

Réponse de Calcul Canada et de CANARIE
au document de consultation TC3+ :
« Tirer profit des données massives : Vers un
nouveau cadre stratégique pour l'évolution de la
recherche numérique au Canada »

Décembre 2013

Introduction

CANARIE et Calcul Canada sont heureux de pouvoir soumettre leurs impressions sur le document de travail intitulé « Vers un nouveau cadre stratégique pour l'évolution de la recherche numérique au Canada » et remercie les conseils subventionnaires ainsi que la Fondation canadienne pour l'innovation d'alimenter le dialogue en la matière.

En tant que piliers de l'infrastructure numérique nationale qui sous-tend la recherche, l'éducation et l'innovation au Canada, Calcul Canada et CANARIE partagent une perspective unique sur l'état actuel de cette infrastructure au pays et sur la manière dont on pourrait la renforcer afin de faire fructifier les deniers publics davantage et plus durablement.

CANARIE et les réseaux provinciaux et territoriaux qui sont ses partenaires exploitent collectivement le réseau canadien de la recherche et de l'éducation tout pilotant son évolution. En plus d'assurer une connectivité ultrarapide à son réseau, CANARIE dispense des services d'accès et de gestion des identités par le biais de la Fédération canadienne d'accès, autorisant ainsi un accès à distance aux données de recherche et à divers services essentiels. Parallèlement, CANARIE orchestre le développement d'une gamme de plateformes logicielles de recherche, notamment une série d'interfaces réutilisables, dans le but de combler le vide qui existe présentement au niveau des logiciels scientifiques et d'aider les utilisateurs à accéder plus facilement aux données nationales et internationales en vue de les analyser.

De son côté, Calcul Canada est l'organisation nationale chargée de dispenser les services de calcul de pointe aux chercheurs et aux innovateurs du pays. Parmi ces services figurent les plus importantes installations de stockage de données destinées aux chercheurs canadiens et des installations de calcul qui vont de grappes d'ordinateurs à haute performance à des services d'infonuagique. Plus important encore, Calcul Canada procure le personnel et le savoir-faire sans lesquels les chercheurs ne pourraient exploiter efficacement une telle infrastructure, en prodiguant des conseils à ces derniers, en créant des logiciels, en procédant à des analyses et en proposant d'autres services.

Calcul Canada et CANARIE font partie du Conseil du leadership sur l'infrastructure numérique qui prépare lui aussi un document dans le contexte de ces consultations. Si la réponse du Conseil du leadership reflètera le point de vue d'une multitude d'intervenants associés à l'infrastructure numérique, le document que voici expose celui de Calcul Canada et de CANARIE, qui sont des chefs de file dans l'infrastructure numérique et la prestation d'une telle infrastructure de longue date.

Réponse aux recommandations

1. Établir une culture de saine gestion

Bien qu'il existe d'excellents modèles de gestion des données dans certains secteurs de recherche, la rapidité avec laquelle les données voient le jour dans maintes disciplines signifie qu'on produit et stocke actuellement une masse colossale d'informations sans véritables directives ni stratégies pour leur préservation. Par conséquent, nous convenons que les bailleurs de fonds comme les fournisseurs de services doivent impérativement instaurer une culture de saine gestion des données dans le milieu de la recherche.

CANARIE et Calcul Canada appuient les recommandations du document de consultation et en formulent quelques-unes à ce sujet :

- a) Développer des mécanismes qui garantiront le respect des normes reconnues en gestion et en préservation des données. On pourrait notamment envisager un « certificat de conformité » attestant que les données ont été correctement marquées et stockées pour que d'autres puissent les consulter et les réutiliser. Les exigences des bailleurs de fonds en matière de gestion des données seraient donc respectées avant le stockage. Pareille approche autoriserait l'implantation de mécanismes ascendant et descendant qui garantiraient la conformité et pourrait faciliter le passage du « protectionnisme » classique des données à un accès libre et plus facile à l'information.
- b) Procurer des logiciels de recherche courants aux chercheurs et aux utilisateurs, et aider ceux-ci à s'en servir afin uniformiser l'archivage et la récupération des données en fonction de normes communément acceptées par la collectivité.
- c) Bien qu'ils appuient l'élaboration de politiques et de normes régissant les données qui en garantiront une gestion plus efficaces, Calcul Canada et CANARIE réitèrent qu'en l'absence d'un financement soutenu de l'infrastructure matérielle qui abrite ces données et en permet la consultation, ainsi que du personnel, des outils et des services sans lesquels les chercheurs ne pourraient exploiter ces données, un tel effort serait voué à l'échec. Il est difficile de voir comment l'horizon de financement quinquennal (au maximum) actuel peut répondre à des besoins d'archivage qui s'étendent sur des décennies. Une solution éventuelle serait que le gouvernement s'engage à financer le renouvellement de l'infrastructure avant l'expiration du mandat pour bien signaler l'intérêt qu'il porte à un archivage à long terme, pour atténuer le climat d'incertitude qui règne dans la collectivité et pour inciter les créateurs de données à recourir aux installations nationales de stockage.
- d) Plus avant, il est capital qu'on exploite au mieux l'expertise existante. Bien que Calcul Canada ait les moyens et le savoir-faire voulus pour stocker les données, l'expertise en matière de gestion des données réside ailleurs (notamment à l'Association des bibliothèques de recherche du Canada, ou ABRC, qui chapeaute le développement d'un réseau canadien de gestion des données de recherche avec le

concours du Réseau canadien de documentation pour la recherche, des bibliothèques universitaires et de Données de recherche Canada).

2. Coordonner la participation des parties prenantes

Bien que les « actions émergent de la base » aient servi la collectivité, nous sommes d'accord qu'il est temps que les intervenants coordonnent mieux leurs efforts. Calcul Canada et CANARIE sont prêts à offrir leur concours. Aucune organisation au Canada ne réunit à elle seule l'expertise requise pour couvrir l'écosystème numérique dans son entièreté. Ce n'est qu'en mettant tous l'épaulé à la roue que nous parviendrons à optimiser le rendement des sommes investies, pour le bénéfice des Canadiens.

En plus d'appuyer les recommandations contenues dans le document de consultation, nous tenons à souligner qu'on pourrait renforcer les mécanismes existants pour harmoniser davantage les activités et les besoins des bailleurs de fonds, des utilisateurs et des fournisseurs d'infrastructure numérique.

- a) On pourrait convier les fournisseurs d'infrastructure numériques aux réunions durant lesquelles les représentants des organismes subventionnaires examinent les grandes initiatives scientifiques ou aux réunions ordinaires où sont échafaudés les plans des projets de longue haleine. Cette simple mesure concourrait considérablement à aligner l'infrastructure avec les besoins futurs des chercheurs et des innovateurs sans que la charge administrative s'en trouve alourdie pour autant.
- b) Face au modèle décentralisé qui la caractérise présentement, on pourrait faciliter une évolution ordonnée de l'infrastructure numérique, soutenir les initiatives propices à l'ouverture des données et voir à ce que les composants deviennent interopérables. Pour y parvenir, il suffirait de développer un jeu de normes auquel chaque composant devrait se conformer, condition préalable pour obtenir des fonds publics. Les dépôts de données, par exemple, pourraient être accessibles à un plus grand nombre d'utilisateurs une fois qu'ils auront accepté d'adhérer à certaines normes de gestion et de préservation des données (que pourraient élaborer l'ABRC et ses partenaires). Ces dépôts seraient hébergés dans une installation nationale de données de recherche (se conformant aux normes de stockage). On assisterait alors à la convergence vers une architecture qui satisferait aux besoins des secteurs public et privé tout en laissant aux entreprises privées la marge de manœuvre nécessaire pour adapter leurs produits et services en vue de proposer une aide en nature dans le sens des initiatives publiques.

3. Accroître les ressources et définir les paramètres de financement futurs

CANARIE et Calcul Canada abondent dans le sens des recommandations et souhaitent répondre à la demande concernant l'identification des lacunes en formulant leurs propres recommandations pour remédier à celles-ci.

- a) Le Canada n'a pas de stratégie cohérente pour les données de recherche. Bien que Calcul Canada compte de nombreux sites qui pourraient collectivement constituer une installation nationale de stockage des données, des problèmes de financement interdisent à l'organisme de mettre sur pied un centre national de données de recherche robuste et durable. La structure de financement actuelle complique la planification stratégique à long terme d'une telle installation.
- b) Le Canada possède une stratégie sur mesure pour les logiciels de recherche. Une stratégie plus cohérente permettrait d'identifier les secteurs pour lesquels il faudrait élaborer des logiciels en priorité, de créer une banque d'interfaces réutilisables et de mieux coordonner les efforts des actuels fournisseurs de logiciels (y compris CANARIE, Calcul Canada, Génome Canada et d'autres).
- c) CANARIE et Calcul Canada sont d'accord avec la recommandation que Données de recherche Canada a formulée dans sa réponse au document de consultation eu égard aux efforts déployés avec la collectivité internationale, et plus particulièrement la Research Data Alliance, en vue d'assurer l'interopérabilité des données de recherche.
- d) Une tactique relativement simple pour mieux faire connaître les infrastructures numériques disponibles serait d'inclure aux demandes de subvention une section qui présenterait aux chercheurs la liste des outils de cette infrastructure que l'on met à leur disposition (par ex., réseau, centres de calcul, logiciels, installations de stockage, etc.). Le chercheur serait tenu d'indiquer les outils dont il aurait besoin en plus de l'infrastructure existante.

Si on leur soulignait l'infrastructure disponible quand ils remplissent leur demande de subvention, les chercheurs comprendraient mieux l'éventail actuel d'infrastructures et d'outils auxquels ils pourraient recourir. Un tel processus ferait en sorte que l'argent destiné à la recherche serve précisément à cette fin plutôt qu'à bâtir l'infrastructure qui la sous-tend. Parallèlement, on identifierait les lacunes au niveau de l'infrastructure numérique, car le processus d'examen des demandes de subvention ferait ressortir les tendances relatives aux infrastructures supplémentaires que réclament les utilisateurs

L'enjeu du financement

Quoique que Calcul Canada et CANARIE aient identifié diverses tactiques susceptibles de rehausser l'efficacité de l'infrastructure numérique canadienne, on pourrait aligner les méthodes de financement actuelles et les comportements qu'elles engendrent pour modifier la stratégie actuelle et en retirer les avantages qu'on imagine.

Les structures de financement existantes nuisent à l'élaboration d'une stratégie de longue haleine et à une planification à long terme des principaux aspects de l'infrastructure numérique. Calcul Canada et CANARIE reçoivent des fonds de services différents d'Industrie Canada, se voient attribuer des objectifs distincts et la durée de leur financement n'est pas la même. C'est pourquoi il est parfois si difficile d'amener ces infrastructures de base à investir en vue de réaliser le

maximum de bénéfiques à quelque moment que ce soit. On pourrait harmoniser davantage les mécanismes de financement et les mandats pour que ces organisations planifient sur un horizon plus lointain, équilibrent les sommes injectées dans le réseau, les centres de calcul, les logiciels et les installations de stockage en fonction des besoins de la collectivité évoluent et proposent une interface plus simple aux utilisateurs desdites infrastructures. Parmi les possibilités de financement qui amèneraient un meilleur alignement de ces infrastructures, il convient de mentionner les suivantes :

- harmoniser l'horizon de financement de CANARIE et de Calcul Canada, et demander à ces organisations de s'aligner dans leur planification stratégique et opérationnelle;
- revoir le modèle de financement des infrastructures nationales, notamment le personnel expert. Par exemple, on ne sait pas si le fait d'exiger un co-financement de 60 % comme on le fait actuellement est le modèle idéal pour garantir la stabilité de fonctionnement à long terme de Calcul Canada;
- une combinaison des solutions qui précèdent.

CANARIE et Calcul Canada conviennent qu'il est inutile de créer un organisme-cadre pour superviser leurs activités. Les deux organisations nourrissent déjà un dialogue constant sur le plan stratégique, dialogue articulé sur la collaboration et cette coopération devrait se resserrer avec le temps.

Une telle relation favorise l'implantation des principaux composants de l'infrastructure numérique requise : CANARIE procure le réseau évolué, les services réseau et les logiciels de recherche (bien que l'organisme ne soit pas le seul acteur dans le domaine des logiciels de recherche), alors que Calcul Canada offre des services évolués de calcul mais aussi de stockage, de consultation et d'analyse des données (tout en admettant qu'il n'existe présentement aucune stratégie pour le stockage des données). Les deux organismes appuient le travail de leurs partenaires dans l'élaboration d'une stratégie qui permettrait la préservation et la réutilisation des données de recherche. Ils s'efforcent de répondre aux besoins toujours changeants de la vaste communauté d'utilisateurs dans les universités, les collèges, les hôpitaux universitaires, les bibliothèques et les laboratoires de recherche du gouvernement.

Ultimes réflexions

Le Canada doit absolument développer la capacité de satisfaire aux besoins actuels et s'adapter à un environnement qui évolue rapidement pour mieux entreprendre des recherches numériques et exceller dans ce domaine. Acquérir une telle capacité est impossible sans le matériel et le savoir-faire technique d'une main-d'œuvre compétente. Calcul Canada et CANARIE réitèrent que, pour s'épanouir, il faut des partenaires stables et un financement prévisible.

Nous pensons aussi que, pour indispensables que soient les pratiques exemplaires et la meilleure technologie, il ne faut pas laisser les meilleures pratiques et technologies internationales limiter notre ambition. Pour que les chercheurs canadiens et leurs partenaires disposent des services, des ressources et des avantages auxquels ils ont droit, il importe d'avoir les compétences et les

ressources voulues pour repousser les frontières et progresser au-delà des normes internationales existantes, de façon à créer les outils et les technologies qui répondront le mieux aux priorités canadiennes.

L'évolution rapide de la technologie rendant vite désuètes divers éléments de l'infrastructure physique, le type d'infrastructure numérique le plus important et le plus durable est en fait constitué de ressources humaines, c'est-à-dire des chercheurs éclairés et déterminés, des experts chevronnés capables d'entamer de nouveaux projets d'importance nationale, et qui œuvrent en tandem pour dispenser une gamme complète de services numériques grâce à une infrastructure matérielle et une infrastructure de gestion des données.

CANARIE et Calcul Canada sont reconnaissants d'avoir pu donner leur point de vue sur les recommandations formulées dans le document de consultation et les deux organismes seraient heureux de donner plus de précisions sur les mesures supplémentaires qu'ils proposent.



Jim Ghadbane
Président et chef de direction
CANARIE



Jill Kowalchuk
Présidente
Calcul Canada