Commissariat
à la protection de
la vie privée du Canada

Métadonnées et vie privée

Un aperçu technique et juridique

Octobre 2014

Table des matières

Introduction

Divers événements récents survenus au Canada et dans d'autres pays nous ont amenés à nous demander si certains organismes gouvernementaux recueillent et utilisent des métadonnées dans le cadre de leurs activités et, le cas échéant, comment ils s'y prennent. Les programmes de collecte de métadonnées aux États-Unis et au Canada ont récemment suscité de grands débats dans les médias. Même si de telles données peuvent être créées et utilisées de façon légale dans le secteur aussi bien public que privé (sous réserve des restrictions et conditions appropriées), il semble y avoir un débat qui persiste sur la nature des métadonnées, ce qu'elles peuvent révéler, et sur le traitement à leur réserver en l'absence d'une disposition législative expresse. Diverses personnes et organisations reconnues continuent de faire valoir que l'on doit établir une distinction entre les métadonnées et le contenu réel des communications et, par conséquent, que les métadonnées sont moins dignes d'être protégées sous l'angle de la vie privée.

De nombreuses sources se sont penchées sur la nature des « métadonnées » et ce qu'elles peuvent révéler. Le Commissariat à la protection de la vie privée du Canada (le Commissariat) a déjà analysé en juillet 2006 les répercussions des métadonnées en lien avec la protection de la vie privée et il a publié une fiche d'information intitulée *Les risques associés aux métadonnées*[1]. En mai 2013, nous avons également publié un rapport de recherche intitulé *Ce qu'une adresse IP peut révéler à votre sujet*[2], qui montre comment la connaissance d'éléments d'information concernant un abonné, notamment le numéro de téléphone et l'adresse IP, peut constituer un point de départ pour retracer les activités en ligne d'une personne. S'appuyant sur ces travaux effectués auparavant par le Commissariat, le présent document propose une analyse technique de ce que peuvent révéler les métadonnées et donne un aperçu de la façon dont les tribunaux ont interprété la notion de métadonnée.

Qu'est-ce qu'une « métadonnée »?

Pour simplifier, disons qu'une métadonnée est une donnée qui fournit de l'information sur une autre donnée. Il s'agit en fait des renseignements qui sont générés lorsqu'on utilise la technologie et qui permettent de situer dans leur contexte (qui, quoi, où, quand et comment) diverses activités. Ces activités peuvent aller de la création d'un document à un appel téléphonique en passant par le clavardage. Dans le contexte des communications, les métadonnées fournissent certaines précisions sur la création, la transmission et la diffusion d'un message. À cet égard, les métadonnées peuvent, par exemple, indiquer la date et l'heure où un appel téléphonique a été fait ou le lieu à partir duquel un courriel a été consulté.

On décrit généralement les métadonnées comme de l'information sur un dossier électronique ou numérique, mais la notion de métadonnée est indéniablement vaste. Étant donné que le débat récent sur la nature et la valeur des métadonnées découle de l'interception de métadonnées associées à des communications, le présent document mettra l'accent sur les métadonnées créées par les communications Internet, filaires et sans fil.

[1] https://www.priv.gc.ca/resource/fs-fi/02_05_d_30_f.asp.
[2] https://www.priv.gc.ca/information/research-recherche/2013/ip_201305_f.asp.

Comme nous le verrons ci-après, la distinction entre une « communication » ou un « contenu », d'une part, et l'information générée par cette communication ou son contenu ou s'y rapportant, d'autre part, n'est pas si claire.

Quelques exemples de métadonnées dans le contexte des communications

Chaque fois que nous communiquons, des métadonnées sont produites. Qu'il s'agisse d'une conversation en face à face avec une personne, de l'envoi de messages textes, de clavardage ou de conversations téléphoniques, certains renseignements concernant cette communication — autres que la communication en soi — sont produits.

En ce qui concerne les communications par Internet ou par téléphone, voici quelques exemples de métadonnées que peuvent générer certaines activités courantes :

Activité	**Métadonnées produites**
	- Numéro de téléphone de l'appelant - Numéro de téléphone composé - Numéro de série unique des appareils téléphoniques utilisés - Heure de l'appel - Durée de l'appel - Emplacement de chaque participant - Numéro de carte d'appel
	- Nom, adresse de courriel et adresse IP de l'expéditeur - Nom et adresse de courriel du destinataire - Renseignements sur le transfert via le serveur - Date, heure et fuseau horaire - Identifiant unique du courriel et des courriels connexes (identifiant de message) - Type de contenu et codage - Dossier de connexion du client de la messagerie avec adresse IP - Format de l'en-tête du client de la messagerie - Priorité et catégorie - Objet du courriel - Statut du courriel - Demande de confirmation de lecture

- Votre nom et les renseignements biographiques indiqués dans votre profil, notamment votre date de naissance, votre ville natale, vos antécédents professionnels et vos centres d'intérêt
- Votre nom d'utilisateur et identifiant unique
- Vos abonnements
- Le lieu où vous vous trouvez
- L'appareil que vous utilisez
- La date et l'heure de l'activité ainsi que le fuseau horaire
- Vos activités, ce que vous aimez, le lieu où vous vous trouvez et les événements auxquels vous assistez

- Votre nom, le lieu où vous vous trouvez, votre langue, les renseignements biographiques indiqués dans votre profil et votre URL
- La date à laquelle vous avez créé votre compte
- Votre nom d'utilisateur et votre identifiant unique
- Le lieu du gazouillis, la date, l'heure et le fuseau horaire
- Le numéro d'identification unique du gazouillis et celui du gazouillis auquel vous répondez
- Le code d'identification des contributeurs
- Le nombre d'abonnés, d'abonnements et de favoris

- Les pages que vous visitez, et quand
- Les données sur l'utilisateur et peut-être les détails de connexion de l'utilisateur avec la fonction de saisie automatique
- Les adresses URL
- Votre adresse IP, votre fournisseur de services Internet, les détails matériels de votre appareil, la version du système d'exploitation et du navigateur
- Les témoins et données en cache provenant des sites Web
- Vos requêtes de recherche
- Les résultats de recherche qui s'affichent
- Les pages que vous visitez par la suite

Ce que peuvent révéler les métadonnées

Selon le contexte, il est parfois difficile de tracer une ligne précise entre une communication et une métadonnée. D'après Michael Morell, ancien haut responsable de la CIA et actuellement membre du groupe de réflexion sur les technologies de communication du président des États-Unis, la distinction entre contenu et métadonnée n'est pas nettement marquée, c'est-à-dire qu'il s'agit plutôt d'un continuum[3]. Comme l'a reconnu en outre l'American Civil Liberties Union, l'information sur le lieu où se trouve une personne obtenue à partir des tours de téléphonie mobile, le nom de l'émetteur ou du récepteur d'un message électronique, ou les achats par Internet, par exemple, [traduction] « ne constituent peut-être pas le contenu de nos communications, mais ils peuvent donner une image très détaillée de notre vie »[4].

La ligne de séparation entre les métadonnées et le contenu réel d'une communication peut dès lors sembler illusoire. La taille, la forme ou la couleur d'une enveloppe sont parfois très révélatrices du message qu'elle renferme. Par exemple, le type d'enveloppe et sa couleur peuvent indiquer s'il s'agit d'un contenu d'ordre personnel ou commercial; l'adresse de l'expéditeur ou le logo figurant sur l'enveloppe peuvent indiquer de qui elle provient; le timbre et le cachet de la poste peuvent révéler la date et le lieu d'expédition; et l'écriture manuscrite, à la différence des adresses produites par ordinateur, peut donner à penser que la correspondance émane d'un particulier plutôt que d'une grande entreprise.

Si l'on transpose cet exemple particulier dans le contexte d'Internet, l'adresse URL représente des instructions de livraison qui précisent l'adresse de la page Web que demande une personne. Autrement dit, il s'agit d'une métadonnée créée lorsque l'utilisateur essaie de consulter un site Web particulier. Toutefois, il peut également s'agir d'un contenu puisque le fait de demander une page Web signifie essentiellement que l'on envoie un message disant « veuillez m'envoyer la page qui se trouve à cette adresse URL ». En outre, une adresse URL révèle exactement quelle page était recherchée et correspond donc exactement au contenu qui a été reçu[5].

> *[traduction] « En fait, l'extraction de métadonnées peut non seulement révéler des renseignements sensibles concernant le passé, mais aussi permettre à un observateur de prédire des actes futurs. Par exemple, des recherches ont montré que l'on peut prévoir le lieu où se trouvera une personne en examinant les constantes dans l'historique des lieux fréquentés par ses amis et connaissances. Un spécialiste de la sécurité a d'ailleurs fait valoir qu'en dressant la liste des appels téléphoniques entre les hauts dirigeants d'une entreprise et une entreprise concurrente, un avocat ou un courtier, on peut avoir une idée d'une éventuelle prise de contrôle de l'entreprise avant que celle-ci ne soit rendue publique ».*
>
> *Metadata : Piecing Together a Privacy Solution, a Report by the ACLU of California,* février 2014, *https ://www.aclunc.org/publications/metadata-piecing-together-privacy-solution*.

[3] S. Ackerman, « NSA review panel casts doubt on bulk data collection claims », *The Guardian*, 14 janvier 2014, http://www.theguardian.com/world/2014/jan/14/nsa-review-panel-senate-phone-data-terrorism.
[4] American Civil Liberties Union of California, *Metadata : Piecing Together a Privacy Solution*, février 2014, p. 3, https://www.aclunc.org/publications/metadata-piecing-together-privacy-solution.
[5] *Ibid*., p. 4.

En effet, les métadonnées peuvent parfois en révéler davantage que le contenu lui-même. À l'époque du numérique, la quasi-totalité des activités en ligne laissent une trace personnelle de quelque nature[6]. Selon l'informaticien Daniel Weitzner, [traduction] « on peut faire valoir que les métadonnées sont plus révélatrices [que le contenu] parce qu'il est en réalité beaucoup plus facile d'établir des corrélations avec des événements du monde réel en analysant les constantes dans un vaste univers de métadonnées qu'en effectuant une analyse sémantique de tous les courriels et de tous les appels téléphoniques d'un individu »[7]. Même les termes saisis dans les moteurs de recherche peuvent être utilisés pour identifier des individus et révéler des renseignements sensibles les concernant. John Battelle a inventé l'expression « base de données des intentions », qu'il décrit comme [traduction] « les résultats globaux de toutes les recherches jamais effectuées, toutes les listes de résultats jamais soumises et toutes les avenues empruntées en conséquence »[8]. Toujours selon Battelle, [traduction] « cette information représente, sous une forme regroupée, le miroir des intentions du genre humain — une énorme base de données sur les désirs, les besoins, les souhaits et les goûts qui peuvent être dévoilés, archivés, retracés et exploités à toutes sortes de fins et dont la production peut être exigée par un tribunal. Il s'agit d'une entité d'un genre nouveau, inconnue auparavant dans l'histoire de la culture, mais il est pratiquement sûr qu'elle est appelée à croître de façon exponentielle jusqu'à la fin des temps »[9].

[traduction] *« Les métadonnées téléphoniques peuvent être extrêmement révélatrices, d'abord au niveau des appels individuels, mais plus particulièrement sous une forme agrégée.*

Même si à première vue ces métadonnées peuvent sembler ne pas représenter beaucoup plus que des renseignements concernant les numéros composés, l'analyse des métadonnées téléphoniques dévoile souvent des renseignements que l'on ne pouvait obtenir autrefois qu'en dépouillant le contenu des communications. Cela revient à dire que les métadonnées sont souvent un substitut du contenu.

Prenons l'exemple le plus simple. Certains numéros de téléphone sont composés dans un but unique, si bien que tout appel révèle des renseignements de base et souvent sensibles à propos de l'appelant. On pense ici notamment aux services d'écoute téléphonique d'aide aux victimes de violence conjugale et de viol. Mais il en existe bien d'autres, comme ceux qui s'adressent aux personnes suicidaires, aux premiers intervenants, aux anciens combattants ainsi qu'aux jeunes gais et lesbiennes. Il y a aussi des services d'écoute téléphonique pour les personnes souffrant de diverses formes de dépendance comme l'alcool, la drogue et le jeu. »

> *Témoignage écrit du professeur Edward W. Felten, Sénat des États-Unis, Committee on the Judiciary, audience sur la surveillance continue de la Foreign Intelligence Surveillance Act, le 2 octobre 2013, http://www.cs.princeton.edu/~felten/testimony-2013-10-02.pdf.*

---

[6] Voir *Klayman c. Obama*, « Brief Amici Curiae of the Electronic Frontier Foundation, the American Civil Liberties Union, and the ACLU of the Nation's Capital in Support of the Appellees » (20 août 2014), USCA, n° 14-5004 aux p. 12 à 15, https://www.eff.org/files/2014/08/20/klayman_amicus_brief.pdf.

[7] E. Nakashima, « Metadata reveals the secrets of social position, company hierarchy, terrorist cells », *The Washington Post*, le 15 juin 2013, http://www.washingtonpost.com/world/national-security/metadata-reveals-the-secrets-of-social-position-company-hierarchy-terrorist-cells/2013/06/15/5058647c-d5c1-11e2-a73e-826d299ff459_story_1.html.

[8] John Battelle, « The Database of Intentions », le 3 novembre 2003, http://battellemedia.com/archives/2003/11/the_database_of_intentions.php.

[9] *Ibid.*

Comme le note aussi un rapport du groupe de réflexion sur le renseignement et les technologies de communication du président des États-Unis, la collecte de métadonnées au fil du temps peut révéler un grand nombre de facettes de la vie privée d'une personne[10]. Le rapport fait observer que [traduction] « la consignation de chaque appel téléphonique fait ou reçu par une personne sur plusieurs années peut être très révélatrice de sa vie privée ». Le rapport recommande au gouvernement des États-Unis de mettre fin à son programme de collecte et de stockage de métadonnées téléphoniques brutes dès que cela est réalisable, notant à cet égard que l'accès par le gouvernement aux dossiers sur les appels téléphoniques peut avoir l'effet d'une douche froide sur les libertés d'association et d'expression, et porter atteinte à la relation entre l'individu et l'État[11].

Enfin, certains poussent le raisonnement encore plus loin et font valoir que les métadonnées ne constituant pas un contenu peuvent même être « une communication privée » au sens du *Code criminel* et de la *Loi sur la défense nationale*. Ces lois renferment d'ailleurs la même définition de l'expression « communication privée[12] » et précisent toutes deux que l'interception d'une communication privée doit respecter certaines conditions pour être légitime. Compte tenu de la jurisprudence de la Cour suprême du Canada, selon laquelle la partie VI du *Code criminel* (se rapportant à l'interception de communications privées) protège non seulement la communication elle-même, mais aussi toute information *connexe* à cette communication qui permettrait d'en dégager la substance ou le sens, certains ont fait valoir que les métadonnées peuvent dans de nombreux cas répondre à ce critère[13].

Ces opinions contrastent avec la position adoptée par certaines institutions gouvernementales. Certains organismes gouvernementaux, en effet, ne considèrent pas que les métadonnées sont apparentées à l'information véhiculée par le contenu ou à une communication. Par exemple, dans un document déposé auprès d'un tribunal juridique en février 2014, le gouvernement a adopté la position selon laquelle l'expression « métadonnées » désigne l'information associée à une télécommunication pour identifier, décrire, gérer ou acheminer cette télécommunication, en tout ou en partie, de même que les moyens par lesquels on l'a transmise, mais elle exclut toute information ou tout élément d'information susceptible de révéler l'objet de cette communication ou la totalité ou une partie de son contenu[14]. De même, dans son témoignage devant un comité parlementaire, le ministre de la Justice a émis l'opinion que les métadonnées relatives à la transmission sont à distinguer du contenu[15].

---

[10] *Liberty and Security in a Changing World*, rapport et recommandations du groupe de réflexion du président sur le renseignement et les technologies de communication, 12 décembre 2013, http://www.whitehouse.gov/sites/default/files/docs/2013-12-12_rg_final_report.pdf.

[11] *Ibid.*, page 117.

[12] L'expression « communication privée » désigne une « communication orale ou télécommunication dont l'auteur se trouve au Canada ou son destinataire se trouve au Canada, et qui est faite dans des circonstances telles que son auteur peut raisonnablement s'attendre à ce qu'elle ne soit pas interceptée par un tiers ».

[13] Craig Forcese, « Law, Logarithms and Liberties: Legal Issues Arising From CSEC's Metadata Program » (à paraître), mars 2014, Les Presses de l'Université d'Ottawa, page 14.

[14] *Ibid.*, page 12.

[15] Canada, Comité permanent de la justice et des droits de la personne (1er mai 2014), 1150 (hon. Peter MacKay).

Les métadonnées en tant que « renseignements personnels »

Les deux lois fédérales du Canada en matière de protection des renseignements personnels définissent l'expression « renseignements personnels » de façon générale comme des renseignements concernant un individu identifiable. Ainsi, la *Loi sur la protection des renseignements personnels et les documents électroniques* (LPRPDE) définit les renseignements personnels comme étant « tout renseignement concernant un individu identifiable, à l'exclusion du nom et du titre d'un employé d'une organisation et des adresse et numéro de téléphone de son lieu de travail ». Quant à la *Loi sur la protection des renseignements personnels,* elle définit les renseignements personnels comme « les renseignements, quels que soient leur forme et leur support, concernant un individu identifiable » et propose divers exemples de ce qui constitue au sens de la Loi des renseignements personnels conformes au caractère général de la définition initiale.

La définition de « renseignement personnel » a donné lieu au fil du temps à une interprétation large[16]. Bien que l'on dispose d'exemples clairs de ce qui constitue un renseignement personnel au sens de ces lois, un renseignement qui ne semble pas à première vue concerner un individu en particulier peut également, combiné avec d'autres renseignements et dans certains contextes, se révéler un renseignement personnel. Par exemple, dans la décision *Gordon c. Canada (Ministre de la Santé)*[17], la Cour fédérale a convenu que le champ « province » dans une base de données concernant les réactions aux effets indésirables de médicaments au Canada constitue un renseignement personnel au sens de la *Loi sur la protection des renseignements personnels*. Selon la Cour, un renseignement concerne un « individu identifiable » lorsqu'il y a une possibilité sérieuse qu'un individu puisse être identifié au moyen du renseignement, que ce renseignement soit pris seul ou en combinaison avec d'autres renseignements disponibles.

Parmi les exemples de renseignements personnels dans le contexte technologique, mentionnons, selon le cas, différents types de renseignements biométriques comme les empreintes digitales et vocales, les renseignements de géolocalisation recueillis grâce au système GPS placé à bord des véhicules des employés et les renseignements recueillis au moyen d'étiquettes d'identification par radiofréquence pour suivre la trace d'objets ou d'individus[18].

Le Commissariat a aussi déterminé que lorsque des renseignements en apparence anodins sont combinés à d'autres renseignements accessibles, ils peuvent constituer des renseignements personnels et parfois donner une idée assez exacte des activités, des idées, des opinions et du mode de vie de la personne concernée. Par exemple, une adresse IP (adresse de protocole Internet) peut s'avérer un renseignement personnel si associée à un individu identifiable, et peut être assez révélatrice des activités d'une personne sur Internet. À vrai dire, comme le souligne le rapport du Commissariat intitulé *Ce qu'une adresse IP peut révéler à votre sujet*, une adresse IP, combinée à des renseignements de base sur un abonné de services de télécommunications, peut révéler entre autres ses intérêts, ses penchants, les personnes qu'il fréquente, de même que les voyages qu'il fait.

Si suffisamment de métadonnées permettent de disposer d'une foule de renseignements fort précieux concernant la même personne, l'accumulation de métadonnées dans certains contextes peut parfois

---

[16] *Dagg c. Canada (Ministre des Finances), [1997]* 2 R.C.S., dissidents, 403 paragr. 68*; Canada (Commissaire à l'information) c. Canada (Bureau canadien d'enquête sur les accidents de transport et de la sécurité des transports),* 2006 CAF 157*; Canada (Commissaire à l'information) c. Canada (Commissaire de la Gendarmerie royale du Canada), [2003]* 1 R.C.S. 66, 2003 CSC 8, paragr. 23.

[17] *Gordon c. Canada (Ministre de la Santé)*, 2008 C.F. 258.

[18] Commissariat à la protection de la vie privée du Canada, Bulletin d'interprétation : *Renseignements personnels* (mis à jour en octobre 2013), https://www.priv.gc.ca/leg_c/interpretations_02_f.asp.

permettre également d'identifier l'individu associé à ces données. Par exemple, dans le cadre d'un processus appelé « analyse des réseaux sociaux » ou « établissement de la chaîne de relations sociales » — qui requiert la création d'un diagramme du réseau humain gravitant autour d'un individu en particulier —, les analystes peuvent identifier toutes les personnes qui se trouvent à un ou deux degrés de distance de la personne concernée. Un diagramme illustrant la chaîne de relations sociales peut montrer comment les gens dans le voisinage du réseau sont reliés les uns aux autres. Même le fait que des gens travaillent à une faible distance d'un suspect identifié peut élargir rapidement le réseau de relations jusqu'à inclure certaines personnes n'ayant peut-être pas connaissance de l'existence du suspect. Un exemple de ce genre cité dans l'article intitulé « Connecting the Dots : Tracking Two Terrorist Suspects »[19], montre qu'il a suffi de surveiller les activités de deux suspects, notamment l'information à propos des appels téléphoniques qu'ils faisaient, des courriels qu'ils envoyaient et des réunions qu'ils organisaient, pour commencer à se faire une idée de leur réseau personnel. Avec suffisamment de renseignements, on a pu identifier la cible et d'autres individus faisant partie du réseau.

Le 4 août 2006, AOL a publié, à des fins de recherche, un fichier renfermant 20 millions de mots clés de recherche utilisés par plus de 650 000 utilisateurs sur une période de trois mois. D'après AOL, il n'y avait aucun renseignement permettant d'identifier un individu dans les données publiées — l'entreprise avait d'ailleurs remplacé le nom des utilisateurs par un numéro d'identification aléatoire. Toutefois, comme ce numéro était le même pour toutes les recherches effectuées par un individu donné, ce qui signifie que les individus pouvaient être associés à leur compte et à leur historique de recherche, on a pu dans certains cas les identifier, parfois en recoupant les données avec d'autres renseignements publics (p. ex. les annuaires téléphoniques). AOL a déclaré par la suite que les données n'avaient pas été examinées adéquatement avant leur diffusion et a reconnu que ce genre de requêtes de recherche en soi pouvait parfois comporter des renseignements permettant d'identifier un individu (p. ex. le nom, le numéro de sécurité sociale, l'adresse et d'autres éléments d'information que les gens peuvent chercher à savoir).

Dans un exemple, après avoir passé au crible les données publiées, le New York Times *a été en mesure d'identifier une utilisatrice. La personne derrière le numéro d'utilisateur 4417749, attribué par AOL pour protéger son anonymat, était en fin de compte une veuve de 62 ans de l'État de la Géorgie, dont l'identité a été dévoilée grâce à une analyse des requêtes de recherche de cette utilisatrice d'AOL supposément anonyme.*

Voir http://www.nytimes.com/2006/08/09/technology/09aol.html?pagewanted=all&_r=0, et http://www.nytimes.com/imagepages/2006/08/08/business/09aol-graphic.html.

Aussi convaincant que soit cet exemple, le fait de prendre des décisions concernant les gens d'après l'information qu'ils recherchent en ligne pourrait toutefois mener à des conclusions erronées sur l'auteur des recherches. En effet, la femme jumelée au numéro d'utilisateur 4417749 d'AOL faisait régulièrement des recherches en ligne sur les problèmes de santé d'autres personnes, notamment pour trouver de l'information sur l'arrêt du tabagisme. Par conséquent, si on avait fait un lien entre ces problèmes de santé et la femme correspondant au numéro d'utilisateur 4417749, les résultats auraient été inexacts.

[19] Valdis Krebs, « Connecting the Dots: Tracking Two Identified Terrorists », orgnet.com, http://orgnet.com/tnet.html.

Le fait de savoir où se trouve/où se trouvait une personne à un moment donné peut également révéler de l'information sur un individu identifiable, y compris des renseignements que cette personne pourrait préférer garder secrets. Par exemple, une photo prise au moyen d'un appareil équipé d'un GPS intégré peut révéler le lieu et l'heure où le cliché a été pris, de même que l'identifiant unique de l'appareil, qui pourrait être un téléphone intelligent. De nombreux appareils permettent cette collecte par défaut, et souvent les utilisateurs ne sont pas au courant de cette pratique, laquelle pourrait avoir de graves conséquences. Par exemple, un dénonciateur, un journaliste ou un dissident politique comptant sur la protection que lui confère l'anonymat pour dénoncer des malversations perpétrées par une entreprise, un criminel ou un gouvernement pourrait être mis en danger par cette collecte de données par défaut.

Les appareils mobiles sont vraiment personnels. Les gens les apportent généralement partout où ils vont et les utilisent pour une foule d'activités (courriels, messages textes, appels téléphoniques, photographies, etc.). Or, ces appareils peuvent transmettre et transmettent effectivement de plus en plus des données de géolocalisation précises. Leurs propriétaires se servent parfois de cette fonction de manière délibérée, comme lorsqu'ils s'inscrivent à un service tel que Foursquare reposant sur la géolocalisation, mais il arrive souvent que les données soient transmises à l'insu ou sans le consentement de l'utilisateur (p. ex. lorsque les téléphones intelligents équipés d'un GPS diffusent leur emplacement). L'emplacement de l'appareil peut également être calculé à partir des données des tours de téléphonie mobile.

Des études ou expériences réalisées au fil des années ont montré le caractère extrêmement sensible et unique des données de géolocalisation, notamment :

a) *En 2010, le site Web « I Can Stalk U » (http://icanstalku.com) a vu le jour. Ce site aujourd'hui fermé analysait les photos affichées en ligne, pour trouver les étiquettes de géolocalisation (métadonnées de localisation) et il indiquait ensuite l'emplacement précis associé au message tweeté;*

b) *En 2012, le programmeur d'antivirus John McAfee a été arrêté au Guatemala après qu'une photo géolocalisée eut révélé l'endroit où il se trouvait;*

c) *En 2013, des chercheurs ont publié l'étude intitulée « Unique in the Crowd : The privacy bounds of human mobility », montrant qu'il suffisait de disposer de quatre points spatio-temporels choisis de façon aléatoire pour identifier un individu (http://www.nature.com/srep/2013/130325/srep01376/pdf/srep01376.pdf).*

Un chercheur a conclu que nos déplacements sont si uniques que quatre points de données de géolocalisation sont suffisants pour identifier 95 % des gens[20] tandis qu'un autre a affirmé que [traduction] « toute série de données livrant assez d'information sur les gens pour retenir l'intérêt des chercheurs en comporte suffisamment pour être dépersonnalisée »[21]. La possibilité d'identifier un individu à partir de métadonnées

[20] Jason Palmer, « Mobile location data 'present anonymity risk' », BBC.com, le 25 mars 2013, http://www.bbc.co.uk/news/science-environment-21923360.
[21] Pete Warden, « Why you can't really anonymize your data », O'Reilly Strata, le 17 mai 2011, http://strata.oreilly.com/2011/05/anonymize-data-limits.html.

s'accroît lorsque plusieurs types de métadonnées sont combinés puis associés à d'autres renseignements accessibles. Des entreprises du secteur privé mettent au point des moteurs de recherche qui exploitent de multiples flux de données, y compris des réseaux de médias sociaux et des sources de données ouvertes des gouvernements et du secteur privé. Dans le cadre d'une expérience, à partir d'un seul gazouillis géolocalisé choisi de manière aléatoire, un journaliste a été capable de déterminer où se situait l'école de l'auteure, de savoir qu'il s'agissait d'une école d'immersion française, mais aussi de connaître les lieux où elle rencontrait ses amis et où elle gardait des enfants, et de déterminer qu'il s'agissait d'une joueuse de soccer et qu'elle faisait probablement du ski ou de la planche à neige[22].

L'étude Metaphone

À partir d'un petit échantillon de données (avec la participation de 546 volontaires sur une période de quelques mois seulement), des chercheurs de l'Université de Stanford ont réussi à démontrer, sans aucune ambiguïté, que l'analyse de métadonnées téléphoniques peut révéler des renseignements extrêmement sensibles sur les individus. En utilisant des sources publiquement accessibles pour identifier leurs contacts et se basant sur des appels téléphoniques uniques, les chercheurs ont été en mesure de déterminer que les individus communiquaient avec des services de santé, des services financiers et des services juridiques de même qu'avec des organisations religieuses, entre autres. Comme on pouvait s'y attendre, les habitudes d'appel sont encore plus révélatrices, puisque les chercheurs ont réussi à conclure à l'existence de problèmes de santé (p. ex. sclérose en plaques, arythmie cardiaque), et également à découvrir qu'une personne possédait des armes à feu. Dans certains cas, ils ont réussi à corroborer ces conclusions en utilisant des sources d'information publiques. Pour en savoir davantage sur l'étude et ses résultats, consultez le site http://webpolicy.org et les messages en ligne concernant le projet Metaphone.

Par conséquent, la nature révélatrice des métadonnées remet de plus en plus en question l'idée que ces renseignements seraient moins dignes de protection sous l'angle de la vie privée du fait qu'il faut les distinguer des renseignements sous forme de contenu et qu'ils seraient en conséquence moins sensibles. Elle remet également en question l'idée voulant que les métadonnées mériteraient moins d'être protégées sous l'angle de la vie privée puisqu'elles sont déjà publiquement accessibles à d'autres sous une forme ou sous une autre. L'interprétation du terme « métadonnées » dans la jurisprudence appuie l'idée que les métadonnées recèlent plus d'information qu'on pourrait le croire à première vue.

Une semaine suffit

Dans le cadre d'une expérience menée récemment, des chercheurs ont été en mesure de dresser un portrait incroyablement détaillé de la vie d'un homme après avoir recueilli des métadonnées d'une application mobile pendant seulement une semaine. Ils ont en effet été en mesure d'attacher une référence temporelle à quelque 15 000 éléments et non seulement de déterminer les habitudes de travail et les intérêts personnels de celui-ci, mais également de connaître son réseau social à partir de l'utilisation qu'il faisait de son téléphone et de ses courriels. Ils ont aussi été en mesure de voir les sites Web qu'il avait consultés et les recherches qu'il avait effectuées, ainsi que de connaître l'objet, l'expéditeur et le destinataire de chacun de ses courriels. Les chercheurs ont également trouvé le mot de passe de ses comptes Twitter, Google et Amazon et auraient pu changer les paramètres de ces comptes et même commander des articles à partir de son compte Amazon (ce qu'ils n'ont pas fait). Pour en savoir plus sur cette expérience et les conclusions de celle-ci, voir : http://www.statewatch.org/news/2014/jul/bits-of-freedom-on-the-metadata-of-your-phone.pdf.

[22] Gillian Shaw, « The withering of secrecy: Technology reveals your life on social media », *Vancouver Sun*, le 31 mars 2014, http://www.vancouversun.com/technology/personal-tech/secrecy+Technology+reveals+your+life+social/9676829/story.html.

Comment les tribunaux considèrent-ils les « métadonnées »?

Un examen de l'interprétation du terme « métadonnées » dans la jurisprudence canadienne révèle qu'il y a relativement peu d'interprétation judiciaire de la nature des « métadonnées » en général, et encore moins dans le contexte des communications.

La plus grande partie de cette jurisprudence peu abondante concerne des procédures au civil dans le cadre desquelles une partie a demandé la production de métadonnées intégrées à un support quelconque, notamment un disque dur d'ordinateur, une photographie, un relevé téléphonique ou un document électronique. Ces affaires portent principalement sur la mesure dans laquelle les métadonnées associées aux documents électroniques ou aux relevés devraient être produites conformément à l'obligation d'une partie de produire tous les documents pertinents pour un litige et sur la question de savoir si les métadonnées électroniques peuvent être considérées comme des « renseignements électroniques » au sens des règles de procédures civiles applicables ou si elles sont pertinentes et constituent des éléments de preuve dans le cadre du litige en question[23].

Toutefois, certaines affaires s'inscrivent dans d'autres contextes, notamment les demandes d'ordonnance concernant la communication publique ou la préservation de documents ou d'information[24], les demandes de mandat pour permettre à la Couronne d'accéder à certains documents ou renseignements et les contestations de la constitutionnalité de perquisitions et de saisies de renseignements[25].

Les tribunaux ont inclus parmi les métadonnées le tampon de la date et de l'heure sur une lettre ou l'en-tête de télécopies. Certains ont décrit les métadonnées comme étant une sorte de talon de paye. Comme l'a mentionné un tribunal, [traduction] « les talons de paye peuvent renfermer des renseignements à propos du nombre de jours de travail effectués, des jours travaillés et du taux horaire, à l'instar des métadonnées qui renferment, par exemple, des renseignements concernant la date et l'heure où un courriel a été créé et envoyé, ou bien la date et l'heure où une personne a consulté un site Web »[26].

Dans d'autres affaires, les tribunaux ont poussé plus loin leur analyse et interprété le terme pour inclure :

- des renseignements concernant les numéros de téléphone utilisés pour envoyer et recevoir des messages textes et l'heure des appels ou de la transmission des messages[27];
- des renseignements associés à une télécommunication qui identifient, décrivent, gèrent ou acheminent cette télécommunication en tout ou en partie[28];

---

[23] Voir, par exemple, *Peter Laushway c. Albert Messervey and Sobeys Group Inc.,* 2014 NSCA 7, confirmée dans 2013 NSSC 47; *Frangione c. Vandongen*, [2010] N° 2337; *Warman c. National Post Company*, 2010 ONSC 3670; *Bishop (Litigation Guardian of) c. Minichiello,* 2009 BCSC 358; *Hummingbird c. Mustafa*, 2007 CanLII 39610; *Spar Aerospace Ltd. c. Aeroworks Engineering Inc.,* 2007 CarswellAlta 1156, confirmée dans 2008 ABCA 47; *Desgagne c. Yuen et al.,* 2006 BCSC 955; *Baldwin Janzen Insurance Services (2004) Ltd. c. Janzen*, [2006] BCJ N° 753; *Ireland c. Low*, 2006 BCSC 393; *Nicolardi c. Daley*, [2002] N° 595; *Reichman c. Toronto Life Publishing Co.,* [1988] N° 1727.

[24] *United States of America c. Fraser,* 2014 BCSC 227; en ce qui concerne la jurisprudence américaine, voir également *O'Neill c. City of Shoreline,* 2010 Washington, LEXIS 870 (Washington. le 7 octobre 2010); *Armstrong c. Executive Office of the President, Office of Admin.,* 303 U.S. App. D.C. 107, 1 F.3d 1274, 1993 U.S. App. LEXIS 20527 (le 13 août 1993).

[25] *R. c. Vu*, 2013, CSC 60.

[26] *Frangione c. Vandongen*, [2010] O.J. n° 2337.

[27] *Canada (Attorney General) c. B. (A.),* 2014 NLCA 8.

[28] *United States of America c. Fraser*, 2014 BCSC 227.

30, rue Victoria, 1er étage, Gatineau (Québec) K1A 1H3 • Sans frais : 1-800-282-1376 • Télécopieur : (819) 994-5424 • ATS : (819) 994-6591
www.priv.gc.ca • Suivez-nous sur Twitter : @priveeprivacy

- des renseignements qui peuvent fournir de l'information convergente concernant un document, y compris ceux qu'une personne a essayé de supprimer ou de caviarder[29];
- des renseignements comme l'heure à laquelle un dossier a été créé, quel utilisateur était connecté au système au moment de la création d'un dossier et la période pendant laquelle un dossier est demeuré ouvert[30];
- des renseignements concernant l'objet présumé d'un document et les circonstances entourant la création de ce dernier[31].

Bien que l'interprétation du terme « métadonnées » dans la jurisprudence canadienne semble dépendre du contexte sous-jacent, il semble y avoir un consensus selon lequel les métadonnées sont considérées comme des données qui fournissent de l'information sur d'autres données et qui, dans de nombreux cas, peuvent permettre de tirer des conclusions sur la conduite ou les activités d'un individu. Certains tribunaux ont amorcé une analyse plus exhaustive de la notion de métadonnées et des raisons pour lesquelles l'accès à ces métadonnées est devenu une question de plus en plus controversée. Dans l'arrêt *Abougoush c. Sauve*[32], la Cour suprême de la Colombie-Britannique a reconnu que la collecte de métadonnées peut parfois révéler des détails intimes du mode de vie d'un individu, même si ces métadonnées sont recueillies au moyen d'un appareil photo numérique :

> On peut avoir une idée des capacités physiques de l'utilisateur d'un appareil photo sur une base quotidienne ou sur plusieurs jours. Plus particulièrement, les métadonnées peuvent permettre de déterminer que le plaignant a la capacité nécessaire, par exemple, pour être actif tout au long d'une journée donnée et faire ensuite une promenade sur la plage le soir, passer une soirée en boîte de nuit jusqu'à une certaine heure, tout en demeurant capable de nager le lendemain matin.

Mais les tribunaux canadiens ne sont pas les seuls à reconnaître que les métadonnées peuvent être très révélatrices. La jurisprudence américaine révèle quelques exemples où des tribunaux juridiques ont eu l'occasion de formuler certains commentaires sur la nature révélatrice des métadonnées.

La Cour suprême des États-Unis a elle-même récemment noté à quel point les métadonnées sur l'emplacement d'une personne peuvent en dire long à son sujet. Dans *United States c. Jones*[33], la Cour a déterminé que l'installation sur un véhicule d'un système de positionnement global (GPS), en vue de surveiller les mouvements de ce véhicule, constitue une recherche protégée par le quatrième amendement de la Constitution américaine. Dans son opinion concordante, la juge Sotomayor note l'ampleur de l'information qui peut être tirée des données sur l'emplacement du GPS : [traduction] « La surveillance par GPS crée un enregistrement précis et complet des mouvements publics d'une personne, qui contient de nombreux renseignements sur ses associations familiales, politiques, professionnelles, religieuses et sexuelles ». La juge Sotomayor cite une décision d'un tribunal de l'État de New York[34] décrivant les conclusions qu'on peut tirer de données GPS :

---

[29] *Warman c. National Post Company*, 2010 ONSC 3670.
[30] *Desgagne c. Yuen et al.*, 2006 BCSC 955.
[31] *Big Pond Communications 2000 Inc., c. Kennedy*, [2004] O.J. n° 820.
[32] *Abougoush c. Sauve*, 2011 BCSC 885.
[33] *United States c. Jones*, 132 S.Ct. 945 (2012).
[34] *People c. Weaver*, 12 N. Y. 3d 433, (2009).

> [traduction]Les données [GPS] divulgueront entre autres… des déplacements dont on imagine sans peine la nature incontestablement privée : déplacement chez le psychiatre, le chirurgien plasticien, la clinique d'avortement, le centre de traitement du SIDA, le club de danseuses nues, l'avocat au pénal, le motel loué à l'heure, la réunion syndicale, la mosquée, la synagogue ou l'église, le bar gai, et ainsi de suite[35].

Par ailleurs, des décisions anciennement secrètes concernant des programmes de surveillance de la NSA aux États-Unis jettent un éclairage sur les raisons qui poussent la United States Foreign Intelligence Surveillance Court (FISC) à autoriser la surveillance gouvernementale des métadonnées, et fournissent des commentaires sur ce que peuvent révéler les métadonnées[36]. À titre d'exemple, le juge dans une décision lourdement expurgée de la FISC, probablement rendue en juillet 2004, confirme la constitutionnalité d'un programme de collecte de métadonnées brutes sur Internet, mais prend acte toutefois que cela impose [traduction] « une collecte d'un type beaucoup plus vaste que les autres demandes portant sur des enregistreurs des numéros de téléphone composés ou des dispositifs d'interception et de localisation[37] ». Dans une autre décision expurgée, un juge de la FISC voit les métadonnées, dans le contexte des communications, comme [traduction] « de l'information au sujet de la communication, et non la communication en elle-même, y compris les numéros composés, la durée de l'appel, les adresses de protocole Internet, les adresses de courriel et de l'information semblable au sujet de la livraison de la communication, plutôt que le message entre deux parties[38] ».

De plus, dans *Klayman c. Obama*[39], où le plaignant contestait la constitutionnalité de la collecte de métadonnées téléphoniques brutes par le gouvernement des États-Unis, la Cour de district du district fédéral de Columbia est allée jusqu'à dire que la collecte de métadonnées [traduction] « révèle une foule de détails sur les relations familiales, politiques, professionnelles, religieuses et sexuelles d'un individu ». La Cour a affirmé que :

> [traduction] La nature des métadonnées n'a pas changé au fil du temps, mais l'omniprésence du téléphone a modifié de façon spectaculaire la quantité de renseignements qui sont maintenant accessibles et, qui plus est, ce que ces renseignements peuvent divulguer au gouvernement sur la vie des gens. (Les communications par téléphone cellulaire et par messages textes sont si généralisées que certaines personnes peuvent les considérer comme des moyens essentiels ou des instruments nécessaires pour s'exprimer, voire s'identifier.)

Ces cas sont une illustration que les métadonnées contiennent des renseignements qui peuvent révéler énormément d'information au sujet des gens, même si elles ne constituent pas une communication en elles-mêmes.

Les métadonnées et les attentes raisonnables en matière de vie privée

La jurisprudence de la Cour suprême du Canada donne à penser que les individus peuvent avoir des attentes raisonnables en matière de protection de la vie privée pour ce qui concerne les renseignements générés par un ordinateur et l'usage d'Internet qui révèlent des renseignements biographiques de base.

---

[35] *Ibid.,* pages 441 et 442.
[36] Bryce Clayton Newell, « The Massive Metadata Machine: Liberty, Power, and ~~Secret~~ Mass Surveillance in the U.S. and Europe » (2014), vol. 10 :2, *I/S : A Journal of Law and Policy for the Information Society*, p. 490 à 492.
[37] [nom du cas expurgé], n° PR/TT [expurgé] (FISA Ct.), p. 80, http://www.dni.gov/files/documents/1118/CLEANEDPRTT%201.pdf.
[38] 37 [nom du cas expurgé], n° PR/TT [expurgé], (FISA Ct.), p. 1, http://www.dni.gov/files/documents/1118/CLEANEDPRTT%202.pdf.
[39] *Klayman c. Obama*, motif et ordonnance datés du 16 décembre 2013; voir également *American Civil Liberties Union c. James R. Clapper*, motif et ordonnance datés du 27 décembre 2013.

30, rue Victoria, 1er étage, Gatineau (Québec) K1A 1H3 • Sans frais : 1-800-282-1376 • Télécopieur : (819) 994-5424 • ATS : (819) 994-6591
www.priv.gc.ca • Suivez-nous sur Twitter : @priveeprivacy

Dans l'arrêt *R. c. Vu*, la Cour suprême du Canada a reconnu l'ampleur des renseignements personnels que l'on peut déduire à partir des métadonnées en faisant valoir que « les ordinateurs stockent d'immenses quantités de données, dont certaines, dans le cas des ordinateurs personnels, touchent "l'ensemble de renseignements biographiques d'ordre personnel" » puis, comme il est indiqué ci-dessus, que les ordinateurs renferment des données qui sont générées automatiquement à partir desquelles des détails personnels sensibles peuvent être tirés. Mais le tribunal a également souligné que les ordinateurs peuvent conserver des fichiers et des données même après que les utilisateurs pensent les avoir détruits et que « les ordinateurs qui sont connectés à Internet servent de portails à une quantité presque infinie de données qui sont partagées entre différents utilisateurs et stockées presque n'importe où dans le monde. De même, un ordinateur connecté à un réseau permettra à la police d'avoir accès à des renseignements se trouvant dans d'autres appareils. »

> *« [I]l arrive souvent que les logiciels de traitement de textes génèrent automatiquement des fichiers temporaires permettant aux analystes de reconstituer l'élaboration d'un fichier et d'avoir accès à des renseignements indiquant qui a créé le fichier et qui y a travaillé. De même, la plupart des navigateurs utilisés pour consulter Internet sont programmés pour conserver automatiquement des renseignements concernant les sites Web que l'utilisateur a visités dans les semaines précédentes, ainsi que les syntagmes de recherche qu'il a utilisés pour y accéder. Normalement, ces renseignements peuvent aider l'utilisateur à retracer ses démarches cybernétiques. Dans le contexte d'une enquête criminelle, toutefois, ils peuvent également permettre aux enquêteurs d'avoir accès à des détails intimes concernant les intérêts, les habitudes et l'identité de l'utilisateur, à partir d'un dossier que ce dernier a créé sans le savoir. O. S. Kerr, « Searches and Seizures in a Digital World » (2005), 119 Harv. L. Rev. 531, p. 542-543. Les renseignements de ce genre ne possèdent pas d'équivalents dans le monde concret qui est celui des autres types de contenants. »*
>
> *R. c. Vu,* 2013 CSC 60.

Dans l'arrêt *R. c. Spencer*, la Cour suprême du Canada a soutenu que le nom et l'adresse d'une abonnée, combinés à une adresse IP particulière, avaient fourni à la police plus qu'un simple nom et une adresse, puisqu'il s'agissait plutôt de « l'identité d'une abonnée aux services Internet à qui correspondait une utilisation particulière de ces services ». La Cour a reconnu que les individus peuvent avoir des attentes raisonnables en matière de vie privée pour ce qui concerne les renseignements reliant leur identité à une métadonnée, en l'espèce une adresse IP. La Cour a déterminé que le fait que la police avait obtenu auprès du fournisseur de services Internet des renseignements sur l'abonnée correspondant à une adresse IP, et ce, sans mandat, constituait pour une fouille non autorisée par la loi, et violait par conséquent l'article 8 de la *Charte*.

> *« […] la Cour a adopté dans le passé, [sur la question de la fouille ou de la perquisition] une approche large et fonctionnelle, en examinant le lien entre la technique d'enquête utilisée par la police et l'intérêt en matière de vie privée qui est en jeu. La Cour a examiné non seulement la nature des renseignements précis recherchés, mais aussi la nature des renseignements qui sont ainsi révélés. »*
>
> *R. c. Spencer,* 2014 CSC 43.

L'arrêt *Spencer* ce veut un prolongement logique de la jurisprudence antérieure de la Cour suprême du Canada en matière de vie privée, d'ordinateurs et de renseignements générés par ordinateur. Dans l'arrêt *R. c. Morelli*, la majorité des juges de la Cour suprême du Canada avaient estimé que les ordinateurs utilisés à des fins personnelles, quel que soit l'endroit où ils se trouvent ou leur propriétaire, « contiennent souvent notre correspondance la plus intime. Ils renferment les détails de notre situation financière, médicale et personnelle. Ils révèlent même nos intérêts particuliers, préférences et propensions, enregistrant dans l'historique et la mémoire cache tout ce que nous recherchons, lisons, regardons et écoutons dans l'Internet ». Dans l'arrêt *R. c. Cole*, la Cour a ajouté : « Cela est particulièrement vrai lorsque, comme en l'espèce, l'ordinateur sert à naviguer sur le Web. Les appareils connectés à Internet "révèlent nos intérêts particuliers, préférences et propensions, enregistrant dans l'historique et la mémoire cache tout ce que nous recherchons, lisons, regardons et écoutons dans l'Internet". Les renseignements de ce genre se situent au cœur même de "l'ensemble de renseignements biographiques" protégés par l'article 8 de la *Charte.* »

Ces décisions judiciaires mettent en évidence le fait que l'usage de l'ordinateur par les individus, en particulier lorsqu'il est connecté à Internet, peut constituer un « renseignement biographique de base » auquel sont associées des attentes raisonnables élevées en matière de protection de la vie privée. Que cette jurisprudence fasse également référence aux syntagmes de recherche, aux adresses URL, à l'historique de navigation, à la mémoire cache et à la création non délibérée de dossiers donne également à penser que des intérêts importants en matière de protection de la vie privée sont associés à certains types de métadonnées.

Conclusion

Tout en prenant acte de l'importance du contexte, les tribunaux ont observé à maintes reprises que les métadonnées peuvent être très révélatrices au sujet d'une personne et appellent une protection sous l'angle de la vie privée.

Les institutions gouvernementales qui recueillent ces renseignements, ou envisagent de le faire, ne peuvent sous-estimer l'ampleur de l'information que les métadonnées peuvent révéler au sujet d'un individu. Il en va de même pour les organisations du secteur privé auxquelles on demande de divulguer de telles données aux institutions gouvernementales, y compris les organismes d'application de la loi. Compte tenu de l'omniprésence des métadonnées et des conclusions convaincantes qui peuvent en découler concernant des individus en particulier, les institutions gouvernementales et les organisations du secteur privé doivent baliser leurs activités de collecte et de communication en fonction de méthodes et de normes appropriées qui sont adaptées au niveau de sensibilité potentiel des métadonnées dans des circonstances données.