PROTECTION DE LA VIE PRIVÉE : PLUSIEURS PAYS PUBLIENT UNE DÉCLARATION COMMUNE SUR L'EXTRACTION DE DONNÉES

Publié le 12 septembre, 2023

Catégories: Perspectives, Publications

Le 24 août 2023, le Commissariat à la protection de la vie privée du Canada et ses homologues de onze autres pays (les « **organismes de réglementation** ») ont publié une déclaration commune sur l'extraction de données et la protection des renseignements personnels (la « **déclaration commune** »)[1]. Cette déclaration commune expose les principaux risques pour la vie privée que pose l'extraction de données et établit la manière dont les entreprises de médias sociaux (« **EMS** ») et les autres exploitants de sites Web qui hébergent des données personnelles accessibles au public devraient protéger les utilisateurs contre l'extraction illégale de données. Elle comprend aussi des mesures que les individus peuvent prendre pour réduire les risques connexes.

L'extraction de données consiste à utiliser un programme ou une application informatique pour obtenir le précieux contenu de la base de données d'un site Web. Les données personnelles ainsi extraites le sont souvent dans un but illicite : cyberattaques, pourriels et marketing non sollicité, vol d'identité, voire utilisation non autorisée à des fins policières ou politiques[2].

Les lois canadiennes sur la protection de la vie privée interdisent généralement aux entreprises d'extraire arbitrairement des renseignements personnels de sources Web. Nous avons d'ailleurs publié plusieurs bulletins sur l'application de la loi fédérale à l'extraction de données[3].

L'extraction de données connaît un regain d'intérêt, attribuable en partie à l'émergence de systèmes d'IA générative entraînés à l'aide de données accessibles publiquement sur Internet. Par exemple, Google a récemment modifié sa politique de confidentialité, qui l'autorise désormais à utiliser des renseignements accessibles publiquement pour entraîner ses modèles d'IA, notamment Google Bard[4]. En outre, plusieurs actions collectives ont été intentées aux États-Unis contre des concepteurs d'outils d'IA générative qui auraient recueilli arbitrairement ce type de renseignement[5].

Dans leur déclaration commune, les organismes de réglementation affirment que les EMS et les autres sites Web sont tenus de protéger les renseignements personnels des individus contre l'extraction illégale de données. Ils invitent donc les entreprises à atténuer les risques en adoptant des mesures de contrôle à couches multiples sur les plans technique et procédural. Voici quelques-unes de leurs recommandations :

- Affecter une équipe au choix et à l'application de mesures de protection des utilisateurs contre l'extraction de données.
- Appliquer à certains comptes une limite de visites (horaire ou quotidienne) en cas d'activité inhabituelle, et surveiller les nouveaux comptes pour détecter toute activité suspecte.
- Prendre des mesures pour détecter les robots, par exemple avec des tests CAPTCHA[6], et bloquer les adresses IP associées à l'extraction de données.
- Prendre les mesures juridiques qui s'imposent si une extraction de données est soupçonnée ou confirmée.
- Aviser les personnes touchées et les organismes de réglementation si nécessaire, dans les cas où l'extraction peut constituer une atteinte à la protection des données.
- Aider les utilisateurs à faire des choix éclairés en leur présentant, en toute transparence, les réglages de confidentialité de la plateforme et leurs incidences potentielles.
- Vérifier que les mécanismes internes de traitement des renseignements personnels sont conformes aux règles de protection de la vie privée.

Bien que ces énoncés soient présentés comme des recommandations, la déclaration commune indique qu'ils peuvent être juridiquement contraignants à certains endroits ou être interprétés comme tels par les tribunaux et les organismes de réglementation. Les organisations dont le site Web présente des renseignements personnels doivent donc s'interroger sur la nécessité de revoir leurs mesures de sécurité des données pour respecter les recommandations. Voici quelques sources de renseignements personnels accessibles publiquement :

- Sections de commentaires, forums et salons de clavardage;
- Profils publics;
- Profils semi-publics auxquels des fraudeurs pourraient avoir accès (p. ex. quand on accepte une invitation sur une application de rencontres ou un réseau social);
- Bases de données comportant des coordonnées ou d'autres renseignements personnels.

En outre, la déclaration commune encourage les individus à agir pour protéger leurs renseignements personnels et réduire le risque d'extraction de données. Ils peuvent notamment (i) lire les politiques de confidentialité des EMS et des autres sites Web; (ii) limiter la quantité de renseignements qu'ils publient en ligne; et (iii) apprendre à utiliser les paramètres de confidentialité pour restreindre les renseignements accessibles au public.

Une approche plus musclée?

Les organismes de réglementation ont transmis la déclaration commune directement aux grandes EMS[7].

Celles-ci sont invitées à démontrer, d'ici un mois, comment elles respectent les recommandations.

Les incidents d'extraction de données ne sont pas rares, et les organismes canadiens de protection de la vie privée enquêtent proactivement sur les entreprises qui extraient illégalement des renseignements personnels de sources publiques. Pensons à l'exemple récent de Clearview AI, une société américaine qui a extrait des milliards d'images sur le Web pour créer une base de données de reconnaissance faciale qu'elle vendait aux forces de l'ordre[8]. Au terme d'une enquête conjointe, en 2021, les organismes de réglementation canadiens ont conclu que l'entreprise avait obtenu et utilisé ces images sans le consentement des personnes y figurant et avait donc enfreint les lois fédérales et provinciales sur la protection des renseignements personnels[9].

À notre connaissance, cependant, aucune autorité canadienne n'a encore enquêté sur le défaut, par un exploitant de site Web public, d'*empêcher les tiers* d'extraire des données.

Ce que les entreprises doivent retenir

- À beaucoup d'endroits, les renseignements personnels « disponibles publiquement » sont protégés par la loi; nul ne peut s'en servir librement.
- Les EMS et les autres sites Web qui hébergent des renseignements personnels doivent protéger activement les utilisateurs contre l'extraction illégale, par exemple en appliquant les mesures techniques recommandées pour protéger les données accessibles publiquement.
- Les individus doivent agir pour réduire le risque d'extraction de données.
- À certains endroits, un incident d'extraction de données constitue une atteinte à déclaration obligatoire.
- Les entreprises ont intérêt à suivre l'évolution du cadre réglementaire et à revoir leurs politiques et procédures en conséquence.

Si vous avez des questions sur l'extraction de données, les lois canadiennes sur la protection des renseignements personnels et la meilleure façon de se conformer aux recommandations, un membre de notre groupe Protection de la vie privée et des données se fera un plaisir de vous aider.

[1] Commissariat à la protection de la vie privée du Canada, Déclaration commune sur l'extraction de données et la protection des renseignements personnels (août 2023). La déclaration commune est signée par les responsables de la protection de la vie privée des pays suivants : Australie, Argentine, Canada, Chine, Colombie, Jersey, Mexique, Maroc, Nouvelle-Zélande, Norvège, Suisse, Royaume-Uni.

[2] Commissariat à la protection de la vie privée du Canada, Conclusions en vertu de la LPRPDE n°2021-001 (février 2021). [**Clearview AI**]
[3] Voir quelques-uns de nos bulletins sur l'extraction de données ici, ici et ici.
[4] Google, Politique de confidentialité de Google (consultée le 30 août 2023).

[5] Voir par exemple l'action collective contre Google et l'action collective contre OpenAI.
[6] Un CAPTCHA est un test de Turing automatisé qui permet de différencier un humain d'un ordinateur.
[7] Les destinataires sont Alphabet Inc. (YouTube), ByteDance Ltd (TikTok), Meta Platforms, Inc. (Instagram, Facebook et Threads), Microsoft Corporation (LinkedIn), Sina Corp. (Weibo) et X Corp. (X, auparavant Twitter).
[8] *Clearview AI*.
[9] *Clearview AI*. Voir les bulletins de McMillan sur la décision et les ordonnances subséquentes des organismes de réglementation provinciaux.

par Robbie Grant et Laurene Oliveira (stagiaire en droit)

Mise en garde