*Cet entretien a été publié dans le numéro 17 d'*Article11 R

Je donne, tu donnes, il/elle donne… nos données. Les data sont le nouvel or noir que se disputent gouvernements et géants du numérique. Parfois, les premiers se servent même tout simplement chez les seconds, comme l'a montré l'existence du programme Prism, l'une des nombreuses révélations d'Edward Snowden - l'ancien sous-traitant de la toute-puissance *National Security Agency* (NSA) américaine. Par une tentative de retournement culotté de la situation, les artisans de la surveillance de masse et de la présomption de culpabilité tentent de se dédouaner en pointant la collecte tentaculaire opérée par Google, Facebook et consorts. Leur argument : *vous – population – filez bien plus à des entreprises privées* ! Lesquelles répondent que *pas du tout*… Un cercle sans fin.
Ces enfumages et faux débats empêchent de penser la transformation de nos vies en données, en signaux infrapersonnels qui ne font sens qu'agrégés et moulinés par millions. En naît un pouvoir d'un genre nouveau, la « gouvernementalité algorithmique », explique la chercheuse belge **Antoinette Rouvroy**1.

Au commencement, il y a les données. Qu'est-ce qu'une donnée ?

« On parle aujourd'hui de plus en plus des données brutes, par opposition aux données à caractère personnelle2. Dans le contexte des *big data*3 ou du *data mining*4, les données traitées sont elles-mêmes ''assignifiantes'' : elles n'identifient pas une personne.
De mon point de vue, elles constituent de nouveaux objets, dans la mesure où elles sont produites par des opérations de décontextualisation, de désindexation, de purification de toute signification personnelle. Elles ne sont pas signifiantes individuellement, mais sont agrégeables avec d'autres données recueillies dans des contextes temporels et géographiques différents. Elles permettent ainsi de créer des profils ou des modèles de comportement qui ne se rapportent à personne en particulier mais désignent des trajectoires possibles.
Les données ne fonctionnent pas comme des signes, les éléments traditionnels de la sémiotique. Pour qu'un signe soit signifiant, il faut qu'il représente la chose qu'il signifie. Par ressemblance (comme les icônes), par convention (comme les signaux routiers) ou par le souvenir d'un rapport physique avec la chose représentée (comme les traces ou empreintes).
Les données brutes ne fonctionnent pas ainsi parce qu'elles ne ressemblent à rien, ne font évidemment pas signe par convention (personne ne décide a priori de la signification à leur accorder) et ne conservent aucun lien matériel avec un geste physique (le doigt qui frappe le clavier, par exemple). Ce sont donc bien de nouveaux objets, qui se rapprochent plus du signal que du signe. Je m'en rapporte ici à la définition d'Umberto Eco : un signal est ce qui est calculable, malgré l'absence de signification. »

C'est donc un potentiel ?

« Comme ces données ne sont pas signifiantes, qu'elles ne conservent aucun lien direct avec les individus, qu'elles sont souvent anonymisées, décontextualisées, elles paraissent triviales, peu menaçantes, y compris pour la vie privée des individus. Mais elles vont servir à produire une forme de modélisation du social et du comportement qui aura des effets très concrets de gouvernement sur les conduites possibles. »

Lors d'une intervention récente, vous avez défini la donnée brute comme « *texture même du capitalisme* »…

« Dans *Mille Plateaux, Capitalisme et Schizophrénie*, Gilles Deleuze et Félix Guattari définissent le capitalisme comme la libération des flux dans des champs déterritorialisés. Cette libération nécessite des opérations de dé-codage. Les signes qui ont de la signification pour nous en ont grâce aux

codages propres à la société dans laquelle nous vivons. Les codes d'intelligibilité nous permettent de les comprendre.
Ici, il y a un dé-codage. C'est-à-dire une sorte d'abstraction sur fond de complète décontextualisation qui supprime des possibilités d'intelligibilité dans l'esprit humain. Le capitalisme fonctionne de cette manière, en extrayant notamment de leur contexte de vie les moyens de production. »

En quoi consiste le phénomène de *big data* ? C'est une simple juxtaposition de données brutes ou une agrégation intelligente (grâce aux algorithmes) des données brutes ?
« Il s'agit d'abord d'un phénomène de perte de contrôle. Il y a *big data* lorsque des seuils de quantité, de complexité (les données sont recueillies dans des contextes très différents) et de rapidité (elles prolifèrent en temps réel) sont atteints, de telle sorte que ces masses de données ne peuvent plus être gérées par les techniques traditionnelles de banques de données. L'existence des *big data* constitue donc une sorte d'aveu de faiblesse, lequel renvoie en creux à notre incapacité à administrer ces masses de données autrement qu'en sous-traitant leur gestion à des algorithmes, dont on connaît de moins en moins le fonctionnement.
Les *big data* offrent de nouvelles manières de rendre le monde signifiant – Michel Foucault parlerait d'un nouveau "*régime de vérité*" – sur la base de la pure induction (et non de la causalité), en temps réel, dans des modèles qui s'affinent en permanence… C'est vraiment un changement de paradigme dans le mode de production de ce qu'on considère comme le réel. Les possibilités ouvertes sont gigantesques dans tous les domaines : le marketing, la connaissance des consommateurs, l'astronomie, l'histoire, l'archivage…
Dans le même temps, les *big data* transforment de nombreuses disciplines, pas toujours pour le pire. Elles permettent de prendre en compte des masses de données qu'on devait ignorer auparavant. Donnent l'impression d'avoir un accès plus exhaustif à la réalité. Satisfont le mythe d'un accès immédiat et total à celle-ci sans passer par quelque représentation que ce soit. Un rêve d'immanence qu'on a l'impression de voir se réaliser. »

L'utilisation des *big data* n'est-elle pas le prolongement technique et logique de la statistique ?
« Les finalités de la statistique traditionnelles sont très différentes des opérations statistiques du *data mining* ou des *big data*. La statistique consiste à objectiver des hypothèses faites a priori : elle est un moyen de preuve. Le *data mining* au contraire fait surgir des hypothèses du réel numérisé lui-même !
Les techniques aussi sont différentes. La statistique ne tient pas compte des points les plus éloignés pour calculer une moyenne, car ils sont perturbateurs, générateurs de bruit. Les capacités de calcul sont limitées. Avec le *data mining*, il est possible, grâce aux puissances de calcul, de tenir compte de tout, y compris des points les plus éloignés. La notion même de normal, de normalité, de moyenne disparaît. »

Ce qui donne l'illusion d'une retranscription totale du réel, voire d'un dédoublement du réel ?
« Le réel numérisé peut en effet sembler être le réel lui-même. Comme il n'y a plus de sélection de données par des acteurs humains, donc faillibles, on a l'impression d'une très grand objectivité. La totalité des données étant actualisée en temps réel, le monde numérisé paraît reproduction exacte

du monde. La notion de médiation, de représentation en fait les frais, comme pour consacrer le mythe de l'accès immédiat au réel tel qu'il est. »

Il y a un risque de confondre le réel avec le réel numérisé ?

« Oui, parce que ce processus présente des aspects extrêmement performatifs. Ce qui n'est pas numérisé ou numérisable, que ce soit la détresse humaine ou des circonstances particulières, ce qui ne rentre pas dans les cases, n'a plus voix au chapitre. Cette évolution, qui paraît très objective, est en fait profondément injuste : on sait très bien que la totalité du réel n'est pas mise en nombres.

Cette impression de repli de la réalité à l'intérieur de la réalité numérisée, cette totalisation – c'est presque un régime totalitaire numérique –, risque de faire disparaître la distinction entre le monde et sa représentation, et donc la possibilité même de la critique. On peut critiquer la représentation, comme on critiquait les anciens objets statistiques, qui n'étaient pas assez représentatifs ou mal construits. Tandis que le réel numérisé va à rebours de l'idée que le savoir est toujours construit, et donc critiquable en fonction des conditions de sa construction. Si ce n'est plus construit, ce n'est plus critiquable. C'est un enjeu fondamental. »

Comment critiquer efficacement les effets de pouvoir engendrés ?

« La critique est d'autant plus difficile que le pouvoir n'apparaît pas comme tyrannique. Pour fonctionner, ce mode d'exercice du pouvoir a besoin que nous nous sentions le plus libre possible de nous exprimer, que nous donnions un maximum de données. Il ne nous contraint pas dans notre expression. Au contraire ! Dès que vous êtes connecté sur Facebook, un champ vous propose de vous exprimer ! C'est exactement l'inverse du monde orwellien dans lequel la liberté d'expression était l'ennemi du pouvoir. Nous sommes en quelque sorte gouvernés par notre liberté, ce qui rend la critique extrêmement difficile. »

Et nous sommes consentants...

« Il faudrait plutôt parler d'un effet d'adhésion. Nos profils sont si multiples, adaptables et adaptés en temps réel que ne pas vouloir être profilé, c'est ne pas se vouloir soi-même. Dans ce mode de gouvernement, l'autorité n'est plus assumée par personne, sinon par nous-mêmes, par les trajectoires qui nous sont assignées à titre d'un destin identitaire conduisant à une clôture de l'individu sur lui-même. Ce qui correspond, dans un certain sens, à l'idéal des années 1960 - ne pas être gouverné par autre chose que nous-mêmes. »

La multiplicité presque infini des profils créés permet de résoudre ce que vous appelez « *l'oxymore de la personnalisation industrielle* »...

« C'est paradoxal : la personnalisation très fine ne s'accompagne d'aucune rencontre entre des individus qui n'ont plus à exprimer leur propre désir. Le président de Google a ainsi expliqué qu'il deviendra bientôt très difficile pour les individus de vouloir quelque chose qui n'a pas été prévu pour eux par sa firme. Dans ce cadre disparaissent l'expression, le façonnement de nos propres désirs, ainsi que notre manière de les exprimer à autrui.

Nous nous exprimons ou rencontrons quelqu'un parce que nous ressentons une sorte de manque, d'insatisfaction. Si tout est prévu par une providence numérique, nous n'avons plus besoin de cet

espace public pour exprimer nos envies, nos désirs. La sphère individuelle est gavée, volontairement ou non, de manière à éviter le commun. »

Personne ne détient ce pouvoir ? Même pas les créateurs d'algorithme ? Ou ceux qui les possèdent, comme Google ?
« C'est très dilué. Ceux qui façonnent les algorithmes ne sont pas détenteurs du pouvoir, car ils n'exploitent pas les données eux-mêmes. D'autant plus qu'ils sous-traitent leur boulot aux algorithmes (le "*machine learning*"), lesquels sont auto-apprenants.
Par contre, ceux qui sont autorisés (ou qui s'autorisent) à exploiter ces données à des fins de prédation détiennent un pouvoir important. Ils visent par exemple le temps d'attention des utilisateurs d'Internet et ciblent leur espace de vie avec la publicité individualisée. Grâce à l'Internet des objets, des publicités pourront s'afficher dans nos espaces privés, sur nos frigos ou sur d'autres objets domestiques. Les industriels et les commerciaux tirent ainsi les marrons du feu. De même que les États surveillants, comme les États-Unis avec la NSA. »

La cible de ce nouveau pouvoir paraît davantage être le futur que le présent...
« Il se concentre sur l'incertitude radicale liée au fait que nous sommes des individus qui nous pensons libres, avec une possibilité de ne pas forcément faire tout ce dont nous sommes capables. Alors que les algorithmes appliqués à la sécurité considèrent que l'individu susceptible de commettre un acte terroriste est déjà un terroriste, comme s'il avait commis l'acte dont il est physiquement capable.

Spinoza expliquait que notre puissance vient de notre capacité à ne pas faire tout ce dont nous sommes capables. Cette possibilité de retenue est essentielle. Même si nous ne sommes pas des êtres absolument libres, rationnels, autonomes, il existe une sorte de marge. Et c'est précisément cette dernière qui est la cible de la gouvernementalité algorithmique. Il s'agit en somme de l'excès du possible sur le probable. »

En quoi ce gouvernement du futur se distingue-t-il de la prévention, qui consiste à empêcher la réalisation d'actes jugés illégaux ou immoraux ?
« La prévention passe par l'identification des causes, puis par une intervention pour que le comportement ne se produise pas. Ici, il n'est plus du tout question de prévention : les causes sont devenues totalement indifférentes. La production de ce mode de "savoir" est purement inductive. Je parlerais de préemption : il faut empêcher (ou faire advenir) certains comportements, sans se soucier des causes. En laissant de côté ces dernières, on se prive de la possibilité de prévenir. »

Est-ce le prolongement de ce que Foucault décrit à propos de l'intériorisation de la contrainte ?
« Non, c'est complètement nouveau ! Dans les dispositions algorithmiques, le sujet n'existe pas. L'effet de ces dispositions est précisément de rendre impossible les opérations de subjectivation. La norme n'est plus connue par personne, elle n'est plus visible, elle évolue en temps réel. Elle est complètement plastique, s'adapte à tout ce qui se passe. Ce système ne connaît que les données infrapersonnelles, les données brutes donc, agrégées au niveau supra-individuel.

L'objectif est très différent de la discipline foucaldienne, qui visait la réforme des comportements, soit une sorte de moralisation. Ici, le caractère moral est complètement absent. Cela fait apparaître

la gouvernementalité algorithmique comme neutre, objective et non-gouvernante. La cible n'est pas la réforme des individus, mais celle des comportements - peu importe par qui ils sont portés.

Lire Foucault permet surtout de comprendre ce qui a glissé depuis ses analyses. Au lieu de la production de corps dociles par les normes, on en vient à la production de normes dociles par les corps. La normativité s'adapte à tous les mouvements. »

Dans quelle mesure la gouvernementalité algorithmique comporte-t-elle des failles ? Vous disiez un peu plus tôt que refuser d'être profilé, c'est se refuser soi-même...
« Tout événement est disséqué, éparpillé en une multitude de données elles-mêmes réagrégées. Mais il en reste. L'affaire Prism a par exemple fait bouger certaines lignes. Pas assez, évidemment, mais elle a quand même constitué une sorte d'électrochoc.

Alain Badiou explique que si on reconnaît un fait comme événement, si on décide que c'est un point de rupture, un point de refus, on peut en tirer les conséquences. Ce qui suppose une décision à l'origine et se traduit ensuite par des décisions concrètes. À cet égard, l'affaire Prism est un non-événement, malgré les possibilités portées par la situation. Elle a certes permis une certaine prise de conscience de ce dans quoi on s'embarque. Mais elle n'a pas débouché sur une forme de "débarquement". »

En quoi le mouvement d'ouverture des données publiques (*open data*) peut-il constituer un contre-pouvoir ?
« C'est tout à fait insuffisant. Les individus n'ont aucune possibilité de traiter ces données d'une manière qui fasse sens pour eux. Ce mouvement ne change rien aux déséquilibres de pouvoir et de moyens. C'est davantage un cache-sexe, qui rajoute une couche au discours dominant faisant l'apologie des données et expliquant que l'avenir est dans leur traitement. »

Les données protégées par les institutions de type CNIL et les gouvernements sont celles à caractère personnelle. Mais si on prend l'exemple de la la NSA, on se rend compte qu'elle se nourrit de métadonnées5...
« Le droit est complètement à côté de la plaque ! Il faut bien sûr continuer à protéger les données à caractère personnel. Mais on rate un large pan de la problématique actuelle si on en reste là. C'est-à-dire qu'on fait alors l'impasse sur tout ce qui a trait au profilage, à la personnalisation, à l'hypertrophie de la sphère privée, à la paupérisation de l'espace public et à la prédation par des sociétés privées des espaces eux-mêmes privés des internautes.

Cet acharnement à camper sur la protection des données à caractère personnel est curieux, voire suspect. Une cécité pareille ne peut être que volontaire. L'objectif serait-il raté à dessein ?

Quoi qu'il en soit, ce processus est éminemment révélateur. Nous sommes tellement imprégnés d'une culture individualiste que nous sommes incapables de voir ce qui se joue au-delà de la question des données à caractère personnel. Cet aveuglement va de pair avec un mépris pour le commun, pour l'espace public et pour le débat public. »

1 Qui oeuvre au sein du Centre de recherche en information, droit et société de l'Université de Namur.

2 Une donnée personnelle comporte des informations renvoyant à l'identité d'une personne, que ce soit directement (son nom) ou indirectement (son numéro de téléphone, son adresse IP, son empreinte ADN, etc.).

3 Le terme big data désigne des ensembles de données très volumineux.

4 Le *data mining* consiste à passer à la moulinette des données, souvent en très grand nombre, pour en extraire du sens. Les moulinettes sont aussi appelées algorithmes.

5 Les métadonnées sont l'ensemble des informations qui entourent les données elles-mêmes. Par exemple, pour un e-mail, il s'agit des adresses IP de l'émetteur et du récepteur, de leurs adresses tout court, de l'objet du message et de sa date d'envoi.