# Mégadonnées, petits secrets? – Pas facile d’assurer la confidentialité des données personnelles à l’ère des réseaux sociaux et de l’open data | Le Net Expert Informatique

## Mégadonnées, petits secrets? – Pas facile d’assurer la confidentialité des données personnelles à l’ère des réseaux sociaux et de l’open data

**Vous croyez qu’il suffit d’enlever les informations nominatives qui figurent dans une banque de données pour en assurer la confidentialité? Détrompez-vous. Non seulement est-ce insuffisant, mais même avec des mesures beaucoup plus serrées, la confidentialité d’une banque de données peut être facilement mise à mal par recoupement avec d’autres sources d’informations, notamment celles que les individus disséminent à tout vent sur le Net. Voilà la leçon de prudence livrée par Anne-Sophie Charest, professeure au Département de mathématiques et de statistique, aux participants du colloque «Big Data, le défi du traitement des données», présenté le 29 octobre sur le campus.**

La professeure Charest compte au nombre de la dizaine de spécialistes que le Centre de recherche en données massives de l’Université Laval et l’Institut technologies de l’information et sociétés avaient réunis pour discuter du potentiel et des défis des mégadonnées, ces banques d’information si volumineuses et si complexes qu’elles exigent des méthodes de traitement particulières. D’entrée de jeu, Anne-Sophie Charest a rappelé les termes du contrat qui lie les chercheurs et les gens qui acceptent de participer à des enquêtes ou à des études. «On assure aux participants que les données resteront confidentielles et qu’elles ne seront utilisées qu’à des fins statistiques. Par contre, les chercheurs rendent publics des études ou des rapports à partir de ces informations et ils partagent même les données avec d’autres chercheurs ou avec la population. Le défi, qui existait même avant l’avènement du Big Data, est de concilier ces deux objectifs contradictoires.»

La solution intuitive, qui consiste à supprimer les informations nominatives, ne suffit pas à blinder une banque de données. À preuve, la professeure Charest a cité le cas du Massachusetts qui avait accepté, à la fin des années 1990, que les dossiers médicaux anonymisés des 135 000 employés de l’État soient mis à la disposition des chercheurs. Le gouverneur William Weld avait alors assuré que la confidentialité était garantie étant donné que les noms, les adresses et les numéros d’assurance sociale des employés avaient été supprimés. Une étudiante-chercheuse du MIT, Latanya Sweeney, aujourd’hui professeure à l’Université Harvard, avait toutefois trouvé une brèche de taille. En recoupant cette banque de données avec la liste électorale, elle a démontré qu’elle pouvait associer une bonne partie des dossiers médicaux à la personne correspondante. «Elle a même fait imprimer le dossier médical du gouverneur et elle l’a fait livrer à son bureau», raconte la professeure Charest.

Des organismes comme Statistique Canada et l’Institut de la statistique du Québec travaillent fort à assurer le respect de leur promesse de confidentialité, souligne Anne-Sophie Charest. La chercheuse explore, elle aussi, de nouvelles façons de compliquer le travail des personnes mal intentionnées qui tentent d’extraire des informations personnelles des bases de données publiques ou des publications qui en découlent. Ces différentes méthodes présentent toutefois un inconvénient important: si on les applique trop rigoureusement, on réduit l’accès aux données, ce qui n’est guère dans le ton en cette ère de l’open data, et on limite l’information utile qu’on peut en tirer.

Autre problème, il ne suffit plus que la confidentialité d’une base de données soit intrinsèquement protégée, il faut qu’elle le soit en tenant compte des recoupements possibles avec les autres sources d’informations. «Il est très difficile de prédire quelle information pourrait causer du tort au répondant si elle était rendue publique. L’approche de la confidentialité différentielle offre toutefois un compromis intéressant, estime-t-elle. Elle promet aux répondants qu’une tierce personne ne pourra rien apprendre de plus sur eux qu’ils acceptent ou non de participer à l’enquête. Cette nuance est importante considérant toute l’information que chaque personne diffuse maintenant sur elle-même. On ne peut plus fonctionner en vase clos.»

Par Jean Hamann
Source: Université Laval

Pour plus d’informations:
Organisation:Université Laval
Adresse:1160 Université Laval
Québec, Québec
Canada, G1V 0A6
www.ulaval.ca

---



---



Source :
http://www.rimq.qc.ca/detail_news.php?ID=551819&titre=Pas+facile+d%27assurer+la+confidentialit%C3%A9+des+donn%C3%A9es+personnelles+%C3%A0+l%27%C3%A8re+des+r%C3%A9seaux+sociaux+et+de+l%27open+data&cat=;71;21