

ELITISME

ELITISME WORKING PAPER SERIES

ANR-ELITISME-2017-004

Big Data : Perspectives

André de Palma



Labex MME-DII

Modèles Mathématiques et Économiques de la Dynamique, de l'Incertitude et des Interactions.



Big Data : Perspectives

André DE PALMA

Les auteurs du XVIII^e siècle en savaient autant que nous sur les problèmes à résoudre. L'évaluation des facteurs qui affectent la santé publique, l'instruction et autres voies de la perfectibilité du genre humain ; les déviations mercantiles ; les pouvoirs faibles ; les informations lacunaires ; les perspectives lumineuses : tout cela leur était familier. En revanche, ils ne connaissaient ni la fibre optique, ni les ordinateurs, ni les données massives. Nous sommes mieux équipés.

On tentera d'abord de cerner les facteurs qui ont déclenché la ruée des données massives sur les machines. On détaillera ensuite leur impact sur les transports, et plus largement sur la nouvelle économie. La deuxième partie de ce texte examinera les ressources, les buts et les stratégies des protagonistes : Puissance publique, entreprises, opérateurs de transport, usagers, citoyens et institutions. Nous y verrons comment les données massives modulent les stratégies des parties prenantes dans la définition de la politique des transports. Pour illustrer l'exposé, on évoquera le parcours de la prise de décision sur l'implémentation de l'écotaxe poids lourds en France, terrain par excellence de la gestion de la datamasse¹ dans la politique des transports.

1. Gérer l'écologie de la datamasse

1.1. *Un tsunami sécrété par des animalcules*

Les masses de données ne datent pas d'hier. Les pratiques du recensement sont très anciennes, comme le montre Jacqueline Hecht dans sa contribution à *l'histoire de la statistique* de François Bédarida (1987). Les gouvernements ou leurs percepteurs des impôts (qui souvent n'étaient pas des fonctionnaires) en ont usé et abusé. Deux passages parallèles de la Bible racontent la tentation qui poussa le roi David à recenser ses troupes :

2 Samuel 24 :1 La colère de Yahvé s'enflamma encore contre les Israélites et il excita David contre eux : « Va, dit-il, fais le dénombrement d'Israël et de Juda » ;

1 Chroniques 21 :1 Satan se dressa contre Israël et il incita David à dénombrier les Israélites.

1. Pour ce terme entré dans l'usage, je renvoie à la première note de l'Introduction.

Il s'agit clairement d'imputations disputées ; et les exégètes ne peuvent s'accorder sur la question de savoir pourquoi c'était une faute, ce qui est remarquable.

Les bons commerçants sont toujours curieux des opportunités et attentifs aux risques. Mais les ouvrages grand public sur l'histoire du risque et des compagnies d'assurance ont la vertu de véhiculer des anecdotes qui confirment le lecteur dans l'idée qu'il sait déjà l'essentiel, alors que les relectures du passé sont riches de problèmes et de remises en cause. Les anecdotes sur la *coffee house* de l'assureur Edward Lloyd et sa *Lloyd's List* ne remplacent pas une analyse s'il s'agit d'expliquer l'institutionnalisation des compagnies d'assurance.

- *Ce qui a changé c'est la production de données*

Durant la seconde moitié du XX^e siècle, les ordinateurs ont fait chuter le coût du traitement des données ; ils se sont intégrés aux processus de paiement, de comptabilité, de gestion du personnel, de décision et de production de valeur. Pourtant, on observe que l'intérêt porté par l'ensemble des acteurs économiques aux données massives est largement postérieur à cette évolution. Par ailleurs, les super-ordinateurs en cours de construction ne seront accessibles qu'à une partie des nombreux acteurs qui se réclament de la vague du *Big Data*. Il faut donc une autre raison à ce déferlement que la seule montée en puissance des machines.

Il se trouve que de nombreux acteurs sont depuis peu capables de produire de telles données, à coût bas, et ce sans négliger leur activité principale (production et vente). Certains entendent faire de la collecte leur cœur de métier. Ils n'ont pas surgi du néant. Cette opportunité nouvelle est venue du développement des dispositifs de mesure et d'enregistrement, capteurs spécialisés ou transactions électroniques laissant une trace numérique exploitable.

- *Une inversion des media de masse*

Nous sommes donc ici devant une situation symétrique de celle que décrivaient les sociologues et les politologues des deux derniers siècles, quand les médias déversaient sur les masses, bientôt rebaptisées grand public, des produits faiblement diversifiés (Beethoven, Vivaldi, les ballets russes, les valse viennoises au Nouvel an). La comparaison est instructive.

Demain, nous aurons affaire à des myriades de petites machines, et souvent très petites, qui mettront en circulation dans l'Internet des flux proprement immenses de données. Une nouvelle culture de masse s'annonce, inverse de celle qu'on a abondamment critiquée. Elle s'est diversifiée dans ses contenus (Les tissages des Andes, les sculptures inuit ou les voix bulgares). Elle se cherche encore dans ses modes de mobilisation sociale autour d'une idée, d'une protestation ou d'une innovation. Parfois l'autorité diffuse d'une nouvelle *vox populi* s'invoque dans l'argumentation des commerçants, des candidats au suffrage universel, des experts auto-proclamés et autres nouveaux rhétoriciens.

Ce fait crucial de la production de données numériques susceptibles d'un traitement automatique se retrouve dans tous les domaines d'application des données massives : dans l'analyse des besoins et du profil des clients via la numérisation de leurs transactions (demandes d'informations ou achats conclus) ; dans le suivi des déplacements ; dans l'automatisation de procédés de surveillance pour la maintenance industrielle prédictive (maintenance, gestion de la production et de la qualité) ; dans la collecte automatique de données biologiques, géologiques, chimiques, physiques, voire extra-terrestres.

- *Données amorphes, brutes et formatées*

Mesurer le réel est une chose. Rendre le réel analysable par un programme informatique, c'est un autre défi, que relèvent le « web sémantique » et les « ontologies » (descriptions logiques d'informations). Que faut-il faire, en effet, pour qu'un ordinateur puisse lire des informations et les organiser ? La question est cruciale, car une fois organisées suivant des schémas logiques, les données peuvent être relues « automatiquement ». Il s'agit d'un passage technologique obligé où l'intelligence artificielle rencontre les données massives. On retrouve le rôle décisif joué ailleurs par les capteurs, tenu ici par de nouveaux logiciels d'analyse de l'information.

Ainsi, les données numériques représentant des textes étaient déjà disponibles, mais sous forme peu exploitable (le texte brut se prête à l'affichage, à l'indexage, à la statistique, mais pas à l'analyse sémantique et logique du contenu) ; il fallait les organiser en fonction de relations schématiques (les ontologies) pour pouvoir parcourir des phrases qui véhiculent une information donnée, et rédiger ainsi des énoncés qui correspondent à des connaissances. Une liste de valeurs de paramètres est un ensemble de données assez brutes, mais un texte non analysé est par comparaison beaucoup plus difficile à lire pour une machine, car au premier abord plus amorphe. Pour nourrir les ontologies il faut des grammaires. Ce programme de travail est en cours : par exemple, des robots extraient de la prose peu normée des pages de Wikipédia des données sémantiques simples du genre fiche d'état civil et parcours scolaire d'un personnage donné¹. On a fait le même type de travail depuis quelques décennies sur les faits divers ou les fiches météo.

1. C'est en repérant des intervalles de dates correspondant à des événements canoniques (naissance, accès au trône et décès d'un roi) que Tatiana Proskouriakoff a entamé le programme de déchiffrement de textes gravés sur des stèles mayas. Nos machines feront de même sur nos documents, avec plus d'atouts que l'archéologue.

On rapprochera ces exemples de celui d'un réseau de caméras qui surveillerait d'abord une foule sans avoir les logiciels qui permettraient d'identifier et de suivre chaque personne. Dans ce cas, ce sont la modélisation et les techniques d'analyse d'images qui couplent les caméras à des machines productrices de données logiques ; il s'agissait de données brutes (des images animées pour nous), et elles attendaient d'être analysées pour se transformer en informations utilisables par un programme informatique. Même à un homme il peut arriver, dans des conditions d'éclairage faible, ou quand il n'a pas encore accommodé sa vision à la distance, de voir sans comprendre ce qu'il voit. De l'amorphe vers les connaissances en passant par le brut et les données élémentaires, il existe un ensemble de procédures, qu'on peut comparer aux programmes de la chimie analytique du XIX^e siècle (Je renvoie à l'esquisse brossée dans le texte *Tour d'horizon et repérages* d'après les exposés de Serge Abiteboul).

1.2. Les réseaux sociaux comme sources de données

Les réseaux sociaux sont souvent associés aux données massives, et ce pour deux raisons : d'un côté, ils produisent énormément d'informations plus ou moins lisibles, et ont donc besoin de systèmes de traitement automatique ; d'un autre côté, ils ont prouvé leur utilité en tant que fournisseurs de données élémentaires et incomplètes, mais peu onéreuses. Certains vendeurs de navigateurs assistés par GPS tentent de développer entre leurs clients des échanges, parfois institutionnalisés sous la forme de réseaux sociaux. On parle parfois de *crowd-cruising*. Des solutions institutionnelles innovantes permettront à moyen terme de stimuler et d'encadrer ces initiatives pourvoyeuses en données.

Henri Verdier, Administrateur général des données auprès du Premier ministre, est bien placé pour montrer que ces processus complexes sont déjà amorcés (2015) ; reprenons-lui donc deux observations :

– « Les données produites par les administrations et les données produites en externe ne sont pas antagonistes, elles peuvent même se compléter. L'INSEE développe ainsi un projet de suivi de l'indice des prix à la consommation basé sur des tickets de caisse anonymes. De nombreux chercheurs ont montré que, soigneusement analysées, des données issues du web social peuvent apporter des éclairages essentiels à des questions extrêmement difficiles – ou coûteuses – à analyser sur le seul fondement de données d'enquête. »

– « Produire une donnée essentielle [bien définie via des dispositifs publics] n'est plus aujourd'hui la seule prérogative de l'État. ... La co-production des données essentielles avec la multitude n'est pas un scénario futuriste : le projet de *Base adresse nationale* est le fruit d'une collaboration entre l'Institut national géographique, La Poste, l'Administrateur général des données et l'association OpenStreetMap France ».

On n'oubliera pas, parmi les réseaux sociaux, le plus (mé)connu d'entre eux, l'Internet. Il est tentant de comparer ce dernier à une bibliothèque, puisque c'est une source d'information, et que c'est ainsi que nous l'utilisons. Mais sa structure première est bel et bien celle d'un réseau de sites web, chacun étant produit par une personne ou une collectivité plus ou moins bien définie (institution, entreprise, association...).

Et d'ailleurs, aujourd'hui encore, la satisfaction de nos recherches documentaires vient en partie du fait qu'une fois trouvé le bon site, c'est-à-dire la bonne source d'information, nous lui faisons confiance : il sera aussi complet que possible. Eugene Garfield, institution vivante de la documentation, expliquait volontiers qu'une recherche bibliographique peut renoncer à parcourir des chaînes d'articles qui se réfèrent aux uns et aux autres, et se contenter de construire des boules de noms propres de personnes se reconnaissant entre elles pour travailler dans le même domaine.

L'Internet a aussi joué un autre rôle, celui de faire apparaître non seulement de nouveaux *business models* inattendus auparavant, mais aussi de nouvelles façons de construire un service ou un produit : qui aurait pu penser que l'écriture de la plus parfaite des encyclopédies serait gratuite, et que les professionnels développeraient des outils pour extraire et structurer des données à partir d'expéditions dans ce nouvel Eldorado ? Il faut aussi noter que les miracles de l'Internet n'ont pas seulement leur valeur intrinsèque : ils ont aussi élevé les seuils d'aversion au risque chez les entrepreneurs et sur les marchés.

Quelles seront les conséquences de l'irruption des données massives dans la vie économique ? Certaines seront particulières à tel domaine d'application, d'autres seront générales.

1.3. Effets sur les transports

La mobilité sera certainement transformée dans les années qui viennent, mais ce sera sans doute autant, voire plus, par le développement des véhicules autonomes que par le traitement des données massives. Ces véhicules vont modifier la façon de se déplacer, et ce de diverses manières. Les véhicules autonomes privés offriront un grand confort et réduiront le stress des conducteurs ; ils pourront augmenter la vitesse moyenne de circulation et diluer les embouteillages. Mais ils vont aussi encourager les choix modaux vers la route, ce qui pourrait aggraver la congestion. Ces effets possibles doivent être étudiés. Pour les véhicules publics, les taxis et les nouveaux véhicules de tourisme avec chauffeurs (VTC), c'est le prix qui sera abaissé, surtout pour les deux dernières activités, qui sont intensives en main-d'œuvre.

Revenons aux données massives, qui joueront aussi un rôle majeur. Les transports en commun bénéficieront sûrement des possibilités offertes par l'usage des données massives, ne fût-ce que parce que leurs opérateurs sont souvent de grande taille, et pourront profiter d'économies d'échelle. Par exemple, l'analyse des déplacements sera bientôt moins chère et mieux documentée, grâce à de nouveaux dispositifs combinant enquêtes, recueil des traces de déplacements et traitement statistique. Ces dispositifs permettront, sur le moyen terme, d'ajuster l'offre à la demande, et de faire apparaître des demandes nouvelles. Dès aujourd'hui, ils permettent d'ajuster une offre adaptée à des événements exceptionnels (en jouant par exemple sur les horaires ou les nombres de voitures de train). On voit tout de suite que l'usage et la présentation des données dépendront souvent de chaque domaine particulier, et que leur interprétation et leur traitement ne seront donc pas nécessairement confiés aux seuls spécialistes informatiques des données massives.

- *Les grands nombres et les transports en commun*

Ceux-ci pourront aussi devenir plus réactifs, en mobilisant plus de véhicules là où la demande est plus grande. Mais l'usager des transports en commun souhaite des itinéraires fixes et des horaires fixes : le gain de temps importe souvent moins que la prédictibilité.

Cette règle de préférence a un précédent : en 1662 Blaise Pascal, mathématicien probabiliste devenu entrepreneur, impose sur le marché des transports parisiens en croissance rapide (500 000 habitants au milieu du XVII^e siècle) le transport en commun que nous connaissons. Il investit avec des capitaux privés. Des entrepreneurs de transport existaient déjà, plus proches de nos taxis. Pascal a évalué les risques, et les répartit selon un schéma innovant (Lundwall, 2000 ; Margairaz, 2012; De Palma, en préparation). La compagnie des « Carrosses à 5 sols » obtient un privilège royal sur cinq itinéraires fixes (dont l'un fait le tour du petit Paris du

XVII^e) ; elle prend le risque de travailler à prix fixe et horaire fixe. Les voitures partiront tous les « demi-quarts d'heure », le cas échéant avec un seul passager, qui paiera donc non pas le prix effectif de la course qui le transporte, mais un prix moyen précis et comparativement bas (cinq sols pour un segment donné ; on ne donnera pas le change aux pièces d'or pour gagner du temps).

Voilà donc un manège qui fonctionnait « quelque petit nombre de personnes qui s'y trouvaient, même à vide s'il ne se présentait personne ». Ce carrousel régi par une règle avait de quoi donner le tournis aux comptables, mais il était fondé en raison. Les pascaliens évoqueront avec Lundwall une hiérarchie du monde des affaires selon les Trois Ordres : le Corps avec le chiffre d'affaires, l'Esprit avec la créativité, la Volonté avec la stratégie. Les calculateurs s'émouvront de voir une machine partie d'un si bon pas.

Ce schéma nous semble banal, mais il était inconnu à cette date, et s'avéra profitable. Nous n'avons pas les commentaires d'exploitation de Pascal, qui devait mourir quelques semaines plus tard d'une longue maladie. Mais nous avons des remarques de sa sœur Gilberte, qui regrette le trop petit nombre de véhicules face au succès de la formule : elle a laissé passer cinq voitures pleines avant de pouvoir monter. L'entreprise échouera quinze ans plus tard, victime des contradictions de l'Ancien Régime¹, et la formule ne reprendra cours qu'au XIX^e siècle. Les grands opérateurs de transport ajustent aujourd'hui leurs investissements et leurs horaires à une étude fine de la distribution de la demande, sans négliger sa dimension probabiliste. On sait que d'autres schémas d'ajustement offre / demande existent. Du Maroc à l'Indonésie, de nombreux pays présentent des formules intéressantes qui sont autant d'expériences à commenter.

Un mot sur les critères de décision touchant les investissements. Aucune firme informatique n'investira sans garantie sérieuse les sommes nécessaires pour assumer sur moyens propres, par exemple, la surveillance du trafic, car pour elle ce marché est marginal, aléatoire, et peu profitable, ce qui n'exclut pas des interventions de démonstration de savoir-faire qui dépassent les effets d'annonce. Le problème sera toujours de transformer à terme une intervention, parfois localement gracieuse, en transaction avec des partenaires solvables, quelle que soit la complexité du modèle économique.

Il reviendra donc aux organismes de gestion du réseau, du trafic ou de la sécurité d'utiliser les données massives pour remplir mieux leurs fonctions propres, probablement en sous-traitant des segments choisis des nouvelles activités. Les médias peuvent publier les dossiers d'un marchand de systèmes de navigation assistés par satellite sur l'engorgement des voies de circulation. Une fois noté que ces données sont en fait souvent assez minces, cela n'enlève rien à leur intérêt. On va sans doute vers des partenariats public-privé.

- *Les marchés bifaces*

La solution du transport partagé sera favorisée par la mise en contact de l'offre et de la demande via les réseaux sociaux (commerciaux ou associatifs). Elle diminuera la quantité de véhicules sur les routes, tout en offrant un confort voisin de celui de la voiture individuelle (contre moins de liberté de choix d'itinéraires et d'horaires).

On peut penser que la demande suffira à susciter l'offre ; mais si aucune entreprise n'investit il reste le pouvoir d'initiative, d'incitation ou d'encadrement de la puissance publique. Le marché existe, et intéresse un certain nombre d'acteurs.

Pour rendre la rentabilité moins mystérieuse, il faut retenir le modèle des marchés bifaces de Jean-Charles Rochet et Jean Tirole: « Une structure de prix particulièrement intéressante apparaît dans les marchés appelés “ bifaces ” (*two-sided markets*), dans lesquels des “ plateformes ” rassemblent de multiples communautés d'utilisateurs souhaitant interagir les uns avec les autres : les joueurs et les développeurs de jeux dans le cas de l'industrie du jeu vidéo ; les utilisateurs et les développeurs d'applications dans le domaine des systèmes d'exploitation ; les utilisateurs et les annonceurs dans le cas des moteurs de recherche ou des médias ; les détenteurs de carte bancaire et les commerçants dans le cas des transactions par carte de paiement. Le défi lancé

¹ Les lettres du roi ne filtraient pas selon la qualité des usagers, mais les juristes, rois anoblis, tenaient à marquer leurs privilèges, et leur Parlement n'accepta d'enregistrer ces lettres qu'en refusant les « soldats, pages, laquais et autres gens de bras ». Après un succès initial notable, l'entreprise succomba à cette coûteuse contrainte.

par les marchés bifaces est de trouver un modèle économique viable garantissant la participation de chacune des communautés d'utilisateurs » (Tirole, 2014).

- *Cycle de l'information sur le trafic*

En ce qui concerne la gestion du trafic, les informations en temps réel servent d'abord les usagers. Elles permettent de gérer les embouteillages, et donnent aux agences de sécurité routière des moyens nouveaux.

Aujourd'hui l'observation directe du trafic intra-urbain fait aujourd'hui appel aux motards ou aux hélicoptères (avec son allure tellement XX^e, le motard a encore un bel avenir, car la taille de son véhicule et sa puissance font qu'il dispose toujours d'une issue, comme s'il échappait au plan auquel sont confinés les autres véhicules). Ces équipes d'observation fournissent des informations aux centrales d'information routière et aux radios locales pour avertir les usagers, lesquels les alimentent en retour *via* le téléphone ou les applications internet.

Ces expériences recourent à divers capteurs fixes ou mobiles, et affectent du personnel compétent à l'évaluation des données ainsi produites, croisées ou non avec des informations OD (origine-destination). L'information ainsi obtenue est diffusée par les centres de contrôle du trafic ou les médias via des panneaux routiers digitaux, ou *via* les assistants électroniques de conduite des véhicules. On verra sans aucun doute se dégager des asymétries fortes dans ce marché en pleine structuration.

- *L'usager décide, mais il est encadré*

Le traitement automatique de données massives amènera sans aucun doute de grands changements dans le domaine des transports. Plusieurs modèles existent, sur la base des compétences de grands gestionnaires de connaissances numérisées comme Google, d'applications de navigation appuyées sur la communauté des usagers (Waze), ou de collaborations plus complexes parfois soutenues par les grands réseaux de transport (RATP).

Mais il sera plus prudent de formuler les objectifs en termes d'une plus grande efficacité des travailleurs du secteur que de chercher à introduire « pour elles-mêmes » des innovations extrinsèques. À l'heure actuelle, ces données et leur traitement n'ont finalement d'intérêt que lorsque l'usager en prend connaissance pour poser ses choix, et c'est très bien ainsi.

Les flottes de véhicules assistés par ordinateur sont pour bientôt. Les données massives traitées dans ces circuits n'empêcheront pas les usagers de prendre leurs décisions d'itinéraire – les machines peuvent même les aider à éviter les comportements grégaires qui piègent les mieux renseignés, selon le paradoxe bien connu.

Enfin, des incitations peuvent tenter de modifier leurs choix de temps de départ, d'itinéraire et de mode de transport. En soi, de telles mesures posent les mêmes questions que l'organisation de périodes décalées des vacances d'été, qui présente en soi de nombreux avantages. On retiendra qu'elles pourraient poser d'autres questions. Depuis la ceinture de sécurité dans les voitures, qui introduisit un précédent, chacun comprend qu'une norme légale peut viser non seulement les dommages à autrui, mais aussi la protection du corps propre de chacun contre les accidents. En ce sens, il n'y a pas de limite à la gouvernance des corps. Faudra-t-il que l'usager signe une décharge de responsabilité s'il sort d'une plage horaire ? Et s'il en accepte une ? Nous reviendrons à ces problèmes dans notre paragraphe 2.3.

1.4. Effets sur l'économie

Au niveau des conséquences générales, les outils de collecte, de transmission et de traitement de données massives permettront d'orienter le chemin du progrès technologique en abaissant le coût de certaines innovations. Ainsi, disposer, par exemple, des capteurs sur un équipement est avantageux en soi, pour abaisser par exemple le coût de la maintenance, de la gestion ou de la sécurité. Mais cela ouvre aussi la voie à d'autres innovations, par exemple à d'autres usages de ces capteurs ; à certaines étapes du processus d'innovation, la maîtrise des outils constituera donc un facteur limitant décisif. Les fabricants de pneumatiques développent des capteurs qui permettent de construire des connaissances sur l'état de la route et la sécurité du véhicule, mais qui peuvent servir à étudier d'autres informations, comme le style de freinage du conducteur qui pourrait intéresser les assureurs.

En ce sens, le traitement technique des supports des données massives représentera une compétence-clé pour profiter des opportunités de croissance actuelles. Pour générer et stocker les données massives, de nouveaux acteurs économiques vont se mettre en avant, dont le cœur de métier sera l'informatique. Ils pourront proposer des services de stockage, de traitement et parfois d'analyse. Mais, dans le cas général, on peut s'attendre, comme ci-dessus, à ce que l'analyse et les décisions qui en découlent soient effectuées par des acteurs directement intéressés qui connaissent leur terrain et qui préféreront garder le contrôle sur la prise de décision à toutes les étapes du traitement proprement dit.

Un autre exemple est celui de la recherche pharmaceutique. Celle-ci est menée par des entreprises, des organismes publics et par des universités. Sauf exception, les entreprises ne publient pas intégralement leurs recherches récentes, et ne vont donc pas se contenter de confier à des contractants l'analyse des bases de données biomédicales, car il s'agit de leur cœur de métier : il y va de leur capital de compétences et de leur indépendance. Les universités, qui ne sont pas uniquement motivées par la concurrence, ont d'autres intérêts ; et elles ont souvent les compétences pour faire ce travail elles-mêmes. Elles sont donc à même d'évaluer les avantages respectifs du traitement direct et de la collaboration, ce qui nous ramène aux énoncés fondamentaux de Ronald Coase (1937) sur le coût des transactions internes et externes à l'entreprise, qui décident de sa raison d'être. Plusieurs formules sont envisageables ici.

Un autre phénomène lié aux données massives, qui se prête moins bien aux prédictions, est le développement de réseaux sociaux à vocation mixte : entre le service aux clients et le partage égalitaire de connaissances. On n'oubliera pas les plateformes (commerciales) de *crowd-funding* et de marchés de brevets qui, en concentrant une information jusque-là dispersée, offrent un service nouveau aux investisseurs.

Les données massives, moyen faiblement défini en quête de fins opportunistes, appellent donc ce qu'il ne faut pas craindre d'appeler un « bricolage institutionnel ». Elles ont, nous le comprenons mieux, un rôle important à jouer dans l'implantation d'innovations technologiques : elles pourraient orienter l'innovation vers telle ou telle direction efficiente et/ou rentable (elles sont peu coûteuses à expérimenter et à mettre en place) ; et la particularité des systèmes de gestion des données massive (SGDM) sera de produire une information compréhensible à partir de données trop riches pour les équipes qui n'auront pas accès aux super-ordinateurs.

Pendant la phase de transition, leur rôle sera souvent limité aux processus d'innovation technique, et elles affecteront moins immédiatement les activités de planification de la production ou de vente. Mais la phase de transition n'aura qu'un temps, et les données massives d'aujourd'hui alimenteront les grands tableaux de gestion demain.

On conçoit que les équipes capables d'articuler les diverses sources d'informations structurées auront un capital de compétences à valoriser lorsque commencera l'exploitation systématique de cet univers complexe. Les données massives sont trop riches pour être laissées à des opérateurs mal préparés aux nécessaires collaborations.

2. Stimuler les interactions entre les parties prenantes

Il est remarquable que nous devions à la littérature des sciences politiques et économiques du monde anglo-saxon de nous avoir appris à identifier les parties prenantes, les *stakeholders*. Mais il reste à les étudier autant qu'il le faudrait.

On néglige parfois d'étudier les interférences de leurs stratégies, ce qui impose un retard à la recherche scientifique (les universitaires ont d'autres travaux en chantier), parfois un manque-à-gagner pour l'État, commanditaire des études et garant de l'intérêt public. Il devrait être évident que le jeu collectif de la prise de décision politique est lui aussi un objet d'étude ; et un pays trop habitué aux confrontations politiques néglige peut-être d'étudier ces stratégies et leurs conséquences. Un ensemble très riche d'interactions passe ainsi inaperçu, alors que dans les discussions de corridors on évoque fréquemment le sujet.

Comme dans d'autres domaines, l'intérêt bien entendu de la politique est de prendre en compte, au moins autant qu'il semble possible, l'ensemble des partenaires concernés par les données massives. À titre d'exemple, nous reprendrons l'un ou l'autre épisode de l'implémentation ajournée de la désormais fameuse écotaxe poids lourds (2013-2017)¹.

2.1. Le cas de l'écotaxe poids lourds

Le dernier rapport annuel de la Cour des comptes critique l'état inabouti de l'implémentation de l'écotaxe poids lourds. Nous ne discuterons pas les arguments, ni les critères et les finalités, ni ceux de la réponse du Premier ministre, qui figure en annexe². Notre approche est centrée sur la concertation des parties prenantes comme critère et garantie de l'efficacité du processus de prise de décision politique.

Ce dossier est pertinent pour nos réflexions en raison du rôle intensif qu'y jouent la collecte et la gestion des données massives, sans lesquelles il n'aurait été ni possible ni économiquement praticable d'introduire une taxation au kilomètre pour les camions qui empruntent les voies visées par l'écotaxe poids lourds. Dans le cas

1. Le lecteur pour qui la discussion sur l'internalisation des externalités négatives évoquerait le marxisme austro-hongrois ou le tantrisme tibétain peut commencer par les dossiers d'un chroniqueur du journal *Les Échos*, Lionel Steinmann, qui a suivi le processus de prise de décision politique dans toute sa complexité ; voir aussi les remarques de la politologue Marianne Ollivier-Trigalo (2013).

2. Nous y renvoyons le lecteur, parce que ses auteurs donnent une évaluation de ce que nous cherchons à comprendre sur le processus de prise de décision politique (Cour des comptes, 2017). Pour le cadre théorique, on peut consulter de Palma *et al.*, 2007. Pour une analyse et une évaluation de la taxation des camions en Europe quant à ses objets (carburant, distance, forfait à terme temporel) et de ses objectifs (financement, report modal, réduction de la pollution et autres), mais aussi de sa dynamique contagieuse de compétition fiscale, voir l'article passionnant de Mandell et Proost (2016).

français, il a été avancé que la taille du réseau et le nombre de véhicules justifient le fait qu'on ait recouru aux services de sociétés spécialisées dans les données massives, dont certaines utilisent leurs propres technologies (Total, Axxès, Eurotoll), au lieu de confier le travail aux administrations des douanes et des transports. Le nouveau type de taxe recourait à des capteurs divers (équipements embarqués, portiques, bornes de contrôle mobile) et à des serveurs informatiques de concentration et de calcul des données.

- *Des objectifs politiques multiples*

Les directives *Eurovignette* (1999/62/EC, 2006/38/EC et 2011/76/EU) définissent progressivement un cadre commun pour la tarification des infrastructures routières fondé sur les principes utilisateur-payeur et pollueur-payeur¹. Cette politique n'est pas d'abord ou uniquement une norme de l'Union Européenne : depuis l'an 2000, plusieurs pays européens complètent la fiscalité sur les carburants en taxant la longueur du parcours effectué sur certaines routes visées. Auparavant, de nombreux pays appliquaient une taxe annuelle forfaitaire aux camions qui parcourent leur territoire. Après 2000, certains ont pris en compte la rentabilité d'une taxe au km calculée en fonction des émissions et du poids à l'essieu, en fait utilisée surtout pour équilibrer le Trésor public, ou plus directement les fonds d'investissement des infrastructures routières et de la gestion du transport. Le mot *écotaxe* ne doit donc pas tromper : il s'agit de fiscalité, de budget, de compétition fiscale, et aussi d'invasion par des camions que dissuade une taxe/km plus élevée dans un État voisin (par exemple, l'Autriche est un pays de transit parallèle à la Suisse). La Cour des comptes rappelle que l'*écotaxe poids lourds* peut constituer à terme un outil de fiscalité environnementale, en incluant progressivement dans son barème le coût des externalités négatives produites par le transport routier de marchandises, mais qu'en début d'exercice, il s'agit surtout de financer les infrastructures, en équilibrant la contribution des camions étrangers.

Un des buts explicites de ces initiatives était de contrôler l'explosion du transport routier et ses externalités négatives : congestion, dégradation des routes, pollution. De nombreuses études, appuyées sur des données de divers types, ont permis de valider cette démarche, qui internalise sur les usagers les plus responsables (camionneurs de transport) le coût de certaines externalités négatives : en particulier, l'usure de la route est une fonction non linéaire du poids par essieu². Parmi les objectifs figurait un report modal vers les voies d'eau ou le chemin de fer, voté par le peuple suisse depuis plusieurs décennies.

- *La contagion de la taxe au kilomètre*

L'explosion récente du transport routier est liée, forcément, à des calculs de rentabilité. La taille des réservoirs permet de payer la taxe diesel dans un petit pays bien situé pour parcourir un autre pays sur 1 000, voire 2 000 km, tandis que les taxes au kilomètre sont payées au pays parcouru. Par suite, le pays qui introduit une taxe au kilomètre peut réduire ses taxes carburant, et menacer ainsi la base fiscale des voisins qui n'auraient pas introduit de taxe au kilomètre (Mandell et Proost, 2016). Car la compétition fiscale sur les carburants est une dure réalité. En 2012, l'Allemagne taxait le carburant diesel à 47 eurocentimes par litre, tandis qu'un petit pays, voisin de

1. Pour l'histoire de la négociation de l'*Eurovignette*, voir Weismann (2013).

2. En France, le ministère des Transports estimait en 2003 que les poids lourds couvrent 66 % de leurs coûts d'usage sur le réseau routier national non concédé, et 23 % sur les réseaux routiers locaux.

l'Allemagne, de la Belgique et de la France, appliquait le minimum communautaire négocié par l'Union : 30 eurocentimes (soit un surcoût de 50% pour le transporteur).

Il existe une dynamique géographique de ces nouvelles taxes, que montre une carte de l'Europe. C'est un cas exemplaire de diffusion. Au départ, les pays périphériques (Royaume-Uni, Pays-Bas, Belgique, Luxembourg, Finlande et pays baltes, Suède, Hongrie, Roumanie, Bulgarie) n'ajoutent rien à la taxe carburants, sinon parfois la vignette. Ce sont des pays du centre (Suisse, Autriche, Allemagne, Tchéquie, Slovaquie, Pologne, Belgique) qui sont infectés : quand un pays (la Suisse au départ) introduit la taxe au kilomètre, il force son voisin à l'adopter malgré ses coûts de transaction plus élevés. Je renvoie le lecteur au modèle de Mandell et Proost (2016), qui ont entrepris leur travail pour lire cette carte, et à leur discussion sur sa confrontation avec la dynamique observable. Dans ce cadre de lecture, l'Eurovignette était en un sens une première réponse aux initiatives de pays comme la France, l'Espagne ou l'Italie, qui soumettaient à péage l'utilisation par les camions de certaines voies.

- *La taxation des camions en France*

Depuis 1968 la France appliquait une taxe annuelle aux camions français, et finançait certaines voies par le péage (les « autoroutes à péage », dont le statut et la gestion évolueront). Elle avait aussi depuis des décennies instauré des taxes dites à l'essieu, qui n'affectaient que les poids lourds immatriculés en France (soit bon an mal an la moitié du trafic camions), et ignoraient forcément les récentes techniques de gestion des données massives.

Nous verrons à présent ce qu'il est advenu d'un programme longuement mûri par des économistes des transports, et porté par des groupes de parties prenantes convoqués par un gouvernement (accords de Grenelle de l'environnement 2008). Il s'agissait donc d'un processus de prise de décision politique exemplaire, sanctionné par un vote quasi unanime du Parlement (loi Grenelle 1 de 2009). On a reconnu les premières parties prenantes au dossier. Nous verrons à présent la précarité de politiques d'intérêt national quand l'État échoue à concilier de manière satisfaisante les intérêts en présence.

Le point qui frappe d'abord est la perte subie par des entreprises françaises qui avaient pris le risque d'un programme de travail spécialisé dans la gestion des données massives. Leur mission était de suivre par satellite 800 000 camions, dont 250 000 étrangers, sur un réseau de 15 000 km, dont 10 000 km de routes nationales (soit 90 % du réseau dont l'État a la charge, hors réseau autoroutier concédé) et 5 000 km de routes locales (soit moins d'1 % du réseau local). Ce contraste, où l'écotaxe poids lourds tarifait l'usage de beaucoup plus de voies que n'en compte le réseau d'autoroutes à péage, est une spécificité française.

L'implémentation d'une écotaxe poids lourds sur une partie du réseau européen était considérée dans les milieux de la Commission comme un cas exemplaire d'application massive des Nouvelles Technologies de l'Information et de la Communication. C'était spécialement le cas, dit la Cour des comptes, avec le projet du gouvernement français. Mais on notera qu'au moment de la confrontation politique, inévitable quand des intérêts distincts sont entrés en compte (2013-2014), le système français n'était pas toujours bien préparé à la compatibilité avec les autres pays de l'Union Européenne (voir l'évolution de la firme Axxès).

2.2. La faiblesse publique

- *La datamasse appelle une autre politique*

Reprenons un peu de recul. L'apparition de la chaîne de collecte, de traitement et d'emploi des données massives permet d'enrichir et de différencier, selon les domaines, la liste des fonctions de la Puissance publique. Reprenons une fois encore le rapport de l'Administrateur général des données (Verdier, 2015, p. 7). Celui-ci donne une liste fournie de domaines où l'analyse prédictive des données massives peut jouer. On y trouve « les vols de voitures ; les temps d'attente aux urgences médicales ; les contrôles douaniers ; le chauffage des immeubles ; les offres d'emploi ; les remplacements d'enseignants ; les feux de circulation et la fluidité du trafic ; le calcul des prix des médicaments ; les pics de consommation d'électricité ; la négociation des achats publics ; les effets d'une réforme fiscale ; les besoins d'investissement médical par l'analyse de la littérature scientifique ».

Nous nous intéressons ici au transport. Mais cette liste donne à réfléchir. Dans tous ces domaines d'intervention, l'État peut inciter, réguler, normer, sanctionner, prendre l'initiative ... Le débat politique sur son rôle en dehors des missions régaliennes est en passe de devenir plus complexe, en raison directe de la complexité de la poursuite des objectifs considérés comme souhaitables par des acteurs de plus en plus nombreux. En particulier, on voit que les interventions de l'État ne seront pas uniquement fonction des politiques sectorielles, puisqu'elles seront aussi liées à une volonté globale d'activer les initiatives des parties prenantes en matière de données massives. On pressent les défis et les enjeux de cette orientation, dont l'ambition forte rappelle les prophéties qui annonçaient dans les années 1970 que l'ordinateur modifierait bientôt tout l'organigramme des entreprises et des administrations. Ollivier-Trigalo (2013) suggère que des actions de ce type peuvent contribuer à créer un modèle étatique nouveau, l'« État en interaction ».

- *Les pouvoirs locaux*

À cette complexité croissante qu'implique le Rapport Verdier, on ajoutera l'évolution de l'intrication des niveaux de pouvoir en fonction du jeu des diverses majorités électorales locales et non locales. Il existe des travaux sur le rôle des politiciens régionaux dans les politiques de transport. Nous renvoyons aux travaux de Stef Proost¹. Dans le cas de Bruxelles qu'il a étudié, les prérogatives des divers exécutifs concernés ne facilitent pas la définition d'une politique cohérente, puisque la ville et les communes fixent la tarification du stationnement le long des voies vénales, la région est compétente pour un péage à l'entrée dans la zone, et le gouvernement fédéral national ajuste les taxes sur les carburants. À cela s'ajoutent les contraintes du calendrier des politiciens, et les limites propres des taxes à la Pigou, qui de manière générale visent à compenser le dommage marginal des externalités négatives. Ce domaine appelle des analyses diversifiées.

Nous voyons que la loi est universelle, mais que son application locale fait l'objet de négociations entre niveaux de pouvoirs et acteurs privés, lesquelles parfois échouent. L'Alsace, soumise depuis les années 2000 à un effet de détournement de camions quittant le réseau allemand, devait servir de région-test à l'écotaxe poids lourds.

1. Voir Proost in : De Palma *et al.*, 2007 ; De Borger et Proost, 2012) ; et le rapport de Stef Proost et Jacques-François Thisse (2015) : *Skilled cities, regional disparities and efficient transport*.

Des décotes tarifaires furent ensuite négociées pour prendre en compte les spécificités de région : les débats ont dans un premier temps accordé 25 % au seul Finistère, puis davantage à la Bretagne dans son ensemble (40 %, puis 50 % en 2013) et enfin 30 % à l'Aquitaine et au Midi-Pyrénées. La rétrocession aux départements a aussi été fortement modulée (40 % en région Centre et en Picardie, 2 % en Nord-Pas-de-Calais).

Bien que le transport soit par définition un enjeu non local, l'intérêt des politiciens locaux peut rencontrer l'intérêt général. C'est le cas ici : des politiciens régionaux ont tenté de relancer une prise de décision nationale paralysée. On retiendra dans ce contexte une tentative récente du député-maire de Boulogne-sur-mer Frédéric Cuvillier (*Les Échos*, 10 août 2105) : « J'ai plaidé dès l'an dernier pour que le principe de l'écotaxe, qui a dû être abandonné au niveau national, puisse renaître dans les régions » (l'auteur a géré le portefeuille des transports dans le Gouvernement national). Cette proposition a rencontré un écho certain dans plusieurs régions, dont l'Alsace et Auvergne-Rhône-Alpes, mais n'a pu s'imposer.

- *Une négociation inaboutie*

On a aussi dû prendre en compte des spécificités de réseau. Le réseau autoroutier concédé aux sociétés de gestion privées, déjà soumis à péage, ne pouvait être partie au dispositif de l'écotaxe poids lourds dans le cadre juridique et politique retenu. Qu'il suffise de dire qu'on n'a pas pu mettre autour de la table l'ensemble des acteurs intéressés d'une manière ou l'autre par le financement des infrastructures routières et la gestion des redevances, alors qu'en 2003 on avait utilisé le produit de la concession par l'État de la gestion des autoroutes aux entreprises retenues par le programme. Cette concertation incomplète fut lourde de conséquences. D'autres équilibres juridico-financiers ont eu raison de l'écotaxe poids lourds. Lors des accords de Grenelle, les organisations de transporteurs routiers avaient conditionné leur accord à la construction d'une parfaite neutralité pour ne pas affecter la compétitivité de leurs membres. Or, l'État n'a pas trouvé le moyen de reporter sur leurs donneurs d'ordres l'augmentation du prix du transport, créant ainsi les conditions de la remise en cause globale du dispositif par les transporteurs.

La suite des déboires politiques, légaux et administratifs, ne nous intéresse pas ici. Une mission d'information de l'Assemblée nationale et les tentatives des administrations des douanes et des transports pour la mise en place d'un péage de transit poids lourds n'ont pu sauver le dispositif, même sous une forme réduite¹.

2.3. Les entreprises et leurs divers objectifs

- *Les sociétés spécialisées*

Le rapport de la Cour des comptes donne un schéma des flux financiers annuels liés au projet de mise en œuvre de l'écotaxe poids lourds. L'agence de financement des infrastructures de transport de France fut le grand perdant. Mais on retrouve parmi les acteurs lésés des sociétés privées spécialisées dans la gestion des données massives, les sociétés habilitées de télépéage. Il s'agissait d'acteurs nouveaux, innovants et flexibles, spécialisés dans la collecte et le traitement des données. Il est important de souligner que ces acteurs faisaient dans le projet

1. On notera entre autres que les serveurs informatiques prévus pour l'exploitation des données massives de gestion de l'écotaxe ont été revendus à 2 % de leur valeur.

du gouvernement l'objet d'une politique volontariste de déploiement industriel « à forte ambition technologique ».

En fait, selon la Cour des comptes, le dispositif français était le premier à « permettre le développement de l'interopérabilité des systèmes de tarification routière à l'échelle européenne : les équipements embarqués français étaient en effet compatibles avec les systèmes de télépéage de plusieurs pays européens ». À ce titre, ajoute-t-elle, cette chaîne industrielle de collecte, de traitement et d'usage de données massives était observée comme un modèle par la Commission européenne. Cette analyse est intéressante pour comprendre certaines motivations de la stratégie industrielle. Mais en réalité le système français n'était alors pas très avancé, et on ne pouvait raisonnablement pas fonder sur lui beaucoup d'espairs au niveau de l'exportation de savoir-faire et de système à l'échelle européenne. Il est remarquable que le gestionnaire de datamasse Axxès compte parmi ses membres fondateurs les plus importantes sociétés concessionnaires d'autoroutes françaises. Il y a bien là une stratégie industrielle et financière.

Le cas de l'affectation des ressources financières dégagées du péage autoroutier par les sociétés françaises d'autoroutes est plus facile à exposer que celui de l'écotaxe poids lourds. Il illustre la conclusion de James K. Galbraith (2009) : certains milieux économiques et politiques considérés jadis comme peu favorables à l'intervention étatique ont appris à s'y accommoder, puisque celle-ci leur assure des rentes qu'ils ne pourraient garantir sans elle. Il convient dans ce contexte de distinguer plusieurs types d'acteurs du privé.

- *Entrepreneurs et investisseurs*

Les uns sont des entrepreneurs. Une fois prise en compte, la localisation de leurs sites de fabrication et de vente ainsi que leur type de marchandise (plus ou moins périssable, volumineuse, valorisable sur le marché du frais, sans négliger la valeur ajoutée par tonne), il est concevable d'évaluer les effets de la taxe à l'essieu sur leurs coûts. Ce sont des parties directement prenantes à la politique que nous évaluons.

D'autres sont des gestionnaires ou des investisseurs. C'est l'occasion de renoncer à des analyses qui verraient dans la maximalisation du profit le critère unique des décisions stratégiques de tous et chacun de ces différents acteurs.

Que la dimension spatiale joue plusieurs rôles complexes dans les externalités des entreprises est évident. La protestation bretonne est liée à des intérêts privés que le gouvernement n'a pu rencontrer. Elle se focalise en 2013-2014 autour des portiques installés pour la gestion de la taxe ; une dizaine sera démontée et une quinzaine sera vandalisée. Puisque nous parlons du point de vue des gestionnaires des données massives, il faut noter que ces portiques ne servaient qu'à l'enregistrement et au contrôle des véhicules¹, et non à l'observation du trafic, qui relève d'un traitement intensif de données massives.

Il convient de ne pas négliger la typologie des acteurs économiques en présence et la diversité, parfois la contrariété de leurs intérêts et de leurs objectifs. Ainsi, certains entrepreneurs seront moins sensibles à la pression de leurs actionnaires, voire tentés de se passer d'eux dans certaines circonstances. Si un consortium industriel se propose de prendre le contrôle d'un autre pour renforcer sa masse propre et négocier d'égal à égal avec ses partenaires, il peut faire l'objet de tentatives de blocage par une partie de ses actionnaires, qui ne voient

1. Contrôle de la présence d'un équipement capteur embarqué dans le poids lourd et contrôle de la cohérence des informations enregistrées dans cet équipement embarqué avec les caractéristiques mesurables du poids lourd.

pas pour eux l'intérêt de l'opération. Pour parler en termes abstraits, on a là des critères de décision distincts qui s'investissent dans des acteurs distincts. Le récent rapprochement de Safran et Zodiac est un cas d'école de cette contrariété.

- *L'écotaxe poids lourds en France, un essai à reprendre*

Pour conclure ce paragraphe, nous voyons combien sont pertinentes les remarques de la Cour des comptes : l'écotaxe poids lourds appelait un déploiement de compétences innovantes et exportables dans la gestion de la datamasse par un secteur compétitif du tissu industriel, et ouvrait un domaine nouveau destiné à la gouvernance publique : là était, et demeure, une opportunité majeure d'intervention dans le nouveau monde des données massives.

Il n'y a aucun doute que cette tentative perfectible figure bientôt dans la liste des nouveaux terrains reprise à l'Administrateur général des données au début du paragraphe 2.2. Dans tous ces domaines, qui sont autant de cas similaires nous l'avons dit, l'effet d'une prise en compte des données massives est susceptible d'améliorer des pans entiers de politiques sectorielles d'intérêt majeur. Elle permettra une montée en puissance de notre savoir-faire politique : une évolution favorable de l'aptitude de l'État à nouer avec les parties prenantes des rapports constructifs.

- *Extrapolations, mobilisations, concertations*

Mais il y a une leçon à tirer des échecs comme tels. La prédiction reste un genre difficile. Mais nous pouvons nous aider d'une prise en considération de trois modalités du discours sur le futur. Nous en avons mentionné deux dans le document *Tour d'horizon et repérages*, qui sont familières à chacun : l'extrapolation d'un passé considéré comme inertiel, et le choix stratégique d'un futur qui cherchera ses moyens dans le présent. Des épisodes comme la tentative ajournée d'implémentation de l'écotaxe poids lourds rappellent qu'il en existe une troisième.

N'est-il pas possible de sortir de cette alternative ruineuse entre l'image du bateau qui court sur son erre jusqu'à heurter le quai, et celle de la manœuvre volontariste, qui mobilise à grands cris les ressources disponibles pour ne pas manquer la cible ? La déduction qui extrapole et le choix stratégique ne sont pas nos seules ressources. Un regard simple sur le présent suggère qu'une meilleure sensibilité à ce qui s'est manifesté lorsque la décision collective n'a pas pu concerter toutes les parties prenantes pour renégocier leurs légitimités respectives pourrait peut-être permettre de rencontrer sans arrière-pensées leur pluralité.

Faut-il prétendre connaître les futuribles ? Peu importe que nous sachions si dans dix ans la datamasse aura noyé la sécurité sociale ; si les motards seront remplacés par des drones ; ou si l'engagement de notre population active dans des entreprises qui vaillent sera devenu impossible. Plus important est de mieux sentir qui nous sommes et où nous sommes au présent. Les nœuds qui sous-tendent les contraintes du futur se serrent sous nos yeux. Renoncer à mettre en présence autour d'une table tous les protagonistes du financement des transports, c'est s'imposer de nourrir le symptôme d'une société bloquée.

2.4. Les citoyens et les institutions

- *Loi, contrat et institution*

Il convient ici de rappeler les grandes lignes d'une classification éclairante des sources de norme, due au juriste Maurice Hauriou¹ : les Lois obligent tous et chacun. Les Contrats engagent des partenaires. Les Institutions agissent et interagissent en fonction de la mission et des moyens qui leur sont assignés.

C'est l'agrément qui valide les institutions. L'acceptation des usagers est toujours cruciale dans la gestion du transport routier, comme il se voit dans le respect tout relatif des normes de sécurité (vitesses excessives, manœuvres perturbantes).

On retiendra donc les usagers et le contribuable parmi les parties prenantes. Or, les modalités usuelles de la prise de décision conduisent au fait, de plus en plus mal reçu par eux, que celle-ci, même quand elle touche leurs droits et leurs intérêts, se déroule quelquefois à leur insu. D'où les insatisfactions multiples que secrète un monde où la gestion des flux de données, comme celles de l'information ou du pouvoir de fixer les prix, correspond à des contraintes asymétriques.

- *Le service au Public*

La loi prévoit pourtant que le particulier a accès aux données le concernant, pour protéger sa vie privée, mais aussi « pour connaître les biais qui peuvent être induits par la collecte d'informations [le] concernant », comme le dit Nicolas Curien dans sa *Préface* ; et aussi pour permettre des interventions qui exploitent la portabilité des données, laquelle « stimule la concurrence entre les acteurs du marché, au bénéfice des consommateurs » (ibid.). Les parties prenantes ont en effet des intérêts parfois convergents dans ce type de situation. Encore faut-il bien faire ses comptes, et bien diffuser les résultats et les conclusions, si l'on veut favoriser une prise de décision toujours difficile.

Un recueil de Dominique Margairaz et Philippe Minard (2008) sur *L'information économique* a montré aux historiens combien il reste à apprendre sur la circulation et les usages de l'information. Sur quatre siècles, du XVI^e au XIX^e, il montre quand l'État a su tenter de mettre à la disposition des acteurs privés l'appareil statistique qu'il a mis en place, sans négliger le cas notable du ministre Jean-Antoine Chaptal, médecin manqué de Montpellier et chimiste entreprenant, pollueur à l'occasion, qui, après la chaptalisation des vins et l'enseignement de la chimie végétale à Polytechnique, devint ministre de l'Intérieur en 1800. Il est des heures où les circonstances poussent certains à vouloir ce que tous ont rêvé.

- *Le paradoxe d'Abraham Moles*

À défaut de garanties soutenues par des institutions dûment mandatées, qui ne voit les abus possibles ? En l'état, il faut le reconnaître, on observe que peu de choses sont systématiquement faites pour modifier l'opinion

1. Le lecteur trouvera dans Millard (1995) un tableau des diverses interprétations récentes de la doctrine de l'institution, et des remarques sur le droit pré-étatique chez Hauriou : commerce, contrat, institutions (ibid. p. 390).

et le comportement des usagers sans prendre en compte leurs intérêts et leurs préférences exprimées. Abraham Moles (1967, p. 332) avait examiné la lenteur de la mise en place de dispositifs de gouvernance manipulateurs¹ :

« Les raisons pour lesquelles l'individu a été relativement protégé des effets de la psychologie et de la sociologie appliquées ... sont :

- 1) le manque de conception d'ensemble suffisamment claire de la part des maîtres du corps social,
- 2) le fait que ceux qui pourraient prendre les choses en main n'ont pas une soif assez grande du pouvoir,
- 3) et surtout le désordre de la société actuelle, qui corrige une propagande par une autre et une publicité par une autre, développant une sorte d'immunisation quant aux fins de ces propagandes mais transformant l'individu en terrain de combat de slogans divers, grossiers, schématiques, exerçant sur l'esprit une action dispersive et contribuant tous à renforcer l'aspect " mosaïque " de la culture. »

Nous voilà rassurés et prévenus. Dans cette perspective, l'homme des données massives subit aujourd'hui le processus de décomposition qui alimente les images juxtaposées sur les écrans de son smartphone, de sa télévision et de sa mémoire. Celle-ci devient un tissu capricieux et versatile qui le pousse à réagir vivement aux moindres épisodes contrariants pour son idéologie ou ses intérêts du moment plus ou moins bien compris. En même temps, il n'est pas assuré que cet homme ait pour autant le contrôle des moyens d'enchaîner aisément et de critiquer les idées en cause dans les dossiers le concernant. La discussion publique et les processus de prise de décision n'y gagneront pas. Voilà pour les acteurs du court terme.

Les progrès de la scolarisation permettent d'envisager une évolution des courants d'opinion moins tumultueuse : plus documentée, mieux argumentée, plus sensible aux enjeux globaux. Les jeunes générations dépendent moins des *mass media* et de leurs campagnes de dénonciations de scandales. Ils sont souvent plus sensibles aux enjeux majeurs : croissance équilibrée, pays en voie de développement, investissements de recherche et d'éducation, environnement naturel. Voilà pour les acteurs du moyen terme.

La confrontation entre ces deux modèles de participation est engagée. Une troisième voie est du reste accessible, celle d'un rôle plus clair des institutions, si leurs mandats leur permettent de solliciter l'expression et la négociation des opinions de chacun.

Conclusions

Promenez-vous en forêt de Fontainebleau. L'immense majorité des espèces qui occupent ce site sont venues des cinq continents depuis les grandes navigations du XVI^e siècle. La forêt est une bibliothèque dont les pépinières et les jardins botaniques constituent les catalogues, on dira aujourd'hui les bases de données.

Les accès d'optimisme qui ont porté nos ancêtres à indexer toutes choses ne seront pas regrettés. Au contraire, nous pourrions sans gêne y succomber dans le contexte nouveau des données massives. L'existence des grandes institutions, cathédrales où sont déposés certains de ces trésors, invite à reprendre la marche.

Ainsi, dans le domaine des transports, un effort de créativité institutionnelle permettra à la République de disposer de sa puissance de collecte et de calcul, sans oublier son pouvoir d'initiative et d'incitation, en créant pour commencer une instance de gestion de la collecte, du traitement et de l'usage des données relatives aux

1. La conclusion de l'ouvrage, quelques pages plus bas, trahit une inquiétude forte.

espaces et à la mobilité¹. « Les opportunités sont réelles : le ministère de l'Écologie, de l'Énergie et de la Mer, chargé des transports, vient ... de se doter en 2016 d'une fonction de superviseur général des données (SGD). L'objectif est notamment de mieux valoriser les données existantes pour gagner en productivité, pour réaliser des économies et pour créer de nouveaux services. », je cite Luc Mathis dans sa *Synthèse étendue*.

On en parlera donc. Le volume des investissements consacrés aux autres facteurs de la mobilité (infrastructures, véhicules, énergie et autres) permet de penser qu'un tel investissement se situerait fort loin de la plage des rendements décroissants. De grands acteurs du secteur, publics, semi-publics et privés, seraient heureux de participer à la définition et à la construction d'un centre d'évaluation des moyens de calcul consacrés à l'espace national. L'avenir resterait à écrire, à calculer aussi.

Quoi qu'il en advienne de ces ambitions, je vous propose de conclure sur l'image d'un dispositif de gestion du trafic où les choses se réguleraient d'elles-mêmes, c'est-à-dire sans calcul explicite en temps réel. Comme il en va de deux fleuves qui, croisant leurs eaux, apportent leurs sédiments vers les basses terres, cette régulation rencontre un phénomène qui a sa dynamique propre. Elle peut avoir été calculée, ce qui aura permis le tracé des courbes et de la pente des rives. Aussi bien, on rencontre dans le corps des ingénieurs des voies hydrauliques des gens pour s'étonner du niveau archaïque de la gestion des données du transport, sachant que nous aurons tôt ou tard de très grands bassins de datamasse.

Le tourbillon dans le fleuve dissipe des énergies. Le croisement giratoire, souvent dit « rond-point », donne l'image d'un monde qui se modère soi-même. Le calcul sur les données achevé², on confie à un architecte convenablement initié le soin de dessiner les parcours en sorte que le flux induit projette sa propre masse sur les orientations de chacun. Le résultat escompté est que les écarts de direction et de force engagée, qui dans d'autres types de carrefours opposeraient les divers véhicules s'amortissent ici réciproquement selon la géométrie du manège. C'est le paradis des navigateurs, qui nous consolera de l'enfer de Raymond Devos³. Cet automate n'épuise pas notre sujet mais il laisse deviner à quoi rêvent les spécialistes des données massives quand ils déposent leur crayon : « enrrouler, entre les jarrets, sur la chaussée, selon l'instrument en faveur, la fiction d'un éblouissant rail continu » (Mallarmé, 1897).

Références

- Beck U. (1986). *Risikogesellschaft: Auf dem Weg in eine andere Moderne*, Suhrkamp.
Bédarida F. (1987). *Pour une histoire de la statistique*, Economica (2 vols.).
Coase R.H. (1937). « The nature of the firm », *Economica*, 4, 16, novembre, p. 386-405.

1. À défaut, des doctorants de l'Arizona pourraient bien tenter d'imposer un programme interactif Mobil@Hexagon 2.0.

2. Ce dispositif est moins susceptible d'accidents que le modèle classique de carrefour, en croix, protégé ou non par des feux ou des priorités. On en a donc récemment construit par dizaines de milliers en Europe, voir Houk, Lasserre et Sultan (1996).

3. Le mouvement circulaire s'approprie sur les escaliers et tapis roulants, comme dans ces ascenseurs sans portes où le circuit vertical des cabines invite les usagers à entrer et sortir sans tergiverser : ce mouvement de chapelet leur vaut le nom de *Paternoster* sur le campus de l'Université de Francfort.

- Cour des comptes (2017). Rapport public annuel. Tome 1. « L'écotaxe poids lourds », consultable sur le site www.cocomptes.fr
- De Borger B., Proost S. (2012). « A Political economy of road pricing », *Journal of Urban Economics*, 71, p. 79-92.
- De Palma A. (en préparation). « Transferts des risques dans le modèle des Carrosses à cinq sols de Blaise Pascal ».
- De Palma A., Lindsey R., Proost S. (2007). *Investment and the use of Tax and Toll Revenues in the Transport Sector*, Elsevier.
- Galbraith J.K. (2008). *The Predator State*, Free Press, *L'État prédateur*, Le Seuil, 2009.
- Houk M., Lasserre V., Sultan V. (1996). « L'incontournable avancée des carrefours giratoires. Analyse de la prise de décision publique », *Politiques et management public*, 14, p. 111-131.
- Lévi-Strauss C. (1971). *L'Homme nu*, Mythologiques 4, Plon.
- Lundwall É. (2000). *Les carrosses à cinq sols - Pascal entrepreneur*, Science Infuse.
- Mallarmé S. (1897). « L'action restreinte », *Divagations*, Fasquelle, p. 255-262.
- Mandell S., Proost S. (2016). « Why truck distance taxes are contagious and drive fuel taxes to the bottom », *Journal of Urban Economics*, 93, p. 1-17.
- Margairaz D., Minard P., eds. (2008). *L'information économique (XVI^e -XIX^e siècle)*, Paris, Comité pour l'histoire économique et financière de la France.
- Margairaz D. (2012). « Entreprise, privilège et service dans les transports publics parisiens XVII^e-XVIII^e siècles », *Concevoir et agir : Énoncés techniques et économie de l'action XVI^e-XX^e siècles*, Garnier.
- Millard E. (1995). « Hauriou et la théorie de l'institution », *Droit et société*, 30/31, p. 381-412.
- Moles A. (1967). *Sociodynamique de la culture*, Mouton.
- Ollivier-Trigalo M. (2013). « L'instauration d'une écotaxe sur les poids lourds en France : endurance technico-économique et impulsions politiques », *Développement durable et territoires*, 4.
- Proost S., De Borger B., Koskenoja P. (2007). « Public finance aspects of transport charging and investments », in : De Palma A., Lindsey R., Proost S. *Investment and the use of Tax and Toll Revenues in the Transport Sector*, Elsevier.
- Proost S., Thisse J.-F. (2015). « Skilled cities, regional disparities and efficient transport: the state of the art and a research agenda », Discussion paper series 10790, London, Centre for economic policy research, in: Matyas L., Blundell R., Cantillon E., Chizzolini B., Ivaldi M., Leininger W., Marimon R., Steen F. (2017). *Economics without Borders, Economic Research for European Policy Challenges*, Cambridge University Press.
- Tirole J. (2014). « Défaillances de marché et politiques publiques », traduction de la Nobel lecture du 8/12/2014, Document de travail, Institut d'économie industrielle, Toulouse.
- Verdier H. (2015). « Les données au service de la transformation de l'action publique ». Rapport au premier ministre sur la gouvernance de la donnée. Administrateur général des données, secrétariat général pour la modernisation de l'action publique, services du Premier ministre, Paris, La Documentation française ; site agd.data.gouv.fr
- Weismann P. (2013). « The genesis of the Eurovignette directive », *European Transport*, 53, Paper no. 2.