la vie des idées.fr

L'emprise des algorithmes

par Sonia Desmoulin-Canselier

Les algorithmes qui exploitent les données pourraient éclairer les décisions publiques au service des citoyens. Mais leurs utilisations actuelles par les GAFA et les entreprises de la finance sont opaques et portent, parfois directement, atteinte aux libertés. Peuvent-elles être réglementées ?

Recensé : Frank Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Information*, Harvard, Harvard University Press, 2015, 320 p.

L'affaire Cambridge Analytica/Facebook n'est qu'un épiphénomène d'un mouvement plus profond. Tel qu'il est actuellement pratiqué, le recours aux Big Data et aux algorithmes a un coût : celui du renoncement progressif à des principes et à des valeurs, comme la protection de la vie privée, l'exercice loyal de la concurrence ou l'accès aux informations permettant de se défendre lorsque des décisions nous affectent. Cette face sombre de l'innovation technologique préoccupe en Europe (comme en témoignent les nombreux rapports récemment publiés, par exemple par l'OPECST ou la CNIL), mais semble produire ses effets les plus néfastes aux États-Unis.

C'est ce qui ressort du livre de Frank Pasquale dénonçant les affres d'une « société boîte noire ». L'auteur y déploie la métaphore dans sa double dimension technique et conceptuelle (p. 3). Désignant le système d'enregistrement embarqué dans les moyens de transport (par exemple les avions), une boîte noire est aussi un dispositif opaque et fermé, inaccessible au regard, selon le sens que les sciences sociales, et notamment B. Latour, ont donné à l'expression. L'auteur constate ainsi que les citoyens étatsuniens — et par extension tous ceux qui ont affaire aux entreprises étatsuniennes — sont plus que jamais pistés et jaugés sans avoir la moindre idée de la façon dont leurs données circulent et sont utilisées ni des conclusions qui en sont tirées. Un système opaque et déséquilibré permet, au contraire, aux agences gouvernementales et aux grandes entreprises de la banque, de la finance et du secteur des TIC d'imposer la confidentialité de pratiques contestables.

En reliant les différents types de secrets (secret des affaires, secrets industriels, secrets d'État, etc.) à l'hermétisme des outils algorithmiques, F. Pasquale veut montrer que la complexité des algorithmes constitue le versant technique d'une opacité plus globale. Dans ce contexte, l'opposition entre pouvoirs publics et sociétés privées n'a plus cours (p. 206). Au nom de la lutte contre le terrorisme, le gouvernement et les autorités administratives des États-Unis font cause commune avec ceux qui détiennent les bases de données qu'ils n'ont pas pu eux-mêmes constituer (p. 21, 49, 51, 155 et 184). Les mêmes outils servent à classer et à hiérarchiser les citoyens ou les entreprises concurrentes. Selon F. Pasquale, cette connivence serait payée au prix fort : par une dégradation des garanties accordées aux citoyens et aux petits acteurs économiques en matière d'atteinte à la vie privée et par la perte progressive de pouvoir et de légitimité des autorités publiques. Il appelle donc à sortir de ce dédale en répartissant plus équitablement les droits et devoirs liés à l'information et au secret, et en restaurant les autorités publiques dans un rôle de défense de l'intérêt général incluant des prises d'intérêt dans les secteurs économiques cruciaux. Pour le lecteur européen, sa critique et ses propositions sont éclairantes à un double titre. Elles montrent les différences qui perdurent d'un continent à l'autre, mais elles indiquent aussi les questions qui nous rassemblent.

Une hydre à deux têtes

Le labyrinthe a été construit au moyen de deux outils : les Big Data et les algorithmes. Sur Internet et par les objets connectés, chacun distille, de manière plus ou moins consciente, des informations sur sa vie, ses goûts, son comportement. Au delà de ce que nous souhaitons faire savoir et sans que nous en soyons toujours informés, des données personnelles sont constamment collectées, archivées, traitées et cédées. Ces informations peuvent être compilées et croisées avec d'autres sources, issues de bases de données privées ou publiques, et le phénomène s'étend avec la numérisation croissante des activités sociales (p. 20-37). Ce flux massif de données fait l'objet d'un traitement automatisé, par le biais d'algorithmes variés : pour la saisie des données, le calcul du résultat, leur affichage, la communication avec d'autres logiciels, etc. (voir la définition de la CNIL). Parmi ceux-ci, les algorithmes de classement, de suggestion et de décision produisent les effets les plus problématiques, car leurs résultats ont une incidence directe sur les décisions et les vies humaines (p. 84-91, 194). Une offre d'emploi ou de service peut ainsi être rendue invisible par un moteur de recherche intégrant un critère dissimulé. Une demande de prêt pourrait être refusée si le calcul de risque inclut des données relatives à la fréquentation du système de santé ou aux statistiques de délinquance dans le quartier du lieu de résidence. Les Big Data ne peuvent être exploitées sans les algorithmes et ces derniers ne peuvent fonctionner sans être nourris de données.

L'attrait exercé par les traitements algorithmiques tient à leur rapidité, à leur efficacité et à leur apparente objectivité (p. 25 et 107). Comment se passer d'un outil qui nous dispense d'un travail long et fastidieux et nous évite d'avoir à fournir des justifications ? Pourtant, cette

préférence pour l'efficacité a un prix, car elle se déploie au détriment de la dignité et de la justice (p. 199). La délégation de nos choix aux traitements algorithmiques a plusieurs conséquences fâcheuses. Pour les recherches sur Internet, l'inférence statistique en fonction de préférences (réelles ou supposées) passées conduit à « personnaliser » des résultats au point de ne plus rendre perceptible le biais induit : nous pensons choisir, mais notre liberté ne s'exerce que dans la limite de ce qui a été sélectionné pour nous (p. 79). Pour les classements et les décisions, l'outil algorithmique évalue et projette à partir des données numériques disponibles sur les faits passés. Il nous rend aveugles à tout ce qui n'est pas quantifiable ou numérisable (p. 8 et 191) et repose sur une logique de reproduction (p. 41, 110). Or, parier sur l'absence de changements n'est ni nécessairement pertinent, ni toujours opportun. D'autant que l'apparente objectivité du résultat ne résiste pas à l'examen approfondi des conditions de son obtention, car sa dépendance aux données (et aux critères de hiérarchisation) le rend vulnérable aux biais. Par exemple, un logiciel d'évaluation des personnes et des zones géographiques « à risque » pénal fonctionnant avec les données issues des contrôles policiers ne donnera un résultat « objectif » que si les pratiques policières sur la période passée ont été véritablement égalitaires, car des contrôles plus fréquents induisent un taux plus important de découverte d'infractions (p. 39-40). Toute éventualité de contrôles au faciès ou de surveillance accrue dans un quartier annihile une prétention sérieuse à la neutralité. De manière plus générale, la logique algorithmique cache la part de choix inhérente à toute décision derrière le paravent du résultat chiffré (p. 24, 41), ce qui est éminemment problématique dans un État de droit, où les justiciables qui pâtissent d'une décision doivent pouvoir en comprendre les raisons pour être en mesure de la contester (p. 149, 164, 198).

En finir avec l'opacité et la passivité ?

Loin d'offrir toujours sa protection aux citoyens en quête d'explications et de justifications, le droit peut être un terrible instrument au service des puissants. Profitant de la protection contractuelle et judiciaire, les secrets industriels et d'affaires se développent de manière inversement proportionnelle à la protection de la vie privée (p. 26, 161 et 183). Le droit des libertés publiques qui peine à protéger les personnes physiques devient une ressource pour les grandes sociétés. Google a mobilisé le Premier amendement de la constitution étatsunienne, faisant valoir la liberté d'expression pour ne pas avoir à justifier un classement de son moteur de recherche (p. 32 et p. 166-167). Sans souci de cohérence, l'entreprise s'est par ailleurs prévalue de son statut d'intermédiaire sans responsabilité éditoriale (et donc d'une posture de transmetteur passif) lorsque des actions en diffamation ou en violation du droit d'auteur ont été engagées contre elle (p. 77-78). Les secteurs de l'information et de la finance ont pris tant d'importance dans l'économie, que les pratiques frauduleuses ou déloyales des grandes sociétés — lorsqu'elles sont détectées — donnent lieu à des procédures souvent confidentielles pour éviter des retombées économiques désastreuses, à des sanctions ridiculement faibles en proportion des gains et à un apurement des dettes au moyen des deniers publics (p. 22, 158).

Même lorsque la violation des règles relatives à la protection des données personnelles ou d'utilisation d'algorithmes discriminatoires est prouvée, la justice pénale semble moins apte aujourd'hui à condamner les fraudeurs que par le passé (p. 176). Derrière l'expression « *too big to fail* » (p. 177), c'est une logique infernale qui est à l'œuvre : en recourant à des moyens illégaux (et/ou immoraux) pour maximiser son profit, une entreprise peut parvenir à occuper une place stratégique hors d'atteinte des régulateurs dès lors que le dépassement d'un seuil de risque aboutit paradoxalement à une solution plus clémente en cas de perte de contrôle (p. 174, 178). Dans ce contexte, l'appel à transparence peut paraître illusoire.

Cet imbroglio est-il inextricable ? Frank Pasquale estime que les États-Unis ont connu par le passé des situations comparables, quand des systèmes opaques et frauduleux permettaient à certaines entreprises capitalistes de s'emparer d'une position dominante sur des services primordiaux et que des solutions avaient alors pu être trouvées. À titre d'exemple de lutte contre la fraude, il cite la réponse réglementaire et technique apportée pour mettre un terme au système sophistiqué de surfacturation organisé par les établissements de santé étatsuniens au détriment du système d'assurance santé Medicare (p. 179 sq.). Les pouvoirs publics étatsuniens pourraient aujourd'hui reprendre la main en imposant aux GAFA et aux banques mondialisées le respect des principes de loyauté et de non-discrimination dans les traitements algorithmiques et en demandant davantage de transparence dans les conditions de collecte et d'utilisation des données personnelles. Leurs capacités de surveillance sont importantes : elles pourraient être réaffectées au contrôle des grandes entreprises, plutôt qu'à celui des individus en raison de leurs engagements politiques (p. 13 et 218). De même, la lutte contre le terrorisme ne devrait plus servir à couvrir le développement de manœuvres commerciales douteuses, au nom d'une improbable alliance entre l'État et les sociétés privées (p. 184). Frank Pasquale nous invite à abandonner l'actuelle attitude passive et victimaire vis-à-vis d'une technologie qui s'imposerait d'elle-même et d'une concurrence internationale qui justifierait de brader tous les principes (p. 171 et 197). Les techniques actuellement déployées pour collecter des informations sur les citoyens-consommateurs à des fins peu compatibles avec leurs intérêts (profilage, calcul de risque, surveillance, etc.) pourraient être utilisées différemment, par exemple en intégrant l'injonction au respect de la vie privée dès la conception (*privacy by design* : p. 157) ou en exigeant que des systèmes d'enregistrement des décisions (enregistrements automatiques, inaltérables et accessibles aux contrôleurs) soient imposés aux organismes de crédit (p. 158). Aux États-Unis, certains organismes de contrôle existent déjà, notamment en matière bancaire, mais davantage de moyens d'agir devraient leur être accordés en développant ce que l'auteur dénomme une « transparence qualifiée » (p. 160 sq.).

Bien que les illustrations et les propositions se réfèrent aux États-Unis et à leur ordre juridique, l'état des lieux et l'analyse critique sont pertinents pour des lecteurs européens. Certes, les droits européen et français intègrent depuis de nombreuses années des textes exigeants en matière de respect de la vie privée et de traitement des données personnelles. Depuis 1978 en

France[1] etla Directive du 24 octobre 1995 dans l'Union européenne[2], les citoyens français et européens se voient accorder des droits d'information, d'accès, de rectification et d'opposition, tandis que les responsables de traitement doivent respecter des procédures de déclaration ou d'autorisation (pour le traitement des données personnelles et particulièrement les données sensibles, comme les données relatives à l'état de santé, aux préférences sexuelles ou aux choix politiques). Lorsque les données personnelles font l'objet d'un traitement automatisé, un certain nombre de conditions doivent être réunies, comme le consentement et/ou l'autorisation légale, accompagnée de garanties en matière d'information et de contestation. De valeur équivalente aux traités fondateurs de l'Union européenne, la Charte des droits fondamentaux du 7 décembre 2000 prévoit en son article 8 que « toute personne a droit à la protection des données à caractère personnel la concernant » et que « ces données doivent être traitées loyalement, à des fins déterminées et sur la base du consentement de la personne concernée ou en vertu d'un autre fondement légitime prévu par la loi ». Le nouveau Règlement européen sur la protection des données personnelles[3], entré en application le 25 mai 2018, a renforcé les pouvoirs accordés aux autorités nationales et accordé quelques droits supplémentaires aux personnes (par exemple, reconnaissance d'un véritable « droit à l'oubli » par la conjugaison du droit à l'effacement, de l'opposition à la publication d'informations sur un site et de la demande de suppression d'un contenu). De plus, ainsi que le souligne F. Pasquale, les juges et les autorités en Europe semblent plus exigeants, plus créatifs et plus efficaces que leurs homologues étatsuniens pour faire assumer aux GAFA leurs responsabilités (p. 69, 163 et 197). Des condamnations ont déjà été prononcées en France, en Allemagne et par des instances européennes (CJUE ou Commission) pour atteinte à la vie privée, non-respect des prescriptions réglementaires en matière de traitement des données personnelles ou atteinte aux principes du droit de la concurrence.

Néanmoins, même si les autorités et les agences ne sont pas les mêmes, un lecteur français s'approprie sans difficulté le constat d'insuffisance des ressources des contrôleurs étatiques (notamment p. 22, 58 et 177, où l'auteur constate que l'expression « trop important pour faire faillite » se complète par « trop pauvre pour contrôler » : « "*Too Big to Fail*" *meets "Too Poor to Regulate*" »). De plus, les données circulent mondialement et les citoyens européens sont exposés aux carences de l'ordre juridique étatsunien. Les failles du Traité Safe Harbor, organisant les conditions de circulation des données personnelles entre l'Europe et les États-Unis, ont conduit à une censure par la Cour de justice de l'Union (CJUE, 6 octobre 2015, Shrems), mais le Privacy Shield Act qui l'a remplacé en juillet 2016 montre déjà ses limites. Le G29, organisation réunissant les autorités type CNIL de l'UE, a rendu un rapport très critique et menacé de saisir les juges européens si des améliorations n'étaient pas apportées[4], notamment au regard des risques d'atteinte à la protection de la vie privée que la loi étatsunienne sur la

[1] Loi informatique et Libertés du 6 janvier 1978 modifiée.
[2] Directive 95/46/CE relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données remplacée par le RGDP.
[3] RGDP : Règlement [UE] 2016/679 du 27 avril 2016.
[4] EU-U.S. Privacy Shield – First annual Joint Review, 28 novembre 2017.

surveillance et le renseignement étranger représente (Foreign Intelligence Surveillance Act, ou FISA).

Le droit européen pourrait d'ailleurs être encore amélioré. Il serait notamment utile de repenser le champ d'application des textes, aujourd'hui limité aux données personnelles, et la catégorisation des données (données de santé, données particulières, etc.) en tenant compte des moyens techniques de ré-identification et de la possibilité d'obtenir des résultats aux effets équivalents à partir de croisement de bases de données (p. 28 et 147). Enfin, les juristes européens découvrent aussi les difficultés posées par l'usage d'algorithmes décisionnaires pour établir individuellement les responsabilités lorsque les règles font référence à des intentions (élément moral de l'infraction en matière pénale, p. 173). Si les propositions de F. Pasquale ne sont pas toutes convaincantes, les plus radicales impliquant une expansion discutable de l'interventionnisme étatique (p. 204-211), et si certaines manquent de précision (par exemple, l'idée d'instaurer un agrément des algorithmes décisionnaires inspiré de celui de la *Food and Drug Administration*, p. 181), le propos n'en est pas moins pertinent et stimulant. Il est effectivement temps de redonner aux personnes qui subissent le système un peu de discernement et de restaurer l'État dans ses fonctions régaliennes au service de l'intérêt général.

Pour aller plus loin :

- Serge Abiteboul et Gilles Dowek, *Le temps des algorithmes*, Le Pommier, 2017.
- Bilel Benbouzid, « À qui profite le crime ? Le marché de la prédiction du crime aux États-Unis », *La Vie des idées*, 13 septembre 2016. URL : http://www.laviedesidees.fr/A-qui-profite-le-crime.html
- Dominique Cardon, *À quoi rêvent les algorithmes. Nos vies à l'heure des big data*, Seuil/La République des idées, 2015.
- CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, décembre 2017. URL : https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de
- Virginie Gautron et David Monniaux, « De la surveillance secrète à la prédiction des risques : les dérives du fichage dans le champ de la lutte contre le terrorisme », *Archives de politique criminelle*, 2016, 38, p. 123-135. URL : https://hal.archives-ouvertes.fr/hal-01446359/document
- Cathy O'Neil, *Weapons of Math Destruction : How Big Data Increases Inequality and Threatens Democracy*, Crown Random House, 2016.
- Bruno Latour, *La science en action. Introduction à la sociologie des sciences*, La Découverte, 1989.
- Antoinette Rouvroy, « Pour une défense de l'éprouvante inopérationnalité du droit face à l'opérationnalité sans épreuve du comportementalisme numérique », *Dissensus*

[En ligne], Dossier « Efficacité : normes et savoirs », n° 4 (avril 2011). URL : http://popups.ulg.ac.be/2031-4981/index.php?id=963.

- Éric Sadin, *La vie algorithmique*, Éditions L'échappée, 2015.

Publié dans laviedesidees.fr, le 20 juin 2018.