

POLICY PAPER

Mai 2014

Faire entrer la France dans la troisième révolution industrielle : le pari de l'innovation

#1 *Le big data*

Observatoire de l'Innovation
de l'Institut de l'entreprise

Faire entrer la France dans la troisième révolution industrielle : le pari de l'innovation

Note thématique #1 :

Le big data

Observatoire de l'Innovation
de l'Institut de l'entreprise
présidé par Christophe de Maistre

Pilote : Delphine Manceau
Rapporteur : Julie Fabbri

Note rédigée à partir des travaux de l'observatoire.

SOMMAIRE

AVANT-PROPOS.....	5
SYNTHÈSE.....	7
<i>BIG DATA</i> , DE QUOI PARLE-T-ON ?	11
LES 4V DU <i>BIG DATA</i>	12
DE NOUVELLES GÉNÉRATIONS DE DONNÉES ET DE TRAITEMENT	13
<i>BIG DATA</i> , LEVIER D'INNOVATIONS	17
LES NOUVEAUX USAGES STIMULÉS PAR LE <i>BIG DATA</i>	17
SOURCE D'INNOVATIONS ORGANISATIONNELLES	27
<i>BIG DATA</i> , FACTEUR CLÉ DE PERFORMANCE	32

L'IMPACT DU <i>BIG DATA</i> SUR LES ENTREPRISES.....	35
LE <i>BIG DATA</i> COMME NOUVEAU MARCHÉ.....	37
LES PRINCIPAUX POINTS DE VIGILANCE.....	39
LES MUTATIONS À PRÉVOIR	46
PISTES DE RECOMMANDATION	56
L'OBSERVATOIRE DE L'INNOVATION	67

AVANT-PROPOS

Cette note est le deuxième item d'une série de trois études issues du travail de l'Observatoire de l'Innovation de l'Institut de l'entreprise, portant respectivement sur le *big data*, l'efficacité énergétique et les nouvelles technologies de production, publiées séparément et précédées d'une note introductive générale.

L'Observatoire de l'innovation est présidé par Christophe de Maistre, président-directeur général de Siemens France et ses travaux ont été supervisés par Delphine Manceau, professeur à ESCP Europe, assistée de Julie Fabbri, doctorante au Centre de Recherche et Gestion (CRG) de l'Ecole Polytechnique, rapporteur de l'Observatoire. Eudoxe Denis, Directeur des études de l'Institut de l'entreprise, a supervisé le cadrage du projet et de l'ensemble des séances de l'Observatoire ainsi que la rédaction de la note introductive.

L'Institut de l'entreprise tient à exprimer sa gratitude :

› À l'ensemble des membres de l'Observatoire de l'innovation pour leur engagement dans ce projet :

Frédéric Allard, Vice-président R&D, IBM

Grégory Blokkeel, Responsable stratégie & open innovation, PSA

Marie-Anne Broodschi, VP Innovation, Veolia

Eric Conti, Directeur Innovation & Recherche, SNCF (également auditionné par l'Observatoire)

Priscille Crozemarie, Chargée de mission - Secrétariat général, RATP

Julie Fabbri, Secrétaire Générale, Institut pour l'Innovation et la Compétitivité i7, ESCP Europe et doctorante au centre de recherche et Gestion (CRG) de l'Ecole Polytechnique

François Gerin, Directeur général adjoint, Siemens France

François Grosse, Senior Vice-President Digital Services, Veolia

Jakob Haesler, co-fondateur, Tynclues (également auditionné par l'Observatoire)

Frédéric Lefebvre-Naré, Consultant, Isée

Pascal Lemoine, Directeur technique et de la recherche, FNTP

Delphine Manceau, Professeur, ESCP Europe

Françoise Mercadal-Delasalles, Directrice des ressources et de l'innovation, Société Générale

Jean-Yves Moreau, Directeur des relations gouvernementales et parlementaires, Sanofi France

Michel Morvan, président et co-fondateur, The CoSMo Company (également auditionné par l'Observatoire)

Jean Richard de la Tour, Directeur des études, de l'innovation et des affaires publiques, Carrefour (également auditionné par l'Observatoire)

Martin Roulleaux-Dugage, Business transformation et Knowledge officer, Areva

Bernard Salha, Directeur recherche et développement, EDF

Jean-Christophe Saunière, Associé, PWC

Vincent Schramm, Directeur général, Symop (également auditionné par l'Observatoire)

Jérôme Weill, Associé, Oliver Wyman

› Aux personnes auditionnées lors des séances de l'Observatoire :

Philippe Berthier, Chef du département CRO, Innovation & Recherche, SNCF

André Bouffioux, président-directeur général, Siemens Belgique-Luxembourg

Eric Carreel, président fondateur de Withings, Sculpteo et Invoxia

Gilles Grapinet, Directeur général adjoint en charge des fonctions globales, Atos

Olivier Meunier, Directeur business development, Siemens Industry

Olivier Siri, VP Head of system design, Cassidian (Groupe EADS)

Claire Viguier-Petit, Directeur des opérations diabète, Sanofi France

› Aux personnes auditionnées par les rapporteurs en dehors des séances de l'Observatoire :

Alexandre Grux, Responsable de la recherche et de l'innovation, Kyriba

Philippe Vappereau, Chairman d'Ixxi, filiale de la RATP

L'Institut de l'entreprise remercie enfin, pour leur collaboration tout au long des travaux de l'Observatoire, **François Gerin**, Directeur général adjoint, Siemens France, **Jérôme Weill** et **Reda Fettah**, respectivement associé et analyste chez Oliver Wyman, ainsi que pour leur précieuse contribution **Frédéric Allard**, **Robin Rivaton**, **Lauriane Contamin** et **Elise Schobel**.

Les opinions exprimées dans le présent rapport n'engagent ni les personnes citées, ni les institutions qu'elles représentent.

SYNTHESE

Avec l'essor des réseaux Internet et Wi-Fi, des *smartphones*, des objets connectés et des réseaux sociaux, de plus en plus de données de formes variées sont générées. En parallèle, le développement d'outils de stockage et d'analyse – notamment le *cloud computing* – ainsi que de nouveaux outils de visualisation permettent la valorisation de ces données non structurées, variées et en très grande quantité : c'est un phénomène que l'on nomme le *big data*. Le volume croissant de données permet de s'abstraire des modèles prédéfinis et d'observer directement des corrélations entre données, qui peuvent ensuite être interprétées. De plus, de nouveaux outils permettent l'analyse simultanée de données nombreuses et variées, débouchant sur la microsegmentation des ensembles étudiés. La visualisation de tendances et de corrélations auparavant invisibles permet d'optimiser les prédictions, de détecter des variations et de prévoir l'impact d'une décision. Elle permet de tracer et cibler, via des outils de *profiling*, et par là d'améliorer des connaissances non plus statistiques mais individualisées. Cette microsegmentation permet aussi une gestion plus précise des risques et de la qualité.

À l'échelle de l'entreprise, les applications du *big data* sont nombreuses pour optimiser la logistique, le marketing via la personnalisation, le développement et le suivi de produits plus proches des besoins des clients, la gestion et la qualité, ou encore les ressources humaines : le *big data* est donc source d'innovations dans tous les secteurs d'activité, mais aussi de performance et de productivité. Le *big data* représente indéniablement un marché : au sens large, il inclut tous les outils permettant le traitement des données (capteurs, espaces de stockage) ; de façon plus restrictive, il comprend les acteurs du *cloud* et les entreprises qui analysent les données. De nouveaux services peuvent aussi voir le jour, comme la revente de données.

De plus, le *big data* peut accompagner de nouvelles démarches : l'*open innovation*, soit un mode d'innovation basé sur le partage et la coopération entre entreprises ; le *yield management* et l'automatisation de microdécisions ; de

nouvelles solutions pour le développement durable ; l'essor des *smart objects* et *smart services* ; ou encore l'utilisation de nouveaux outils pour la recherche.

Toutefois, le *big data* a beau être vecteur d'opportunités, il demeure avant tout un outil technique au service d'objectifs préalablement déterminés.

La valorisation des données est encore peu effective en France. Si de plus en plus de données sont stockées, elles ne sont pas toujours complètes, de qualité, accessibles facilement et rapidement, interopérables et exploitables par des outils informatiques. Peu de démarches sont mises en œuvre pour valoriser les données qui pourraient l'être. Cet état de fait peut s'expliquer par plusieurs limitations qui trouvent leur source dans les entreprises elles-mêmes : des difficultés de coordination entre les différents services de l'entreprise, impliquant souvent la coûteuse mise en place d'un département de *data scientists* ; la complexité de l'évaluation des bénéfices en terme de productivité et de croissance des stratégies liées au *big data*, face à des besoins d'investissements importants ; ou encore la mise à niveau des compétences en statistique, informatique et management pour nombre de cadres. Les craintes liées à la sécurité et à la confidentialité des données ainsi qu'à une exploitation abusive des données personnelles doivent aussi être prises en compte.

Certains de ces enjeux demandent l'implication des pouvoirs publics, comme les risques pour la vie privée, la formation, l'établissement de normes et de standards, et enfin la prévention contre l'éventuelle utilisation à mauvais escient d'analyses prédictives.

Si le soutien public à l'innovation par le *big data* est présent en France, par l'ouverture de données publiques et des aides financières aux *start-up* et PME, deux erreurs sont commises : celle de considérer le *big data* comme une filière propre, alors qu'au contraire il doit être développé dans tous les secteurs d'activité, et celle de ne concevoir ces politiques de soutien que selon une approche purement technologique, alors que ce sont les usages qui doivent servir de point de départ.

Pour lever les obstacles au développement du *big data*, l'accentuation de la politique d'*open data* publique est en premier lieu nécessaire. Ensuite, des mesures devraient être prises par les pouvoirs publics pour repenser l'encadrement des risques et la protection de la vie privée. De nouvelles expertises et institutions de régulation pourraient voir ainsi le jour : les démarches *big data*

seraient soumises à des audits d'algorithmes et le respect des données personnelles par les institutions ainsi que leur transparence à propos de leur sécurité pourraient être notés par des tiers certificateurs. Cette transparence accrue pourrait s'accompagner d'un droit à l'expérimentation, permettant de déroger aux règles actuelles en matière de protection des données personnelles, tandis que de nouveaux droits tels que le droit à l'oubli ou le droit au refus du profilage pourraient voir le jour. Face à l'explosion du nombre de données échangées et au partage par tout un chacun du même bien commun qu'est le web, des normes internationales deviennent de plus en plus nécessaires.

Du côté des entreprises, le développement du *big data* sera possible par une meilleure prise de conscience des gains offerts par celui-ci et de l'avantage compétitif apporté par les démarches de valorisation des données. La formation d'un département de *data scientists* dans les entreprises, département responsable du *big data* qui aura à identifier les démarches utiles d'ouverture ou d'exploitation interne des données et à quantifier les bénéfices en termes d'innovation, de productivité et d'image, en est une étape essentielle.

Condition de cet essor, la culture de la donnée devra être diffusée, aussi bien dans les entreprises que dans les administrations.

BIG DATA, DE QUOI PARLE-T-ON ?

Avec l'avènement des ERP, du CRM, du web, des réseaux sociaux, mais aussi la diffusion des capteurs, des *smartphones* et autres objets connectés¹, nous faisons face aujourd'hui à une explosion du nombre des données disponibles. Ces données prennent des formes variées, comme des mails, des photos, des vidéos, des fichiers, des commentaires sur les réseaux sociaux, des signaux GPS, des transactions bancaires... 90 % des données dans le monde ont été créées ces deux dernières années !

Cette surcharge d'informations ou « infobésité » est parfois difficile à gérer pour les organisations. Néanmoins, selon 75 % des responsables Études interrogés lors de la première édition du *Big Data Index EMC/IDC* (septembre 2012), « *le big data pourrait contribuer à renforcer d'une manière ou d'une autre la capacité d'innovation de leur organisation* ». 40 % indiquent que le *big data* rend possible de nouveaux modèles économiques basés sur la réactivité et l'interactivité en temps réel. Techniquement, nous vivons un véritable phénomène de rupture, que le McKinsey Global Institute va jusqu'à qualifier de « nouvelle frontière de l'innovation de la compétition et de la productivité ».

« Avec les big data, les données d'une organisation deviennent stratégiques. C'est la découverte d'une mine d'or gigantesque dont les ressources étaient jusque-là insoupçonnées. »

Gilles Grapinet, directeur général adjoint d'Atos²

1. D'après Intel, 31 milliards d'objets seront connectés à Internet en 2020. Dans les pays occidentaux, un cercle familial composé de quatre personnes possède aujourd'hui dix objets connectés en moyenne : ce sera vingt-cinq en 2017 et cinquante en 2022 (source : « Les objets connectés : la nouvelle génération d'Internet ? », bulletin électronique 341, ambassade de France aux États-Unis).

2 Entretien réalisé par l'Observatoire de l'innovation

Illustrations de la diversité du big data

1- Décider l'arrêt ou la poursuite de l'exploitation d'une plate-forme pétrolière est une décision stratégique qui peut être prise, grâce au *big data*, sur la base des caractéristiques géologiques de la zone étendue, des données météorologiques passées et à venir, des rendements réalisés et estimés, et des marqueurs sociologiques et politiques de la population locale.

2- Les responsables marketing peuvent se servir d'informations glanées sur les réseaux sociaux – via Facebook par exemple – pour faire des mailings finement ciblés, par exemple auprès de femmes asiatiques de moins de 50 ans pratiquant une activité sportive extérieure.

3- Les organismes de crédit peuvent coupler un système Hadoop avec un *data warehouse* et un système de détection de fraude pour assurer le suivi en temps réel des transactions.

LES 4V DU BIG DATA

Le *big data* est une démarche (un ensemble de technologies, d'architectures, d'outils et de procédures) consistant à collecter puis à traiter en temps réel, ou presque, des données à la fois très nombreuses et très hétérogènes.

Le *big data* se définit autour des 4V suivants :

- Volume : la masse de données en constante croissance (ex : plus de 2 000 petabytes en Europe) ;
- Vitesse : la collecte, l'analyse et l'exploitation des données en temps réel (ex. : 50 millions de tweets et 247 milliards d'e-mails envoyés par jour ; 100 heures de vidéos postées chaque minute sur Youtube) ;
- Variété : l'hétérogénéité des formats et des sources (*people to people, people to machine, machine to machine*) ;
- Valeur : la valorisation financière des données³.

3. À ce sujet, Viktor Mayer-Schönberger souligne bien la difficulté de quantifier la valeur d'une donnée dans son ouvrage *big data, A Revolution That Will Transform How We Live, Work and Think*. En effet, la valeur d'une donnée augmente au cours du temps et dépend des multiples utilisations qui en sont faites : elle ne se résume pas à la valeur ajoutée issue de sa première utilisation. Sa valeur peut aussi se déprécier au cours du temps, et dépend grandement de la véracité de l'information, parfois difficile à évaluer.

L'interprétation de ce large volume de données constitue le principal défi du *big data*. Car l'information générée peut être à la base d'une connaissance nouvelle entraînant des opportunités de développement d'activités. Mais les dimensions « interprétabilité », « pertinence » ou « véracité » des données collectées ne doivent pas pour autant être mises de côté. C'est l'augmentation du volume de données qui modifie profondément les analyses possibles. On évolue ainsi d'un modèle probabiliste confirmé par l'étude d'un nombre restreint de données vers un modèle dans lequel les données sont exhaustives. Partant des données, des corrélations sont observées puis éventuellement interprétées. De plus, grâce au développement des outils d'analyse, le traitement de ces données peut être qualitatif quand auparavant des approximations et des catégorisations étaient nécessaires : les ensembles étudiés peuvent donc être « microsegmentés ».

DE NOUVELLES GÉNÉRATIONS DE DONNÉES ET DE TRAITEMENT

De multiples évolutions techniques parallèles ont convergé depuis vingt ans et ainsi permis l'éclosion du *big data*. Selon Gilles Grapinet, directeur général adjoint d'Atos, quatre phénomènes permettent aujourd'hui de **traiter de la donnée structurée et non structurée de façon industrielle** :

- 1) Les technologies de **stockage** de grandes quantités de données permettent d'envisager des volumes de stockage qui étaient auparavant inconcevables et d'offrir des accès toujours plus rapides à la donnée, à un prix de plus en plus abordable⁴.
- 2) L'essor du *cloud computing*⁵ permet le passage à l'informatique sur demande (ou à la demande) et de dégager de nouvelles capacités de traitement. Au lieu de dimensionner les infrastructures pour les pics de stockage ou de traitement, les organisations peuvent désormais ajuster la taille et donc le coût de leurs infrastructures de calcul et de stockage au gré de l'évolution de leurs besoins.
- 3) La **mobilité** et l'**ubiquité** génèrent de plus en plus d'informations car les connexions sont quasi continues. En 2020, la connectivité du monde sera

4. Le prix d'un gigaoctet pour un disque dur est passé d'environ 16 dollars en 2000 à 0,06 dollar en 2013 (source : <http://ns1758.ca/winch/winchest.html>).

5. Le *cloud computing* permet d'accéder librement à des ressources informatiques partagées via un réseau de télécommunications.

décuplée avec près de 8 milliards de *smartphones* et 50 milliards d'objets connectés.

4) Enfin, les **réseaux sociaux** fournissent une nouvelle source de données personnelles et communautaires, qui peuvent être utilisées à des fins économiques.

Ces nouveaux types de données transforment les façons de traiter ces dernières. Auparavant, une donnée était créée pour un usage monodirectionnel avec une finalité unique, sans réutilisation possible. La nouvelle génération de « données en vrac » comprend de multiples sources de données (*smartphones*, comptes bancaires, passeports biométriques, titres de transport équipés de puces RFID, communications électroniques) qui sont utilisées par de multiples utilisateurs, potentiellement en même temps. De nouvelles techniques analytiques (*pattern-based analysis* ; développement du web sémantique), de nouvelles technologies de bases de données et de *processing* (essor des bases de données relationnelles et des outils de gestion des *big data* tels que Cassandra, Hadoop, GoogleFS, BigTable, Hive ou HBase), de nouvelles techniques de visualisation (par exemple le site Many Eyes d'IBM permettant de charger des jeux de données importants pour en tirer des visualisations innovantes) sont nécessaires pour exploiter commercialement ou non ces importants volumes de données.

On accumule actuellement tous les dix-huit à vingt-quatre mois d'activité humaine presque autant de données que depuis les débuts de l'humanité. Si l'on imagine que cette croissance se stabilisera probablement un jour, le flux de déversement des données continue pour le moment de s'accélérer. Nous sommes dans une phase de stockage illimité de données, et ce sur des périodes longues⁶ et sous tous les formats (image, son, conversation), afin de constituer un patrimoine informationnel. Les applications pour valoriser ce patrimoine restent encore majoritairement à inventer...

6. Si certaines données sont stockées sur de longues durées, ce n'est pas le cas de toutes. De plus, de nouveaux algorithmes en streaming permettent d'analyser les données à mesure qu'elles sont créées et ne nécessitent que très peu de mémoire (de l'ordre d'1 ko de mémoire vive).

Éléments de définition

La définition du *big data* proposée est très ouverte, car le terme est de plus en plus largement utilisé. On pourrait cependant essayer de distinguer deux démarches du *big data* : d'une part, l'utilisation plus exhaustive et rapide, ou la valorisation plus systématique et transversale, de données existantes en améliorant et automatisant les outils analytiques et prédictifs ; d'autre part, l'utilisation de nouveaux outils d'analyse de gros volumes de données pour obtenir des informations autrefois impossibles à obtenir. De plus, il est important de comprendre le sens et l'implication en termes de *big data* des notions suivantes :

Les « facilitateurs » du *big data*

- **L'Internet des objets** décrit les interactions croissantes entre les objets de la vie courante et leurs représentations virtuelles, à l'origine via des étiquettes électroniques passives qui contiennent une information numérique relative à l'objet physique, et de plus en plus au moyen de capteurs qui connectent directement ces objets à l'Internet sans fil. Le développement de l'Internet des objets est notamment possible grâce aux puces RFID.

- **Le cloud computing**, ou l'Internet en nuage, transforme les infrastructures de stockage et la puissance de calcul en services, par l'intermédiaire d'entreprises qui possèdent des serveurs et louent leurs capacités. Cette approche permet de mutualiser les besoins et d'accorder aux usagers une grande flexibilité de stockage et de rapidité de traitement de données. L'essor du *cloud* et l'émergence de nouveaux outils *open source* tels MapReduce ou Hadoop a permis le traitement de grands ensembles de données, sur plusieurs serveurs en parallèle.

- **Le data mining** est un processus d'analyse informatique qui regroupe des méthodes d'intelligence artificielle, d'apprentissage automatique et de statistiques dans le but d'extraire des informations compréhensibles d'un ensemble de données. Le *data mining* traditionnel s'est développé et enrichi pour permettre l'analyse des nouvelles sources de données à l'ère du *big data*.

Un mouvement parallèle

- Les données ouvertes, l'*open data*, sont des données dont la licence permet le libre partage et le réemploi. Pour que ce réemploi soit réel, elles doivent être interopérables et structurées. L'*open data* qualifie aussi le fait de rendre accessibles et réutilisables les données produites et détenues par des institutions, dans un objectif de transparence ou de création de valeur. C'est déjà le cas des données publiques dans le cadre de l'initiative Etalab et de la directive européenne 2003/98/CE, une nouvelle version étant actuellement discutée par l'Union européenne pour adapter l'accès aux données publiques aux évolutions récentes.

Parmi les données concernées par le mouvement de l'*open data*, certaines ne représentent pas des volumes suffisants pour nécessiter un traitement statistique complexe et être qualifiées de *big data* à proprement dit, mais leur diffusion et leur exploitation sont rendues possibles par les nouveaux outils de visualisation développés face à l'augmentation du nombre de données. Vice versa, les *big data* ont tout à gagner à une large diffusion des données, notamment publiques. Les deux mouvements sont donc liés.

BIG DATA, LEVIER D'INNOVATIONS

Les applications du *big data* sont nombreuses et constituent un facteur de renforcement de la capacité d'innovation des organisations, en jouant sur les deux dynamiques que sont l'exploration et l'exploitation.

En tirant parti de données jusqu'ici non disponibles, la démarche *big data* rend possible le lancement de nouveaux produits et services, de nouveaux processus, et même de nouveaux modèles économiques. On parle de **dynamique d'exploration** ou encore de développement *greenfield*. Mais les *big data* peuvent également apporter des réponses à des défis *business* non résolus jusque-là permettant d'offrir des solutions d'amélioration et d'optimisation de l'activité traditionnelle de l'entreprise. On parle alors de **dynamique d'exploitation** ou de développement *brownfield*.

LES NOUVEAUX USAGES STIMULÉS PAR LES *BIG DATA*

Qu'ils permettent d'exploiter plus, mieux ou plus vite les activités existantes ou de générer de nouvelles opportunités de développement d'activité, les *big data* se révèlent un facteur clé d'innovation pour les petites comme pour les grandes entreprises de tous secteurs.

Le *big data* permet de passer d'une analyse partielle à une **économie prédictive**. L'analyse actuelle des données, en mode réactif aux événements passés sur la base de lois statistiques à partir d'échantillons limités, offre aujourd'hui une représentation tronquée de la réalité. Il est désormais possible d'évoluer vers un modèle prédictif, fondé sur une analyse exhaustive des informations de différentes sources, afin de trouver des relations permettant de prédire un événement futur. Gilles Grapinet illustre ce propos en expliquant que la révision obligatoire des 15 000 kilomètres pourrait ne plus être utile dans un monde *big data* où chaque voiture pourrait être automatiquement rappelée au garage dès qu'elle en présenterait le besoin.

- **Détecter et optimiser** : le croisement en temps réel d'un grand nombre de données diversifiées autorise une meilleure connaissance des activités, de l'environnement et de l'écosystème d'affaires. Un contrôle accru est alors possible, permettant une meilleure aide au pilotage et à la prise de décision.

Exemple : la société danoise Vestas Wind Systems, parmi les plus importants fabricants de turbines d'éoliennes au monde, utilise les solutions IBM *big data analytics* et IBM Systems pour décider de la localisation des éoliennes en croisant en quelques heures (vs quelques semaines) des données variées telles que des données météorologiques et géospatiales, les phases de la lune et de la marée, des cartes de déforestation et des images satellites. L'emplacement des éoliennes est un enjeu crucial pour optimiser la production d'énergie. Si les compagnies d'électricité installent les éoliennes à des endroits mal adaptés (pas assez ou trop de vent), elles risquent de ne pas produire assez d'électricité pour justifier les investissements dans l'énergie éolienne. Vestas espère ainsi conquérir de nouveaux marchés en accélérant la diffusion de l'énergie éolienne au niveau international.

Exemple : l'*Ingestible Event Marker* est une nouvelle technologie, conçue par la société Proteus Digital Health et approuvée par la FDA en juillet 2012, qui permet d'insérer un capteur miniaturisé dans une gélule devenant alors une « *smart pill* ». Un message d'alerte en cas de non-respect de la prescription (délai ou dosage) peut ainsi être transmis à un récepteur externe, lorsqu'il est activé par les sucs gastriques, rendant possible un suivi et un contrôle en temps réel de la prise de médicaments.

-
- **Tracer et cibler** : les *big data* permettent d'analyser la situation et le contexte de milliers de personnes en temps réel, favorisant ainsi une meilleure compréhension des réactions du marché et la proposition de messages et d'offres personnalisés.

Exemple : de nombreux véhicules possèdent aujourd'hui des systèmes GPS et télématiques qui permettent aux constructeurs automobiles de collecter et monétiser une multitude de données sur les habitudes de conduite de

leurs clients. En 2007, le système télématique OnStar de General Motors Corporation et GMAC Insurance se sont associés pour créer une nouvelle offre se servant des données télémétriques afin de proposer des primes d'assurance aux clients qui roulent moins. Aujourd'hui, la tarification classique des assurances impose en effet le même prix à tous les clients, qu'ils conduisent bien ou mal, peu ou beaucoup... Ou, au contraire, discrimine des franges entières de population en fonction de l'âge (ex. : premium de plus de 50 % pour les plus jeunes conducteurs) ou du sexe par exemple, sans faire de détails à l'intérieur de ces groupes. La personnalisation de la prime d'assurance constitue une source de satisfaction client accrue et une innovation disruptive, d'autant qu'elle intervient sur un marché de reconduction.

Exemple : comment utiliser les *big data* dans un aéroport ? D'un côté, les passagers doivent obligatoirement passer un certain nombre de *checkings* de sécurité (ce qui génère du stress, des temps d'attente) ; de l'autre côté, ils peuvent flâner dans les boutiques et autres restaurants occupant l'espace. Une application de géolocalisation dans l'aéroport pourrait offrir aux passagers, en retour de leurs données émises, un guide interactif du lieu, des informations sur les temps d'attente aux différents postes de sécurité, sur les horaires des vols et portes d'embarquement. L'aérogare serait alors en mesure de communiquer en temps réel avec des milliers d'individus, de diminuer le niveau de stress des passagers en leur permettant de mieux gérer leur temps et probablement par conséquent de les amener à consommer plus (on pourrait même parler « d'attente intelligente »). Les services présents dans l'enceinte de l'aéroport pourraient également avoir l'opportunité, via cette application, de « hooker » ces clients « captifs » avec des coupons promotionnels customisés en fonction de leur position dans l'aéroport, de leur situation de voyage et de leurs préférences.

- **Prévoir et prédire** : l'analyse prédictive permet de faire des projections ultra-réalistes pour identifier des nouvelles sources d'opportunités (ou des menaces) et ainsi anticiper les réponses adaptées à la situation réelle.

Exemple : Flu Trends est un modèle de Google qui permet d'estimer la propagation des cas réels de grippe en fonction des recherches sur Internet

faites autour de ce mot – prenant en compte le fait que tous les internautes qui tapent « grippe » sur leur clavier ne sont pas nécessairement malades. Au final, l'évaluation que fournit Google en temps réel sur la propagation de la maladie est étonnamment précise (cf. article publié dans la fameuse revue scientifique *Nature*) et souvent plus rapide que les réseaux de veille sanitaire. Les estimations du niveau d'activité grippale permettent d'évaluer l'ampleur de l'épidémie et de commencer à prendre les mesures qui s'imposent le cas échéant. Dans le même esprit, les Nations unies ont développé un programme anticipant les épidémies et les retournements de conjoncture économique à travers des mots-clés échangés sur Twitter.

La démarche *big data* ouvre de nouveaux champs du possible en établissant des relations entre des ensembles de données hétérogènes pour trouver des informations pertinentes et être proactif, dans tous les secteurs d'activité. À ce titre, les gagnants sont les médias, l'électronique grand public, les services financiers et les administrations, et ensuite la distribution, le tourisme et le transport, les *utilities* et les acteurs du système de santé. La capacité d'un secteur à exploiter les *big data* varie en fonction des contextes culturels, réglementaires ou de compétences. En la matière, les services financiers, la distribution, le tourisme et le transport semblent plus en avance que le secteur public, l'énergie ou la santé.

Exemple : dans la grande distribution, la géolocalisation en magasin ou la reconnaissance faciale dans des vidéos peuvent permettre à un distributeur d'analyser les parcours clients ou encore le temps passé dans chaque rayon et simuler en conséquence différentes options de *merchandising*.

Exemple : dans la banque/assurance, l'analyse des flux de transactions peut permettre de détecter en temps réel des événements aberrants et de déclencher automatiquement des décisions telles que le blocage d'une carte de crédit, par exemple, en cas de suspicion de fraude bancaire.

Exemple : dans le secteur de la santé, l'analyse prédictive permet de repérer des schémas récurrents de certaines maladies et d'accélérer la prise en charge des patients concernés, même si ces derniers ne sont pas en me-

sure de décrire leurs symptômes, ou encore de fluidifier le parcours d'hospitalisation des patients en améliorant l'échange d'informations et la coordination entre les différents services (entre les ambulanciers, les urgentistes, le médecin traitant, le chirurgien, la sécurité sociale).

Les voies du *big data* à la SNCF

Avec plus de 10 millions de clients par jour, la SNCF évolue de fait dans un univers de *big data* ; c'est pourquoi cette thématique a été naturellement identifiée comme l'un des axes majeurs de l'innovation au sein du groupe. Le soutien du *top management* est un des facteurs clés de succès de la démarche, avec des projets portés aujourd'hui par les hauts responsables de différentes entités comme SNCF Transilien ou SNCF Voyages.

Les exemples d'application du *big data* au sein de la SNCF sont nombreux et concernent notamment :

* **La maintenance curative et préventive** des caténaires, des 52 800 km de voies et de la signalisation à partir des données collectées par des capteurs ou des systèmes de mesure embarqués. Ainsi, aux trains métrologiques, qui scrutaient les voies à intervalles réguliers, vont se substituer progressivement l'ensemble des trains commerciaux, équipés de capteurs, permettant d'établir des points de comparaison dans le temps, entre lignes ou composants, grâce à des traitements statistiques massifs.

* **L'efficacité énergétique.** La SNCF est le premier consommateur français d'énergie, depuis que les industries du nucléaire ont baissé leur consommation en améliorant les techniques de centrifugation. La SNCF utilise aujourd'hui les compteurs qui équipent les trains, et qui doivent être généralisés⁷, pour établir la facturation de sa consommation énergétique. À l'avenir, ces mêmes compteurs pourraient être utilisés pour optimiser sa

7. La SNCF prévoit ainsi d'équiper toute sa flotte de compteurs embarqués qui vont brasser une quantité de données gigantesque.

consommation, sa facture et son bilan carbone, trois enjeux aujourd'hui parfois divergents.

* **La régulation des flux de voyageurs.** Le recours aux *big data* relève ici de l'aide décisionnelle, avec un comptage des voyageurs entrant et sortant à chaque arrêt commercial qui permet de mesurer en temps réel la charge des trains et la répartition des voyageurs. À l'avenir, on pourrait imaginer des applications mobiles permettant aux voyageurs de calculer la ponctualité de leur train ou de les orienter vers les voitures les moins remplies dans les rames.

* **La personnalisation de la relation client et un meilleur ciblage marketing,** à travers l'analyse du trafic sur le site voyages-sncf.com, premier site de e-commerce en France avec 11 millions de visiteurs uniques (le e-commerce représente aujourd'hui 10 % du chiffre d'affaires du groupe). L'analyse fine de ses clients va amener la SNCF à proposer de nouvelles offres personnalisées, en se positionnant comme intégrateur global de la mobilité face aux défis de l'intermodalité. Des applications SNCF sur *smartphone* ou Internet, comme Mytripset, permettent déjà d'organiser ses voyages en Europe en train, en avion ou en voiture.

Les obstacles au déploiement du *big data* sont de plusieurs ordres, actuellement, au sein de la SNCF : a) internes (liés à la complexité de l'organisation ou à la culture du groupe) ; b) externes (rôle de la réglementation et des autorités régulatrices, rôle de l'État actionnaire) ; ou c) liés aux limites mêmes du phénomène *big data*.

a) La SNCF est un groupe de mobilité organisé autour de trois grands métiers : le transport de voyageurs, le transport de marchandises et les infrastructures. Ces trois segments impliquent des périmètres de données différents, qu'il n'est pas toujours évident de faire dialoguer. « *Ce sont d'énormes chantiers pour faire se parler les DSI, créer les passerelles* », explique Éric Conti, directeur de l'innovation et de la recherche de la SNCF. De plus, il importe de souligner les contraintes liées à l'interopérabilité et à la compatibilité des systèmes de production de données entre les différentes zones géographiques sur lesquelles opère la SNCF : « *Quand le train passe de France en Belgique, ce n'est pas la même chose. Eurostar porte sept systèmes de signalisation à bord !* » « La » donnée est donc souvent

complexe, même si tout l'enjeu de la démarche *big data* est bien de prendre des données hétérogènes et de les amalgamer. Le premier enjeu du *big data* à la SNCF relève donc de la structuration des données et de « *la construction d'une culture collective de la donnée* ». Cette construction implique « *un vrai changement culturel pour que la prise de conscience de la donnée s'infilte dans l'entreprise* ». Aujourd'hui, en effet, « *la donnée est souvent un consommable, voire un jetable quand elle est acquise par des moyens non informatiques* ». Au-delà des techniques, il faut user de pédagogie pour faire naître auprès des salariés la prise de conscience que la donnée est une richesse. Une autre transformation à mener est liée à la mise en œuvre de compétences (en termes d'analyse statistique, de *data visualisation*, d'optimisation) qui ne sont pas toujours disponibles en interne, ou alors en quantité insuffisante ; au-delà des réponses en termes de politique de recrutement, cette difficulté est aujourd'hui en partie levée grâce aux partenariats noués par le groupe auprès des *start-up*, notamment au travers du Camping, l'incubateur de Silicon Sentier.

b) À ces transformations nécessaires en interne s'ajoute la complexité du paysage institutionnel. Ce dernier n'est pas encore complètement dessiné car le processus de libéralisation du rail est en cours au niveau européen. De plus, il existe une multiplicité d'acteurs dans le domaine du transport qui ne permet pas toujours d'avoir une relation directe avec le client. Dans certains cas, l'autorité régulatrice peut être propriétaire de la donnée (ex. : Stif). La stratégie actuelle de la SNCF consiste justement à devenir « un industriel des services B2C ». « *Nous ne maîtrisons pas toute la chaîne des services, il faut des montages institutionnels très variés. Sncf.com fait du B2C, mais face au Stif nous sommes en B2B. Comment avoir une même politique de la donnée ?* », s'interroge Éric Conti. Par ailleurs, la spécificité du statut de la SNCF – un Epic, établissement public à vocation industrielle et commerciale – induit un positionnement ambigu vis-à-vis de l'exploitation des données. Ainsi, en matière d'*open data*, la dimension « établissement public », qui renvoie à des missions d'intérêt général, incite à l'ouverture des données, tandis que la dimension « à vocation industrielle et commerciale » renvoie davantage au secret des affaires, et donc à une moindre transparence. La réglementation peut également se révéler être un obstacle à la mise en œuvre d'une démarche *big data*. L'usage de la vidéosurveillance, dans les gares notamment, pourrait être élargi au-delà des questions liées à la sécu-

rité des voyageurs. La vidéo est en effet une source très riche de data, et les caméras pourraient être utilisées pour compter et réguler les flux de passagers. Un tel usage se heurte toutefois aux limites de la réglementation actuelle en matière de conservation des images.

c) Il convient également de noter certaines limites propres au *big data*. Ces limites sont de deux ordres. La démarche *big data* est souvent associée au développement des analyses prédictives et en temps réel. Ces dernières ont de nombreuses retombées bénéfiques, par exemple sur le réseau Transilien, où des décisions à prendre sur l'affectation des trains ou des circulations doivent l'être en moins d'une minute. Toutefois, dans certains cas, elles peuvent se révéler contre-productives si elles ne correspondent pas à un besoin réel des consommateurs. Ainsi, dans le domaine de l'information voyageur, l'actualisation trop fréquente des données peut être une source de stress pour les clients. Le *big data* ne doit donc pas être perçu seulement comme une technique qui devrait être poursuivie pour elle-même, mais comme un outil qui trouve sa valeur auprès des usages qui peuvent en découler.

Les nouvelles offres permises par le *big data* viennent compléter une offre de service préalable, plus qu'elles ne viennent s'ajouter à cette dernière. La mobilité est d'abord physique et il est difficile de séparer la question du transport physique du service immatériel offert au client. Autrement dit, l'offre nouvelle qui pourrait naître du *big data* suppose que l'offre de base – celle du transport physique – soit correctement remplie. « À chaque fois qu'on a voulu offrir de nouveaux services, dès lors que le service de base n'était pas assuré, cela a été un échec », plaide Éric Conti. L'utilisation du *crowdsourcing* dans la maintenance des équipements a pu ainsi connaître quelques ratés dès lors que l'organisation physique pour tenir compte des alertes clients – soit réparer dans des délais très courts les dysfonctionnements signalés – était sous-dimensionnée. La SNCF reste également vigilante à ce que ce potentiel d'automatisation des processus et des décisions n'aboutissent pas à de « *nouveaux services déshumanisés, sans personne pour dialoguer en face* ».

En conclusion, la SNCF produit aujourd'hui un volume important de données ; l'exploitation de ces dernières avec des nouvelles données exogènes (météo, événementiel, etc.) constitue un gisement important d'accroisse-

ment de ses performances techniques et services. Le dilemme pour l'entreprise est de parvenir à concilier cette exploitation avec la réalité du ROI pour allouer au mieux ses investissements.

Le big data et l'automobile

Les voitures sont de plus en plus équipées de capteurs et de logiciels embarqués, qui leur permettent d'analyser leur environnement et d'agir en conséquence. Certaines voitures possèdent déjà des technologies qui permettent une conduite partiellement autonome, comme une option pour se garer automatiquement, ou une assistance aux embouteillages pour maintenir la distance nécessaire vis-à-vis des autres véhicules.

La voiture se personnalise en intégrant et utilisant les données, devenant connectée, voire sans conducteur. Ces services personnalisés sont un nouveau moyen de fidéliser les clients, qui prend de l'ampleur par rapport au produit en lui-même.

La voiture connectée

Celle-ci repose sur un échange de données issues des capteurs présents dans le véhicule avec l'extérieur. Selon les chiffres d'ABI Research, 80 % des véhicules qui seront produits et commercialisés en Europe et aux États-Unis en 2017 seront connectés.

Les interfaces homme-voiture, à l'origine les GPS et les *smartphones*, offrent déjà de plus en plus de services en lien avec les données collectées par ces capteurs. Ensuite, de nouveaux services exploitent les données issues des capteurs et les croisent avec d'autres sources : services de guidage qui allient le GPS avec la connaissance du trafic et les positions des parkings ou stations-service ; suivi de l'usure des pièces et télédiagnostic en cas de panne, autorégulation de la circulation.

Le regroupement d'une multitude de données sur les habitudes de conduite des clients peut aussi être monétisé par les fabricants : si le système télématique OnStar de General Motors Corporation, par exemple, propose des

services de sécurité, d'information et de diagnostic des véhicules aux automobilistes, il recueille également des données télémétriques. En 2007, OnStar et GMAC Insurance se sont associés pour créer une nouvelle offre qui se sert de ces données afin de proposer des primes d'assurance aux clients qui roulent moins que les autres : c'est le concept du « *pay as you drive* ».

Pour répondre aux défis du *big data*, le Car Connectivity Consortium (CCC) regroupe plusieurs constructeurs dont PSA, BMW et General Motors, des équipementiers et des acteurs des technologies informatiques comme Samsung ou Nokia. Il a notamment lancé le projet Genivi, qui développe un socle commun pour les services d'information ou de divertissements dans les véhicules, en mutualisant les efforts dans le développement des logiciels enfouis et la standardisation des équipements. Les développements issus de ce projet sont sous licence *open source*. Genivi a aussi développé Mirrorlink, un standard commun permettant de reporter, via un raccord USB ou par Bluetooth, le contenu du *smartphone* sur l'écran de bord.

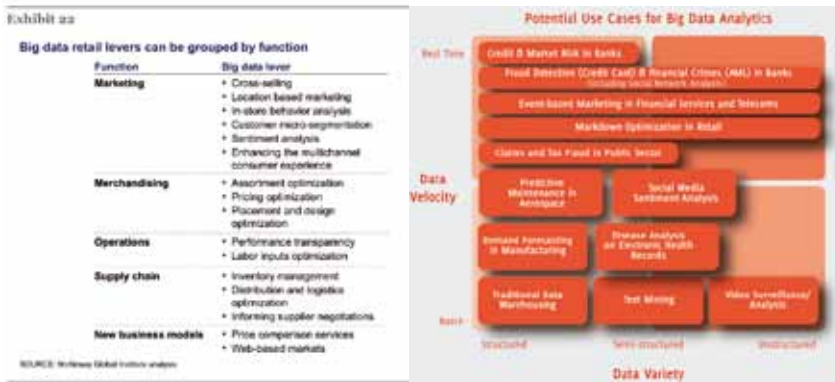
La voiture sans conducteur

Si des esquisses de voiture autonome ont existé depuis environ trente ans, comme le projet Eureka Prometheus de l'Union européenne, les nouveaux outils de collecte et de traitement des données permettent leur concrétisation. Ainsi, la Google Car, entièrement autonome, est testée depuis 2010 sur des routes américaines. Le logiciel qui permet sa conduite a été perfectionné par des outils d'apprentissage automatique, grâce aux données récoltées lors des centaines de milliers de kilomètres de test. De nombreux constructeurs dont Volvo, BMW, Daimler, Tesla Motors et Nissan soutiennent des projets de R&D sur les voitures autonomes. À terme, selon l'organisme américain de sécurité routière, les voitures autonomes permettraient d'éviter 80 % des accidents.

En théorie, la voiture autonome n'a pas besoin d'être connectée à Internet : Mitsuhiro Yamashita, responsable en R&D chez Nissan, a expliqué que Nissan ne s'appuie quasiment pas sur les infrastructures (capteurs extérieurs et Internet) et que toute la technologie se trouve dans le véhicule. Cependant, une connexion Internet pourrait être utile pour obtenir des informations complémentaires, par exemple lorsque le relief ou la route sont masqués par la neige.

SOURCE D'INNOVATIONS ORGANISATIONNELLES

La démarche *big data* met la donnée au cœur des activités des organisations, au-delà des fonctions informatiques ou statistiques. Les trois domaines privilégiés des initiatives *big data* sont l'**analyse financière** (58 %), la **connaissance client** (53 %) et l'**identification de tendances à long terme (prévisions)** (47 %), selon le *big data Index EMC/IDC*. Cette démarche transversale que sont les *big data* n'est pas l'apanage d'une direction en particulier.



Source : VINT big data research report

Les implications *big data* pour les fonctions **logistique** et **marketing** apparaissent de façon assez évidente. La collecte et l'exploitation d'informations sur des produits connectés ou les comportements des consommateurs permettent d'améliorer le fonctionnement opérationnel de l'entreprise et de prévoir l'évolution des marchés. Les prévisions temporelles des ventes croisées avec les données de stocks permettent de calculer précisément les besoins en réapprovisionnement et d'en optimiser la logistique. Le développement de nouveaux produits et services et leur adaptation au plus près des attentes et des usages sont facilités par l'apparition de **microsegmentations**. Combinées à la mobilité, les *big data* réinventent l'interaction avec l'utilisateur, cherchant à délivrer le bon service, à la bonne personne, au bon endroit et dans le bon *timing*.

Exemple : Lokad, *start-up* française éditrice de logiciels, a développé des algorithmes qui aident les distributeurs à optimiser quotidiennement leurs niveaux de stock en analysant les tickets de caisse et en se basant sur l'historique des ventes. Trouver le juste niveau de stock (ni trop, ni trop peu) permet aux clients d'optimiser leur marge. En adoptant les solutions Microsoft telles que la plate-forme Azure, Lokad a démultiplié ses capacités de stockage et de prévision statistique. « Nous comptons sortir cent fois plus de prévisions et atteindre les 100 millions de prévisions à l'heure », explique le fondateur de l'entreprise Joannès Vermorel. La précision des corrélations statistiques permet aujourd'hui à Lokad de détecter la saisonnalité d'un produit en vente depuis moins de trois mois. La société a ainsi pu diversifier son portefeuille de clients en s'adressant à de petites sociétés et à des e-commerçants caractérisés par de petits stocks à faible rotation, grâce à la technologie de « prévisions quantiles ». Lokad envisage désormais de s'attaquer à des marchés non traditionnels, comme celui de la gestion de projets de construction, qui pourraient bénéficier de son expertise d'optimisation des prévisions.

Exemple : créée en 2010 par David Bessis et Jakob Haesler, Tinyclues pratique le « *profiling* prédictif » au sein de bases de données fournies par ses clients, e-commerçants ou médias digitaux, pour comprendre le profil et les préférences du consommateur à partir de « signaux faibles » (identité, adresse e-mail, mots-clés tapés et pages Internet consultées) et être ainsi en mesure de lui recommander des produits en ligne avec ses précédents comportements d'achat et ses envies actuelles. Le modèle économique de Tinyclues repose sur une logique d'abonnement mensuel basée sur le volume de la base de données traitée. L'ambition de cette *start-up* française serait d'égaliser les performances de son aînée Criteo, moteur prédictif en temps réel lancé en 2005 pour analyser et prévoir les comportements des consommateurs et aujourd'hui leader mondial de la publicité en ligne. La stratégie de conquête des nouveaux marchés de ce spécialiste du *retargeting*, ou reciblage publicitaire basé sur un algorithme de recommandation, est confirmée par une récente étude qui affirme que 20 % des internautes cliquent sur les bannières publicitaires ciblées.

La **gestion des risques** et de la **qualité** est également plus fine. Les capacités de réaction en cas de pannes, d'incidents, de pics, de pandémies... peuvent alors être raccourcies et affûtées. D'autres fonctions, comme les **ressources humaines**, peuvent être favorablement impactées par l'essor du *big data*, via l'optimisation du recrutement, grâce à la lecture sémantique des CV par exemple.

Exemple : PredPol (predictive policing) est un algorithme conçu pour prédire où et quand des délits ont le plus de probabilités de se produire. Grâce à une base de données recensant les infractions passées, les forces de police peuvent, comme dans le film de science-fiction de Steven Spielberg *Minority Report* (2002), se rendre sur les lieux avant même que les actions soient commises. Utilisé également pour prévenir les catastrophes naturelles comme les séismes, ce type d'analyse est une nouvelle façon de hiérarchiser les priorités et d'allouer les ressources humaines et financières.

Exemple : Evolv, un spécialiste du recrutement et de l'optimisation des performances des employés, a récemment réalisé une enquête auprès de 20 000 employés. Sa principale conclusion : il n'existe aucune corrélation sérieuse entre la réussite d'un candidat dans le travail qui lui est proposé et l'expérience accumulée par ce même candidat dans les postes qu'il a occupés précédemment. Et inversement, ce n'est pas parce qu'un candidat a été inscrit au chômage pendant une période relativement longue (plus de six mois) qu'il ne sera pas capable de répondre aux attentes de son futur employeur. Evolv affirme en conséquence que les anciennes méthodes de recrutement sont dépassées. L'analyse de certaines interactions (ex. : les employés entretenant des relations avec au moins trois collègues restent plus longtemps dans leur entreprise que ceux qui n'y ont pas d'amis) ou de certaines pratiques dans les réseaux sociaux (ex. : un employé actif sur les réseaux sociaux est plus enclin à rester fidèle à son entreprise et donc à y rester longtemps) serait plus performante que l'étude du parcours du candidat. L'entreprise enjoint donc les recruteurs à s'intéresser aux nouvelles technologies analytiques pour améliorer l'efficacité de leurs employés et de leur recrutement. Selon cette même étude, l'utilisation des techniques analytiques et l'exploitation du *big data* renforceraient la relation de confiance entre employeurs et employés : la fidélité des employés recrutés via l'analytique augmenterait de 35 % et leur performance grimperait de 30 %.

« MagicBands », les bracelets Disney source de *big data*

Depuis le printemps 2013, les parcs d'attractions et hôtels Disney équipent leurs clients de bracelets, les « MagicBands ». Ceux-ci sont munis de puces RFID et servent à la fois de tickets tout-en-un, de moyen de paiement sur place et de clé pour ouvrir sa chambre dans un *Disney Resort*. La RFID (*radio frequency identification*) est une technologie qui transfère à distance les données contenues dans une étiquette ou une puce émettant des ondes radio. Elle permet donc d'identifier et de suivre les objets équipés de ce système. Longtemps utilisée comme simple outil de suivi des inventaires, la RFID est aujourd'hui plus largement utilisée et équipe désormais les passeports, les badges d'entreprises et, depuis peu, les « MagicBands » de Disney. Dans ce dernier cas, la puce RFID peut contenir des informations sur le porteur du bracelet, comme son identité, ses préférences de consommation ou ses coordonnées bancaires.

Ces bracelets nouvelle génération font partie d'une initiative plus large de Disney appelée « MyMagic+ » – développée depuis février 2011 – qui vise à rendre l'expérience consommateur plus interactive et personnalisée. L'investissement consenti est évalué entre 800 millions et 1 milliard de dollars – avec l'espoir qu'un client plus satisfait dépensera davantage... Thomas O. Staggs, président de Disney Parks and Resorts, explique : « *We continually push the boundaries of creativity and innovation to deliver the best possible experience for our guests. [...] Over the past few years, we've devoted considerable time and resources to create a more immersive, more seamless and more personal experience for each and every guest who spends time with us.* »

« MyMagic+ » rassemble :

- « My Disney Experience », qui permet de planifier en avance ses vacances et qui se décline sous forme d'application mobile et de site Internet ;

- « Fastpass+ », qui permet de choisir trois coupe-files dans les attractions de son choix ou de réserver des places VIP pour certains événements. En cas d'hésitation, le programme peut utiliser les données entrées par l'utilisateur pour lui faire des recommandations ;

- Le « MagicBand », permettant aux utilisateurs de payer ou de participer à des attractions sur simple présentation de leur bracelet, mais aussi à l'équipe Disney de suivre leur temps d'attente aux attractions et leurs consommations en temps réel au sein du parc. Ce bracelet sert également de clé d'hôtel, permettant ainsi une meilleure continuité des services entre les différents univers.

Avec près de 42 millions de visiteurs dans ses parcs d'attractions en 2012, Disney a accès à un immense flux de données. Ces données portent sur l'identité des visiteurs, leurs parcours dans le parc, leurs achats et préférences (via les coupe-files choisis par exemple). Ces informations sont collectées à partir de leur bracelet et via le site Internet et l'application « My Disney Experience ».

Ces données seront utilisées par exemple pour faire face à des problèmes structurels de Disney comme des temps d'attente trop longs dans le parc pour les clients ou une difficulté pour Disney à allouer efficacement ses employés. Jusqu'ici les visiteurs entraient dans le parc sur présentation de tickets papier et circulaient ensuite librement dans le parc. Des kiosques vendaient des coupe-files sur place mais, pris d'assaut, ils étaient rapidement à court. Les temps d'attente pour la plupart des visiteurs se révélaient alors interminables. En traçant les parcours des visiteurs, Disney pourra être à même de réduire les pics d'affluence et d'améliorer la fluidité dans ses parcs en proposant par exemple à un client qui a réalisé la plupart des montagnes russes d'aller en faire une nouvelle car, à cette heure, le temps d'attente est inférieur à celui de l'attraction pour laquelle il patiente présentement. De même, le « Fastpass+ » et « My Disney Experience » proposent aux visiteurs de planifier davantage leur séjour en réservant à l'avance le bracelet, un dîner ou encore des coupe-files. Une fois sur place, les visiteurs peuvent encore utiliser l'application sur leur *smartphone* pour changer leur plan, se diriger dans le parc et recevoir des alertes sur la programmation (heure et lieu de passage des parades) ou les temps d'attente. En couplant ces données avec l'analyse du nombre de visiteurs par heure, Disney tentera ainsi d'optimiser l'allocation et le planning de ses employés.

Mais, au-delà de l'optimisation de la gestion des ressources humaines et du temps d'attente, cette initiative offre de nouvelles perspectives au groupe.

Ce dernier pourra offrir une expérience plus interactive à ses clients. Les personnages de Disney parcourant le parc seront à même, grâce aux données contenues dans les bracelets, de saluer les enfants par leurs prénoms en les croisant ou de leur souhaiter un joyeux anniversaire le cas échéant. Ces informations sont également un atout précieux pour le lancement d'opérations efficaces de marketing direct. Par exemple, si l'heure du déjeuner approche et que le temps se couvre, le visiteur recevra un coupon de réduction pour le restaurant le plus proche sur son *smartphone*. Une utilisation intelligente de ces *big data* générerait donc des revenus additionnels et une meilleure fidélisation des clients.

Néanmoins, Disney doit faire face à des défis afin de faciliter l'adoption de ces nouveaux services, notamment l'acquisition de nouvelles compétences en interne, la sécurisation des données et la gestion des réticences des consommateurs désireux de protéger leur vie privée. Une formation sera dispensée à près de 60 000 employés. Le port du bracelet ne sera pas obligatoire et les visiteurs pourront choisir le niveau d'informations qu'ils souhaitent révéler. Les informations transmises à Disney par le bracelet seront sources de contre-valeur pour les clients. Par exemple, les visiteurs qui accepteront de communiquer leur numéro pourront recevoir des coupons de réduction sur leur *smartphone*.

BIG DATA, FACTEUR CLÉ DE PERFORMANCE

Dans une étude publiée en 2011, le McKinsey Global Institute démontre que les *big data* induisent des gains substantiels dans de nombreux secteurs d'activité. **Le temps de développement pourrait être réduit de 25 %** dans l'industrie du fait d'une meilleure productivité au niveau du design, de la production et de la distribution. Dans le *retail*, les revendeurs seraient plus en phase avec les goûts des consommateurs ; ils optimiseraient leurs stocks et pourraient **accroître leurs marges nettes de 60 %**. Dans le secteur public, ces techniques d'analyse des données fluidifieraient les processus administratifs, avec, à la clé, une baisse des coûts de 15 à 20 %, soit de 150 à 300 milliards d'euros rien qu'en Europe... Les *big data* devraient **créer d'ici 2015, 4,4 millions d'em-**

plois dans le monde et 1,9 million d'emplois dans le secteur informatique aux États-Unis, selon la société Gartner⁸. En comptant les emplois indirects générés, ce sont même près de 6 millions d'emplois qui pourraient être générés à cet horizon pour les seuls États-Unis.

Exhibit 1

Big data can generate significant financial value across sectors



SOURCE: McKinsey Global Institute analysis

Les *big data* sont susceptibles d'accroître la performance globale des organisations car ils permettent de :

- Réduire les délais et les coûts grâce à des prévisions ajustées.
- Améliorer la satisfaction client grâce à un *feed-back* continu sur les produits et l'état des relations client, à des communications marketing personnalisées et à une optimisation des offres commerciales.
- Gérer mieux les risques en les détectant plus tôt et en ayant plus d'informations pour y remédier de façon adaptée.
- Accélérer le développement de produits, services et processus innovants en renforçant les processus de décision.

8. « Gartner Says Big Bata Creates Big Jobs : 4.4 Million IT Jobs Globally to Support Big Bata By 2015 » (<http://www.gartner.com/newsroom/id/2207915>).

Big data et productivité

Deux chercheurs du Massachusetts Institute of Technology (Sloan Business School), Erik Brynjolfsson et Heekyung Kim, en partenariat avec Lorin Hitt, chercheur à l'université de Pennsylvanie (Wharton Business School), ont publié une étude menée auprès de 179 entreprises cotées. L'étude montre que les organisations qui s'appuient sur l'analyse des données pour développer leur stratégie augmentent leurs niveaux de productivité de 5 à 6 %⁹.

Une étude de l'université du Texas portant sur 150 entreprises de la liste « Fortune 1 000 » tente d'étudier les gains de performance financière, opérationnelle et commerciale qui peuvent être suscités par des investissements dans les données, donc dans l'amélioration de leur qualité. Les résultats montrent qu'une bonne utilisation des données favorise la productivité dans trois cas :

- L'amélioration du caractère utilisable des données de 10 % permet des gains de productivité des employés : par exemple 15 % de ventes supplémentaires.
- L'amélioration de la mobilité des données de 10 % entraîne une augmentation du ROIC (*Return on Invested Capital*) de 1,4 %, et une augmentation de la productivité des actifs de 7,28 % via une meilleure gestion et utilisation de ces actifs.
- L'amélioration de l'intelligence des données de 10 % génère une augmentation de 18,5 % dans la précision des prévisions et de la programmation. Si l'accessibilité augmente aussi de 10 %, une augmentation moyenne des revenus est constatée, de 0,81 % grâce au développement de nouveaux produits, et de 0,7 % grâce à de nouveaux consommateurs¹⁰.

9. « Strength in Numbers : How Does Data-Driven Decisionmaking Affect Firm Performance ? », *Social Science Research Network Working Paper Series*, avril 2011.

10. Les résultats obtenus varient beaucoup selon les secteurs. Au niveau des bénéfices financiers et des bénéfices en termes de consommation, ils sont souvent plus élevés pour les entreprises des télécommunications, du conseil et de l'industrie pétrolière (aux investissements lourds).

L'IMPACT DU *BIG DATA* POUR LES ENTREPRISES

Le *big data* est aujourd'hui une réalité émergente au sein des entreprises françaises selon le *Big Data Index France EMC/IDC* réalisé à l'été 2012 auprès de 160 entreprises de plus de 200 salariés. Le terme *big data* n'évoque rien à 35 % des répondants et 91 % déclarent que leur entreprise n'a pas d'initiatives ou de réflexions dans ce domaine.

Selon une enquête plus récente, réalisée par Steria à l'été 2013, 42 % des entreprises françaises considèrent le *big data* comme susceptible d'améliorer leurs processus de planification et de prévisions et 34 % comme susceptible d'accroître la connaissance et la maîtrise de leur activité. Toutefois, seules 7 % des personnes interrogées dans le cadre de cette enquête (au niveau européen cette fois-ci) considèrent le *big data* analytique comme un sujet très important pour elles.

Pour tirer profit de la démarche *big data*, le *top management* doit comprendre et adhérer à cette révolution. Les décideurs et les managers doivent définir comment en tirer parti puis adapter les outils et processus *big data* à leurs ambitions.

Les défis pour les organisations consistent à :

- Considérer tous types de données, même celles dont on ne voit pas encore l'utilité ou dont tous les paramètres ne sont pas encore maîtrisés.
- Stocker et sécuriser de gros volumes de données, et ce dès maintenant.
- Analyser et donner du sens à ces données.
- Exploiter et valoriser commercialement les nouvelles informations tirées de ces données.

« En termes d'acquisition et de conservation des données, nous sommes au point ; en revanche, la valorisation de ces données ainsi que leur circulation dans le groupe restent à faire. »

Directeur innovation d'un groupe du CAC40

Le *big data*, un outil au service de la « transformation digitale » pour la Société Générale

Au sein de la Société Générale, le *big data* est considéré comme une source d'opportunités pour la lutte antifraude et l'amélioration de l'expérience client. Selon Françoise Mercadal-Delasalles, directrice des ressources et de l'innovation à la Société Générale, « *l'exploitation des données comptables et commerciales recèle un potentiel de croissance colossal* ».

La démarche *big data* comporte deux avantages majeurs pour l'institution bancaire : faciliter la détection des fraudes et être un levier de lutte contre la cybercriminalité. Les données de l'activité bancaire d'un client pourraient par exemple être croisées avec ses données personnelles communiquées lors de l'ouverture du compte pour prévenir une utilisation frauduleuse de ses coordonnées bancaires.

Surtout, le croisement de ces mêmes données permet d'affiner l'analyse comportementale des clients, en y intégrant une dimension prédictive, susceptible à terme de transformer en profondeur la relation des établissements bancaires avec ces derniers. Fin 2013, le groupe avait lancé deux projets pilotes d'analyse prédictive sur son réseau en France.

Pour autant, le but de la Société Générale n'est pas de commercialiser les données. Une telle activité risquerait de nuire à ses deux atouts majeurs face à ses nouveaux concurrents proposant des services financiers exclusivement en ligne : la sécurité et la confiance. Pour la directrice des ressources et de l'innovation à la Société Générale, « *les banques seront toujours plus prudentes que les autres prestataires de services dans l'utilisation des données de leurs clients* ».

Les freins de mise en œuvre de solutions intégrant les *big data* sont nombreux dans un groupe tel que celui de la Société Générale :

* **des obstacles culturels** : les dirigeants actuels ne font pas partie de la génération des hyperconnectés et ne sont pas forcément convaincus des bénéfices potentiels du numérique ;

* **des difficultés de coordination** : les données collectées par le groupe ne sont pas ou très peu partagées à l'heure actuelle entre la banque d'affaires et la banque de détail, tout comme entre le siège français et les filiales étrangères ;

* **des difficultés informatiques, qui découlent souvent des précédentes** : la Société Générale ne dispose pas de plate-forme de gestion de la relation client (*Customer Relationship Management*) ni de référentiel unique. Des efforts sont cependant en cours sur ce point ; la rationalisation des référentiels d'organisation et de structure devrait permettre l'analyse et la valorisation des données.

Les banques sont désormais engagées dans une course au premier entrant en matière de *big data*. Si Wells Fargo avec sa banque virtuelle et HSBC font pour l'heure figure de précurseurs, la Société Générale a de vrais atouts à faire valoir. Un poste de *chief data officer* a été créé et l'entreprise peut s'appuyer sur ses compétences internes dans le domaine de l'informatique. En effet, ce dernier représente 11 000 à 12 000 collaborateurs en interne et 3 milliards d'euros de budget annuel. Cela lui a permis de conserver de précieuses compétences sur la collecte et l'analyse des données, contrairement à certains de ses concurrents qui ont pu pousser trop loin l'externalisation et perdre ce savoir-faire.

LE BIG DATA COMME NOUVEAU MARCHÉ

De nouvelles entreprises ou *business units* peuvent être créées qui s'appuient sur le *big data* :

- Le *big data* au sens large : fabricants de capteurs¹¹, services de location d'espace de stockage et de capacité de calcul, développeurs d'outils analytiques¹², de plates-formes de mise à disposition des données et de services de visualisation des données...
- Le stockage de la donnée.

11. Harbor Research estime que le marché mondial des *Internet enabled devices* sera de 10 milliards de dollars en 2014, contre 4,3 milliards de dollars en 2009.

12. D'après IDC, le marché mondial des programmes analytiques passera de 25,5 milliards de dollars en 2010 à 34 milliards de dollars en 2014.

Les modèles économiques de valorisation des données sont pour l'instant embryonnaires et restent difficiles à maîtriser pour ceux qui débutent sur ce marché. Deux types d'entreprises se distinguent :

- Les entreprises qui analysent les données.
- La société Gartner prédit l'émergence de « revendeurs d'informations » spécialisés, à qui la distribution pourra être sous-traitée, et de « gestionnaires de produit information » (internes) pour piloter cette nouvelle activité¹³. En effet, alors que les organisations produisaient et utilisaient jusqu'à maintenant leurs propres données, des *data brokers* revendent aujourd'hui les données d'entreprises ou encore de l'État à divers acteurs¹⁴. On estime ainsi que la société américaine Acxiom, spécialisée dans le recueil et la vente d'informations, et qui a dégagé un revenu de 1,15 milliard de dollars en 2012, posséderait en moyenne 1 500 données sur 700 millions d'individus dans le monde¹⁵.

Exemple : Ixxi, filiale du groupe RATP en charge de l'information voyageurs et des billets, a développé l'application Android « Your Tube Summer » lors des Jeux olympiques de Londres en 2012 pour que les touristes français puissent anticiper, préparer et organiser leur voyage dans la capitale londonienne, et ce grâce aux données librement partagées par le gouvernement britannique. Cette application propose de visualiser les stations situées autour de la position de l'utilisateur et les prochains départs des trains en temps réel, d'afficher les temps d'attente par station, ou encore d'effectuer des recherches par station, par lieu touristique et par manifestation sportive.

13. Voir « Gartner : vers la monétisation des *big data* », *C'est pas mon idée !*, blog de Patrice Bernard, 11 janvier 2013 (<http://cestpasmonidee.blogspot.fr/2013/01/gartner-vers-la-monetisation-des-big-html>).

14. Dans un rapport publié en 2012, la Commission fédérale américaine du commerce s'est souciée de l'essor de la profession d'*information broker*. Elle demande entre autres à ce que les citoyens puissent avoir accès aux informations que ces « vendeurs d'informations » ont sur eux (Federal Trade Commission, *Protecting Consumer Privacy in an Era of Rapid Change*, mars 2012).

15 Voir « *Data brokers* : aux États-Unis, votre vie privée est en vente », ZDNet.fr, 12 avril 2013 (<http://www.zdnet.fr/actualites/data-brokers-aux-etats-unis-votre-vie-privee-est-en-vente-39789295.htm>).

LES PRINCIPAUX POINTS DE VIGILANCE

Comme tout grand sujet de changement dans les organisations, les enjeux organisationnels et culturels associés au *big data* nécessitent **le soutien et l'implication de la direction générale**. Selon le *Big Data Index France*, 73 % des répondants n'associent pas de bénéfices au *big data* et 77 % estiment que ce n'est pas vraiment un sujet pour leur entreprise au-delà de nouveaux investissements et de nouvelles politiques de stockage et d'archivage. Or, nous avons vu que l'intérêt du *big data* est avant tout de croiser des données hétérogènes (météo, logistique, géolocalisation, trafic automobile) pour améliorer le rendement des activités existantes et en tirer des offres inédites. Les entreprises souhaitant prendre le virage stratégique du *big data* doivent combattre les doutes et les craintes associés à une telle démarche, assouplir les exigences de retour sur investissement en la matière pour permettre l'amorçage du changement et accorder les ressources nécessaires – ressources humaines aussi bien que financières – à un tournant d'une telle envergure.

Le buzz du *big data* étant largement centré sur les aspects technologiques, l'enjeu de **mise à niveau des compétences** est souvent sous-estimé, particulièrement en France. 60 % des répondants à une étude mondiale réalisée au printemps 2012 sur les enjeux du *big data* identifient le déficit de compétences comme le défi principal, tandis que 6 % seulement des répondants au *Big Data Index France* mettent en avant la transformation des compétences comme un impact très fort du *big data*. Les compétences pertinentes à l'ère du *big data* sont variées : ce sont des compétences techniques, statistiques, analytiques et en gestion de projets, afin de développer, déployer et/ou maintenir les technologies et logiciels associés au *big data*. Michel Morvan, ancien VP Strategic Intelligence and Innovation de Veolia explique que « *pour faire réellement parler les données, il faut des compétences spécifiques pour les nettoyer, établir des statistiques, modéliser* ». Mais il faut aussi des compétences d'analystes métier, connaissant parfaitement le modèle économique de l'entreprise et son environnement, pour être à même d'interpréter les résultats. De nouveaux métiers hybrides apparaissent, comme les *creative technologists*, mélangeant technologie, créativité, marketing et communication, pour faire le lien entre les différents services d'une entreprise.

Les entreprises seront probablement amenées à revoir leur organisation pour donner au *big data* la place stratégique qu'il requiert. De nouveaux métiers pour-

raient prendre de l'ampleur comme celui de *chief data officer* ou *chief information officer*. La question de l'environnement informatique ne sera pas non plus à occulter afin de se préparer aux exigences des grands volumes de données et de l'ère du *cloud computing*. Le rôle du directeur des systèmes d'information (DSI) sera alors amené à être revu et repensé, comme le préconise Gilles Grapinet : « la " commoditisation " du DSI ces dernières années et la pression sur les coûts ne doivent pas l'empêcher d'être un acteur central de la démarche big data ».

Les Data Scientists dans l'entreprise, vus par l'Ensaë

« Comment récolter les fruits du *big data* en minimisant les risques ?

- Pensez *data science* avant de penser *big data*.
- Embauche une équipe de *data scientists*.
- Placez-les sous la responsabilité d'une autorité légitime. »

Le *big data* n'est qu'un élément d'un nouvel ensemble d'outils et de techniques appelé *data science*. Une compréhension globale des données d'une entreprise et de leur potentiel – ainsi que la maîtrise des moyens d'exploitation – peut devenir un nouveau vecteur de performance.

La *data science* permet d'effectuer une analyse fine de l'information qui permettrait d'être plus efficace dans les processus (conception, marketing, relation client), d'identifier les données disponibles et de distinguer les données utiles, présentes ou non. Elle permet l'étude des méthodes de traitement de l'information les plus efficaces pour aboutir à des réponses concrètes et opérationnelles aux questions posées par les pilotes des processus.

À ce titre, le *data Scientist* doit avoir des compétences en mathématiques : l'analyse des données fait appel à des modèles statistiques ou des algorithmes complexes. Il doit aussi être capable d'extraire les données des systèmes informatiques, de programmer les algorithmes de traitement et de permettre la visualisation rapide des résultats, ce qui requiert des compé-

tences informatiques. Enfin, il ne peut se résumer à une simple fonction technique : au service de l'efficacité et de la rentabilité, il doit avoir des compétences d'expert métier et connaître les enjeux économiques de ses travaux.

Si certaines entreprises explorent déjà les capacités du *big data* par des initiatives cloisonnées des départements informatiques, marketing, financiers, la création d'une équipe transversale de *data scientists* permet une démarche globale de valorisation des données. Face au risque de résistance des dirigeants de l'entreprise, confrontés à des mécanismes qu'ils ne peuvent entièrement maîtriser, un tel projet devrait être mis entre les mains d'une autorité explicitement reconnue par la hiérarchie de l'entreprise, par exemple en le rattachant à la direction générale. C'est l'esprit de la création du rôle de *chief data officer* aux États-Unis¹⁶.

Enfin, les enjeux liés à la **sécurité**¹⁷, à la **disponibilité**, à l'**accessibilité** et à la **confidentialité** des données seront également de grands défis pour les organisations. La protection de la propriété intellectuelle ou encore la protection du consommateur constituent autant de zones d'évolution du droit, de la réglementation et de la norme dans un scénario de croissance continue du *big data*.

Des efforts de normalisation seront nécessaires pour permettre le décloisonnement des données en toute sécurité dans les entreprises et les organismes publics. Pour encourager les échanges, des normes de sécurité des échanges sont progressivement mises en place entre les administrations et les usagers, comme le Référentiel général de sécurité (RGS) ou la Plate-forme d'échange de confiance (PEC). Des outils comme le Référentiel général d'interopérabilité (RGI), qui fixe les règles techniques permettant d'assurer l'interopérabilité des systèmes d'information, encouragent aussi le partage. Depuis 2010, le Centre d'accès sécurisé distant (CASD) donne accès aux chercheurs, de façon très

16. Source : *Variance*, n° 46, février 2013 (http://www.ensae.org/docs/2013162028_var46_web-planches_31-01.pdf).

17. Dans son rapport de novembre 2013, *Analyse des big data. Quels usages, quels défis ?*, le CGSP rappelle ainsi que la plupart des fournisseurs des technologies du *cloud* se ménagent des *back doors* qui leurs permettent l'accès aux données stockées sur leurs serveurs.

encadrée, à des données individuelles possédées par l'Insee et les services statistiques ministériels.

Les craintes sur le non-respect de la vie privée constituent un autre obstacle important au développement du *big data* que les pouvoirs publics ne peuvent ignorer. Si le *big data* permet de mieux prévoir le futur proche pour y adapter au mieux les politiques publiques et optimiser la dépense publique, comme le fait par exemple le programme Global Pulse développé par les Nations unies¹⁸ en analysant les mots-clés échangés sur Twitter, certains redoutent le phénomène « *Big Brother* » de la surveillance à outrance. Certains experts estiment qu'à l'heure des *big data*, « l'anonymat est devenu algorithmiquement impossible ». Certaines entreprises ont été contraintes de prévoir une offre alternative pour répondre aux craintes de certains consommateurs. La formule Navigo Découverte de la RATP (coûtant approximativement 5 euros) permet depuis quelques années au possesseur de la carte de ne pas être inscrit dans un fichier client, contrairement à la formule classique. Le forfait est chargé sur une carte à puce et accompagné d'une carte nominative cartonnée où le voyageur indique lui-même son nom, son prénom et colle sa photo d'identité.

Les risques croissants portés par le *big data* sur la vie privée créent également de nouvelles inégalités, entre ceux qui maîtrisent les outils informatiques – par exemple les *privacy enhancement technologies* (PET) –, ceux qui contrôlent ou non leur dévoilement sur les réseaux sociaux, ceux qui peuvent payer pour des services protégés, et les autres.

Ces enjeux sont d'autant plus sensibles que la plupart des fournisseurs de l'infrastructure physique du *cloud* et des services en ligne qui génèrent massivement des données d'utilisateur sont des entreprises privées américaines. Si la CNIL a engagé une procédure contre Google, dont les politiques de confidentialité ne correspondent pas aux normes françaises, son pouvoir punitif est limité à 150 000 euros. Face à ces enjeux, l'État français a financé, à hauteur de 150 millions d'euros, deux *cloud computing* nationaux sous la forme de partenariats public-privé : Numergy et Cloudwatt. Cependant, les moyens restent

18. Global Pulse permet d'anticiper et de suivre l'impact des crises socio-économiques locales et mondiales de façon rapide. C'est le cas par exemple pour les épidémies, qui sont ainsi détectées plus rapidement qu'avec les réseaux de veille sanitaire classiques. Global Pulse scrute également le chômage ou encore les crises alimentaires.

limités ; le programme de R&D Quaero, lancé en 2009 pour développer des outils de recherche d'information en ligne avec un financement public-privé d'environ 200 millions d'euros, n'a pas permis le lancement d'un moteur de recherche français indépendant.

Au-delà, les débats sur la **justice prédictive** (cf. exemple de PredPol en page 26) font redouter à certains que demain des personnes puissent être punies pour leurs penchants, leurs relations, plutôt que pour leurs actions, ce qui remettrait en question un des principes fondateurs de la justice, celui de la présomption d'innocence. Cela amènera probablement les décideurs publics à redéfinir la notion de justice pour garantir la liberté de l'homme à agir. Plus largement, l'utilisation croissante d'algorithmes pour déterminer les préférences et prédire les comportements soulève la question de l'influence des données sur la liberté de choix, et l'**importance croissante accordée aux données** et aux corrélations qu'elles révèlent sans qu'une interprétation, un modèle humain y soient associés. En effet, la démarche *big data* inverse les démarches traditionnelles d'un modèle confirmé par les données, lui substituant l'observation de corrélations auxquelles peuvent être apportés des éléments explicatifs. Or, corrélation n'est pas interprétation, et encore moins science.

La restitution des données personnelles aux États-Unis et au Royaume-Uni

Dans plusieurs pays, le gouvernement engage des démarches pour redonner aux citoyens l'accès aux données personnelles que les entreprises stockent sur eux. Cette initiative peut avoir plusieurs avantages : créer des opportunités de dialogue et augmenter la confiance entre les entreprises et les individus ; faire apparaître de nouveaux services d'exploitation de ces données pour permettre aux particuliers d'optimiser leur consommation.

MiData

C'est le cas du programme MiData, lancé en novembre 2011 par le gouvernement britannique, qui rassemble de grands acteurs privés ainsi que des associations de consommateurs, sur la base du volontariat. Il réunit actuel-

lement plus de vingt-cinq entreprises, dont BarclayCard, MasterCard, HSBC, British Gas, Scottish Power, Google, EDF Energy.

Après une phase de réflexion, le midata Innovation Lab (mIL) a été lancé le 4 juillet 2013. Il propose aux particuliers de stocker volontairement sur le site les données que les entreprises possèdent sur eux, de manière sécurisée. Ces données permettent à des développeurs de réfléchir à des moyens innovants de les exploiter, ces derniers étant ensuite disponibles sur le site – par exemple, aider le consommateur à choisir son forfait de téléphone, à organiser ses courses. Ce projet devrait permettre une réflexion très concrète sur les moyens de mettre à disposition les données tout en les protégeant.

En France, le projet MesInfos, lancé en janvier 2012 par la Fondation Internet nouvelle génération (Fing), procède de même. L'expérimentation a été lancée en 2013 et regroupe 300 testeurs volontaires.

Blue Button

La création d'une plate-forme sécurisée pour stocker des données personnelles n'est pas nouvelle : aux États-Unis, l'initiative Blue Button du gouvernement permet depuis fin 2010 aux bénéficiaires de Medicare et aux anciens combattants de consulter en ligne et de télécharger leur dossier médical. Ils peuvent ensuite le transmettre aux divers organismes de santé. Des développeurs en ont profité pour proposer des services utilisant ces données, comme une application de Northrop Grumman qui permet aux anciens combattants de recevoir des conseils de santé adaptés à leur dossier sur leur téléphone. De nombreux organismes de santé se sont ensuite engagés à fournir la même offre à leurs clients.

L'initiative Green Button a quant à elle été lancée en janvier 2012 à l'initiative de fournisseurs (eau, électricité, gaz). Elle permet aux consommateurs de télécharger leurs informations de consommation sous un format standardisé, afin de les étudier et de les optimiser à l'aide d'applications. Trente-cinq entreprises proposent actuellement ce service.

LES MUTATIONS À PRÉVOIR

L'Open Innovation. Le *big data* s'inscrit dans une démarche d'innovation ouverte qui incite les organisations à collaborer avec d'autres acteurs en dehors des frontières traditionnelles de la firme sur des sujets stratégiques ou inattendus (*open innovation outside-in*) et à réfléchir à la valorisation des projets et des données produites sur lesquels l'entreprise n'envisage pas de capitaliser directement dans son cœur d'activité (*open innovation inside-out*).

Les spécificités de l'open data dans les entreprises

Si de nouvelles entreprises apparaissent pour exploiter les données mises à disposition par le gouvernement, pour l'instant peu d'entreprises ont adopté elles-mêmes le mouvement de l'*open data* en publiant certaines de leurs données en accès libre. Pourtant, elles sont productrices de données par leur activité, et souvent consommatrices de données externes qui leur permettent d'analyser l'économie ou d'échanger avec leurs partenaires (facturation, prospection).

Pour elles, l'ouverture des données est plus dépendante des opportunités en termes commerciaux ou d'innovation que d'enjeux de transparence, bien que cela leur permette aussi d'améliorer leur image.

L'ouverture des données peut avoir les avantages suivants pour les entreprises : à destination des collaborateurs et des individus, elle peut accélérer le développement de produits et stimuler l'innovation en créant un écosystème dynamique de partenaires tout en améliorant la satisfaction client ; à destination de concurrents, elle permet par exemple de mieux comprendre les évolutions d'un marché.

Cependant, il est difficile d'évaluer les opportunités commerciales directes ou indirectes de cette démarche, ce qui ralentit les investissements privés dans la valorisation des données. À ce frein s'ajoutent l'absence de cadre juridique de référence qui établirait les engagements de l'entreprise envers les utilisateurs de

ses données (disponibilité, qualité, sécurité) et une crainte face à l'absence de contrôle quant à la réutilisation. De plus, il n'existe pas non plus de régime juridique unifié qui définisse la propriété de ces données « privées » et, selon les contextes, le réutilisateur doit aller vérifier dans des régimes spécifiques s'il existe des clauses concernant la diffusion et la réutilisation des informations¹⁹.

Enfin, pour que la démarche soit un succès, elle demande un réel investissement de l'entreprise. En effet, il n'est pas évident d'assurer la qualité des données mises à disposition : celles-ci doivent être complètes, non altérées, accessibles à tous aussi rapidement que possible (en temps réel à l'idéal) et exploitables facilement.

Exemple : le groupe Poulit prend part à un projet d'échange de données de marché avec d'autres producteurs de produits alimentaires concurrents. Ce système leur permet d'obtenir des informations de marché très précises, tout en restant semi-ouvert en étant restreint à un certain nombre de partenaires choisis. L'échange peut passer par un tiers qui regroupe les données pour plus de sécurité.

Exemple : PSA a développé le service d'applications d'infodivertissement embarquées « Peugeot Connect Apps ». Les données spécifiques du véhicule ont été mises à disposition pour que des tiers puissent aussi développer des applications. Parrot, Ford et General Motors proposent des services similaires²⁰.

Exemple : les badges de télépéage commercialisés par Vinci Autoroutes contiennent des morceaux de données sur les automobilistes, la fréquence de leurs trajets ou la catégorie de leur véhicule. Ces informations sont une mine pour les partenaires commerciaux des sociétés d'autoroutes, comme les stations-service, restaurants et hôtels des aires d'autoroutes, et les villes et villages aux abords des sorties. L'utilisation de ces informations pose évidemment le problème de la protection des données personnelles, mais aussi surtout celui de la mise en place de partenariats intelligents et gagnants-gagnants.

19. Dossier de synthèse de la journée d'étude du GFII, « *big data*, exploiter de grands volumes de données : quels enjeux pour les acteurs du marché de l'information et de la connaissance ? », 3 juillet 2012.

20. Source : « *Open Data*, quels enjeux et opportunités pour l'entreprise ? », Bluenove, novembre 2011

« *Open Innovation is the formal discipline and practice of leveraging the discoveries of unobvious others as input for the innovation process through formal and informal relationships.* »

Professeur Frank Piller, MIT smart customization & RWTH,
Aix-la-Chapelle

- La démocratisation du *yield management*, soit l'analyse en temps réel de la demande pour moduler des prix algorithmiques, ajuster les promotions et les approvisionnements en fonction des évolutions (notamment dans les secteurs du *retail* et du tourisme). Le *big data*, via des étiquettes électroniques, permet par exemple de faire varier les prix en quelques minutes en fonction de la taille, de la couleur, de la localisation, de la météo, des stocks, etc. Aux magasins d'adapter en temps réel leur gestion des approvisionnements et des démarques ! Les retombées du *big data* sur les activités de *pricing* peuvent être à l'origine de nouveaux modèles économiques.

L'automatisation de la prise de décision

L'une des richesses du *big data* est sa granularité : si le rendement d'une information supplémentaire est souvent rapidement décroissant²¹, posséder un grand nombre de données permet une analyse plus fine des sous-ensembles. Ainsi, le nombre permet de calculer un plus grand nombre de modèles.

Le *big data* permet donc d'appréhender les hétérogénéités des ensembles, et d'ajuster l'analyse en fonction des données en temps réel pour permettre une réponse plus appropriée. L'augmentation du nombre des données analysées dans des contextes très spécifiques sous-tend une prochaine étape dans la gestion des décisions, l'évolution d'une vision stratégique d'aide à la prise de décision vers une vision tactique, l'automatisation des microdécisions en temps réel. Cette automatisation nécessitera le développement d'une plate-forme décisionnelle, c'est-à-dire une capacité de production de règles de décision, des moteurs d'inférence, des simulateurs, en supplé-

21. Les réponses à une question oui/non d'un échantillon non biaisé de 1 000 personnes suivent déjà une distribution à faible variance, c'est-à-dire que si l'on interrogeait un plus grand nombre de personnes, on ne gagnerait quasiment pas en précision de réponses.

ment de la plate-forme de collecte et de stockage des informations. Si ces microdécisions permettent des gains très faibles à l'unité, l'accumulation de ces gains peut justifier un tel investissement.

L'automatisation des décisions et la disparition de certaines tâches routinières ne sont pas sans risques. Sans se substituer à la stratégie de moyen et long terme, les *big data* sont surtout utiles pour améliorer des microdécisions opérationnelles. De plus, des mesures de sécurité pourront être mises en place, comme la possibilité de reprendre le contrôle des décisions à tout moment²².

« *Managing treasury in the cloud* », nouveau *business model* permis par les *big data*²³

Le développement des *big data* s'accompagne de nouvelles opportunités de modèles d'affaires que certains ont su saisir à l'instar de Kyriba. Créée en 2000, cette entreprise commercialise des solutions d'optimisation de gestion de la trésorerie (*cash management*). Kyriba travaille avec plus de la moitié des entreprises du CAC 40 et emploie 200 personnes dans le monde. Sa technologie repose sur un hub de communication bancaire qui intègre et uniformise les informations recueillies quotidiennement auprès des banques. Intégrant les créances et les diverses opérations comptables, l'entreprise fournit un monitoring global mais aussi des prévisions de liquidité et des flux financiers. Aujourd'hui, Kyriba s'affirme comme un acteur incontournable du marché des services applicatifs Internet de gestion de trésorerie à la demande. En 2012, la société réalise 35 milliards de dollars de chiffre d'affaires grâce à ses 700 clients groupes, ce qui représente plus de 30 000 entités servies.

Le *big data* est au cœur de son activité : Kyriba gère en effet un flux de 15 000 envois par seconde quand Twitter ne comptabilise « que » 5 000 tweets par seconde. Alexandre Grux, responsable de la recherche et de l'innovation de Kyriba, explique que la société met à profit le *big data* en utilisant à

22. Sources : « *big data*, small decisions and smart organization », Christophe Benavent, septembre 2012 ; « *big data* – un œil sur les organisations », Christophe Benavent, mai 2011.

23. Ce cas a été recueilli avec la collaboration du cabinet Oliver Wyman.

la fois le *data mining*, c'est-à-dire la collecte et l'exploration de données, et le *cloud computing*, permettant de grandes capacités de stockage et de calcul, à prix abordables, pour traiter les données recueillies en temps réel.

Description du service

Kyriba propose un outil de gestion de trésorerie. L'entreprise donne à ses clients une image consolidée de leur situation bancaire et de leurs liquidités en temps réel. Il peut être en effet complexe pour une entreprise multinationale, qui a des participations dans d'autres sociétés et parfois des milliers de comptes différents dans plusieurs devises et institutions bancaires, d'avoir une vision globale de sa situation bancaire. Et ce, plus encore en temps réel ! Alexandre Grux témoigne : « *Avant, pour avoir une image consolidée de tous les comptes, les trésoriers devaient remettre à zéro les comptes tous les jours et faire converger les cash flows sur un seul compte. Ce qui revenait à faire un reporting qui pouvait prendre une semaine voire plus selon la complexité des comptes.* »

Kyriba a su répondre à cette problématique en s'appuyant sur la technologie du *cloud computing* pour automatiser la collecte des relevés de comptes auprès des différentes banques de son client et leur consolidation. Kyriba communique ensuite l'image consolidée de l'état de trésorerie à son client avec différents niveaux de droit. De plus, ces informations sont disponibles en plusieurs devises dont les cours sont mis à jour en temps réel.

Parallèlement à cette offre, Kyriba a développé un panel important de services qui vont du pilotage des flux bancaires et de la liquidité de l'entreprise à des outils d'aide à la décision en passant par le contrôle des conditions bancaires. Ces services sont accessibles via un logiciel en ligne, qui n'implique aucun déploiement et ne nécessite qu'une simple connexion internet. Ce format permet à un collaborateur d'utiliser les solutions Kyriba à distance, au cours d'un voyage d'affaires par exemple.

En plus de la visualisation de leur situation bancaire, l'intérêt des données collectées par Kyriba pour les entreprises clientes est double.

1) Tout d'abord, ces données ouvrent de nouvelles perspectives pour les prévisions de flux de trésorerie (*cash flow*). Les directeurs financiers et tré-

soriers, qui bénéficient désormais d'une image quotidienne et précise de leurs comptes, pourraient intégrer les données de la trésorerie dans l'élaboration d'un plan stratégique à trois ans aux côtés des données traditionnellement prises en compte comme les ventes.

2) Dans un second temps, les clients pourraient bénéficier des résultats des recherches et des analyses faites par Kyriba à partir des données collectées. En effet, grâce aux clauses de confidentialité, Kyriba peut exploiter les données des clients et mettre son expertise à leur service en leur proposant par exemple de nouveaux indicateurs clés de performance pour piloter leur activité.

Enjeux

Si Kyriba est désormais leader dans son domaine, les défis auxquels l'entreprise doit faire face sont nombreux. Ces derniers sont liés aux risques traditionnels de la démarche *big data* : réticences des clients, nécessaire acquisition de nouvelles compétences, rôle ambigu de l'État.

L'enjeu du *big data* est avant tout celui de la sécurisation et de la protection des données. Il s'agit donc d'identifier les risques liés aux données collectées et de mettre en place des solutions adaptées en fonction des supports utilisés (serveurs, ordinateurs personnels, mobiles, tablettes) ; leur démultiplication augmentant les risques d'infractions. Alexandre Grux explique que Kyriba possède ses propres serveurs de stockage et en assure la sécurité. Les données sont systématiquement dupliquées pour assurer une continuité d'activité en cas de panne.

Kyriba doit également faire face à certaines réticences de la part de ses clients. La société a mené une enquête auprès de ses clients qui révèle que l'appétence pour les solutions informatiques chez les trésoriers est faible. Nombre de clients voient cette utilisation des données d'un mauvais œil, même si celle-ci est faite dans le cadre d'un accord de confidentialité : « *Le client peut nous reprocher de faire du business avec ses données et cela même si on lui propose le service gratuitement* », raconte Alexandre Grux. Proposer aux clients des améliorations sur leur pilotage financier à partir de l'analyse de leurs données, comme un nouveau KPI, se révèle encore délicat

pour Kyriba. L'entreprise doit donc développer des solutions simples et intuitives pour pallier cette difficulté. Il faut informer les dirigeants d'entreprise et leur faire prendre conscience des bénéfices liés à l'ouverture de leurs données dans un cadre légal bien défini. Une évolution de la culture d'entreprise, traditionnellement attachée au culte du secret, est nécessaire.

Pour développer son offre, Kyriba doit recruter des collaborateurs compétents pour mettre en place des services liés aux *big data*. L'entreprise peine à trouver des profils à l'aise avec ces enjeux et capables de prendre des initiatives dans l'exploitation de données qui pourraient potentiellement générer de nouveaux services. Cette pénurie s'explique selon Alexandre Grux par une forte demande de telles compétences mais aussi par leur rareté. La formation universitaire doit donc évoluer pour s'adapter à cette nouvelle demande sur le marché de l'emploi.

Enfin, Kyriba se doit de composer avec l'État dont la position vis-à-vis du *big data* est ambiguë. D'une part, l'État a lancé Etalab, une mission qui vise à mettre à la disposition de tous des informations publiques, exception faite des données à caractère personnel protégées par un droit de propriété intellectuelle ou de celles mettant en jeu la sécurité publique. En combinant ces données publiques avec celles dont l'entreprise dispose, Kyriba pourrait imaginer d'autres modèles d'affaires. L'État est donc source de nouvelles opportunités. D'autre part, il tient un discours souvent alarmiste sur la protection des données. L'État n'a pas encore différencié de façon claire les données individus des données entreprises, ni statué sur leur protection et utilisation respective, ce qui constitue un frein au développement des sociétés dont l'activité est basée sur le *big data* telles que Kyriba.

- L'avènement du **développement durable** : le *big data* constitue une source privilégiée d'optimisation de l'exploitation et de la consommation des ressources limitées (cf. la note de l'Institut de l'entreprise sur l'efficacité énergétique), même si le *big data* est lui aussi un grand consommateur d'énergie.

Exemple : entre 20 % et 30 % de la consommation en eau est perdue, par fuite ou par gaspillage. L'introduction de capteurs embarqués dans des

compteurs intelligents (cf. *smart grids*), tant sur les réseaux qu'au niveau des foyers, permet ainsi de récolter l'information en permanence et d'être potentiellement alerté en temps réel lors d'une consommation qui dévie de la norme, par exemple au travers d'une application mobile.

Exemple : les *data centers* consomment énormément d'énergie pour refroidir leurs serveurs. Pour renverser cette logique, l'entreprise Qarnot Computing fabrique des radiateurs numériques qui contiennent des processeurs et se branchent simplement sur les prises internet. L'entreprise chauffe ainsi gratuitement des logements ou bureaux et vend la capacité de calcul nécessaire pour atteindre la température voulue par l'utilisateur. Une centaine de logements seront équipés de ces radiateurs de calcul d'ici fin 2013.

- L'essor des *smart objects* et des *smart services* : la mise en place de nouveaux produits, services et processus connectés, « intelligents ».
-

Exemple : Sanofi propose en France depuis 2011 un lecteur de glycémie connecté, l'iBGStar, qui permet à la personne souffrant de diabète de mesurer son taux de sucre dans le sang, de l'enregistrer et de communiquer ses données en temps réel à son médecin. Cette application s'inscrit dans un projet plus large de Sanofi, Diabeo, solution de télémédecine innovante pour accompagner les patients diabétiques et les professionnels de santé dans une prise en charge globale. Le traitement du diabète nécessite de la part du patient un ensemble d'aménagements qui vont de la connaissance de la maladie et de son traitement à des compétences d'auto-observation, d'autosurveillance et d'autoadaptation du traitement en fonction des circonstances mêmes de sa vie. Pour l'aider dans cet effort, une application simple à utiliser au quotidien (pas de calculs, pas de retranscription papier, nombre limité d'alertes au patient et au médecin) a été développée pour automatiser une partie de la prise de décision. En identifiant à distance les patients en difficulté, ceci permet aussi de mieux utiliser le « temps médecin » et de personnaliser la prise en charge (ex. : adapter la fréquence des consultations plutôt que de décréter obligatoire une visite de contrôle tous les six mois).

Exemple : la nouvelle proposition de valeur faite par Lectra, numéro un mondial des solutions technologiques intégrées (logiciels, équipements de CFAO et services associés), à ses clients (entreprises qui utilisent des textiles, du cuir, des tissus industriels ou des matériaux composites dans la réalisation de leurs produits), grâce aux *smart services*, est l'application du concept de *lean manufacturing* (amélioration continue et réduction du gaspillage), issu de l'industrie automobile japonaise, à la salle de découpe. Les machines intelligentes de Lectra permettent d'optimiser et de sécuriser l'ensemble des opérations de production liées à l'activité de découpe automatique, quels que soient les besoins de production (de masse, diversifiée, petites et moyennes séries), grâce à un système de diagnostic et d'analyse en temps réel de la consommation de matière. Les experts des *call centers* internationaux de Lectra peuvent intervenir à distance en cas de problème ou pour former et conseiller les utilisateurs chez les clients, qui réalisent des économies significatives en termes de délai de fabrication et de consommation de matière, accélérant ainsi la mise sur le marché de leurs produits.

Exemple : les vêtements « intelligents ». Le projet européen « EASY-IMP – Développement collaboratif de produits portatifs et intelligents dans le *cloud* » a été lancé à la mi-septembre au Centre allemand de recherche sur l'intelligence artificielle (DFKI) de Kaiserslautern (Rhénanie-Palatinat). Financé sur trois ans à hauteur de 4,4 millions d'euros, ce projet met en place une équipe interdisciplinaire d'experts de différents domaines : capteurs, informatique, programmation, biomécanique, rééducation et sport. Douze partenaires académiques et industriels prennent part au projet, notamment l'université Lumière Lyon-II pour la partie française. Un exemple typique d'application est le *smartphone*, qui propose une plate-forme à partir de laquelle différentes applications peuvent être personnalisées. Dans le cadre du projet, ce concept serait appliqué aux vêtements. Par l'intégration de capteurs, de nouveaux produits pourront être conçus. On peut imaginer la configuration d'un vêtement, grâce au *smartphone*, de manière à enregistrer les signaux vitaux d'un patient, pour son entraînement sportif ou dans le cadre d'une rééducation. Grâce à leurs capteurs, ces vêtements seraient ainsi capables de mesurer la respiration, le rythme cardiaque, mais aussi le nombre de calories brûlées, l'état émotionnel, ou encore, de manière plus classique, le nombre de pas effectués ou l'activité physique.

- D'autres tendances importantes se dessinent avec l'avènement du *big data*, comme la *big research*, qui risque de transformer les pratiques scientifiques toutes disciplines confondues, de la biologie à la sociologie, ou encore la *big creativity*, qui démultiplie les occasions de faire des ponts inattendus entre différentes idées, personnes ou faits, et dope ainsi la capacité d'innovation des organisations.

« Les nouvelles technologies nous ont condamnés à devenir intelligents et imaginatifs car c'est tout ce qui nous restera. »

Michel Serres

Le big data, la révolution de l'assurance

Le principe des assurances est de partager un risque à l'intérieur d'un groupe. Son fonctionnement est assuré par des systèmes probabilistes, qui permettent d'affecter un prix « juste » aux risques assumés. Cependant, ce principe ne peut fonctionner qu'avec une évaluation imparfaite du risque, tout au moins à l'échelle individuelle : les mouvements d'ensemble doivent être aussi bien évalués que possible, mais les risques individuels doivent rester flous, sans quoi l'individu n'a plus intérêt à la mutualisation.

Impact sur l'efficacité commerciale

La collecte et le traitement de données de plus en plus nombreuses permettent aux compagnies d'assurances d'améliorer leur efficacité commerciale, en assurant :

- De nouvelles méthodes de segmentation, et donc de tarification.
- Des créations de produits plus efficaces (meilleure étude de marché, tarification).
- Une connaissance précise des habitudes des clients, d'où une meilleure gestion des campagnes publicitaires ou de prévention et du service personnalisé au client.
- Une limitation du risque de fraude à l'assurance.
- Une meilleure gestion des sinistres.

L'impact des données individualisées sur l'actuariat

Pour certains, les nouvelles connaissances issues du *big data* affectent l'économie des assurances et mettent en péril son fonctionnement même : trop d'informations conduit à une sélection forte et peut détruire la mutualisation. Néanmoins, d'autres opinions moins pessimistes ne voient dans le *big data* qu'un nouvel outil puissant pour les actuaires, du moment qu'il respecte les normes éthiques sur l'anonymat des données : l'anonymisation des données sera un enjeu d'équilibre. Le *big data* pourrait aussi améliorer la qualité de l'actuariat : l'actuaire serait « un certificateur de la qualité, de la pertinence des données, et un vérificateur de leur usage²⁴ », et l'augmentation de la démutualisation pourrait supprimer les « bon risques », assurés pour l'instant à prix élevés.

Les réglementations dans l'usage du big data

Les systèmes d'assurance sont soumis à de nombreuses réglementations²⁵ ; l'utilisation des données personnelles est soumise au contrôle des individus et la segmentation des assurances est limitée par la protection contre les discriminations : ces considérations éthiques limitent l'utilisation des données personnelles et l'étendue de l'impact du *big data* sur les systèmes assurantiels²⁶.

De nouveaux acteurs dans la collecte de données

L'abondance et la qualité des données étant des facteurs de différenciation compétitive pour les assureurs, à l'ère du *big data*, un nouveau rapport de force s'établit entre les compagnies d'assurances et les fournisseurs de données. Ces derniers sont en position de pouvoir pour capter une part importante de la valeur des données ; ils pourraient aussi eux-mêmes devenir assureurs grâce à l'avantage compétitif que leur confère l'exclusivité de leurs données.

24. Patrick Tourot, François Ewald, « *big data*, défis et opportunités pour les assureurs », *Banque & Stratégie*, no 315, juin 2013.

25. Le droit du contrat d'assurance, le droit des sociétés et le droit de la santé et de la Sécurité sociale. L'utilisation des données est notamment soumise au respect des libertés publiques (CNIL), du droit pénal (les discriminations), du droit communautaire et des règles européennes et internationales (ONU).

26. « *big data*, défis et opportunités pour les assureurs », art. cit.

PISTES DE RECOMMANDATIONS

Afin de tirer tous les bénéfices de la révolution qu'est le *big data*, il est important que les pouvoirs publics, en partenariat avec les acteurs télécom, s'intéressent à l'attractivité numérique des territoires et cherchent à déployer l'**ultra-haut débit fixe et mobile**, en généralisant par exemple la 4G et en assurant une couverture Wi-Fi optimale, même dans les transports (train, avion, métro). Toutefois, les effets sur la santé des ondes électromagnétiques, bien que non avérés scientifiquement²⁷, restent un sujet de préoccupation pour une partie de l'opinion publique qui doit être pris en compte.

Les freins au développement du *big data* doivent être bien identifiés. Les freins inhérents au *big data* ont été cités plus haut : pour les entreprises, l'évaluation des apports du *big data* et le soutien de la direction générale, la nécessaire et coûteuse mise en place d'un département de *data scientists*, la mise à niveau des compétences, ainsi qu'une attention particulière portée à la sécurité, à la disponibilité, à l'accessibilité et à la confidentialité des données.

Certains de ces enjeux concernent aussi les pouvoirs publics, comme les risques pour la vie privée, la formation, l'établissement de normes et de standards, et enfin l'éventuelle utilisation à mauvais escient d'analyses prédictives.

Des freins plus génériques s'opposent aussi à certains développements : des freins liés aux spécificités des tissus économiques nationaux et à l'autonomie

27. L'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (Anses) a ainsi rendu public en octobre 2013 un avis sur les effets des ondes électromagnétiques sur la santé. À la lumière de plus de 300 études scientifiques internationales publiées depuis 2009, l'Anses ne conclut pas à un « effet sanitaire avéré chez l'homme » et ne propose donc pas de « nouvelles valeurs limites d'exposition de la population ». Elle constate toutefois certains effets biologiques chez l'homme et chez l'animal – cassures de l'ADN, stress oxydatif susceptible d'altérer les cellules –, modifications qui « semblent être rapidement réparées ».

de décision dont disposent les managers de terrain²⁸ ; des freins d'ordre réglementaire, ou inhérents aux politiques de protection des consommateurs²⁹. Des freins fiscaux peuvent aussi émerger : conscients de la valeur de leurs actifs, certains détenteurs de *big data* pourront décider de ne pas autoriser l'accès à leurs données à des tiers, limitant ainsi les possibilités de croisements qui sont le vecteur de la création de richesses. Enfin, la concurrence déloyale du secteur public, qui détient d'importants gisements de données et peut être tenté d'en restreindre l'accès, pourrait ralentir le développement de *start-up* de valorisation des données.

Propositions

Deux lignes directrices peuvent être distinguées concernant le développement du *big data*.

Premièrement, il importe de mettre le client ou l'utilisateur au centre des politiques de soutien public au *big data*. Le *big data* bénéficie en France d'une attention soutenue des pouvoirs publics. Le rapport Lauvergeon identifie ainsi le *big data* comme l'un des sept domaines d'innovation que la France se soit de soutenir, et c'est l'une des 34 filières de la Nouvelle France Industrielle. Par ailleurs, le ministère de l'Innovation et de l'Économie numérique a annoncé en juin 2013 un plan de soutien à la filière du *big data*, inspiré des recommandations de l'Afdel (association française des éditeurs de logiciels et solutions internet). Dans le cadre des investissements d'avenir, plusieurs projets traitant des *big data* ont aussi été sélectionnés pour recevoir des financements. Si cette attention est évidemment bienvenue, le risque est toutefois que les pouvoirs publics n'envisagent le déploiement du *big data* que sous le seul angle de la technologie, quant ce sont les usages qui importent réellement.

28. Cette autonomie varie fortement selon les entreprises. Or, pour que le *big data* prenne toute sa mesure, les managers de terrain doivent se voir accorder suffisamment d'autonomie. « Le *big data* nécessite donc le développement de nouvelles compétences, et sans doute plus encore un véritable *empowerment* des unités opérationnelles : plus de données, plus de contexte dans les modèles et les indicateurs, donc plus d'autonomie dans la décision » (http://technologies.lesechos.fr/business-intelligence/big-data-c-est-le-chef-de-rayon-qu-il-faut-former_a-41-506.html).

29. Le rapport de la mission gouvernementale concernant la fiscalité numérique a été rendu public le 18 janvier 2013. Il préconise de taxer les entreprises en fonction du volume de données personnelles qu'elles collectent et exploitent, ainsi que du respect des libertés individuelles, et notamment du degré d'accès des utilisateurs aux données collectées.

Par ailleurs, il est nécessaire d'envisager le potentiel du *big data* dans toutes les filières. Les initiatives gouvernementales font souvent l'erreur d'identifier le *big data* comme une filière en soi. Au contraire, le *big data* est source d'innovation dans tous les secteurs d'activité, et il faudrait l'intégrer systématiquement dans les politiques d'innovation.

Entreprises

a) Développer une culture de la donnée dans les entreprises

Le développement du *big data* nécessite le soutien de la direction générale, qui doit se traduire par la nécessaire et coûteuse mise en place d'un département de *data scientists* dépendant de la direction, mais aussi par la mise à niveau des compétences de l'ensemble des salariés.

Il importe pour l'entreprise de comprendre que les données ne sont pas seulement l'affaire des spécialistes, mais celle de tous, et que les pratiques de chaque salarié s'en trouveront modifiées. Pour développer une culture de la donnée dans l'entreprise, il faut donc s'interroger sur les moyens de capter, conserver et stocker les données, et développer ces moyens en conséquence.

b) Donner davantage d'autonomie de décision aux managers de terrain

Cette autonomie varie fortement selon les entreprises. Or pour que le *big data* prenne toute sa mesure, et que les données disponibles soient comprises mais aussi débouchent sur des actions, les managers de terrain doivent se voir accorder suffisamment d'autonomie.

Puissance publique

a) Repenser la protection des données personnelles

La loi informatique et libertés de 1978 protège des dangers liés à l'exploitation de données personnelles, mais toutes les nouvelles « *data breach* »

pourraient ne pas être toutes prises en compte à temps. Une politique de la donnée semble nécessaire aussi bien dans les entreprises qu'au niveau des pouvoirs publics. La propriété intellectuelle, les politiques de réutilisation des données, l'utilisation secondaire des données générant en général davantage de valeur que l'utilisation primaire, les droits d'accès et de rectification des utilisateurs, la durée de conservation des données, autant d'exemples de mesures auxquelles les responsables nationaux se doivent de réfléchir, tout en veillant à ne pas couper court à l'innovation promise par le *big data*.

Pour autant, la protection par la réglementation reste souvent perçue comme une contrainte, parfois à cause du retard du droit par rapport aux évolutions technologiques, tandis que la définition d'un levier punitif est malaisée puisque les données n'ont pas de valeur intrinsèque. De plus, certains prédisent des pressions croissantes, de la part de certaines entreprises et de la société, contre les mouvements de protection des données.

1. Instaurer un droit à l'expérimentation qui permette de déroger aux règles actuelles en matière de protection des données personnelles

Aujourd'hui, la France manque d'un cadre conceptuel pour faire face au développement du *big data*. Elle hérite d'une réglementation sur les données personnelles excessivement restrictive qui pénalise le déploiement de nouvelles applications.

Comme le suggère le rapport Lauvergeon, on pourrait imaginer la consécration d'un droit à l'expérimentation, sous l'égide « d'un observatoire des données ». Il s'agirait d'évaluer l'efficacité et le bien-fondé de certaines pratiques permettant l'exploitation des données, avant de réfléchir à un possible encadrement législatif.

2. Responsabiliser et encadrer

Afin de garantir la transparence de l'utilisation des données personnelles par les entreprises et les administrations, on pourrait former des tiers-certificateurs qui appliqueraient un système de labellisation,

imposer davantage de transparence sur les brèches de sécurité et soumettre entreprises et administrations à des « audits d'algorithmes ».

Le projet de règlement européen en cours d'adoption vise aussi à imposer diverses obligations aux responsables de traitement informatique, comme celle de prendre en compte la protection des données dès la conception des systèmes (*privacy by design*).

3. « *Empowerment* » et transparence

Afin de réduire l'asymétrie d'information entre les institutions et les individus, on pourrait créer une plate-forme d'accès individuel à ses données personnelles.

On pourrait par ailleurs développer un système d'assurance pour éviter et ensuite résoudre les problèmes d'utilisations abusives des données.

4. Adapter la réglementation et les normes

« À l'international, l'UE reste perçue comme un territoire offrant un haut niveau de protection pour les données », ce qui pour certains lui procurerait un « avantage compétitif ». ³⁰ Ainsi il faudra adapter la réglementation tout en conservant ce haut niveau de protection.

La définition de « donnée personnelle » nécessite d'être revue, l'anonymisation ³¹ des données rencontrant aujourd'hui ses limites. D'une part, en effet, l'anonymisation parfaite rend les données quasiment muettes, donc inintéressantes, d'où l'utilisation croissante de bases de données anonymisées relativement. D'autre part, de nouvelles technologies permettent aujourd'hui de réidentifier des données au départ anonymes en agrégeant plusieurs jeux de données. C'est ce que l'on nomme des *inference attacks*.

30. Compte-rendu: « Midis du GFI », Déjeuner-débat du 22 mai 2012 autour d'Isabelle Falque-Pierrotin, présidente de la CNIL

31. Selon la CNIL, l'anonymisation des données consiste dans la suppression de tout lien entre une donnée et une personne physique.

De nombreuses discussions sont aujourd'hui en cours au sein de l'UE pour définir un nouveau statut de la donnée qui soit intermédiaire entre la donnée totalement anonymisée et la donnée personnelle identifiante. Cela permettrait une réutilisation plus simple des données, tout en préservant leur lien avec une personne physique.

En parallèle, la CNIL envisage la définition de nouveaux droits³² : droit à la personnalité, droit à l'oubli³³, à la mémoire, à la portabilité des données, à un environnement non pollué par la publicité numérique, et au refus du profilage.

De plus, il faudrait définir des normes internationales, autant pour encadrer les échanges que pour s'accorder sur la gestion du réseau internet qui constitue un bien commun international, un « *global networked commons* »³⁴.

b) Ouvrir les données publiques pour permettre l'essor de nouveaux business models dans le domaine de la protection sociale

Le sujet de l'*open data*, lié à celui de la *big data*, concerne en premier lieu les administrations publiques (obligation légale³⁵) qui ouvrent les données qu'elles produisent dans le cadre de leur activité quotidienne (fonds de cartes, statistiques, mesures, horaires, dépenses). Lancée depuis les années 2000, cette dynamique de mise à disposition des données numériques accessibles en ligne, sous des formats ouverts et exploitables, accompagnées de licences, autorise la réutilisation des données par tous.

32. *Vie privée à l'horizon 2020*, 2012.

33. Une charte du droit à l'oubli numérique intitulée *Droit à l'oubli numérique dans les sites collaboratifs et les moteurs de recherche* a été signée par les représentants du secteur et des acteurs de la société civile sous l'égide du Secrétariat d'Etat à la prospective et au développement de l'économie numérique le 13 octobre 2010. Ni Google ni Facebook n'ont signé cette charte (<http://www.alain-bensoussan.com/avocats/charte-droit-a-loubli-numerique-dans-les-reseaux-sociaux/2010/10/30>).

34. "Data, data everywhere, a special report on managing information", *The Economist*, février 2010.

35. Pour les données publiques, le principe de « disponibilité » est inscrit dans la loi depuis 1978, auquel s'ajoute le principe de « réutilisation » (décret du 30 déc. 2005, en transposition de la Directive européenne 2003/98/CE) et le principe de « gratuité » (décret du 26 mai 2011 et circulaire du 26 mai 2011).

Pourtant, la France est à la traîne en matière d'ouverture des données publiques. Le rythme de diffusion de ces dernières dépend en effet entièrement des autorités publiques : il n'existe pas à l'heure actuelle de droit « opposable » à l'*open data*.

Mais la retranscription en 2015 dans le droit français de la directive européenne de 2003 sur les informations publiques devrait permettre de modifier cet état de fait. Ce texte réaffirmera un certain nombre de grands principes, notamment ceux de la gratuité des données et du droit à la réutilisation, étendant le périmètre à toutes les données publiques, et pas seulement à celles ayant déjà été publiées. La directive doit également évoquer un droit opposable à la publication, mais encore imparfait. Celui-ci obligerait les administrations à publier toutes les données demandées par les citoyens.

Axelle Lemaire, secrétaire d'Etat au numérique, a par ailleurs conclu les deux jours de débats de la Conférence de Paris sur l'*open data* et le gouvernement ouvert organisée par Etalab fin avril en faisant part de la volonté du gouvernement d'aller « *plus loin, en matière d'ouverture et d'accessibilité* », et en expliquant que la transposition de la directive offrait « *un cadre juridique et une opportunité majeure d'évolution* ». Elle a indiqué souhaiter que soit inscrite dans la loi, à cette occasion, le principe d'une ouverture par défaut des données publiques, et que toute fermeture soit explicitement expliquée, justifiée et réversible³⁶.

Cette ouverture serait particulièrement heureuse au vu des applications prometteuses liées au *big data* dans le domaine de la santé et de la protection sociale³⁷. L'industrie de la santé se fonde sur quatre sources de données : les données de R&D d'entreprises pharmaceutiques ou académiques (essais cliniques par exemple) ; les données cliniques d'hôpitaux (dossiers médicaux électroniques) ; les rapports d'activité et les données de coûts

³⁶ Ces déclarations rejoignent les propositions de l'Institut de l'entreprise

³⁷ Voir notamment le *Rapport sur la gouvernance et l'utilisation des données de santé*, octobre 2013. http://www.sante.gouv.fr/IMG/pdf/Gouvernance_et_utilisation_des_donnees_de_sante_septembre_2013.pdf. Le rapport Lauvergeon propose par ailleurs de lancer des programmes de valorisation par licence de cinq « stocks » de données massives dont l'analyse pourra apporter une plus-value à l'ensemble de la société : Pôle emploi, Sécurité sociale, éducation nationale, enseignement supérieur et aides à la valorisation du patrimoine touristique.

(données des caisses d'assurances maladies) ; le comportement des patients, données détenues par divers acteurs, en dehors de l'industrie de la santé, voire par le patient directement (préférences du consommateur, historique d'achats). Or les innovations issues de l'exploitation de ces données sont fondamentales. Elles permettraient d'évoluer vers une prévention plus ciblée, de mettre en place des soins appropriés tendant vers la médecine personnalisée, d'optimiser le médicament pour obtenir l'impact clinique attendu, de réduire les coûts pour une qualité de soin égale ou supérieure et enfin de favoriser la sécurité. Aux États-Unis, l'impact du *big data* sur la réduction des coûts de santé serait de l'ordre de 300 à 450 milliards de dollars (jusqu'à 17 % de baisse), principalement grâce à la prévention et à la médecine personnalisée.

Il subsiste toutefois des obstacles soit techniques, soit liés à la propriété des données. Ainsi, si les deux premières sources de données sont en passe d'être ouvertes du fait de la demande de la puissance publique, cette dernière ne semble pas encore disposée à ouvrir ses propres données et celles des patients. Or il semble difficile de demander de tels efforts à l'industrie pharmaceutique si l'État ne prend pas sa part dans la démarche.

c) Améliorer les procédures d'appels d'offres liés à une innovation dans les données

Il pourrait être pertinent de diviser les procédures d'appel d'offres en plusieurs marchés distincts de taille plus petite. Ainsi, les petites entreprises - qui participent souvent aux procédures d'appels d'offres en tant que sous-traitants de grandes entreprises - seraient moins dépendantes des grandes et pourraient participer aux programmes de marchés publics.

Ces appels d'offres pourraient aussi mettre davantage l'accent sur l'utilisation de certaines méthodes ou procédés. Il s'agirait d'obliger les grandes entreprises à adopter et inclure la contribution d'une petite entreprise innovante dans leurs propositions lors d'une procédure d'appel d'offres.

d) Développer des formations adaptées, en utilisant le potentiel démultipliateur de l'enseignement en ligne

Pour que les besoins des entreprises en spécialistes du *big data* soient comblés, il faut promouvoir l'enseignement supérieur en mathématiques, en technologie de l'information et en sciences. La formation continue doit également être encouragée. Cela permettrait aux individus de mettre à jour les compétences nécessaires pour l'analyse de données.

De nouvelles compétences et de nouveaux profils sont indispensables à l'essor du *big data*. Les politiques publiques de formation devront donc s'adapter aux nouveaux besoins en la matière et accélérer la diffusion de profils de type « *data scientists* », alliant compétences techniques et managériales. Ces compétences croisées sont encore très rares aujourd'hui alors qu'elles sont cruciales pour imaginer les nouveaux *business models* issus des *big data*. Henri Verdier expliquait néanmoins que la France disposait « *de très bons designers pour inventer de nouveaux modes de visualisation et [d'] une école mathématique de premier plan mondial* »³⁸.

Selon le ministère de l'innovation et de l'économie numérique, on estimait début 2014 à 300 000 le nombre de *data scientists* nécessaires à l'Europe dans les années à venir. Cela implique de mettre en place dès aujourd'hui des filières de formation adaptées à ces profils. À ce jour, il existe encore peu de formations françaises en ce domaine : le Mastère Spécialisé « *big data* : gestion et analyse des données massives (BGD) » de Telecom Paris-Tech, a ouvert à la rentrée 2013, un Mastère Spécialisé en *big data* a été lancé par l'Ensimag (Grenoble INP) et l'EMSI Grenoble (GEM) pour la rentrée 2014. Par comparaison, plus d'une vingtaine d'universités américaines ont lancé ou devraient lancer des formations *big data*. L'Université Columbia (New York) a ainsi créé au sein de l'IDSE (Institute for Data Sciences and Engineering) un nouveau diplôme intitulé « *Certification of Professional achievement in Data Sciences* ». Ce programme, ouvert dès l'automne 2013, propose 4 matières principales : probabilités et statistiques, algorithmes pour *big data*, *machine learning* et exploration des données. L'Université de

38. «Quand 'big data' rime avec 'big business'», *Les Echos* du 20 Mars 2012.

Stanford délivre quant à elle un cours en ligne depuis 2013 orienté vers les *big data* : « *Mining Massive Data Sets* ».

Pour toucher davantage de personnes, on pourrait envisager de développer un MOOC français sur le sujet, sur la plateforme FUN lancée par le ministère de l'enseignement supérieur et de la recherche. Ce MOOC pourrait être sponsorisé par des entreprises.

Cette formation pourrait valoir aussi pour le service public lui-même, où la présence de *data scientists* aiderait à valoriser les données publiques.

e) Diffuser une « culture de la donnée »

Pour ce faire, la culture de la donnée doit être diffusée dès la période de formation. Le renforcement d'une « culture quantitative » serait nécessaire dans les programmes de formation commerciale et marketing, afin que les futurs managers sachent interpréter les résultats des modélisations et comprendre le sens caché des données. La CNIL est à l'origine depuis l'été 2013 d'un collectif visant à déclarer l'éducation au numérique cause nationale. La révolution *big data* implique selon elle que chacun devienne un véritable « citoyen numérique », à la fois acteur informé et responsable du numérique, capable d'exercer de manière effective ses droits et devoirs dans cet univers.

f) Repenser l'encadrement du risque

De nouvelles formes d'expertises et d'institutions seront probablement nécessaires pour encadrer la révolution du *big data*, dont l'ampleur dépasse la compréhension humaine ordinaire et heurte notre besoin d'« explicabilité » selon le terme utilisé dans le domaine de l'intelligence artificielle.

Alors qu'il permet de mieux tracer et cibler les usagers et consommateurs, le *big data* peut être perçu comme une « boîte noire » peu exploitable et inspirant la méfiance. Mayer-Schönberger et Cukier³⁹ en appellent ainsi dans leur livre *big data: A Revolution That Will Transform How We Live*,

39. <http://blogs.lesechos.fr/internetactu-net/comment-les-big-data-vont-transformer-notre-societe-a13014.html>

Work, and Think à la naissance d’“algorithmistes”, spécialistes capables de comprendre les données pour contrôler les entreprises de l’extérieur comme de l’intérieur – « *tout comme les entreprises disposent en interne de comptables et d’auditeurs externes pour surveiller leurs comptes.* » Car les auteurs rappellent que les chiffres ne sont pas infallibles, qu’ils peuvent être manipulés, mal analysés ou utilisés à tort. Il existe selon eux un risque de discrimination et de catégorisation des populations auquel les pouvoirs publics devraient prêter la plus grande attention pour préserver les principes du libre arbitre.

L'OBSERVATOIRE DE L'INNOVATION

LE PRÉSIDENT



Christophe de Maistre est Président de l'Observatoire de l'Innovation de l'Institut de l'entreprise.

Président de Siemens France, Christophe de Maistre est diplômé de l'école d'ingénieur Institut Supérieur de Mécanique de Paris. Il a obtenu un D.E.A. de l'Ecole Normale Supérieure de Cachan et un eMBA (Duke University).

Christophe de Maistre intègre le Groupe Siemens en 1991 où il occupe diverses fonctions marketing en Allemagne et en France avant d'être nommé General Manager Automation & Drives Electrical Technologies de Siemens Ltd China (1998/2002). Il est ensuite successivement Corporate Account Manager Group de Siemens AG pour les Groupes de Distribution REXEL et SONEPAR (2002/2005), General Manager A&D Low Voltage (Systèmes & Produits) de Siemens Ltd China (2005/2008) et, enfin, Senior Vice-President Siemens Building Technology North-East Asia (2008/2010). Début 2011, Christophe de Maistre est nommé président de Siemens France. Il a exercé également les fonctions de Président Europe Sud et Ouest de Siemens entre 2011 et 2013.

LE PILOTE



Delphine Manceau est Pilote de l'Observatoire de l'Innovation de l'Institut de l'entreprise.

Professeure à ESCP Europe, elle y a fondé l'Institut pour l'innovation et la Compétitivité i7 qui analyse les nouvelles pratiques d'innovation des entreprises. Elle est aujourd'hui Directrice Europe de la Division Corporate de ESCP Europe qui rassemble

les activités de formation continue (programmes sur mesure et sur catalogue), l'Executive MBA et les relations avec les entreprises.

Spécialiste de marketing et d'innovation, elle a réalisé en 2009 avec Pascal Morand le rapport *Pour une nouvelle vision de l'innovation* commandé par Christine Lagarde, alors Ministre de l'Economie, de l'Industrie et de l'Emploi, sur la capacité d'innovation des entreprises françaises et européennes. Elle est également l'auteur de l'ouvrage de référence *Marketing Management* avec Philip Kotler et Kevin Keller et de *Marketing de l'innovation* (avec Emmanuelle Le Nagard). Titulaire d'un doctorat en sciences de gestion et du diplôme d'habilitation à diriger des recherches, elle a été *Senior Fellow* à la Wharton School (University of Pennsylvania). Elle a également occupé les fonctions de Directeur académique de ESCP Europe et de Directeur du programme Master in Management Grande Ecole entre 2005 et 2008.

LE RAPPORTEUR



Julie Fabbri est Rapporteur de l'Observatoire de l'Innovation de l'Institut de l'entreprise.

Secrétaire Générale de l'Institut pour l'Innovation et la Compétitivité i7 de ESCP Europe depuis 2011, elle organise et coordonne les événements et travaux de recherche d'i7 sur les nouvelles pratiques d'innovation des entreprises.

Diplômée de ESCP Europe (Master in Management) et titulaire d'un Master Recherche en Gestion et Dynamique des Organisations de l'Université Paris X Nanterre, elle est aujourd'hui doctorante au Centre de Recherche en Gestion de l'Ecole Polytechnique (CRG) où elle s'intéresse au rôle de l'espace physique de travail et des tiers lieux (espaces de *coworking*, *fab lab*) dans les processus d'innovation de petites entreprises et d'entrepreneurs

LES DERNIÈRES PUBLICATIONS DE L'INSTITUT DE L'ENTREPRISE

Les réseaux sociaux d'entreprises : entre promesses et illusions

Par Denis Moneuse (avril 2014)

Royaume-Uni, l'autre modèle ? La Big Society de David Cameron et ses enseignements pour la France

Par Eudoxe Denis avec Laetitia Strauch (mars 2014)

Assurance chômage : six enjeux pour une négociation

Par Bruno Coquet (janvier 2014)

Smart Cities. Efficace, innovante, participative : comment rendre la ville plus intelligente ?

Par l'Institut de l'entreprise (novembre 2013)

Entreprises et territoires : pour en finir avec l'ignorance mutuelle

Par l'Institut de l'entreprise (octobre 2013)

Mettre enfin la fiscalité au service de la croissance

Par l'Institut de l'entreprise et l'Institut Montaigne (septembre 2013)

Réformer vraiment la formation professionnelle

Par Jacques Barthélémy et Gilbert Cette (septembre 2013)

Allemagne : miracle de l'emploi ou désastre social ?

Par Alain Fabre (septembre 2013)

Service public 2.0

Par Elisabeth Lulin (juillet 2013)

Toutes nos publications sont téléchargeables
sur notre site internet : www.institut-entreprise.fr

Créé en 1975, l'Institut de l'entreprise est un think tank indépendant de tout mandat syndical ou politique. Association à but non lucratif, l'Institut de l'entreprise a une triple vocation : être un centre de réflexion, un lieu de rencontre et un pôle de formation. Profondément ancré dans la réalité économique, il concentre ses activités sur la relation entre l'entreprise et son environnement. L'Institut de l'entreprise réunit plus de 130 adhérents (grandes entreprises privées et publiques, fédérations professionnelles et organismes consulaires, institutions académiques, associations...). Ses financements sont exclusivement privés, aucune contribution n'excédant 2 % du budget annuel.

THINK TANK

- La réflexion de l'Institut de l'entreprise s'organise autour de 5 thématiques prioritaires : compétitivité et innovation, emploi et prospective sociale, management, finances publiques et réforme de l'action publique.
- Dans cette réflexion, la vision de l'entreprise – conçue à la fois comme organisation, acteur du monde économique et acteur de la société – tient une place prépondérante. Pour réaliser ses études et élaborer ses propositions, l'Institut de l'entreprise met à contribution un vaste réseau d'experts (universitaires, hauts fonctionnaires, économistes, politologues, dirigeants d'entreprise, think tanks partenaires étrangers...). La diffusion de ses idées s'appuie sur la parution régulière de rapports et de notes et sur la publication d'une revue annuelle, *Sociétal* – qui propose également des débats en ligne sur les questions d'actualité économique via la page *Sociétal - Le Blog*, intégrée au site internet de l'Institut de l'entreprise. Résolument tourné vers l'international et partenaire fondateur du *Réseau International des Think Tanks Economiques* (www.isbtt.com), l'Institut de l'entreprise intègre systématiquement dans sa réflexion l'analyse de modèles étrangers susceptibles d'inspirer les politiques publiques françaises.

RENCONTRES

Ouvertes à un large public ou réservées aux adhérents, les manifestations organisées par l'Institut de l'entreprise ont pour objectif d'animer le débat public et de stimuler la réflexion sur des sujets d'intérêt collectif, liés à l'entreprise. Dirigeants d'entreprise, personnalités politiques, experts issus de l'entreprise ou du monde universitaire sont invités à s'exprimer à l'occasion de déjeuners, de conférences et de débats.

FORMATION

L'Institut de l'entreprise propose des programmes pédagogiques visant à sensibiliser les publics appartenant à l'écosystème de l'entreprise aux enjeux économiques et sociaux. Dans ce cadre, l'Institut s'adresse prioritairement aux enseignants de Sciences économiques et sociales (SES), avec le Programme Enseignants-Entreprises; aux jeunes « hauts potentiels », avec l'Institut des Hautes Études de l'Entreprise (IHEE) et Le Cercle; aux représentants politiques avec le programme Elus & Entreprises.

Pour en savoir plus : www.institut-entreprise.fr



29, rue de Lisbonne, 75008 Paris
Tél. : +33 (0)1 53 23 05 40 / Fax : +33 (0)1 47 23 79 01
www.institut-entreprise.fr

