Données. Quand le numérique produit et gouverne la ville

Matthieu Adam, Hervé Rivano

▶ **To cite this version:**

Matthieu Adam, Hervé Rivano. Données. Quand le numérique produit et gouverne la ville. Matthieu Adam; Émeline Comby. Le Capital dans la cité, Éditions Amsterdam, pp.125-137, 2020, 9782354802172. halshs-03027296

HAL Id: halshs-03027296
https://shs.hal.science/halshs-03027296

Submitted on 20 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adam, Matthieu, et Hervé Rivano. « Données. Quand le numérique produit et gouverne la ville ». In *Le capital dans la cité. Une encyclopédie critique de la ville*, édité par Matthieu Adam et Émeline Comby, p.125-37. Paris : Éditions Amsterdam, 2020. Version auteurs.

Données

Quand le numérique produit et gouverne la ville

Des bouchons s'annoncent sur le périphérique. Vous lancez *Waze*, que vous avez installé comme 130 millions de personnes à travers le monde (dont 14 millions en France en décembre 2019). L'application, propriété de *Google*, vous suggère un itinéraire qui fait gagner 45 minutes en empruntant des routes secondaires. Elle vous indique aussi les zones où les contrôles de police sont fréquents et vous signale que vous allez passer devant un fast-food et une boutique de pneus.

Waze s'appuie sur les données de ses utilisateurs et de ses utilisatrices pour connaître l'état du trafic en temps réel. Une fois activée, l'application récupère leurs positions et leurs vitesses. Elle en déduit l'existence de bouchons et calcule des itinéraires optimisés avec une précision dépendant du nombre de personnes connectées à un endroit donné. En l'absence d'informations sur une voie, *Waze* considère qu'elle est libre et y envoie automobilistes et chauffeurs routiers, sans considération pour les plans de déplacement définis par les agglomérations. En cherchant à vous faire échapper à un embouteillage, il arrive donc à *Waze* d'en générer un nouveau.

À durée équivalente, *Waze* privilégie l'itinéraire qui passe devant ses annonceurs les plus généreux, car ses bénéfices viennent surtout de l'exposition à la publicité, achetée par les commerces à la régie publicitaire *Google Ads*. L'application ne vous fait cependant pas faire un détour pour passer devant un magasin : il est dans son intérêt de satisfaire ses usagers et usagères, afin que leur nombre croisse et que les données de mobilité affluent. Celles-ci sont une mine d'or pour les collectivités territoriales ou les sociétés d'assurance, qui s'en servent pour ajuster leurs politiques ou leurs tarifs. *Waze* vend par exemple des données aux villes (cartographie en temps réel, statistique d'usage des tronçons, identification des sites attractifs, etc.), en arguant du fait qu'elle connaîtrait mieux qu'elles l'utilisation de leurs routes. Aux collectivités qui sont sous la pression de politiques d'austérité budgétaire et ne peuvent s'offrir un tel service, *Waze* propose désormais de transmettre « gratuitement » ses données. En échange, ces villes doivent mettre « gratuitement » à disposition leurs propres données, auxquelles l'application n'aurait pas eu accès sans ce troc (location de vélos en libre-service, transports en commun, parkings). Ce faisant, *Waze* fait fructifier les traces numériques de ses usagers et de ses usagères, dans une démarche parfois décrite comme gagnant-gagnant, mais dont l'application reste le principal bénéficiaire. Cet exemple illustre le fonctionnement de l'économie de la donnée : « Si c'est gratuit, c'est que vous êtes le produit. »

Individualisation et optimisation, l'eldorado de la donnée

À l'orée des années 1990, les capteurs en mesure d'enregistrer des grandeurs physiques – températures, concentrations, pressions, vitesses, poids – sous un format numérique, et non plus analogique, commencent à se multiplier. Les outils informatiques s'installent dans le quotidien d'une part croissante de la population mondiale et la donnée pénètre ainsi progressivement la vie quotidienne. La fin des années 2000 voit l'apparition des smartphones, les années 2010 leur généralisation : il y a aujourd'hui plus d'un téléphone actif par personne en France, dont les utilisateurs génèrent en permanence une grande quantité de données.

Une donnée est une valeur numérique qualifiée, c'est-à-dire un chiffre ou un ensemble de chiffres dont on connaît l'origine, le mode de mesure, la précision et ce qu'il représente. Brutes, elles ont rarement de la valeur, et leur nature peut être de plusieurs ordres. Par exemple, les traces de nos activités numériques : avec qui nous interagissons sur les réseaux sociaux, messageries ou

téléphones, quels contenus nous cherchons ou consultons sur le web, etc. Nos activités physiques génèrent également des données : nos déplacements sont enregistrés et utilisés par les systèmes d'exploitation de nos smartphones (iOS, Android), par des fournisseurs d'applications ou encore par notre opérateur téléphonique. Nous laissons aussi des traces sur les images de vidéosurveillance, par nos relevés de consommations électriques ou par nos paiements. D'autres données sont *a priori* non personnelles, d'ordre environnemental ou économique par exemple : mesures de température, de précipitations ou de pollution, données des réseaux urbains de fourniture d'eau ou d'énergie, informations sur les transactions immobilières, etc.

À l'image des matières premières, les données constituent un gisement de ressources exploitables dont l'extraction puis la transformation (croisements, traitements algorithmiques) en informations créent de la valeur. Ces informations sont ensuite utilisées afin de réguler des systèmes opérationnels (par exemple des capteurs dans un bâtiment qui pilotent la ventilation, les volets et le chauffage), de fournir des services (comme une sélection personnalisée de film sur une plateforme de vidéo à la demande) ou d'identifier des comportements humains (par exemple pour du profilage commercial et de la publicité ciblée). Savoir où et quand une personne se déplace a par exemple un intérêt limité. En revanche, identifier où elle s'arrête, où elle ralentit, puis croiser cela avec la localisation des commerces permet progressivement de décrire ses habitudes et de savoir quand elles changent. Si une personne se met à ralentir systématiquement devant les magasins de vélo et qu'en parallèle elle fait des recherches sur différents modèles, les algorithmes des sociétés spécialistes du ciblage publicitaire, comme *Google* ou *Facebook*, vont l'exposer à des publicités pour des vélos, maximisant ainsi les chances qu'elle clique dessus, et donc le profit réalisé.

La grande diversité, l'immense quantité et l'actualisation continue des données disponibles permettent de proposer des services inédits. Pouvoir être guidé en temps réel en fonction du trafic, comme le propose *Waze*, était inimaginable il y a quelques années. Identifier un restaurant d'un type spécifique, dans un quartier donné, et savoir qu'il est apprécié par des personnes qui ont les mêmes goûts que soi, comme le permettent *TripAdvisor* ou *Yelp*, ne l'était pas davantage. En rendant possible de confronter ses propres envies aux opportunités offertes par un territoire, la donnée conduit ainsi à une forme inédite d'individualisation des expériences.

À un autre niveau, avoir des données sur la consommation d'eau, d'électricité ou de gaz des ménages, connaître le trafic automobile, le nombre de passagers des bus ou le volume de déchets ramassés de façon précise, régulière et géolocalisée permet d'optimiser les systèmes urbains : les données disponibles servent à calibrer, parfois en temps réel, les réseaux de distribution de fluides et d'énergie, le séquençage des feux ou le ramassage des ordures. Le compteur *Linky*, par exemple, est conçu pour optimiser le réseau de distribution d'électricité. En mesurant la consommation quelques fois par jours, il permet de savoir quand vous consommez de l'électricité. En la mesurant plusieurs fois par seconde, il permettrait d'identifier pour quels usages et avec quels appareils. La donnée améliore la connaissance de ces systèmes et la capacité de réaction, ce qui permet de limiter les coûts ou les pertes de temps et de ressources.

De l'extraction des gisements de données à leur exploitation

La donnée est désormais partout et son économie incontournable : le marché français de la donnée (hébergement, traitement et autres activités) a généré 8,6 milliards de dollars de revenus en 2018, avec des projections à 9,3 milliards en 2024 (Statista Eurostat 2020). La valeur économique des données se crée au fur et à mesure d'un processus que l'on peut découper en trois grandes étapes – extraction, transformation, exploitation – et qui implique une grande diversité d'acteurs économiques.

[intégrer schéma données flèches]

Schéma de principe de l'économie de la donnée (Matthieu Adam, 2020)

La première étape consiste en l'extraction des données afin de constituer des gisements numériques. Cette extraction est réalisée par des capteurs, des systèmes de mesure d'observation ou d'instrumentation, embarqués dans nos *smartphones*, véhicules ou ordinateurs personnels, intégrés dans les services que nous utilisons ou déployés dans l'environnement urbain. Les citoyens et citoyennes n'ont pas toujours conscience que leurs données sont extraites ou collectées, même si les régulations telles que le RGPD (Règlement général sur la protection des données entré en vigueur en 2018 dans l'Union européenne) poussent à davantage les informer. Plusieurs types d'entreprises sont mobilisées. Celles dont le commerce de la donnée est le cœur de métier, comme *Waze*, *TripAdvisor* ou les réseaux sociaux, peuvent fournir des services pour collecter des données afférentes aux conditions d'usage de ces services. Elles travaillent alors à améliorer leurs services : plus ils sont satisfaisants, plus ils attirent des usagers et des usagères sources de données. D'autres entreprises tirent leurs revenus d'activités différentes, mais génèrent des données pour leurs propres besoins. Le gisement de données ainsi créé peut avoir un tel potentiel commercial qu'il devient une opportunité de diversification et de croissance. Par exemple, Veolia utilise des capteurs pour optimiser la gestion de leurs réseaux d'eau. Ils génèrent ainsi une grande masse de données, que la société peut ensuite valoriser. Il existe aussi des situations hybrides ou l'activité secondaire de collecte de données devient la source de revenus principale. C'est probablement le cas des opérateurs de mobilité à bas coût – vélos (*Uber*), trottinettes (*Lime*, *Bird*) ou VTC (*Uber*, *Lyft*) –, dont certains ont déjà cessé leur activité primaire non rentable dès que la maintenance devenait nécessaire.

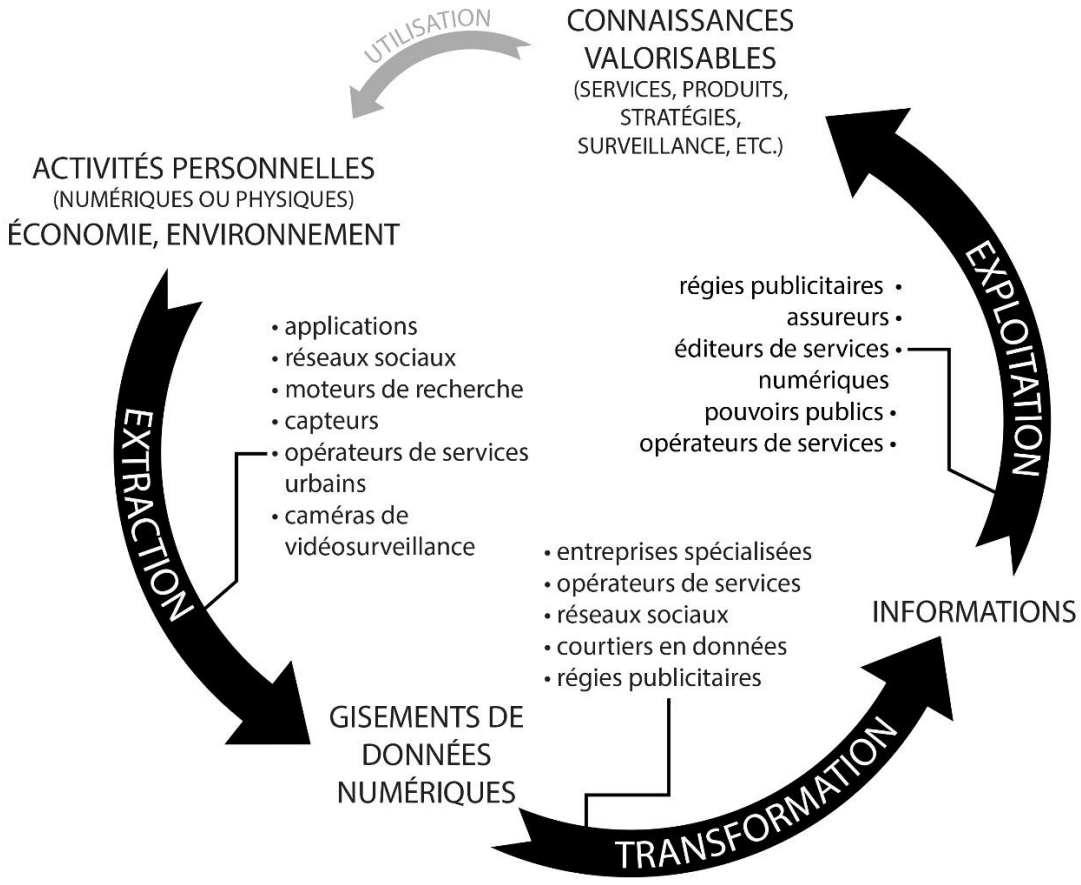

Fig. 1 : Schéma de principe de l'économie de la donnée (2020)

La donnée doit ensuite subir une étape de transformation. Entrent alors en scène des algorithmes d'intelligence artificielle, de classification et d'apprentissage qui permettent de croiser des données d'origines et de natures variées pour en déduire des informations de plus grande valeur. Dans cette phase, ou à l'interface entre celle-ci et la précédente, on trouve les *data brokers* (courtiers en données). Ces intermédiaires achètent ou récupèrent des gisements auprès d'entreprises ou de collectivités, puis les transforment avant de les revendre. L'exemple le plus connu est *Cambridge Analytica*, célèbre pour avoir aspiré les données de dizaines de millions de comptes *Facebook* dans le but d'influencer les présidentielles étasuniennes de 2016 et le référendum sur le Brexit.

La troisième et dernière étape de l'économie de la donnée est l'exploitation des informations tirées des gisements. Le secteur commercial le plus avide de données est la publicité. *Google* et *Facebook* sont des régies publicitaires 2.0 : elles vendent aux publicistes l'accès à des clientèles ciblées grâce à leurs données, censées être intéressées par leurs produits. Leur diversité d'action (réseaux sociaux, OS, applications, etc.) vise à maximiser les moments durant lesquels les usagers et usagères sont exposés aux publicités et fournissent de la donnée. C'est ce que théorise « l'économie de l'attention » : face à une offre abondante, la ressource à valoriser devient le temps et l'attention des consommateurs et consommatrices (Citton 2014). Les sociétés d'assurance sont elles aussi friandes de données : en améliorant leurs connaissances des comportements et des contextes, elles leur permettent de préciser leurs évaluations des risques, d'individualiser les tarifs des polices, d'optimiser leurs stratégies et donc d'améliorer leurs revenus. En pleine croissance, le secteur de la sécurité s'appuie largement sur des données individuelles pour identifier les comportements délinquants ou déviants. La police prédictive, ou *predpol*, s'appuie par exemple sur des algorithmes qui ingèrent quantité d'informations statistiques et de données personnelles pour identifier des zones et des individus supposément à risque (Benbouzid 2015). Enfin, les collectivités territoriales et les autres acteurs de l'urbanisme, de l'immobilier et des transports commencent eux aussi à investir dans la donnée, dans le but d'optimiser leurs politiques, leurs réseaux ou leurs investissements.

Individualisation ou uniformisation des comportements ?

L'individualisation des services conduit en théorie à un capitalisme plus horizontal. Pour qu'un restaurant intègre le Guide *Michelin*, il faut qu'il soit identifié par les inspectrices et inspecteurs, qu'ils y mangent et qu'ils rédigent leur avis. En s'appuyant sur le jugement d'une multitude d'individus, *TripAdvisor* ou *Google Maps* promettent qu'un restaurant puisse émerger, grâce à un bouche-à-oreille numérique. Or, la quantité d'adresses indexées trahit cette promesse : un restaurant répertorié sur la dixième page de *TripAdvisor* a peu de chance de voir sa popularité grandir par le seul truchement de l'accumulation d'avis. Ce fonctionnement profite ainsi surtout à qui peut payer pour être mis en avant. Cette horizontalité affichée ne freine donc ni la publicité ni la concentration des richesses.

La volonté d'individualisation de l'expérience urbaine conduit en fait à une uniformisation des comportements de consommation : le fait que les goûts soient des marqueurs sociaux est renforcé par les algorithmes dits de *clustering*. En agrégeant des données en fonction de leurs caractéristiques communes, ils tendent à enfermer les individus dans des bulles informationnelles qui leur font toujours les mêmes propositions commerciales, gastronomiques ou culturelles (Pariser 2011). Dans des espaces urbains de plus en plus semblables (Berdet 2013), l'individualisation grâce à la donnée se réduit finalement pour l'essentiel à l'expérience faussement unique d'un monde standardisé.

Libération de la donnée ou libéralisme des usages ?

Toutes les utilisations de la donnée ne sont pas *a priori* marchandes. En permettant de les objectiver, par la mesure et la représentation, les données ont contribué à transformer des phénomènes impalpables en sujets politiques. La pollution de l'air, par exemple, est indolore, le plus souvent incolore et inodore, mais demeure néanmoins toxique. Sa mesure citoyenne, grâce à des capteurs peu coûteux, ou la récupération des données publiques, a contribué à imposer le débat avec les responsables politiques. En 2016, dans le 1er arrondissement de Lyon, quartier gentrifié dont une partie de la population a un capital économique et culturel élevé, des parents d'élèves se sont saisis des données publicisées par *Atmo*, l'association chargée de la surveillance de la qualité de l'air. Elles leur ont permis de mettre en évidence la surexposition d'une école à la pollution atmosphérique et de légitimer leur combat aux yeux des élus et de la presse locale. En 2019, *Greenpeace* a produit une carte montrant des problématiques similaires sur toute la ville (https://www.greenpeace.fr/pollution-de-lair-a-lyon/).

Ce cas témoigne des possibilités offertes par l'*open data*, c'est-à-dire la mise à disposition libre et gratuite de données collectées en direction des populations, des institutions et des entreprises. C'est un lent mouvement qui commence, pour la France, en 1978 avec la loi CADA (Commission d'accès aux documents administratifs), et qui s'accélère mondialement à partir de la fin des années 2000 avec des initiatives comme la licence « Creative Commons » et la théorisation par Elinor Ostrom (prix Nobel d'économie 2009) des biens communs informationnels. Sous l'impulsion d'une directive européenne de 2013, les données publiques des collectivités et des états membres deviennent obligatoirement ouvertes, parmi lesquelles des données urbaines (économiques, de transport, environnementales ou foncières). L'*open data* répond à deux objectifs. Le premier est politique : il s'agit d'offrir une capacité de contrôle citoyen, en rendant disponibles des documents administratifs. L'*open data* est alors un facteur de légitimation politique, destiné à renforcer les liens entre l'institution et la population. Cela fonctionne plus ou moins bien : les plateformes de mise à disposition des données sont souvent peu ergonomiques, les données sont généralement difficiles à exploiter et de ce fait utilisables seulement par des spécialistes. Le second objectif est économique : il s'agit de permettre à des entreprises de créer de nouveaux services et de valoriser des gisements d'informations jusqu'ici peu exploités. L'*open data* sert alors plus ou moins sciemment le projet néolibéral de mettre le public au service des intérêts privés.

L'*open data* n'est pas généralisée. Certaines données, jugées sensibles, ne sont pas ouvertes. C'est le cas des données de transport en commun en temps réel, que certaines collectivités territoriales se réservent afin de garder la main sur leurs politiques de transport. Cela pointe une seconde limite de l'individualisation par la donnée : bénéficier de services ajustés à une métrique individuelle peut parfois nuire à l'intérêt général. Certaines collectivités tentent d'inverser la tendance en imposant aux opérateurs privés de partager à leur tour leurs données de mobilité ou de suivre une politique d'usage de l'espace publique. Los Angeles a ainsi suspendu l'autorisation de circuler des Jump, les vélos et trottinettes en libre-service d'*Uber*, car l'entreprise ne se conformait pas à l'obligation de partage des données.

Cet exemple illustre aussi une limite intrinsèque à l'*open data* : l'accessibilité et la gratuité des données n'est pas un gage d'égalité. L'*open data* profite d'abord à qui dispose des moyens financiers et de l'ingénierie nécessaire à leur exploitation. L'exemple de la mobilisation des parents d'élèves du 1er arrondissement de Lyon, trois ans avant une mobilisation militante à l'échelle de la ville, illustre bien l'inégalité des citoyens face à la capacité de s'emparer des données. Le cas de Google, à travers son système de cartographie *Maps*, mais aussi à travers *Waze* ou *Sidewalk Labs*, sa filiale dédiée à l'urbanisme, est aussi remarquable : l'entreprise s'appuie notamment sur l'*open data* pour nourrir ses

applications, avant de revendre les informations tirées des données de ses utilisateurs et utilisatrices… aux collectivités qui lui ont fourni une partie de ses informations de base (Vraiment Vraiment 2019). Les « géants du web », dont les plus célèbres permettent de créer les acronymes GAFAM (*Google, Apple, Facebook, Amazon* et *Microsoft*) et, depuis 2015, NATU (*Netflix*, *Airbnb*, Tesla et *Uber*, se manifestent par leur puissance économique et technique : la concentration des marchés de la publicité numérique et du traitement des données vont donc bon train (Guignard 2019). Ils apparaissent comme les grands gagnants de l'ouverture des données urbaines.

Données et justice sociale

Les professionnels de l'immobilier d'entreprise sont friands de données. Le nombre quotidien de requêtes *TripAdvisor* dans un quartier est par exemple un élément qui pourrait permettre de juger du potentiel de plus-value d'un secteur ou de fixer les prix d'un programme. Sur le marché des particuliers, les agences s'intéressent aux données qui permettent d'évaluer les établissements scolaires, les commerces de bouche d'un quartier ou encore le montant moyen des transactions immobilières. Ces données sont aussi utilisées par des institutions bancaires ou des opérateurs touristiques pour établir des classements qui hiérarchisent les villes entre elles, ainsi que par les collectivités pour informer leurs habitants. La mairie de San Francisco a par exemple mis en place un système de données ouvertes (datasf.org) où l'on peut trouver des informations sur la qualité de l'environnement, sur la desserte en transports en commun, sur les transactions immobilières, mais aussi sur le niveau fiscal moyen des ménages. Le système d'*open data* national français, *Etalab*, rend publiques les données de 15 millions de transactions, permettant de savoir à quel tarif le mètre carré se vend. Ces dispositifs ont notamment pour objectif, en mettant à disposition un maximum d'informations, d'éviter que des bulles économiques ne se forment trop vite sur des effets de réputation. Le risque est qu'ils permettent aussi d'objectiver les différences sociales entre quartiers et faciliter les mécanismes ségrégatifs : seuls les ménages qui ont accès à ces informations, qui sont en mesure de les interpréter et surtout d'investir dans les quartiers les plus prisés vont être à même d'en tirer bénéfice. Les ménages moins dotés vont quant à eux s'installer là où ils peuvent.

La donnée ne change pas complètement la donne, ce type de mécanismes ségrégatifs existant aussi sans elle. L'évolution vient du fait que des données *a priori* factuelles succèdent au bouche-à-oreille, limité à une communauté. En ce sens, la donnée promet de concrétiser le rêve d'Adam Smith : la liberté de choisir fondée sur une information pure et parfaite. Bien sûr, l'information tirée des données numériques n'est ni pure ni parfaite, mais l'illusion libérale opère : les promoteurs de la donnée avancent qu'un libéralisme informé serait déjà plus juste qu'un libéralisme sans information.

L'accès aux informations pose lui-même un problème de justice sociale : pour profiter des possibilités ouvertes par la donnée, il faut savoir interagir avec un système numérique. L'illectronisme, soit le fait d'avoir des difficultés ou d'être dans l'impossibilité de le faire, dessine alors une fracture qui n'est pas seulement générationnelle : il touche aussi des personnes jeunes non formées aux outils numériques ou que la dyslexie empêche de lire un écran. Leur capacité à effectuer des actions administratives ou à trouver des informations en ligne est limitée, tout comme leur accès à l'emploi. La question se pose aussi pour les personnes déjà en poste. Être chauffeur de bus, intervenir sur un réseau urbain d'énergie ou entretenir un parc public implique désormais d'être capable d'interagir constamment avec un écran tactile. La donnée change donc la nature des métiers, voire conduit à leur disparition. Là où les capteurs et les algorithmes informent, la connaissance fine d'un système (de transport, de distribution, de biodiversité) est moins une compétence reconnue, voire jugée nécessaire. Par exemple, le suivi en temps réel des réseaux d'énergie grâce aux compteurs dits intelligents fait progressivement disparaître le relevage manuel

et les emplois correspondants. Dans le même temps, un prolétariat numérique extrêmement précaire, chargé d'entraîner les algorithmes d'intelligence artificielle (en insérant à la chaîne des données) ou de réaliser les tâches que ceux-ci ne sont pas encore capables de faire (conduire un VTC *Uber*, recharger une trottinette électrique), voit le jour (Casilli 2019).

La révolution numérique change ainsi profondément les processus de production. Comme lors des révolutions industrielles qui l'ont précédé (Noiriel 2011), tout le corps social n'est pas armé de la même manière pour l'affronter ou s'y adapter.

De l'eldorado de la donnée à l'e-panoptique

Les professions qualifiées souffrent elles aussi d'une concurrence numérique, car la donnée change la manière de penser. Un grand nombre d'observations numérisées suffit aux algorithmes d'intelligence artificielle pour avoir une description d'un phénomène et pour en déduire des propositions de régulation, sans pour autant comprendre ou expliquer comment fonctionne ce phénomène. Les systèmes de transport en commun sont par exemple de plus en plus dits *data driven* (pilotés par les données). Dans toutes les métropoles ayant un réseau de transport en commun dense, une équipe d'ingénieurs ayant une compréhension théorique, mais aussi pratique et sensible du réseau était nécessaire pour proposer une solution optimisée en matière d'horaires, de fréquences, de type de véhicules, etc. Désormais, à partir de données de fréquentation détaillées et innombrables, un algorithme est capable de faire de même, grâce à diverses opérations mathématiques (optimisation combinatoire, classification, réseaux de neurones, etc.).

Passer de la compréhension à la description est également problématique sur le plan politique. Les individus se retrouvent eux aussi réduits à leurs traces numériques, c'est-à-dire à leurs actions indépendamment de leurs intentions. Le risque totalitaire ne se limite alors pas seulement aux possibilités de fichage et de contrôle que la donnée ouvre et à ce que des régimes autoritaires pourraient en faire. Il tient désormais en partie à l'eldorado de la donnée, un fantasme d'omniscience et d'omnipotence : celui de la connaissance totale des pratiques et des attentes, et de la possibilité de les interpréter, de les contrôler et de s'y adapter en temps réel, de façon individualisée et optimisée. D'une part, cet eldorado repose sur le mythe de l'*homo œconomicus* qui raisonnerait de façon rationnelle (Rist 2010) et à qui les données permettraient de faire des choix éclairés et optimaux. De l'autre, il s'appuie sur l'illusion de l'infaillibilité des systèmes numériques.

La collecte de données permanente et généralisée est dangereuse parce qu'elle limite à leur portion congrue les espaces de dissidence et d'intimité, mais également parce qu'elle véhicule la croyance qu'il est possible de comprendre et de juger automatiquement, uniquement à partir de descriptions. Quel que soit le régime politique en place, disposer d'un système de surveillance supposé total et fondé à juger est une perspective totalitaire. Ainsi, Michel Foucault définit le panoptique comme le modèle d'une société disciplinaire dont le contrôle social serait basé sur sa capacité d'observation totale (Foucault 1975). À travers la possibilité d'avoir des traces numériques de toutes les activités humaines, la donnée rend crédible la mise en œuvre d'un e-panoptique. Les régimes de gouvernement urbain contemporains sont à bien des égards technocratiques, dans le sens où ce sont des personnes expertes qui décident des destinées collectives, au nom de leurs compétences et de leurs connaissances. La croyance dans le fait que les ordinateurs ne se trompent jamais peut conduire à une étape supplémentaire, qui verrait la prise de décisions déléguée à des systèmes numériques, entrant alors en rupture avec la notion de démocratie.

Comme l'illustre le dynamisme du salon international de la sécurité *Milipol* (qui a lieu en alternance entre Paris et Doa au Quatar), les intérêts économiques se marient parfaitement avec les tentations

totalitaires. Le développement de la reconnaissance faciale découle autant d'attentes policières que de lobbying industriel. Si les caméras de vidéosurveillance scrutent des proportions croissantes des territoires urbanisés, il est encore difficile de parler en France de surveillance généralisée, dans la mesure où les moyens humains ne permettent pas d'en exploiter le plein potentiel (Lemaire 2019). Toutefois, la donnée change les choses : des algorithmes sont désormais capables de reconnaître des individus, de détecter leurs comportements et de croiser ces informations avec d'autres données. Le système de surveillance britannique suit les automobilistes en temps réel. En Chine, le système de reconnaissance faciale va jusqu'à suivre les individus partout et relever leurs infractions. Les services de police et de gendarmerie français fantasment ouvertement sur la généralisation de la reconnaissance faciale dans l'espace public. Dans une note de 2019, un officier, chercheur au CREOGN, l'organisme de recherche de la gendarmerie, présente même cela comme un progrès démocratique : « Sous réserve d'algorithmes exempts de biais, elle pourrait mettre fin à des années de polémiques sur le contrôle au faciès puisque le contrôle d'identité serait permanent et général. » Orwellien, ce discours est surtout infondé, puisqu'il a été observé que les algorithmes reproduisent les biais, notamment racistes ou sexistes, des organisations qui les créent (Bertail *et al.* 2019).

E-panoptique et économie de la donnée vont de pair. Faute de moyens économiques ou technologiques, les pouvoirs publics confient souvent le travail de surveillance à des sociétés privées, comme *Palantir* ou *Amesys*. Ces entreprises assurent la conception, la construction, l'installation et la maintenance de caméras, micros et capteurs en tous genres. Surtout, elles développent les algorithmes d'analyse des comportements numériques, de reconnaissance faciale et d'identification utilisés par les services de police. Qu'elles cherchent à influencer les élections ou les décisions prises ou qu'elles travaillent comme sous-traitantes, les entreprises de la donnée ont donc aujourd'hui une influence politique grandissante.

Du fichage pour l'optimisation à la reprise de contrôle

Les systèmes *smartgrids* correspondent à des réseaux urbains capables de s'adapter en direct à la production et à la demande et d'absorber des énergies renouvelables intermittentes. Ils ont besoin pour fonctionner de contrôler les consommations à une fréquence très haute, jusqu'à plusieurs fois par seconde. Si l'optimisation passe par l'obtention de données personnelles, elle permet forcément un fichage plus ou moins sophistiqué. Cette sophistication dépend de la conception des systèmes, de leur régulation et des garanties techniques et politiques qui conditionnent le croisement de données hétérogènes. Qu'un fournisseur d'énergie puisse savoir quand quelqu'un allume sa télévision ou ses lampes à UV dans un placard est un danger limité si cette donnée est immédiatement anonymisée ou détruite et s'il est impossible de la croiser avec d'autres bases de données.

Aujourd'hui, ces garanties sont très faibles. En premier lieu parce que, compte tenu de leur poids économique, les multinationales du numérique sont souvent en capacité d'imposer leurs règles aux institutions publiques. Ensuite parce que les pouvoirs publics ne se donnent pas les moyens de limiter les croisements ou sont les premiers à profiter des possibilités de contrôle qu'ils ouvrent. Enfin parce que la majorité de la population n'a pas conscience des traces qu'elle laisse et des données qu'elle produit et transmet, ni des risques inhérents. Pour tendre vers une reprise individuelle et collective du contrôle sur les données, il s'agit d'abord de faciliter la prise de conscience qu'il est possible d'être observé, par qui, quand et comment, et la compréhension de ce que cela permet d'apporter ou de surveiller. Reprendre le contrôle requiert un changement de paradigme : plutôt que d'offrir ses données à des entreprises en échange de services, en rester maître suppose de choisir quand on les produit, lesquelles on produit, et quand on accepte de les prêter,

contre garanties, pour un temps et un usage limités. C'est la démarche *self-data* (voir par exemple : mesinfo.fing.org), passant par une implication plus forte, donc plus exigeante, dans ses pratiques numériques et numérisables. L'enjeu de la donnée est donc aujourd'hui moins technique que politique.

Matthieu Adam & Hervé Rivano

Voir également : austérité ; éducation ; marketing ; smart city ; réseaux techniques.

Références

Benbouzid, Bilel. 2015. « De la prévention situationnelle au predictive policing. Sociologie d'une controverse ignorée », *Champ pénal*, XII.

Berdet, Marc. 2013. *Fantasmagories du capital. L'invention de la ville–marchandise*, Paris : La Découverte.

Bertail, Patrice, Bounie, David, Clémençon, Stephan et Werlbroeck, Patrick. 2019. « Algorithmes : bias, discriminations et équité ».

Casilli, Antonio A. 2019. *En attendant les robots. Enquête sur le travail du clic*, Paris XIXe : Éditions du Seuil.

Citton, Yves (dir.). 2014. *L'économie de l'attention. Nouvel horizon du capitalisme ?*, Paris : La Découverte.

Foucault, Michel. 1975. *Surveiller et punir. Naissance de la prison*, Paris : Gallimard.

Guignard, Thomas. 2019. « Données personnelles et plateformes numériques : sophistication et concentration du marché publicitaire », *tic&société*, Vol. 13, N° 1-2, p. 43-69.

Lemaire, Élodie. 2019. *L'oeil sécuritaire. Mythes et réalités de la vidéosurveillance*, Paris : La Découverte.

Noiriel, Gérard. 2011. *Les ouvriers dans la société française : XIXe-XXe siècle*, Paris : Points.

Pariser, Eli. 2011. *The filter bubble: what the Internet is hiding from you*, New York : Penguin Press.

Rist, Gilbert. 2010. *L'économie ordinaire entre songes et mensonges*, Paris : Presses de Sciences Po.

Vraiment Vraiment. 2019. « Espace public : Google a les moyens de tout gâcher, et pas qu'à Toronto », *Medium*, URL : https://medium.com/@vvraiment/https-medium-com-vvraiment-espace-public-google-a-les-moyens-de-tout-gacher-2ab92ac11df4, consulté le 6 novembre 2019.