

Merci à Jeanne Fine, pour cet envoi.

Le Monde, 8 janvier 2013

Les données, puissance du futur

L'Europe reste à la traîne dans l'exploitation des données électroniques. Des géants étrangers dominent ce secteur stratégique et déterminant

Nous vivons une période de rupture, celle de la numérisation de tout : l'homme, la société, les organisations, le savoir, les interactions, le corps biologique de chacun, etc. Tout devient donnée numérique, et par là même a vocation à une existence dans le monde virtuel en émergence. L'écart entre le monde des données et le monde réel, bien anticipé par Jorge Luis Borges dans sa nouvelle visionnaire *De la rigueur de la science*, évoluera pour devenir de plus en plus subtil. Il y a un avant et il y a un après-numérisation. Il nous est aussi difficile d'imaginer les évolutions de la société de l'information qu'il l'était de prévoir au milieu du XIXe siècle les possibilités permises par l'électricité. Les données constituent les briques de base de la société de l'information. Leur quantité est en croissance exponentielle. Alors que la loi de Moore, prédisant le doublement de la capacité de calcul des ordinateurs tous les dix-huit mois, touche à sa fin, les données semblent avoir usurpé à leur profit cette prédiction. Désormais qualifiées de *Big Data* dans le monde anglo-saxon, elles représentent déjà des masses considérables pour lesquelles on recourt à de nouvelles mesures. L'infrastructure physique de la société de l'information, les systèmes de télécommunications, les centres de stockage et de traitement des données, les nouveaux services en ligne, constituent des secteurs industriels qui connaissent une croissance inégalée. Les centres de données consommeront bientôt autant d'électricité qu'un pays comme la France.

Les données en elles-mêmes offrent un potentiel extraordinaire que l'on commence à exploiter. Elles permettent de générer des connaissances, qui étaient soit hors d'atteinte, soit inexistantes, parce que hors du domaine du pensable. Une nouvelle médecine se développe, qui, grâce aux données personnelles, sera en mesure de proposer des traitements adaptés à chacun et non plus calibrés pour des populations. De nouveaux champs scientifiques sont ouverts avec des découvertes réalisables sur les masses de données accessibles.

Une nouvelle économie émerge, qui exploite les données pour des services à valeur ajoutée. Par exemple, Google réalise, grâce à ses données, un chiffre d'affaires de plus d'un milliard d'euros en France. Dans un autre registre, le cabinet de consultants McKinsey estime le potentiel économique annuel du Big Data pour le système de santé américain à 300 milliards de dollars, soit 1 000 dollars par habitant et par an !

La société de l'information est apparue au début des années 1990 avec l'émergence de la Toile. La multiplication rapide des pages Web a très vite nécessité l'invention de nouveaux outils pour accéder aux pages pertinentes : les moteurs de recherche. Deux révolutions majeures accompagnent l'introduction de ces systèmes : le développement d'une industrie du traitement des données à des échelles ambitieuses, et l'établissement du modèle économique qui domine la Toile, un service gratuit en échange de données personnelles. Les moteurs de recherche exploitent ainsi les requêtes faites par les utilisateurs, dont ils peuvent dériver des profils qui permettent de cibler des publicités, mais également de dégager des connaissances globales, comme l'extension des épidémies de grippe au niveau mondial.

Cette révolution prend son essor dans les années 2000 avec le développement des systèmes interactifs et collaboratifs comme les réseaux sociaux, qui permettent aux utilisateurs de contribuer. Les données personnelles, tant celles produites par les usagers (textes, photos, vidéos, etc.) que celles générées par les systèmes que nous utilisons souvent à notre insu, sont au cœur de l'économie de la société de l'information, et donc de l'économie.

Les données sont devenues une ressource, peu différente des matières premières, comme le charbon ou le minerai de fer, et dont l'importance économique dépassera celle du pétrole. La comparaison avec le pétrole illustre une caractéristique essentielle de l'économie des données personnelles : la concentration. Le pétrole est, pour des raisons géologiques, concentré dans des puits, d'où il est extrait et acheminé après transformation dans les foyers des particuliers sur toute la surface de la terre. Les

données personnelles du Web 2.0 sont à l'inverse récoltées chez les particuliers sur la surface de la terre, pour être acheminées vers les centres de données de multinationales, qui ont le monopole de leur traitement, comme Google, Facebook ou Amazon, et qui concentrent 80 % de ces données aux Etats-Unis. La décision rendue le 3 janvier par la *Federal Trade Commission* aux Etats-Unis de ne pas poursuivre Google pour la manière dont son moteur de recherche favorise ses propres outils de vente en ligne contribuera au renforcement de sa position dominante et montre la tendance lourde à la concentration dans l'industrie de la donnée.

Pour le pétrole, comme pour les autres matières premières, nous nous sommes attachés à développer une chaîne industrielle cohérente allant de la prospection à la production de produits finis, en passant par l'exploitation, le transport, la transformation... Cela n'a pas été simple. Il a fallu acquérir des connaissances, construire de grandes entreprises, mettre au service de ces enjeux une volonté politique forte. Qu'en est-il pour les données numériques ? Dans ce domaine, l'Europe a fait l'impasse. En ne construisant pas d'industrie du Web 2.0, elle s'est privée de l'accès à la ressource, y compris à celle provenant de son territoire. Pour des raisons historiques et politiques, l'Europe a peur des données. Elle voit dans la société de l'information une menace qu'il convient de circonscrire, et qui semble inhiber toute vraie ambition.

La faiblesse en données et en capacité de traitement de la donnée résulte d'une asymétrie d'information. L'économiste Joseph Stiglitz a montré les conséquences sur les marchés de cette asymétrie. Grâce à l'analyse des requêtes sur son moteur, Google sait, ou est en capacité de savoir, plus de choses que l'Insee sur la France.

La maîtrise de la donnée permet aussi la maîtrise de certains marchés qui transitent déjà dans certains domaines par les outils de commerce électronique américains. Faute de développer cette industrie, il est probable qu'à brève échéance nous achèterons de nombreux biens et services, produits et consommés en France comme nos billets de train ou notre électricité, à un prestataire étranger qui dégagera une part importante de la valeur ajoutée et aura le contrôle de la chaîne industrielle.

La maîtrise de la société de l'information donne une puissance qu'on soupçonne encore peu et qui dépasse de loin les secteurs de l'économie marchande. A titre d'exemple, les systèmes de cours en ligne, qui connaissent une croissance rapide aux Etats-Unis, et qui finiront par révolutionner l'enseignement, permettront aussi une maîtrise des ressources humaines au niveau mondial, stratégique au moment où les pays industrialisés feront face à un déficit croissant de personnels scientifiques. La société de l'information pénètre des domaines moins visibles. Certains services régaliens, comme l'identité des personnes, pourraient même être assurés demain par des multinationales. Le réseau social Facebook permet une authentification des personnes qui n'a pas d'égale en qualité et pourrait devenir incontournable. Le Royaume-Uni envisage d'ailleurs de l'utiliser pour l'accès aux services publics en ligne. La maîtrise des données est en fin de compte liée à la sécurité et à l'indépendance nationales. L'absence de l'Europe des systèmes du Web 2.0 est-elle une fatalité ? La captation de données est la priorité absolue des Etats-Unis, qui dominent la société de l'information à l'échelle planétaire. Mais d'autres pays ont des vues stratégiques dans ce domaine. La Chine tout d'abord, qui détient 16 % des 50 premiers sites mondiaux, à côté des 72 % américains. Dans ces deux pays, les données nationales restent sous contrôle de l'industrie nationale. Et tous deux ambitionnent de récolter la donnée à l'international. Quant aux pays plus petits, certains parviennent à équilibrer leur industrie. La Corée du Sud s'appuie sur un petit tiers de sites nationaux, un tiers de sites américains et un tiers de sites chinois. D'autres pays aussi différents que le Brésil, la Russie ou l'Iran sont en avance sur l'Europe en la matière.

En Europe, il y a un usage monolithique des sites américains. En France, Google détient l'une de ses plus grosses parts de marché mondial (92 %) et il en va de même pour Facebook. Quant aux premiers sites français, aucun n'accumule de la donnée. Cette situation est plus qu'alarmante. Nous disposons pourtant en France d'un moteur de recherche, Exalead, développé par Dassault Systèmes. Pourquoi ne fait-on pas de son développement une cause nationale ? Il existe par ailleurs des moteurs qui ne conservent pas de données personnelles, comme Ixquick.

Sans entrer dans la compétition internationale, sans disposer de géants du stockage et du traitement des données et des services, et en particulier d'un moteur de recherche, de réseaux sociaux, de systèmes de blogs et de micro-blogs, de *cloud*, détenant des parts de marché importantes, non seulement en Europe, mais également dans le reste du monde, il est peu probable que la voix de l'Europe puisse se faire entendre et qu'elle contribue à façonner les grandes orientations de la société de l'information.

Nos pays sont condamnés à un rôle subalterne, à la protection illusoire d'industries, dont les modèles sont en pleine évolution, et que nous sommes incapables de réinventer. Trois aspects fondamentaux font le succès des systèmes américains : la qualité de service, le génie des applications et la pertinence du modèle économique associé. Pour réussir, l'Europe doit promouvoir un modèle original compatible avec ses principes, mais qui répond autant à des besoins que les systèmes que les Américains ont développés et qui ont changé le monde.

Stéphane Grumbach¹ Stéphane Frénot²

1 Directeur de recherche à l'Institut national en recherches informatiques avancées et ancien conseiller scientifique en Chine. Il travaille sur les nouveaux équilibres induits par la société de l'information

2 Professeur des universités au laboratoire CITI de l'Institut national des sciences appliquées de Lyon. Ses recherches portent sur les systèmes informatiques communicants au cœur du Web et de ses applications.

Marc Bourdeau, 2113.01.12 - Petite note. Quant à citer Borgès, on pourrait aussi citer sa nouvelle «La Grande bibliothèque», où tous les livres possibles sont rangés dans un ordre méticuleux, tous de même format... Tous, absolument tous. Tous les écrits de tant de pages, chacune ayant le même nombre de lignes, etc. Même ceux qui sont aléatoires. Tous, absolument tous les livres rangés dans des rayons identiques de la Grande bibliothèque, elle-même une merveille d'architecture alvéolaire, une grande ruche à stocker/produire du savoir.

On comprend que tout, sauf de rarissimes exceptions, perdues dans de rarissimes livres, un vrai mot par çà par-là, n'est que du charabia aléatoire, et pourtant tout Shakespeare est là quelque part, et tout Shakespeare avec une coquille, deux coquilles... On comprend aussi que tous ceux qui fréquentent ce monde, sont à la recherche de sens, désespérément, comme nous tous n'est-ce pas ?

Ils ne trouvent rien, beaucoup se suicident de désespoir dans les rayonnages infinis. Ceux qui n'osent pas tentent désespérément de trouver ne fût-ce qu'un petit epsilon de sens, en lançant des cornets à dés, dont les dés sont des lettres...

N'est-ce pas aussi là une métaphore de ce déluge, que dis-je, de ce tsunami d'informations inutiles, incontrôlables, que nous fournissent les logiciels des immenses bases de données bibliographiques—non prévues dans le Borgès, mais cela ne vient qu'ajouter au malaise ! Tout n'est-il pas dans Wikipedia, tout, n'est-ce pas, se trouve sur les murs des Facebook de tout un chacun. Une métaphore de notre nouveau monde intellectuel. Ce monde, le monde est devenu trop complexe, trop étourdissant. Vivement le cornet à dés !