

FEUILLE-INFO

SUR LA PROTECTION DE LA VIE PRIVÉE

Les mégadonnées et le droit à la vie privée

Grâce à de nouveaux outils permettant de combiner et d'analyser des informations, les chercheurs peuvent désormais découvrir, dans de grands ensembles de données, des tendances et associations qui étaient auparavant impossibles à percevoir. Ces grands ensembles de données et les outils et pratiques analytiques employés pour déceler les tendances s'appellent les « mégadonnées ». Les entreprises du secteur privé analysent souvent des mégadonnées à des fins de marketing et de développement de produits; les organismes du secteur public, quant à eux, s'y intéressent en vue d'améliorer l'élaboration des politiques et des programmes afin que ceux-ci s'appuient sur de meilleures données probantes.

La présente feuille-info vise à renseigner le public sur les mégadonnées et leur incidence possible sur la vie privée des particuliers.

Les mégadonnées pourraient permettre aux pouvoirs publics de mieux évaluer la qualité et l'efficacité de services et de programmes tels que les soins de santé, les services sociaux, la sécurité publique et les transports. Cependant, elles suscitent également des inquiétudes concernant la vie privée et la protection des renseignements personnels des particuliers.

Le Bureau du commissaire à l'information et à la protection de la vie privée de l'Ontario (CIPVP) est chargé de surveiller l'application de la **Loi sur l'accès à l'information et la protection de la vie privée**, de la **Loi sur l'accès à l'information municipale et la protection de la vie privée** et de la **Loi sur la protection des renseignements personnels sur la santé**. Les organismes régis par ces lois, tels que les ministères du



Information and Privacy
Commissioner of Ontario

Commissaire à l'information et à la
protection de la vie privée de l'Ontario

gouvernement, les municipalités, les services de police, les fournisseurs de soins de santé et les conseils scolaires, doivent recueillir, utiliser et divulguer des renseignements personnels dans le respect de ces lois.

La présente feuille-info vise à renseigner le public sur les mégadonnées et leur incidence possible sur la vie privée des particuliers.

QUE SONT LES MÉGADONNÉES?

Le terme « mégadonnées » s'applique à différents concepts et peut être difficile à définir. En général, il désigne les ensembles de données qui ne sont pas faciles à gérer ou à interpréter par des moyens traditionnels en raison du volume, de l'irrégularité ou de la complexité des données. Les mégadonnées comprennent souvent trois caractéristiques :

- **Volume** – Il y a tellement de données et de points de données qu'il peut être difficile de les parcourir et de les comprendre.
- **Vitesse** – Les données sont générées et ajoutées rapidement à l'ensemble de données et peuvent être mises à jour en temps réel.
- **Variété** – Les données comprises dans l'ensemble de données peuvent se présenter sous de nombreuses formes et provenir d'un grand nombre de sources. Par exemple, un ensemble de mégadonnées peut comprendre des données structurées, telles que celles que l'on retrouve dans une base de données, et des données non structurées, comme des ensembles de messages publiés dans les médias sociaux ou des billets de blogue.

L'« analytique des mégadonnées » désigne les méthodes et outils divers qui sont employés pour générer des perspectives à partir des mégadonnées. Les grands ensembles de données peuvent être soumis à des algorithmes afin de relever des tendances et des associations, établir des règles pour automatiser la prise de décisions et même prévoir les résultats d'une démarche, y compris le comportement d'une personne.

QUELLE EST L'INCIDENCE DES MÉGADONNÉES SUR LA VIE PRIVÉE?

Les mégadonnées peuvent comprendre des renseignements personnels, c'est-à-dire des renseignements consignés sur une personne qui peut être identifiée, comme le nom, l'adresse, les opinions ou les antécédents médicaux, et c'est souvent le cas. Comme dans le cas de tout ensemble de données qui contient des renseignements personnels, les lois ontariennes sur la protection de la vie privée prévoient des précautions que les institutions des pouvoirs publics doivent prendre afin que ces renseignements personnels soient recueillis, utilisés, conservés et divulgués de façon appropriée.

Les mégadonnées peuvent comprendre des renseignements personnels, c'est-à-dire des renseignements consignés sur une personne qui peut être identifiée, comme le nom, l'adresse, les opinions ou les antécédents médicaux, et c'est souvent le cas.

Cependant, les mégadonnées soulèvent des questions uniques sur le plan de la vie privée en raison de leur volume et de leur complexité, ainsi que des procédés analytiques employés pour les interpréter.

Surveillance

Les renseignements que recueillent les pouvoirs publics peuvent fournir des indications très détaillées sur les citoyens, comme leur lieu de résidence, les personnes qu'ils fréquentent et leurs convictions politiques ou religieuses. Lorsque des projets de mégadonnées font intervenir la collecte et la mise en commun de grandes quantités de données (qu'elles soient accessibles au public ou d'usage interne dans des organismes), il est possible d'apprendre beaucoup de choses sur les gens et leur vie. Lorsque c'est le cas, et lorsqu'on fait le suivi de ces données sur une certaine période, il s'agit d'une forme de surveillance, c'est-à-dire une activité qui contrôle les comportements, actions ou communications. La surveillance peut être autorisée dans bien des cas, comme à des fins de sécurité publique et de recherche dans le secteur de la santé. Cependant, les mégadonnées peuvent donner lieu à des activités de surveillance injustifiées si elles sont employées de façon irresponsable. Chaque fois que les pouvoirs publics décident de mener des activités liées à des mégadonnées contenant des renseignements personnels, les avantages de ces activités doivent l'emporter sur le risque possible pour la vie privée.

Exactitude des résultats

La collecte et le regroupement d'une grande quantité de renseignements personnels qui ne sont pas nécessairement apparentés font en sorte que l'analytique des mégadonnées peut dégager des tendances qui sont en fait de simples coïncidences. Cela peut se produire en raison de l'abondance de données analysées. À cause de la multiplication des variables en cause, certains éléments de données qui semblent être associés ne le sont pas vraiment.

Comprendre la différence entre corrélation et causalité est essentiel pour évaluer l'exactitude de l'analytique des mégadonnées. Il y a corrélation lorsqu'il existe un lien statistique entre les valeurs de deux variables d'un ensemble de données; par exemple, ces valeurs ont tendance à augmenter ou à diminuer en parallèle. Il y a causalité, un lien plus fort, lorsque ces deux variables sont nécessairement associées, et que si l'une change, l'autre doit changer aussi. Lorsqu'il y a corrélation, cependant, un changement qui se produit dans une variable ne cause pas nécessairement un changement dans l'autre variable; leur association peut être due au simple hasard. Par exemple, la découverte, dans un ensemble de données, d'une tendance selon laquelle les personnes aux cheveux roux qui conduisent une voiture rouge sont impliquées dans des accidents de la route ne signifie pas qu'avoir les cheveux roux et conduire une voiture rouge ont causé ces accidents. Ces faits peuvent être en corrélation, mais cela ne suffit pas pour établir un lien de causalité.

Le biais dans les ensembles de données

On dit parfois que les mégadonnées sont plus objectives et moins biaisées que les analyses de données traditionnelles car « toutes » les données sont analysées. Cependant, même s'il était possible de recueillir toutes les données pertinentes, l'analytique des mégadonnées peut présenter des biais. Les pratiques de collecte ou de production de données peuvent comprendre des biais implicites qui excluent certains groupes de personnes ou leur accordent une importance accrue.

Par exemple, si les pratiques d'embauche d'un organisme ont fait en sorte que des personnes ayant des antécédents semblables ont été embauchées plus souvent, les données sur ces embauches refléteront simplement ces décisions. Si ces données sont analysées afin d'en dégager des caractéristiques communes en vue de sélectionner des candidats futurs, les biais des pratiques d'embauche antérieures pourraient être renforcés. La prise de décisions ou la sélection automatique des candidats effectuée en fonction de cette analyse de données pourrait exclure d'office des personnes qualifiées en raison d'une caractéristique qui n'a aucune pertinence.

Les pratiques de collecte ou de production de données peuvent comprendre des biais implicites qui excluent certains groupes de personnes ou leur accordent une importance accrue.

Utilisation non autorisée

Les pouvoirs publics ou toute institution assujettie aux lois ontariennes sur la protection de la vie privée à qui l'on remet des renseignements personnels doivent indiquer à quelle fin ces renseignements serviront et pourquoi. Les projets fondés sur les mégadonnées, cependant, sont plus complexes. Des renseignements recueillis à une fin précise peuvent être combinés à des renseignements recueillis à d'autres fins. Ce regroupement et les analyses auxquelles les données combinées seront soumises peuvent faire partie d'une étude ou d'un programme dont n'ont jamais entendu parler les personnes qui ont fourni leurs renseignements. En règle générale, en vertu des lois ontariennes sur la protection de la vie privée, les renseignements personnels doivent être utilisés uniquement aux fins auxquelles ils ont été recueillis ou à une fin que la loi autorise. À moins d'être bien gérées, la collecte, l'utilisation et la divulgation d'ensembles de mégadonnées pourraient enfreindre ces lois.

QUE FAIT LE CIPVP POUR PROTÉGER MA VIE PRIVÉE?

Le CIPVP est résolu à protéger votre vie privée. Les mégadonnées posent de nouveaux défis pour les institutions ontariennes tout en leur offrant de nouvelles occasions à exploiter. Nous sommes persuadés qu'avec de la diligence, une bonne planification et de l'information sur les questions uniques que soulèvent les mégadonnées et leur analytique, les mégadonnées peuvent être utilisées dans le respect de la vie privée. Le CIPVP continuera de collaborer étroitement avec les institutions des pouvoirs publics et publiera des documents d'orientation sur les pratiques exemplaires en matière de protection de la vie privée et la

conformité aux lois pertinentes dans le cadre des projets de mégadonnées.

QUELS SONT MES DROITS EN VERTU DES LOIS ONTARIENNES SUR LA PROTECTION DE LA VIE PRIVÉE?

Si vous avez des questions sur les renseignements personnels que recueille une institution des pouvoirs publics ou leur utilisation, consultez le site Web de l'institution ou communiquez avec elle pour obtenir de plus amples renseignements. Pour recevoir une copie des renseignements qu'une telle institution possède à votre sujet, vous pouvez déposer une demande d'accès à l'information auprès de l'institution.

Pour en savoir davantage sur la présentation d'une demande d'accès à l'information et sur les mesures à prendre pour protéger votre vie privée, visitez le site Web du CIPVP à www.ipc.on.ca.