Pre-print du chapitre publié dans D. Bourcier et P. De Filippi (dir.),
Open Data & Big Data. Nouveaux défis pour la vie privée, Paris, Mare & Martin, 2016, p. 245-262

Vers une économie politique des données : le pouvoir à l'aune des *data*

Benjamin Loveluck
Télécom ParisTech et CERSA (CNRS/Paris 2)

L'extension de l'environnement numérique a entraîné une inflation de *traces numériques* produites par les individus autant que par les institutions – et, de plus en plus, par les objets eux-mêmes à travers la prolifération de capteurs, de compteurs et de dispositifs « intelligents ». L'afflux de données quantifiées atteint aujourd'hui les proportions d'un *data deluge* qu'il est devenu trivial de signaler (*The Economist* 2010). En tant qu'outils d'observation scientifique et de mise au jour de phénomènes inédits, mais aussi comme instruments d'étude de marché et d'intelligence économique, d'analyse et de gestion des risques, de prospective et d'aide à la décision etc., la collecte et l'analyse de ces données a acquis une place prépondérante dans tous les domaines de la vie sociale, économique et politique.

On parle aujourd'hui plus volontiers de « données » que d'« informations », mais cette évolution terminologique traduit avant tout le stade le plus récent d'un processus initié au début du XX^{e} siècle lors du basculement vers une conception mathématique et scientifique de la notion d'information (Aspray, 1985 ; Segal, 2003). Alors qu'elle désignait jusqu'à lors avant tout un fait, une occurrence ou une « nouvelle » susceptible d'être communiquée, l'information devient un matériau universel capable d'être encodé et ainsi quantifié, traité et mis en circulation à l'aide des ressources de la logique mathématique et d'instruments techniques. Derrière cette évolution, l'on peut également identifier un idéal de libre circulation qui a animé toute l'histoire des réseaux informatiques, si bien que l'on peut parler à cet égard d'un « libéralisme informationnel » mu par l'idée que la numérisation et la mise en flux de l'information pourrait, en soi, servir de fondement au « pouvoir d'agir » (*empowerment*) des individus d'un côté et à l'auto-régulation socio-politique de l'autre (Loveluck, 2015). Ce faisant, de nouveaux rapports de force entre l'individu, la société civile et la puissance publique sont apparus, dont l'*économie politique des données* – leurs modes de production, d'appropriation, de partage – se présente comme un révélateur. En particulier, l'articulation entre la transparence et l'ouverture d'un côté et la protection des libertés individuelles – notamment la vie privée – de l'autre, qui est au cœur du libéralisme classique, se trouve posée à nouveaux frais.

Différents ordres de données peuvent être identifiés, que nous examinerons successivement dans ce chapitre. Il peut s'agir 1) de données « ouvertes », en particulier celles qui ont été produites ou collectées par des administrations publiques et qui sont mises à disposition du public (*open data*) ; il peut également être question 2) de données « massives », produites ou collectées avant tout par des entités privées, en particulier à travers les traces des actions et des interactions réalisées au sein d'un environnement de plus en plus susceptible d'en conserver des enregistrements (*big data*) ; mais il peut également s'agir 3) de données « fuitées », c'est-à-dire de données personnelles ou institutionnelles à caractère confidentiel qui ont été divulguées de manière involontaire ou illicite (*leaked data*). L'articulation entre ces trois ordres des données pose un certain nombre de questions nouvelles pour des catégories structurantes du politique, et sur

les possibilités émancipatoires autant que les modalités du contrôle. Nous voudrions ici explorer quelques-unes de leurs implications, en revenant sur ce que ces trois ordres des données révèlent des dispositifs de savoir/pouvoir à l'œuvre – une question qui peut se résumer de la manière suivante : en matière de données, *qui accède à quoi et avec quels effets ?*

1. Ce que les pouvoirs publics dissimulent, ce qu'il doivent montrer : les dimensions de l'*open data*

Du point de vue de l'État ou du secteur public et dans le contexte français, comme l'a par exemple souligné Jacques Chevallier dans ce volume, la notion de « donnée » contribue tout d'abord à bouleverser les termes du compromis juridique sur le principe de transparence trouvé en 1978 (loi « CADA », qui constitue le pendant du *Freedom of Information Act* de 1966 aux États-Unis), et qui renvoie aux notions de « document administratif » et de « donnée publique ». Ce compromis est lui-même le résultat d'un processus historique par lequel la puissance publique, sous l'effet du déploiement du libéralisme politique à l'époque moderne, a été soumise à un principe de responsabilité devant les citoyens (*accountability*), qui implique un droit de regard sur ses activités ainsi que sur les données qu'il collecte à ses propres fins, et qui est un élément constitutif des « démocraties de surveillance » modernes (Rosanvallon, 2006). De manière plus générale, on peut en effet associer ce mouvement à l'impératif de *publicité* qui est l'un des axiomes centraux des démocraties représentatives, et qui entre en tension avec la part de *secret* que l'administration de l'État entend préserver – l'idée de soumettre les activités politiques à l'appréciation publique étant structurellement liée à l'apparition de la notion de « raison d'État » (Gauchet, 2005). En outre, les arcanes du pouvoir (*arcana imperii*) se justifient par la nécessité de stabiliser l'État : en théorie, si elles constituent un lieu de suspension de la loi qui ouvre à toutes sortes de dérives potentielles (arbitraire, corruption, violence), elles permettent également le maintien et l'effectivité de la loi, des libertés individuelles et de la sphère publique (Horn, 2011), ce qui illustre les rapports éminemment complexes qu'entretiennent la transparence et le secret et le caractère historiquement changeant d'une telle « économie du visible » (Senellart, 1995).

Si tout ne peut pas être divulgué au citoyen, du moins sait-il que certaines choses lui sont cachées, et il pourra chercher à se tenir aussi informé que possible des activités de la puissance publique, afin éventuellement de pouvoir les contester dans les arènes juridique, politique voire médiatique. L'*open data* va cependant plus loin, d'abord parce qu'il est ancré dans une philosophie plus large de partage des connaissances, et ensuite parce qu'il s'agit de mettre en œuvre une mise à disposition *pro-active* et *suivie* de l'information, dans des formats techniques et sous des licences juridiques « ouverts » qui permettent leur *réutilisation.* La perspective est d'encourager la participation des citoyens ainsi que leur « pouvoir d'agir », en favorisant aussi bien les initiatives économiques que les actions politiques (notamment la production de contre-expertises) au sein de la société civile. L'*open data* se présente ainsi parfois comme allant « au-delà de la transparence » (Goldstein & Dyson, 2013). S'agissant de l'État et des autorités locales, le bilan demeure cependant assez limité jusqu'à présent, soit parce que les pratiques administratives elles-mêmes n'ont pas été modifiées en profondeur, soit parce que les données publiées s'avèrent finalement assez inoffensives, bien qu'elles puissent apporter des éléments contextuels importants (Goëta & Mabi, 2014). Par ailleurs, l'ambition dans certains cas de pouvoir convertir les données ainsi « libérées » en problèmes publics voire en « affaires » ne va pas du tout de soi, d'abord parce qu'elle néglige la nécessité de mettre en récit ces données et de les faire surnager dans une « économie de l'attention » de plus en plus saturée, et ensuite parce qu'il est parfois très délicat de s'appuyer sur des éléments purement quantitatifs pour jauger de certaines réalités[1].

[1] Par exemple, certains sites proposent de mesurer l'activité parlementaire à partir de données extraites automatiquement des sites des différentes assemblées. C'est le cas de NosDeputes.fr ou de NosSenateurs.fr pour

En outre, d'autres objectifs peut-être plus implicites peuvent également être associés à l'*open data*, qui mettent en lumière une certaine ambivalence dans les effets qui en sont attendus. Parmi les arguments avancés pour justifier le développement et l'extension de la « libération des données », se trouve d'abord l'idée que cela pourra générer un surcroît d'activité économique, dans la mesure où l'*open data* constituerait une « information liquide au service de l'innovation et de la performance » (Manyika *et al.*, 2013). Le risque ici est de voir les données ouvertes être avant tout réutilisées par ceux qui disposent des ressources financières et des compétences nécessaires à leur valorisation (Gurstein, 2011). Mais surtout, d'un point de vue politique, l'*open data* se trouve également justifié par une volonté non seulement de contrôler si les activités et les actions de la puissance publique sont *justes*, mais également de pouvoir évaluer si elles sont *efficaces* – par exemple, les nids de poule dans la voirie ont-ils bien tous été rebouchés (FixMyStreet.com) ? L'*open data* peut en effet se présenter comme une mesure de performance qui s'inscrit aisément dans l'optique du *new public management*, d'un meilleur contrôle et d'une mise en concurrence des services publics, et d'une ouverture qui peut se traduire dans les faits par des privatisations de biens publics et par l'extension du marché (Bates, 2014). Ainsi le « gouvernement ouvert » (*open government*) associé à l'*open data* peut-il faire figure d'« État-plateforme » entendu comme un « gestionnaire de place de marché », soit une nouvelle déclinaison du *small government* cher aux libéraux conservateurs (O'Reilly, 2010), qui va de pair avec une extension de la bureaucratie.

Par ailleurs, et au-delà du compromis – historiquement changeant – autour de la transparence, il est bien connu que la nature politique de l'information est directement liée à la naissance de la statistique en tant que « science de l'État », organisant la collecte de données sur sa population afin d'orienter son action (Desrosières, 2010). En mesurant ainsi différents phénomènes – allant de la démographie au chômage en passant par la croissance et la criminalité – il s'agit de représenter la réalité, afin éventuellement d'engager des actions visant à affecter cette réalité. À ce titre les statistiques construisent des objets sociaux qui peuvent également être *débattus*, aussi bien sur le plan des objets eux-mêmes que sur celui des conventions qui président à leur construction. Ces débats mais aussi ces hésitations et ces conflits d'interprétation ont traversé toute l'histoire des statistiques, et se perpétuent aujourd'hui – par exemple sous une forme très explicite à travers un « statactivisme » appelant à « lutter avec les nombres » en utilisant autrement les chiffres fournis par l'État, en choisissant de mesurer autre chose que ce qui est présenté, ou en le mesurant différemment (Bruno, Didier & Prévieux, 2014).

Cependant, ce mouvement témoigne également d'un déplacement de la rationalité qui repose sur l'objectivité supposée des faits ainsi quantifiés, et qui puise ses racines à des sources pré-modernes (Berns, 2009). De ce point de vue en effet, l'une des propensions du pouvoir est de tendre à éviter l'édiction de normes explicites qui apparaîtraient comme l'émanation d'une entité surplombante et coercitive, ce qui permet au gouvernement de se dédouaner de toute responsabilité dans l'exercice de la norme pour la déplacer sur les sujets eux-mêmes et leurs comportements, et de cette manière de « gouverner sans gouverner » : « Le modèle n'est plus celui de commandements édictés dans des termes généraux et définitifs par des autorités considérées comme souveraines et douées de contrainte, mais celui d'une multiplicité de dispositifs de contrôle, de classement et d'évaluation, souvent quantitatifs, qui émergent de la

---

les députés et sénateurs français. C'est aussi le cas au niveau européen, mais il existe deux sites – VoteWatch.eu et MEPranking.eu – qui, à partir des mêmes données, proposent chacun des indicateurs différents, et leurs classements des « députés les plus actifs » ne se recoupent pas. En s'appuyant sur VoteWatch, la rubrique « Les Décodeurs » du journal *Le Monde.fr* avait épinglé l'eurodéputé Jean-Luc Mélenchon pour sa faible assiduité, et celui-ci s'était défendu en ayant recours à MEPranking qui, lui, tient compte des « explications de vote » – alors que celles-ci peuvent n'être que des courriers électroniques justifiant une prise de position (*LeMonde.fr*, 16 avril 2014 [http://www.lemonde.fr/les-decodeurs/article/2014/04/16/les-astuces-de-m-melenchon-pour-paraitre-assidu-au-parlement-europeen_4402075_4355770.html, consulté le 12 janvier 2016]).

réalité même qu'il s'agit de réguler, qui sont spécifiques à un champ d'activité (…) et qui accompagnent de manière constante et adaptable les acteurs concernés, sans même réclamer l'appui d'une sanction activée de l'extérieur » (Berns, 2009, p. 7-8).

2. Ce que les entreprises et les institutions peuvent voir, comment elles orientent le regard : *big data*, protection de la vie privée et politique des algorithmes

Parallèlement à ces évolutions, la production et la collecte de données par des acteurs économiques connaissent aussi une expansion vertigineuse. L'idée de *big data* désigne avant tout des ensembles de données – ouverts ou non – qui sont non seulement de plus en plus *exhaustifs*, mais qui sont également de plus en plus captés *en temps réel* et qui sont de plus en plus *relationnels*. En vertu notamment du rôle médiateur des technologies numériques, une part croissante des traces de l'activité quotidienne des individus – qui va de l'utilisation d'une carte de paiement, aux trajets effectués à l'aide d'un GPS, ou aux communications échangées par téléphone ou sur les médias sociaux – peut désormais être captée, et ce largement à l'insu des utilisateurs qui ont rarement conscience de l'étendue du sillage numérique qu'ils laissent derrière eux. Le plus souvent, l'anonymat ou le pseudonymat ne fournit qu'une protection toute relative, et de nombreuses données et métadonnées peuvent permettre de reconstituer leur parcours. Or ils n'ont généralement pas donné leur accord explicite pour que ces informations soient conservées – voire exploitées à des fins qui demeurent le plus souvent opaques. Centralisées par les fournisseurs de services dans de vastes « fermes de données » (les serveurs sur lesquels elles sont stockées), ces informations sont, pour filer la métaphore, nourries, entretenues, croisées entre elles, en somme *cultivées* afin d'en moissonner les fruits.

Dans la mesure où elles permettent de cerner plus précisément les comportements individuels et collectifs, ces données s'intègrent d'abord dans un vaste dispositif de *surveillance*. Celui-ci peut être mobilisé à des fins commerciales privées, pour les entreprises cherchant à mieux connaître les préférences des consommateurs afin d'adapter leur offre, proposer des services plus pertinents, fournir de la publicité ciblée etc. Ces techniques, issues du marketing, sont également mobilisées de manière croissante par les partis politiques et leurs responsables, qui disposent parfois de connaissances très précises de leur électorat pour mener des campagnes *data-driven* (Howard, 2006 ; Issenberg, 2012). Enfin ces masses de données constituent également une mine potentielle pour les institutions et les pouvoirs publics, qu'il s'agisse de maintien de l'ordre ou de politiques publiques de santé, ou encore par les agences de renseignement dont on sait désormais qu'elles les exploitent massivement en s'appuyant sur des arguments sécuritaires et notamment la lutte contre le terrorisme. Mais nous le verrons, ces données revêtent également des dimensions politiques parce qu'*en elles-mêmes* et à travers la mise en œuvre de règles automatisées et adaptatives d'agrégation, de tri et de sélection qui leur sont appliquées (les *algorithmes*), elles participent à *orienter* la perception et les actions et ont à ce titre des effets normatifs et des incidences sur l'autonomie des individus.

Les dangers inhérents au déploiement du *big data* et des traitements auxquels ces données sont susceptibles d'être soumises ont fait l'objet de nombreuses réflexions, en particulier s'agissant des menaces en termes de vie privée (*privacy*) et des responsabilités juridiques et morales qui incombent aux acteurs impliqués (voir par exemple la contribution de Primavera De Filippi à cet ouvrage). De ces réflexions ont découlé dans certains cas des régulations et « bonnes pratiques » destinées à encadrer la production, la conservation, le partage et l'utilisation de ces données, mais qui sont confrontées à un territoire mouvant et à des déplacements dans les catégories d'appréhension de cette réalité. On annonce ainsi depuis longtemps déjà la « mort de la vie privée » (Froomkin, 2000 ; Garfinkel, 2000), à laquelle les entrepreneurs du numérique, qui sont aussi des « entrepreneurs de morale » œuvrant dans l'intérêt de leur *business*, ont largement

contribué[2]. Dans les faits cependant, et s'agissant par exemple d'un média social tel que Facebook, la vie privée fait l'objet d'intenses négociations collectives entre la plateforme et ses utilisateurs voire d'une « guerre culturelle » pour déterminer les contours des pratiques acceptables en la matière (par exemple s'agissant des paramètres par défaut), tenant compte des formes contemporaines de sociabilité qui poussent les utilisateurs à dévoiler des informations personnelles en ligne (Casilli, 2013 ; Casilli, 2014). Ces constats font écho à l'approche en termes d'« intégrité contextuelle » de la vie privée (Nissenbaum, 2010), selon laquelle la *privacy* ne peut être définie dans l'absolu et par opposition au « public » en général, mais que le degré d'ouverture ou de fermeture associé à une information doit systématiquement être fonction du contexte dans lequel l'information a initialement été produite.

Implicitement cependant, c'est donc toujours la notion de *consentement éclairé* de l'individu considéré comme sujet autonome qui semble guider la réflexion. La tradition philosophique moderne est en effet traversée par la notion de sujet moral et souverain, dont Kant a fourni l'édifice le plus abouti (Schneewind, 2001), et à laquelle l'idée de consentement est étroitement associée, ainsi que la nécessité de préserver certains droits politiques limitant les incursions de la puissance publique. D'un point de vue plus strictement politique en effet, dans une société démocratique la protection de la vie *privée* (qui peut être entendue comme la protection de la vie intime et familiale, du domicile et de la correspondance) est une condition essentielle à l'exercice de la vie *publique* (liberté d'expression et d'association), bien qu'elle rentre parfois en conflit avec ces droits fondamentaux. Or comme l'a classiquement montré Hannah Arendt, à l'époque moderne se déploie un processus où la différence entre les domaines du public et du privé tend à s'effacer, tandis que tous deux se trouvent « résorbés dans la sphère du social », et où le privé est remplacé par l'intime « de façon précaire » (Arendt, 1961, p. 59-121). Mais pour Arendt le privé, en tant que condition d'accession au public, est beaucoup plus que l'intime : l'homme est véritablement libre de se consacrer à des activités publiques lorsqu'il n'est pas soumis à la nécessité et lorsqu'il peut se retrancher hors du monde commun pour « vivre sans être vu, sans être entendu ». Défendre la démocratie implique donc de continuer à revendiquer une nette séparation entre ces deux sphères.

Par là également, Arendt réaffirme l'importance de la propriété (qu'elle distingue nettement de l'*accumulation*), étant entendu que « la seule manière efficace de garantir contre le grand jour de la publicité l'ombre des choses qui ont besoin du secret, c'est la propriété privée, un lieu que l'on possède pour s'y cacher » (p. 113). Or s'agissant de la circulation massive des données dans l'environnement numérique, c'est aussi en termes de *propriété* des données (et de notions dérivées telles que le consentement explicite, la portabilité, le droit à l'oubli, voire même le droit de rétribution) que les solutions ont souvent été formulées pour fournir une base à la protection de ces dernières – notamment dans le cas américain, où la *privacy* est très faiblement protégée face à une liberté d'expression toute puissante (1er Amendement). Cependant, il est en pratique très difficile d'identifier les cas, nombreux, où des informations à caractère personnel sont recombinées, réutilisées ailleurs, migrent vers d'autres bases de données ou sont revendues à des tiers (par exemple des annonceurs ou des assureurs), ce qui rend souvent caduque la possibilité de définir les contours d'une telle propriété et donc les possibilités de recours, même collectifs. En outre, la capacité toujours accrue de pouvoir *inférer* ou *déduire* de nouvelles informations (y compris concernant des individus tiers) à partir des données initialement collectées rend vaine

[2] Par exemple Mark Zuckerberg, fondateur et dirigeant de Facebook, déclarait en janvier 2010 : « Les gens sont de plus en plus à l'aise à l'idée de partager non seulement plus d'information et de différentes sortes, mais de manière plus ouverte et avec plus de personnes. C'est une norme sociale qui a évolué avec le temps. Notre rôle est d'innover constamment et de mettre à jour notre service afin qu'il reflète l'état présent des normes sociales. » En novembre 2013 Vint Cerf, co-inventeur du protocole TCP/IP et désormais *Chief Internet Evangelist* chez Google, disait lui aussi que la préservation de la vie privée pourrait n'avoir été qu'une parenthèse historique : « La vie privée (*privacy*) est peut-être une anomalie (…). Il sera de plus en plus difficile de protéger sa vie privée. »

toute tentative de cerner exactement ce dont les individus seraient propriétaires, et à quoi leur consentement éclairé pourrait s'appliquer.

À travers l'application de traitements mathématiques sous la forme d'algorithmes, des corrélations statistiques de plus en plus poussées peuvent être calculées. Il devient ainsi possible d'établir des *patterns* ou « signatures », et donc 1) de *classifier* en opérant des regroupements catégoriels sous forme de profils, 2) de *détecter* d'éventuelles singularités ou anomalies, et enfin 3) dans certains cas d'*anticiper* des comportements futurs en extrapolant à partir de ces informations. Sans même évoquer les données les plus sensibles (financières ou de santé), un exemple très simple est celui des capteurs dont sont équipés beaucoup de véhicules récents, et qui peuvent fournir des indications sur la conduite de leur propriétaire : un assureur hollandais, Fairzekering, propose ainsi d'attribuer des scores à ses clients et d'adapter leur police d'assurance, en l'augmentant si la conduite est considérée « à risque », ou au contraire en leur accordant des remises si elle est évaluée positivement (Meyers, 2015). Outre qu'il s'agit d'une intrusion très poussée dans la vie quotidienne des individus, de telles initiatives interrogent profondément les notions de solidarité et même de justice, en personnalisant ainsi un service au nom de l'équité (*fairness*) tout en faisant peser sur l'individu une responsabilité directe et cependant *médiée par la technique*. Ces déplacements nous incitent donc à penser la « politique des algorithmes » qui trouve ainsi à se manifester (Cardon, 2013), et qui va au-delà d'une simple exigence de « transparence » quant à leur fonctionnement précis, pour saisir de manière critique ce qui se joue lorsque des décisions de différentes natures sont déléguées aux machines. À ce titre, certains enjeux ont depuis longtemps été pointés, qui ont trait notamment à la segmentation des populations et au problème des « classifications sociales » discriminatoires construites à partir de données personnelles (Gandy, 1993 ; Lyon, 2003 ; Custers *et al.*, 2013), permettant dans un contexte commercial d'exclure certains types de consommateurs ou même de faire varier dynamiquement les prix en fonction du profil de l'acheteur, et dont les conséquences par exemple dans le domaine du maintien de l'ordre public sont potentiellement préoccupantes (Harcourt, 2007).

Plus généralement, ces mécanismes d'interprétation automatique des données fonctionnent de manière de plus en plus indépendante, mettant en place leurs propres critères de sélection et d'organisation de l'information. Une forme d'immanence qui fait dire à Antoinette Rouvroy et Thomas Berns : « Plutôt qu'une menace pour les droits individuels au respect de la vie privée et à la protection des données personnelles, l'enregistrement massif et systématique de données, le "data mining" et le profilage sont les instruments d'une transformation des rationalités, stratégies et tactiques de gouvernement » (Rouvroy & Berns, 2010, p. 88). Cette « gouvernementalité algorithmique » s'exerce sur les individus – qu'il s'agisse de citoyens, de consommateurs etc. – en s'appuyant sur les traces de leurs activités passées d'un côté, afin de saisir les caractéristiques des comportements et même des *intentions* qui peuvent leur être associés, et afin également de circonscrire le champ des actions qu'il leur sera possible d'effectuer dans le futur, c'est-à-dire de préempter ces comportements et ces intentions. En retour, les individus intègrent progressivement l'ubiquité de ces mécanismes et s'auto-disciplinent en fonction des effets attendus, et sont traités différemment en fonction de comportements qu'ils sont susceptibles (ou non) de manifester. Mais l'*immanence* apparente de ces dispositifs ainsi que leur *invisibilité* remet radicalement en cause la capacité, aussi bien individuelle que collective, à comprendre leurs finalités et à questionner leur légitimité (Rouvroy & Berns, 2010, p. 90).

Dans ce contexte, le sujet lui-même change de nature puisque « l'unité à laquelle s'adresse le pouvoir n'est plus l'individu unitaire, figure centrale du libéralisme, doué de capacités d'entendement et de volonté, identifié à un territoire corporel – cet individu-là n'intéresse plus (directement) le pouvoir » (Rouvroy & Berns, 2010, p. 94). La distinction entre la description de la réalité d'un côté, et l'édiction de la norme de l'autre, se brouille. Les normes morales ou les règles juridiques ne sont plus explicites ni assorties de sanctions, et ne sont plus dirigées vers un individu unitaire mais vers des *fragments* de celui-ci ou « dividuels » (Deleuze, 1990). Les capacités

de jugement et de réflexivité politique s'en trouvent directement affectés, puisqu'il devient difficile voire impossible de se reconnaître en tant que sujet de ces formes de pouvoir – et donc de les contester – ce qui rejoint les analyses évoquées précédemment dans le cadre des évolutions de l'État. Dans les deux cas, on le voit, les fondements sur lesquels ont été bâties les libertés civiles se trouvent profondément ébranlés.

3. Ce que la société civile dévoile de force : l'irruption des *leaked data*

Enfin, les *leaked data* correspondent aux données qui ont été « fuitées », c'est-à-dire qui n'ont pas été mises à disposition volontairement comme c'est le cas de l'*open data*, mais qui présentent une valeur économique ou politique et qui ont été arrachées de force à l'escarcelle privée ou publique où elles résidaient. Les données se présentent ici comme un *mode de révélation* qui prend la forme d'une dénonciation publique, et qui rejoint donc également les dimensions « contre-démocratiques » de la surveillance, de l'empêchement et du jugement (Rosanvallon, 2006). L'exemple paradigmatique est celui de WikiLeaks, qui a notamment mis à disposition successivement au cours de l'année 2010 des milliers de rapports militaires américains classés « secret » concernant la guerre en Irak et en Afghanistan (*War Logs*), puis des centaines de télégrammes diplomatiques (*Cablegate*). WikiLeaks, dont le slogan initial était « *We open governments* », se présente comme un acteur majeur d'une « économie de la fuite » (Bieber, 2013) c'est-à-dire d'un contre-pouvoir susceptible de compléter voire de remplacer les institutions traditionnellement en charge de dévoiler les « affaires » d'intérêt public, telles que le journalisme. L'organisation est donc considérée par certains comme un nœud essentiel d'un nouveau « quatrième pouvoir en réseau » (Benkler, 2013). Très explicitement, l'ambition de WikiLeaks est d'inverser le rapport de force entre les individus et les grandes organisations économiques et politiques (les États mais également certaines grandes compagnies privées), en se servant des fuites d'information comme d'une arme pour mettre au jour les abus et les injustices, en comptant avant tout sur le sens moral d'individus au sein de ces organisations (*insiders*), indignés par des pratiques contraires à l'éthique et qui joueraient ainsi le rôle de « lanceurs d'alerte » (Chateauraynaud & Torny, 1999).

À travers son médiatique chef de file Julian Assange, WikiLeaks prône une défense radicale de la vie privée assortie d'une non moins radicale exigence de transparence pour les grandes organisations – *privacy for the weak and transparency for the powerful* – qui implique notamment un usage bien compris de la cryptographie (Assange, 2012). Mais le cas de WikiLeaks est pour partie inédit, dans la mesure où la divulgation de données a été associée à une forte charge narrative et une large couverture médiatique. Or les éclairages de la sociologie soulignent, comme c'était déjà le cas pour l'*open data*, que les bases de données ne se convertissent pas aisément en controverses ou en « machines à scandales » (Parasie, 2013). Pour cela, il faut que se constitue *un public* d'une importance suffisante et assez indigné pour produire un jugement collectif, et par là des effets sur les normes collectives et les institutions ou organisations concernées. Lorsque c'est le cas cependant, l'une des particularités de ce mode de dénonciation est son caractère potentiellement plus démocratique, au sens où « au lieu qu'une personnalité publique prenne le public à témoin pour désigner des victimes et des coupables, la base [de données] permettrait à n'importe quel citoyen de construire localement sa propre indignation – suscitant éventuellement, de proche en proche, la constitution d'une indignation partagée par un public plus vaste » (Parasie, 2013, p. 131).

Dans ce contexte, on peut également citer un certain nombre d'actions entreprises sous le label des Anonymous, un collectif militant aux contours incertains, des « hacktivistes » qui ont pour caractéristique récurrente de monter des attaques ciblées dont les motivations sont diverses mais qui impliquent généralement la divulgation de données – parfois même la publicisation de données personnelles, une pratique connue sous le nom de *doxxing* (Coleman, 2013). Cependant,

les dérives sont nombreuses et les *leaked data* peuvent également être l'expression d'une forme de justice expéditive ou vigilantisme (Dennis, 2008). Elles peuvent aussi se réduire à des phénomènes de délation, de dénigrement ou de harcèlement dirigés contre des individus en particulier ou des groupes de personnes, et dont les femmes sont très souvent la cible (Citron, 2014), comme ce fut le cas au mois d'août 2014 quand des centaines de photos intimes de célébrités furent volées depuis les serveurs d'Apple (iCloud) et largement diffusées[3]. Elles peuvent enfin viser à déstabiliser de grandes entreprises : en novembre 2014 Sony Pictures Entertainment a vu de vastes quantités de données internes – informations personnelles des dirigeants et employés, emails, documents stratégiques, films pas encore sortis en salles – être mises à disposition du public par un groupe de *hackers* auto-désigné comme les « *Guardians of Peace* », dont les motivations sont toujours opaques[4]. En outre, les fuites sélectives d'information constituent une méthode bien connue permettant de conforter des intérêts particuliers à travers la manipulation des médias, et sont à ce titre sujettes à caution.

Qu'elles soient considérées comme servant l'intérêt général et le bien commun, ou qu'elles visent des objectifs moins reluisants, les fuites de données font cependant partie intégrante de l'environnement informationnel contemporain, et participent de son équilibre général – ou de ses déséquilibres. Cette part d'instabilité et d'insécurité est inhérente à la libre circulation de l'information et à la mise en flux des données, dont les « anomalies » et les « accidents » doivent également être pris en compte (Parikka & Sampson, 2009). Elles viennent parfois jouer un rôle bienvenu de perturbation des mécaniques trop bien huilées du libéralisme informationnel, du fantasme des « échanges sans frictions » et des aspirations à des formes de contrôle cybernétique qui peuvent lui être associés.

Conclusion

Ce tour d'horizon de trois grands ordres de l'économie politique des données – ouvertes, massives, fuitées – ne prétend aucunement à l'exhaustivité. Il permet cependant de rappeler comment les *data* sont mobilisées aussi bien pour gouverner les sociétés que pour créer de la valeur économique, et dans quelle mesure elles peuvent être tout à la fois facteur de liberté, de contrainte et d'instabilité, selon les modalités des agencements sociotechniques dans lesquels elles s'inscrivent. À un certain niveau, ce tableau très général est révélateur de l'extension de deux tendances conjointes des formes contemporaines du capitalisme libéral : la quête d'*efficacité* et de productivité d'un côté, la demande de *sécurité* et de minimisation des risques de l'autre.

Ces deux aspirations semblent converger dans un idéal profondément ambivalent de libre circulation de l'information animé par des agencements de boucles de rétroaction, visant au fond à se défaire du politique. En pratique, elles se résolvent en effet bien souvent en une même réduction de la démocratie à l'administration et à la gestion des comportements collectifs. C'est ainsi que l'action *politique* et *visible*, qui désigne en principe (et de manière schématique) dans les démocraties représentatives avant tout la prise de décision et la mise en place de règles de droit en conformité avec la volonté collective (idéalement le fruit d'une délibération), peut se muer en un ensemble d'actes *techniques* et *invisibles* qui seraient l'émanation du réel lui-même, et en cela difficilement contestables sur le plan des valeurs. Ce mouvement peut donc se définir comme une extension du domaine de la *technocratie*. Si le lieu du pouvoir demeure « vide », comme l'exige le libéralisme démocratique selon Claude Lefort (Lefort, 1986, p. 28), l'« institutionnalisation du conflit » qui est censé l'accompagner se trouve ici en quelque sorte désamorcé, dans la mesure où

[3] https://en.wikipedia.org/wiki/ICloud_celebrity-photo_leaks

[4] https://en.wikipedia.org/wiki/Sony_Pictures_Entertainment_hack

il ne subsiste plus de lieu « visible » où les tensions et les divergences d'intérêt inhérentes au social puissent chercher à se résoudre, à travers une renégociation permanente du pacte social.

La technocratie consiste en effet à apporter des solutions techniques à des problèmes d'ordre politique et social, et constitue un dévoiement de l'autonomie démocratique. Ce que Arendt appelait également « la victoire de la société aux temps modernes, substituant d'abord le comportement à l'action et éventuellement la bureaucratie, la régie anonyme, au gouvernement personnel » (Arendt, 1961, p. 84) se poursuit à travers l'extension des mesures de performance, et le déploiement de mécanismes de plus en plus automatisés d'auto-régulation immanents agissant en vertu de ces mesures. Ceux-ci défont pour partie l'édifice politique moderne visant à concilier l'autonomie individuelle et l'autonomie collective, qui était bâti sur le règne de la loi mais auquel une « gouvernance par les nombres » tend à être substitué, comme l'a montré Alain Supiot : « Le renversement du règne de la loi au profit de la gouvernance par les nombres s'inscrit dans l'histoire longue du rêve de l'harmonie par le calcul, dont le dernier avatar – la révolution numérique – domine l'imaginaire contemporain. Cet imaginaire cybernétique conduit à penser la normativité non plus en termes de législation mais en termes de programmation. On n'attend plus des hommes qu'ils agissent librement dans le cadre des bornes que la loi leur fixe, mais qu'ils réagissent en temps réel aux multiples signaux qui leur parviennent pour atteindre les objectifs qui leur sont assignés. » (Supiot, 2015, p. 23)

Des solutions d'ordre juridique (licences libres) ou technique (architectures distribuées, cryptographie) peuvent aider à prévenir une part des excès que nous avons pointés, et sous certaines conditions les données lorsqu'elles sont « ouvertes » voire « fuitées » peuvent même contribuer de manière bienvenue au bon déroulement de la vie démocratique. Cependant, et comme nous avons essayé de le montrer ici, la réduction croissante des clivages idéologiques, des divergences d'intérêt et des injustices sociales à des problèmes de traitement de données est une invitation pressante à penser de manière critique les déplacements souterrains affectant les catégories qui informent notre compréhension du politique.

Références bibliographiques

Arendt, H. (1961), *Condition de l'homme moderne,* Paris, Calmann-Lévy.

Aspray, W.F. (1985), « The scientific conceptualization of information: a survey », *IEEE Annals of the History of Computing* vol. 7 n° 2, p. 117-140.

Assange, J. (2012), *Cypherpunks. Freedom and the Future of the Internet,* New York, OR Books.

Bates, J. (2014), « The strategic importance of information policy for the contemporary neoliberal state: the case of Open Government Data in the United Kingdom », *Government Information Quarterly* vol. 31n° 3, p. 388-395.

Benkler, Y. (2013), « WikiLeaks and the networked fourth estate », in *Beyond WikiLeaks. Implications for the Future of Communications, Journalism and Society,* sous la direction de B. Brevini, A. Hintz, & P. McCurdy, Basingstoke and New York, Palgrave Macmillan, p. 11-34.

Berns, T. (2009), *Gouverner sans gouverner. Une archéologie politique de la statistique,* Paris, Presses universitaires de France.

Bieber, C. (2013), « Lessons of the leak. WikiLeaks, Julian Assange, and the changing landscape of media and politics », in *A Companion to New Media Dynamics,* sous la direction de J. Hartley, J. Burgess, & A. Bruns, Malden, MA, Oxford and Chichester, Wiley-Blackwell, p. 322-335.

Bruno, I., Didier, E., & Prévieux, J. (dir.), (2014), *Statactivisme. Comment lutter avec des nombres,* Paris, Zones.

Cardon, D. (2013), « Politique des algorithmes. Présentation », *Réseaux* n° 177, p. 9-21.

Casilli, A.A. (2013), « Contre l'hypothèse de la "fin de la vie privée". La négociation de la *privacy* dans les médias sociaux », *Revue française des sciences de l'information et de la communication* n°3.

Casilli, A.A. (2014), « Quatre thèses sur la surveillance numérique de masse et la négociation de la vie privée », in *Le numérique et les droits fondamentaux,* sous la direction de J. Richard & L. Cytermann, Paris, La Documentation française, p. 423-434.

Chateauraynaud, F. & Torny, D. (1999), *Les Sombres précurseurs. Une sociologie pragmatique de l'alerte et du risque,* Paris, Ed. de l'EHESS.

Citron, D.K. (2014), *Hate Crimes in Cyberspace,* Cambridge, MA and London, Harvard University Press.

Coleman, G. (2013), « Anonymous and the politics of leaking », in *Beyond WikiLeaks. Implications for the Future of Communications, Journalism and Society,* sous la direction de B. Brevini, A. Hintz, & P. McCurdy, Basingstoke and New York, Palgrave Macmillan, p. 209-228.

Custers, B. *et al.* (dir.), (2013), *Discrimination and Privacy in the Information Society. Data Mining and Profiling in Large Databases,* New York, Springer.

Deleuze, G. (1990), « Postscriptum sur les sociétés de contrôle », in *Pourparlers 1972-1990,* Paris, Minuit, p. 240-247.

Dennis, K., « *Keeping a close watch* – the rise of self-surveillance and the threat of digital exposure », *The Sociological Review* vol. 56, n° 3, août 2008, p. 347-357.

Desrosières, A. (2010) [1993], *La Politique des grands nombres. Histoire de la raison statistique,* nouvelle édition, Paris, La Découverte.

Froomkin, A.M. (2000), « The death of privacy? », *Stanford Law Review* vol. 52, p. 1461-1543.

Gandy, O.H. (1993), *The Panoptic Sort. A Political Economy of Personal Information,* Boulder, CO, Westview Press.

Garfinkel, S. (2000), *Database Nation. The Death of Privacy in the 21st Century,* Beijing and Cambridge, O'Reilly.

Gauchet, M. (2005), « L'Etat au miroir de la raison d'Etat », in *La Condition politique,* Paris, Gallimard, p. 205-260.

Goëta, S. & Mabi, C. (2014), « L'open data peut-il (encore) servir les citoyens ? », *Mouvements* n° 79, p. 81-91.

Goldstein, B. & Dyson, L. (dir.), (2013), *Beyond Transparency. Open Data and the Future of Civic Innovation,* San Francisco, CA, Code for America Press.

Gurstein, M.B. (2011), « Open data: empowering the empowered or effective data use for everyone? », *First Monday [online]* vol. 16 n° 2.

Harcourt, B.E. (2007), *Against Prediction. Profiling, Policing, and Punishing in an Actuarial Age,* Chicago, IL and London, University of Chicago Press.

Horn, E. (2011), « Logics of political secrecy », *Theory, Culture & Society* vol. 28 n° 7-8, p. 103-122.

Howard, P.N. (2006), *New Media Campaigns and the Managed Citizen,* Cambridge, Cambridge University Press.

Issenberg, S. (2012), *The Victory Lab. The Secret Science of Winning Campaigns,* New York, Crown Publishers.

Lefort, C. (1986), « La question de la démocratie », in *Essais sur le politique. XIXe-XXe siècles,* Paris, Seuil, p. 17-32.

Loveluck, B. (2015), *Réseaux, libertés et contrôle. Une généalogie politique d'internet,* Paris, Armand Colin.

Lyon, D. (dir.) (2003), *Surveillance as Social Sorting. Privacy, Risk and Automated Discrimination,* London and New York, Routledge.

Meyers, G. (2015), « Personal data tracking in insurance: from solidarity to fairness? », présenté lors du colloque international *Profile, Predict and Prevent. Data-Driven Policies, Markets and Societies*, organisé par le CERSA (CNRS-Paris 2), Paris.

Nissenbaum, H. (2010), *Privacy in Context. Technology, Policy, and the Integrity of Social Life,* Stanford, CA, Stanford University Press.

Manyika, J. *et al.* (2013), *Open Data: Unlocking Innovation and Performance with Liquid Information*, rapport du McKinsey Global Institute.

O'Reilly, T. (2010), « Government as a platform », in *Open Government. Collaboration, Transparency, and Participation in Practice,* sous la direction de D. Lathrop & L. Ruma, Sebastopol, CA, O'Reilly, p. 11-39.

Parasie, S. (2013), « Des machines à scandale. Eléments pour une sociologie morale des bases de données », *Réseaux* n° 178-179, p. 127-161.

Parikka, J. & Sampson, T. (2009), « On anomalous objects of digital culture. An introduction », in *The Spam Book. On Viruses, Porn, and other Anomalies from the Dark Side of Digital Culture,* sous la direction de J. Parikka & T.D. Sampson, Cresskill, NJ, Hampton Press, p. 1-18.

Rosanvallon, P. (2006), *La Contre-démocratie. La politique à l'âge de la défiance,* Paris, Seuil.

Rouvroy, A. & Berns, T. (2010), « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps "numériques" », *Multitudes* n° 40, p. 88-103.

Schneewind, B. (2001), *L'Invention de l'autonomie. Une histoire de la philosophie morale moderne,* Paris, Gallimard.

Segal, J. (2003), *Le Zéro et le Un. Histoire de la notion scientifique d'information au 20e siècle,* Paris, Syllepse.

Senellart, M. (1995), *Les Arts de gouverner. Du regimen médiéval au concept de gouvernement,* Paris, Seuil.

Supiot, A. (2015), *La Gouvernance par les nombres. Cours au Collège de France, 2012-2014,* Paris, Fayard.

The Economist (2010), « The data deluge », 25 février.