

BIG DATA le gouvernement invisible

Author : helene-molinari

Date : 23 décembre 2013

Imaginez un peu ! Des machines qui nous gouverneraient à partir de données récoltées dans la masse, sans qu'on s'en rende compte. Des machines qui nous attribueraient un profil, dont on ne saurait comment se défaire. Un monde où vos désirs et vos envies n'émaneraient plus de vous mais vous seraient continuellement suggérés ; où votre vie consisterait à cocher des cases pour mieux être redirigé ; où l'erreur, la discussion et l'opposition n'existeraient plus... Vous trouvez ça flippant ? Nous aussi ! Parce que le futur, c'est maintenant.

Rien de nouveau, me direz-vous : « *Mais oui, on parle de Big Brother !* » Ce gros œil qui nous observe, ce pouvoir central qui espionne notre vie privée. Mais non, ici nous parlons de Big Data, et lui, ce que tu dis sur le fond, ça ne l'intéresse pas vraiment. Ce n'est pas tant qu'il espionne notre « vie privée », il est plutôt du genre à collecter nos métadonnées. Big Data s'imposerait alors comme un outil de gouvernement des masses par le profilage.



Antoinette Rouvroy a nommé ce phénomène la « gouvernamentalité algorithmique ». Juriste de formation, elle est chercheuse du FNRS au Centre de Recherche en Information, Droit et Société (CRIDS). « Ce à quoi je suis arrivée », dit-elle, *«c'est : un mode de gouvernement qui est nourri essentiellement de données brutes, c'est-à-dire insignifiantes pour les humains, mais quantifiables, et qui affectent les comportements humains sur le mode de l'alerte, qui provoquent du réflexe, plutôt que sur le mode de la réflexivité. »*

La première étape pour bien comprendre ce qui se cache derrière ce concept tient bien dans le fait que ce type de gouvernement ne se base pas sur l'individu et la surveillance, mais sur la traçabilité et le profil.

Signes vs signaux

Chaque jour, nous laissons des traces. Nous produisons, volontairement ou involontairement, des métadonnées. Par exemple, à chaque fois que nous sommes sur Facebook : à quelle heure j'ai posté tel message contenant tant de mots-clés, avec qui je suis en lien, le fait que je « like » autant de fois des choses que tel autre a postées, etc. Tout cela fait partie d'une gigantesque masse de données disponibles par défaut. Et Facebook ne représente qu'une infime partie de tout ce que vous émettez chaque jour. On parle de « cloud computing », tout ce qu'on trouve dans « les nuages », sans forme et sans historicité.

Là-dessus, « on » (le « on » crucial qui reste encore à définir) fait tourner des algorithmes. Un algorithme équivaut à une recette de cuisine, une liste d'opérations à faire, qui a pour but d'attribuer à ces masses de traces insignifiantes une nouvelle signification, mais de façon purement quantitative, en détectant les occurrences simultanées de différents types de données. « *Ça vous donne des sortes d'images, des profils ou des patterns de comportements – c'est-à-dire que vous allez avoir des modélisations de comportements humains, d'achats, de tout ce que vous voulez, en fait !* » explique Antoinette Rouvroy.

Des graphes relationnels peuvent alors être établis, permettant d'identifier des réseaux, des comportements, voire des formes de vie. « *Pour les moteurs de recherche, ça fonctionne de la même façon. Par exemple PageRank sur Google, cet algorithme qui hiérarchise les contenus lorsque vous introduisez une requête. Ça ne fonctionne pas par une évaluation du contenu des documents, mais simplement par une évaluation du nombre de citations. Et on s'en fout complètement des contenus. Pourtant, chaque fois que vous faites une référence dans un article, vous le faites toujours dans un certain contexte, pour critiquer, pour dire que vous êtes d'accord ou que c'est complètement absurde, etc. Tout ce contexte disparaît complètement au profit de la seule citation. Ça vaut pour les personnes, les documents, ça vaut pour tout.* » Les applications pour le marketing fonctionnent de la même façon. « *Vous pouvez avoir un mot dans votre statut Facebook qui va appeler telle ou telle publicité. Indépendamment du fait que votre propos était justement de dire que ça ne vous intéressait pas du tout.* »

Reprenons. Vous allez sur Internet et vous commentez quelque chose, ou vous postez un article. En faisant cela, vous laissez des traces, mais vous ne savez pas exactement lesquelles. Car tout est retiré de son contexte, explosé en une multitude de données brutes qui ne veulent rien dire toutes seules : ce sont les métadonnées, stockées et disponibles, qui fonctionnent comme des signaux. Quelque part, une machine, entraînée à repérer ces signaux les récupère, applique un (ou plusieurs) algorithme(s) et fait ressortir un profil. Ce profil correspond en fait à un réassemblage de données, il constitue une forme de vie à laquelle vous allez ou non être par la suite réassocié, mais qui n'a plus rien à voir avec vous au départ. À l'arrivée, peu importe le contenu ou le sens. Vous avez été profilé.

« On ne sait jamais comment on va être interprété par ces dispositifs. On perd complètement à la fois la maîtrise des trajectoires des données qu'on émet, et les codes d'intelligibilité. C'est à la fois massif et personnalisant. Par exemple, dans le domaine du marketing : ça s'adapte à vos goûts, au fur et à mesure, et à la limite, même, vos désirs vous précèdent. On vous envoie des pubs à l'avance. Le système sait des choses sur vous que vous-même ignorez. Vous n'avez plus besoin d'exprimer vos désirs, ni même de les façonner, pour que vous soient présentées des solutions. »

La « gouvernementalité algorithmique » ne passe donc pas par le prisme de la conscience humaine et ne présuppose pas non plus des individus capables d'entendement et de volonté. Nous ne parlons donc plus d'individu, mais de « profil ». *«Le profil, c'est quoi ? Ça ne se rapporte pas à une personne identifiée ou identifiable, mais à une forme de vie. Ça cible les comportements possibles et vise une certaine forme de dangerosité, de risque, d'opportunité. C'est une projection dans l'avenir qui ne représente personne en particulier. On va par exemple essayer de vous personnaliser comme consommateur sur internet, on va adapter les offres à votre profil individuel. Mais celui-ci est construit sur une base de modèles et de données qui n'ont rien à voir avec vous, récoltés absolument partout et qui ont servi à construire des modèles de comportement dans lesquels, pour une raison ou pour une autre, vous, vous tombez. C'est à la fois massif, individuel et personnalisant. »*

Les drones, de la théorie à l'application

Les drones incarnent parfaitement la gouvernementalité algorithmique. Ils représentent son bras armé, la police du futur qui s'expérimente dans le ciel du Pakistan. Grégoire Chamayou met tout cela en lumière dans son ouvrage *Théorie du drone*. Il y montre comment des personnes sont profilées sous la forme de vie « terroriste ». Les cibles abattues ne sont pas identifiées mais elles ont été profilées. On a récolté leurs métadonnées, établi un profil et estimé que telle personne constituait un nœud dans un réseau « terroriste ». Ce n'est pas monsieur untel qu'on va tuer parce qu'on sait qu'il a fait ça ou qu'il veut faire ça : on va tuer une personne dont on ne connaît pas l'identité mais qui a été identifiée par les trajectoires qu'il a suivies, par les mots-clé qu'il a introduits sur Internet, ou parce qu'il a vendu du pain à quelqu'un qui, lui, aurait pu être terroriste et s'est déplacé d'une manière qui ne correspond pas au profil « vendeur de pain »... *«Et comme Grégoire Chamayou l'explique, on n'élimine pas seulement le nœud, mais aussi ses ramifications. C'est-à-dire que les personnes qui, suite à la frappe du drone, viennent*

porter secours à la personne qui a été touchée, vont être identifiées aussi comme potentiellement liées à cette personne. Et puis après, troisième frappe, on va suivre le cortège funéraire et on va le bombarder, comme ça on est sûrs de tous les avoir ! » ajoute Antoinette Rouvroy.

On tombe alors dans une auto-justification : le comportement correspond à certains égards au comportement d'un terroriste, donc je suis terroriste. Et peu importe la vérité : c'est fiable, mais ce n'est pas vrai. Aujourd'hui au Pakistan, demain à votre porte.

Big Data n'est pas Big Brother, il est bien pire

À ce stade, l'inquiétude commence à vous envahir. Vous n'êtes pourtant qu'à la moitié du problème. Les machines dont nous vous parlons ont tendance à devenir auto-apprenantes. L'humain, incapable de suivre leurs calculs et leur logique, devient optionnel. *« Au début, on tague les informations pour entraîner la machine et puis après, progressivement, la machine apprend d'elle-même, précise Antoine Rouvroy. C'est ce qu'on appelle le machine learning : elles apprennent à détecter ce qui est signifiant et ce qui ne l'est pas. »* Elles seront capables de générer leurs propres algorithmes, sans action humaine préalable.

Il n'y a donc plus de contrôle. *« L'autorité n'est plus assumée par aucune figure concrète, il n'y a pas un Big Brother qui vous impose une fois pour toute une norme rigide. »* Big Data n'impose pas de limite bien identifiable une bonne fois pour toutes. Votre profil est-il déviant ? Vous pourrez répondre à cette question avec certitude seulement le jour où vous serez attaqué par un drone.

Alors là, vous nous direz : *« Est-ce qu'il n'y aurait pas comme un petit problème politique majeur ? »* Mais c'est fort possible. S'il n'y a pas de règle établie, il n'est pas seulement difficile de savoir quand on les enfreint, il est surtout impossible de les discuter. Le plan de Big Data n'est pas de surveiller le monde, c'est d'être le monde et de le dessiner.

Sur cette question, Antoinette Rouvroy ne nous rassure pas : *« Ce n'est pas parce que vous êtes noir, supporter de foot, femme ou que vous portez un turban que vous allez être arrêté à la douane : c'est parce que vous avez été classé par un algorithme aveugle et vous ne savez pas pourquoi. Donc, vous ne pouvez pas vous opposer non plus. Vous ne connaissez pas les autres qui sont profilés avec vous. Ce sont des gens qui ne vous ressemblent pas. Ici, il n'y a plus de discours possible. On est à la fois dans l'infra-individuel, sur base de données numérisées qui sont des fragments partiels de vos existences quotidiennes – parce que ce ne sont pas toutes vos données qui vont servir pour chaque profil, mais seulement quelques données – et c'est supra-individuel, parce que c'est une recombinaison qui concerne beaucoup d'individus. »* On ne sait même plus avec qui être solidaire. Bonne chance pour construire un mouvement social !

Le sens contre les traces, « don't feed the troll »

« On est foutus ! » aurait-on tendance à se dire. Mais il y a tout de même des moyens de

contourner Big Data. Il existe un réseau alternatif, comme TOR, ou des systèmes de brouillage. *« Mais en même temps, si vous allez sur TOR, vous allez être profilé comme étant potentiellement un dangereux terroriste ou quelqu'un qui fait de la pédopornographie... À partir du moment où vous y entrez, on ne sait plus ce que vous y faites. C'est pour ça que les hackers demandent que tout le monde utilise tout le temps TOR. Y compris pour écrire à votre grand-mère. Chaque fois que vous n'avez rien à cacher, utilisez TOR ! »* insiste Antoinette Rouvroy.

Et puis, nous devons « faire face ». *« Nous sommes des performances identitaires et donc faire face a un sens vraiment théâtral : trouver des lieux et des scènes, du temps interruptif, du vrai temps humain, sur lesquels rendre compte de nous-même, rendre compte de ce qui nous fait agir, et donc se subjectiver. Faire face, c'est aussi assumer les décisions. C'est ce dont nous dispense précisément la gouvernementalité algorithmique, qui nous permet de croire qu'on est en passe d'éradiquer la notion d'incertitude radicale. Or celle-ci donne toute sa dignité à la notion de décision. On ne décide réellement que quand on décide dans des situations d'incertitude et qu'on assume la possibilité de se tromper. Faire face, c'est ça aussi, c'est assumer la décision, y compris la décision politique, parce qu'on est dans l'incertitude et qu'on sait qu'on a de fortes chances de se tromper mais tant pis, on y va quand même. L'envoûtement, c'est le confort en fait, le confort de n'avoir rien à décider, parce que finalement les chiffres parlent d'eux-mêmes, c'est-à-dire qu'ils parlent de nous, qu'ils prétendent parler de nous, et que la réalité numérisée prétend se substituer à ce dont elle n'est en quelque sorte qu'une image complètement tronquée et réduite. »*

Alors, soyons récalcitrant ! Car *« nous sommes, dit-elle, des animaux autobiographiques. C'est nous qui rendons compte de ce que nous sommes. Nous ne sommes pas des ensembles de données ! »* Raconter sa vie, voilà une des plus sûres manières de ne pas laisser aux algorithmes le monopole de la production du réel.