

## LA VALEUR LUCRATIVE DES DONNÉES EST UNE VIEILLE HISTOIRE

[Christophe Masutti](#)

Alternatives économiques | « L'Économie politique »

2022/2 N° 94 | pages 23 à 35

ISSN 1293-6146

Article disponible en ligne à l'adresse :

<https://www.cairn.info/revue-l-economie-politique-2022-2-page-23.htm>

Distribution électronique Cairn.info pour Alternatives économiques.

© Alternatives économiques. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

# La valeur lucrative des données est une vieille histoire

Par Christophe Masutti

**E**N JUIN 2013, GRÂCE AUX RÉVÉLATIONS D'EDWARD SNOWDEN et avec le soutien courageux des journalistes Glenn Greenwald et Laura Poitras, une série d'articles dans le *Guardian* mettait en lumière les programmes de surveillance mondiale des réseaux de communication. Emblématiques, les programmes Prism et XKeyscore attestaient de la connivence entre les autorités du renseignement et de la défense étasuniennes, d'une part, et les grandes entreprises internationales et monopolistiques, elles aussi étasuniennes, d'autre part.

De nombreuses publications ont alors questionné le rôle des Gafam (Google, Apple, Facebook, Amazon, Microsoft) dans la surveillance de masse et l'acquisition des données personnelles. Parmi les travaux les plus marquants figurent ceux de Shoshana Zuboff, fine observatrice des mécanismes d'analyse comportementale que des firmes comme Google et Facebook n'ont eu de cesse de perfectionner depuis le début des années 2000 <sup>[1]</sup>. Dans un article de *Monthly Review* paru durant l'été 2014, le sociologue John B. Foster et le politiste Robert McChesney replaçaient, eux, les révélations de Snowden dans l'histoire longue du capitalisme de leur pays. Forgeant l'expression

**Christophe Masutti**

Philosophe, historien des sciences et des techniques, laboratoire SAGE, Université de Strasbourg, coadministrateur de Framasoft. Auteur de *Affaires privées*.

*Aux sources du capitalisme de surveillance* (C&F Editions, 2020).

[1] Voir entretien avec Shoshana Zuboff, p. 8.



« capitalisme de surveillance », ces chercheurs voyaient dans les récentes révélations de l'ancien employé de la National Security Agency (NSA) une gigantesque démonstration de ce que signifie vraiment l'hégémonie (voir encadré ci-dessous).

### Le capitalisme de surveillance selon John Bellamy Foster et Robert McChesney

Shoshana Zuboff a popularisé l'expression « capitalisme de surveillance » avec son livre paru en 2019, *L'âge du capitalisme de surveillance* (Zulma, 2020). Elle l'emprunte (sans les citer) au sociologue John B. Foster et au politiste Robert McChesney (2014). Pour l'ancienne professeure à Harvard, il s'agit de démontrer que l'utilisation par Google et Facebook du « surplus comportemental » a perverti le capitalisme. Si elle décrit avec précision les mécanismes à l'œuvre, sa critique se borne au rôle déterminant de certaines entreprises au tournant des années 2000 (les révélations faites par Snowden sont ramenées au rôle des Gafam), et ses préconisations se limitent au vœu pieux d'une régulation du marché. Tout autre est la perspective de Foster et McChesney, pour qui les révélations de Snowden mettent en lumière de manière paradigmatique la façon dont fonctionne le capitalisme étasunien depuis les années 1960. Celui-ci repose sur l'intrication du complexe militaro-industriel, de la finance et du marketing et répond à un double objectif : asseoir un marché intérieur en créant une société de consommation grâce à toutes les possibilités que permet la surveillance et renforcer le marché extérieur

grâce à l'effort militaire (guerre de Corée, guerre du Vietnam, guerre du Golfe, etc.), qui repose lui aussi en partie sur la surveillance<sup>[4]</sup>, afin d'assurer l'hégémonie des Etats-Unis au niveau international. Contrairement à Zuboff qui confie à l'Etat le rôle de protéger le libre marché contre le monopole des Gafam afin de soigner un capitalisme « malade », Foster et McChesney utilisent le concept de capitalisme de surveillance pour souligner l'extrême faiblesse d'une telle approche face aux enjeux de pouvoir géopolitiques et aux luttes hégémoniques. Ils ont raison. Mais en considérant que la surveillance a d'abord été inventée au cœur du complexe militaro-industriel puis transposée dans les sphères civiles (la société de consommation et le marketing), ils semblent oublier que, au-delà des seules innovations, la transformation numérique des organisations, les routines et les pratiques ont été, de manière incrémentale, un moteur suffisant dans la course aux données.

[4] On peut se rapporter au livre du journaliste Duncan Campbell, *Surveillance électronique planétaire* (Allia, 2001), rédigé à partir de son rapport auprès du Parlement européen sur le thème des programmes de surveillance étasuniens (qu'il a contribué à révéler depuis les années 1980, en particulier le programme Echelon, bien avant les révélations d'Edward Snowden).

Pour riches et stimulantes que soient ces critiques, l'une arrimée aux valeurs du libéralisme, l'autre à tendance marxiste, elles font l'impasse sur l'histoire des technologies – au moins dans la méthode. Les révélations de Snowden n'ont en effet surpris que ceux qui, par naïveté sans doute, se sont laissés surprendre. L'histoire montre qu'il n'y a rien de nouveau dans le penchant inquisiteur des gouvernements et des entreprises de services numériques, avec ou sans collusion des intérêts, quel que soit le pays <sup>[2]</sup>.

Ce penchant se déploie durant la seconde moitié du XX<sup>e</sup> siècle et marque de son empreinte l'histoire de l'informatique. Contrairement au postulat de Zuboff (qui s'intéresse davantage aux Gafam qu'au capitalisme) et à celui de Foster et McChesney (peu loquaces sur cette question), les technologies numériques ne sont pas neutres, elles sont issues de choix structurels et fonctionnels dans l'économie. Elles ne sont pas de simples outils dont Google et Facebook auraient fait un usage dévoyé. Elles ne sont pas non plus des mediums comme les autres dont se serait emparée la politique économique capitaliste pour mener à bien, de manière inéluctable, son projet hégémonique de surveillance. Au sortir de la Seconde guerre mondiale, ces technologies, qui sont devenues centrales dans le fonctionnement du capitalisme contemporain, ont été façonnées par des usages et des pratiques qui sont eux-mêmes le reflet de stratégies collectives, politiques ou d'entreprise.

## La donnée, une nécessité structurelle de l'économie

L'existence de réseaux de communication au sein des entreprises et entre les entreprises ou administrations ne date pas d'internet <sup>[3]</sup>. Dès les années 1940, des moyens de communication de plus en plus sophistiqués sont développés, à l'instar du réseau Télex, ancêtre du fax, qui sert de support dans bien des domaines pendant plus de trois décennies <sup>[4]</sup>. Dès 1960-1963, en particulier grâce à l'entreprise Bell, apparaissent les premiers modems <sup>[5]</sup> commercialement distribués à des milliers d'exemplaires qui permettront de faire de l'informatique en réseau.

En parallèle, à la fin des années 1950, l'ordinateur, qui est essentiellement un serveur (avant l'invention de l'ordi-

**[2]** Pour ce qui concerne la France, on peut se rappeler de l'affaire Safari (Système automatisé pour les fichiers administratifs et le répertoire des individus) en 1974. La vive opposition qu'elle suscita fut l'un des motifs de la création de la Commission nationale de l'informatique et des libertés (Cnil) en 1978 (voir la contribution de François Pellegrini et d'Elia Verdon, p. 36).

**[3]** Même si le projet est conçu de manière expérimentale à la fin des années 1960, Internet, qui est littéralement un « réseau de réseaux », est mis en place dans les années 1980 lorsque l'on commence à choisir, normaliser et harmoniser les protocoles de communication que nous utilisons aujourd'hui.

**[4]** Certains réseaux Télèx sont encore en service au début des années 2000. Les entreprises développent aussi leurs propres réseaux puis viendront Télèx, Videotex, Transpac, et au-dessus de tout cela, Arpanet et ses protocoles, etc.

**[5]** Modulateur-démodulateur (modem) sur réseaux téléphoniques commutés (RTC) : pour convertir un signal analogique en signal numérique pour l'ordinateur, et inversement. Durant les années 1960, on retient notamment la percée de Bell avec son système Data-phone, modems Bell 101 et Bell 103.



**PENDANT LES ANNÉES 1960, LE RECOURS  
ACCURU AUX MACHINES ET L'INFORMATISATION  
QUI ACCOMPAGNE CE MOUVEMENT  
CHANGENT RADICALEMENT LES PRATIQUES**

nateur personnel dans les années 1970), sort du laboratoire expérimental et devient un produit industrialisé. Très vite, terminaux, modems et « mini-ordinateurs » <sup>[6]</sup> telle la série IBM system/360 à 370 équipent les banques, les grandes administrations publiques, l'aviation civile et les assurances. Ils y rendent de nombreux services : solutions de calcul partagé à des fins comptables, programmes sur mesure dotés de macrocommandes compréhensibles moyennant quelques heures d'apprentissage de langage de haut niveau, et enfin services à distance sur des ordinateurs plus gros (*mainframe*).

L'utilisation des machines a d'emblée pour effet d'entraîner la constitution de bases de données qui sont depuis la fin des années 1940 de plus en plus stockées sur des bandes magnétiques, remplaçant les fiches perforées en attendant l'apparition des disques durs. La question de la sécurité du stockage et de la communication de données est alors secondaire aux yeux des entreprises, malgré l'avis des sociétés d'assurances : les points les plus importants et qui guident les investissements sont plutôt la vitesse de transfert, la capacité de stockage et l'accessibilité de l'information <sup>[7]</sup>. Dans les entreprises, il devient de plus en plus courant de voir des terminaux télétypes (ou téléimprimeurs) à côté du bureau ou au guichet, même après l'arrivée des terminaux disposant d'un support vidéo. Très vite, on s'habitue à s'entendre répondre : « *Patientez un instant, j'interroge l'ordinateur.* »

Pendant les années 1960, le recours accru aux machines et l'informatisation qui accompagne ce mouvement changent radicalement les pratiques. Interrogations à distance, analyses statistiques et évaluation du marché « en temps réel » sont de plus en plus perçues comme indispensables dans le contexte très concurrentiel de la consommation. Aux Etats-Unis, les sociétés d'évaluation de crédit et les multiples agences couvrant le territoire pratiquent la centralisation des données dans le but d'évaluer les risques de solvabilité et de faire des études statistiques sur les comportements de groupes de consommateurs. Cela suppose en retour une mise en réseau de toutes

[6] Malgré leur nom, il s'agit de grosses machines qui préfigurent l'arrivée des micro-ordinateurs (avec leurs microprocesseurs). Les mini-ordinateurs sont moins chers que les ordinateurs centraux (*mainframe*), ils fonctionnent généralement en réseau avec des systèmes d'exploitation spécifiques, souvent configurés sur mesure pour l'usage en entreprise. IBM et DEC ont été les producteurs les plus en vogue fournissant le marché avec, respectivement, les gammes system/360 et 370 et PDP 8 à 11.

[7] La question de la sécurité préoccupait en revanche beaucoup l'Agence pour les projets de recherche avancée de Défense (Darpa).

ces données (un vendeur dans un magasin peut effectuer des requêtes à distance, et l'agence peut évaluer les débiteurs) <sup>[8]</sup>.

La définition de l'ordinateur – comme un système de stockage et de récupération de l'information, et de gestion de bases de données rendu pertinent par sa mise en réseau – apparaît au cours des années 1960. C'est de là que viendra ce que les historiens de l'informatique appellent la seconde révolution informatique (Cerruzi, 2003) : celle des services (*utility computer*) et de la gestion de l'information <sup>[9]</sup>.

Vers la fin de la décennie, une étape supplémentaire est franchie, celle de la généralisation de l'automatisation des processus décisionnels et de leur analyse <sup>[10]</sup>. Le marché voit apparaître plusieurs gammes de produits spécialisés dans ce domaine. Les premiers logiciels dédiés à la prise de décision en situation d'incertitude économique sont développés. Pour fonctionner, ces logiciels ont bien évidemment besoin de données. Ils recourent à des techniques de programmation en statistique, à des données qualitatives et quantitatives sur lesquelles des modèles mathématiques peuvent s'appliquer. Deux logiciels, créés au MIT en 1969, sont d'excellents exemples de cette évolution. Le premier, *Sprinter* <sup>[11]</sup>, permet de déterminer les risques d'échec d'un nouveau produit introduit sur le marché du commerce de détail et d'inférer des stratégies efficaces. Le second, *Mediac* <sup>[12]</sup>, permet de générer des segments de marché pour mettre en place une stratégie de démarchage publicitaire. Pour ce faire, il croise plusieurs programmes et bases de données afin de traiter des informations clientèle hétérogènes (état civil, habitudes médiatiques...).

Confiées en pâture aux machines, les données prennent une importance centrale dans la société de consommation des années 1960 et 1970. La majorité des décideurs des entreprises comprennent ce que représentent les bases de données et leur emploi à grande échelle à des fins de segmentation de marché, de marketing et d'influence comportementale, comme nous allons le voir dans la partie suivante.

Du reste, très vite surgissent des discussions sur la légitimité de ces bases de données. La presse et les archives

**[8]** En France, l'entreprise pionnière dans l'analyse de la consommation est Cetelem, équipée dès 1953 de machines IBM et s'inspirant du modèle étasunien. La décennie suivante, Cetelem améliore ses équipements non seulement pour gérer la croissance des dossiers de crédit mais aussi étudier et planifier l'état de la consommation en France. C'est la création du Service des études de marché de Cetelem, qui devient en 1985 L'observatoire Cetelem.

**[9]** La première correspond, schématiquement, à l'innovation technoscientifique.

**[10]** On parle aujourd'hui de *business intelligence*. Cette informatique décisionnelle est ancienne.

**[11]** *Sprinter* pour *Specification of Profits with Interdependencies*. Ce logiciel avait été développé au MIT par Glen L. Urban.

**[12]** *Mediac* pour *Media Planning Calculus*, créé lui aussi en 1969 au MIT par John D. C. Little et Leonard M. Lodish.



**« LA CAPACITÉ DE L'ORDINATEUR À STOCKER  
ET À RÉCUPÉRER D'ÉNORMES QUANTITÉS  
DE DONNÉES MENACE D'ENGLOUTIR LE DROIT  
À LA VIE PRIVÉE DE L'INDIVIDU COMME  
JAMAIS AUPARAVANT »**

ANNETTE HARRISON

des débats du Congrès des États-Unis et des différentes commissions parlementaires montrent combien ces questions figurent au premier plan politique durant toute cette révolution numérique entre 1955 et 1975. En 1967, le rapport du *think tank* Rand Corporation ne recense pas moins de 300

publications, monographies et articles portant sur la question de la vie privée à l'heure de l'informatisation des entreprises et des services publics. Il affirme : « Avec l'avènement de l'ordinateur, la

*menace pour la vie privée – individuelle et collective – a pris des proportions encore plus grandes. La capacité de l'ordinateur à stocker et à récupérer d'énormes quantités de données menace d'engloutir le droit à la vie privée de l'individu comme jamais auparavant »* (Harrison, 1967).

Quant à la protection de la vie privée (« *the right to privacy* »), une tradition se développe dans ces mêmes années aux États-Unis à l'encontre des grands projets fédéraux de bases de données et contre l'espionnage des citoyens<sup>[13]</sup>, sujet particulièrement sensible à l'époque des mobilisations pour les droits civiques ou contre la guerre du Vietnam. Les débats du Congrès précédant le vote du Privacy Act de 1974 illustrent l'intention générale d'une défense de la liberté individuelle. L'Europe, elle, est sensibilisée de manière différente, laissant plus de place au concept de dignité dans la vie privée (Whitman, 2004), entre le scandale de la revente des données sur les foyers par le Bureau du recensement britannique et l'affaire Safari<sup>[14]</sup> en France.

À la lumière de ce tableau historique, les bases de données – quelle que soit la définition qu'on en donne – s'imposent d'emblée comme un problème public relatif à la vie privée (Atten, 2013). Comme le prédisait l'un des plus grands spécialistes en droit sur ces questions, Alan Westin (1966) : « *Si la société devait suivre les possibilités de l'ingénierie technologique et sociale, il est tout à fait possible que les informations de base sur chaque aspect majeur de la vie de l'individu soient collectées dans divers systèmes à mémoire centrale.* » Et après avoir rappelé que les institutions ont toujours cherché à limiter les pouvoirs privés de surveillance à l'encontre des citoyens,

[13] On peut citer les programmes de renseignements informatisés divulgués dans la presse au tournant des années 1970, et dont l'objectif était la surveillance de masse : Cointelpro (FBI, débuté en 1956), Minaret (NSA, débuté en 1962), ConusIntel (US Defense, débuté en 1965).

[14] Voir note numéro 2.

il soulignait que ces mêmes autorités « *semblent être déconcertées par les protestations contre les utilisations actuelles ou futures des nouvelles techniques de surveillance* ».

### L'avènement du courtage de données, un tournant

Le recours croissant aux bases de données optimise la gestion d'entreprise et étend les possibilités commerciales. Au tournant des années 1970, la nouvelle frontière consiste à faire communiquer entre elles différentes bases de données – c'est la question de leur harmonisation – et à inférer ainsi des informations au départ non disponibles sur les consommateurs et les processus de production<sup>[15]</sup>. Dans le champ de la statistique, l'analyse exploratoire de données (AED) est un secteur en plein essor qui trouve grâce à l'informatique de nombreuses applications. Il s'agit de méthodes qui permettent d'étudier un ensemble de données afin de trouver leurs caractéristiques respectives, de les manipuler pour tester des hypothèses, de définir des modèles, des relations systématiques entre des variables (pour élaborer des profils de consommateurs par exemple) ou encore d'identifier des anomalies.

À la fin des années 1960, le consortium industriel Codasyl (Conference on data system language) impliqué dans l'harmonisation des bases de données et leur traitement émet les premières spécifications des bases de données réseau, celles qui permettent le traitement de liaisons et d'associations entre différentes données. Des chercheurs jouent alors un rôle décisif dans les progrès réalisés par l'AED. L'informaticien britannique Edgar Codd travaille sur les bases de données relationnelles. Par opposition au modèle hiérarchique d'un tableau de données, il se penche sur les relations entre différents tableaux et sources plus ou moins hétérogènes. Les avancées qu'il obtient en termes de modèles informatiques révolutionnent la gestion des bases de données. Un autre chercheur, le mathématicien et statisticien John Tukey, travaille sur la relation entre technique statistique et analyse exploratoire de données : la méthode statistique utilisée pour obtenir des informations sur le réel est-elle la bonne ? Comment comprenons-nous les variations entre différentes données apparemment homogènes et les relations (par exemple la causalité) entre elles ? La portée de ces travaux est elle aussi considérable.

**[15]** L'exploitation du « surplus comportemental », que Shoshana Zuboff attribue à Google (voir entretien p. 8), c'est-à-dire le fait d'inférer des informations sur un consommateur à partir de multiples données disponibles par ailleurs, est ainsi bien antérieure aux années 2000.





Les applications de cette gestion de l'information sont diverses : si la gestion administrative, la comptabilité, les systèmes de réservation et la location de temps de calcul étaient depuis près de quinze ans entrés dans les routines des services informatiques, les avancées en gestion de bases de données permettent de fournir des services beaucoup plus complets avec des savoir-faire spécifiques, par exemple le traitement de listes de clientèle et le publipostage, auxquels on adjoint d'autres services tels que les méthodes d'inférence statistique appliquées à la création et la vente de listes, ou les modèles de régression linéaire appliqués au marketing. Tant et si bien que l'AED est considérée aussi comme un outil permettant d'inférer à partir des agents économiques beaucoup plus d'informations qu'auparavant, ce qui accroît encore la quantité de données.

Soulignons que l'analyse exploratoire de données et la production d'analyses de profils n'ont pas été des innovations à proprement parler. Ces services existaient déjà et leur qualité dépendait essentiellement de la manière dont on utilisait les ordinateurs, leurs systèmes d'exploitation et la programmation bas niveau ou à partir de langages déjà anciens (Fortran, Cobol, Lisp, etc.), tout en appliquant les récentes avancées en mathématiques et statistiques. Sur ce créneau, l'essentiel n'était pas tant d'avoir une machine plus performante que les autres que de savoir s'en servir <sup>[16]</sup>.

La convergence de plusieurs technologies rend bientôt possible la mise en place de systèmes automatisés de gestion de bases de données qu'appellent de leurs vœux les décideurs d'entreprises, dans l'espoir d'obtenir sur leurs concurrents un avantage décisif. On assiste, d'une part, à l'amélioration notable des langages informatiques de haut niveau connus et éprouvés comme le Fortran et, d'autre part, au développement de modèles de plus en plus performants de stockage de données (surtout chez IBM) qui sont désormais compatibles et qui permettent de hautes vitesses de transfert (requêtes/réponses). Ces dispositifs n'ont cependant pas conditionné le nouveau secteur de la marchandisation des données. Il fallait pour cela que des entreprises en comprennent tout l'intérêt stratégique et décident d'utiliser leurs savoir-faire en gestion de bases de données à d'autres fins que la gestion interne ou le développement des processus de

**[16]** Rappelons que le nombre d'ordinateurs et de mini-ordinateurs est encore relativement bas à la fin des années 1960, bien que plusieurs entreprises (IBM, DEC, Sperry Corporation, HP, Xerox, Honeywell, Texas Instruments) se soient lancées dans ce secteur.

production et d'ingénierie. Cette prise de conscience caractérise une vague de repositionnements spectaculaire et l'invention de ce nouveau métier qu'est le courtage de données.

Plusieurs entreprises se lancent ainsi dans l'activité d'évaluation des crédits bancaires qui implique d'agréger, de croiser et bientôt de vendre des bases de données. Certaines sociétés y voient des opportunités de diversification hautement rentables en faisant jouer la convergence entre leurs capacités technologiques et l'amélioration des systèmes de stockage et de communication. L'Union Tank Car

**L'ANALYSE EXPLORATOIRE DE DONNÉES  
ET LA PRODUCTION D'ANALYSES DE PROFILS  
N'ONT PAS ÉTÉ DES INNOVATIONS  
À PROPREMENT PARLER**

Company, spécialisée dans la location de wagons et dont les équipements sont à la pointe des technologies informatiques, crée Trans Union en 1968. Pour cela, elle mène une politique offensive de rachat d'agences de crédit avec pour objectif de rassembler des millions de fiches clients. En 1972, elle crée un système d'information de stockage sur bandes et de transfert d'information depuis et vers les guichets locaux appelés Cronus (Credit Reporting Online Network Utility System). Il s'agit de permettre aux succursales (vendeurs de voitures, agences de location, bref tous les commerces qui proposent des crédits) de récupérer l'information sur le consommateur en temps réel. Autre exemple : Thompson Ramo Wooldridge Inc. (plus tard nommée TRW), spécialisée dans l'aérospatiale et les composants électroniques, rachète Credit Data Corporation en 1968 pour se lancer dans l'évaluation de crédit. Quant à la Retail Credit Company, la plus grande agence d'évaluation de crédit des Etats-Unis, elle s'informatise complètement en 1970 mais est obligée de changer d'image et de nom pour Equifax en 1975, après plus de cinq années d'enquêtes administratives sur ses pratiques de *reporting* (activité de collecte, de formatage et de traduction des données dans un format répondant à un besoin) qui portent atteinte à la vie privée.

Une autre activité, le profilage des consommateurs, s'avère hautement nécessaire au développement de ces secteurs. Le profilage est issu d'un autre type de convergence (Masutti, 2021) : l'automatisation du service de gestion de listes de clientèle sur mesure à des fins de marketing direct et le ciblage



de la clientèle à l'aide de données géodémographiques et psychographiques (les « styles de vie »).

Dans le premier domaine, l'entreprise pionnière en 1969 est Demographics (alias Acxiom, aujourd'hui leader mondial du courtage de données), qui invente en 1972 un système permettant d'automatiser le traitement de listes de clientèle, le LOFS (*List order fulfillment system*). Travaillant au départ sur le profilage électoral et le publipostage, son premier grand

client après le Parti démocrate est Direct Media Inc., l'équivalent des pages jaunes pour les Etats-Unis, qui se spécialise dans le marketing direct. Demographics embauche des ingénieurs venant

### ON ATTRIBUE SOUVENT LA CITATION SUIVANTE

AU SOCIOLOGUE JONATHAN ROBBIN :

« SI VOUS ME DITES LE CODE POSTAL  
DE QUELQU'UN... JE PEUX PRÉDIRE CE QU'IL  
MANGE, BOIT, CONDUIT ET MÊME PENSE »

d'IBM sous la houlette du directeur Charles Morgan. Profitant de leurs connaissances des machines IBM system/360 et 370, ils deviennent des experts de l'entrepôt de données (les collections de données, de véritables trésors de guerre, sont issues tant des organismes publics que des entreprises tous secteurs confondus, elles sont vendues, dupliquées et souvent accaparées par ces nouveaux acteurs) et du *data mining* (terme désignant toutes les techniques de fouille de données et d'extractions de connaissances).

Dans le second domaine, s'agissant du ciblage clientèle à l'aide de données géodémographiques et psychographiques, la société Claritas, créée en 1971 par le sociologue Jonathan Robbin, s'illustre par son savoir-faire. Grâce aux données des services publics, comme le recensement, Robbin invente une méthode d'analyse des groupes démographiques pour les diviser en un nombre limité de blocs répartis sur la base de leur code postal. Il aboutit ainsi à une quarantaine de segments géodémographiques répartis en 14 blocs, permettant une analyse fine des comportements des habitants ainsi caractérisés, l'analyse PRIZM (*Potential Rating Index for ZIP Markets*). Claritas vend ses analyses en croisant ces segments avec d'autres données psycho-sociographiques, c'est-à-dire des sous-catégories sociales et comportementales (par exemple : hommes mariés + habitant la zone postale xxx + propriétaires de leur logement + vote conservateur + possède un chien + inscrit au club de

golf, etc.). On attribue souvent la citation suivante à Jonathan Robbin : « *Si vous me dites le code postal de quelqu'un... je peux prédire ce qu'il mange, boit, conduit et même pense.* »

On a là quelques exemples des évolutions technologiques et sociales qui ont lieu dans les années 1960 et 1970. Les plates-formes des années 1990 et 2000, héritières des plates-formes des années 1970 et 1980, n'ont fait que suivre le même cheminement : ce qui change dans les années 2000, c'est l'échelle, pas le principe. Au début du XXI<sup>e</sup> siècle, quoi de plus normal qu'une jeune start-up comme Google ait fini par succomber à la pression de l'actionnariat pour traiter les données personnelles des utilisateurs de son service de moteur de recherche comme autant de jetons à faire valoir sur le marché publicitaire ? L'histoire de l'exploitation des bases de données ne se réduit pas à une innovation qui aurait mal tourné pour finalement donner naissance à un capitalisme immoral qui dicte nos conduites et nous « dépossède » de notre vie privée, même si ce résultat est bien présent. Il s'agit d'un virage pris par l'informatisation des organisations dès lors que l'industrie informatique a trouvé des débouchés pour rentabiliser la production d'information et en automatiser le traitement. En matière de profilage et d'influence, il aura certes fallu attendre l'existence de solutions de stockage performantes et les premiers travaux sur la normalisation des bases de données. Cependant, si des savoir-faire sont nés en matière de *data mining*, c'est parce que des besoins ont été créés dans une économie de consommation qui ne se satisfaisait plus de la rationalité supposée de ses agents et misait tout sur les capacités de surveillance et d'influence induites par l'exploitation de ce nouveau pétrole que sont les données numériques.

En nous plaçant sur le temps long de la seconde moitié du XX<sup>e</sup> siècle, on peut dès lors définir à notre tour le capitalisme de surveillance comme un mode d'être du capitalisme, dont la caractéristique est de reposer sur l'activité de surveillance entendue au sens large, c'est-à-dire comme la surveillance des processus de production, des configurations du marché et des fonctions économiques comme la consommation. En ce sens, il est fortement dépendant de la technologie. Dans cette histoire, le capitalisme de surveillance n'est toutefois



**[17]** Une telle évolution est inséparable du système capitaliste, de plus en plus caractérisé par la recherche de la satisfaction des besoins personnels et l'injonction au développement individuel plus qu'au bien-être collectif (Muhlmann, 2021), qui réduit les êtres humains à une somme de données extractibles. Pour reprendre la formule de J. B. Foster et R. McChesney, chaque personne est « placée dans des catégories relatives à son style de vie, inférées de sa classe, de ses habitudes de dépense et de son emplacement géographique ».

**[18]** Dès le début de son autobiographie, le directeur d'Acxiom y fait lui-même référence (voir Morgan, 2015).

**[19]** Le CRM (*customer relationship management*) *onboarding*, qu'on peut décrire comme du reprofilage, consiste à retrouver et prospecter des clients sans qu'ils soient pour autant identifiés (sur un site Internet, par exemple) en comparant leur profil (comme leur comportement) à une base clientèle préexistante (les données « matchent » entre elles). La résolution d'identité consiste à trouver un faisceau suffisant de correspondances entre des bases de données hétérogènes, afin d'identifier exactement un client et de reconstruire son profil.

pas déterminé par la technologie mais par le développement des pratiques qui ont saisi l'intérêt de la donnée numérique et du *monitoring* social en termes de rentabilité et de profit.

Cette conception de la donnée numérique comme un capital a façonné les usages des technologies numériques, en orientant une grande part de leurs développements comme un « travail de la donnée » à visée commerciale, un levier d'optimisation de la production par le marketing, et comme modèle d'affaires. Cela a significativement changé la représentation de la société du point de vue des fonctions économiques. Depuis les années 1970, le profilage commercial a profondément redéfini la notion de communauté dans la structure sociale étasunienne : l'idée que les groupes sociaux ont surtout en commun leurs niveaux de revenus et leurs comportements de consommateurs s'est imposée comme repère essentiel de la construction de l'image de soi (Reich, 1992) <sup>[17]</sup>.

Ce tableau se situe bien sûr assez loin de l'idéal d'émancipation sociale des premières décennies de l'informatique. Il entre en confrontation directe avec le *story telling* des grandes figures de la Silicon Valley. Croire aujourd'hui en une forme de maturité des débats serait présomptueux. En Europe, beaucoup s'enorgueillissent du Règlement général sur la protection des données (RGPD) : il n'est pas vraiment sérieux de prétendre que les limites du capitalisme de surveillance se bornent à celles du consentement individuel. Les entreprises multinationales comme Acxiom (qui a récemment fusionné avec Liveramp) ou Nielsen (qui a racheté Claritas) disposent de près de cinquante ans de données et d'expérience en analyse marketing. Leur collaboration active avec des autorités étasuniennes dans la surveillance de masse au nom de luttes contre le terrorisme n'est un secret pour personne <sup>[18]</sup>. Ces sociétés deviennent de plus en plus spécialistes du reprofilage (*onboarding*) et de la résolution d'identité (*identity resolution*) <sup>[19]</sup>. Associées aux entreprises du Web, elles transforment les structures de nos échanges numériques en les réduisant au jeu de l'influence, de l'accumulation et du profit. D'autres choix sont possibles, basés sur les logiciels libres, sur des technologies émancipatrices

comme la Fédiverse (Mansoux et Abbing, 2021), sur des protocoles et des algorithmes permettant des communications égalitaires et sécurisées à disposition de tous, sur la neutralité des échanges et le respect de nos intimités numériques. Plus encore que nos vies privées, le plus grand larcin du capitalisme, c'est de nous avoir privés de notre capacité à contribuer aux technologies que nous voulons et à décider du sens que nous leur donnons. ■

## Bibliographie

**Atten M., 2013**, « Ce que les bases de données font à la vie privée. L'émergence d'un problème public dans l'Amérique des années 1960 », *Réseaux* n° 178.

**Campbell D., 2001**, *Surveillance électronique planétaire*, Allia.

**Castagnino F., 2018**, « Critique des surveillances studies : éléments pour une sociologie de la surveillance », *Déviance et société* n° 42.

**Ceruzzi P., 2003**, *A History of Modern Computing*, MIT Press.

**Harrison A. (dir.), 1967**, *The Problem of Privacy in the Computer Age: An Annotated Bibliography*, The Rand Corporation.

**Mansoux A. et Abbing R. R., 2021**, « Sept thèses sur la Fédiverse et le devenir du logiciel libre », trad. Framablog, disponible en ligne sur <https://framablog.org>

**Masutti C., 2021**, « En passant par l'Arkansas. Ordinateurs, politique et marketing au tournant des années 1970 », *Zilsel* n° 9.

**McChesney R. W., 2013**, *Digital Disconnect. How Capitalism Is*

*Turning the Internet Against Democracy*, The New Press.

**Morgan C., 2015**, *Matters of Life and Data. A Memoir. The Remarkable Journey of a Big Data Visionary Whose Work Impacted Millions (including you)*, Morgan James Publishing.

**Muhlmann D., 2021**, *Capitalisme et colonisation mentale*, PUF.

**Reich R., 1992**, *The Work of Nations: Preparing Ourselves for 21<sup>st</sup> Century Capitalism*, Vintage.

**Turner F., 2012**, *Aux sources de l'utopie numérique. De la contre-culture à la cyberculture. Stewart Brand, un homme d'influence*, C&F Editions.

**Westin A. F., 1966**, « Science, Privacy, and Freedom: Issues and Proposals for the 1970's. Part I – The Current Impact of Surveillance on Privacy », *Columbia Law Review* n° 66.

**Whitman J., 2004**, « *The Two Western Cultures of Privacy: Dignity v. Liberty* », *The Yale Law Journal* n° 113.