

PAR PAUL HERMELIN (72)



P-DG de Capgemini, cochef de file du plan big data, dans le cadre de l'initiative La Nouvelle France industrielle

Une opportunité pour la France dans l'économie de l'innovation

Les milliards de données que les citoyens créent dans l'infosphère constituent une mine dont l'exploitation va révolutionner la société autant que les entreprises. Forte de son potentiel d'informaticiens et de mathématiciens, la France a là une opportunité de repartir de plain-pied dans l'économie de l'innovation à condition de desserrer les contraintes qui pourraient freiner l'essor de cette activité.

Une révolution pour nos sociétés

Le monde crée aujourd'hui en deux jours autant de données que toute l'humanité en a créées pendant deux mille ans. C'est l'avènement des *big data* qui émergent sous nos yeux en combinant trois caractéristiques : volume, vitesse, variété. Le volume des données générées augmente de 30 % à 40 % chaque année. Les vitesses de traitement sont proches aujourd'hui de l'instantané dans de nombreux cas (no-

Le volume des données générées augmente de 30 % à 40 % chaque année

REPÈRES

En avril 1812 les « luddites », des artisans tisserands, attaquent une fabrique de draps dans le Yorkshire en Angleterre. C'est le premier conflit industriel violent de notre temps, qui opposa artisans de l'ancien monde et nouveaux industriels. Au cours du siècle qui suivit, la révolution industrielle du Royaume-Uni puis celle du monde entier procurèrent à l'homme une prospérité matérielle inconnue jusqu'alors. Bien sûr, cette révolution ne se déroula pas sans heurt, tandis qu'elle engendrait une transformation fondamentale des métiers et des activités humaines. Nous sommes à l'aube d'une telle révolution, et la responsabilité incombe au politique d'agir sans attendre pour gérer au mieux la transition sociétale qui s'ouvre devant nous.

Lutter contre la fraude

Le groupement Visa estime aujourd'hui déjouer deux milliards de dollars de tentatives de fraude par an. De 40 critères en 2005, Visa est passé à 500 critères aujourd'hui. Et les *big data* ont marqué la fin de l'échantillonnage : désormais, 100 % des transactions sont analysées, contre 2 % en 2005.

tamment dans le commerce en ligne). La variété des données s'est fortement accrue ; celles-ci sont issues de sources toujours plus nombreuses. La multiplication des équipements (smartphones, capteurs divers) et des opportunités de création de données (blogs, réseaux sociaux) créent une richesse et une complexité des données croissantes.

Un « thesaurus » de milliards de données

Il y a longtemps, on prenait les données et on les mettait dans des registres, dans des bases de données. Maintenant, on exploite toutes les données que laisse l'utilisateur dans l'infosphère, en utilisant son téléphone portable, en se connectant à un site Web, en payant ses achats, etc. On récupère ainsi des milliards de données, qui deviennent un « thesaurus » de données donnant des indications sur ce que les gens font, sur ce que les gens ont envie de faire, peut-être aussi ce qu'ils veulent éviter de faire. Cela pose des problèmes d'anonymat à l'évidence, mais c'est un matériau très riche qu'il nous faut exploiter pour créer de la valeur en France et pour la France.

Naissance de nouvelles technologies

Les technologies sont prêtes pour exploiter les données recueillies en masses considérables et dans des formats toujours plus variés. Ce sont ainsi de nouveaux modes de « rangement » ou de visualisation. Google a jeté les bases d'un système de fichiers distribué qui a



© REUTERS

Le siège de la NSA. L'affaire Snowden a servi de révélateur.

inspiré la mise au point de Hadoop et MapReduce, technologies sous l'égide de la Fondation Apache permettant de répartir le traitement des données structurées ou semi-structurées sur de multiples serveurs à bas coûts. D'autres solutions de *big data* sont apparues, comme les bases dites NoSQL, dont la particularité est d'aller au-delà des limites imposées par le SQL, le langage standard d'interrogation des bases de données structurées (NoSQL signifiant *Not only SQL*, soit « pas seulement SQL »).

D'autres solutions orientées temps réel ont aussi fait leur apparition, et notamment la technologie dite « SAP HANA », qui permet le traitement des données en instantané grâce à un stockage en mémoire (*In Memory*) plutôt que sur un disque. Ces différentes technologies représentent des réponses à l'évolution des besoins fonctionnels des entreprises.

Cinq enjeux pour les entreprises

Les enjeux des *big data* peuvent être considérés selon cinq angles d'approche possibles. La première approche consiste à affiner les analyses stratégiques, en fournissant des éléments analytiques inaccessibles auparavant – notamment en considérant les phénomènes non plus sous une forme moyenne mais en tenant compte de leur diversité. Une deuxième approche est fondée sur l'amélioration des processus cœur de métier. Par exemple, les opérateurs du secteur des télécommunications ont là des outils pour réduire le taux de déperdition sur leur base de clients.

Une troisième approche est celle qui vise à accélérer la prise de (meilleures) décisions grâce à de nouveaux systèmes d'information d'entreprises (par exemple le *pricing* dynamique en temps réel des produits de grande

consommation). Une quatrième approche vise à exploiter l'évolution de la chaîne de valeur d'une industrie sous l'effet des données *via* la transformation du *business model* existant. C'est le cas de l'intermédiation des hôteliers développée par Booking.com. Enfin, une dernière approche est la création de revenus entièrement nouveaux.

Un rôle particulier pour la France

Comment mettre ces données intelligentes au service du progrès et de la croissance ? Comment les exploiter tout en respectant la vie privée ? Je porte sur ce sujet à la fois le regard d'un acteur des *big data* à qui le gouvernement a confié la coordination d'un plan industriel sur le sujet, et le regard d'un consommateur et d'un citoyen. Le marché des *big data*, au sens le plus strict du terme (c'est-à-dire la vente de logiciels ou de services spécialisés), est estimé en France à environ 500 millions d'euros par an aujourd'hui, et près de 2 milliards d'ici trois ans. Mais cela n'est que la (toute petite) pointe de l'iceberg. Le véritable enjeu est ailleurs : il est dans la transformation numérique des entreprises et de notre société. Transformation qui se produira de toute façon, soit volontairement, soit contrainte et forcée. Reprenons l'exemple de l'hôtellerie en France : elle représente environ 185 000 emplois directs. Or Booking.com capte 23 % des revenus hôteliers pour toute réservation faite *via* son intermédiaire. En faisant l'hypothèse (certes un peu simpliste), de la proportionnalité entre revenus et emplois directs, 42 000 emplois seraient

100 %

des métiers
du marketing
direct touchés
par les big data

Innovation publicitaire

Criteo, une start-up française fondée en 2008 désormais cotée au Nasdaq, est aujourd'hui leader devant Google du « recyclage publicitaire ». Par exemple, Criteo offre aux sites marchands du Web la possibilité d'adresser un message personnalisé aux internautes qui quittent leur site sans faire d'achats, afin de les faire revenir sur le site.



© iStock

La lutte contre la fraude fait de plus en plus appel aux *big data*.

- ▶ directement menacés. Aujourd'hui, Booking.com référence 420 000 hôtels dans le monde (alors que la France compte au total un peu moins de 17 000 hôtels) et a une part de marché de 70 % sur les réservations *online* en France.

Mobiliser informaticiens et mathématiciens

La première révolution industrielle a vu notre société devenir un monde d'usines. Le monde de demain sera un monde de plateformes. Lors de la première révolution industrielle, c'est l'éducation des masses qui permit de traiter le problème des ruptures des métiers, et qui fit passer en deux siècles notre société d'une situation où 75 % des travailleurs étaient dans les champs, à 2 % aujourd'hui. De nombreux métiers actuels vont disparaître dans les décennies à venir. Mais de nombreux nouveaux métiers vont apparaître.

Or, la France possède un atout unique par rapport à ces nouveaux modèles : son système éducatif. Informaticiens et mathématiciens français sont plébiscités dans le monde. Les

Le monde de demain sera un monde de plateformes

Une révolution multiforme

De nombreux secteurs ont été, ou sont en train d'être révolutionnés. Les majors du disque disparaissent au profit des plateformes de *streaming* comme Deezer ou Spotify. Les agences de voyages disparaissent au profit de Booking.com ou Expedia. Demain, les assureurs verront Google débarquer dans leur pré carré. Toutes les industries traitant des clients finals évoluent vers une intermédiation par de nouveaux acteurs, Google étant le premier d'entre eux.

big data, c'est l'alliance de ces deux disciplines. Là doit porter l'effort de notre pays, pour construire un avantage concurrentiel, créer de nouveaux Criteo, et gagner la bataille mondiale des *big data*.

Aider les entreprises à accélérer

Deux contraintes majeures pèsent aujourd'hui sur l'essor des *big data* : la disponibilité des ressources humaines d'un nouveau type – nommées *data scientist* –, et les règles de manipulation des données, héritées d'un monde où les réseaux sociaux n'existaient pas.

Concernant la formation, les fondements sont là, grâce à l'esprit même de notre système éducatif. Mais l'État doit augmenter dès la rentrée 2014 les contingents de spécialistes, pour fournir à l'industrie *hi-tech* française des ressources, rares dans le monde entier, qui leur permettront d'aller plus vite, plus loin que leurs concurrents internationaux.

Revisiter la loi Informatique et Libertés

Sur l'aspect réglementaire, nous vivons aujourd'hui avec des principes, formalisés dans le cadre de la loi Informatique et Libertés de 1978. Récemment, l'affaire Snowden a confirmé la nécessaire prudence dont nos sociétés doivent faire preuve dès lors qu'il s'agit d'autoriser (ou non) l'usage des données informatiques. Mais ne faudrait-il pas revisiter ces principes, au risque sinon de prendre du retard sur les pays qui, eux, ne s'embarassent pas de telles précautions ? Bien sûr, il ne s'agit pas d'ouvrir les données à tous vents sans contrôles ni contreparties. Il s'agirait plutôt d'adopter une démarche d'expérimentation, fondée sur la notion de réciprocité, et qui permettrait à nos *start-ups* d'exprimer toute leur créativité et de pouvoir conquérir le monde.

Un levier de croissance

Les *big data* sont plus que naissants, mais ils ne forment pas encore un marché structuré. Je considère qu'ils constituent une chance formidable pour notre pays de repartir de plain-pied dans l'économie de l'innovation. C'est un levier de croissance majeur.

L'État doit aider nos entreprises à accélérer, en menant des actions fortes de formation. Et par ailleurs engager sans attendre la réflexion sur l'évolution de notre cadre réglementaire. ■