HAL
open science

Des données à l'intermédiation,une révolution économique et politique

Stephane Frenot, Stéphane Grumbach

To cite this version:

Stephane Frenot, Stéphane Grumbach. Des données à l'intermédiation,une révolution économique et politique. Big data : nouvelles partitions de l'information, De Boeck, 2014, 2804189155. hal-01091812

HAL Id: hal-01091812

https://inria.hal.science/hal-01091812v1

Submitted on 6 Dec 2014

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

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

Des données à l'intermédiation, une révolution économique et politique *

Stéphane Frénot
Université de Lyon

Stéphane Grumbach
INRIA

12 août 2014

Résumé

L'accumulation massive de données sur l'activité humaine a permis l'émergence de nouveaux services qui transforment progressivement notre environnement. L'intermédiation, c'est-à-dire la mise en relation, entre personnes ou services, est en général au coeur de ces services. Les moteurs de recherche ou les réseaux sociaux font partie des nouveaux types d'intermédiation impensables dans l'ère pré-numérique. Progressivement, les systèmes d'intermédiation pénètrent les secteurs traditionnels d'activité. Leur puissance augmente très rapidement, plaçant les premiers d'entre eux à l'égal des géants de l'industrie du pétrole. Si la disruption économique se fait déjà sentir dans divers secteurs d'activité, rendant caduques bien des entreprises, leur impact devrait aller croissant et toucher les services de l'administration et de la gouvernance. Ils opèrent un déplacement du pouvoir à la fois vers le bas, vers les utilisateurs, et vers le haut, vers les plateformes elles-même. Leur concentration géographique dans quelques pays seulement, parmi lesquels aucun pays européen, pourrait être un sujet d'inquiétude.

1 Introduction

La révolution numérique change le monde. La majorité de la population mondiale utilise quotidiennement des services numériques, moteur de recherche, media sociaux, etc. encore inexistants il y a seulement une quinzaine d'années. De surcroît, un nombre croissant de services traditionnels, publics ou privés, banques, écoles, impôts, commerces, etc. encouragent leurs utilisateurs à un mode d'interaction numérique également. Ces transformations touchent des aspects fondamentaux de nos sociétés et de nos vies, comme l'accès au savoir ou les relations entre les personnes, qui déstabilisent de nombreuses institutions traditionnelles, et font déjà l'objet de travaux de sociologie [1] [2]. Les changements sont donc extrêmement visibles. Mais peut-on pour autant parler de révolution ?

Au coeur de ces changements, se trouvent les entreprises de technologie qui offrent les nouveaux services du monde numérique. Ces entreprises ont connu au cours de la dernière décennie, une croissance unique dans l'histoire économique, se hissant pour les plus grandes d'entre elles au premier rang des capitalisations mondiales, rivalisant ainsi avec l'industrie pétrolière, une industrie traditionnelle, absolument centrale pour l'économie et reposant sur des infrastructures massives.

*à paraitre dans "Big data : nouvelles partitions de l'information", actes du colloque Inria, octobre 2014.

1. Pierre Merklé, Sociologie des réseaux sociaux, Paris, Éd. La Découverte, coll. Repères, 2004
2. Doreian, Patrick, and Frans Stokman, eds. Evolution of social networks. Routledge, 2013.

Les services numériques sont devenus aussi incontournables que l'énergie. Leur usage est devenu universel, et le nombre d'utilisateurs des plus grosses sociétés dépasse le milliard.

Ces nouveaux services numériques bien qu'ils couvrent une large diversité de domaines d'application ont un point commun essentiel, ils assurent une intermédiation. Leur rôle est de mettre en relation des personnes entre elles, des personnes avec des services, et plus généralement des objets quelconques entre eux. Le moteur de recherche, premier grand système d'intermédiation, met en relation des personnes avec des connaissances. Les réseaux sociaux mettent en relation des personnes dans leur environnement.

L'intermédiation est une activité absolument fondamentale qui est loin d'être nouvelle. Elle est à la base de notre organisation sociale, et est assurée dans tous les secteurs par de nombreuses structures, privées ou publiques, qui vont des banques aux bibliothèques, des organisations syndicales aux agences de voyage, des services consulaires à la presse, des commerces aux services de la poste, etc. qui toutes permettent de fournir un service à des usagers ou clients en les mettant en relation avec les services ou les produits recherchés, ou avec les opérateurs les offrant.

Le nombre des organisations dont le rôle est d'intermédier, c'est-à-dire par exemple de transmettre de l'information, de mettre en relation, ou de servir de tiers de confiance pour des transactions, est très important. L'intermédiation est un des services de base de n'importe quelle société, qui se decline sous un très grand nombre de formes différentes.

Les systèmes d'intermédiation numérique introduisent une révolution majeure dans nos sociétés. D'abord ils ont permis l'émergence de nouveaux services comme le moteur de recherche ou les réseaux sociaux, qui étaient irréalisables, pas même pensables sans le numérique. Mais au-delà de ces nouveaux services, les systèmes d'intermédiation pénètrent progressivement la sphère historique de l'intermédiation traditionnelle, c'est-à-dire des services offerts à leurs utilisateurs par des institutions ou des entreprises dont la mission essentielle est la fourniture de ces services.

Quelques exemples simples permettent de comprendre comment les services numériques s'insinuent dans des secteurs traditionnels. Dans la presse par exemple les systèmes d'intermédiation peuvent s'abstraire de la production d'articles pour juste intermédier entre les lecteurs et les rédacteurs, offrant ainsi une plateforme d'accès à l'actualité. Le fonctionnement de telles plateformes déstabilise de manière profonde l'économie de la presse. Dans les transports, les systèmes d'intermédiation intermédient entre les passagers et les transporteurs individuels, taxi par exemple, s'abstrayant des sociétés qui gèrent des flottes de véhicules, qui peuvent assez vite devenir caduques.

Tous les services traditionnels qui intermédient d'une manière ou d'une autre entre des utilisateurs et des services seront affectés par les systèmes d'intermédiation numérique. On ne perçoit de cette révolution que les prémisses dans des secteurs comme ceux que nous avons mentionnés ci-dessus comme la presse ou les transports. Mais la transformation avancera subrepticement dans tous les secteurs. La plupart seront radicalement transformés. D'autres disparaitront simplement.

C'est probablement à ce niveau, celui de la maîtrise de l'intermédiation entre les personnes, et au-delà entre les objets connectés, que ce situera l'impact le plus profond de la révolution numérique, la disruption la plus radicale, celle qui fera s'effondrer le plus grand nombre d'établissements, aussi solidement installés soient-ils, et qui aujourd'hui assurent ce type de service, c'est-à-dire tous ceux qui d'une manière ou d'une autre interagissent avec le monde extérieur de manière ouverte, autrement dit la majeur partie.

Les systèmes d'intermédiation procèdent à un déplacement du pouvoir d'une part vers le bas, cela a déjà abondamment été décrit et retenu comme un argument en faveur des systèmes numé-

riques avec l'empowerment des citoyens en particulier, mais également vers le haut, vers le système d'intermédiation lui-même qui maîtrise des champs entiers de l'économie globale, dépouillant potentiellement complètement les corps intermédiaires, devenus impuissants et caduques, comme nous le montrons dans la Section 5 au travers de quelques exemples importants dans différents secteurs d'activité.

Le secteur public sera également touché. On le voit dès aujourd'hui. Ce que les grands systèmes d'intermédiation assurent n'est autre qu'un service essentiel satisfaisant les grandes règles du service public. La disponibilité progressive des données publiques permettra le développement de nouveaux services mis en oeuvre en particulier par les grands systèmes ou tout au moins reposant sur leurs infrastructures, qui rendront superfétatoires certains services administratifs. La gestion de l'identité, du ressort des Etats, évoluera également pour s'adapter à la fois à l'authentification en ligne, mais également à des critères d'identification renouvelée comme nous le montrons dans la Section 6. La fiscalité devra également être complètement adaptée pour prendre en compte les échanges sur les systèmes d'intermédiation.

Les systèmes d'intermédiation reposent tous sur la même architecture. Ils s'attaquent à des données qui sont produites à l'extérieur du système (les pages Web pour le moteur de recherche, ou les données personnelles pour les réseaux sociaux). En analysant et transformant ces données, ils sont en mesure de développer des services. L'utilisation de ces services par les utilisateurs génèrent à son tour des données, les traces d'utilisation, qui permettent à leur tout de générer de nouveaux services, parmi lesquels des services génériques, comme les tendances sur le moteur de recherche, et des services personnalisés, qui exploitent la fine connaissance qu'ils ont de leurs utilisateurs.

Le nombre d'utilisateurs et la quantité de données maîtrisées jouent un rôle fondamental dans l'efficacité de l'intermédiation. Plus il y a d'utilisateurs, plus il y a de données, plus le système est pertinent. La taille de la population décuplent de manière quadratique, suivant une loi de Metcalfe, la capacité du système. La force des grands nombres expliquent l'émergence de systèmes qui finissent par devenir universels et détiennent une véritable suprématie. Dans le domaine de l'intermédiation, la concentration est la règle.

Les grands systèmes suivent une logique de *plateformes*, c'est-à-dire de systèmes ouverts sur lesquels viennent se greffer des fonctionnalités offertes par d'autres. C'est un aspect fondamental de cette industrie, qui génèrent des écosystèmes autour desquels se développent d'autres industries. Les services de base qu'offrent les plateformes à leurs utilisateurs incluent en particulier l'authentification, la capacité de stockage de données et la gestion de l'accès à ces données, ainsi que des outils de communication. Au développeurs, les plateformes offrent une API, c'est-à-dire une interface qui permet de programmer des applications utilisant les services de base de la plateforme.

L'ouverture non pas seulement aux utilisateurs mais aux entreprises des plateformes est une des conditions essentielles de leur développement, qui permet l'emergence de ces écosystèmes avec différents types de partenaires. Ces différents partenariats sont gérés par les plateformes suivant les règles d'une économie biface, où les utilisateurs bénéficient d'un service gratuit, sponsorisé par les clients, à qui des services payants sont proposés. Les principaux clients sont aujourd'hui les annonceurs, qui trouvent sur les plateformes une efficacité décuplée par rapport aux media traditionnels.

Comme les systèmes de la finance haute fréquence, qui contrôlent l'économie mondiale en gérant les flux d'informations capitalistiques à très haute fréquence, les plateformes d'intermédiation gèrent une partie des flux mondiaux de données à haute fréquence également. Les données sont devenues une resource, comparable au pétrole, pour une industrie en pleine émergence qui en tire déjà de

considérables revenus. Les plateformes qui offrent les services pour la multitude récoltent une partie importante des données personnelles de part le monde.

Les plateformes ne changent pas seulement l'équilibre des entreprises ou des services de l'Etat, elles bouleversent également les équilibres internationaux à cause de leur concentration géographique qui donne aux pays qui les hébergent un avantage évident que les révélations d'Edward Snowden ont pour une part permis de mettre au grand jour. L'Europe se trouve à ce titre dans une relation de dépendance vis-à-vis des Etats-Unis qui augmente rapidement. Cette économie suscite également de nombreuses interrogations sur l'usage possible qui peut-être fait des données personnelles qui sont récoltées, sur la possibilité d'une surveillance globalisée des individus par les Etats, sur l'usage légitime que peuvent en faire les entreprises. Ce débat ne fait que commencer.

L'article est organisé de la manière suivante. Dans la Section 2, nous rappelons les changements technologiques qui avec le développement de l'Internet et du Web ont rendu possible l'émergence des plateformes d'intermédiation. Le fonctionnement des systèmes d'intermédiation numérique est présenté dans la Section 3, avec les différents types de données sur lesquels ils reposent. La Section 4 est consacrée à l'économie des plateformes d'intermédiation. Dans la Section 5, nous présentons quelques exemples de secteurs traditionnels qui subissent la disruption des plateformes. Nous montrons dans la Section 6, que cette disruption touchera le domaine de l'administration publique et l'organisation politique. Enfin, la Section 7 est consacrée aux enjeux internationaux des échanges de données au niveau planétaire.

2 De nouveaux possibles

La Révolution numérique est rendue possible par le développement des capacités technologiques, en particulier pour la gestion et le traitement de données complexes, peu structurées, en très grande quantité, émanant de tous les secteurs d'activité. Elle est portée par l'évolution des technologies, dont les progrès conduisent régulièrement à de nouveaux services innovants. Elle contribue en retour au développement d'un secteur transversal, ouvrant de nouveau champ du possible, qui justifient les investissements considérables alloués à l'industrie du numérique.

2.1 La croissance des capacités de stockage et de traitement des données

C'est incontestablement avant tout la croissance continue des capacités de stockage et de calcul, qui a permis l'extraordinaire pénétration du numérique dans tous les secteurs d'activité, aujourd'hui au contact de chaque individu et embarqué dans un nombre toujours croissant d'appareils interconnectés.

Cette croissance est décrite par une loi, la loi de Moore, qui s'est imposée durablement à ces technologies, et a gouverné l'économie de l'industrie du semi-conducteur et de toutes les industries en aval. La Loi de Moore, énoncée à la fin des années 1960, prédisait le doublement tous les deux ans de la capacité des calculateurs. Plus précisément, la loi de Moore stipulait le doublement du nombre de transistors dans un circuit intégré. Cette croissance exponentielle a touché également d'autres aspects des calculateurs, outre l'intégration toujours plus forte, parmi lesquels la vitesse des processeurs et la capacité de mémoire. Elle s'est accompagnée également d'une baisse drastique des prix qui a permis la pénétration généralisée des composants électroniques dans tous les secteurs.

Depuis que les premiers microprocesseurs, produit par Intel, ont fait leur apparition au début des années 1970, leur capacité a connu une croissance ininterrompue. La vitesse des processeurs

est passée d'à peu près 1 MHz en 1970, à plusieurs GHz en 2010, soit plus d'un facteur mille. La capacité des disques durs est passée de 1 MB en 1980 à 1 TB en 2010, soit un million de fois plus. Quant aux supercalculateurs, reposant sur quelques microprocesseurs au tournant des années 1970, ils en comptent aujourd'hui des dizaines de milliers. Le nombre d'opérations par seconde grimpe de 80 MFLOPS avec le Cray-1 en 1982, soit des dizaines de milliards d'opérations par seconde, à des dizaines de PFLOPS, soit des dizaines de millions de milliards d'opérations par seconde, la aussi une croissance de six ordres de grandeur.

Après plus de cinquante ans de croissance exponentielle, la loi de Moore touche à des limites physiques, qui mettront progressivement fin à ce mode de croissance. Les efforts se poursuivront dans d'autres directions, parmi lesquels des modèles de calcul radicalement différents, exploitant les propriétés de la matière à des échelles différentes, comme les calculateurs quantiques, basés sur la superposition et l'intrication d'états quantiques, ou encore les bio-calculateurs, reposant sur les reactions chimiques de la biologie. Mais c'est incontestablement du coté du logiciel que les avancées se produisent aujourd'hui.

2.2 Des algorithmes toujours plus puissants

La capacité des algorithmes a suivi la croissance des machines, permettant progressivement de résoudre des problèmes insolubles auparavant. La taille des programmes a connu une croissance du même ordre que le hardware sous-jacent, conduisant aux systèmes les plus complexes que l'humanité ait jamais réalisé. Les plus gros programmes atteignent aujourd'hui des centaines de millions de lignes de code [3], rendant leur vérification extrêmement complexe.

L'algorithme du Page Rank, utilisé par Google pour son moteur de recherche, et développé au milieu des années 1990 à partir d'algorithmes pour la bibliométrie, fait partie de ces défis algorithmiques qui paraissaient hors de portée quand ils ont été initialement proposés. La taille des données, l'ensemble des pages du Web, et la complexité du calcul de la matrice étaient des ordres de grandeur au-delà de ce qui se faisait à l'époque. Avoir su relever ce défi a conduit au développement d'immenses capacités de calcul pour une industrie nouvelle, celle du Web et des media sociaux.

La capacité à résoudre des problèmes dont la complexité algorithmique est élevée, a aussi fait d'immense progrès avec les techniques d'approximation en particulier, qui permettent de traiter de manière pratique des problèmes d'optimisation théoriquement difficile. Mais ce ne sont pas seulement les frontières liées à la taille des données ou à la complexité algorithmique des calculs, qui sont progressivement repoussées, mais également, le type même de problèmes auxquels les algorithmes s'affrontent qui évolue.

La défaite de Kasparov en 1997 face à la machine d'IBM, Deep Blue, une machine spécialement conçue pour ce défi, marque un tournant dans les rapports entre l'homme et la machine, et le début d'une longue série de succès des algorithmes face aux capacités intellectuelles humaines. Le défi que constitue la réalisation par l'homme d'une machine plus puissante que lui a hanté la littérature depuis des temps reculés. Le film de Stanley Kubrick, 2001 Odyssée de l'Espace, produit en 1968 montre l'étonnante rivalité entre une machine poursuivant ses propres objectifs et l'homme. Les capacités respectives de l'homme et de la machine sont un sujet d'inquiétude qui suscitera bien des passions dans les décennies à venir.

La capacité que des algorithmes, c'est-à-dire de programmes, déterministes ou non, mais dont la série d'instructions est déterminée a priori, puisse rivaliser avec la liberté d'imagination du cerveau

3. `http://www.informationisbeautiful.net/visualizations/million-lines-of-code/`

humain laisse perplexe. La question n'a pas fini d'être débattue. Nous découvrirons dans les années à venir qu'un nombre croissant de tâches, autrefois considérées comme relevant de l'intelligence humaine, seront assurées par des machines. La rédaction de textes intelligibles, comme des articles de presse, fait partie de celles-ci.

2.3 L'émergence de l'à-peu-près

Le passage d'un monde exact, de forte inspiration scientifique, d'objets mathématiques idéaux, à un monde, qui tout en gardant une incontestable rigueur, tolère le flou et une certaine approximation, est une étape fondamentale de l'évolution de la pensée numérique. Cette capacité à penser l'imprécis, l'à-peu-près n'est pas nouvelle, et bien évidemment elle n'est pas propre au numérique. Les disciplines des sciences naturelles, des sciences du vivant à la physique, connaissent l'approximation et savent modéliser l'incertitude. Mais l'informatique s'est dans son histoire surtout consacrée à s'abstraire de l'incertitude pour garantir les calculs. Sa récente pénétration universelle dans toute la société contraint le numérique à apprehender l'à-peu-près, à en faire un atout majeur et non plus à chercher à l'éviter.

Les fondations de l'informatique, basées sur la logique mathématique et des objets abstraits modélisant le calcul, ont fait émerger un monde d'exactitude, proche des mathématiques, avec des résultats théoriques puissants sur la complexité du calcul, ou même sur ce qui était calculable. La programmation à l'inverse est une activité humaine, considérée par certains comme un art, dont l'objet est de produire des programmes qui satisfont un cahier des charges, mais sans que l'on puisse offrir de garantie formelle de résultat.

La correction des programmes que de nombreuses applications requièrent a donc été un enjeu considerable pour l'informatique. Le domaine bancaire, secteur d'application de toute première importance, a conduit au développement de techniques contraignantes qui permettent de faire des transactions en toute sécurité. La capacité à garantir l'exécution correcte des programmes, en permettant la résilience en cas de panne, est essentielle dans de nombreuses applications, en particulier dans les domaines critiques associés à des risques importants.

L'un des défis majeurs de l'informatique est de pouvoir prouver mathématiquement la correction des programmes. Même si des avancées majeures ont été réalisées dans la preuve automatisée, à ce jour les grands systèmes qui nous entourent ne sont pas certifiés. De plus, même si les programmes sont certifiés, l'ensemble des machines et réseaux sur lesquels ils tournent ne le sont pas, diminuant d'autant la capacité à garantir le fonctionnement d'un système. Ce qui est extraordinaire dans ce domaine de la preuve, c'est que désormais, même les mathématiques sont confrontées, de par la longueur des preuves qui dépassent les mille pages, à l'incapacité de garantir un résultat [4].

La nature des données et des calculs impliqués dans les applications émergentes du numérique, en particulier celles des media sociaux qui concernent les activités humaines, est radicalement différente de ce qui prévalait dans l'histoire de l'informatique. Des données peu structurées, multimedia, sur lesquels sont exécutées des operations souvent statistiques, des données floues donc avec des calculs flous eux aussi, en complete rupture avec les transactions bancaires.

Ce n'est pas seulement la nature des données et des calculs qui force l'informatique à prendre en compte l'à-peu-près mais également l'ambition de faire des systèmes universaux, utilisés par un nombre très large d'utilisateurs et traitant des quantités de données et de calculs immenses. De tels systèmes imposent une organisation distribuée, reposant sur des réseaux d'un très grand nombre

4. Jeremy Avigad, John Harrison, Formally verified mathematics, CACM April 2014

de machines. Au début des années 2000, un nouveau résultat d'impossibilité a été démontré, le théorème de Brewer, qui stipule qu'un système distribué ne peut garantir à la fois les trois conditions suivantes : (i) la cohérence, garantissant que les noeuds du système ont les mêmes données au même moment, (ii) la disponibilité, garantissant une réponse aux requêtes, et finalement (iii) le morcellement, garantissant que le système continue à fonctionner même si le réseau est morcelé. Les grands systèmes de la toile ont fait le choix de la disponibilité et du morcellement, tolérant donc l'incoherence de leurs données, pour satisfaire les besoins de leurs utilisateurs.

2.4 Hic et nunc

Les services numériques contrôlent aujourd'hui de manière continue un nombre croissant de secteurs. A titre d'exemple, on peut mentionner l'automobile qui intégrera progressivement un système global d'assistance, qui a commencé par le contrôle mécanique par l'ordinateur de bord, l'assistance à la navigation du conducteur, se développe avec le contrôle globalisé des véhicules, la capacité à intervenir à distance sur un véhicule, et vise à terme à automatiser la conduite même, un défi dont Google a contribué à démontrer la faisabilité avec sa Self-driving car.

Au-delà du contrôle automatisé d'un nombre croissant de systèmes, les grands opérateurs en ligne offrent des services aux utilisateurs, moteurs de recherche, media sociaux, etc. avec une très grande qualité de service qui sont disponibles en continu et essentiellement sans faille. Ces nouveaux services essentiels viennent completer la gamme des services publics, comme l'approvisionnement en eau et en énergie, les transports, etc, avec la même continuité de service.

La rapidité de fonctionnement des services numériques est essentielle. Si, le temps est un critère évident pour les systèmes temps reel, il en va de même pour les systèmes offrant des services aux utilisateurs. Les grands industriels ont pu mesurer qu'un légère baisse de célérité dans leur réponse avait un impact immédiat sur leurs ventes. Pour Amazon par exemple, une augmentation d'un dixième de seconde dans le temps de réponse provoque une chute de 1% des ventes, alors que pour Google, une demie seconde de délai provoque une chute du trafic de 20%, un impact très important [5].

Le défi de l'industrie n'est plus de gérer des données et de faire des calculs sur les données, mais de gérer des flux de données qui sont transformées en continu. Comme dans la finance haute-fréquence, où un retard de quelques mili-secondes sur la concurrence entraine des pertes en millions de dollars, les flux sont extrêmement rapides et permettent d'offrir des services extrêmement à jour dans leurs interactions avec leurs utilisateurs.

Les réseaux qui maillent de plus en plus profondément notre environnement depuis le démarrage d'Internet modifient l'espace et le temps dans lequel nous vivons. Le temps est un temps continu supporté par un flux d'interaction extrêmement rapide. L'espace est déterritorialisé, et peut s'affranchir des contraintes imposé par la distance dans le monde physique, ce qui conduit certains à parler d'espace non métrique.

L'espace est profondément modifié par les outils de la mobilité qui permettent à notre espace virtuel de nous suivre dans nos déplacement, et de toucher ce que l'on veut où qu'on soit. Mais l'espace est aussi radicalement transformé dans sa composante physique, permettant le ce que je veux, quand je veux, où je veux. Amazon déploie des efforts impressionnants pour diminuer les délais de livraisons des biens physiques. Mais c'est surtout les imprimantes 3D qui permettent de reproduire des objets matériels n'importe où à la demande, aussi facilement qu'on reproduit des

5. `http://highscalability.com/latency-everywhere-and-it-costs-you-sales-how-crush-it`

biens multimedia, qui changent profondément les caractéristiques de notre espace.

3 L'intermédiation

La révolution numérique a permis l'émergence de nouveaux services, qui n'existaient pas dans le passé. Le moteur de recherche en est certainement à la fois le premier exemple de taille majeure et la meilleure illustration. Il met essentiellement en relation des personnes avec des connaissances qu'elles recherchent. Son incroyable succès — le premier d'entre eux, Google, traite à lui tout seul près de six milliards de requêtes par jour, essentiellement une par terrien par jour — en a fait un élément incontournable de notre accès à l'information, transformant les structures classiques d'intermédiation dans bien des domaines comme nous le verrons dans la Section 5.

L'intermédiation est un défi extraordinaire et l'époque que nous vivons nous confronte à la fois au meilleur avec l'efficacité souvent déconcertante des moteurs de recherche ou des systèmes de recommandation par exemple, mais également au pire avec les nombreuses interactions invasives, visibles ou non, qui envahissent nos appareils.

De toutes les transformations induites par le numérique sur la société, l'automatisation de l'intermédiation aura probablement les conséquences les plus profondes et les moins anticipées. Avant de considérer les modèles économiques de l'intermédiation et les transformations sociétales induites, il est indispensable de s'arrêter sur la machinerie de l'intermédiation réalisée au moyen des systèmes numériques.

3.1 L'art de l'intermédiation

Le grand défi des systèmes d'intermédiation est de réaliser cette mise en relation avec pertinence et efficacité, c'est-à-dire de fournir aux utilisateurs des liens qui répondent exactement à leurs attentes, ou bien ce qui est encore plus intéressant, qui devancent leurs attentes et leurs ouvrent de nouveaux horizons. C'est précisément cette capacité d'assurer une intermédiation d'un autre type qui constitue la plus grande révolution de nos sociétés. Les systèmes d'intermédiation reposent tous sur la même alchimie :

1. Collecte de données primaires ;
2. Analyse et transformation des données ;
3. Génération de services à partir des données primaires ;
4. Collecte des traces d'utilisation par les utilisateurs ;
5. Génération de nouveaux services à partir des données secondaires ;
6. Collecte des traces de fonctionnement de la plateforme et adaptation.

Les données initiales sont de manière générale des données extérieures à la plateforme, comme celles constituées par les pages Web pour le moteur de recherche. Ces données sont créées par des entités indépendantes de la plateformes sans effort de cette dernière pour leur génération. La plateforme doit par contre être en mesure de récolter ces données, et différents modèles sont possibles. Dans le cas du moteur de recherche, les données sont récoltées en crawlant le Web continûment et en copiant les pages découvertes. Dans le cas du réseau social, la plateforme offre des services d'hébergement des données, que les utilisateurs chargent eux-mêmes pour profiter des services.

Une fois les données obtenues, il faut les analyser, les indexer et les classer suivant des critères dont la pertinence rendra possible les services associés. Le premier algorithme à avoir été utilisé sur

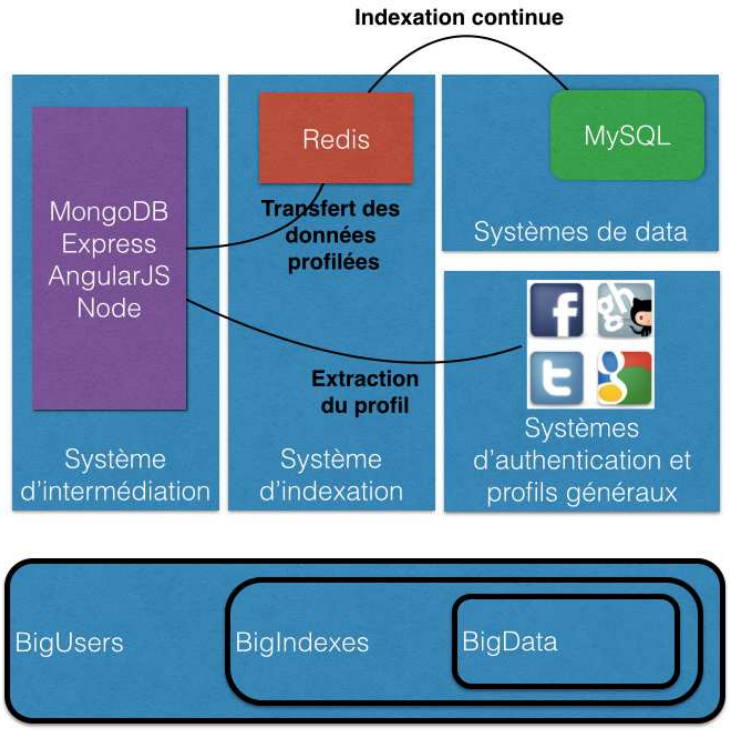

FIGURE 1 – Des données aux utilisateurs

une échelle dans les ordres de grandeur du Web, est le PageRank de Google, qui permet d'indexer les données et de leur associer un indice de pertinence.

Les premiers systèmes de navigation dans le Web ont été des portails offrant des classements thématiques, comme des encyclopédies en quelque sorte. Les algorithmes de pertinence ont complètement révolutionné l'accès. L'idée de base est relativement simple. L'indice de pertinence du PageRank [6] attribué à une page Web dépend du nombre de liens vers cette page Web. Essentiellement, plus ce nombre est grand plus la page est visible et donc pertinente. La pertinence repose donc sur un concept, le lien, qui est à la base du Web, et constitue donc une vraie nouveauté par rapport à la connaissance dont on pouvait disposer avant.

Le PageRank est en fait un peu plus subtil qu'un simple décompte du nombre de liens. D'abord, les liens sont pondérés, d'une manière inversement proportionnelle au nombre de liens émis par une page, comme si la page partageait sa notoriété en autant de part qu'elle a de liens sortants. Une page hérite donc d'une part de notoriété d'autant plus petite que la page source a de liens sortants.

Pour obtenir une estimation plus précise, on itère le calcul. On considère le graphe des pages avec leur liens. On attribue initialement à chaque noeud une notoriété égale, disons de 1, que le noeud partage et distribue de manière égale aux noeuds aux extrémités de ces liens sortants. Chaque

6. Anand Rajaraman, Jeffrey D. Ullman. Mining of Massive Datasets

noeud se retrouve donc avec ce qu'il a obtenu sur ces liens entrant. On répète cette opération jusqu'à l'obtention d'un point fixe avec des valeurs stables sur chaque noeud.

Ce calcul est toutefois loin d'être évident. Mathématiquement, on peut montrer, sous des hypothèses raisonnables pour le Web, que ce point fixe existe bien et peut-être obtenu en itérant le calcul. Du point de vue du calcul, la taille de la matrice constitue un défi majeur. Le génie des fondateurs de Google a été de ne pas avoir eu peur du caractère démesuré de l'aventure.

Le PageRank est un algorithme très complexe dont Google garde le secret, qui repose sur plusieurs centaines de paramètres, bien plus donc que les simples liens entre pages, et qui permet d'offrir un service, le moteur de recherche, dont les résultats donnent une satisfaction qui lui assure une suprématie particulière sur le secteur au niveau mondial.

3.2 Des données primaires aux données secondaires

Les données secondaires sont les données extraites des traces de l'activité sur les services de la plateforme. La capacité d'extraire des traces précises appartient au monde numérique. Elle était impensable avant, et tous les services découlant des données secondaires sont donc nouveaux.

Il y a essentiellement deux types de traces, celles qui sont associées à un utilisateur particulier et constituent donc des données personnelles, et celles qui correspondent plus largement à une population d'utilisateurs et indiquent des tendances dans les usages.

Les requêtes sur un moteur de recherche par exemple contiennent les mots clés permettant de sélectionner les pages Web désirées. Cette activité correspond à leur *utilisation primaire*. La séquence de requêtes d'un utilisateur permet de dresser un profil de cet utilisateur, qui peut servir à mieux cibler la publicité à son intention, ou plus généralement à adapter les réponse à ses requêtes en fonction de son profil, de fournir des réponses qui correspondent mieux à son attente. Il est possible d'activer ou de désactiver cette personalisation sur certains moteurs.

Les requêtes peuvent être également groupées par sujet et non par utilisateur. Ainsi la variation du nombre de requêtes, les tendances, constitue un indicateur qui peut être exploité au profit d'intérêts particuliers. La variation des requêtes concernant les symptômes de grippe permet de mesurer l'étendu de l'épidémie, son augmentation ou sa diminution. La pertinence de ces analyses de tendance, si elle fait débat aujourd'hui quant à son efficacité ne fait aucun doute sur le long terme.

Les requêtes d'une région permettent de dresser un profil de la région. Les moteurs de recherche, comme de nombreux autres systèmes d'intermédiation disposent de données extrêmement riches permettant de dériver des informations dans un nombre de domaines insoupçonnable a priori. A titre d'exemple, les données des systèmes d'aide à la navigation automobile s'avèrent être aux Etats-Unis des indicateurs temps réel de l'activité économique.

Nous verrons dans la Section 4, que de surcroît, le modèle économique des systèmes d'intermédiation repose de manière générale sur les données secondaires et les activités dérivées.

La question de l'utilisation des données secondaires personnelles par les systèmes d'intermédiation est donc complexe. Les applications possibles sont peu prévisibles, et de surcroit le modèle économique repose de manière essentielle sur ces données. Deux positions s'affrontent sur ce sujet. Une restriction à un usage prédéfini des données par les systèmes d'intermédiation, ou une liberté d'usage contrôlée a posteriori. La situation actuelle reposant sur un accord explicite par l'utilisateur de conditions d'utilisation que personne ne lit n'est pas satisfaisante.

3.3 La gouvernance des algorithmes

Les algorithmes jouent un rôle clé dans les systèmes d'intermédiation, qui révolutionne les capacités d'intermédiation et ce non seulement à cause de la puissance des systèmes numériques et des quantités de données, mais plus encore par l'introduction de formes de raisonnement et d'adaptations radicalement nouvelles. Comme pour le big data en général, les algorithmes cherchent des corrélations dans les données sans avoir de théorie a priori, sans chercher à démontrer des causalités. La conséquences sont importantes. D'abord par la capacité à faire émerger des correlations peu prévisibles. Ensuite par la capacité à agir sur la base non de causalités bien établies, mais de simples corrélations.

Aux débuts d'Amazon, Jeff Bezos avait fait appel à des critiques littéraires pour recommander les livres vendus sur la plateforme, avant de se rendre compte que les algorithmes de recommandation obtenaient de meilleurs résultats de vente. La recommandation automatique fonctionne sur un principe très simple. A partir de profil utilisateurs offrant de grandes similarités, suggérer à chaque utilisateur des éléments de la différence avec les autres profils. D'un certaine manière, c'est comparable dans la sphère non numérique aux recommandations que peuvent se faire des collègues ou amis ayant les mêmes goûts ou les mêmes centres d'intérêt entre eux. Les plateformes d'intermédiation reposent toutes essentiellement sur des variantes plus subtiles de ce principe de recommandation automatique.

Les algorithmes permettent aussi de tester l'efficacité de la plateforme en temps réel, en analysant les résultats. L'A/B Testing par exemple repose sur la mise en concurrence de services légèrement différents pour comparer leurs avantages respectifs et choisir le meilleur. Cette méthode a été abondamment utilisé pour la deuxième campagne présidentielle d'Obama. Ce sont des systèmes organiques, qui s'adaptent en permanence au réel, sur le modèle des systèmes vivants. Leur conception n'est pas celle de systèmes qui durent avec dès le depart toutes les bonnes fonctionnalités, mais au contraire de systèmes qui pourront évoluer de manière continue,

Les algorithmes sont également étroitement associés à l'intelligence collective qu'ils exploitent soit indirectement en exploitant les profils comportementaux de leurs utilisateurs, soit en les mettant à contribution explicitement en leur demandant de valider la qualité de résultats, soit de fournir un travail, comme dans le cas des Captcha par exemple. La logique des plateformes d'intermédiation est de combiner le travail des utilisateurs de la plateforme à la puissance des algorithmes.

4 L'économie des plateformes

Les plateformes sont des systèmes ouverts qui permettent le développement de services externes, soutenus par la plateforme (stockage de données, authentification, etc.), et partageant la création de valeur avec la plateforme. La plateforme met en place un écosystème qui lient les utilisateurs et les entreprises au travers de la plateforme. Contrairement à une entreprise traditionnelle qui propose des produits ou des services, les plateformes ne produisent rien, elles sont simplement des facilitateurs pour le développement d'entreprises tierces.

La meilleur démonstration du basculement de nos économies vers les plateformes est sans doute donnée par l'industrie des terminaux de la mobilité, les smartphones. Nokia et Blackberry qui dominaient largement le marché ont rapidement perdu leur suprématie au profit d'Apple et des terminaux Android, qui tous deux reposent non pas sur de meilleures technologies ou des fonctionnalités plus avancées, mais sur un écosystème ouvert aux développeurs extérieurs qui permet à ces outils d'évoluer en permanence.

Pour relever l'extraordinaire défi de l'intermédiation, les plateformes doivent d'une part proposer des services riches susceptibles d'attirer les utilisateurs, et d'autre part très bien connaitre leurs utilisateurs, qui sont leur seule richesse. Le nombre d'utilisateurs joue un role considérable. Essentiellement, une plateforme avec n utilisateurs conduit à un nombre quadratique de connections possibles entre ses utilisateurs. C'est la fameuse loi de Metcalfe initialement développée pour les telecommunications.

Les premières plateformes, comme Facebook ou LinkedIn, ont développé des outils de connexion entre leurs utilisateurs, qui ont petit à petit développé leurs contenus personnels, sur lesquels la plateformes ont développé de nouveaux services, parfois payant comme LinkedIn. Aujourd'hui l'abondance des outils de connexion entre utilisateurs, remonte l'importance des contenus qui sont proposés sur les nouvelles plateformes avant même les capacités d'interconnexion. C'est le cas par exemple de Pinterest ou d'Instagram par exemple.

Deux principes fondamentaux gouvernent l'architecture des plateformes d'intermédiation :

1. L'ouverture aux applications extérieures. La plateforme n'assure que l'infrastructure pour le développement des services offerts par des entreprises extérieures.
2. Le lien direct avec les utilisateurs. C'est un aspect essentiel pour récolter les données des transaction utilisateurs, qui sont le carburant des plateformes et qu'elles exploitent pour créer de nouveaux services pour leurs clients.

Le lien direct avec les utilisateurs est stratégique. Il permet à la plateforme de jouer son rôle d'intermédiaire, sans être à son tour intermédiée par un autre acteur. L'intermédiation assure à celui qui la realise, la mainmise et en général l'exclusivité sur les données d'activité, qui permettent de connaître les utilisateurs et de faire évoluer dynamiquement les services.

L'ouverture aux applications extérieures, c'est-à-dire celles développées par d'autres acteurs, que ce soient des développeurs indépendants, ou de grandes entreprises pour accéder à leurs services, est essentielle pour offrir une pluralité de services qui répondent aux besoins illimités des utilisateurs.

4.1 Le modèle économique biface

Dans une économie biface, une distinction est faite entre deux types d'utilisateurs ou clients, qui sont traités différemment. Les cartes de paiement par exemple reposent sur un modèle biface avec deux types d'utilisateurs, d'une part des commerçants qui payent un pourcentage sur les transactions, et des acheteurs, qui payent une cotisation annuelle fixe. Ce système incite bien sûr les acheteurs à utiliser ce moyen de paiement que les commerçants sont obligés d'accepter favorisant ainsi la société assurant le service du paiement.

Si divers exemples existent dans l'économie traditionnelle, c'est sur ce modèle, qui a contribué massivement à leur succès, que reposent toutes les plateformes d'intermédiation. En general les utilisateurs de base accede gratuitement au service des plateformes. L'usage du moteur de recherche ou des réseaux sociaux est gratuit pour les utilisateurs. Les clients de la plateforme ont accès à des services payants dérivés en général des données secondaires de la plateforme. Dans l'économie actuelle, les clients sont surtout des annonceurs.

Ce modèle vise à développer l'element le plus stratégique pour les plateformes, à savoir la base d'utilisateurs. Son financement est assuré au depart par le capital risque, puis une fois le système bien établi, par les clients. La gratuité du service pour les utilisateurs ne durera que tant que cette économie est dans l'intérêt de la plateforme. Certains systèmes offrent deja des services premium de qualité bien supérieure aux services gratuits.

Les plateformes sont plus généralement des systèmes complexes multifaces avec diverses categories d'utilisateurs, accédant à des services différents au rang desquels l'extraction de connaissances, la gestion de la propriété intellectuelle, l'interaction avec les services de l'Etat, parmi lesquels la fiscalité, la surveillance, etc.

4.2 La concentration

Les plus grandes capitalisations comptent désormais parmi les pétroliers des plateformes d'intermédiation au premier rang desquels Google et Apple. C'est dire l'importance de cette industrie inexistante il y a seulement une quinzaine d'années. La concentration qui la caractérise est pour une part liée à la force des grands nombres qui voit la qualité des services des plateformes augmenter avec leur taille.

A titre d'exemple, la faible variété dans le domaine des moteurs de recherche est révélatrice de cette situation. Quelques moteurs de recherche se partagent l'essentiel du marché mondial [7]. La Table 1 présente les parts de marché des moteurs de recherche dans quelques pays représentatifs. Google a une suprématie absolue au niveau mondial dans ce secteur d'activité, et plus largement sur l'intermédiation en general comme nous le verrons dans la Section 7.

Pays	Premier	marché	Second	marché
USA	Google	65%	Bing / Yahoo	15%
Chine	Baidu	73%	Qihoo / Sogou	8-9%
Japon	Yahoo ! Japan	51%	Google	36%
Russie	Yandex	60%	Google	25%
UK	Google	91%	Bing	5%
France	Google	92%	Bing	3%
Rép. Tchéque	Google	53%	Seznam	37%

Table 1 – Les premiers moteurs de recherche par pays en 2013

La logique de l'intermédiation est celle du "winner takes all", c'est-à-dire que la rivalité entre plateformes de tailles inégales aboutit à la disparition de la dernière au profit de la première. Les rachats ou les offres de rachats sont souvent la règle y compris pour des investissements colossaux de milliards de dollars pour des sociétés ne dégageant pas encore de benefices comme ça a été le cas pour Snapchat en 2013.

5 Un modèle disruptif

Les plateformes d'intermédiation bouleversent profondément notre organisation économique. Elles ont la capacité de se glisser entre les utilisateurs et les producteurs de services, écartant les sociétés dont l'activité principale est de fournir ces services initiaux et dont la raison d'être même se trouve très affaiblie. La raison pour laquelle les plateformes parviennent à évincer les structures en place tient à la qualité du service qu'elles peuvent mettre en place, qui est bien supérieure à celle initialement fournie. L'exemple du transport est édifiant à ce sujet.

7. `http://searchengineland.com/google-worlds-most-popular-search-engine-148089`

Les taxis font partie des industries les plus menacées aujourd'hui. Il ne fait pas de doute que l'organisation actuelle disparaitra au profit de plateformes. Si les organisations de taxis font partie des premières victimes c'est pour une part parce que le service qu'elles rendent est extrêmement décevant pour l'utilisateur. Dans beaucoup de villes du monde réserver un taxi tient de la gageure. La faible qualité de leur service n'explique toutefois que la rapidité avec laquelle ce secteur est touché. Un service de grande qualité n'empêcherait pas l'arrivée des plateformes, qui bouleverseront radicalement ce type d'organisations, en donnant à l'utilisateur un choix sur le service bien plus grand.

Bien sûr ce bouleversement n'est pas sans conséquence sociale ce qui explique à la fois la colère des organisations de taxi et la résistance des service de l'Etat. Mais l'une et l'autre ne peuvent espérer que retarder une échéance inéluctable. De la part des services de l'Etat, il est important de mettre en place les cadres légaux pour encadrer ces nouveaux services, défendre l'intérêt commun et favoriser l'emergence de plateformes sur le territoire national.

5.1 Un nouveau mode d'organisation

Les plateformes révolutionnent les modèles actuels sur deux points fondamentaux qui se déclinent de manière différente dans les différents secteurs économiques.

1. Le degré d'abstraction des services.
2. L'indépendance des utilisateurs.

L'abstraction du service augmente. La plateforme permet de se focaliser sur des services de haut niveau en s'abstrayant d'un moyen particulier de les réaliser. Dans le secteur du transport par exemple, le service consiste en un déplacement donné à un moment donné, et non plus dans le moyen comme le train ou l'avion, laissant à la plateforme le soin de proposer une gamme de choix.

L'une des propriétés essentielles des plateformes est de mettre au même niveau les consommateurs et les fournisseurs de services, les chauffeurs de taxi et leurs clients par exemple. La plateforme assure l'intermédiation entre les deux. C'est ce que faisait également la société de taxi, mais avec un service très limité. La plateforme généralise le service de multiples manières. D'abord, elle est en mesure de proposer un taxi ou tout autre moyen. Ensuite, elle le personnalise au moyen de la connaissance qu'elle a de ses utilisateurs, en leur proposant le service qui leur conviendra le mieux en fonction par exemple du prix, de la rapidité, etc. Finalement, elle introduit la recommandation, qui permet aux différents utilisateurs de signifier leur appréciation, conduisant à une augmentation de la qualité de service et au développement d'une certification apportée par la multitude des utilisateurs et non plus par l'institution qui en était garante.

Conceptuellement, la mise au même niveau sur la plateforme des consommateurs et des fournisseurs de services à une consequence majeure, ne pas les distinguer, faire de tout utilisateur un consommateur potentiel et un fournisseur potentiel de services. Ce glissement est une révolution majeure dont nous ne mesurons sûrement que très imparfaitement les consequences aujourd'hui.

Pour continuer avec l'exemple du transport, elle met le co-voiturage au même niveau que tout autre mode de transport. Le succès de ces nouveaux modes est deja évident aujourd'hui et l'usage s'en généralise, grâce en particulier à la capacité d'offrir des formes nouvelles de garantie de sécurité, assurée par la recommandation et non plus par les institutions. Elle met donc le particulier au même niveau qu'un professionnel de la conduite, une révolution qui pose toute sorte de défis à l'Etat sur lesquels nous allons revenir dans la Section 6.

Ce mouvement de mise en capacité d'offrir des services par tout un chacun sans être un professionnel reconnu, assermenté, est général. Il touche de très nombreux secteurs. Le logement saisonnier par exemple, qui permet de louer tout ou partie de son propre logement et concurrence sérieusement les professionnels de l'hôtellerie. Ils bénéficie des mêmes garanties offertes par les systèmes de recommandation.

5.2 Les secteurs du savoir

Dans le domaine de la création intellectuelle et du savoir, le phénomène est le même, mettant tout utilisateur en capacité de produire des contenus, de les distribuer et de les voir reconnus par la multitude, sans avoir besoin de la légitimité apportée par une institution.

La presse a été la première touchée par ces changements. D'abord sont venus les agrégateurs d'information qui offrent une vue sur des articles produits par des journaux. Ils permettent de développer des plateformes d'information sans staff, sans journalistes rédigeant les articles. Dans certains cas, le choix rédactionnel de ce qui est affiché est fait par des personnes, dans d'autres par des algorithmes. Il convient de souligner le potentiel des algorithmes qui sont utilisés également pour la redaction d'articles dans des domaines spécifiques aujourd'hui comme la bourse ou le sport.

Le potentiel des plateformes est également de fournir à chaque utilisateur, le "journal" qui lui convient, taillé sur mesure en fonction de ses intérêts, de ses compétences. Très clairement la plateforme Twitter se rapproche d'un tel système, combinant différents mode de transmission des informations, le suivi d'auteurs, la communication entre utilisateurs, et les recommandations automatiques de la plateforme.

L'enseignement et plus généralement la transmission du savoir sont touchés par les plateformes d'une manière absolument fondamentale. Tout d'abord l'accès au savoir est révolutionné par les moteurs de recherche qui permettent d'accéder à un corpus illimité de connaissances. Cette capacité ouverte à tous, élève comme professeur, change la nature de leurs rapports. Elle destabilise le professeur qui est amené à reconsidérer son apport à l'élève.

Cette révolution déstabilisera également les grandes institutions académiques. Leur légitimité tient aujourd'hui à leur capacité à recruter les meilleurs scientifiques et à concentrer les moyens matériels de la recherche. Pour les universités, le parcours obligé au travers d'un cursus local va être battu en brèche. Le mouvement pour la mobilité des étudiants, qui est deja bien développé en Europe, sera démultiplié par la capacité de suivre des cours en particulier à distance. Les universités devront évoluer vers une économie de plateforme permettant l'interaction renouvelée entre professeurs et étudiants, dans laquelle les étudiants ne sont plus des simples receptacles de connaissance, mais également des contributeurs. Pour une université devenir une plateforme, c'est être en capacité d'intermédier sur des savoirs produits hors de la plateforme, avec en particulier des contributeurs indépendants de toute institution.

6 La révolution politique

La disruption des plateformes touchera probablement tous les secteurs d'activité ayant une interaction avec des utilisateurs quels qu'ils soient, clients, partenaires, ou dans un autre mode de relation. Les changements induits poseront différents problèmes sociaux, liés au devenir des entreprises rendues caduques et à l'emploi. Ils poseront de sérieux défis au législateur. Si différents pays tentent de ralentir l'évolution au moyen de mesures légales qui, n'ayant d'autres justifications

que ce ralentissement, auront des effets peu durables, le véritable enjeu est d'adapter les outils de la gouvernance, de reconsidérer le rôle de l'Etat et ses moyens d'action face à des puissances déterritorialisés.

Les plateformes opèrent un déplacement du pouvoir. D'une part vers le bas, vers les utilisateurs dont les potentialités sont décuplées et qui peuvent avoir des prétentions que seule une institution ou une entreprise bien établie aurait pu se permettre autrefois. D'autre part vers le haut, vers la plateforme elle-même qui, de la quantité d'information extraordinaire qu'elle récolte, tire un pouvoir peu imaginable dans le passé.

La concentration d'information qui permet l'émergence des nouveaux services, s'accompagne d'un bouleversement considérable des équilibres de pouvoir en place. La disruption des plateformes d'intermédiation atteindra au-delà des secteurs d'activité traditionnelle de l'économie, le coeur de l'activité de l'administration et du gouvernement, dont une part importante de l'action est tournée précisément vers l'interaction avec la population.

Les plateformes sont de surcroît des entités opérant au niveau global, dont l'impact sur les territoires se fait par des moyens déterritorialisés. L'action du législateur, qui précisément porte sur un territoire, est parfois limitée face aux plateformes. Nous verrons de surcroît dans la Section 7, que les plateformes, qui opèrent sur l'ensemble de la planète, sont concentrées dans seulement quelques pays au premier rang desquels les Etats-Unis. L'Europe par exemple ne dispose d'aucune grande plateforme et dépend complètement des plateformes américaines.

6.1 Surveillance et transparence

La disponibilité de masses considérables de données a changé le monde de manière irréversible. Les révélations d'Edward Snowden à partir de juin 2013 ont fait prendre conscience à l'opinion publique internationale d'un facette de ce changement, la surveillance étatique généralisée. La capacité des Etats à étendre le spectre d'application de la surveillance des services de sécurité d'une infime partie de la population à essentiellement toute la population connectée, et ce quelque soit son territoire, est un changement radical de paradigme, auquel la science fiction avait consacré certains de ces meilleurs ouvrages un demi siècle avant l'émergence de cette surveillance numérique. L'utilisation par les entreprises des données personnelles pose aussi des problèmes qui feront l'objet de débats continus dans la période à venir, tant est peu satisfaisante la solution du consentement personnel aux conditions générales d'utilisation.

Dans le même temps que les Etats se dotaient de moyens de surveillance très extensifs, le mouvement pour les données ouvertes de l'administration prenait de l'ampleur, conduisant à l'adoption sous l'impulsion du Royaume-Uni d'une charte par le G8 en 2013 prônant une politique de données ouvertes par défaut, ainsi que dans le même temps l'adoption d'une directive européenne renouvelée sur le sujet. Sont exclues du champ d'application les données personnelles, les données protégées par la propriété, ainsi bien sûr que les données touchant à la sécurité nationale.

On peut facilement imaginer que les données ouvertes de l'administration permettront le développement de nombreux services aux utilisateurs qui deviendront incontournables. Même si de nombreuses sociétés émergent dans ce secteur, il est probable que les meilleurs services seront offerts par des sociétés concentrant une masse de données ouvertes considérables, et qui généreront des données secondaires, résultants des traces d'utilisation, qui leur assureront une vraie suprématie dans une nouvelle forme de gouvernance et de régulation algorithmiques.

Ce qui assure la suprématie des grands moteurs de recherche, ce n'est en effet pas l'accès aux pages Web qui est ouvert à tout le monde, mais leur exclusivité sur les données secondaires.

La situation des données ouvertes de l'administration est similaire, les données ouvertes seront accessibles à tous, mais ce sont les services innovants qui généreront un flux de données secondaires à l'origine de nouvelles applications qui très certainement rivaliseront avec les services de l'Etat, mais dont l'Etat n'aura pas la maîtrise. La situation dans le secteur de la gouvernance a peu de raison de différer des secteurs économiques déjà touchés par les plateformes.

La propriété des données secondaires des grandes plateformes qui aujourd'hui est détenue par les dites plateformes pourrait faire l'objet d'un débat. En faire des biens communs pourrait constituer une alternative au modèle actuel, même si à ce stade cela poserait de nombreux problèmes de protection des données qui relèvent aujourd'hui de la responsabilité des plateformes.

6.2 Nouveaux services publics et fiscalité

Les grandes plateformes assurent désormais des services qui sont devenus incontournables et dont ni les particuliers ni les entreprises ne sauraient se passer aujourd'hui. Ces services ne sont pas sans rappeler les services publics que l'Etat assure, ou ceux que des entreprises assurent dans le respect de missions de service public définies avec l'Etat. Les grands principes du service public que sont la continuité de service, la non discrimination, et l'adaptation du service aux évolutions font partie des principes garants de la qualité de service offerts par les grandes plateformes à leurs utilisateurs.

Ces services sont pour la plupart offerts gratuitement aux utilisateurs, ce qui implique que leur financement est assuré ou sponsorisé par des tiers, aujourd'hui selon le modèle de l'économie biface. On pourrait même considérer que les plateformes lèvent en quelque sorte un impôt planétaire pour assurer un service essentiel de qualité.

La question de la fiscalité est essentielle. Laissons de coté les questions d'optimisation fiscale des grands groupes qui ne sont dues qu'à une organisation sous-optimale des Etats et qui ne sont pas propres aux plateformes. L'intermédiation est souvent associée dans l'économie traditionnelle à un échange qui est fiscalisé. C'est le cas de tout échange marchand en particulier.

Les plateformes introduisent une nouvelle forme d'économie de l'échange qui est d'ailleurs extrêmement pratique pour leurs utilisateurs qui voient certaines de leurs transactions arrangés à l'avance par la plateforme, comme pour le co-voiturage par exemple. Ces échanges peuvent se faire dans une monnaie propre à la plateforme, qui échappe ainsi à la fiscalité traditionnelle, mais pas à celle de la plateforme qui prend un pourcentage des transactions. Introduire de nouvelles formes de fiscalité des échanges non monétisés de manières traditionnelles est donc un enjeu important.

L'un des enjeux, et pas des moindres, auquel font face nos sociétés est celui d'un développement plus parcimonieux dans l'usage des ressources et plus respectueux des équilibres naturels qui semblent affectés par l'activité humaine. Un meilleur usage des ressources supposent une meilleure connaissance de comment les partager mieux. Les exemples abondent où une intermédiation efficace apporterait un gain important. On pense bien sûr immédiatement au co-voiturage si difficile à réaliser et dont les plateformes ont permis l'émergence. Mais c'est également l'emploi, l'intermédiation entre l'offre et la demande d'emploi. Cette intermédiation se jouera également au-delà des personnes, entre les objets producteurs et consommateurs d'électricité par exemple qui optimiseront l'efficacité énergétique au niveau global. Les potentialités sont immenses qui formateront le monde dans lequel nous vivrons, définies par une industrie, celle des plateformes.

6.3 L'identité et le fonctionnement démocratique

La connaissance qu'acquièrent les plateformes de leurs utilisateurs permet le développement de nombreux services, parmi lesquels les assistants personnels qui sont en mesure d'assister leurs utilisateurs comme le feraient une personne, si ce n'est peut-être mieux. Ces connaissance permettent également de reconnaître un utilisateur, en particulier dans la mobilité, par le croisement de différents types d'information, comme son environnement numérique, ses connaissances propres vérifiables, son comportement, combinés potentiellement à des indicateurs biométriques.

Cette capacité des plateformes à reconnaître avec une grande probabilité un utilisateur dans le vaste monde leur permet de leur offrir un service d'authentification fiable et utilisable vers tous les services extérieurs à la plateforme qui le souhaitent. L'authentification Facebook ou Google par exemple est largement adoptée y compris par des services administratifs.

la gestion de l'identité est un défi fondamental et stratégique, assurée de plus en plus dans un espace déterritorialisé. Elle revêt potentiellement des enjeux commerciaux. En juin 2014, Flickr prévenait par exemple ses utilisateurs que la possibilité de se connecter avec un identifiant Facebook ou Google serait supprimée. Mais au-delà, il s'agit bien de la maîtrise de la population. L'identité est nécessaire à l'organisation par l'Etat de nombreuses activités essentielles à son fonctionnement, comme la conscription, la levée de l'impôt, la sécurité ou les élections en particulier.

Les moyens dont disposent l'Etat dans la gestion de l'identité sont rudimentaires comparés à ceux des plateformes. Il ne fait pas de doute, que les services de l'Etat devront recourir aux plateformes pour de nombreuses questions. Celle de la sécurité a déjà été amplement démontrée en particulier par les révélations d'Edward Snowden. Le contrôle des passagers dans le transport aérien est déjà en partie centralisé. La gestion des visa au niveau international le sera certainement, tant le dispositif est aujourd'hui coûteux pour les services consulaires et rudimentaires.

Le fonctionnement démocratique ou simplement politique évoluera sous la pression des plateformes, tant la quantité de pouvoir qu'elles concentrent change la donne. Nous avons abordé différents aspects ci-dessus, la maîtrise de l'information, la fiscalité, les services numériques essentiels, la gestion de l'identité déterritorialisée, qui tous affectent en profondeur notre organisation politique. La concentration des plateformes dans quelques pays que nous décrivons dans la section soulève un autre aspect des nouveaux équilibres mondiaux.

7 Les équilibres internationaux

Les plateformes d'intermédiation suivent comme nous l'avons vu ci-dessus une loi de concentration dictée par la croissance forte de leur efficacité en fonction du nombre d'utilisateurs. Le très faible nombre de plateformes ayant une influence internationale est un indicateur explicite de cette concentration. En effet, moins d'une cinquantaine de plateformes parviennent à figurer dans les 25 premiers sites de plus d'un pays, c'est-à-dire dans un pays en plus du pays d'origine de la plateforme. La Figure 2 montre le nombre x de plateformes présentent dans le Top 25 de y pays. On voit que la décroissance est très rapide. Seules 14 plateformes sont présentes dans le Top 25 de 10 pays alors que 21 plateformes sont présentes dans le Top 25 de 5 pays.

Si on considère plus en détail les 21 plateformes présentes dans le Top 25 de 5 pays, montrées en Figure 3, on découvre que toutes, à l'exception de la dernière qui est chinoise, sont américaines. La concentration ne résulte donc pas seulement de leur modèle économique, mais également clairement des politiques nationales.

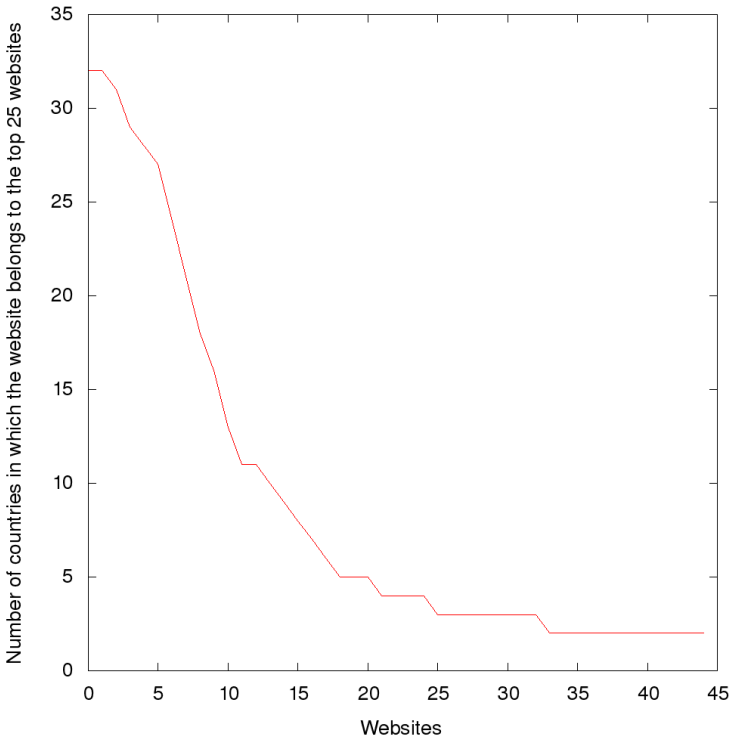

FIGURE 2 – Influence internationale des premières plateformes mondiales

Une étude[8] de la société Pingdom qui inspecte le Web montre l'extreme concentration des serveurs du Web sur la planète. Trois villes, Houston, Mountain View et Dallas, hébergent 10% du premier million de sites Web du monde, suivant le classement établi par Alexa. Les Etats-Unis occupent une position absolument centrale en hébergeant globalement 43% de ces sites, dont la moitié d'ailleurs dans seulement une dizaine de villes. Si on considère maintenant, les 100 premières villes mondiales qui hébergent le plus de sites, 69 sont américaines, 7 sont chinoises, et 4 seulement sont dans l'Union européenne, Amsterdam, Berlin, Paris et Londres.

Cette concentration se retrouve également dans la capacité de stockage et de traitement des données. Le stockage et la gestion des données sont désormais assurés de manière croissante par des compagnies proposant des services d'hébergement, comme Microsoft ou Amazon par exemple, et non plus par les compagnies qui produisent et exploitent les données. Cette dématérialisation s'accompagne également d'une concentration des données par les compagnies qui offrent les services de nuage et dont les plus efficaces se trouvent aux Etats-Unis.

Cette concentration a bien sûr d'importantes conséquences sur l'activité économique dans ces grands centres du Web. Mais au-delà, c'est plus généralement la conception même de la société de l'information qui se dessine dans ces centres, et comme l'a bien souligné le blog de Technollama[9], le cadre légal dominant également.

8. `http://royal.pingdom.com/2013/03/07/hosting-locations-2013/`
9. `http://www.technollama.co.uk/where-is-the-worlds-content-hosted`

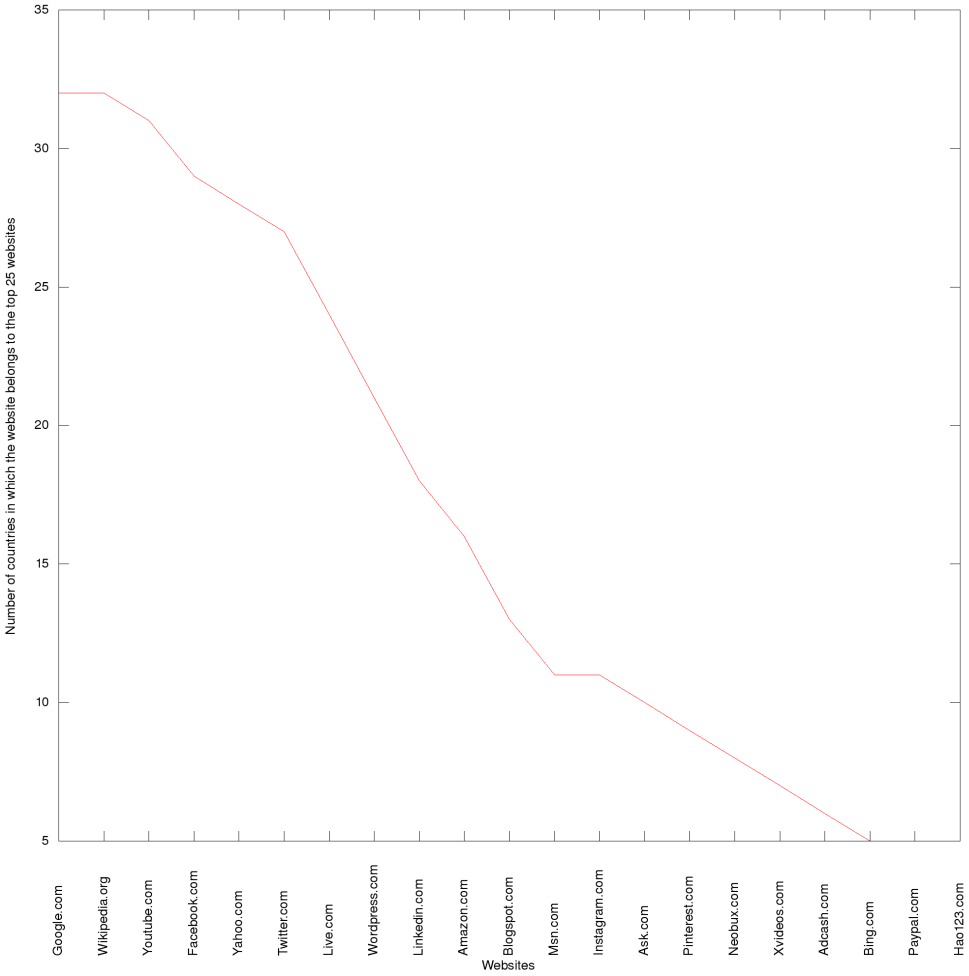

FIGURE 3 – Influence internationale des 21 premières plateformes mondiales

Nous avons essayé de mesurer les flux internationaux de données saisies en ligne [10]. La majeur partie des données récoltées en France est exportée aux Etats-Unis. La carte de la Figure 4 montre l'importance de ces flux entre quelques pays représentatifs. Les données sont extraites des flux sur les dix plus grands sites de chaque pays.

8 Conclusion

Les plateformes d'intermédiation sont en train de changer radicalement l'organisation et la gouvernance du monde. Jamais auparavant un tel degré de concentration d'information sur les activités humaines à l'échelle de la planète n'avait été possible. Cette concentration d'information ne peut que résulter dans une concentration de pouvoir du même ordre de grandeur. Les plateformes d'in-

10. Stéphane Frénot et Stéphane Grumbach. Les données sociales, objets de toutes les convoitises. La Découverte Hérodote 2014/1 - n 152-153

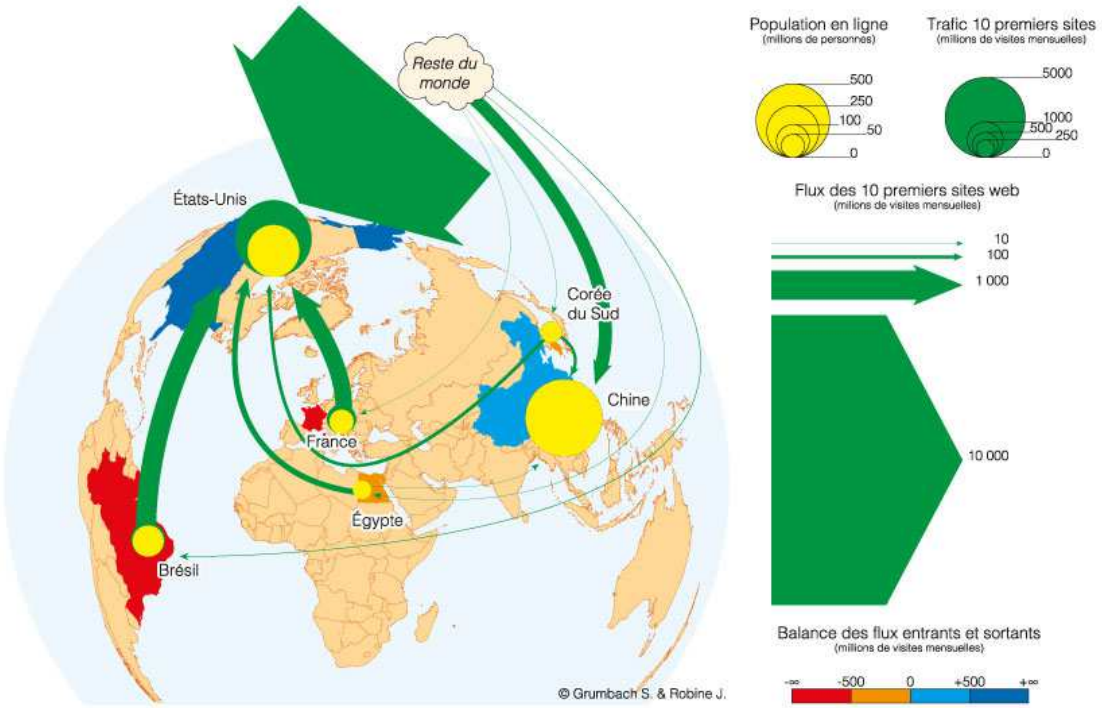

FIGURE 4 – Carte des flux mondiaux entre quelques pays

termédiation parce qu'elles assurent au niveau de la planète entière les nouveaux services publics de l'ère numérique, établissent un lien direct avec leurs utilisateurs, particuliers ou entreprises, qui fait fi des frontières et bien souvent des réglementations locales. Les plateformes concentrent également la richesse, et d'une certaine manière, dans la mesure où elles assurent un service public gratuit, il n'est pas absurde de considérer qu'elle lève désormais un impôt sur un espace déterritorialisé. Les plateformes sont en mesure d'établir une identité numérique d'une très grande fiabilité en regard de celle offerte par les services des Etats, qui joue déjà un rôle important pour la sécurité. Avec l'ouverture des données de l'administration, les plateformes devraient être en mesure d'offrir des services innovants qui remplaceront probablement avantageusement ceux offerts par les administrations. Le déclin des finances publiques qui touche de nombreux pays, associé à l'extraordinaire croissance des capacité financières des plateformes devraient favoriser un transfert de compétence des Etats vers les plateformes pour des services pour lesquels les plateformes présentent une claire plus-value. Ce déplacement de compétence poussera inéluctablement vers la périphérie les administrations territoriales, quelque soit le statut des territoires. Enfin la concentration géographique des plateformes sur un très petit nombre de pays devraient conférer à ceux-ci une claire prééminence dans la gouvernance mondiale.