

Géopolitis : Big Data : que fait-on de nos données ?

Transcription

5 LE CONTEXTE

10 Bienvenue sur Geopolitis. On connaissait Big Brother. Désormais, il y a Big Data, c'est un peu la même famille ! Disons que c'est surtout la même peur face à des technologies qui parfois nous dépassent, parce qu'on ne les maîtrise pas bien soi-même. Le problème avec Big Brother, c'était : « qui nous surveille ? » L'inquiétude avec Big Data, c'est : « qui utilise nos données personnelles ? » et pourquoi ? De tout cela, il sera question à Davos lors du prochain Forum Économique Mondial, dont le thème 2015 est « le nouveau contexte global ».

15 Alors c'est quoi Big Data, en anglais « grandes données ». Et bien, cela consiste à rassembler, à partir de tout ce qui circule dans le monde numérique, un maximum d'informations, ensuite à les traiter, puis, c'est le but, à croiser ces informations pour en tirer des conclusions, des analyses, qui sont d'autant plus pertinentes qu'elles sont le fruit d'un maximum de sources. Tout de suite, un exemple : en 2009, année où éclatait la grippe H1N1, Google, pour ne pas le citer, avait mis au point un système de détection basé sur des mots-clés tels que « médicaments, toux et fièvre ». Résultat, en croisant 20 toutes ces données, Google avait pu dresser l'état des lieux de la contagion, avant même les autorités sanitaires qui devaient attendre, pour leur part, que l'information remonte des hôpitaux ou des cabinets médicaux. C'est un beau résultat, mais acquis sur des données qui étaient bien des données personnelles. D'où la question que pose Géopolitis : que Big Data constitue un système performant, certes, mais à quel prix, et avec quelles garanties pour nous et pour nos précieuses données?

25 Vous vous souvenez peut-être de ce film, *Minority Report*, de Steven Spielberg, avec Tom Cruise. Un scénario où l'on arrête un homme non pas pour ses actes, mais pour des délits ou des crimes qu'on pense qu'il va commettre. Dans la réalité, voilà qu'un professeur américain de criminologie affirme, en cette année 2014, que, grâce à sa méthode statistique, méthode qu'il a mise au point en croisant des 30 milliers et des milliers de données, et bien, il est en mesure de prédire si un détenu libéré sur parole sera impliqué dans un nouveau crime ou délit. Problème : le taux d'erreur serait de 25 %. On imagine les conséquences ! Mais le fait est là : plus la masse de données recueillies, croisées et traitées, est grande, plus on a de chance de définir un profil ou une tendance ou même un comportement. Les géants de la grande consommation et du marketing le savent bien, qui offrent des publicités ciblées, 35 presque en temps réel, en fonction du Big Data qu'elles récoltent auprès de leurs clients. On l'aura compris : parmi les géants du numérique, c'est la lutte au couteau, pour recueillir le maximum de nos données.

40 Voici, par exemple, des images du tout premier centre de traitement de données de Facebook, en dehors du sol américain. C'est en Suède, dans la bourgade de Lulea, non loin du cercle polaire. C'est là que Facebook a installé ses serveurs informatiques géants, dans une région où la température moyenne est de 2 degrés, ce qui facilite le refroidissement des ordinateurs. C'est dans de tels endroits que se stockent toutes les données. Le contrôle de ces données, c'est le nerf de la guerre du Big Data.

LE REPORTAGE

Un gigantesque trafic

5 Chaque jour, il s'échange, dans le monde, 145 milliards de courriels. Quotidiennement, on compte, sur Google, 4 milliards et demi de recherches. Et, tous les jours, sur YouTube, on met en ligne 104 000 heures de vidéo. Et il y a en moyenne 552 millions d'utilisateurs de Facebook. Vous êtes friands d'autres statistiques ? Et bien, quand on évoque Big Data, l'unité de mesure, c'est le téra-octet. Un téra-octet, c'est à peu près ce qu'il y a dans 570 000 livres contenant, chacun, 300 pages. Donc, un
10 téra-octet, c'est l'équivalent de 171 milliards de pages ! Ajoutons enfin que, d'après les spécialistes, ce flot de données va quasiment doubler tous les deux ans.

Applications en tout genre

15 C'est un exemple tout récent, il concerne la progression du virus Ebola : un groupe de chercheurs et informaticiens de Boston a développé un algorithme qui dresse effectivement une image de la maladie se déplaçant à travers l'Afrique de l'Ouest. 9 jours avant qu'Ebola soit déclaré à un stade épidémique en Guinée, le groupe avait déterminé, à partir de tout ce qui circule sur Internet, que ce pays allait
20 être touché. Les compagnies d'assurance, elles aussi, ont recours au Big Data, pour développer des modèles de prédiction, sur l'anticipation de certaines maladies ou sur la gestion des coûts de santé. Voilà qu'émerge un nouveau concept, celui de cybersanté, source de biens des débats. Moins de débats en revanche à New York lorsque les services municipaux engrangent toutes sortes de données et déterminent, par exemple, quels sont les immeubles présentant des risques d'incendie.

25

La Suisse : un coffre-fort numérique

30 Et voici que la Suisse convertit ses abris militaires en coffres-forts. Pas d'argent, de bijoux ou de lingots d'or dans ces gigantesques grottes et cavernes, mais des serveurs par milliers ; ces machines hébergent et actualisent en permanence les données numériques de toute personne ou toute entité qui souhaite par exemple garder un double de ses informations, par peur d'une panne informatique, d'un tremblement de terre ou d'une attaque commise par un hacker ou un groupe terroriste. C'est ainsi qu'une société privée a récupéré l'ancien quartier général souterrain de l'armée suisse, un bunker
35 de 15 000 m² situé à 200 m à l'intérieur de la montagne de granit proche d'un village au coeur du pays, le tout se trouve à 1000 mètres sous la cime de la montagne, protégé par des portes anti-nucléaires. Depuis l'affaire Snowden et le scandale de la NSA, les clients se pressent au portillon, enfin, devant les portes en acier de 4 tonnes chacune.

40