

**Séminaire  
Ressources technologiques  
et innovation**

organisé avec le soutien de la Direction générale de la compétitivité, de l'industrie et des services (ministère du Redressement productif) et grâce aux parrains de l'École de Paris :

Algoé<sup>2</sup>  
ANRT  
CEA  
Chaire "management de l'innovation"  
de l'École polytechnique  
Chambre de Commerce  
et d'Industrie de Paris  
CNES  
Conseil Supérieur de l'Ordre  
des Experts Comptables  
Crédit Agricole SA  
Danone  
EADS  
Erdyn  
ESCP Europe  
Fondation Charles Léopold Mayer  
pour le Progrès de l'Homme  
Fondation Crédit Coopératif  
Groupe ESSEC  
HR VALLEY<sup>2</sup>  
HRA Pharma  
IDRH  
IdVectoR<sup>1</sup>  
Institut de l'entreprise  
La Fabrique de l'industrie  
La Poste  
Lafarge  
Mairie de Paris  
MINES ParisTech  
Ministère du Redressement productif,  
direction générale de la compétitivité,  
de l'industrie et des services  
NEOMA Business School  
OCP SA  
Orange  
PSA Peugeot Citroën  
Saint-Gobain  
SNCF  
Thales  
Total  
UIMM  
Ylios

<sup>1</sup> pour le séminaire  
Ressources technologiques et innovation  
<sup>2</sup> pour le séminaire Vie des affaires

(Liste au 1<sup>er</sup> avril 2014)

**PEUT-ON CRÉER UN ÉCOSYSTÈME  
FRANÇAIS DU BIG DATA ?**

par

**François BOURDONCLE**  
Copilote du plan Big Data

Séance du 6 février 2014  
Compte rendu rédigé par Pascal Lefebvre

**En bref**

Les barbares sont-ils à nos portes ? Domaine aussi mystérieux pour le profane qu'enthousiasmant pour ceux qui y voient de nouveaux territoires économiques à conquérir, le monde des Big Data fait brutalement irruption dans les activités les plus inattendues. Déplaçant la valeur créée dans les entreprises de la technicité pure vers la gestion de la relation client, les quelques géants américains qui dominent le secteur ouvrent des brèches dans les situations les mieux établies, menaçant de transformer en simples sous-traitants ceux qui naguère s'enorgueillissaient de leurs succès techniques. Est-il déjà trop tard ? Certes non, mais il est urgent de prendre la mesure des enjeux, économiques mais aussi démocratiques et, encore plus, d'agir de façon offensive, aucune autre stratégie ne pouvant nous protéger du flot montant des données massives.

*L'Association des Amis de l'École de Paris du management organise des débats et en diffuse des comptes rendus ; les idées restant de la seule responsabilité de leurs auteurs.  
Elle peut également diffuser les commentaires que suscitent ces documents.*

## EXPOSÉ de François BOURDONCLE

Je suis ingénieur au Corps des mines, docteur en informatique, et une partie de ma carrière s'est déroulée, en France et aux États-Unis, dans la recherche et l'industrie. En 2000, j'ai créé la société Exalead, qui réalise des moteurs de recherche et qui, en 2010, a été rachetée par Dassault Systèmes. Aujourd'hui, j'en suis le directeur technique et j'incarne ainsi, en quelque sorte, la partie internet de ce grand Groupe.

### Le plan Big Data

Au printemps 2013, on m'a demandé de faire partie de la commission Innovation 2030, présidée par Anne Lauvergeon. Le but de cette commission est d'identifier les pistes pour lesquelles la France dispose d'atouts solides et où elle est capable de générer des leaders mondiaux, présents à l'export et créateurs d'emplois dans notre pays. Sept ambitions ont été identifiées, dans les ressources marines, l'économie du vieillissement ou le stockage d'énergie, par exemple, et l'une d'entre elles concerne le Big Data. À l'examen, cette piste est apparue non seulement comme une filière verticale, mais aussi comme un axe transverse, irriguant un grand nombre de secteurs.

Un concours mondial d'innovation a été lancé à la suite de ce premier travail et, sur les six cent trente réponses reçues à ce jour, plus du tiers concerne le Big Data, ce qui atteste de la surreprésentation de ce sujet parmi toutes ces nouvelles problématiques. J'ai ensuite été nommé co-chef de file du plan Big Data avec Paul Hermelin, le PDG de Capgemini. D'une certaine façon, ce dernier représente les gens du service et des grands groupes alors que je représente plutôt les gens du logiciel et des start-up, ces deux mondes très différents cohabitant pourtant dans l'univers du Big Data.

Ce que le ministère du Redressement productif attend de nous, c'est d'abord d'identifier les verrous à lever pour qu'une industrie du Big Data puisse se développer en France, puis de déterminer les secteurs dans lesquels l'État pourrait être prescripteur ou, éventuellement, prendre des initiatives. Pour cela, nous avons entrepris un travail essentiellement de terrain, en rencontrant les acteurs impliqués afin de comprendre leurs problèmes et en nous efforçant de repérer, industrie par industrie, les secteurs concernés. Notre approche est donc tirée par l'aval et vise à faire émerger le marché. Nous partons du principe qu'il ne sert à rien de pousser l'innovation, comme on l'a systématiquement fait depuis trente ans, en finançant les entreprises innovantes mais en les laissant ensuite se débrouiller seules pour trouver leur marché. Nous avons en conséquence privilégié une approche business en délaissant les aspects de recherche pure ou ceux concernant les technologies.

Dans le domaine d'action de l'État, quelques secteurs nous sont alors apparus comme porteurs d'enjeux colossaux pour nos sociétés : il en va ainsi de la santé, de l'emploi ou de la lutte contre la fraude. Le défi est de faire rentrer les organismes qui gèrent ces domaines dans l'ère de la donnée et dans une logique d'innovation ouverte. Nous travaillons donc désormais, sur des aspects très opérationnels, avec le patron de Pôle emploi, celui de la Caisse nationale d'assurance maladie (CNAM), etc., afin de définir les priorités et les expérimentations à mener. Les informations actuellement à la disposition des conseillers de Pôle emploi étant notoirement insuffisantes, quantité de choses peuvent y être réalisées afin de réconcilier offres et demandes.

Notre système de santé, quant à lui, est confronté au défi majeur de devoir passer progressivement d'une logique curative à une logique préventive. Dans cette perspective, tout ce qui concerne le *Quantified Self*, c'est-à-dire les objets connectés – bracelets, capteurs ou autres – qui mesurent en temps réel vos données personnelles, est très prometteur car ils permettent, par exemple, de "monitorer" les personnes à risques afin d'éviter qu'elles ne tombent malades. Dans les assurances, les enjeux sont également considérables et, à la demande du cabinet du ministère du Travail, nous nous plaçons désormais dans une perspective de création d'un écosystème d'innovation autour de ces métiers.

Dans le domaine des entreprises privées, notre démarche est quelque peu différente mais nous cherchons, là aussi, à identifier les principaux verrous et à déterminer les secteurs potentiellement les plus impactés.

### **Qu'est-ce que le Big Data ?**

Selon le modèle défini par la société d'analystes américains Gartner, parler des Big Data implique de prendre en compte les données selon la règle des trois V : quand leur *Volume* est important, quand elles arrivent avec *Vélocité* et qu'elles sont *Variées*. Cela définit le *Quoi*, mais pour le *Comment* et le *Pourquoi*, les choses sont plus complexes.

Si l'on prend du recul face à la multiplicité de domaines relevant du Big Data, cela ressemble à une histoire vieille comme le monde : comment les modernes peuvent-ils prendre la place des anciens ? La seule innovation est désormais la nature du levier utilisé : la donnée.

Si vous considérez les géants du numérique que sont IBM, Google, Oracle, Facebook, Amazon, etc., vous constaterez que ces entreprises, toutes américaines, entretiennent, en particulier dans le *B to C (Business to Consumer)*, une relation commerciale avec plusieurs centaines de millions de clients, sur lesquels ils détiennent des masses de données précieuses. Pour l'instant, ils ne les utilisent que pour faire du commerce, proposer différents types de services, faire de la publicité, etc. Mais, progressivement, leur appétit étant sans bornes et leurs moyens illimités, ils captent une part croissante de la valeur créée dans des secteurs économiques toujours plus nombreux.

Ainsi, Google rachète, pour trois milliards et demi de dollars, Nest, un fabricant de thermostats qui ne faisait pourtant que deux cents millions de dollars de chiffre d'affaires. On peut se demander pourquoi : s'intéressent-ils à la domotique ou aux équipes de design de Nest venues d'Apple ? Ou n'est-ce pas plus simplement parce que cela va lui permettre d'acquérir des données supplémentaires sur les usagers lui permettant ensuite de développer des offres de nouveaux services ? Google investit également dans des sociétés de taxis et, il y a peu, a racheté, de façon qui pourrait paraître surprenante, une société d'assurances.

Bien que les dirigeants de Google aient aussitôt annoncé que c'était une erreur stratégique, dans le même temps, ils ont pris contact avec des assureurs français. Tout en les assurant qu'ils n'étaient pas intéressés par un rachat, vue la complexité du métier d'assureur et de sa réglementation, ils leur ont cependant suggéré que leurs algorithmes pourraient améliorer la rentabilité de leur chaîne de création de valeur en y identifiant les zones d'inefficacité. Mais, avant tout, ils se sont présentés comme intéressés par la seule commercialisation des produits d'assurance, activité généralement déléguée à des courtiers par les assureurs car n'étant pas au cœur de leur métier.

### **L'attaque des barbares**

Imaginons maintenant une suite à cette histoire, qui n'est pas encore mise en œuvre mais qui pourrait bientôt l'être. Un jour, sur votre téléphone Android, Google fait apparaître une offre d'assurance 30 % moins chère que la vôtre. Intéressé, vous y répondez. Apparaît alors un bouton vous demandant de valider des conditions générales d'utilisation qui vous imposent d'accepter que votre position GPS et les données de votre accéléromètre soient captées en permanence. Sans vous y attarder, vous acceptez. Dès lors que l'opérateur a accès, en permanence, à ces données sur votre personne, il sait si vous êtes oisif ou pas, si vous faites du jogging et combien de kilomètres par jour, si vous conduisez bien ou mal, en ville ou à la campagne, tous les jours ou occasionnellement, en semaine ou le week-end, etc. Il peut donc établir votre profil de risques, réaliser un *pricing* sophistiqué parfaitement adapté à ce profil et vous établir un contrat sur mesure.

Quand vous savez faire un tel contrat d'assurances, vous pouvez proposer des produits à très bon marché aux clients à faibles risques. Le corollaire est que ces bons clients vont venir chez vous en masse alors que, dans le même temps, vous en dissuaderez les mauvais clients, qui ne

trouveront chez vous que des produits sur mesure, eux aussi, donc hors de prix. Ils resteront alors chez leurs assureurs traditionnels, qui eux, n'accèdent pas aux mêmes données que vous. Ces derniers subiront alors une double peine : ne pouvant assurer que des mauvais clients, leur taux de sinistralité montera en flèche et, comme ils seront face à un concurrent pratiquant des tarifs plus bas, ils seront obligés de baisser leurs prix ! La pression sur leurs marges se fera donc par le haut et par le bas.

Tout cela n'est pas nouveau et s'est déjà produit chez les opérateurs des télécommunications, dans la presse, la musique, etc. Néanmoins ce scénario reflète ce qui pourrait bientôt se passer dans ce secteur-là. Mais nous ne sommes qu'au début d'une tragédie en trois actes qui pourrait s'intituler *L'attaque des barbares*.

Durant le deuxième acte, Google va se retourner vers les assureurs en leur demandant de prendre en charge le *back office* de ce produit nouveau, créé grâce à sa maîtrise des données mais qu'il ne sait pas, ou ne veut pas, gérer techniquement. Les plus gros ne s'abaisseront sans doute pas à devenir les sous-traitants de Google, mais un plus petit, mis en difficulté par cette concurrence sur les produits, se précipitera sur ce qui lui semblera être une aubaine. Cet assureur-là, déjà en survie, sera alors traité comme les équipementiers automobiles le sont par leurs donneurs d'ouvrage : rapidement pris à la gorge, il ne pourra refuser de céder ses données sur les sinistres, sa dernière ressource, précipitant ainsi sa chute.

Au troisième acte, Google, déjà en possession des données sur les personnes et désormais de celles sur la sinistralité des contrats, croisera les deux et fera ce qu'il a déjà fait pour la traduction automatique. *Google Translate* est en effet la résultante d'un processus d'apprentissage machine sur des corpus dits alignés, l'équivalent de ce qu'a été la pierre de Rosette pour Champollion. Ces corpus ont été élaborés par des linguistes au fil des ans, tel mot d'une langue correspondant à tel mot d'une autre langue, telle phrase à telle autre. La machine a ensuite appris ces correspondances et, en apprenant, elle a généralisé tout comme Champollion l'a fait à partir des hiéroglyphes présents sur la pierre de Rosette. En appliquant ce genre d'algorithme à des corpus croisés entre des données personnelles sur le nombre de kilomètres que vous parcourez journalièrement, les maladies dont vous avez souffert, etc., Google affinera de plus en plus sa gestion des sinistres et pèsera toujours plus sur la marge de ses concurrents en s'attaquant, cette fois-ci, aux opérations. Et là, ce sera la triple peine pour les assureurs conventionnels !

### **De la technicité à la gestion de la technicité**

Ce scénario peut se dupliquer sur les avions, les voitures et sur une foule d'autres secteurs industriels qui vont très vite se trouver dans le cœur de la tourmente. La grille de lecture sous-jacente à ces situations est la même partout : il s'agit d'appliquer un effet de levier sur un métier traditionnel pour le réduire à sa technicité pure et déplacer la production de valeur dans la relation client. Ce levier, c'est la capacité qu'ont certaines firmes d'accéder à des données que les autres n'ont pas et de savoir les croiser. Là réside la spécificité des Big Data. Les Big Data ne sont qu'une modalité, pas une fin en soi, la vraie finalité étant, pour ces nouveaux opérateurs, de capter des clients dans des domaines qui, historiquement, ne sont pas les leurs.

Dans le domaine de la gestion d'énergie, le cas de l'île de Malte est particulièrement intéressant : comme cette île doit dessaler l'eau de mer pour faire face à l'afflux de touristes en été, elle est confrontée à la difficulté de fournir simultanément de l'électricité et de l'eau douce. Ne pouvant construire de centrale thermique pour des raisons environnementales, les autorités ont lancé un appel d'offres international qu'IBM a remporté, GDF Suez étant son sous-traitant. GDF Suez assure se tailler la part du lion dans ce contrat, IBM y perdant, selon eux, plus qu'il n'y gagne. Soit. Il me semble cependant que, la gestion de l'énergie tout comme celle de l'eau étant appelées à monter en puissance, la problématique sera à l'avenir davantage centrée sur l'intégration de systèmes que sur la technicité des réseaux de distribution. Là encore, la valeur va se déplacer de la technicité pure à la gestion de cette technicité et à la relation client. J'ai donc tendance à penser que l'avenir appartient aux gens,

comme IBM ou Thales, qui font de l'intégration de systèmes et mènent la guerre, plutôt qu'aux opérateurs historiques de la distribution des eaux qui ne se sentent pas vraiment concernés par cette guerre.

Un autre exemple très intéressant est celui des moteurs d'avions. Pratt and Whitney considère désormais qu'IBM a confisqué leur relation client. En effet, un modèle qui tend à s'imposer progressivement consiste à "forfaitiser" la maintenance des moteurs et des avions (Airbus le fait sur ses A380 pour tester l'approche avec son programme *Flight Hour Services*). En faisant cela, ces deux industries manufacturières se réinventent sans le savoir en industries de services, et pour pouvoir apprécier le coût des services de maintenance qu'ils commercialisent, il est essentiel d'amasser autant de données que possible sur les coûts de maintenance des moteurs et des avions, sans quoi le risque industriel, mal évalué, peut conduire à vendre à perte. C'est très exactement le métier des actuaires dans les assurances. D'ailleurs, pour les hélicoptères, le modèle a été poussé encore plus loin car les armées ont indiqué aux constructeurs leur intention d'acheter des heures de vol et non des appareils. Cette expérience a paraît-il été très douloureuse pour les constructeurs qui, ayant mal apprécié le coût de ce service, y ont laissé quelques plumes. Tout l'enjeu dans ce domaine (comme pour l'évolution du logiciel vers le Cloud), une fois "internalisée" la maintenance par le constructeur, est donc d'arriver à dégager un maximum de marge, et le moyen d'y arriver est d'analyser un maximum de données, comme par exemple les 20 téra-octets de données que produit chaque heure un moteur d'avion. Tout ceci explique que les compagnies aériennes ont très naturellement commencé à davantage parler à IBM, qui traitait ces données, plutôt qu'à Pratt and Whitney, relégué à sa seule technicité d'industriel. On pourrait même alors imaginer qu'à terme, ce soit désormais IBM qui loue des moteurs aux compagnies aériennes, en négociant directement avec les constructeurs qui eux, seront ainsi relégués au second plan. Il est à ce titre significatif que General Electric, qui est le partenaire de Snecma sur le programme CFM56, investisse 1 milliard de dollars pour créer un centre de R&D spécialisé sur le Big Data et 105 millions de dollars chez l'éditeur logiciel Pivotal. Tout aussi intéressant, General Electric, affiche un chiffre d'affaire de 3 milliards de dollars dans le logiciel pur (comparable à celui de Dassault Systèmes) et devrait monter à près de 5 milliards de dollars d'ici deux à trois ans.

L'analyse de la relation client chez les opérateurs des télécommunications est également riche d'enseignements. Si vous êtes un usager exigeant et un gros consommateur, vous allez d'abord choisir votre terminal, un iPhone par exemple, et ensuite choisir l'opérateur qui vous proposera les meilleures conditions pour l'obtenir. La relation client pour ce type d'usager appartient donc à Apple, qui dispose ainsi d'un observatoire des usages unique, condition indispensable pour proposer des produits innovants. Si vous êtes en revanche un utilisateur basique, pour lequel un terminal en vaut un autre et pour qui changer d'opérateur paraît être un casse-tête, vous resterez fidèle au vôtre pour qui vous serez cependant un mauvais client. Le cas de figure est identique à celui des assureurs : les opérateurs traditionnels gardent les mauvais clients, les bons étant captés par les modernes qui, eux, maîtrisent la relation client. Ces pratiques n'ont rien de nouveau mais leur côté systématique est quelque peu angoissant.

De leur côté, il semblerait que les constructeurs automobiles allemands envisagent de se réserver la vente de leurs berlines haut de gamme, et choisissent à l'avenir de louer celles du moyen de gamme sur le modèle de l'Autolib, afin de garder la relation client directe. Les constructeurs français, positionnés sur le moyen de gamme, voient cette évolution d'un très mauvais œil, mais cela suffira-t-il pour que cela ne se produise pas ? Le risque pour eux serait alors qu'ils se laissent transformer en sous-traitants par un spécialiste de la relation client qui opérerait ce service à leur place, Hertz par exemple.

L'assurance est, d'ores et déjà, un secteur très convoité et ses dirigeants, fort inquiets, sont très mobilisés face à la menace. Si les exemples ne manquent pas dans de nombreux autres domaines, la grille de lecture reste toujours la même : il s'agit d'une guerre entre anciens et modernes, ces derniers mettant en coupe réglée tous les segments de l'industrie et des services qui leur ont échappé jusqu'à présent en faisant effet de levier sur leur relation client. Cela est vrai, non seulement sur les secteurs grand public, mais aussi sur le *B to B*.

## Big Data et respect de la vie privée

Dans le cadre de notre mission, nous travaillons sur la loi “Informatique et libertés”. Cette loi a été conçue à une époque où détenir des données personnelles était quelque chose d’exceptionnel, limité à des activités très particulières, alors qu’aujourd’hui nous parlons, potentiellement, de l’ensemble de l’industrie et des services. Cette loi, telle qu’elle est rédigée, avec comme principe le respect de la finalité de la collecte initiale dans tout usage ultérieur, est donc exactement orthogonale à la définition du Big Data, qui réutilise ces données et les croise pour faire émerger de nouveaux services. Pour rester efficace, la loi doit donc impérativement évoluer car on ne peut continuellement s’interdire d’utiliser le levier que les autres utilisent pour soulever le monde, sauf à courir le risque d’être très vite mis hors-jeu.

Pour autant, la protection de la vie privée est un enjeu essentiel et je suis convaincu que nous pouvons en faire un label et un atout à l’exportation en proposant des solutions Big Data respectueuses de la vie privée. Nous avons donc convaincu le ministre du Redressement productif de remettre la loi en chantier. Une commission consultative a été créée à cet effet ainsi que des ateliers de travail qui réunissent la Commission nationale de l’informatique et des libertés (CNIL) et des industriels afin de réfléchir à des pistes d’évolution.

La première de ces pistes consiste à remplacer le principe de finalité par un principe de réciprocité : si on collecte mes données, je dois en retirer un bénéfice économique, sous forme de services par exemple. Si on les réutilise à d’autres fins, je dois en être informé et en retirer, là aussi, une contrepartie. Pour contrer le caractère déclaratif de la loi, irréaliste compte tenu de l’échelle du phénomène, toutes les entreprises ne pouvant déclarer tout ce qu’elles font en matière de collecte de données, une autre piste serait l’adhésion volontaire à une politique de labellisation. Ainsi, en matière d’assurances automobiles, si une entreprise souhaitait à l’avenir faire du *Pay As You Drive* et, pour cela, envisageait de recueillir des données privées sur vous, comme votre biorythme, et de les croiser avec votre géolocalisation, elle serait alors invitée à en anonymiser certaines clairement identifiées en utilisant tel algorithme et tel *process* industriels certifiés. Cela lui conférerait, en contrepartie de cette contrainte, un label qualité. La réflexion commence sur ces sujets et je ne pense pas que l’on puisse faire l’économie d’une évolution à court terme du cadre réglementaire et législatif.

Nous réfléchissons également au thème de la formation des *data scientists*, c’est-à-dire au cas de ces gens qui sont à la fois informaticiens, statisticiens et ont une compétence métier particulière. Nous avons identifié deux aspects différents sur ce point. Tout d’abord, il y a un important aspect de formation continue, les entreprises ayant besoin de faire très vite monter en compétence certains de leurs salariés sur ces thématiques.

Il existe ensuite un certain nombre d’initiatives en formation initiale, comme les masters Big Data de ParisTech ou de l’École nationale de la statistique et de l’administration économique (ENSAE) ou la chaire Big Data de l’université Paris Dauphine et le DUT de Grenoble. Nous avons besoin de tous les niveaux de compétences, depuis le technicien jusqu’à l’ingénieur haut de gamme. Si d’aucuns pensent que ces nouveaux profils de *data scientists* requièrent d’avoir trois cerveaux, il nous semble qu’il s’agit davantage de compléter des formations techniques classiques par une formation culturelle permettant de comprendre les autres logiques mises en œuvre.

Ainsi, un statisticien pourra ne pas comprendre d’emblée que, si on lui demande de faire sur un milliard de personnes l’analyse qu’il a l’habitude de faire sur cent mille, son algorithme pourra ne plus être adapté. Il est donc important de familiariser à minima chacun des spécialistes à la culture des autres afin de constituer des équipes pluridisciplinaires sur ces sujets. Voici un autre exemple : depuis quelques années, on voit apparaître un nouveau domaine de l’apprentissage machine, le *deep learning*, qui commence à battre les humains sur des domaines très complexes. Dans le monde du Big Data, cette approche semble largement s’imposer. Il paraît cependant important que ce domaine, qui relève principalement des mathématiques, ne devienne pas le pré carré des informaticiens et que de la science y soit adjointe.

Toute une réflexion doit donc être menée sur la façon dont on doit produire, en masse et sur le long terme, ce type de compétences dans ces divers métiers. La France, qui possède des formations d'excellence dans ces divers domaines, est plutôt bien placée dans cette course.

Avec le plan Big Data, nous ne sommes pas engagés dans un plan industriel classique, dans la mesure où il ne s'agit pas de construire un objet ou un secteur industriel. Nous sommes plutôt dans la perspective de faire émerger un écosystème, en identifiant des blocages, en parcourant le terrain pour sensibiliser les acteurs et en expliquant ces enjeux. Il consiste à faire passer le message qu'il s'agit d'abord d'action et que c'est de là que naîtra la réflexion. Nous ne nous plaçons pas dans une perspective défensive mais bien dans une démarche offensive. C'est de la confrontation au réel, aux usages, aux clients que naîtront ces nouveaux métiers, dans le prolongement des métiers historiques, pour contrer cette offensive des modernes.

Comment, dans les assurances par exemple, résister à l'érosion des marges annoncée ? L'unique solution sera pour les assureurs classiques de justifier des prix plus élevés par une meilleure qualité du service rendu aux clients : il leur faudra alors apprendre à obtenir ces informations personnelles qui leur permettront à leur tour d'être plus compétitifs. Comme il semble improbable qu'ils inventent ces nouvelles pratiques en interne, cela passera peut-être par des accords avec La Poste ou avec telle ou telle autre entreprise dont le métier est de recueillir ces informations. Il n'a jamais été aussi important, pour les grands groupes français, de travailler dans une logique d'innovation ouverte, vers des métiers de services qui, traditionnellement, ne sont pas les leurs.

Mais cela implique de mettre en œuvre, de façon urgente, une vraie logique d'accompagnement du changement.

# DÉBAT

## L'Europe et le Big Data

**Un intervenant :** *Je ne vous ai pas entendu parler de l'Europe, or il me semble que rien ne pourra se faire dans ce domaine en dehors d'elle...*

**François Bourdoncle :** La dimension européenne est en effet critique. Nous avons déjà réussi à étendre au niveau européen la gestion des données telle que la conçoit la CNIL. Mais les barbares sont à nos frontières et agissent chaque jour, en achetant des entreprises, en investissant, etc. Il nous faut donc réagir très vite et si nous essayons d'abord de gravir la montagne européenne, nous serons très vite débordés.

Ceci étant, quand je parle d'innovation ouverte, je ne limite pas cette démarche aux entreprises françaises et des alliances européennes sont souhaitables. Mais, historiquement, les grands groupes français n'ont guère l'habitude de travailler avec des start-up : ils conçoivent mal de payer une première fois pour les aider à se développer et une seconde pour les racheter quand elles ont réussi. Il me semble pourtant qu'ils le devraient : il est en effet très difficile d'intégrer une start-up dans un grand groupe sans qu'elle s'y dissolve instantanément si elle n'a pas atteint une certaine surface. Il faut donc laisser ces petites entreprises grandir et accepter qu'elles coûtent plus cher quand elles seront achetées car alors, ces groupes n'achèteront pas seulement une technicité innovante mais aussi une relation client, un réseau à l'export, etc. Il faut que ce message soit compris et que les mentalités évoluent en ce sens, en raisonnant évidemment au niveau européen.

**Int. :** *General Electric a investi un milliard de dollars dans le Big Data de façon à maîtriser ce que n'a pas réussi à faire Pratt and Whitney. N'est-ce pas l'exemple à suivre pour des grandes entreprises européennes comme Airbus afin de s'assurer du contrôle de ces données ?*

**F. B. :** Sans doute, mais nous peinons, dans notre pays, à faire émerger de nouveaux géants, 100 % modernes. Faisons donc en sorte que nos anciens deviennent modernes ! Rien n'est perdu d'avance. Il faut certes faire feu de tout bois pour assurer notre défense mais il faut aussi que chaque secteur industriel ait sa propre approche pour faire sauter les verrous. Il s'agit évidemment d'investir dans tout ce qui relève des Big Data, du *machine learning*, etc. Aujourd'hui, même Goldman Sachs peine à recruter des talents pour faire du Big Data, ceux-ci préférant désormais aller vers les modernes : c'est bien le signe que le monde change et qu'il faut réinventer ses métiers en allant vers la relation client et en s'appuyant pour cela sur les données. Les assureurs en sont parfaitement conscients et sont très inquiets de ces évolutions.

**Int. :** *Le Big Data se développe dans la logique d'un modèle américain sans nuances. Ne devrions-nous pas alors faire jouer notre réglementation sur l'utilisation des données relevant de la vie privée et notre culture de la nuance pour structurer une approche européenne dans laquelle le respect de la vie privée serait un outil compétitif en durcissant les normes ?*

**F. B. :** La loi Informatique et libertés, en son état actuel, ne permet pas l'activité industrielle dont nous parlons. Il faut donc rééquilibrer le principe de précaution, trop orienté vers le consommateur et trop peu vers l'industrie. Durcir ce qui existe aujourd'hui, et qui ne marche pas, n'aurait donc pas de sens : ce qu'il faut, c'est changer de référentiel. C'est ce à quoi travaille la commission Lauvergeon. Ce n'est pas tant d'argent dont nous avons besoin mais d'une réflexion intelligente sur un dispositif qui pourrait être à la fois protecteur en interne et un atout à l'export. Il faut surtout que les entreprises passent à l'acte, commencent à expérimenter, travaillent avec des start-up pour inventer de nouveaux métiers dont elles peuvent penser qu'ils seront, en gérant les données, dans la connexité de leur propre métier et qu'ils modifieront ainsi leur relation client, leur permettront de faire des produits à la demande, etc.

**Int. :** *Y a-t-il une prise de conscience, dans les écoles de commerce par exemple, de la nécessité de monter des formations orientées vers les business models liés aux Big Data ?*

**F. B. :** Je viens de visiter “42”, l’école d’informatique gratuite sans condition de diplôme créée par Xavier Niel dont l’approche est très intéressante. Concernant les écoles de commerce, j’ai récemment rencontré la responsable des études de l’ESCP Europe et le président souhaite désormais former l’ensemble des étudiants, quel que soit le cursus, à la programmation. On sent que les choses évoluent et rejoignent le discours de 42 qui pense qu’il y a un type d’informatique, en particulier sur le web, que l’on peut faire sans nécessairement maîtriser les mathématiques. Sensibiliser l’ensemble des cursus commerciaux aux problématiques du numérique en général et à la programmation me paraît désormais essentiel.

### **Problème de technique ou défi stratégique ?**

**Int. :** *Quel rôle les directions de services informatiques (DSI) peuvent-ils jouer dans la transformation des entreprises ?*

**F. B. :** Le problème n’est plus un problème de productivité des entreprises, comme ce fut le cas jusqu’au début des années 2000. Depuis, l’innovation numérique s’est déplacée vers le grand public, en particulier avec la révolution des réseaux sociaux. Pour moi, la troisième révolution, celle des Big Data, partie du *B to C*, va essentiellement toucher le *B to B* et impacter les *business models*, les produits, la relation client, etc., de beaucoup d’entreprises traditionnelles. C’est donc un problème qui concerne avant tout les directions générales, les DSI ayant ensuite la charge de faire évoluer leur métier dans cette perspective.

**Int. :** *Le Big Data me semble beaucoup plus structurant que vous ne le laissez entendre et ne se limite pas à ses aspects marketing. Le Big Data est aussi ce qui a permis au coût du séquençage du génome humain de passer de plusieurs milliards d’euros, en 1992, à quelques centaines, aujourd’hui, et quelques dizaines, demain.*

**F. B. :** Une attaque en coupe réglée comme celle sur les *business models*, les produits et les relations clients des assureurs est, selon moi, bien plus structurante qu’une simple approche technique. Face à une telle offensive, il faut gérer les aspects sociaux, former des salariés, en licencier d’autres, etc.

J’ai ensuite volontairement éliminé les aspects technologiques de mon discours sur le Big Data. Y a-t-il trois V, comme l’a dit Gartner ou quatre V, cinq V ? De telles considérations me paraissent stériles. Quant à la génomique, vous avez parfaitement raison : nous réfléchissons aux implications dans le domaine de la Santé, avec la CNAM et d’autres acteurs, et c’est l’un des sept axes prioritaires de la commission Lauvergeon.

Mais tous les secteurs économiques sont concernés. Ainsi, dans le domaine de l’hôtellerie, les réservations en ligne en France ne concernent que 20 % du total, mais 90 % d’entre elles passent par les trois principales plateformes qui régissent leurs prestations par des clauses léonines. Ainsi, si un hôtelier accorde une ristourne au comptoir à l’un de ses clients, il est contractuellement tenu d’en informer Expedia, qui se réserve le droit de la répercuter sur le prix de toutes les nuitées qu’elle commercialise pour cet hôtel. Même une société comme Accor a du mal à négocier avec les plateformes de réservation en ligne, car elle ne peut rivaliser avec elles en termes de pouvoir de négociation.

Perdre sa relation client, c’est très concrètement perdre son autonomie en termes de politique commerciale : on ne maîtrise plus ni sa marge, ni sa politique tarifaire. C’est pour cela que ces aspects me paraissent plus que structurants : si l’ensemble de nos industriels est transformé en sous-traitants, quel qu’en soit le rythme, cela pourrait s’avérer gravissime pour nos économies. La pression que feraient peser ces pratiques sur les marges de nos industries et de nos services aurait des répercussions sévères. On pourra toujours gagner un peu de temps en attaquant les responsables pour pratiques abusives, mais d’être sur la défensive ne suffira pas à brève échéance. Il est donc urgent de passer à l’offensive.

**Int. :** *Comme la politique du Centre national d'études spatiales (CNES) est que les données publiques doivent être gratuites, Google s'est rapproché de l'Agence spatiale européenne (ESA), l'été dernier, en lui proposant de traiter ces données. Comment alors ne pas être défensif ?*

**F. B. :** Cela me semble une énorme erreur. On donne à l'adversaire l'arme avec laquelle il va nous battre !

**Int. :** *Que faut-il alors penser du couplage possible entre Big Data et Open Data ?*

**F. B. :** A priori, cela n'a rien à voir. Mais cela n'empêche pas que l'on puisse prendre des données ouvertes pour ajouter la géolocalisation, par exemple, à une application Big Data. Mais ce n'est rien de plus que l'un des inputs d'un Big Data. Faut-il promouvoir l'Open Data dans une logique citoyenne ? Évidemment oui, car c'est l'une des garanties à long terme que la démocratie ne dérapera pas. Est-ce qu'il faut, d'un point de vue purement économique, intégrer des données ouvertes ? Certaines, oui, mais d'autres, non.

Quand l'Institut national de l'information géographique et forestière (IGN) a vendu, pour quelques millions d'euros, son fond de cartes à Google, celui-ci l'a utilisé pour démarrer son processus de cartographie qu'il gère désormais très bien tout seul. Nous avons juste ainsi obéré nos chances de pouvoir reprendre la main sur ce sujet. Il faut savoir que, depuis huit ans, Google Map est une usine logicielle de cinq mille ouvriers du XXI<sup>e</sup> siècle, cols bleus indiens qui, devant leur poste de travail ont, d'un côté les photos prises par les véhicules de Google Street Map et, de l'autre, les panneaux Google et qui vérifient la concordance des deux. C'est un travail long et fastidieux, mais une fois que l'investissement initial a été fait, c'est irrattrapable par les concurrents. Construire une carte à partir de données topographiques brutes n'est donc pas du tout la même chose que de les introduire, alors qu'elles sont déjà élaborées par autrui et qu'elles ont été obtenues gratuitement, dans ses propres processus afin de renforcer sa position dominante. Il ne faudrait pas réitérer cette erreur en donnant les données de la CNAM, par exemple, à Google.

### **Démocratie et principe de réalité**

**Int. :** *Y a-t-il un espoir d'équilibre entre le respect de la vie privée et les préoccupations commerciales ?*

**F. B. :** Nous sommes pris entre trois principes : le principe de précaution, celui d'innovation et celui de réalité. Dans le cadre de la mission qui m'a été confiée, ce dernier principe m'interdit à l'évidence de penser à abolir la loi Informatique et libertés ! Notre approche actuelle recueille cependant un certain consensus pour commencer à réfléchir à ce sujet. Je suis convaincu que l'économie numérique ne se développera pas sans une économie de la confiance. Sur ce point, nous sommes plutôt mieux partis que les Américains.

Notre récente loi de programmation militaire autorise cependant, sous couvert de prendre en compte un certain nombre de dispositifs techniques nouveaux, des pratiques comme l'interception du contenu de courriers électroniques en supprimant, dans le même temps, l'autorisation du juge pour y substituer une simple requête administrative, privilège désormais également accordé à l'administration fiscale. La contradiction sur ce point avec la loi Informatique et libertés est flagrante et quelque peu antinomique à la confiance nécessaire au succès d'un Big Data à la française !

**Int. :** *Trop peu de décideurs politiques sont sensibilisés aux enjeux du numérique ce qui explique que certains aspects litigieux de cette loi aient pu être votés.*

**Int. :** *Assurer la confidentialité de mes données suffira-t-il alors que l'on peut prédire mon comportement à partir de celui de mes amis ? Peut-on imaginer que les acteurs individuels puissent vendre leurs propres données ?*

**F. B. :** Une ligne Maginot ne suffira pas à nous protéger. C'est pour cela que je pense qu'il faut imposer aux industriels une réciprocité. Un modèle proche commence à émerger, consistant à faire payer ceux qui souhaiteraient garder l'anonymat et dont le pouvoir d'achat

le leur permettrait, la gratuité étant la règle pour les autres. D'un point de vue démocratique, ce n'est guère satisfaisant mais cela se tient du point de vue économique.

**Int. :** *Dans les grands magasins, le comportement des clients commence déjà à être analysé par le biais de leur smartphone et s'ils acceptent de donner leur identité, ils bénéficient d'avantages divers. La réciprocité y est donc déjà mise en œuvre et les gens le comprennent facilement.*

**Int. :** *Si ce que vous nous annoncez se réalise, cela veut dire qu'à terme, 80 % des gens n'auront plus droit aux assurances, au système de santé, etc. Comment la société pourrait-elle résister à un tel choc ? Et comment lutter contre un tel mouvement déjà largement entamé qui réduit le système mondial à une poignée d'opérateurs économiques géants ?*

**F. B. :** L'enjeu démocratique est gigantesque. Si nous ne nous occupons pas de notre futur, c'est lui qui s'occupera de nous et ce sera fait avec des standards qui ne sont pas les nôtres ! Se greffent sur cela des débats sur le transhumanisme, sujet tabou chez nous mais qui ne l'est ni aux États-Unis, ni en Chine. On ne peut donc pas simplement se voiler la face et faire semblant d'ignorer que les choses avancent. Pour ma part, je pense simplement que, puisque nous ne savons pas faire émerger des modernes, il nous faut faire en sorte que nos anciens deviennent modernes.

Le Big Data est une arme de guerre. C'est le levier avec lequel le capitalisme est en train de soulever le monde. Alors, premièrement, sachons-le, et, deuxièmement, arrêtons de nous congratuler dans les conseils d'administration en nous disant que, jusque-là, tout va bien. Il faut que, sur ce thème, nos conseils d'administration mettent la pression sur les dirigeants. Il faut vite faire évoluer nos standards et, ensuite, agir tout aussi vite.

Présentation de l'orateur :

François Bourdoncle : est l'un des pionniers du marché du logiciel de recherche ; polytechnicien, ingénieur du Corps des mines ; il a été chercheur dans l'industrie ; professeur associé à l'École polytechnique et l'École normale supérieure, et il est cofondateur en 2000 de la société Exalead, leader européen des moteurs de recherche en entreprise, aujourd'hui filiale de Dassault Systèmes.

Diffusion avril 2014