

Big data et smart cities : perspectives et problèmes

Sommaire

Le *big data* : un phénomène à la fois technique, cognitif, économique, social et politique

1. Un savoir de type nouveau

2. Un puissant facteur de mutation économique et sociale

3. *Big data* et régulation

a. la question de la protection des données

b. la question de la prédiction des comportements

4. *Big data* et gestion démocratique

a. l'*open data*

b. les *smart cities*

En guise de conclusion...

Bibliographie

Le *big data* : un phénomène à la fois technique, cognitif, économique, social et politique

La masse des « données » générées de multiples manières (par la géolocalisation, par les objets connectés, par la navigation sur Internet, par l'utilisation des smartphones et celle des cartes de crédit, etc.) est aujourd'hui gigantesque. L'exploitation de ces données ouvre une nouvelle époque. Même si le processus a commencé bien avant la digitalisation contemporaine, il représente un changement considérable aussi bien au niveau technique et cognitif, qu'au niveau économique, social et politique.

Le *big data* existait en effet avant la *digitalisation* (c'est-à-dire la transformation de l'information analogique en code binaire, en une série de zéros et un, et le traitement de cette information par l'ordinateur). Cette dernière ne fait que démultiplier considérablement la puissance de la *datafication*, comme on dit en anglais, qui consiste à collecter, à stocker et à exploiter des données pour en tirer une information. Au début du XIXe siècle, déjà, un statisticien de la marine américaine avait considérablement raccourci les routes maritimes en demandant à chaque navire de reporter sur le carnet de bord tous les événements météorologiques rencontrés, en notant l'heure exacte et le lieu. Tous ces renseignements recueillis pendant des années – l'équivalent d'un GPS d'aujourd'hui, en moins précis – ont permis de dresser des tableaux de probabilité des événements météorologiques et de définir de nouvelles routes en fonction de ces probabilités. C'est un peu l'ancêtre d'OpenStreetMap (une base cartographique en *open source*, construite par plus d'un million d'internautes, qui est la propriété d'une fondation, à la différence de Google Maps).

Le *big data* ne date donc pas d'hier, mais il est clair qu'aujourd'hui, grâce aux capacités de stockage et de calcul, aux nouveaux outils pour recueillir et traiter les données, nous sommes en train de passer à une autre échelle, pas seulement quantitative mais aussi qualitative. À la différence des périodes antérieures, en effet, la collecte et l'exploitation des données ne sont plus seulement le fait des États mais aussi de grandes compagnies (Google, Amazon, Apple, etc.), d'entreprises spécialisées et, de plus en plus, d'organisations de la société civile, voire de simples citoyens.

Un exemple permettra de mesurer l'ampleur de ce changement qu'est le *big data*. En 2009, une épidémie de grippe frappe les États-Unis. Des chercheurs décident d'utiliser les recherches des internautes sur Google pour dresser en temps réel la carte de l'épidémie, de manière à disposer d'informations beaucoup plus rapides et fiables que celles qui étaient recueillies par les services sanitaires (qui se fondaient sur des rapports faits par les médecins, les hôpitaux, etc. et dressaient des cartes sur cette base, disponibles une semaine après la saisie des données). Ces chercheurs créent un algorithme avec 45 mots clés (par exemple « médicament pour toux et fièvre », etc.), le font fonctionner sur des données anciennes (épidémies de 2007 et 2008), le comparent aux données des services sanitaires et s'aperçoivent qu'ils obtiennent des cartes beaucoup plus précises que celles des services en question. Ils ont complètement changé de méthode : plus de rapports des médecins, plus de travail des statisticiens sur des échantillons, mais une analyse d'une masse immense de données (les clics sur des sites *via* Google).

À partir de cet exemple, les auteurs d'une des meilleures synthèses sur le *big data*, Mayer-Schonberger et Cukier, le définissent ainsi : « *big data* renvoie à des choses que nous pouvons faire à une très grande échelle, que nous ne pouvons pas faire à une échelle moindre : extraire de nouvelles informations, créer de la valeur d'une nouvelle manière, avec pour conséquence que cela change les marchés, les organisations, les relations entre les citoyens et les gouvernements, et bien d'autres choses encore ».

J'aborderai successivement les changements d'ordre cognitif (intellectuel), économique et social, et enfin politique. La perspective qui me guidera est la suivante : il me semble qu'il faut éviter les approches enthousiastes et inévitablement vagues (« changement de civilisation », « nouvelle ère de l'humanité », comme dit Gilles Babinet – ce qui ne veut strictement rien dire) tout comme les approches technophobes (la menace du Big Brother) mais analyser sobrement les opportunités offertes et les nouveaux problèmes qui se posent. Les immenses gains de productivité permis par le *big data*, en particulier dans les services, sont bien connus (deux exemples : Ushahidi, une plate-forme, née en Afrique, permet de signaler tous les lieux dangereux à éviter, quasiment en temps réel, ce qui est d'une grande utilité en Afrique ; certains cours en ligne – Moocs – permettent l'individualisation des étudiants grâce au *feed-back* des données, qui aide à repérer ceux qui sont en difficulté, d'adapter les exercices à leur situation, etc.). L'exploitation des données, on le sait, touche tous les domaines de l'économie et de la société : l'*e-business* bien sûr, avec le ciblage individuel du marketing ; mais aussi la production, distribution et consommation de l'énergie, par exemple grâce aux

capteurs connectés (nous n'aurions pas la moindre idée du réchauffement climatique sans la récolte d'un très grand nombre de données) ; la vie des entreprises, avec les effets sur les processus de production, sur la logistique (gestion des stocks, etc.) et l'amélioration de l'organisation interne (l'analyse des données en interne fait souvent tomber des baronnies, en mettant en évidence l'inefficacité de tel ou tel département) ; les transports, bien évidemment ; la sécurité ; mais aussi la politique et la citoyenneté *via* les plateformes d'*open data* ; enfin la gestion par les individus de leur santé, de leur nutrition et de leur activité, avec ce que l'on nomme le *quantified self*, l'image de soi construite à partir de multiples données, par exemple celles fournies par les *wearable*, c'est-à-dire les smartphones, montres ou vêtements connectés (cf. l'application *Santé* d'Apple), etc.

Mais ces succès ne doivent pas nous conduire à occulter les nombreux problèmes de tous ordres que fait naître le *big data*, qui appelle une régulation permettant d'en tirer tous les avantages, pour les individus et les collectivités, tout en maîtrisant les inévitables effets moins désirables.

* * *

1. Un savoir d'un type nouveau

Le *big data* représente un grand changement cognitif. On a souvent insisté sur la diffusion et la démocratisation de l'information et de la connaissance, comme conséquences de la digitalisation, mais beaucoup moins sur le mode de production de celles-ci. Je voudrais en dire quelques mots.

Mayer-Schonberger et Cukier s'attachent particulièrement à montrer le changement intellectuel que représente le *big data* par rapport à la science classique. Celle-ci cherchait à établir des *relations de causalité*, alors que le *big data* ne s'intéresse le plus souvent qu'à des *corrélations*, qui peuvent laisser dans l'ombre les relations causales, même si parfois, mais pas toujours, on imagine assez facilement celles-ci. On comprend aisément pourquoi on peut

compter les automobilistes sur une autoroute simplement en géolocalisant les smartphones, ou compter précisément, par la même méthode, le nombre de participants à une manifestation. Mais, dans d'autres cas, le lien avec la causalité est réel mais reste mystérieux. On a découvert, aux États-Unis, en exploitant un grand nombre de données, que les voitures d'occasion de couleur orange étaient plus fiables que les autres... Les raisons sont inconnues et peuvent être multiples (lien entre couleur orange et marque plus fiable ? lien entre couleur orange et clients plus fortunés entretenant mieux leur voiture ? lien entre la couleur orange et clients valorisant davantage leur voiture et plus soigneux ? voitures orange plus visibles sur la route et ayant eu moins d'accidents ?, etc.). Peu importe... L'essentiel est que la corrélation existe, elle est très utile aux futurs acheteurs et aux vendeurs de voitures d'occasion. Le *big data* ne cherche à produire qu'une seule chose : des probabilités (probabilité qu'un mail soit du spam, que les trois lettres « lse » frappées dans un traitement de texte soient en fait « les », que la trajectoire du piéton et sa vitesse lui permettent d'échapper à la voiture qui s'avance vers lui, que le prix d'achat d'un billet soit moins cher à tel moment, etc.).

La recherche des causalités, soutiennent les auteurs, était le moyen le plus efficace pour extraire de l'information dans un environnement avec peu de données, mais elle n'est plus du tout adaptée lorsqu'on possède des masses gigantesques de données : « *big data* s'intéresse au comment, non au pourquoi ». Des *hedge funds* peuvent investir en se fiant à l'analyse de millions de tweets, par exemple. Cela ne fait pas disparaître la science et les experts, ceux qui recherchent la causalité, mais relativise considérablement leur pouvoir, dans de nombreux domaines (management, administration, éducation, etc.). L'*expérience*, dans un tel contexte, est fortement dévalorisée, alors qu'elle était au centre de la science et de la technique classique. L'expérience était liée au maniement de données peu nombreuses, entre lesquelles il fallait établir des relations de causalité, elle n'est plus nécessaire avec le *big data* (d'où l'âge des créateurs de *start-up* de *big data*, qui sont souvent très jeunes, encore étudiants parfois). C'est le savoir mathématique des constructeurs d'algorithme et l'ingéniosité des informaticiens qui prend le relais, ainsi que la foule des profanes (*crowd*) qui collaborent avec les experts.

L'innovation méthodologique du *big data* apparaît très bien dans le cas de *Google traduction* : Google, pour mettre au point ce service, a rompu avec la méthode traditionnelle de construction des machines à traduire (qui faisaient appel à des linguistes pour construire des algorithmes où figuraient les règles syntaxiques). Google est parti sur un tout autre principe : après avoir scanné des millions de livres dans toutes les langues, il a construit un algorithme

qui compare des millions de livres anglais et leur traduction française, par exemple, et qui propose une traduction approximative des phrases que l'internaute veut traduire (approximative, mais bien meilleure que toute autre machine à traduire reposant sur les anciens principes).

Le *big data* est ainsi la meilleure preuve que la force d'une société réside dans la diversité des talents, non dans l'intelligence de quelques individus. La première source de l'innovation vient de la mise en commun des données et des idées. La science contemporaine, par exemple, utilise massivement l'exploitation des données : la mise en évidence du boson de Higgs n'aurait pas été possible, au CERN, sans l'exploitation de données relatives à 10 milliards de collisions (10 gigaoctets par seconde y sont recueillis !). Le *big data* peut même, au-delà, on le sait, rendre possible une démocratisation de la science, avec le mouvement de l'*open science*, qui permet de publier et de croiser des résultats en provenance de tous les coins de la planète (le CERN a proposé au public de suggérer l'amélioration de certains algorithmes qu'il utilise : évidemment, chacun ne peut pas faire cela, mais c'est une ouverture importante de la science sur la société, en direction des *data scientists*, qui ne sont pas des physiciens). La science devient plus collaborative, elle sort des laboratoires... Elle laisse une place aux « amateurs éclairés »¹.

Il est clair aussi que le *big data* fait reculer, grâce aux données quantitatives, les rumeurs et toutes les affirmations sans le moindre fondement objectif (par exemple dans le domaine politique). La prudence et l'expérimentation sont souvent de règle dans le *big data*. Il existe des laboratoires, les *Living Labs*, qui étudient au sein d'une population réduite le partage des données et les effets de celui-ci. Orange, en Côte d'Ivoire, procède à des expérimentations pour rendre les systèmes des transports et télécommunications plus efficaces, le système de santé plus robuste, pour identifier les zones à risque, en recueillant un très grand nombre de données anonymisées fournies par les téléphones portables. Les expériences sont ensuite généralisées ou non.

Il n'est donc pas question de sous-estimer le changement, et le progrès cognitif, que représente l'exploitation des données. Le *big data* nous contraint à une véritable mutation intellectuelle, par exemple à un autre rapport à l'exactitude. Parce qu'il traite des données en

¹ Cf. Patrice Flichy : *Le sacre de l'amateur. Sociologie des passions à l'heure numérique* (Seuil, Kindle Edition, 2014). Un exemple d'articulation du travail des experts et des amateurs, dans le domaine des biotechnologies : *La Paillasse* <http://lapaillasse.org/>

très grand nombre, il ne recherche plus l'exactitude, alors que celle-ci était un idéal de la science classique. Quand vous disposez d'une masse gigantesque de données, peu importe l'inexactitude des résultats : il peut y avoir de très nombreuses erreurs, mais elles sont finalement submergées par les résultats corrects. Un dernier exemple de ce changement intellectuel considérable que représente le *big data* : celui de l'évolution de la taxinomie, c'est-à-dire de la classification hiérarchique traditionnelle des données (qui étaient rangées dans des catégories hiérarchisées, comme dans les bibliothèques d'autrefois). La technique des multiples *tags* et *métadonnées* qu'on attache à une donnée a remplacé la hiérarchie. Cette technique permet une recherche beaucoup plus rapide et efficace des données et une exploitation de celles-ci, elle permet de modifier sans cesse la classification, de faire émerger, lors de recherches sur la base de données, de nouveaux *tags* qui seront attachés à la donnée, etc.

Tous ces acquis cognitifs sont peu contestables. Mais il importe aussi de prendre du recul par rapport à ces innovations intellectuelles. Toute avancée scientifique et technique, on le sait, suscite un engouement, des généralisations hâtives qui érigent souvent une méthode en « nouveau paradigme »... Dans le cas du *big data*, de nombreux commentateurs attirent l'attention sur le fait que la corrélation sans relation causale peut susciter une sorte de fétichisation de la donnée. Les épistémologues connaissent bien cela dans le cas de la science empirique préscientifique du passé. Si vous avez mis en évidence une corrélation, par exemple que telle plante (qui contient ce que l'on appelle aujourd'hui de l'aspirine) est un antalgique, sans pour autant comprendre la relation causale entre la plante et la disparition du mal de tête, vous pouvez être tenté d'établir une sorte de relation magique entre les deux. Le savoir de la Renaissance, par exemple, établissait des tableaux de correspondance entre de multiples réalités (astres, plantes, animaux, parties du corps humain, événements de la vie humaine, etc.) où entraient ces corrélations de la médecine empirique, mais aussi quantité de corrélations fantaisistes (entre les astres et les événements de la vie humaine, par exemple). Et le succès pratique réel ou imaginé, lorsqu'il y a une corrélation sans relation causale, peut conduire à un dogmatisme intellectuel qui transforme cette corrélation en vérité absolue, d'autant plus absolue qu'elle n'est pas comprise et qu'elle peut être assez fragile (il peut toujours y avoir une variable cachée qui est la véritable cause de la corrélation entre deux phénomènes).

C'est un peu ce à quoi nous risquons d'assister avec le *big data*, si nous ne prenons pas de recul par rapport à cette révolution intellectuelle. Mayer-Schonberger et Cukier ne sont pas tendres avec certaines utilisations du *big data*, avec ceux qui tombent sous le charme d'une

méthode qu'ils fétichisent, qui les dispense de réfléchir, de faire preuve de jugement et d'intuition. Ils rappellent que Steve Job a conçu et amélioré les produits Apple sur la base de rapports nombreux constitués à partir de multiples analyses de données, mais que son acte créateur ne pouvait se déduire de ces données... Cette fétichisation des données peut conduire à accorder une confiance absolue aux décisions automatisées résultant du *big data*, et par là aboutir à des décisions erronées irréversibles et à se priver de certaines innovations – avec le risque supplémentaire de confier le pouvoir à ceux qui extraient cette information...

C'est pourquoi il est important de rappeler qu'aucune décision ne peut être prise exclusivement par un robot sur la base d'une étude de données (sauf, bien sûr, lorsqu'il s'agit de processus automatisés à un niveau qui ne peut pas relever d'une décision humaine). On ne peut pas se fier aveuglément aux machines, il est nécessaire de juger les décisions produites par une étude de données avant de donner le feu vert à celles-ci (ce qui, on le verra, revient à dire, lorsqu'il s'agit des *smart cities*, que la politique ne peut être écartée par les machines et ceux qui les font fonctionner). Cela est d'autant plus nécessaire qu'il y a un risque inhérent au *big data* : que les bases de données et les algorithmes deviennent des boîtes noires qui produisent des résultats parfois vrais, parfois erronés, mais toujours sans aucune responsabilité car sans aucune possibilité de traçabilité de la corrélation (le chemin des données au résultat ne peut être retrouvé). Il est extrêmement difficile, dans la plupart des cas, de savoir après coup comment une corrélation a été établie. Pourquoi *Google traduction* propose-t-il de traduire « light » par « léger » plutôt que par « lumière », à tel moment ? Personne ne peut le savoir. Il en résulte une opacité propre au *big data*, qui est nouvelle (alors qu'on peut savoir pourquoi un ordinateur effectue tel calcul erroné, par exemple, parce qu'il y a une erreur dans le programme, ou une donnée mal entrée ou interprétée, etc., cela n'est plus possible quand un algorithme utilise des données de très grande taille). On le voit, la libération des données et l'exploitation des données ne sont pas nécessairement synonymes de transparence, comme on le répète un peu trop vite.

Il est donc nécessaire de réintroduire une responsabilité humaine pour que les décisions ne soient pas aveuglément confiées aux algorithmes, de ne pas oublier le facteur humain dans la décision, avec tout ce qui le caractérise : « l'instinct, la prise de risque, l'accident, l'erreur » (Mayer-Shonberger et Cukier). Car il est illusoire de croire que la prise de décision automatisée éliminerait risques et erreurs. Il faut souvent redonner sa place à « l'intuition, au bon sens, à la *serenpidity* » (un mot anglais que l'on pourrait traduire par « le fait de réaliser une découverte scientifique ou une invention technique de façon inattendue, accidentelle, à la suite d'un

concours de circonstances fortuit et très souvent dans le cadre d'une recherche concernant un autre sujet »). Car ces facultés humaines peuvent révéler ce que les machines ne peuvent pas révéler : non pas le « what is » (la corrélation probable) mais le « what is not » (la corrélation nouvelle inventée, improbable, à laquelle personne n'avait jamais pensé, comme Edison inventant la lampe à incandescence – qui consiste à éclairer en empêchant un fil de tungstène de brûler – en rompant avec une pratique ancestrale consistant à éclairer par la combustion d'une matière). Ces facultés humaines produisent « le non-dit et le non-encore pensé » : « l'étincelle de l'invention fait apparaître ce que les données ne disent pas » (*Ibid*). Si Henry Ford avait questionné les algorithmes du *big data* avant de commercialiser la Ford T, il aurait sans doute eu comme réponse que les Américains souhaitaient seulement « un cheval plus rapide ».

La prise de décision « probabiliste » cherche donc aujourd'hui à éviter deux obstacles : la prise de décision au pied-levé, basée sur l'humeur du moment, mais aussi la décision totalement automatisée. Une étude passionnante montre que dans le domaine du sport, où les données sont de plus en plus essentielles, certains sportifs « se reposent trop sur les nombres sans réellement comprendre ce qu'ils signifient ni le contexte dans lequel ils devraient être interprétés »². L'art de coacher reste un art, même si les données sont très utiles (elles peuvent aider à prévenir les blessures, mais difficilement à prévoir les performances). D'autres études insistent également sur le danger d'une prise de décision sans le recul nécessaire, d'un raccourcissement du cycle de décision qui se paie, finalement, par une efficacité moindre. On voit très bien cela avec la prise de décision politique qui a les yeux fixés, en temps réel, sur les réactions de l'électorat (établies à partir des données issues des réseaux sociaux, par exemple). Un spécialiste des villes intelligentes écrit que le *big data* est un « moyen d'optimiser les systèmes, de réduire les coûts, de faciliter le quotidien », mais qu'il « n'aura pas d'effet sur les mauvais choix urbanistiques ou les erreurs de gestion »³. Nous y reviendrons. En d'autres termes, il faut éviter le danger de ce que l'on a nommé de manière humoristique la *quantophrénie* : « une recherche guidée par les données sans avoir de cadre conceptuel et théorique sous-jacent »⁴. Ceci parce que les données ne sont jamais brutes, qu'elles décrivent le réel par une multitude d'entrées, et que le résultat peut-être incomplet, partiel et partial,

² *Big, fast & open data...*, page 41.

³ *Ibid.*, page 104.

⁴ *Ibid.*, page 176.

grossier, incertain – ce qui ne signifie pas que l’analyse des données soit à rejeter... Un exemple : si nous avions le droit d’accéder aux données sur la population carcérale en fonction des origines ethniques (la loi française ne l’autorise pas) nous verrions que certains groupes ethniques sont davantage représentés. Mais ce constat ne fait qu’ouvrir le travail d’interprétation. Celui qui en déduirait que telle ethnie est naturellement plus violente qu’une autre oublierait d’autres hypothèses : la partialité possible des juges, l’activité de la police plus développée en direction de telle ethnie, les différences de formes de délinquance selon les groupes sociaux, le Code pénal plus sévère pour certaines délinquances que pour d’autres, la répartition des groupes ethniques pas uniformes, etc.⁵

Un autre risque cognitif, apparenté au précédent, est pointé depuis longtemps en sciences humaines : celui du behaviourisme. Il consiste à croire que l’on peut comprendre les individus uniquement à partir de leurs comportements objectifs observables, sans faire le détour par leurs motivations (croyances et désirs). Ces motivations sont certes reflétées dans les données, mais de manière très partielle, d’où des erreurs d’interprétation, lorsque l’on construit le système de motivations seulement à partir des actions observables. Pour ne prendre qu’un exemple, peut-on seulement mesurer le sentiment de tranquillité publique, comme on le fait dans certaines villes, par la vitesse de la marche des piétons dans ces villes (information que les smartphones peuvent nous fournir avec une grande précision, en mode anonymisé) ?

C’est pourquoi Adam Greenfield, dans *Against the smart city* (un livre par ailleurs souvent très contestable) a raison de dénoncer l’idée positiviste selon laquelle « the data is the data », c’est-à-dire l’idée selon laquelle les données fournissent un « savoir parfait », qui n’a pas besoin d’interprétation, « transcendant, limpide et exempt de tous les défauts inhérents à la faiblesse humaine »⁶. Il montre en particulier qu’il y a mille manières de manipuler les données, par exemple en mettant les capteurs mesurant la pollution à telle hauteur, en rangeant les données de criminalité dans telles ou telles catégories pour les faire ressortir ou les masquer, etc. Les données ne sont ni neutres ni scientifiques. Certains projets de *smart cities* supposent trop rapidement que les données sont neutres, parce qu’ils présupposent que les fins des citoyens sont les mêmes, alors que les fins, dans les villes modernes complexes et diverses, sont

⁵ Cf. l’excellente analyse d’Henri Verdier, *Ibid.*, page 183.

⁶ Page 481.

très différentes et doivent être établies démocratiquement (voulons-nous des caméras contrôlant les rues ? etc.). Nous y reviendrons dans la dernière partie, à propos des *smart cities*.

Il semble donc nécessaire de trouver un équilibre entre l'intervention humaine et la robotisation. L'aide au diagnostic, en médecine, est devenue indispensable, étant donné la masse des données à prendre en compte pour un bon diagnostic, mais ce n'est qu'une aide au diagnostic.

2. Un puissant facteur de mutation économique et sociale

Je l'ai dit en introduction, le *big data* transforme l'économie existante. Gilles Babinet, entre autres, insiste sur la révolution qu'il représente dans la production elle-même, *via le feedback* que rend possible la connexion des objets, qui informent le producteur sur l'utilisation du produit. De plus en plus, « le service devient donc une partie intrinsèque du produit ».

Mais de nombreuses études attirent aussi l'attention sur les conséquences économiques et sociales de ce que l'on nomme, en anglais, « on-demand economy », un type d'économie qui est en train d'émerger, très rapidement, dans certains pays, en particulier aux États-Unis. Les entreprises de cette économie à la demande et leur modèle économique commencent à être connus : Handy, *via* une application de smartphone, met en relation clients et artisans *freelance*; Fancy Hands met en relation clients et agents *freelance* faisant les démarches auprès des administrations ou entreprises (pour avoir le câble, par exemple) ; Spoon Rocket vous fait livrer des repas de qualité fabriqués par des individus *freelance* spécialisés, etc. ; Uber assure, avec Lyft, près 50% des courses en taxi à San Francisco ; le *crowdfunding* commence à modifier même les métiers de la banque.... D'innombrables services (traducteurs, développeurs, médecins, etc.) relèvent de cette économie à la demande, dont le principe est de mettre en relation des clients et des prestataires de service *freelance*, souvent individuels, pour utiliser à plein les compétences non exploitées et le temps libre disponible dans nos sociétés. Ces intermédiaires ont construit un modèle très efficace, pour garantir la qualité des services (contrôle par notation par le client, assurances en cas de malfaçon) et se prémunir contre les

retards de paiement ou le non-paiement du côté des clients... La sélection des personnes offrant les services est souvent très sévère, et les clients indécents sont vite éliminés. Ce nouveau type de travail a été évidemment rendu possible par la digitalisation et la *datafication* (bases de données immenses, algorithmes de recherche, etc.) que permettent les outils informatiques contemporains (sites Internet, applications de smartphone, messageries instantanées, etc.). Les chiffres mettent en évidence la progression très rapide de ce type d'économie partout dans le monde (aux USA, 4,5 milliards de dollars ont été investis dans ces entreprises dans les cinq dernières années).

Il s'agit d'une rupture importante par rapport à l'organisation du travail que nous connaissons. De plus en plus, l'innovation se fait transversalement (« open-innovation », « innovation insolente », « inventeurs candides ») : un spécialiste d'un domaine (nanotechnologie, par exemple) invente un procédé dans un domaine qui n'est pas le sien (industrie pétrolière, par exemple : comment séparer l'eau et le pétrole ?). Potentiellement, il y a là une remise en cause des experts des grandes entreprises, mais aussi des centres de recherche gérés par l'État... Ces changements s'accompagnent souvent, au niveau du discours, de ce que l'on pourrait nommer un « populisme du savoir » : les journalistes, par exemple, seraient condamnés à être remplacés par les citoyens s'informant eux-mêmes, etc.⁷ Mais au-delà de ces excès verbaux, il y a une véritable opportunité, celle d'une appropriation des métiers et des compétences traditionnelles par les citoyens.

Les grandes entreprises font de plus en plus partie des utilisateurs. Talent Group, par exemple, loue des cadres de haut niveau pour les entreprises voulant monter un projet sans pour autant embaucher un nouveau cadre. Colgate a lancé récemment un concours pour demander aux internautes de construire une vidéo publicitaire de 30 secondes, qu'elle a payée 17 000 \$, beaucoup moins cher que si elle avait fait appel à un publicitaire. Amazon's Mechanical Turk fournit quantité de services aux entreprises de cette manière (construction de sites, écriture de textes, repérage et marquage de contenus illicites) : les travailleurs *freelance* se connectent au site et choisissent le travail qui leur convient en fonction de ce qui est demandé et du prix (il peut y avoir des enchères). Les avantages pour les entreprises sont évidents : pas de locaux, pas de contrats à plein temps ou durée limitée, pas d'ordinateurs à entretenir... La loi de Coarse joue

⁷ Voir le n° 776-777 de *Critique* (2012).

à plein : les entreprises ont intérêt à faire appel aux marchés de ce genre quand cela revient moins cher que d'organiser ces services en interne.

Ce nouveau type de travail marque la fin d'un certain modèle de travail et d'un certain modèle social. L'emploi à plein temps dans une entreprise s'est développé au XIXe siècle et est resté en vigueur jusqu'aux années 80 du siècle dernier, date à laquelle il a commencé à se déliter. On a assisté alors à ce qui ressemble, à première vue, à un retour en arrière. On sait que les grandes entreprises capitalistes, à la fin du XVIIIe, s'étaient constituées en intégrant des artisans indépendants, qui ont très vite perdu leur statut pour devenir des ouvriers. Le mouvement actuel semble un retour à ce statut. Mais il y a une grande différence entre l'artisan ancien et le *freelance* d'aujourd'hui, malgré ce que disent les détracteurs de ce modèle (qui parlent d'un retour aux « bras » d'autrefois, par exemple les dockers, employés à la journée) ne serait-ce que parce que, comme on l'a vu, l'entreprise qui sert d'intermédiaire présente des garanties (pour les clients et pour les *freelance*). Ceci dit, le changement du modèle du travail et du modèle social est important et suscite de nombreuses questions :

(a) On assiste à une nouvelle division du travail, une parcellisation plus fine qu'il y a quelques années, qui contredit une autre tendance de notre société, celle de l'automatisation. Automatisation et *freelance* sont les deux grandes tendances de nos sociétés. L'une conduit à un relatif enrichissement des tâches (quelquefois aussi à une nouvelle parcellisation), l'autre généralement à une plus grande parcellisation (encore qu'on assiste aussi parfois à un mouvement inverse, quand un *freelance* est capable de faire des travaux très différents). Topcoder, par exemple, met en relation 300 000 développeurs *freelance* avec les grandes entreprises, et réduit les coûts de 75% en parcellisant à l'extrême les tâches effectuées par des développeurs aux quatre coins du monde ; Pfizer fait la même chose en confiant de nombreuses tâches à des *freelance*, ce qui permet de décharger ses travailleurs à plein temps très qualifiés de 20 à 40% de tâches relevant de la routine. CloudCounsel fait la même chose pour le conseil juridique en confiant la rédaction des documents (brevets, accréditations, etc.) à des *freelance* effectuant des tâches très parcellisées.

(b) Le modèle entre souvent en contradiction avec le droit du travail existant. Aux USA, la situation est très variable d'un État à l'autre, mais de nombreuses entreprises de ce type craignent qu'on leur impose une requalification rétrospective des contrats de travail en contrats réguliers, lorsqu'ils existent (analogues aux CDD français). Plusieurs entreprises font face à des

difficultés du type de celles rencontrées par Uber (conflits avec les professions réglementées). Un débat s'engage un peu partout où s'affrontent ceux qui dénoncent le « Platform Kapitalism », en Allemagne par exemple, et ceux qui pensent que ce modèle est un modèle d'avenir. Les entreprises de ce type se heurtent souvent à la question de leur taille : plus elles sont de taille réduite (par exemple sur la ville de San Francisco) mieux elles fonctionnent, plus elles s'étendent (à l'ensemble des USA et au-delà) plus elles rencontrent des difficultés d'organisation, de motivation, de fidélisation (par exemple les chauffeurs Uber travaillent aussi pour Lyft) et se heurtent à des obstacles réglementaires ou à des résistances politiques. Il est à remarquer que ces conflits sont complexes : du côté des partisans il y a bien sûr les consommateurs, mais aussi souvent les travailleurs indépendants *freelance* qui tirent un bénéfice de ce système, de sa souplesse, de l'absence de démarches administratives, etc. (même s'il arrive que des conflits sociaux naissent entre l'entreprise intermédiaire et les *freelance*, comme cela a été le cas dans certains pays avec Uber). De nombreux étudiants soutiennent ce système, tout comme des jeunes femmes voulant élever leur enfant et avoir un travail à temps partiel, des retraités, des travailleurs d'élite voulant être totalement indépendants, des bohèmes *freelance* travaillant tout en voyageant aux quatre coins du monde, etc. Le paysage est donc contrasté : d'un côté une remise en cause évidente du droit du travail existant, de l'autre des opportunités nouvelles qui séduisent de nombreux travailleurs.

(c) La question de la protection sociale se pose évidemment partout, tout particulièrement aux USA, qui n'ont pas un modèle de cotisation universelle pour la retraite et l'assurance maladie (les travailleurs américains souscrivent retraite et assurance par l'intermédiaire de leur entreprise). Ce modèle n'est évidemment pas compatible avec le modèle de l'économie à la demande, qui suppose une totale portabilité des droits à la retraite et à l'assurance maladie. Ce type de travail accentue une tendance déjà présente dans de nombreux pays : les cotisations pour la retraite et l'assurance maladie sont déconnectées de l'entreprise, l'individualisation de la protection est complète, sur le mode de l'assurance privée. La *self reliance* (compter sur soi-même) devient la norme.

(d) La question de la fiscalité se pose évidemment : il est particulièrement difficile, pour les États, de mettre sur pied une imposition adaptée à ce type de travail.

Il est évident, enfin, que ce modèle est en relation avec la structure inégalitaire des sociétés où il se développe : les clients sont des jeunes urbains qui ont de l'argent mais peu de

temps, ceux qui offrent le service (qu'ils soient sur le territoire national ou n'importe où dans le monde) ont des compétences, du temps mais ont besoin d'argent (ce ne sont pas toujours, loin de là, des gens pauvres : il peut s'agir de jeunes médecins voulant arrondir leur mois, de retraités, etc.). Les services sont parfois facturés au prix fort : Medicast, aux USA, prend 200 \$ pour la visite d'un médecin moins de deux heures après l'appel.

Les questions que pose ce type d'économie sont donc nombreuses, mais on peut cependant écarter une critique. Le *big data* ne va forcément, comme on le croit souvent aujourd'hui, dans le sens de l'hyperconsommation. Cette nouvelle économie, tout comme l'exploitation des données dans le domaine des transports, de l'énergie, de l'architecture, etc. peut aller dans le sens d'une décroissance bien comprise. Quand on pense *big data* on pense évidemment au marketing par individualisation du client. C'est oublier ce que disent plusieurs spécialistes, qui ne sont pas pour autant des idéologues de la décroissance : « Je pense que l'on consommera mieux, et surtout moins » (Gilles Babinet⁸).

J'attire l'attention, pour finir, sur le décalage qu'il y a, fréquemment, ici comme ailleurs, entre la réalité (indéniable, comme on vient de le voir) et les discours. Le *big data* se prête sans doute moins aux discours idéologiques et aux utopies que les innovations antérieures (l'apparition d'Internet), mais ces discours ne sont pas pour autant absents. Dans le domaine économique, le *big data* s'accompagne parfois du déploiement d'utopies (qui peuvent être utiles, les techniques nouvelles font rêver et ces rêves contribuent à leur développement). Le *big data* entraînerait automatiquement la redéfinition des rapports de production, la fin du capitalisme exploiteur, la collaboration entre petits producteurs indépendants (*Do it yourself*, *Fablabs*), la résolution de tous les problèmes, écologiques en particulier. Derrière tout cela, il y a la vieille idée, qui a la peau dure, même si elle a été maintes fois critiquée, du *déterminisme technique* : une technique nouvelle entraînerait nécessairement une nouvelle forme d'économie et de rapports sociaux (pour le meilleur ou pour le pire, les deux versions de ce déterminisme existent : le déterminisme technique conduit soit à l'utopie soit à la dystopie). Gilles Babinet, dans *L'ère numérique*, décrit bien l'atmosphère utopique qui a accompagné, il y a quelques années, l'émergence, aux États-Unis, des *Fablabs* :

L'idée de Neil est alors de faire ce qu'il appellera de l'ingénierie sociale : repenser intégralement la façon dont nous concevons les produits, dans un

⁸ *Big, Fast, Open data...*, page 142.

contexte beaucoup plus horizontal et surtout beaucoup plus ouvert, afin que chacun puisse s'emparer des outils de production industrielle. Il constate que, si la plupart des machines nécessitent des formations complexes pour être mises en œuvre et que les documentations sont souvent mal conçues, il doit être possible d'en simplifier considérablement l'accès au travers d'un apprentissage ouvert et collaboratif. Il rassemble toute une série de machines et les met à disposition de tous : ses étudiants, mais aussi quiconque à Boston, où se situe le MIT, souhaiterait accéder à des équipements dernier cri, permettant de fabriquer presque n'importe quoi. Le premier Fablab – laboratoire de fabrication – venait de naître. L'initiative du MIT ne mit pas longtemps à devenir extrêmement populaire. L'idée que l'on puisse fabriquer localement pratiquement n'importe quoi a fait fantasmer plus d'une université ou association. Il est difficile de retracer ce qui a provoqué l'éclosion de Fablabs un peu partout. En juin 2008, on dénombrait trente-quatre Fablabs, répartis dans dix pays. Quatre ans plus tard, l'International FabLab Association en recensait cent quarante-neuf et, en 2013, il devient impossible de les compter. Il y en a des milliers, probablement des dizaines de milliers si l'on s'en tient au nombre d'appareils de stéréolithographie au sein des Fablabs.

L'idée maîtresse, c'est que le numérique permet de distribuer à coût zéro les fichiers qui permettent de reproduire tous types d'objets tout en facilitant la collaboration entre l'ensemble des acteurs, designers, électroniciens, ingénieurs ou même musiciens, médecins, chercheurs. La notion d'ouverture et de partage gratuit de la propriété industrielle, l'open source, devient un maître mot.

Cette vague participe à l'émergence du Do it yourself (...) comme projet de société pour le XXI^e siècle : l'idée sous-jacente est qu'en « le faisant eux-mêmes », les individus seront plus respectueux de l'environnement, leurs innovations seront plus largement partagées, ils apprendront à réparer et l'on assistera à l'émergence d'une culture industrielle cocréée et collaborative, plus généreuse, plus écologique.

Peu de modes de développement sont plus représentatifs de l'économie cellulaire : chaque Fablab peut améliorer les équipements dont il dispose et mettre en ligne ces améliorations, en accès open source. La collaboration est donc au cœur du fonctionnement des Fablabs : ils se transmettent les noms des logiciels de conception, mettent en ligne des milliers de fichiers 3D permettant d'imprimer un grand nombre d'objets, mettent en relation designers et concepteurs, leur donnant ainsi un pouvoir de collaboration créative impossible à envisager avant l'émergence de l'Internet. C'est l'ère du DIWO (Do It with Others), alimentée par une multitude de forums, de sites collaboratifs, rassemblant des milliers de passionnés qui discutent, échangent et partagent idées, créations, conseils et plans.

Derrière le fait d'ouvrir au tout-venant les techniques industrielles les plus pointues, il y a un souhait marqué de favoriser l'éclosion d'inventions rupturistes, d'astuces qui court-circuiteraient le modèle d'invention incrémentielle qui prévaut dans les grandes entreprises. C'est l'idée fortement défendue par l'économiste Eric von Hippel : l'innovation ascendante (user innovation) ou impulsée par les utilisateurs eux-mêmes, que l'on peut aussi dénommer « proams » ou professionnels-amateurs. Von Hippel a passé un

temps considérable à essayer, par-delà le phénomène de mode, de mesurer l'efficacité de ce modèle d'innovation ascendante. Il aurait trouvé deux à trois fois plus d'innovations de la part des makers qu'il n'y en aurait dans le monde économique traditionnel. Il observe en particulier que, lors de l'émergence d'un marché, il existe très peu d'utilisateurs, ce qui donne peu de raisons d'innover au fabricant, mais beaucoup aux utilisateurs.

Dans un monde où l'innovation de rupture serait la norme et où la capacité des nouveaux entrants à pouvoir se développer serait facilitée, l'exploitation des travailleurs ou même des humains par le capital ne serait ni inéluctable ni même vraisemblable.

Mis bout à bout, l'ensemble de ces gains d'opportunité dessinent un monde nouveau où le gâchis et la surconsommation seraient fortement réduits...

3. *Big data* et régulation

Je me contenterai d'attirer l'attention sur deux questions. La première, celle de la protection des données, est mieux connue que la seconde, celle de la prédiction des comportements, qui commence pourtant à soulever des inquiétudes et susciter des débats.

a. la question de la protection des données

Comme on vient de le voir sur l'exemple de l'économie à la demande, le droit a toujours du retard par rapport à l'innovation technique, économique et sociale. La question de la régulation du *big data* est une question essentielle aujourd'hui, dont je n'aborde ici que quelques aspects principaux.

La compatibilité des innovations avec la législation existante se pose dans de nombreux domaines d'activité. Le microcrédit, par exemple, rendu possible par le *big data*, a fait l'objet de critiques (taux très intérêt très élevés, multiples arnaques, etc.). Sa compatibilité avec la législation en vigueur, qui doit évidemment évoluer, est au centre de ces critiques.

Mais c'est surtout la question de la protection des données qui fait aujourd'hui problème. On sait depuis quelque temps que le principe « j'ai lu et j'accepte » ne fonctionne plus. D'abord, parce que personne ne lit plus ces clauses de confidentialité, ensuite parce qu'il y a de nombreux moyens, par croisement des données anonymisées, de faire apparaître des données de la vie

privée sans que personne ne viole ces clauses... Il faut ajouter à cela le fait que l'exploitation des données évolue sans cesse et qu'on ne peut demander à quelqu'un d'accepter un usage de ses données dont il n'a pas la moindre idée. Quant à lui demander d'accepter tous les usages futurs, c'est évidemment contradictoire avec l'idée même de consentement (on consent à un usage connu) ; et lui demander de refuser toute exploitation inconnue de ses données détruirait le principe même du *big data* et tous les avantages qu'il peut représenter pour les individus et la collectivité.

La solution la moins mauvaise semble à chercher dans trois directions simultanément : la législation visant à encadrer l'exploitation des données ; les mesures visant à favoriser l'auto-régulation, la responsabilité, des entreprises exploitant les données ; enfin l'éducation et la responsabilité des utilisateurs.

Les données sont encadrées, en France, par la Commission nationale de l'informatique et des libertés (CNIL), qui veille à l'application de la loi « Informatique et libertés », une loi européenne (1978) qui est devenue quasiment mondiale par une directive européenne de 1995 (les données ne peuvent être collectées hors de l'Union européenne que si le pays offre une « protection adéquate » ; si ce n'est pas le cas, le collecteur doit signer une convention où il s'engage à respecter la loi européenne). Cette directive a inspiré la législation ou les projets de législation de plusieurs autres pays (Chine, Tunisie, Maroc, etc.). En mai 2014, le Parlement européen a voté un règlement devant entrer en vigueur en 2015 et être appliqué effectivement deux ans après, imposant, partout où cela est possible, que les données soient non seulement proportionnées mais non nominatives. Mais il reste bien des incertitudes : on ne sait pas trop, par exemple, comment un tel principe peut être appliqué avec le développement de l'internet des objets connectés...

La CNIL veille à faire respecter quatre exigences à l'endroit des « données à caractère personnel » : consentement, opposition, accès direct, rectification. Il s'agit de garantir à l'internaute qu'on stockera sur le réseau des informations le concernant avec son accord explicite et qu'il pourra en contrôler l'accès et les modifier ou effacer quand il le souhaite. Une ambition, disent Fogel et Patino, « qui tient de l'édification d'un barrage face à un océan ».

On peut mesurer la complexité de la situation si l'on considère qu'aujourd'hui qu'il n'existe nulle part, officiellement, de droit de propriété des données. Et pourtant Facebook

demande aux utilisateurs de lui accorder la licence d'exploiter ses données ! Il construit donc un contrat sur ces données alors que le droit de propriété de celles-ci n'existe pas. D'où la nécessité de clarifier juridiquement le statut des données et, même en l'absence de droit de propriété, d'établir des règles assurant à chacun un certain contrôle sur ses données.

Une autre difficulté réside dans le fait que la CNIL est censée garantir que l'exploitation des données n'est pas en contradiction avec la finalité initiale de la collecte des données. Le problème est qu'avec le *big data* il n'y a pas vraiment de finalité précise, en raison du croisement des données... La loi s'adapte plus ou moins à cette situation, certains cas sont clairs (on ne peut chercher à la fois les habitudes de lecture et la religion du lecteur), d'autres le sont moins... on reste dans le flou juridique. Il n'en demeure pas moins que la protection en Europe est supérieure à ce qu'elle est dans le reste du monde. Il est interdit, par exemple, de croiser les données des banques et assurances (pour refuser une assurance à quelqu'un, sur la base de sa consommation par exemple : l'analyse des relevés de compte est subordonnée à l'accord du client, etc.). Mais il est clair que les nouveaux usages et les nouvelles pratiques sont toujours en avance sur le cadre juridique, et qu'il y a un énorme travail juridique à faire...

Depuis mai 2014, l'Union européenne a établi, dans un cas opposant un citoyen espagnol à Google, un « droit à l'oubli » (décision de la Cour de Justice). Mais la difficulté de sa mise en œuvre est grande, car Google ne fait que référencer l'information figurant sur des sites ne lui appartenant pas. Et le droit à l'oubli peut entrer en conflit avec le droit à l'information (les Américains parlent de « right to know ») : « le bourreau tue deux fois, la seconde fois par l'oubli » (Elie Wiesel).

La deuxième voie, bien aussi importante que la régulation par la loi dont on vient de parler, est la responsabilisation des grands collecteurs de données. On va de plus en plus vers des entreprises qui demanderont à l'utilisateur son degré de consentement d'utilisation des données, avec évidemment plus ou moins de services à la clé... On peut envisager des comités d'éthique, qui permettent de réguler, avec tous les acteurs, l'utilisation des données, sans paralyser celle-ci. Il ne faut pas la paralyser, car la plupart des utilisations du *big data* qui ont été découvertes récemment l'ont été à partir de données qui n'avaient pas été recueillies à cet effet. C'est un trait caractéristique de l'exploitation des données qu'on peut toujours tirer de nouvelles informations de données anciennes, souvent en les couplant avec des données récentes. D'où la tendance des entreprises de ce secteur (Google, etc.) à vouloir conserver les

données à l'infini, tendance qui entre en contradiction, souvent, avec le souhait des usagers (qui réclament un droit à l'oubli).

Plusieurs spécialistes de la protection des données proposent de renforcer, par la loi mais aussi par l'éthique – qui peut aller au-delà de la loi – un engagement des entreprises responsables chargées de l'exploitation des données. Cela suppose une régulation dans ce domaine, avec un corps de contrôleurs extérieurs spécialisés (capables d'analyser les algorithmes) ainsi qu'un département, à l'intérieur de chacune de ces entreprises, relativement indépendant de la direction, chargé d'appliquer les règles concernant le respect de la vie privée. Les entreprises traitant les données sont les seules à savoir exactement ce qu'elles font, c'est elles qu'il faut contraindre à appliquer les lois, en les contrôlant de l'extérieur et de l'intérieur (pour qu'elles n'aient pas à révéler trop d'algorithmes confidentiels à l'extérieur). Un des aspects de cette régulation est évidemment la durée de conservations des données, qui doit trouver un équilibre entre le respect de l'innovation et le respect de la vie privée. La protection des données passe aussi par des innovations techniques (par exemple la *differential privacy*, des outils qui permettent de rendre floues les données) visant à rendre possible une exploitation des données, mais impossible ou excessivement coûteuse l'identification des individus.

Quelques mots, pour finir, sur la responsabilité de l'utilisateur. La réalité est que les citoyens ont de très nombreux droits (savoir si les données existent, les modifier, les supprimer, etc.) mais qu'ils n'en font pratiquement pas usage. Soit par ignorance, soit parce qu'ils n'estiment pas que l'exploitation des données est dangereuse... Les citoyens, tout au plus, pilotent la collecte de leurs données (par Google, par exemple), mais tous ne le font pas... Et sont-ils prêts à payer 10 euros par mois pour que Facebook n'exploite pas leurs données ? On peut en douter. Les internautes ne devraient pas oublier que des entreprises comme Google fonctionnent selon le principe « service contre données ». Il y a une schizophrénie du consommateur : il a le souci de protéger sa vie privée et la tentation permanente de l'hyperconnexion... Il veut des services gratuits, mais sans céder ses données, en oubliant qu'offrir plus de données, c'est recevoir un meilleur service. Il convient également de signaler l'importance de l'apprentissage, dès l'école, des langages de programmation. Quand un individu sait coder, il est beaucoup plus sensible à ces enjeux de protection des données...

Je signale enfin que d'autres mesures font l'objet d'intenses réflexions, en particulier celles qui visent à faire des données un *bien public* ou un *bien commun*. Je me contenterai de

dire que ces solutions, qui semblent appropriées à certaines données (cf. l'*open data*) semblent difficilement généralisables. Fogel et Patino, dans *La condition numérique...*, font une bonne analyse de ces solutions :

La transformation des données en bien public est une hypothèse qui demeure politiquement envisageable mais difficile à mettre en œuvre. Elle promeut l'idée d'une utilisation contre redevance, mais suppose un contrôle par un organisme représentant l'intérêt général : organisation internationale, État, fondation, association, etc. La diversité des solutions traduit la difficulté de cette proposition aux conséquences économiques radicales. Une autre solution serait de faire des données un bien commun. On rejoint là les travaux du prix Nobel d'économie Elinor Ostrom qui propose des « arrangements institutionnels » pour que des organismes collectifs gèrent ensemble et de façon optimale des écosystèmes. Une expression fructueuse de cette solution se trouve dans les *creative commons*, un système de droits d'auteur créé par un professeur de droit, Lawrence Lessig, et utilisé sur le réseau afin de préciser l'utilisation de chaque contenu : libre, avec mention de l'auteur, avec interdiction à l'usage commercial, avec autorisation préalable, etc.

Ces réflexions et expérimentations (par exemple les *creative commons*) nourrissent la réflexion juridique. Chaque jour de nouvelles solutions juridiques apparaissent (par exemple les licences du type Open Database Licence⁹, le *copyleft*¹⁰, qui permet à chacun de profiter des données sans pour autant les accaparer, etc.). Certains mouvements animent la réflexion sur ces solutions, par exemple My Data¹¹. On peut noter que sur ce point comme sur d'autres l'avenir est ouvert. D'autres propositions vont en effet dans le sens, inverse du précédent, de la privatisation des données, c'est-à-dire d'un contrôle de celles-ci par les individus prenant la forme d'une vente (aux enchères parfois) de leurs données. Un marché des données est en train de se constituer.

Pour en finir avec la question de la protection des données, il convient d'ajouter que la protection des données individuelles n'est qu'une partie du problème, peut-être pas la plus importante. L'espionnage cybernétique visant les entreprises et les États est une menace peut-être beaucoup plus grave. Une entreprise qui se fait voler ses données subit un énorme préjudice, tout comme un État dont les données « confidentiel-défense » ou même simplement « sensibles » sont pillées (on peut paralyser une armée à partir d'un vol de ses données, plus par éradication des disques durs que par exfiltration des données). Les données de l'État

⁹ Cf. http://fr.wikipedia.org/wiki/Open_Database_License,

¹⁰ Cf. <http://fr.wikipedia.org/wiki/Copyleft>

¹¹ Cf. <http://mydatalabs.com/Accueil>

français sont bien protégées, mais on ne peut en dire autant, vous le savez, des données des entreprises.

b. la question de la prédiction des comportements

L'analyse des innombrables données relatives à un individu peut permettre des prédictions concernant son comportement (risque d'infarctus, de cancer, de ne plus pouvoir rembourser son crédit, de violer quelqu'un, de braquer une banque, de participer à un réseau terroriste, etc.). Les compagnies d'assurance, les banques, la police, sont évidemment intéressées par ces prédictions. On peut imaginer – nous n'en sommes pas très loin – des arrestations préventives, à la limite des peines préventives !

Le *big data*, qui peut être très utile, peut présenter aussi une face inquiétante, et les deux sont intimement mêlées. Certaines compagnies d'assurance (qui aux USA ont la fonction qui est celle de la Sécurité Sociale chez nous) utilisent le *big data* pour déterminer la probabilité qu'une personne prenne ses médicaments (en fonction de critères relatifs à son mode de vie, sa résidence, son âge, son travail, etc.). À partir de cette information, elles peuvent moduler la prime d'assurance mais peuvent aussi mieux cibler les adhérents auxquels elles doivent rappeler qu'ils doivent prendre tel médicament qui vient de leur être prescrit. Aviva, une grande compagnie d'assurance, étudie la possibilité d'utiliser les informations fournies par la carte de crédit et la visite des sites Internet pour conseiller l'analyse de sang et d'urine en vue de détecter l'hypertension ou le diabète (en se fondant sur les loisirs, le temps passé devant la télévision, etc.). La prédiction des comportements humains peut être aujourd'hui très précise et se faire par des corrélations très indirectes : la mesure des changements dans le mouvement d'un smartphone qui est dans la poche de quelqu'un peut révéler des symptômes (mal au dos modifiant sa marche et sa posture, etc.), la voix au téléphone peut révéler la grippe avant même qu'elle ne se déclare. Certains hôpitaux et maternités, dans certains pays, installent des capteurs sur les malades ou des caméras pour devant leur lit pour obtenir des données qui permettent de détecter une infection, ou une aggravation d'un symptôme, ou un risque d'épidémie bien avant que le corps médical ne repère le problème. Un capteur attaché à un spray anti-asthme permet, grâce à la géolocalisation, de repérer certaines causes de l'asthme (proximité de certaines céréales ou pollutions, etc.). Mais d'autres usages pourraient sembler plus contestables : certains États américains ont déjà essayé de moduler la libération conditionnelle à partir d'analyses de données sur le comportement ancien des prisonniers...

Mayer-Schonberger et Cukier, dans l'ouvrage déjà cité, estiment que face à ces problèmes le droit doit poser le principe de la *responsabilité individuelle* sur la base des seules actions ayant effectivement eu lieu (non des actions potentielles dont on peut mesurer la probabilité par l'étude des données). La police et la justice pourraient concentrer leurs efforts de surveillance sur les secteurs et individus dangereux mais non les inculper sur cette base. D'autres auteurs estiment qu'il n'est pas scandaleux d'envisager une libération conditionnelle sur la base de données, à la condition que le juge garde la main sur la décision. Le grand maître d'échecs assisté par l'ordinateur bat tout programme d'échecs...

La même question se pose au niveau du recrutement pour tel ou tel emploi, dans la mesure où l'entreprise peut désormais disposer de très nombreuses données sur le passé des candidats. Là encore, il n'est pas scandaleux d'utiliser des données relatives aux individus, même si celles-ci sont souvent peu fiables (comportements anciens du candidat, sortis de leur contexte). C'est pourquoi de nombreux juristes plaisent pour une législation qui, sans interdire l'exploitation de certaines données, encadre celle-ci, en imposant par exemple la transparence des bases de données et des algorithmes utilisés par l'entreprise, ainsi que la possibilité d'une contre-expertise lorsqu'un employé s'estime victime d'une discrimination à l'embauche sur la base d'une exploitation de données, etc.

Certains juristes parlent, à ce propos, de la nécessité d'un « droit des algorithmes ». Il existe par exemple, en droit français, une interdiction de prendre des décisions de justice sur la base d'algorithmes (article 10 de la loi du 6 janvier 1978 modifiée). Cette législation pourrait évoluer et instaurer un droit de prendre certaines décisions sur des personnes avec assistance d'un robot, à condition qu'intervienne *in fine* la main humaine... Les problèmes juridiques ne manquent pas, cependant : les personnes doivent-elles dans tous les cas connaître ce l'on pense d'elle par exploitation des données ?

4. *Big data* et gestion démocratique

a. *l'open data*

On assiste, depuis quelques années, au développement d'une sorte d'*open data* partiel : certains gouvernements et entreprises acceptent de rendre publiques certaines données pour un certain nombre d'acteurs. Les questions relatives à l'*open data* étant mieux connues que celles relatives au *big data* et aux *smart cities*, je me contenterai de quelques remarques.

L'administration française s'y met, depuis 2011. Elle a construit un site, Etatlab, dépendant du Premier ministre, consacré à l'ouverture des données publiques, qui a fait naître quantité de projets, d'applications, de services, de métiers, de startups, etc. Un exemple : l'application Tanquillien, qui utilise les données de la SNCF pour que chacun puisse mieux gérer son voyage (savoir dans quel wagon il y a de la place, par exemple). Mais selon le responsable d'Etatlab, l'*open data* français est un « gruyère plein de trous ». Il y a une résistance de nombreuses administrations. Nombre d'entreprises et associations ont pourtant besoin d'exploiter les données non sensibles de l'administration, par un accès à des bases lisibles et homogènes. Les résultats des élections, par exemple, sont difficilement exploitables en raison de l'absence de données complètes sur les localisations des bureaux de vote et les découpages du système électoral. La mise au point de la cartographie, autre exemple, est essentielle pour la vie économique et certaines associations (OpenStreetMap) font aussi bien, voire mieux, que l'IGN. L'État a sa propre responsabilité dans la collecte et le traitement des données, mais il peut être aidé aussi par la société, qui peut trouver de nouveaux usages, particulièrement innovants, des bases de données (celles concernant les accidents de la route ont été récemment libérées, ce qui a permis d'établir des corrélations particulièrement utiles). Mais l'ONISEP (Office national d'information sur les enseignements et les professions) n'avait pas encore, en 2014, libéré ses données... Nous sommes encore, avec l'*open data*, trop souvent dans le *small data*... Les administrations perçoivent souvent assez mal l'enjeu de la libération des données. Quand la ville de Paris a demandé aux Affaires culturelles si elles avaient des données à partager, la réponse a été négative : les œuvres ne sont pas des données, elles sont couvertes par les droits d'auteur, etc. En insistant, on a découvert une base : la gestion des emprunts et retours

dans les bibliothèques de Paris, qui est une source très riche sur la consommation culturelle par secteur géographique et sociologique...

Les spécialistes de l'*open data* insistent sur l'utilité, pour les individus et la société dans son ensemble, de la libération des données. Ils insistent également sur le fait que cette libération des données permet de mieux gérer l'État et de mieux comprendre comment il fonctionne. La mise en évidence des dysfonctionnements de l'État, ou tout simplement du fait que certains services ne peuvent plus être rendus de manière satisfaisante par l'État, a fait naître deux tendances au sein de nos sociétés : d'une part la méfiance des administrations et leur réticence à libérer les données qui peuvent l'être, d'autre part une idéologie selon laquelle la libération des données pourrait et devrait mettre en extinction l'État. Gilles Babinet se fait clairement le héraut, dans *L'ère numérique*, d'une telle idéologie, en affirmant, sans la moindre donnée factuelle permettant d'étayer ses affirmations, que

Pour remplir des missions essentielles, comme la sécurité intérieure ou extérieure, l'émission de monnaie, la réduction des inégalités, l'organisation du marché du travail, le fonctionnement du système de santé, la planification urbaine ou de nombreuses autres fonctions, les systèmes digitaux pourraient être incomparablement plus efficaces que ceux que nous connaissons, plus réactifs, plus personnalisés, plus préventifs et, évidemment, plus économiques.

Dans la pratique, l'*open data* se développe en faisant reculer peu à peu ces deux tendances, le refus de la libération des données tout comme la vulgate libertarienne qui voit dans l'*open data* une panacée. Un certain nombre de fonctions qui relevaient de l'administration tendent certes à disparaître avec le *big data*, elles sont confiées aux citoyens (informations touristiques, système de guidage des handicapés dans la ville à Nantes, réalisé par des applications faites par les handicapés eux-mêmes, etc.). Mais cela ne fait pas disparaître la fonction politique. De nombreuses études montrent qu'il ne faut pas fantasmer sur la démocratie directe, qui selon certains serait une conséquence inéluctable de la libération des données. À Londres, il y a dans certains quartiers une application, Youchoose, qui permet aux citoyens d'avoir accès aux données relatives au budget du quartier et de modifier en ligne la répartition des lignes de celui-ci, en fonction de ce qu'ils estiment être les priorités. Mais c'est tout au plus un sondage en ligne, les autorités élues gardent la main sur la décision... Et on ne peut passer sous silence le risque de confiscation de la parole par les citoyens les plus actifs et les plus aisés.

Comme nous allons le voir en étudiant le fonctionnement des villes intelligentes, la libération des données, loin de signifier la fin de la politique, s'articule avec elle, à tous les niveaux, au niveau de la ville comme au niveau national et supranational.

b. les *smart cities*

Le terme de *smart city* recouvre des réalités différentes, de vastes projets de cités connectées créées ex nihilo (Songdo, en Corée du Sud, et d'autres villes au Portugal, Brésil, etc.) mais aussi la connexion de villes déjà existantes. Ces derniers projets sont moins utopiques. Je parlerai surtout d'eux, mais il est intéressant de considérer la perspective sous-jacente aux projets plus utopiques, dont je dirai aussi quelques mots.

On connaît le principe des *smart cities* : les données fournies par de multiples bases privées ou publiques, sont exploitées par divers acteurs publics, privés ou associatifs, voire par de simples citoyens, tous ces acteurs étant munis des outils nécessaires (l'outil Citizen Collaboration, ou encore Intelligence Operation Center d'IBM, aux USA), cela dans le but d'améliorer la vie urbaine. Quelques exemples : l'anticipation des risques (hydrauliques, à Montpellier), la rationalisation des transports, les économies d'énergie (détection des fuites d'eau par capteurs sur les canalisations par la mise en place de *smart meters* comme à Beaune et Lyon avec Veolia), la mise en évidence de besoins débouchant sur la création de nouvelles entreprises, ou l'innovation dans les entreprises existantes...

Une des caractéristiques de la *smart city* est de transformer l'*usager* des services publics en *acteur* : si vous avez des données concernant votre production de déchets, la qualité de votre tri, le prix de la collecte pour la communauté ou pour vous-mêmes, vous agirez différemment, vous interviendrez sur votre production de déchets (en devenant un « consomm-acteur » : avec les *smart grids* de l'énergie, l'utilisateur intervient par ses décisions pour influencer l'offre, en produisant de l'électricité solaire, par exemple, et la demande, en modifiant sa consommation grâce aux données ; c'est l'objectif du compteur Linky). Le *big data*, plus généralement, a tendance à rendre l'État ou l'administration municipale plus modeste, il permet de diminuer mais surtout de modifier ses tâches. L'administration n'est plus un guichet ouvert, mais un lieu de partage des expériences et bonnes pratiques, d'animation de réseaux d'acteurs. Le pouvoir politique crée par exemple des écolabels, des écocités, mais fait intervenir, pour leur conception et pour les réalisations, de nombreux acteurs dont il anime et coordonne l'activité. Le pouvoir

politique municipal est chargé de la co-construction d'une vision commune, plutôt que de la construction d'une vision nationale ou municipale. Un exemple : Lyon Urban Data, un *living-lab* très actif du Grand Lyon, qui mutualise des données publiques et privées, en partant du principe que l'*open data* ne suffit pas. On peut en effet ouvrir les données sans que rien ne se produise ; il faut, pour que l'ouverture des données soit efficace, créer des structures pour l'exploitation de celles-ci. On l'a vu récemment dans un autre domaine, avec l'affaire des comptes en Suisse d'HSBC : faire parler les données est un métier complexe, ce n'est pas le premier venu qui peut le faire, même si les outils informatiques nécessaires sont assez simples (il faut en particulier avoir une idée de ce que l'on va chercher). L'ouverture des données, à Lyon, permet de faire de l'engineering des nouveaux services urbains en associant industriels, entrepreneurs (Veolia, EDF, Orange, etc.), université, laboratoires, pôles de compétitivité, usagers. Les projets ne prennent pas la forme d'une commande publique mais surgissent de la confrontation des points de vue entre ces acteurs. Un exemple de ces projets : Lyon Smart Community, issu d'une alliance avec le Japon (l'agence NEDO, équivalent de l'ADEME chez nous) : 60 millions d'euros, consacrés au déploiement d'un outil de management énergétique pour un quartier (construction d'un lot de bâtiments à énergie positive, mise en place d'un service d'autopartage alimenté par panneaux photovoltaïques, programme de rénovation, gestion de l'énergie par les industriels et les particuliers). Il y a, au total, 31 ans acteurs différents : la structure Lyon Smart Community est un acteur parmi les autres qui organise l'interface entre les partenaires, et en particulier pose et tente de résoudre les problèmes juridiques (par exemple le vide juridique pour les voitures sans chauffeur) et, point important, est chargée d'expliquer le projet à la population et à faire qu'elle s'engage activement dans celui-ci. La ville intelligente n'est pas seulement un ensemble de réseaux techniques, c'est une politique nouvelle, qui ne consiste pas à *faire* la ville mais à *faire faire* par des agents multiples, des métiers très divers (la ville intelligente brise ce qu'on nomme les silos, les métiers verticaux sans rapports les uns avec les autres, dans l'administration comme dans le privé ; on sait que les DSP, les délégations de service public, pour l'eau, le gaz, etc., sont organisés en silo). Dans l'expression « ville intelligente », l'intelligence n'est pas à comprendre uniquement au sens computationnel du terme.

Cela dit, il est important de voir que le terme de *smart city* recouvre divers discours et diverses pratiques. Antoine Picon, un architecte et professeur d'architecture à Harvard, auteur d'une des meilleures études sur les *smart cities*, fait remarquer que ce qui caractérise la ville intelligente est l'invisibilité de l'intelligence en question (personne ne remarque la présence de

puces de radio-identification – RFID – dans les 120 000 arbres de Paris, qui permettent aux jardiniers de garder trace de leurs interventions, d'intervenir plus efficacement, etc.). La technique est intégrée au social, comme elle l'est déjà depuis assez longtemps au vivant. C'est une excellente chose : plus la ville devient complexe, plus elle nous apparaît simple. Mais il y a aussi l'illusion de la transformation du social en un système automatique, pensé soit sur le modèle de la machine cybernétique autorégulée et surveillée depuis une salle de contrôle, soit sur le modèle de l'organisme vivant, comme en témoigne la métaphore fréquente du métabolisme de la ville. Si la ville intelligente était un tel système, elle nous dispenserait de l'intervention humaine consciente, c'est-à-dire de la politique.

Picon insiste, en prenant appui sur les meilleures expérimentations en matière de ville intelligente, sur le fait que ces métaphores du vivant et de la machine sont trompeuses. Il vaudrait mieux comparer la ville intelligente à un humain, c'est-à-dire à un vivant doté de conscience de soi : la ville, comme certains l'ont dit, devient sensible (en anglais on dit *sentient city*, capteur se dit *sensor* en anglais ; on a même parlé, dans un jeu de mots, de *senseable city* : une cité qui est capable de sentir, *sense-able*, mais on peut prononcer aussi *senseable* comme le mot anglais *sensible*, qui signifie accessible par les sens). La ville semble consciente de soi, elle ressemble à un humain. Mais même cette métaphore plus appropriée est à éviter, car elle pourrait laisser croire qu'il y a un contrôle central, elle pourrait laisser croire que les citoyens sont seulement les sens, et le pouvoir central le cerveau. Or la ville connectée, insiste Picon, n'est pas un système cybernétique commandé depuis une salle de contrôle (on a rêvé d'une ville de ce type, composée d'un système autorégulé couplé avec un centre de contrôle du type salle de la Nasa, dans les années 70-90 ; IBM voulait transformer Rio en une telle ville¹²). Il est clair qu'il y a eu et qu'il y a encore une tentation technocratique néo-cyberticienne dans l'idée de *smart city*, le fantasme d'un contrôle total par un centre sur la base des innombrables données, selon un modèle militaire et dans la ligne d'une tradition de la pensée architecturale utopiste, voire totalitaire. En réalité, la *smart city* est bien différente, elle n'élimine pas le pilotage politique (qui n'est pas un pilotage technique analogue à celui d'une salle de contrôle de la Nasa) mais l'articule avec de nouveaux types d'intervention d'autres acteurs, dont les citoyens

¹² Picon, *Smart cities...*, page 38-39 ; un article paru dans le New York Times illustre bien le propos de Picon : *Mission Control, Built for Cities. I.B.M. Takes 'Smarter Cities' Concept to Rio de Janeiro* (NYT, Natasha Singer, March 3, 2012) : http://www.nytimes.com/2012/03/04/business/ibm-takes-smarter-cities-concept-to-rio-de-janeiro.html?pagewanted=all&_r=0

(même si pour la gestion de certains risques majeurs, le modèle néo-cybernétique est pertinent¹³). Les *smart cities* sont plutôt à penser sur le modèle de sous-systèmes partiellement automatisés (gestion de la circulation, distribution de l'eau, ramassage des déchets, par exemple) qui supposent l'intervention d'opérateurs humains à différents niveaux : acteurs privés, associations et citoyens (ces derniers intervenant en tant qu'amateurs éclairés) et responsables politiques, ces derniers mettant en place ces systèmes et animant et coordonnant l'action des autres acteurs. Tous ces acteurs humains sont en dialogue permanent avec les robots, si ce terme convient encore pour désigner de tels sous-systèmes partiellement automatisés... On est loin du système cybernétique à commandement central, encore plus loin du rêve de la « singularité technique » où les robots prendraient définitivement le pouvoir. Le modèle est plutôt conversationnel que cybernétique. En un certain sens, la notion de *cyborg city*, qui a été avancée par certains, pourrait être adéquate, dans la mesure où elle attire l'attention sur l'hybridation de la machine et de l'homme – mais elle a des inconvénients aussi, dans la mesure où elle tend à concevoir les dispositifs techniques de la ville sur le modèle de la machine totalement automatisée et qu'elle laisse dans l'ombre la nature de l'intervention humaine, qui on l'a vu n'est pas celle d'un centre de contrôle. Le risque est de penser la ville comme « pilotée », par les humains ou les robots, ou les deux ensemble, à l'inconvénient de rabattre le social sur le technique, de concevoir finalement la société comme une machine (dirigée par les robots ou les hommes, ça ne fait pas grande différence : la direction est toujours centrale) et de faire oublier la dimension polycentrique et conversationnelle du social. Platon, déjà, préférait, pour parler du gouvernement, la métaphore du *tisseur* à celle du *pilote*. Il n'y a pas d'algorithme central prenant la décision politique, ni d'institution humaine centrale. On en vient à rappeler une saine banalité : les villes et les sociétés sont dirigées *politiquement*, par des acteurs humains multiples (en raison de la division des pouvoirs) en interaction avec de nombreux autres centres (organisations professionnelles, société civile, etc.), chacun dialoguant avec les autres et avec les multiples systèmes partiels et partiellement automatisés.

Antoine Picon écrit qu'à « à la ville réglée et pilotée comme un système s'oppose celle qui est censée naître de la coordination spontanée de multiples interactions individuelles à la façon d'un *wiki* ou d'un *mashup* géant »¹⁴. Il est important de noter l'expression « qui est censée naître » : car une société n'est pas non plus gouvernée seulement par les multiples interactions dont parle l'auteur, elle articule la logique spontanée de ces multiples interactions avec le

¹³ Comme le fait remarquer l'auteur, à la page 42.

¹⁴ *Ibid.*, page 50.

gouvernement politique, articulation que la notion de « gouvernance » tente de saisir. Si la ville n'est pas pilotée, elle n'est pas non plus le produit d'un immense *wiki* (un travail collaboratif à l'image de Wikipédia, par exemple) ou un compendium obtenu par *mashup* (en bon et long français : une application composite composant des contenus issus de différentes applications).

L'engouement pour les *smart cities* suscite en ce moment des critiques qui ne sont pas toutes justifiées (certaines reposent sur une technophobie et un anticapitalisme primaires, alimentés il est vrai par les élucubrations du transhumanisme, qui rêve d'une mutation de l'homme en cyborg, voire d'une prise de pouvoir par les robots). Mais Adam Greenfield, dans *Against the smart city*, insiste à juste titre sur le fait que les données nécessaires pour le bon gouvernement d'une ville ne sont pas toutes de type quantitatif, qu'elles doivent être fournies par les opinions argumentées des citoyens (qui ne sont que très mal mesurées par les enquêtes d'opinion). Le caractère technocratique du projet des *smart cities* peut être questionné au nom de la démocratie et de l'intervention politique des citoyens. Si l'on n'y veille pas, le *big data* peut conduire à une nouvelle forme du *nanny state*.

D'autres critiques, de manière plus précise, attirent l'attention sur les nombreux problèmes juridiques que posent les *smart cities*.

À qui les données publiques appartiennent-elles ? Certaines villes en conservent la propriété pour les exploiter (sous la forme de la *régie*, en France), d'autres les cèdent sous contrat à des entreprises... Le citoyen a-t-il un droit sur les données publiques ? sur toutes ? gratuitement ? quel type d'exploitation faut-il autoriser, par quel type de licences ? L'exploitation transversale des données pose en effet de nombreux problèmes juridiques. Quand il y a de très nombreux acteurs qui partagent les données, la question se pose aussi de la responsabilité de l'exploitation : quelles données, jusqu'à quel point peut-on les exploiter, pendant combien de temps, etc. ? La nécessité d'une maîtrise et d'une coordination de l'exploitation des données par les collectivités locales est à l'évidence une nécessité. La transversalité est en fait un nouveau métier, un ensemble de nouveaux métiers, et elle doit être encadrée, juridiquement assurée... La secrétaire d'État française au numérique envisage actuellement la création de données « d'intérêt général » : des données publiques, qui viendraient de l'administration, mais aussi peut-être d'entreprises privées (transport, par exemple), auxquelles on pourrait acheter ces données (devraient-elles les vendre

obligatoirement ?). Elle envisage aussi une forme d'action de groupes contre les entreprises violant le droit des données.

D'autres problèmes juridiques concernent les objets connectés domestiques (bandes de fréquence utilisées, propriété et maîtrise de ces données, car ces objets sont très mal protégés, ils présentent de nombreuses failles exploitables par des hackers : un réfrigérateur connecté pourrait permettre de connaître votre religion, ou que vous souffrez du diabète) et, en particulier les objets connectés *wearable* (qui fournissent des données relatives à la santé de chacun). La solution serait d'obtenir le consentement des clients, en leur appliquant les directives européennes, mais c'est bien difficile lorsqu'il s'agit d'une balance connectée ou, oui ça existe !, une brosse à dents connectée (dont l'intérêt reste à démontrer). C'est pourquoi l'enthousiasme pour l'Internet des objets est en train de refroidir, comme en témoigne un dossier publié dans le Financial Times en février 2015.

On est passé progressivement du *un pour tous* (top-down) au *tous pour un* (bottom-up), où l'utilisateur est actif, au *tous pour tous*, où l'utilisateur devient acteur, co-acteur... (avec les autres citoyens, avec des entreprises privées, avec les pouvoirs publics, qui par exemple mettent en place le système de gestion des places de parking, autorisent les places, les capteurs, etc. et passent ensuite la main à des entreprises privées et aux citoyens). Le pouvoir politique fait à la fois face à un défi (conurrencé par d'autres acteurs) et à de nouvelles tâches plus complexes. Il ne faut pas lire ce processus seulement comme une réduction de la fonction politique.

En guise de conclusion...

Le tableau ci-dessous, extrait d'un article du sociologue Dominique Bouillet¹⁵, montre bien que le *big data*, loin d'être un phénomène unifié, consiste en fait en tendances qui s'équilibrent plus ou moins bien aujourd'hui. Il y a encore, d'abord, la bonne vieille gestion de la ville (en bas à droite, *Good Old City*) ; il y a aussi la conception cybernétique dont j'ai parlé plus haut, celle d'IBM par exemple, la ville intelligente sur le modèle de la « cabine de pilotage » (en bas à gauche) ; il y a aussi la *data city* du type Google, la production de données

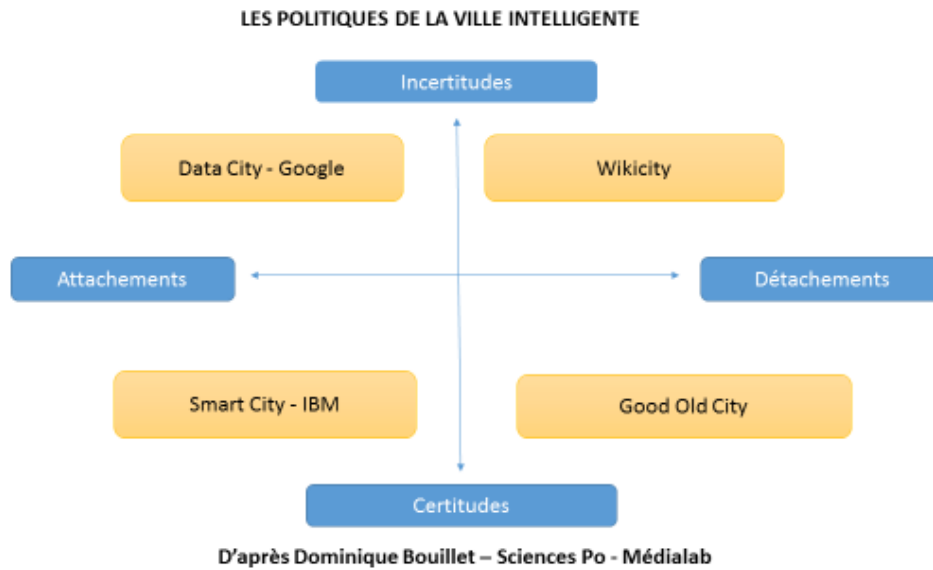
¹⁵ *Ville intelligente, ville démocratique ?*, page 112.

et utilisation par une multitude d'acteurs privés, dont l'activité est à finalité commerciale (en haut à gauche) ; il y a enfin la *wikicity* du *tous pour tous* (par exemple l'application SeeClickFix, où les citoyens rapportent en ligne à leur mairie tout ce qui ne va pas, un lampadaire qui ne fonctionne plus, un trou dans la rue, etc.).

Le schéma fait apparaître deux axes. L'axe vertical est celui du *degré de certitude* : la bonne vieille gestion et la gestion intelligente type IBM maîtrisent totalement l'utilisation, elles reposent sur la certitude du centralisme, au prix de réduire l'innovation ; la *data city* et la *wikicity* sont plus incertaines et ouvertes, mais plus créatrices. La *wikicity* fait quelquefois peur aux administrations, en raison de cette incertitude (chacun sait que ça marche bien dans le cas de Wikipédia, mais cela peut produire aussi du désordre, déposséder les élus de leur pouvoir). L'axe horizontal est celui des *attachements* (un terme de Bruno Latour) : les modèles sur la droite du tableau sont plus attachés aux citoyens, à leurs pratiques et traditions, les modèles sur la gauche du tableau sont davantage détachés des citoyens (ils font intervenir des entreprises ou administrations spécialisées).

Il est normal que les quatre modèles coexistent. Il y a des domaines de l'activité urbaine qui se prêtent mieux à tel ou tel type d'exploitation des données. Mais il y a aussi un enjeu politique, qui est de savoir lequel ou lesquels des modèles doivent devenir dominants. Il me semble qu'on pourrait ajouter qu'un cinquième modèle ne figure pas sur le schéma : c'est celui qui permet aux autres de jouer simultanément. Ce modèle, c'est la politique. Dominique Bouiller dit cela de manière dense et frappante : « il faut être capable d'organiser un lieu de confrontation et de montrer qu'il y a des possibles et qu'il n'y a pas que le visionnaire et la fatalité »¹⁶. Le visionnaire, c'est celui de la *smart city* utopique, type IBM ; la fatalité, celle de la vieille politique. Là encore, Platon avait dit l'essentiel : le modèle de l'activité politique, c'est le tissage.

¹⁶ *Villes intelligentes, villes démocratiques ?* page 116.



Mayer-Schonberger et Cukier concluent excellemment leur étude du *big data* par ces lignes :

Les données que nous sommes capables de recueillir et de traiter ne seront jamais qu'une petite fraction de l'information existant dans le monde. Elles ne peuvent être qu'un simulacre de réalité, comme les ombres sur le mur de la caverne de Platon. Parce que nous n'aurons jamais d'information parfaite, nos prédictions sont par définition faillibles. Cela ne signifie pas qu'elles soient fausses, seulement qu'elles sont toujours incomplètes. Dire cela n'est pas refuser les perspectives nouvelles que nous offre le *big data*, c'est simplement le mettre à la place qui lui convient : celle d'un outil qui n'offre pas de réponses définitives, seulement des réponses suffisamment bonnes pour nous aider jusqu'à ce que nous en trouvions d'autres, meilleures, par d'autres méthodes. Cela signifie également que nous devons utiliser cet outil avec beaucoup d'humilité... et d'humanité.

Bibliographie

Sur le big data

Viktor Mayer-Schonberger, Kenneth Cukier : *Big Data: A Revolution That Will Transform How We Live, Work and Think* (2014, Kindle Edition).

Lejeune, Yannick (sous la direction de) : *Big, Fast & Open Data. Décrire, décrypter et prédire le monde : l'avènement des données*. FYP Editions (2014)

Fogel, Jean-François et Patino, Bruno : *La condition numérique* (Grasset, Kindle Edition, 2103)

Babinet, Gilles : *L'Ère numérique, un nouvel âge de l'humanité : cinq mutations qui vont bouleverser notre vie* (Le passeur, 2014, Kindle Edition).

Sur les smart cities

Picon, Antoine : *Smart Cities, Théorie et critique d'un idéal autoréalisateur* (Editions B2, 2103).

Ville intelligente, ville démocratique ? Actes (Colloque Berger-Levrault – Chaire MADP de Sciences Po, 13 février 2014).

Greenfield, Adam : *Against the Smart City (The City is Here for You to Use, Book 1)* (2013, Kindle Edition)

Sur l'économie à la demande

There's an app for that, The Economist (3/01/2015).

Approches plus générales

Flichy, Patrice: *Le sacre de l'amateur. Sociologie des passions à l'heure numérique* (Seuil, Kindle Edition, 2014).

Innerarity, Daniel : *Démocratie et société de la connaissance* (trad. Serge Champeau, Presses Universitaires de Grenoble, 2015).

Populismes : *Critique*, n° 776-777 (2012).