 1 Paper: Surveying the U.S. National Postdoctoral Association Regarding Software Use and Training in Research », dans *Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE 5.1)*, 2017, <https://doi.org/10.6084/m9.figshare.5328442.v3>.

² <https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers>, <https://zenodo.org/record/1183562>.

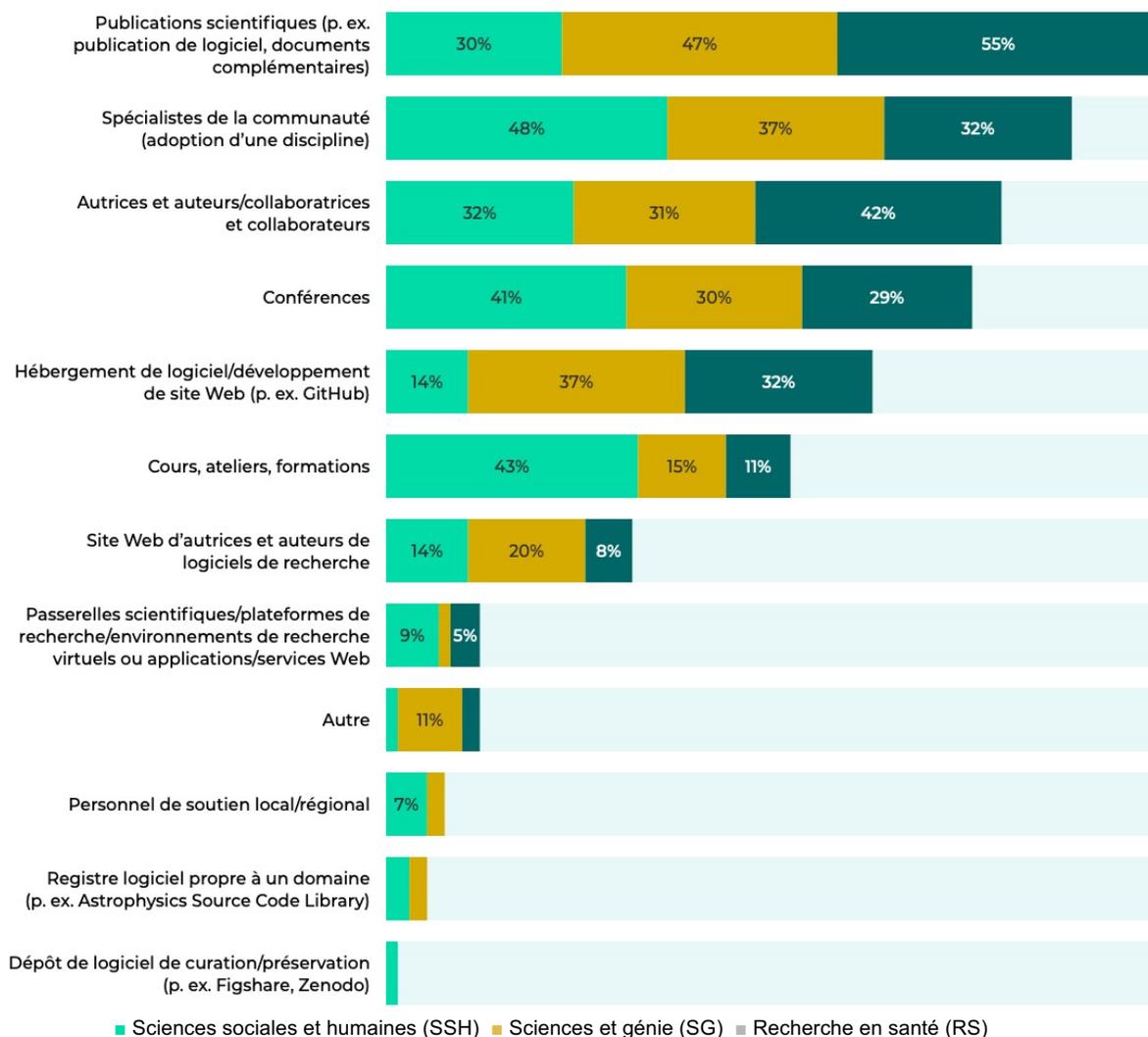


Figure 4. Question [E1] : Choisissez les trois moyens les plus fréquemment utilisés pour découvrir des logiciels de recherche développés par d'autres personnes et que vous intégrez dans vos propres activités de recherche. Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 30; RS : n = 28; SG : n = 84; répondantes et répondants uniques : 142.

Les facteurs qui motivent les chercheuses et chercheurs à utiliser un outil logiciel ou une plateforme en particulier sont très similaires dans tous les domaines et ont une importance relativement assez égale [E14] (figure 5). Il s'agit notamment des interfaces conviviales et de l'efficacité, du coût (y compris les logiciels libres), de la pérennité et de la documentation. Toutefois, l'efficacité était un facteur moins important en SSH (55 %) qu'en SG (75 %) et en RS (82 %).

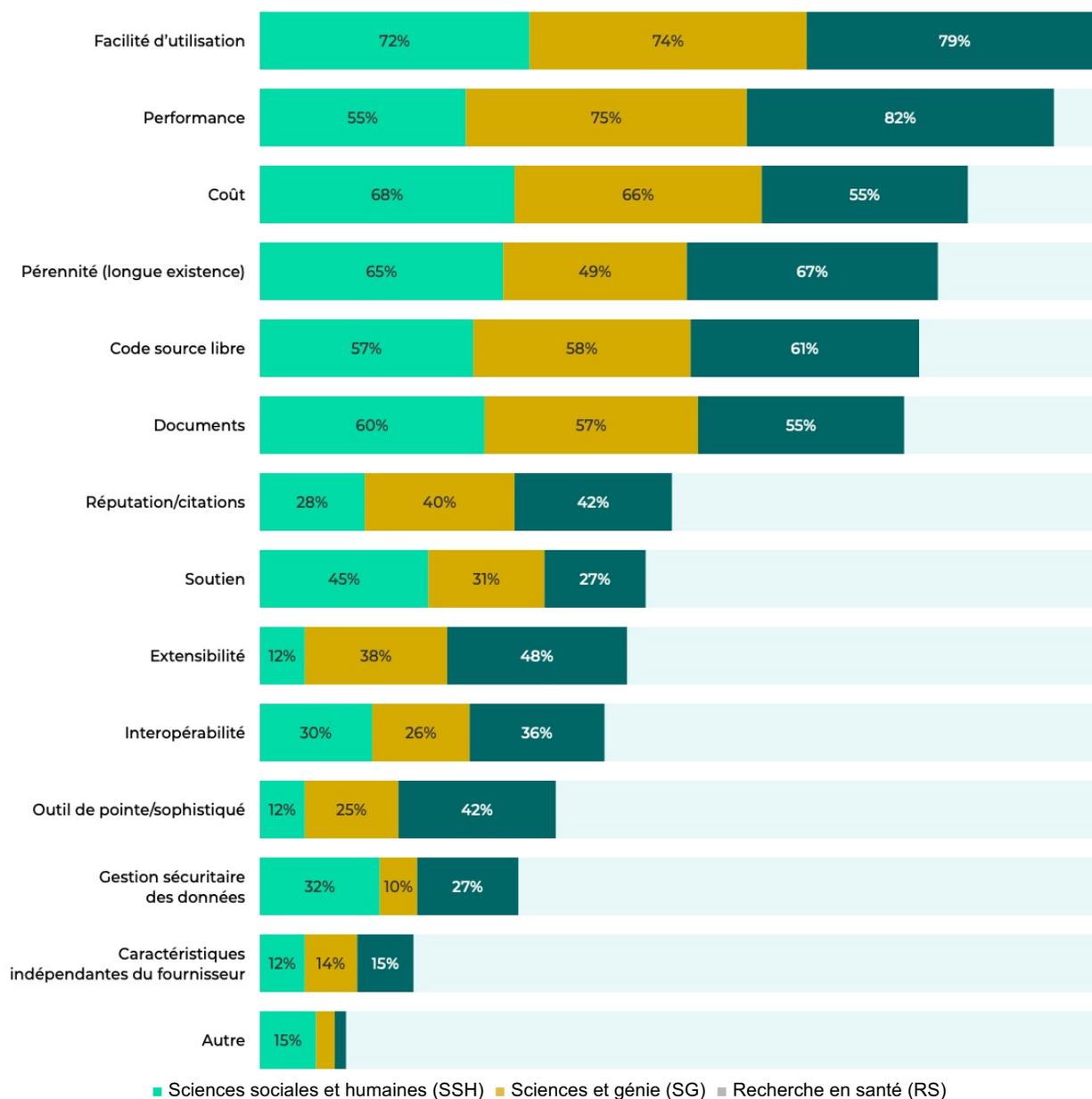


Figure 5. Question [14] : Quels facteurs prenez-vous en compte dans le choix d'un logiciel de recherche? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 40; RS : n = 33; SG : n = 77; répondantes et répondants uniques : 150.

Les répondantes et répondants ont signalé un éventail de difficultés techniques à l'utilisation de logiciels d'autres chercheuses et chercheurs [G1], notamment un manque de documentation (SSH – 59 %; SG – 79 %; RS – 74 %), des difficultés à trouver un logiciel (SSH – 56 %; SG – 38 %; RS – 45 %), des problèmes d'installation (SSH – 33 %; SG – 45 %; RS – 58 %), des problèmes de compatibilité (SSH – 41 %; SG – 34 %; RS – 45 %), un manque de formation (SSH – 54 %; SG – 27 %; RS – 39 %) et des difficultés à maintenir le logiciel (SSH 1 – 8 %; SG –



26 %; RS – 42 %); un petit nombre (moins de 13 %) a aussi signalé d'autres difficultés. S'ajoutent des difficultés non techniques [G2], notamment le manque de documentation (SSH – 51 %; SG – 68 %; RS – 71 %), le manque de soutien/reproductibilité à long terme (SSH – 59 %; SG – 49 %; RS – 61 %), le manque de fonctionnalités (SSH – 46 %; SG – 49 %; RS – 52 %), le manque de soutien (SSH – 16 %; SG – 25 %; RS – 52 %) et le manque de soutien linguistique (SSH – 16 %; SG – 7 %; RS – 13 %), ainsi qu'un petit nombre d'autres difficultés.

Les plateformes les plus courantes sur lesquelles les répondantes et répondants exécutent leurs logiciels de recherche sont principalement leur ordinateur portable ou de bureau (SSH – 97 %; SG – 88 %; RS – 68 %) et leurs ordinateurs de laboratoire (SSH – 47 %; SG – 54 %; RS – 59 %) [E3]. Les ressources de calcul informatique de pointe (CIP) (SSH – 21; SG – 54 %; RS – 50 %), les ressources infonuagiques (SSH – 26 %; SG – 39 %; RS – 35 %) et les ressources de nuage communautaire (SSH – 24 %; SG – 25 %; RS – 26 %) sont aussi couramment utilisées (figure 6). De même, ce sont les ordinateurs portables ou de bureau (SSH – 82 %; SG – 74 %; RS – 55 %) et les ordinateurs de laboratoire (SSH – 45 %; SG – 51 %; RS – 50 %) que les chercheuses et chercheurs utilisent principalement pour accéder aux logiciels de recherche. Les interfaces à ligne de commande à distance (SSH – 27 %; SG – 62 %; RS – 53 %) ainsi que les passerelles scientifiques/environnements de recherche virtuels (SSH – 20 %; SG – 32 %; RS – 37 %) sont largement utilisés dans tous les domaines, mais la première catégorie est beaucoup plus courante en SG [E4]. Parmi les répondantes et répondants, 44 % ont déclaré utiliser un logiciel de recherche installé et maintenu sur des plateformes de CIP gérées et prises en charge par l'Alliance [E5].

Les répondantes et répondants ont indiqué que leurs flux de travail avec les logiciels de recherche comprenaient une gamme d'outils comme des logiciels de contrôle de version (p. ex., GitHub : SSH – 43 %; SG – 67 %; RS – 58 %), des fournisseurs de stockage commerciaux (p. ex., Dropbox : SSH – 59 %; SG – 43 %; RS – 58 %), des logiciels de données libres (p. ex., NumPy : SSH – 36 %; SG – 59 %; RS – 47 %) et des outils de bureau (p. ex., MS Office : SSH – 61 %; SG – 27 %; RS – 37 %). D'autres outils ont également été mentionnés, mais ils étaient beaucoup moins utilisés (c.-à-d. <25 %), entre autres les plateformes d'intégration continue (p. ex., Travis-CI), les plateformes de collaboration infonuagique (p. ex., OSF), les outils de gestion de projet (p. ex., Asana), les logiciels de données commerciales (p. ex., Tableau), les logiciels de saisie de données sécurisée (p. ex., Redcap) et les blocs-notes de laboratoire électronique [E8].

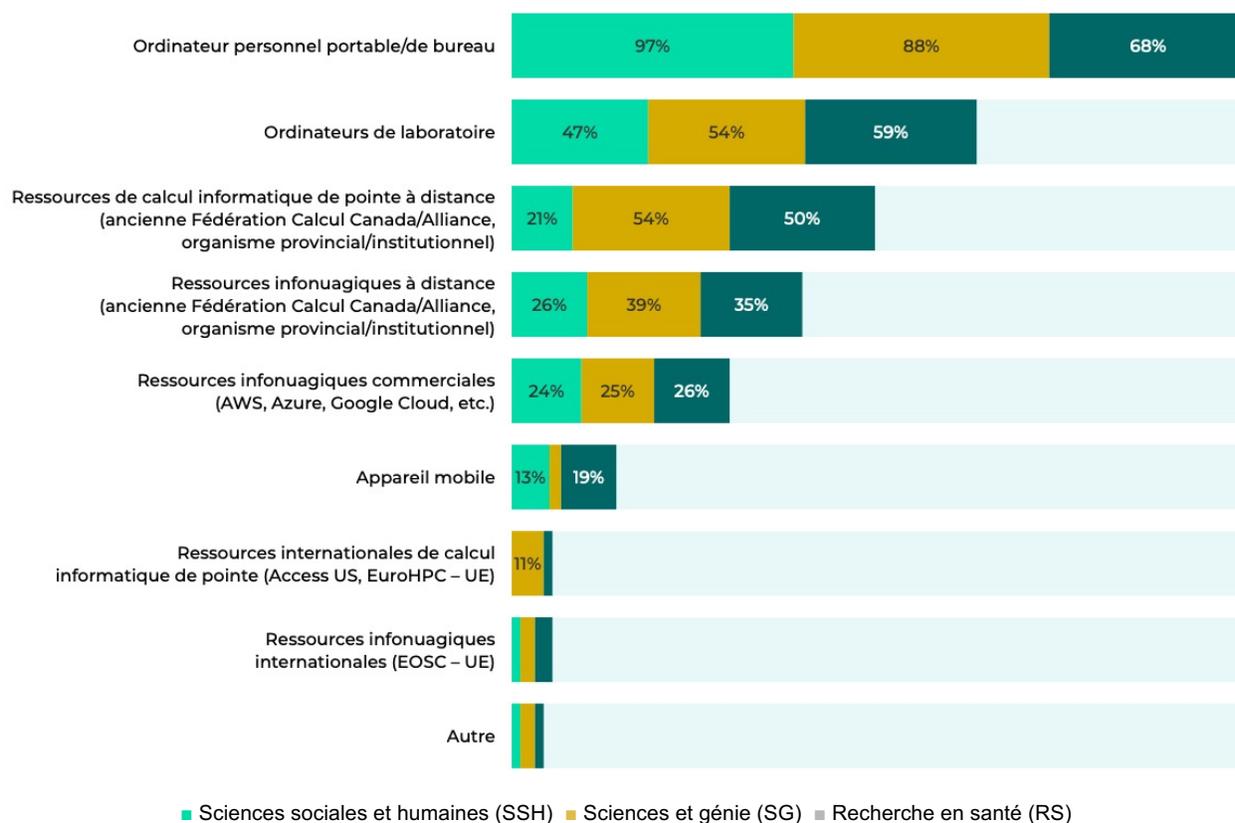


Figure 6. Question [E3] : Sur quelle(s) plateforme(s) utilisez-vous des logiciels de recherche? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 44; RS : n = 38; SG : n = 91; répondantes et répondants uniques : 173.

Lorsqu’interrogés sur le sujet, 76 % des répondantes et répondants ont dit citer le logiciel de recherche qu’elles ou ils utilisaient [E15]. De ce nombre, la moitié cite la publication décrivant le logiciel de recherche tandis que l’autre cite directement le logiciel de recherche [E16].

Un nombre limité de répondantes et répondants (29 %) ont déclaré avoir utilisé des services de gestion de la recherche fournis et soutenus par l’Alliance, comme l’Assistant PGD et la CCDB [E6]. De même, un petit nombre de répondantes et répondants (18 %) ont déclaré utiliser des plateformes de logiciels de recherche soutenues par l’Alliance [E7].

Soutien à l’utilisation des logiciels de recherche

Plus de la moitié des répondantes et répondants (SSH – 50 %; SG – 57 %; RS – 55 %) ont dit participer directement au soutien des logiciels de recherche qu’elles ou ils utilisent [E9] (c.-à-d. celles et ceux qui ont indiqué « moi » comme responsable du soutien des logiciels de recherche utilisés). Près de la moitié de ces personnes en SG et en SSH étaient des professeures et



professeurs (49 % et 43 % respectivement), ce qui est beaucoup plus élevé que leur représentation globale parmi les répondantes et répondants (~33 %). En revanche, les professeures et professeurs en RS ne représentent que 25 % des répondantes et répondants qui soutiennent l'utilisation de logiciels de recherche (figure 7). Cela peut sous-entendre qu'en SG et en SSH, les professeures et professeurs jouent un rôle prépondérant dans le soutien aux logiciels de recherche, plus encore que les ingénieures et ingénieurs et les développeuses et développeurs de logiciels de recherche. En RS, cette tâche est répartie de manière plus équilibrée entre différents rôles (figure 8). Le nombre d'employées et employés à temps plein (ETP) chargés de l'installation et de la gestion des logiciels est toutefois resté faible dans tous les domaines : 53 % des répondantes et répondants ont déclaré consacrer moins de ¼ d'ETP à ces tâches, et seulement 18 % ont indiqué que leurs équipes consacraient plus d'une ou un ETP à la gestion des logiciels [E10].

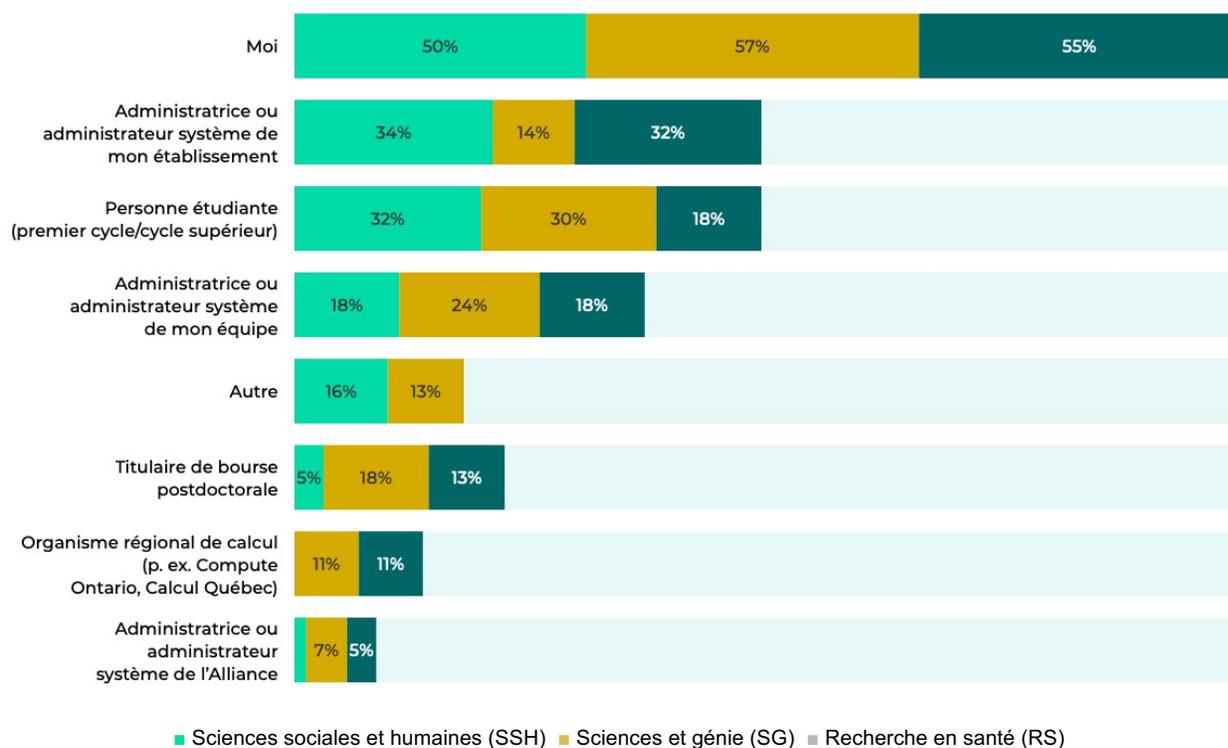


Figure 7. Question [E9] : Qui, dans votre équipe, assure le soutien des logiciels de recherche que vous utilisez habituellement? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 44; RS : n = 38; SG : n = 91; répondantes et répondants uniques : 173.

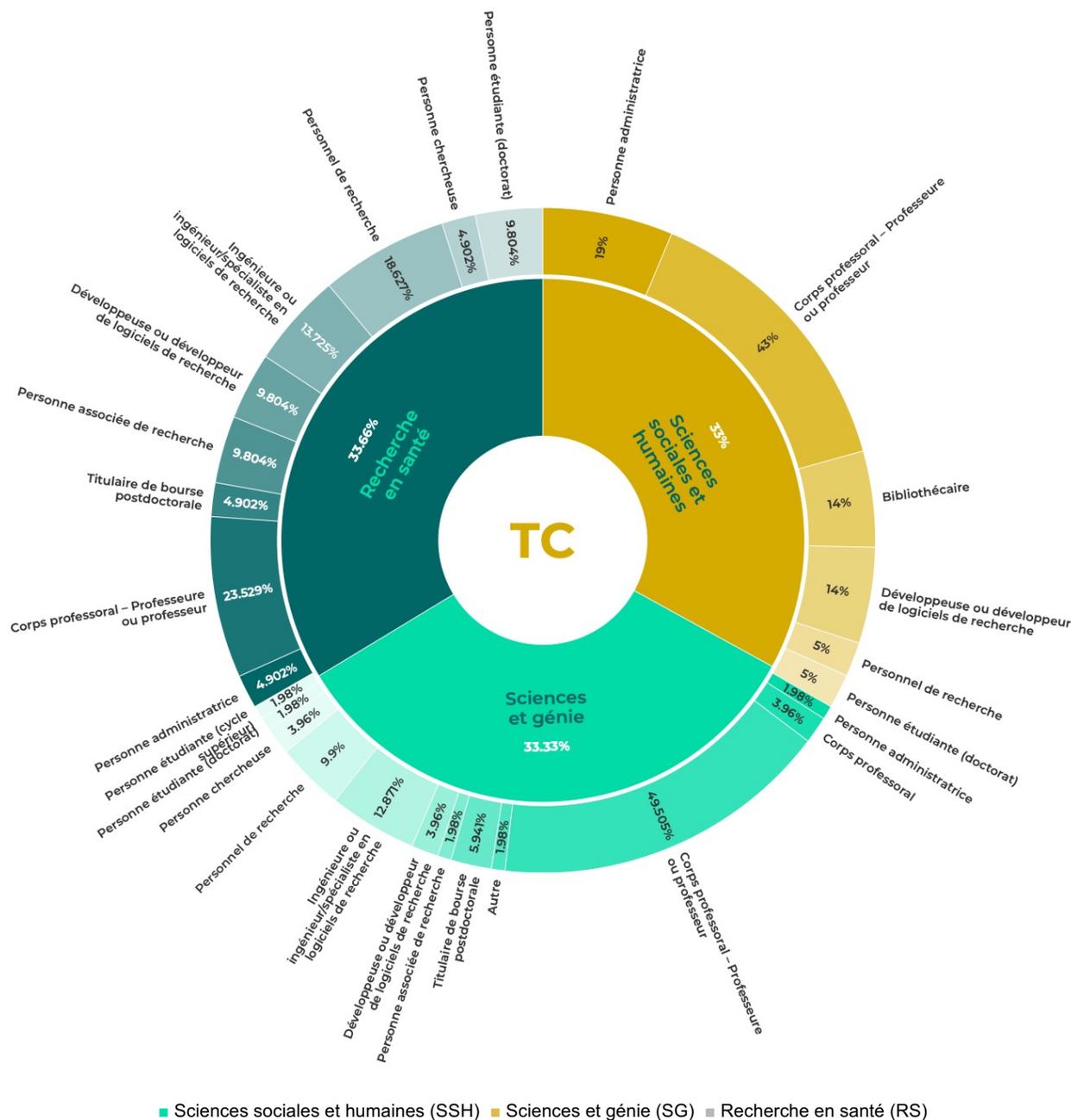


Figure 8. Question [E9] : Qui, dans votre équipe, assure le soutien des logiciels de recherche que vous utilisez habituellement? Répartition des rôles parmi les personnes ayant répondu « Moi » à la question E9.

La participation active des chercheuses et chercheurs au soutien des logiciels de recherche n'est pas sans embûches. Près de la moitié des répondantes et répondants ont déclaré n'avoir parfois pas pu réutiliser les logiciels de recherche d'autres personnes en raison du manque de ressources pour les installer, les utiliser ou en assurer le soutien [E18]. D'autres facteurs empêchent l'utilisation des logiciels de recherche, notamment le financement, l'expertise, le manque de documentation et les systèmes de soutien nécessaires pour les installer, les utiliser



et les maintenir efficacement [E19]. Ces défis montrent l'importance de la participation active des chercheuses et chercheurs et de la mise à disposition de ressources adéquates pour la réutilisation des logiciels de recherche.

En ce qui a trait au type de données stockées et analysées par les répondantes et répondants, 21 % ont déclaré que des données étaient stockées/recueillies par/sur les communautés autochtones, 45 % ont déclaré que certaines données stockées/recueillies impliquaient des renseignements personnels, et 28 % ont déclaré que certaines données stockées/recueillies nécessitaient un niveau élevé de sécurité pour d'autres raisons.

Financement de l'utilisation de logiciels de recherche

Les répondantes et répondants couvrent principalement les coûts d'utilisation des logiciels de recherche avec leurs subventions de recherche [E11], bien que le financement des établissements soit encore plus important pour les chercheuses et chercheurs en SSH. En SG, le financement par des subventions de recherche était toutefois plus important que le financement des établissements (figure 9).

Les chercheuses et chercheurs en SSH déclarent plus souvent utiliser des fonds personnels pour des logiciels de recherche (31 %) que celles et ceux en SG (13 %) et en RS (10 %). Parmi les répondantes et répondants qui ont indiqué le montant de leurs dépenses personnelles, 44 % en SSH, 30 % en SG et 19 % en RS dépensent entre 0 et 1 000 \$ pour leur logiciel de recherche [E12]. Ces résultats s'apparentent à ceux décrits pour l'infonuagique, où les chercheuses et chercheurs en SSH ont également utilisé des fonds personnels pour couvrir les dépenses liées à l'infonuagique commerciale (voir le [Rapport sur le sondage de l'Alliance sur l'infonuagique 2023](#)).

En ce qui concerne les dépenses des groupes de recherche, on observe la même tendance : 23 % des répondantes et répondants en SG, 19 % en RS et 38 % en SSH ont dépensé entre 1 et 1 000 \$ pour l'achat, l'installation ou la gestion de ressources de logiciels de recherche ou l'obtention de licences logicielles, tandis que 13 % en SSH, 28 % en SG et 19 % en RS n'ont rien dépensé [E13]. Plusieurs répondantes et répondants ont déclaré avoir dépensé entre 5 000 et 50 000 \$ pour des logiciels de recherche (SSH – 10 %; SG – 5 %; RS – 13 %), tandis qu'un petit nombre a déclaré avoir dépensé plus de 50 000 \$ (SG – 5 %; RS – 3 %).

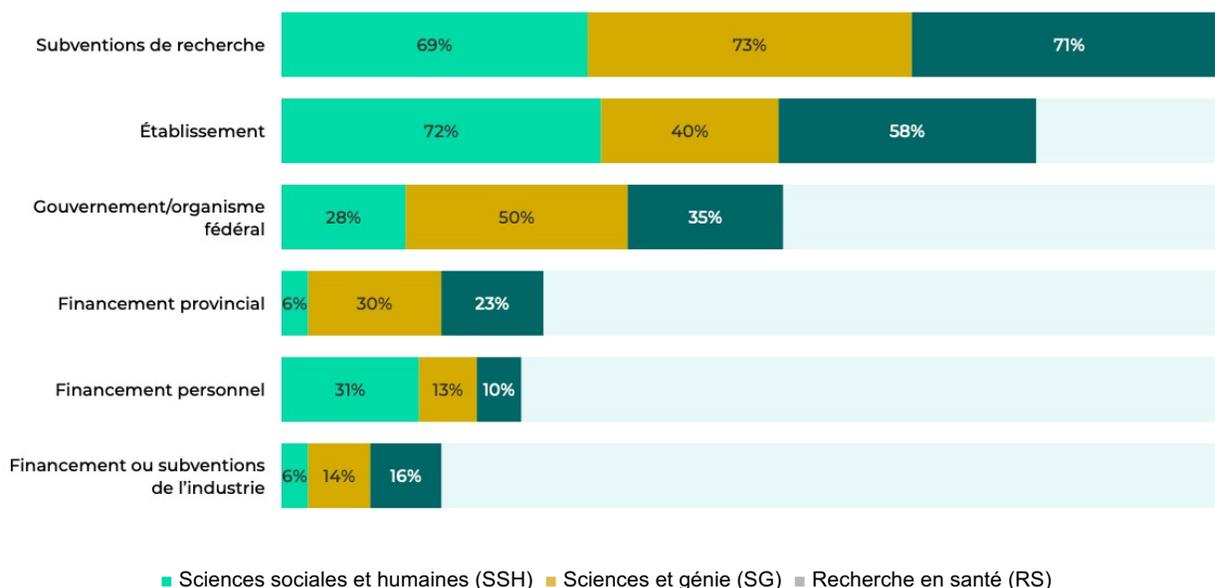


Figure 9. Question [E11] : Comment votre groupe finance-t-il les logiciels de recherche qu'il utilise? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 36; RS : n = 31; SG : n = 70; répondantes et répondants uniques : 137.

Le développement de logiciels de recherche

Interrogés sur leur expérience dans la direction de projets de logiciels de recherche, environ 52 % des répondantes et répondants, principalement des professeures et professeurs (26 %) et d'autres chercheuses et chercheurs (12 %), ont indiqué avoir dirigé ou diriger actuellement un projet de développement de logiciels de recherche [D1]. Il semble que les répondantes et répondants en SG participent davantage à de telles activités que celles et ceux d'autres domaines (ont précédemment dirigé – oui = SSH – 38 %; SG – 53 %; RS – 43 %) (dirigent actuellement – oui = SSH – 36 %; SG – 58 %; RS – 45 %). La majorité des répondantes et répondants (~70 %) jouent plusieurs rôles dans ces projets : conceptrices et concepteurs, chercheuses et chercheurs ou utilisatrices et utilisateurs finaux des logiciels de recherche développés [D5]. Les répondantes et répondants en SG (67 %) et en RS (56 %) participent plus activement au développement de logiciels de recherche qu'en SSH (38 %).

Dans les équipes de recherche, on semble se partager la tâche de développement des logiciels de recherche entre différents membres. Par exemple, les personnes étudiantes au doctorat sont les principales développeuses de logiciels en SG et en RS, probablement parce que leurs recherches les y obligent, alors que la proportion est de seulement 22 % en SSH. Comme pour le soutien aux logiciels de recherche, le corps professoral, en particulier en SG, joue un rôle de premier plan dans le développement de logiciels (SSH – 33 %; SG – 43 %; RS – 28 %). Les membres du personnel traditionnellement responsables du développement professionnel des logiciels de recherche, comme les développeuses et développeurs de logiciels de recherche ainsi



que les ingénieures et ingénieurs ou spécialistes en logiciels de recherche, jouent un rôle dans leurs équipes, mais dans une bien moindre mesure que les étudiantes et étudiants des cycles supérieurs et le corps professoral (figure 10).

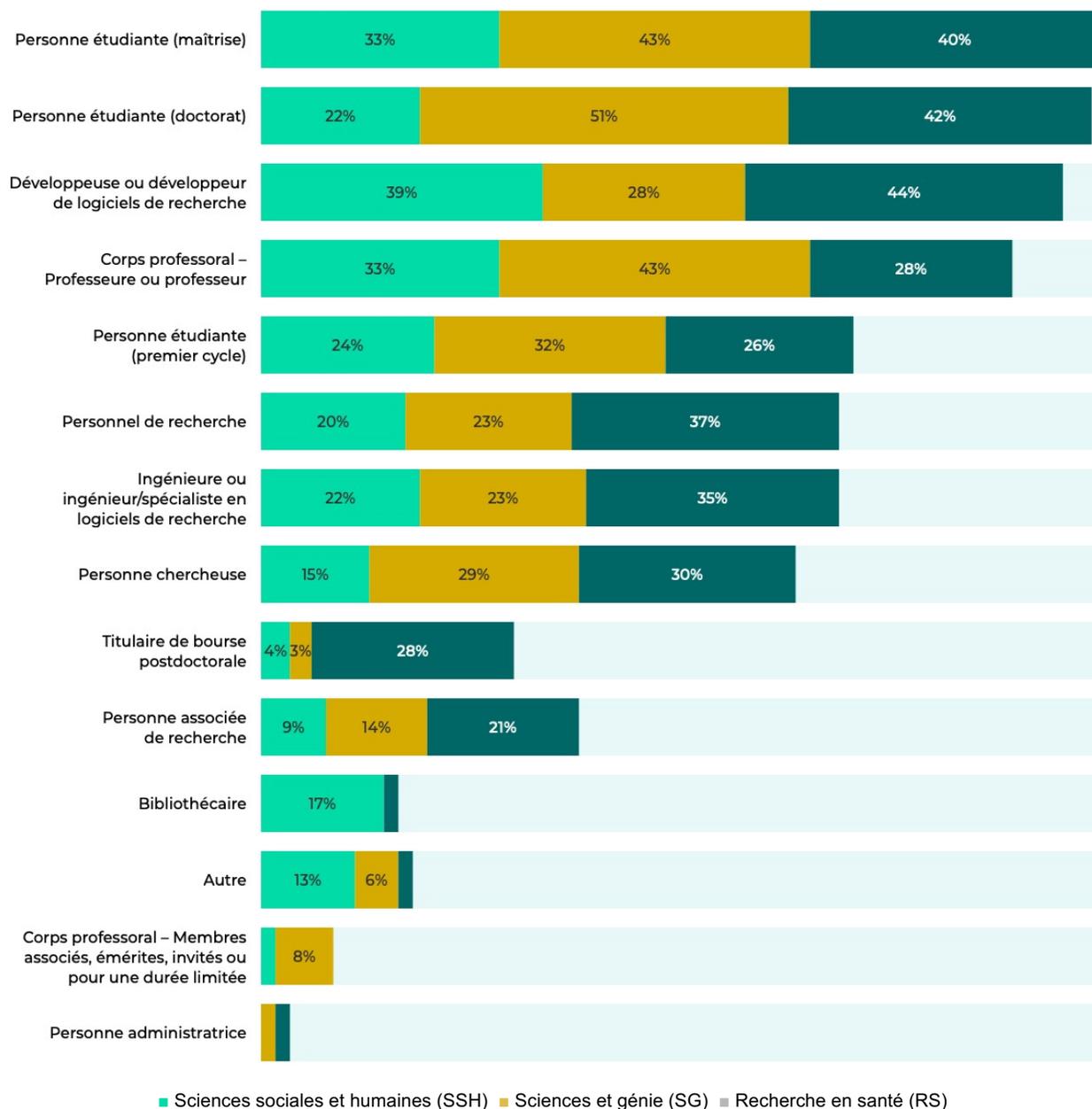


Figure 10. Question [D6] : Dans votre groupe, qui développe les logiciels? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 46; RS : n = 43; SG : n = 98; répondantes et répondants uniques : 165.

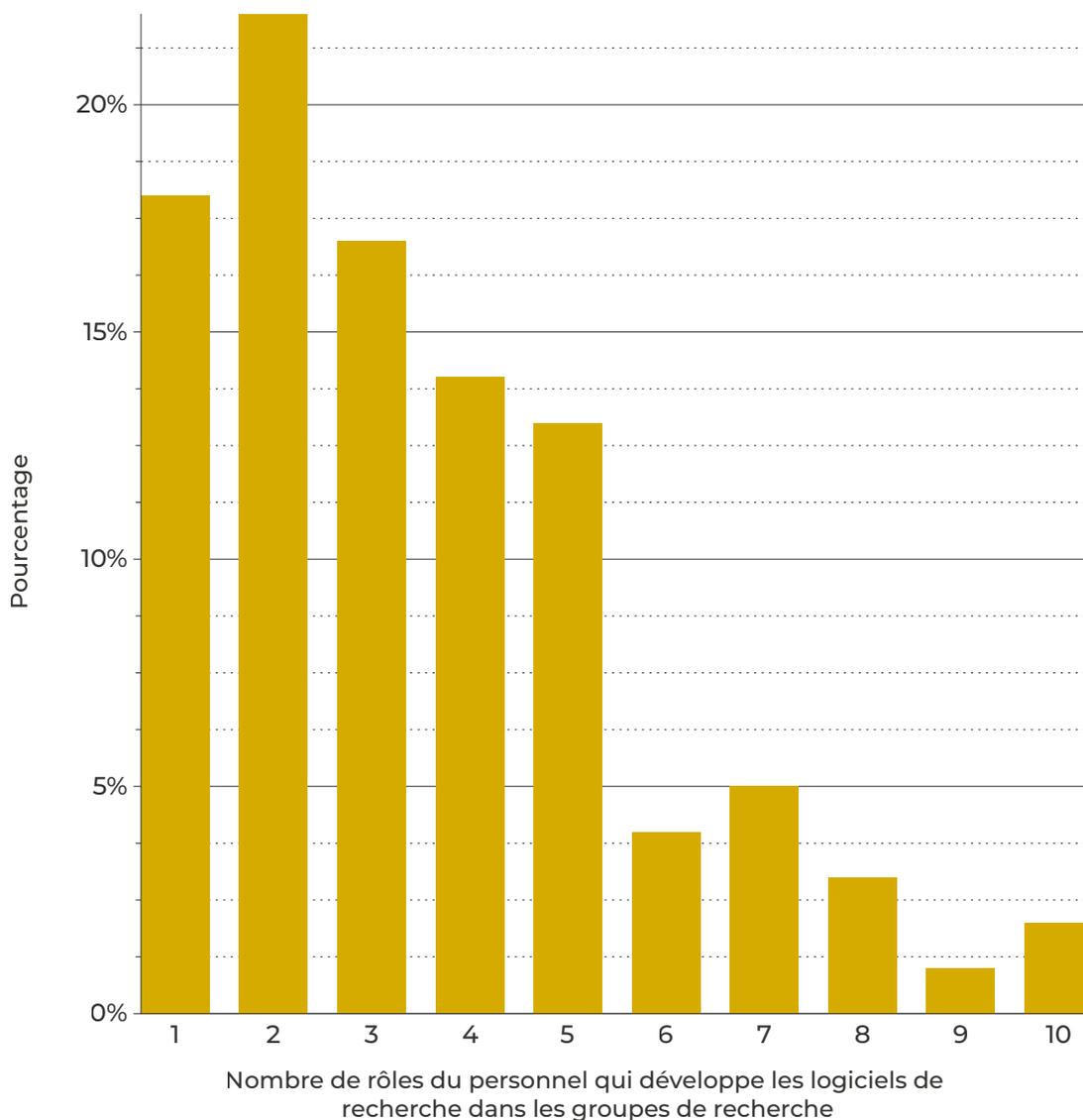


Figure 11. Diversité des rôles du personnel (p. ex., professeure ou professeur, ingénieure ou ingénieur en logiciels de recherche, étudiante ou étudiant de cycles supérieurs, etc.) qui développe les logiciels de recherche dans les groupes de recherche.

La taille et la composition des équipes de développement de logiciels de recherche varient considérablement d'un domaine à l'autre et au sein d'un même domaine. En moyenne, ces équipes étaient composées de 3,5 personnes occupant différents postes (sur les 14 postes répertoriés – professeure ou professeur, ingénieure ou ingénieur en logiciels de recherche, personne étudiante à la maîtrise ou au doctorat, etc.); le nombre maximum de postes dans un groupe était de 10 sur 14 (c.-à-d. de grandes équipes diversifiées) et le nombre minimum, de 1



sur 14 (un seul poste effectuant tout le développement), 40 % des groupes déclarant avoir 1 ou 2 postes de développeuses ou développeurs dans leurs équipes (figure 11).

Le sondage a révélé que, dans tous les domaines de recherche, l'expérience en développement de logiciels est bimodale. Dans tous les domaines, plus de 50 % des répondantes et répondants ont plus de 6 ans d'expérience en développement de logiciels, tandis que l'autre moitié a moins de 5 ans d'expérience. Cela dit, certaines différences importantes entre les domaines sont évidentes, en particulier chez les personnes moins expérimentées. En SSH et en RS, 33 % et 24 % des répondantes et répondants ont déclaré avoir moins d'un an d'expérience dans le développement de logiciels de recherche, contre seulement 9 % en SG [D4]. Ce sondage ne permet pas de déterminer les raisons précises de cet écart, mais il pourrait être attribuable aux différences dans les cours et les projets de premier cycle d'une discipline à l'autre, les étudiantes et étudiants de premier cycle en SG étant plus susceptibles de recevoir une formation ou des cours sur le développement de logiciels que celles et ceux en RS et en SSH.

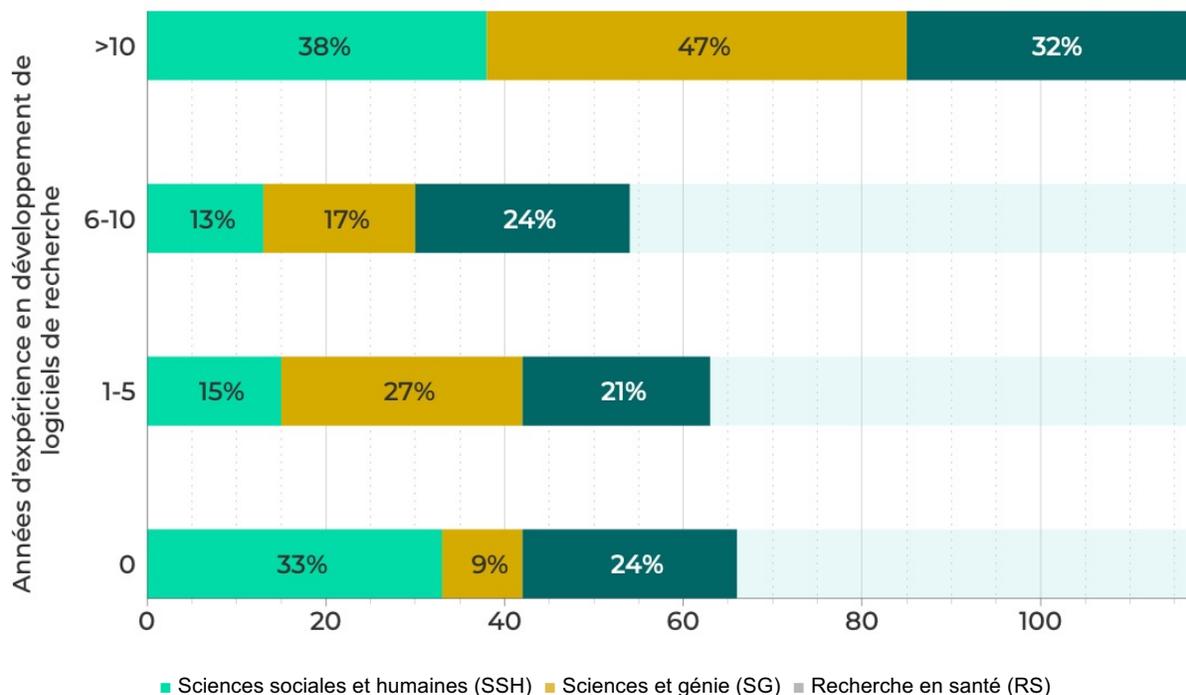


Figure 12. Question [D4] : Combien d'années d'expérience avez-vous en développement de logiciels de recherche? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 39; RS : n = 38; SG : n = 93; répondantes et répondants uniques : 170.

Les années d'expérience en développement de logiciels étaient en quelque sorte liées à la compétence. Lorsqu'on a demandé aux répondantes et répondants quel était leur niveau personnel de compétences en programmation ou en codage [D26], la moitié en SSH a déclaré avoir une expérience limitée ou des connaissances de base, contre 11 % et 27 % respectivement en SG et en RS. En contrepartie, 30 % des répondantes et répondants en SSH se considèrent comme étant à un niveau avancé ou expert, contre 48 % en SG et 45 % en RS (figure 13).

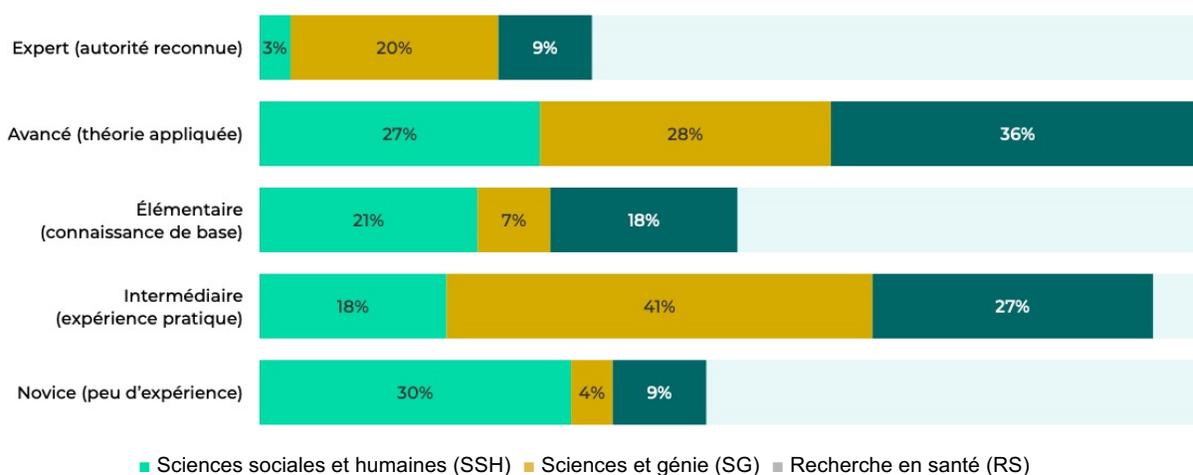


Figure 13. Question [D26] : Comment décririez-vous votre niveau personnel de compétences en programmation ou en codage? Réponses mesurées en pourcentage dans les domaines des trois organismes. Réponses mesurées en pourcentage dans les domaines des trois conseils. SSH : n = 33; RS : n = 33; SG : n = 90; répondantes et répondants uniques : 156.

Le temps que les répondantes et répondants consacrent au développement de logiciels est assez semblable d'un domaine de recherche à l'autre, même si, comme pour les questions précédentes, les chercheuses et chercheurs en SG ont tendance à y consacrer un peu plus de temps. Par exemple, 68 % des répondantes et répondants en SSH consacrent peu ou pas de temps au développement de logiciels (c.-à-d. entre 0 % et 25 % de leur temps), tandis que 58 % en RS et 40 % en SG entrent dans cette catégorie [D27]. Il est courant d'avoir des membres du personnel qui se consacre au développement de logiciels (c.-à-d. qui y passe de 75 à 100 % de leur temps) dans tous les domaines, mais moins en RS (figure 14). Ces résultats indiquent que les répondantes et répondants en SSH ont moins de formation initiale en développement de logiciels [D4] et passent moins de temps à développer des logiciels en tant que chercheuses et chercheurs [D27] que celles et ceux en SG et en RS.

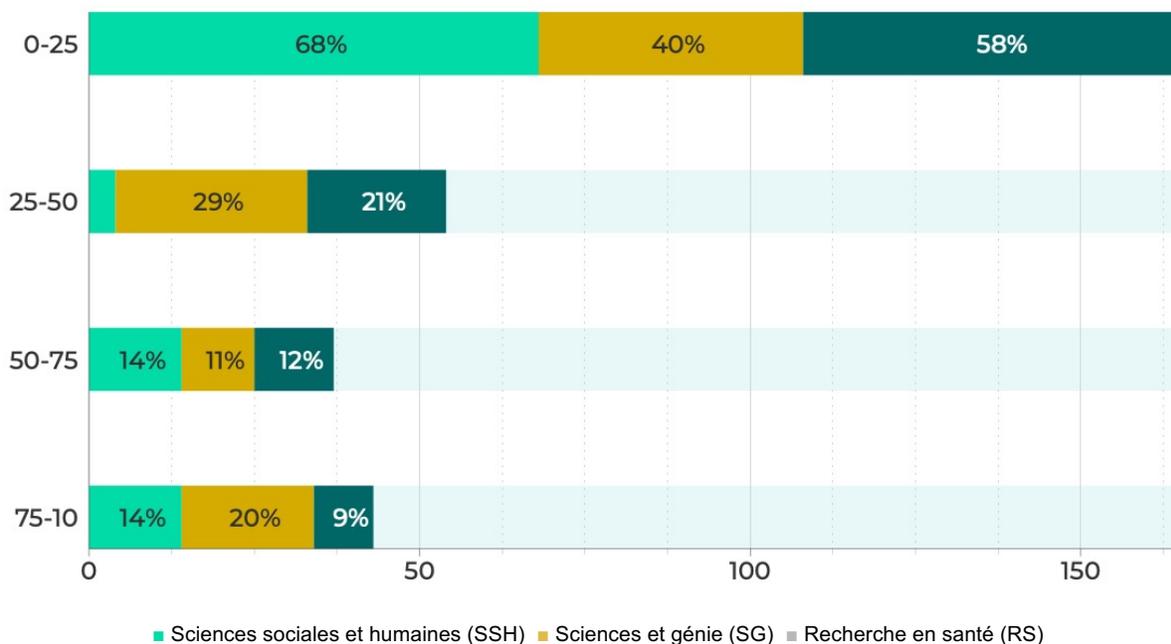


Figure 14. Question [D27] : Combien de temps consacrez-vous en moyenne à la programmation ou au codage par projet (pourcentage [%] du temps total du projet)? Réponses mesurées en pourcentage dans les domaines des trois conseils. Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 28; RS : n = 33; SG : n = 80; répondantes et répondants uniques : 141.

Interrogés sur le temps consacré par les différents groupes de recherche au développement de logiciels, 30 % des répondantes et répondants ont déclaré que leur équipe déployait un effort important (plus de 2 développeuses ou développeurs à temps plein), tandis que 33 % ont indiqué un effort minimal (< ½ ETP) [D7]. Parmi les répondantes et répondants, 37 % ont décrit leur équipe comme consacrant un effort modeste au développement de logiciels (entre ½ ETP et 2 ETP). Les répondantes et répondants en SG semblent consacrer un peu plus de temps à la recherche et au développement de logiciels que celles et ceux d'autres domaines (SSH – 24 % ; SG – 34 %; RS – 27 %; figure 15).

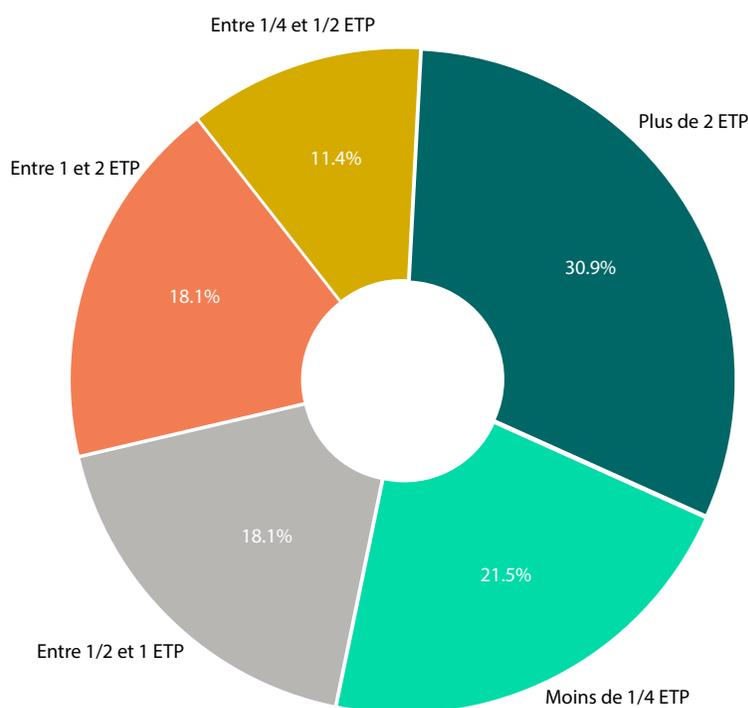


Figure 15. Temps consacré au développement de logiciels de recherche par les membres de l'équipe.

Types de logiciels de recherche

Les types de logiciels développés varient considérablement, les 27 catégories de logiciels du sondage ayant été sélectionnées par au moins 5 répondantes et répondants. Toutefois, Python était le langage de programmation le plus utilisé dans les trois domaines (SSH – 43 %; SG – 76 %; RS – 67 %) [D9]. Les résultats montrent également la diversité des langages de programmation dans les logiciels de recherche, avec 21 langages utilisés par les répondantes et répondants. Ces personnes développent principalement des logiciels libres, des codes sources, des outils gérés par des chercheuses et chercheurs et des plateformes utilisées par des utilisatrices et utilisateurs externes [D10, D14]. À la question demandant de fournir une liste de logiciels développés et activement maintenus par leurs groupes de recherche, plus de 100 répondantes et répondants ont fourni une liste de progiciels activement maintenus [D11]. Un tiers des répondantes et répondants ont indiqué qu'ils disposaient de progiciels qu'ils avaient développés, mais qu'ils ne maintenaient plus [D12]. De plus, 54 % ont déclaré que leur équipe participait au codéveloppement ou au développement collaboratif d'une plateforme de logiciels de recherche [D15].

Les répondantes et répondants ont indiqué une large adoption du contrôle des versions (81 %) pour le développement de logiciels [D23]. Cela dit, les normes ne soient exploitées/utilisées que par certains développeuses et développeurs de logiciels de recherche (38 %), et une proportion



importante d'entre elles et eux ne les utilise pas du tout (62 %) [D24]. De plus, 54 % des répondantes et répondants ont dit utiliser les pratiques exemplaires en matière de sécurité dans le développement de leurs logiciels [D25].

Moins de la moitié des répondantes et répondants documentent systématiquement leurs logiciels (documentent toujours : SSH – 31 %; SG – 38 %; RS – 44 %) [D13]. Comme la bonne documentation est l'un des facteurs clés qui déterminent l'utilisation d'un progiciel par une chercheuse ou un chercheur (SSH – 60 %; SG – 57 %; RS – 55 %) [E14], il s'agit d'une lacune importante qui doit être comblée.

Les pratiques de gestion et de partage des logiciels de recherche sont principalement influencées par les attentes de la communauté de recherche et des collaboratrices et collaborateurs, avec une grande influence observée en SSH (57 %), en SG (61 %) et en RS (49 %) [D16]. Les politiques des bailleurs de fonds et des établissements les influencent également (politiques des bailleurs de fonds : SSH – 30 %; SG – 29 %; RS – 33 %), mais les politiques des éditeurs, elles, ont une moindre influence (SSH – 9 %; SG – 12 %; RS – 19 %).

Pour le partage des logiciels de recherche, les répondantes et répondants s'appuient généralement sur des dépôts de logiciels ouverts tels que GitHub, sur l'accès direct au logiciel et sur des sites Web personnels ou de groupes de recherche [D17]. Ces plateformes constituent des canaux essentiels pour la diffusion des logiciels de recherche dans la communauté de recherche universitaire.

Un nombre considérable de répondantes et répondants (54 %) ont déclaré qu'une publication de recherche décrivait leur logiciel de recherche et 49 % ont inclus une référence à une publication qui indique la reconnaissance scientifique et l'intégration de leur logiciel de recherche dans leurs travaux [D18]. Parmi les répondantes et répondants, 58 % ont dit ne pas connaître les principes FAIR4RS ni les principes connexes pour le développement de logiciels de recherche [D19], et 63 % ont déclaré n'avoir jamais publié de version de leur logiciel de recherche [D20]. Cependant, 52 % des répondantes et répondants ont déclaré pouvoir mesurer l'influence de leur logiciel à l'aide d'autres mesures (à l'exclusion des citations d'articles ou de publications de logiciels de recherche décrites aux points D18 et D20), les principales mesures étant le nombre d'utilisatrices et utilisateurs et le nombre de téléchargements [D21].

Participation au développement de logiciels de recherche

La majorité des répondantes et répondants (75 %) ont reçu une formation sur l'utilisation ou le développement de logiciels de recherche [F1], mais la plupart sont autodidactes (SSH – 84 %; SG – 85 %; RS – 88 %) et apprennent grâce aux ressources en ligne telles que StackOverflow et Software Carpentry (SSH – 61 %; SG – 72 %; RS – 62 %) [F2]. De plus, certaines répondantes et certains répondants ont mentionné la formation par les pairs ou les cours parascolaires comme d'autres moyens d'acquérir des connaissances et des compétences sur les logiciels.



La plupart des répondantes et répondants apprécient la formation sur le développement de logiciels de recherche (72 %) [F4], mais sont moins intéressés (45 %) [G13] par la formation sur la pérennité des logiciels. Cela indique une forte motivation à améliorer les compétences liées aux logiciels et à rester à jour dans le domaine, mais une motivation moindre à développer des compétences pour garantir la maintenance et l'utilisation de leurs logiciels au fil du temps. La quasi-totalité des répondantes et répondants (97 %) indique ne pas faire partie d'une association d'ingénierie des logiciels de recherche [F5], mais une proportion notable (37 %), notamment en SG (41 %) et en RS (38 %), a exprimé son intérêt à rejoindre de telles communautés ou de tels programmes. En revanche, seulement 24 % des répondantes et répondants en SSH ont exprimé leur intérêt [F6]. Les personnes intéressées cherchent principalement à apprendre des pratiques exemplaires en développement logiciel, à recevoir de la formation ainsi qu'à participer à des activités de collaboration et de réseautage dans la communauté d'ingénierie des logiciels de recherche [F7].

Soutien au développement de logiciels de recherche

En ce qui concerne le soutien au développement de logiciels, 76 % des répondantes et répondants ont souligné l'importance (très important = 49 %; plutôt important = 27 %) d'avoir accès à un tel service pour leur recherche [C6]. Néanmoins, 57 % des répondants et répondantes ont déclaré ne pas avoir accès à un tel soutien (42 %) ou être dans l'incertitude (15 %) quant à sa disponibilité [C4].

Pour celles et ceux qui y ont accès (43 %), c'est l'établissement local qui en est le principal pourvoyeur. Plus de 90 % des répondantes et répondants des SSG, 53 % des SG et 92 % de la RS ont indiqué que leur établissement était leur principale source d'assistance (figure 16) [C5]. Ce constat montre le rôle crucial que jouent les établissements pour faciliter le développement de logiciels de recherche en offrant aux chercheuses et chercheurs les ressources et l'expertise nécessaires. L'Alliance (SSH – 26 %; SG – 28 %; RS – 24 %) et les communautés disciplinaires (SSH – 17 %; SG – 38 %; RS – 20 %) ont également été nommées comme fournisseurs de soutien au développement de logiciels.

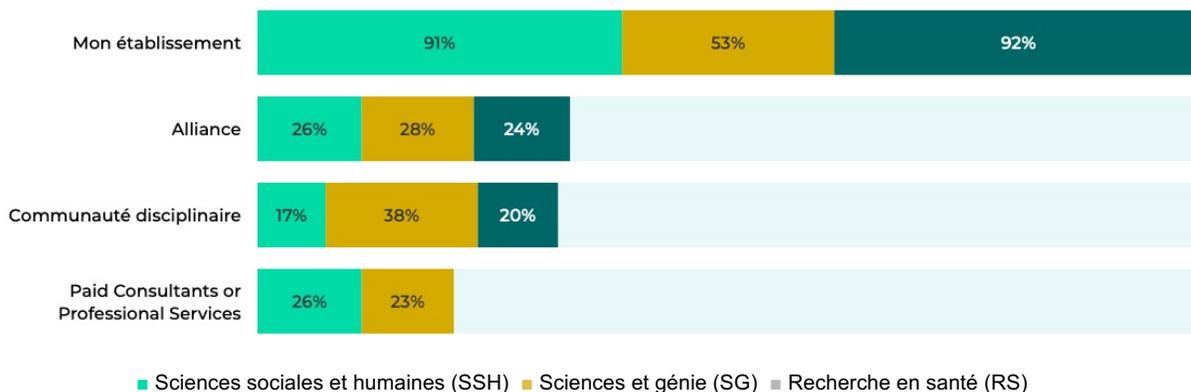


Figure 16. Question [C5] : Qui fournit ce soutien? Réponses mesurées en pourcentage dans les domaines des trois conseils. Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 23; RS : n = 25; SG : n = 53; répondantes et répondants uniques : 101.

Parmi les répondantes et répondants, 48 % ont déclaré qu'il existait des plateformes ou des services logiciels qui seraient utiles s'ils étaient soutenus en tant que service national [C11] et 31 % ont indiqué que les plateformes ou les services logiciels qu'elles et ils sont en train de développer ou de codévelopper seraient, à leur avis, utiles en tant que services nationaux [C13].

Financement du développement de logiciels de recherche

Seulement 26 % des répondantes et répondants ont déclaré avoir reçu du financement dans le cadre d'appels de financement spécifiquement pour le développement de logiciels de recherche, la majorité des répondantes et répondants (63 %) utilisant un financement non spécifique pour leurs activités de développement [D2]. Parmi celles et ceux ayant reçu un financement spécialement pour les logiciels de recherche et en ayant précisé la source, les organismes de financement étaient en majorité CANARIE et la Fondation canadienne pour l'innovation (FCI) [D3]. Le financement non spécifique des logiciels de recherche (couvrant à la fois le développement et l'utilisation) provient d'une multitude de sources [B8; figure 17], notamment des établissements. Une proportion importante de répondantes et répondants a indiqué ne pas recevoir de financement pour le développement de leurs logiciels de recherche, encore plus en SSH (36 %). Les trois conseils étaient la principale source de financement des répondantes et répondants au sein de chaque conseil (SSH – CRSH; SG – CRSNG; RS – IRSC), bien que certains financements croisés existent également. La FCI et les gouvernements provinciaux figurent également parmi les principaux bailleurs de fonds des logiciels de recherche (figure 17).

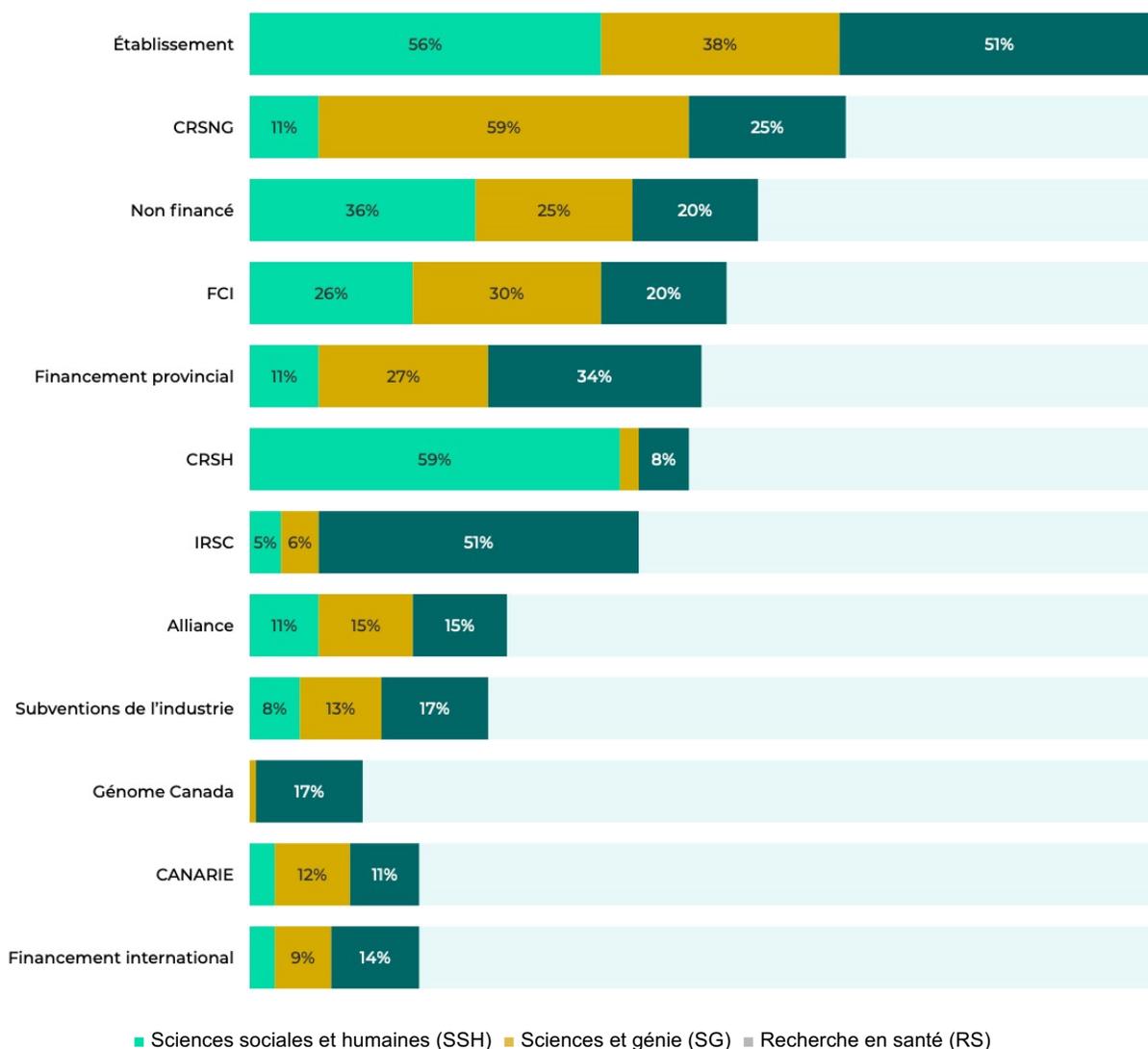


Figure 17. Question [B8] : D'où proviennent les fonds du budget consacré aux logiciels de recherche de votre groupe? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 73; RS : n = 65; SG : n = 136; répondantes et répondants uniques : 274.

Les répondantes et répondants déclarent consacrer chaque année une partie importante de leur fonds de recherche au développement de logiciels de recherche [D8] : 33 % déclarent dépenser plus de 50 000 \$ par an et 15 % plus de 250 000 \$ par an. À l'inverse, 20 % des répondantes et répondants ont déclaré n'avoir rien dépensé pour le développement de logiciels, tandis que 24 % ont dit ne pas connaître le total annuel de leurs dépenses pour ces activités (figure 18).

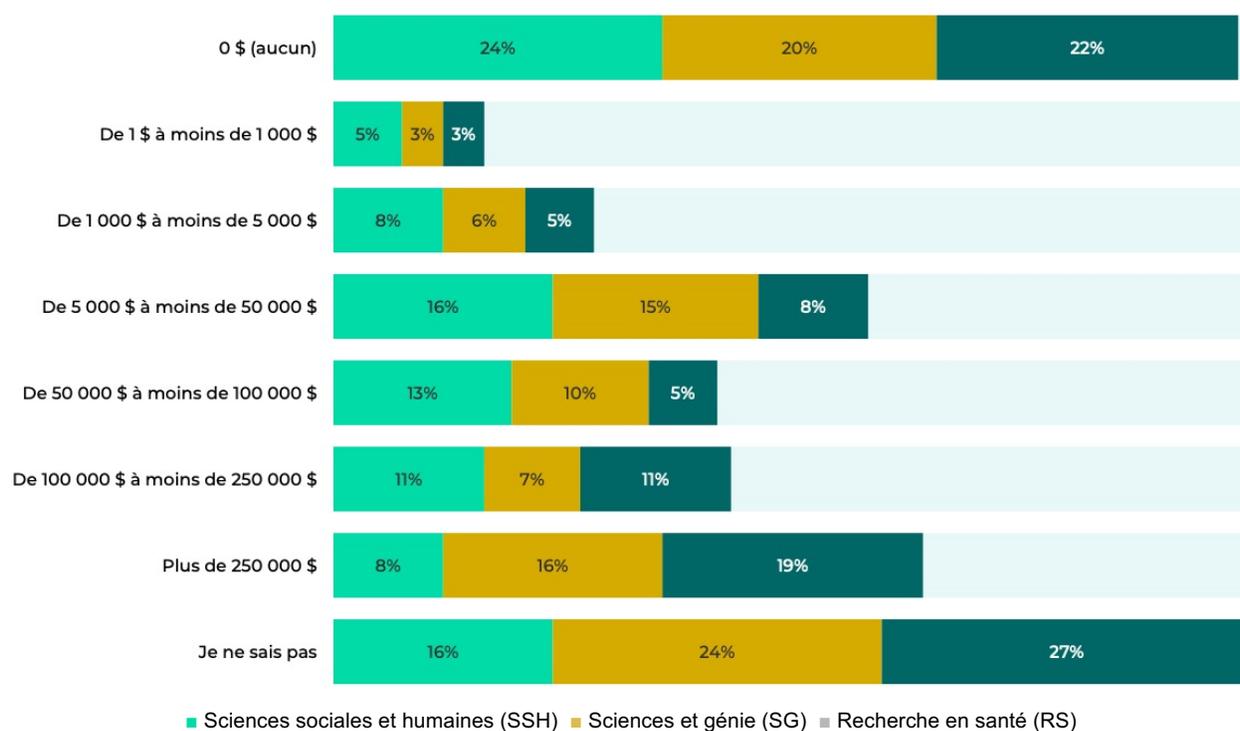


Figure 18. Question [D8] : Environ quel montant (en dollars canadiens) de vos fonds de recherche votre groupe de recherche a-t-il consacré au développement de ressources logicielles de recherche au cours de la dernière année civile? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 38; RS : n = 37; SG : n = 89; répondantes et répondants uniques : 164.

Le financement est considéré comme le principal obstacle à la pérennité des logiciels de recherche dans tous les domaines [D30] : 78 % des répondantes et répondants en SG, 83 % en RS et 67 % en SSH indiquent que le financement est un problème majeur. Le recrutement et la rétention de personnel constituent le principal obstacle à la pérennité des logiciels [D31] (figure 19).

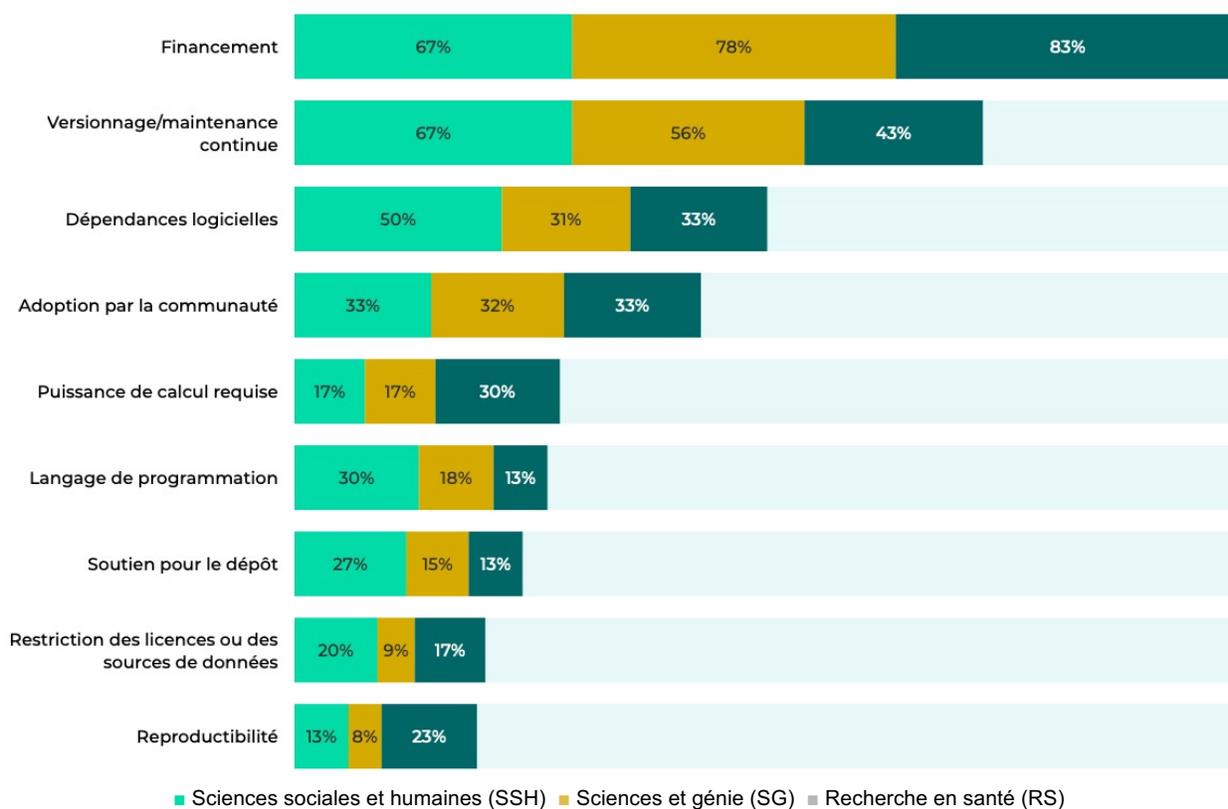


Figure 19. Question [D30] : En ce qui concerne votre propre logiciel, quels sont les principaux obstacles à la pérennité? Réponses mesurées en pourcentage dans les domaines des trois organismes. SSH : n = 30; RS : n = 30; SG : n = 78; répondantes et répondants uniques : 138.

Ces résultats soulignent la disponibilité limitée des possibilités de financement pour le développement de logiciels de recherche et les difficultés rencontrées par beaucoup de chercheuses et chercheurs dans l'obtention d'un soutien financier pour leurs projets logiciels.



La pérennité des logiciels de recherche

La majorité des répondantes et répondants accordent une grande importance à la préservation des logiciels de recherche (94 %) [G9], mais ils ont des opinions divergentes sur la responsabilité de la pérennité des logiciels de recherche. Certaines et certains suggèrent qu'elle revient à la communauté des utilisatrices et utilisateurs, tandis que d'autres estiment qu'elle incombe à la chercheuse principale ou au chercheur principal, à la développeuse principale ou au développeur principal ou à l'établissement [G4]. Les répondantes et répondants s'entendent que la pérennité des logiciels de recherche est importante pour faciliter la réutilisation (SSH – 80 %; SG – 79 %; RS – 83 %), faciliter la reproductibilité (SSH – 54 %; SG – 74 %; RS – 87 %), soutenir les principes de la science ouverte (SSH – 74 %; SG – 63 %; RS – 63 %) et créer une communauté (SSH – 63 %; SG – 47 %; RS – 57 %) [G5]. Plus de 30 % des répondantes et répondants ont bénéficié d'un accès à des logiciels préservés [G6], mais 25 % d'entre eux ont également déclaré ne pas savoir, ce qui suggère un manque de connaissances quant à la préservation ou non des logiciels. De même, un tiers des répondantes et répondants ont développé des logiciels qui ne sont plus maintenus, l'absence de maintenance étant principalement due à des problèmes de financement, de version ou de maintenance même [D12].

Les répondantes et répondants se heurtent en outre à de nombreuses difficultés quand vient le temps d'utiliser les logiciels d'autres utilisatrices et utilisateurs. Les principaux obstacles : le manque de ressources financières (financement), d'expertise et de temps. De plus, les répondantes et répondants ont souligné l'importance de disposer de fonds adaptés à leurs besoins et de recevoir un soutien dans des domaines tels que la formation et la maintenance.

La plupart de répondantes et répondants connaissent les dépôts de logiciels courants pour la préservation des logiciels comme GitHub ou BitBucket (SSH – 96 %; SHG – 91 %; RS – 87 %), mais sont moins familiers avec les dépôts généraux comme Zenodo (SSH – 26 %; SG – 19 %; RS – 35 %) ou les dépôts de conteneurs logiciels comme DockerHub (SSH – 26 %; SG – 16 %; RS – 26 %) [G10]. Les dépôts institutionnels de logiciels et les archives de Software Heritage sont peu utilisés.

En ce qui concerne la publication et la préservation des logiciels, la majorité des répondantes et répondants, tous domaines confondus, considèrent comme utiles l'examen du code par les pairs, l'aide à la rédaction de la documentation et les consultations sur la pérennité des logiciels [G12]. Toutefois, l'importance de ces services varie d'un domaine à l'autre. Par exemple, la consultation en personne est très importante en SSH (70 %), mais l'est moins pour en SG (38 %) et en RS (22 %). Inversement, l'exécution de tests de compatibilité est plus importante en RS (61 %) et en SG (52 %) qu'en SSH (33 %).



Conclusion

Les logiciels de recherche sont devenus un outil essentiel de la recherche moderne dans tous les domaines. Néanmoins, de nombreux facteurs culturels et techniques ont limité l'adoption élargie des pratiques exemplaires en génie logiciel (p. ex., développement et documentation des logiciels) et le partage des logiciels (p. ex., FAIR4RS). Les résultats présentés dans ce sondage sont conformes à ceux observés à l'international, où 95 % des répondantes et répondants utilisent des logiciels de recherche et 50 % estiment qu'il s'agit d'un produit principal de leur recherche.

Le développement de logiciels de recherche dans différents groupes semble être partagé par différents rôles ou types de répondantes et répondants. Néanmoins, les développeuses et développeurs de logiciels de recherche les plus courants sont des étudiantes et étudiants de cycles supérieurs, plutôt que des développeuses et développeurs de logiciels de formation ou des ingénieures et ingénieurs en logiciels de recherche. Pourtant, ce sondage n'était pas nécessairement destiné à ce groupe. En effet, les étudiantes et étudiants (~7 %), les chercheuses et chercheurs au postdoctorat et les assistantes ou assistants de recherche (~7 %) sont très peu représentés. À l'inverse, une grande partie des répondantes et répondants étaient des professeures et professeurs, ce qui permet de bien comprendre comment ce groupe élabore, utilise et finance ses activités consacrées aux logiciels de recherche. Le corps professoral en SG consacre systématiquement plus de temps au développement de logiciels de recherche et y joue plus de rôles que celui en SSH et en RS. Si cela n'est pas surprenant, étant donné que la recherche dans ces domaines nécessite souvent des calculs et des codes lourds, cela montre que les professeures et professeurs en SG consacrent probablement plus de temps au soutien des logiciels de recherche qu'elles et ils développent ou utilisent.

Les établissements jouent un rôle de premier plan dans le soutien et le financement des logiciels de recherche. En effet, la plupart des répondantes et répondants s'appuient sur le soutien de leur établissement pour les logiciels, et celles et ceux qui n'ont pas accès à un tel service ont souligné à quel point il leur serait utile. De même, les établissements sont une source de financement des logiciels de recherche tout aussi importante pour les répondantes et répondants que leur organisme de financement respectif des trois conseils. À mesure que les universités élargissent leurs catalogues de logiciels vérifiés et sécurisés pour la recherche, il est probable qu'elles continueront à jouer un rôle de plus en plus important dans le soutien de leur corps professoral et de leurs communautés de recherche.

Si la majorité des répondantes et répondants reconnaissent la grande importance de la conservation et du partage des logiciels de recherche, beaucoup peinent à réutiliser les logiciels d'autres personnes. Parmi les facteurs les plus importants qui empêchent la réutilisation, il y a le manque de documentation et de soutien appropriés. Pourtant, ce sont les mêmes aspects qui posent problème aux répondantes et répondants qui ont mis au point leur propre logiciel, et bon nombre ne connaissent pas les pratiques exemplaires de développement de logiciels, comme les principes FAIR4RS. La formation et la sensibilisation semblent être les éléments les plus importants pour opérer un changement culturel et une adoption à grande échelle des pratiques exemplaires en développement de logiciels.