Recherche génomique dans le Nuage : Protection de la vie privée

Version 1.0, le 21 juillet 2015

Citation recommandée : Adrian Thorogood, Howard Simkevitz, Mark Phillips, Edward S Dove et Yann Joly « Recherche génomique dans le Nuage : Protection de la vie privée » (21 juillet 2015).

Auteurs: Adrian Thorogood[*], Howard Simkevitz[†], Mark Phillips[‡], Edward S. Dove[§] & Yann Joly[**]

Ce mémoire sur les orientations stratégiques recommande que des moyens de protection soient mis en place contre les risques liés à une atteinte à la vie privée concernant l'archivage et l'analyse de données génomiques et des autres données associées à la santé dans le nuage informatique (ci-après Nuage). Ce mémoire a été préparé dans le cadre du Programme des Contributions 2014-2015 du Commissariat à la protection de la vie privée du Canada. Les recommandations de ce mémoire s'appuient sur une analyse comparative de la législation canadienne, américaine et européenne sur la protection de la vie privée et des données personnelles, ainsi que sur une étude de cas concernant les Conditions d'Utilisation offertes par six fournisseurs de service dans le Nuage (ci-après Fournisseur). La partie I expose le caractère sensible des données génomiques et les défis pour la protection de la vie privée des participants engendrés par la recherche génomique dans le Nuage. La partie II analyse les obligations juridiques et éthiques des chercheurs utilisant le Nuage. La Partie III identifie les lacunes parmi les mesures de protection de la vie privée offertes par les Fournisseurs, dans leurs Conditions d'Utilisation et, propose des recommandations pour l'amélioration de celles-ci. Une liste de vérification à l'attention des chercheurs (voir *Cloud Contract Checklist* à l'annexe 1) fait état des principales clauses contractuelles que les chercheurs devraient considérer dans leurs contrats avec des Fournisseurs. La Partie IV identifie les barrières légales et règlementaires à l'adoption de services dans le Nuage ainsi que les lacunes en matière de protection de la vie privée.

[*] Attachés Affaires universitaires, Centre de génomique et politiques, Faculté de Médecine, Université McGill, Montréal, Canada.
[†] Conseiller Général et Commissaire à la Protection de la Vie Privée, Ontario Institute for Cancer Research, MaRS Centre Toronto, Canada.
[‡] Assistant de recherche, Centre de génomique et politiques, Faculté de Médecine, Université McGill, Montréal, Canada.
[§] Candidat au Doctorat, Faculté de droit, Université d'Édinbourg, Royaume-Uni.
[**] Professeur Agrégé, Centre de génomique et politiques, Faculté de Médecine, Université McGill, Montréal, Canada.

I. Contexte

Génomes et protection de la vie privée : La recherche génomique implique la comparaison des génomes et des données de santé associées d'individus à travers un grand groupe populationnel afin de déterminer comment les facteurs génétiques, médicaux et environnementaux interagissent pour causer la maladie. Le génome est l'entièreté de l'information génétique d'un individu, laquelle est en majeure partie héritée de la parenté biologique, particulièrement les parents d'un individu. Les données génomiques ont un caractère sensible parce qu'elles comprennent les renseignements sur la santé d'une personne et sur ses divers traits personnels. L'accès non autorisé aux données génomiques et aux autres données associées à la santé peut causer des préjudices individuels et sociétaux. Ces préjudices comprennent la discrimination par des tiers tels les employeurs et les assureurs, l'atteinte à la dignité, les dommages psychologiques, et la perte de confiance du public dans la recherche médicale et les soins de santé.

Le visage changeant de la recherche génomique : La mondialisation et une plus grande connectivité permettent aux chercheurs de collaborer et de partager des données de recherche à travers le monde. Les avancées dans la technologie de séquençage ont permis le décryptage du génome d'un individu de façon plus rapide et abordable. La croissance de la puissance informatique et les nouvelles techniques de bioinformatique permettent aux chercheurs d'assembler des génomes individuels et de comparer les données génomiques de plusieurs individus dans le but de découvrir des variations causant la maladie. Par conséquent, les chercheurs archivent une quantité grandissante de données génomiques et d'autres données associées à la santé sur les participants ou accèdent à ces données à partir de registres déjà existants. Ils stockent des données à long terme (sinon indéfiniment) et les exploitent afin de répondre à multiples questions de recherche d'une grande diversité. L'accélération des

découvertes ainsi que l'amélioration de la recherche et des soins de santé vont requérir une puissante infrastructure informatique capable d'emmagasiner et d'analyser de gigantesques ensembles de données et, de les partager à travers le monde.

Génomique dans le Nuage: Plusieurs gigaoctets de données génomiques peuvent être associés à un seul participant; les projets impliquant plusieurs dizaines de milliers de participants peuvent donc rapidement grossir ces quantités de données jusqu'à l'échelle des pétaoctets. Une quantité non négligeable de ressources et de réseaux informatiques peut être requise pour archiver, analyser et partager ces énormes ensembles de données. L'infonuagique présente un bon rapport coût/efficacité permettant au milieu de la recherche en santé publique de faire face à ces enjeux.

L'infonuagique est un assemblage complexe de plateformes informatiques, souvent intégrées avec des services d'affaires, lesquels permettent un accès échelonné, sur demande et à distance, à des ressources informatiques incorporées dans un réseau. Un Nuage peut être déployé par une organisation privée ou communautaire (comme une université) et peut être accessible à travers un réseau interne. Des entreprises commerciales peuvent offrir des services de Nuage au public, permettant aux clients d'emmagasiner et d'analyser les données comprises dans ce Nuage, à distance et à travers Internet. L'écosystème de l'infonuagique est varié et complexe. Des entreprises spécialisées en infonuagique, ou des sous-traitants de ces entreprises, peuvent, par exemple, fournir un accès à du matériel informatique, à une plateforme sur laquelle un chercheur peut développer des applications logicielles, ou à des applications logicielles intégrées que les chercheurs peuvent faire fonctionner à même le Nuage.

Les défis de la protection de la vie privée dans le Nuage: La recherche génomique dans le Nuage présente des risques à la protection de la vie privée. Le partage à grande échelle par des chercheurs à travers le Nuage augmente le risque d'accès non autorisé et d'utilisation de données

génomiques sensibles et des autres données associées à la santé des participants et de leur famille. Les données des participants transférées hors du Canada peuvent faire l'objet d'une divulgation obligatoire sous le régime de lois étrangères et le droit à la protection de la vie privée en vertu des lois canadiennes peut être difficile à faire respecter dans ce contexte. De plus, la complexité du stockage et de l'architecture de l'infonuagique crée de nouveaux risques de cybersécurité possiblement inconnus, particulièrement quand les utilisateurs du Nuage manquent d'expérience. Les chercheurs utilisant le Nuage pourraient devoir renoncer à une part du contrôle sur leurs données de recherche et aux moyens de les protéger. Cette situation soulève des préoccupations de responsabilité.

II. Le cadre éthique et juridique de la recherche génomique dans le Nuage

En général, les lois concernant la protection de la vie privée au Canada permettent aux chercheurs de stocker et d'analyser, dans le Nuage, des données génomiques et des autres données associées à la santé, sous réserve de quelques conditions et d'exceptions. Cependant, certaines approches dont l'objectif est d'accroitre la protection de la vie privée, comme les techniques de dé-identification, peuvent ne pas être suffisantes pour garantir l'anonymat des données génomiques. Les données génomiques sont par nature identifiables et doivent être traitées comme étant des « renseignements personnels sur la santé ». Ces renseignements personnels sur la santé sont protégés par les lois fédérales et provinciales sur la protection de la vie privée, et aussi par les exigences éthiques contenues dans l'*Énoncé de politique des trois Conseils*, lesquelles sont des conditions d'approbation requises pour la plupart des projets de recherche provenant de financement fédéral.

Les chercheurs demeurent responsables des informations transférées à des tiers: le principe juridique de la responsabilité requiert généralement que les chercheurs, en tant que détenteurs de

renseignements personnels sur la santé, respectent leurs obligations juridiques même après un transfert de données à un Fournisseur. Les chercheurs doivent:

1. Limiter les utilisations et divulgations de données à ceux étant autorisés par le participant (ou par un comité d'éthique de la recherche);
2. Prendre des mesures raisonnables pour sécuriser les renseignements personnels sur la santé contre l'erreur, la perte, le vol ou une utilisation ou une divulgation non autorisée;
3. Désigner une personne ressource pour répondre aux demandes d'information associées à la protection de la vie privée ainsi qu'aux plaintes;
4. Être transparent relativement aux utilisations de l'information et aux mesures de sécurité; et
5. Rendre compte aux autorités, et possiblement aux participants, des atteintes à la vie privée et/ou à la sécurité.

Les chercheurs doivent minimalement utiliser des contrats pour s'assurer que les Fournisseurs n'utilisent les données des participants que dans le but envisagé, et qu'ils emploient des mesures de protection solides pour la protection de la sécurité et de la vie privée.

Conditions pour le transfert transfrontalier: Les données des participants, transférées vers le Nuage ou parmi des centres de distributions de données dans le Nuage, peuvent circuler à travers des juridictions étrangères. Ceci est permis par les lois canadiennes s'il y a consentement explicite du participant. Sans ce consentement, les chercheurs ont la permission de transférer des informations à un Fournisseur étranger seulement si le participant est avisé de l'endroit où les données vont circuler et du risque de divulgation obligatoire en vertu d'un régime de lois étrangères. La conformité aux normes dans le contexte du Nuage est douteuse lorsque le Fournisseur n'informe même pas le chercheur du lieu où les données peuvent être transférées.

Toutefois, dans certaines provinces, les chercheurs du secteur public font face à une interdiction de transférer des données de recherche vers un Fournisseur étranger sans un consentement explicite.[††]

Obligations éthiques : L'*Énoncé de politique des trois Conseils* requiert une divulgation complète des risques associés à la protection de la vie privée des participants, incluant l'information sur qui aura accès aux données; comment la confidentialité sera-t-elle préservée; qui pourrait avoir l'obligation de divulguer les données; et à qui cette divulgation pourrait-elle être faite. Il s'en suit que les chercheurs devraient faire preuve de transparence envers les participants sur les limites du processus de dé-identification et sur les risques associés au stockage dans le Nuage. Si certaines de ces informations sont initialement inconnues, le processus d'obtention du consentement devrait être bonifié par une supervision éthique continue supplémentaire, la mobilisation du public et l'offre de moyens aux participants de se retirer.

III. Utiliser des contrats pour la protection de la vie privée dans le Nuage

Dans le but d'offrir un service informatique sur demande et abordable à plusieurs clients, la plupart des Fournisseurs commerciaux offrent des Conditions d'Utilisation standards. Notre examen de ces conditions offertes par six Fournisseurs révèle nombre de questionnements concernant la transparence et la sécurité de leurs pratiques de traitement des données. Par exemple, les conditions sont typiquement non négociables et sujettes à une modification unilatérale par le Fournisseur; le lieu où sont entreposées les données, les mesures de sécurité et les pratiques de sous-traitance ne sont pas clairement définis; les Fournisseurs tendent à s'exclure de toute responsabilité relative aux données des clients; et la notification d'une atteinte à la sécurité peut être discrétionnaire, voire absente. À la lumière de ces constats, les chercheurs

[††] Consentement explicite ou autorisation en vertu de la loi ou d'un ministre, voir par ex., Colombie-Britannique *Freedom of Information and Protection of Privacy Act* [RSBC 1996] c 165, s. 33.1; Nouvelle-Écosse *Personal Information International Disclosure Protection Act* [2006] c 3, s. 9.

devront attentivement évaluer si le Nuage est adéquat pour un projet en particulier et comparer les risques et bénéfices des services de différents Fournisseurs. Dans le but de faciliter la conformité aux lois canadiennes et aux standards éthiques, les Fournisseurs devraient contractuellement être contraints :

1. D'offrir des mesures précises pour la protection de la vie privée et de la sécurité, incluant une panoplie complète de garanties techniques, administratives et physiques qui seraient conformes aux standards de l'industrie et sujettes à des vérifications périodiques par un expert indépendant en sécurité informatique.
2. De maintenir un système de surveillance de façon à ce qu'il soit possible de démontrer la conformité aux politiques, aux procédures et aux meilleures pratiques reconnues. S'assurer que tous les individus impliqués dans la prestation de service (incluant les sous-traitants) soient liés par des ententes de non-divulgation ou de confidentialité ayant force exécutoires pendant et après la relation contractuelle.
3. D'utiliser un protocole pour répondre à toute atteinte à la vie privée ou à la sécurité, lequel inclurait d'aviser le chercheur.
4. D'établir des politiques et des procédures afin de garantir l'intégrité des données de recherche et la disponibilité du service durant la vie complète du stockage, de l'analyse et de l'accès (c.-à-d. des procédures de sauvegarde, un plan de recouvrement en cas de désastre, un plan pour la continuité des opérations, le retour des données à la cessation du service).
5. De fournir un préavis suffisant en cas de modification des Conditions d'Utilisation, au cas où les chercheurs aient besoin de transférer leurs données, particulièrement pour des

changements de lieu du stockage des données, de la sécurité et des garanties pour la protection de la vie privée ou pour une suspension de service.

6. D'assumer la responsabilité pour les dommages résultant d'une négligence ou d'une faute du Fournisseur, notamment les atteintes à la vie privée, la sécurité ou l'utilisation de données à mauvais escient.

IV. Améliorer la règlementation de la recherche génomique dans le Nuage

Les législateurs et les décideurs de politiques publiques au Canada peuvent aussi améliorer la protection de la vie privée et promouvoir l'utilisation responsable de la génomique dans le Nuage, en traitant des enjeux suivants:

1. La définition légale des termes comme « identifiable », « codé », « anonyme », « dé-identifié », « renseignement personnel sur la santé», « consentement », « recherche », et « dépositaire » diffèrent d'une province à l'autre. Pareillement, les obligations imposées aux chercheurs, ex., celles d'aviser les autorités chargées de la protection de la vie privée d'une atteinte à la protection des données ou d'imposer des clauses contractuelles spécifiques quand il y a transfert de renseignements personnels sur la santé à un Fournisseur, peuvent aussi varier. Les divergences légales entre les provinces et les pays constituent des barrières à la collaboration pour la recherche dans le Nuage. Les autorités canadiennes chargées de la protection de la vie privée devraient coopérer afin d'aider au développement de directives uniformes pour la protection de la vie privée dans le Nuage et applicables dans tous les provinces et secteurs, et à l'international.
2. Les comités d'éthique chargés d'approuver et de suivre la recherche peuvent manquer d'expertise et de ressources pour examiner les projets de recherches, lorsque celles-ci sont associées à la circulation d'un grand volume de données par le biais de

collaborations internationales impliquant des technologies comme le Nuage. Dans ces cas, des comités d'éthique distincts devraient être créés pour revoir les enjeux liés à la protection de la vie privée et à la sécurité. Une expertise au niveau de ces enjeux devrait minimalement être présente sur les comités d'éthique.

3. La participation à la recherche, particulièrement pour la recherche internationale dans le Nuage, est découragée par le manque perçu de transparence et de contrôle sur le processus de surveillance gouvernementale, autant au Canada qu'à l'étranger. Les gouvernements devraient préciser l'objectif et l'étendue de la surveillance des renseignements personnels sur la santé pour que les Canadiens puissent faire des choix éclairés à propos de leur participation à des recherches.

Remerciements : Gratien Dalpé et Shahad Salman du Centre de génomique et politiques pour la traduction vers français.

Références primaires

1. E.S. Dove et al., “Genomic Cloud Computing: Legal and Ethical Points to Consider” *European Journal of Human Genetics*, 2014.

2. E.S. Dove, Y. Joly, B.M. Knoppers, “International Genomic Cloud Computing: ‘Mining’ the Terms of Service”, in A.S.Y. Cheung, R. Weber, eds. *Privacy and Legal Issues in Cloud Computing*. Cheltenham: Edward Elgar, 2015: 237-260.

3. M. Naveed et al., “Privacy and Security in the Genomic Era”, arXiv.org (2014).

4. “Accessing Health and Social Data in Canada: The Expert Panel on Timely Access to Health and Social Data for Health research and Health System Innovation” Council of Canadian Academies, 2015.

5. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (2014).

6. M. Power, *The Law of Privacy*, (Ontario: LexisNexis: 2013).

7. C. Kuner, *Transborder Data Flow Regulation and Data Privacy Law* (Oxford: Oxford University Press, 2013).

8. W. Jansen, T. Grance, “Guidelines on Security and Privacy in Public Cloud Computing” National Institute of Standards and Technology, Special Publication 800-144 (2001).

9. D. Krebs, “Regulating the Cloud: A Comparative Analysis of the Current and Proposed Privacy Frameworks in Canada and the European Union” (2012) 10 *Canadian Journal of Law and Technology* 29.

10. G. Gunasekara, “Paddling in unison or just paddling? International trends in reforming information privacy law” *Int J Law Info Tech* (Summer 2014) 22 (2).

11. NIH, “NIH Security Best Practices for Controlled-Access Data Subject to the NIH Genomic Data Sharing (GDS) Policy Security Best Practices” 2015, available at <http://www.ncbi.nlm.nih.gov/projects/gap/pdf/dbgap_2b_security_procedures.pdf>.

12. K. El Emam, *Guide to the De-identification of Personal Health Information*, (Taylor & Francis: New York, 2013)

13. Office of the Privacy Commissioner of Canada, “Reaching for the Cloud(s): Privacy Issues Related to Cloud Computing” (2010).