



2014

Le Big Data et la Conduite du Changement

BENOUARI Neyla
CANOVAS Stéphanie
DIA Maimouna
MARSAUD Audrey
MAUGUIN Tiphaine

TABLE DES MATIERES

INTRODUCTION	4
<u>I. UN PHENOMENE QUI N'EN EST QU'A SES BALBUTIEMENTS MAIS QUI POSE DEJA DE VRAIES PROBLEMATIQUES AU SEIN DES ENTREPRISES</u>	
A. DE L'AVENEMENT D'INTERNET A LA DEMATERIALISATION, LES PREMISSES DU BIG DATA : UNE MINE D'OR POUR LES ENTREPRISES ET UNE NOUVELLE MANIERE D'EXPLOITER LES DONNEES	5
1. Croissance de l'information et nouvelles technologies : les prémices du Big Data	5
2. Le Big Data : une définition complexe	7
3. La donnée : une mine d'or pour les entreprises	10
B. LE DEFI DE LA GESTION ET DE LA BONNE MAITRISE D'UNE QUANTITE D'INFORMATIONS ET DE DONNEES NON STRUCTUREES SANS CESSER RENOUVELEE	12
1. Des sources diverses de données	12
2. Le Big Data et les 3V	14
3. Données structurées et données non structurées	16
C. LE DEFI DE LA SECURITE DE L'INFORMATION :	18
1. Le risque industriel et la menace de la cybercriminalité	18
2. La conduite du changement pour une meilleure gestion de la sécurité des données de l'entreprise	20
3. Entre création de valeur et éthique, l'entreprise doit maintenir ses Best Practice face au risque d'atteinte aux libertés individuelles	21
<u>II. UN PHENOMENE QUI IMPLIQUE PAR CONSEQUENT DE NOUVEAUX ENJEUX COMMERCIAUX ENTRE LES DIFFERENTS ACTEURS DU MARCHE ET UN NOUVEAU MODELE ORGANISATIONNEL</u>	25
A. DE NOUVELLES RELATIONS COMMERCIALES	25
1. Le marché du Big Data	25
2. Nouvelle relation B to B : les données en vente	26
3. Comment la DSI des entreprises choisit-elle les outils adéquats parmi les nombreuses offres disponibles ?	27
B. COMMENT ACCOMPAGNER LE CHANGEMENT DANS L'ORGANISATION DE L'ENTREPRISE EN MUTATION	28
1. Un renforcement des relations professionnelles entre informaticiens et « marketers »	28
2. Evolution du rôle des DSI	29
3. Plusieurs types de « data-driven organization »	29
4. Nouvelles compétences liées au Big Data	30
5. Les nouveaux métiers du Big Data	31
6. Challenge technologique du Big Data pour l'entreprise	32
C. COMMENT UTILISER CES DONNEES POUR AMELIORER LA PRISE DE DECISION ET L'EFFICACITE ORGANISATIONNELLE	33
1. Des décisions prises en temps réel et plus adaptées	34
2. Des outils « Big Data » adaptés au projet de l'entreprise	35
3. Une équipe de travail « Big Data »	35
4. Des professionnels « Big Data » au cœur des décisions	36
5. Une organisation à repenser dans son ensemble	36
6. Un changement qui doit être organisé et progressif	38
7. US Xpress, un exemple de « best practices »	38

III. UN PHENOMENE QUI AMENE LES ENTREPRISES A REPENSER LEUR STRATEGIE	40
A. COMMENT UTILISER LES BIG DATA POUR AMELIORER L'EXPERTISE CLIENT ? EFFICACITE DE L'ACTIVITE DES ENTREPRISES.	40
1. Les Big data, des informations précises sur le client	40
2. Un marketing transformé	41
3. Une offre qui répond et anticipe la demande	43
4. Impact sur la publicité	44
B. COMMENT LES ENTREPRISES PEUVENT-ELLES EXPLOITER LES BIG DATA POUR DEFINIR UNE NOUVELLE STRATEGIE ET UN NOUVEAU BUSINESS MODEL, CECI AFIN DE DEVELOPPER UN AVANTAGE CONCURRENTIEL	45
1. La pertinence des Big Data	45
2. De nouveaux business models	45
C. COMMENT PROCEDER A LA CONDUITE DU CHANGEMENT « BIG DATA » ?	47
1. Stratégie	48
2. Acceptation	49
3. Organisation	50
4. Projet Pilote	50
5. Déploiement	50
BIBLIOGRAPHIE	51

Introduction

Big Data, Big Brother, Big problems ...

En 2012, Andrew McAfee, directeur associé du Centre pour le Digital Business at the MIT Sloan School of Management, définissait le monde comme étant un problème Big Data. Utilisé pour la première fois par le cabinet d'études Gartner en 2008, le terme Big Data, mégadonnées en français, répond à de nombreuses définitions et est encore sujet à interprétation. Le Big Data est à la fois présenté comme un phénomène qui bouleverse les structures de la société, comme un ensemble de données mais aussi comme un ensemble de techniques et de technologies. Faisant référence à l'explosion des données numériques, le Big Data pose les enjeux d'un ensemble nouveau de technologies, d'architecture, d'outils et de procédures permettant à une organisation de très rapidement capter, traiter et analyser de larges quantités et contenus hétérogènes et changeants, et d'en extraire les informations pertinentes à un coût accessible.

Défi majeur pour le secteur tant privé que public, les Big Data aideraient à générer de nouvelles sources de profit, de nouvelles valeurs de différenciation ou d'innovation en organisant de nouvelles interactions entre les données présentes dans le serveur informatique d'une entreprise ou d'une administration et celles qui sont délocalisées en masse sur le web.

Chaque minute, ce sont 98 000 tweets échangés, 695 000 contenus Facebook partagés, 698 500 recherches Google effectuées, 72h de vidéos mises en ligne sur Youtube, 168 millions de mails envoyés. En une journée, 2.5 trillions d'octets de données sont créés si bien que l'on considère aujourd'hui que 90% des données existantes actuellement ont été créées au cours des deux dernières années. En 2020, le Big Data devrait représenter 8% du PIB européen. Mais cette croissance exponentielle des données disponibles ne va pas sans risques et, aux difficultés de gestion liées à ce volume d'information nouveau, s'ajoutent les craintes des citoyens face à la manipulation de leurs données personnelles dans le cadre du Big Data.

Entre richesse de l'information et risque d'atteinte aux libertés individuelles, qu'implique le phénomène Big Data et quels sont les enjeux organisationnels, commerciaux et stratégiques des entreprises aujourd'hui dans l'exploitation de ces données ?

Dans notre étude, nous nous focaliserons sur le secteur privé et sa manière d'aborder le changement lié au phénomène des Big Data. En effet, quand en 2014 une étude du cabinet de conseil Accenture montre que, sur un panel de 1000 responsables d'entreprises mondiales, 60% sont engagés dans un projet de

traitement massif de données, l'étude MARKESS sur les « Meilleures approches pour tirer parti du Big Data » montre que seulement 5% des organisations publiques interrogées conduisent actuellement des projets du même type.

Malgré un nombre d'avancées significatives grâce à de nouvelles législations ainsi que la politique d'ouverture des données publiques, le secteur public demeure en retard sur l'exploitation de ces données. Afin d'étudier et de comprendre l'importance de la conduite du changement, il nous semblait plus pertinent de nous axer sur un secteur où de nouvelles méthodes et de nouveaux instruments ont déjà émergé pour mieux les analyser et préconiser les meilleures solutions quant à la conduite du changement nécessaire dans les entreprises face à ce phénomène nouveau.

I. Un phénomène qui n'en est qu'à ses balbutiements mais qui pose déjà de vraies problématiques au sein des entreprises

A. De l'avènement d'internet à la dématérialisation, les prémisses du Big Data : une mine d'or pour les entreprises et une nouvelle manière d'exploiter les données

1. Croissance de l'information et nouvelles technologies : Les prémices du Big Data

Depuis 1941, le terme « explosion de l'information » est référencé dans l'*Oxford English Dictionary*. En effet, la popularité autour du terme Big Data est certes actuelle, mais découle d'un phénomène de croissance des données engendré depuis près de 70 ans. En 1944, Fremont Rider de l'université de Wesleyan estime que la taille des bibliothèques américaines double tous les seize ans. D'après son postulat, les fonds de Yale en 2040 devraient s'élever à 200 millions de volumes.

En 1961, dans *Science Since Babylon*, Derek Price conclut que le nombre de revues scientifiques a augmenté de manière exponentielle, doublant en l'espace de quinze ans. Il définit ainsi la loi de l'augmentation exponentielle qu'il explique comme le fait que chaque avancée scientifique génère une nouvelle série d'avancées.

En 1975, le Ministère des Postes et des Télécommunications japonais a recensé les flux d'informations circulant dans le pays ainsi que leur volume. L'étude a permis de montrer que l'offre en information augmente bien plus vite que la consommation de ces données. Il souligne aussi que la demande pour les informations fournies par les médias de masse stagne quand la demande pour les informations fournies par les médias de télécommunications a augmenté drastiquement.

En 1980, les premières interrogations quant au stockage et à la bonne gestion de la donnée existante apparaissent. IA Tjomsland décrit ainsi le fait que les données s'étendent pour remplir totalement l'espace existant. Il s'interroge sur le caractère obsolète de certaines données et les complications liées à un stockage systématique et non différencié des données produites devant le volume qu'elles représentent.

En 1990, Peter J. Denning estime dans *Saving All the Bits* que la croissance et la vitesse de croissance des informations est trop importante et dépasse les capacités des réseaux actuels, des logiciels de stockage et de sauvegarde, mais dépasse aussi les capacités d'analyses humaines. L'avenir passe pour lui par la construction de machines capables de surveiller le flux d'informations et d'en proposer un résumé statistique. Il définit ainsi les nouveaux défis qui attendent les systèmes informatiques ainsi que les systèmes d'informations à savoir élever les capacités de mémoire des ordinateurs pour s'adapter aux nouveaux volumes de l'information.

En 1998, K.G. Coffman et Andrew Odlyzko mettent en exergue dans *The Size and Growth Rate of the Internet* le rôle majeur de la croissance et de l'accès à Internet dans le phénomène d'avènement de l'information avec une augmentation de près de 100% de son trafic par an.

Si l'accélération technique (optimisation des composants systèmes comme la mémoire ou les processeurs, montée en puissance et gigantisme d'internet) est une garantie de bon fonctionnement, elle n'est pas suffisante à elle seule pour expliquer l'arrivée du phénomène Big Data. Par exemple, s'il y a bien une augmentation de l'espace de stockage proposé, les données s'étendent systématiquement jusqu'à le combler.

Le Big Data est le fruit d'une rupture technologique engendrée par les problèmes rencontrés par Google et Yahoo, deux géants de l'internet. En effet, ces moteurs de recherche devaient faire face à un volume de données manipulées gigantesque, au caractère non structuré de ces informations, ainsi qu'à la rapidité nécessaire à exécuter les demandes des utilisateurs et donc à la vitesse d'exécution des algorithmes de traitement.

En 2000, les premiers travaux de recherche destinés à remplacer les approches traditionnelles de gestion de bases de données et d'informatique décisionnelle sont conduits afin d'inventer un modèle pouvant régler simultanément ces problèmes. Les premières solutions ont ouvert la voie à un nouveau mode de traitement analytique de l'information qui stocke les données et déploie des calculs algorithmiques.

L'idée est de décomposer les problèmes complexes en plusieurs calculs traités de manière simultanée et associe un mécanisme de fichiers qui distribue, réplique et compresse les données à un modèle de programmation parallèle qui exécute les

traitements localement sur les machines de stockage. Les composants n'étant pas partagés entre les unités de traitement, on parle de parallélisme. Cette parallélisation permet un traitement rapide de données complexes et une adéquation en temps quasi réel à la demande.

Le développement de framework adapté pour le traitement de volumes de données importants tel que Hadoop et son écosystème ont permis l'émergence du Big Data. Utilisé par des entreprises ayant de très fortes volumétries de données à traiter comme Twitter, eBay ou Amazon, le framework répartit l'exécution d'un traitement intensif de données sur plusieurs nœuds ou grappes de serveurs. L'avantage de ce type de cadre est de permettre une meilleure analyse des données en incorporant un support pour les données de différents types, en captant les données en temps réel, en utilisant ce qui existe déjà dans l'environnement de l'entreprise, en supportant le NoSQL et toutes les nouvelles formes d'accès aux données qui sont plus rapides, en fournissant un stockage au meilleur prix, etc.

Hadoop est aujourd'hui le plus populaire et a tellement le vent en poupe que presque tous les acteurs du secteur des bases de données (Oracle, Microsoft, IBM, Teradata,...) ou de l'analytique (SAS, R, Micro Strategy, ...) ont maintenant annoncé des solutions autour de ce nouvel écosystème.

2. Le Big Data : une définition complexe

Comme nous l'avons évoqué dans l'introduction, le terme Big Data est encore aujourd'hui soumis à des questionnements quant à sa définition. Si la plus commune veut que le Big Data soit l'ensemble de données devenu tellement volumineux qu'ils en deviennent difficiles à travailler avec des outils classiques de gestion de base de données ou de gestion de l'information, les points de vue sont nombreux et dégagent d'autres dimensions à cette première approche.

En 1977, le terme Big Data est utilisé pour la première fois dans un rapport de scientifiques de la NASA. Le Big Data désigne alors un problème d'adéquation entre des données trop importantes, trop larges, qui, cannibalisant les capacités et la mémoire des processeurs, et les ressources disponibles.

En 2008, le cabinet d'études Gartner remet le terme Big Data au goût du jour et définit ce dernier comme étant des actifs informationnels volumineux, véloce et variés qui requièrent des dispositifs de traitement de l'information innovants, performants et rentables afin d'améliorer leur processus de décisions. Le terme sera popularisé par le Computing Community Consortium et les publications d'ingénieur informatique prédisant que le « Big Data informatique transformera les activités des entreprises, de la recherche scientifique, des praticiens généralistes et les activités de défense et de renseignement d'une nation ». Pour autant, le terme n'a pas été défini clairement.

En 2011, une étude réalisée par le cabinet de conseil McKinsey souligne l'importance qui existe à définir correctement le terme Big Data alors que la première

définition de la NASA semble atteindre ses limites. En effet, des termes comme « large » sont trop vagues pour rendre correctement le phénomène d'accroissement des données, d'autant plus que l'environnement écologique a évolué. McKinsey caractérisera alors le Big Data comme des « fichiers de données dont la taille dépasse les capacités des logiciels de gestion des bases de données ordinaires pour saisir, enregistrer, stocker, gérer et analyser les données ». Le cabinet précise que cette décision est intentionnellement subjective et floue quant à savoir quelle quantité de données est nécessaire pour être considérée comme faisant partie du Big Data.

En 2012, Viktor Mayer-Schönberger, Professeur de gouvernance et régulation à l'université d'Oxford, et Kenneth Cukier publient *Big Data : A Revolution That Transforms How we Work, Live, and Think* dans lequel ils expliquent qu'il n'existe pas de « définition rigoureuse du Big Data ». Ils en proposent néanmoins une qui met en valeur ce qui peut être fait avec les données et qui explique pourquoi leur taille importe : « c'est la capacité d'une société à exploiter l'information de manière à produire des idées, des biens et des services de valeur » et « ce sont les choses qui peuvent être réalisées à grande, mais pas à petite échelle pour extraire de nouvelles idées et créer de nouvelles sources de valeur ».

Tom Davenport, consultant senior à Deloitte, conclut qu'en raison du « problème de définition » du Big Data, ce terme devrait connaître une expansion limitée. En effet, aujourd'hui encore, la définition reste un sujet sensible lorsque l'on étudie les Big Data et les différences de définitions sont source de confusion et d'interprétation trop variées. Ainsi en 2014, une étude menée par Berkley School of Information a proposé à 40 dirigeants des secteurs de l'édition, de la mode, de l'agroalimentaire, de l'automobile, du domaine pharmaceutique et du marketing de définir le Big Data. Les résultats obtenus ont montré à la fois des grandes disparités dans les termes, mais aussi des idées communes comme le montre le word cloud si dessous.

3. La donnée : une mine d'or pour les entreprises

À l'expression usuelle des trois V du Big Data s'ajoute un cinquième qui correspond à la valeur de ces données. Dès lors, comme l'exprime Peter Sondergaard du cabinet Gartner, « l'information est le pétrole du 21^e siècle » et la donnée une matière première, un capital dont l'entreprise regorge, mais qu'elle doit réussir à exploiter correctement afin de pouvoir la valoriser. En effet, il semble normal que les organisations qui mettent en œuvre des solutions de Big Data cherchent à identifier, comment le Big Data peut être source de valeur et leur procurer une valeur ajoutée.

Dès lors, devant cette masse de données récoltées par les entreprises géantes du secteur, il convient de voir en quoi elles contribuent à la valorisation de l'entreprise et constitue pour cette dernière une véritable mine d'or d'informations.

Dès les années 2000, la question de la valeur des données issues du Big Data apparaît et concerne d'abord les bénéfices dans certains domaines scientifiques. Mais ce n'est qu'avec la montée en puissance des acteurs majeurs d'internet que sont Google, Facebook, Amazon, Twitter, Linkedin et leur succès économique considérable sur cette décennie, que les données ont commencé être perçues comme ayant une valeur intrinsèque (que ce soit dans leur volume, leur variété et leur vitesse de production et donc d'adaptation au marché) monétarisable comme l'illustre notamment Facebook.

Facebook a pour objectif la mise en contact d'individus se connaissant, parce qu'issus du même cercle familial ou amical. Permettant de mettre en relation les hommes à travers le monde, le réseau fonctionne sur un ensemble d'algorithmes. À chaque fois qu'une personne partage un lien, écoute une musique, « Like » une page ou change sa situation de célibataire à en couple, une multitude de données sont ajoutées à la base de données de Facebook. Ce sont autant de données qui s'accumulent et qui permettent d'établir le profil complet des individus présents sur le réseau en passant par son nom, son image, ses goûts, ses préférences et ses centres d'intérêt.

Si Facebook est considéré si profitable aujourd'hui, ce n'est pas parce qu'il est la somme de ses 1.2 milliard d'utilisateurs de par le monde, mais parce qu'à chaque clic sa base de données capture et s'agrandit. Facebook propose aux annonceurs un service de publicité par centres d'intérêt ou caractéristiques des profils pour que leur publicité atteigne la juste cible. Aujourd'hui, un utilisateur Facebook vaut \$81, chaque ami \$0.62 et une page de profil peut s'élever jusqu'à \$1 800.

Outre le processus de monétarisation, l'idée que ces données ont une valeur complémentaire liée aux analyses et aux corrélations avec d'autres données et qu'elle participe à la création de valeur de l'entreprise est apparue en même temps et a été largement diffusée sur le modèle de l'algorithme de recommandation d'Amazon. Chaque utilisateur, après avoir consulté au moins un article, voit des suggestions s'afficher comme le montre l'image suivante.

Produits fréquemment achetés ensemble



Prix pour les trois: EUR 16,52

[Ajouter ces trois articles au panier](#)

[Afficher la disponibilité du produit et le mode de livraison](#)

- ✓ Cet article : Fahrenheit 451 de Ray Bradbury Broché EUR 4,84
- ✓ Le Meilleur des mondes de Aldous Huxley Poche EUR 4,27
- ✓ 1984 de George Orwell Poche EUR 7,41

Les clients ayant acheté cet article ont également acheté

Page :



Aujourd'hui, on estime que 30% du chiffre d'affaires total d'Amazon est généré grâce à son algorithme de recommandation. Le développement de cet outil se base sur le postulat que le client aime être conseillé par des personnes ayant des goûts similaires aux siens. Ainsi, Amazon a mis au point un algorithme, appelé item-to-item collaborative filtering, de contenu collaboratif qui, plutôt que de mettre en relation un utilisateur avec des clients similaires, met en rapport chacun des articles achetés et classés par les utilisateurs avec des articles semblables et ensuite une liste de recommandations. La priorité est donc à la personnalisation à l'extrême pour chaque utilisateur du site en fonction de ses intérêts propres.

Aujourd'hui, le processus d'évaluation d'entreprise, c'est-à-dire l'estimation de la valeur de marché potentielle ou de la valeur patrimoniale d'une entreprise en fonction de critère objectif tend de plus en plus à intégrer la dimension de l'information. En effet, les données possédées par l'entreprise sont de plus en plus créatrices de valeur. Ce ne sont pas les données en tant que telles qui sont source de valeur, mais leur exploitation, leur manipulation, leur croisement, leur consolidation et bien sûr leur analyse qui permettent de créer et de proposer de nouveaux services ou d'améliorer les existants comme l'a expliqué Olivier Lallement, manager en conseil et technologie chez Deloitte.

Les données sont aujourd'hui un atout majeur pour les entreprises. En effet, l'ensemble des informations relatives à l'usage de l'utilisateur comme son écosystème, son environnement, sa localisation, etc. est une synthèse de son comportement qui permettra aux entreprises de mieux connaître son client et de mieux cibler ses attentes et les offres de l'entreprise. La valeur ajoutée liée au Big Data se situe donc dans la capacité à croiser ces données comportementales avec des données externes afin d'offrir de nouveaux services ou d'améliorer ceux existants précédemment.

Au-delà du traitement des données et de leur analyse dans un but de reporting qui vise à rendre compte du passé, les informations liées au Big Data ont aussi un caractère prédictif et permettent une projection plus claire et plus sûre dans le futur. L'analyse n'est plus réactive, mais prédictive, ce qui est source de changement majeur dans les processus de l'entreprise en remettant le client au centre de l'analyse et des transactions.

L'analyse de la masse d'information est une aide à la prise de décision pour les entreprises, car elle permet de profiter des énormes volumes de données transactionnelles jusqu'alors non structurées comme des emails ou des contenus de réseaux sociaux qui étaient jusqu'alors inexploitées par les programmes standards. Mais les technologies Big Data permettent aussi un meilleur suivi de l'activité ou de la chaîne logistique, la détection d'erreurs, la maintenance prédictive, etc., en somme, l'optimisation des processus, du chiffre d'affaires ainsi que la réduction des coûts. Les entreprises doivent prendre conscience que les données sont une richesse, aidant à la prise de décisions et doit être valorisée.

B. Le défi de la gestion et de la bonne maîtrise d'une quantité d'informations et de données non structurées sans cesse renouvelée

1. Des sources diverses de données

Toute la complexité du Big Data tient dans la quantité de données disponibles dont la croissance est exponentielle. L'adjectif anglais « Big » traduit explicitement le volume impressionnant des informations que fournit la Data. Appelée à juste titre « données massives » ou « méga données », la Data offre des perspectives d'évolution qui sont encore pour partie insoupçonnées. Les données se trouvent partout. Tout est donnée : les informations contenues dans nos ordinateurs, nos recherches sur internet, les documents, emails au sein d'une entreprises, les panneaux publicitaires ... A tel point que les machines (hardwares) et logiciels (software) ne sont plus en mesure de traiter les gigantesques volumes de données. Les données ne sont pas homogènes et il existe plusieurs manières de distinguer les différents types de données notamment par leur provenance et leur source. Il existe alors des données que l'on peut qualifier d'internes, d'autres données externes et celles regroupant les deux types pré-cités.

Les données internes sont générées par deux types de sources :

- les archive : on y trouve l'ensemble des documents numérisés, déclarations, archives médicales, formulaires d'assurances, dossiers médicaux, archives papiers, mails d'échanges entre l'entreprises et ses clients.
- Le « Data storage » ou stockage de données informatiques dans lequel nous regrouperons les appareils et supports qui conservent des données numériques d'ordinateurs tels Hadoop dont nous aurons l'occasion de parler, les méthodes de traitement SQL, NoSQL.

Les données externes correspondent à un ensemble d'informations disponibles sur internet que l'on qualifie de « Public wEB » que fournissent les gouvernements, la météorologie, les satellites, le trafic, la finance publique, la Bourse, les données Wikipédia.

Enfin les données dites « mixtes », ont à la fois une provenance interne et externe sont les données issues :

- des médias de tout type : images, photos, documents audio, podcasts ...
- des applications – entreprises ou « Business Apps » : des services d'automatisation du marketing sur le Web, CRM, Google documents, portails ...
- des réseaux sociaux tels que Twitter, LinkedIn, Facebook, Tumblr, Youtube, Instagram ...
- Sensor data ou données de capteurs fournies par les compteurs électriques, les satellites, les dispositifs d'enregistrement de trafic aérien, automobiles, les machines, les jeux vidéos, les appareils électroménagers ...
- Machine log data : ensemble de données des serveurs, applications GPS de localisation sur téléphones mobiles, applications sur Smartphones ...

Depuis cinquante ans, l'expansion d'internet, du digital et des technologies diverses a fait exploser la production de données. Chaque minute, de la donnée est récupérée et stockée en masse. Il est important de noter que si des sociétés virtuelles telles que Amazon ou encore Google sont les principaux responsables du stockage de la donnée et de sa redistribution, les réseaux sociaux - notamment Facebook en 2003, Twitter en 2006 ou encore les sites de partage de musique tel que Youtube créé en 2006, sont un phénomène récent et qui pourtant, compte tenu de la vitesse de son développement, représente aujourd'hui l'une des sources principales de production de données. En avril 2010 Facebook lançait son Open Graph, un protocole qui permet à des sites tiers d'interagir avec les informations d'un profil Facebook et avec les relations de ce dernier. Ces informations sont appelées objets, puisque sur Facebook, les utilisateurs sont connectés à leurs relations sociales, mais également à des objets, par exemple à des pages Facebook, des articles de presse ou encore des morceaux de musiques etc. Deux ans plus tard, Marc Zuckerberg, fondateur et PDG du réseau social mondialement connu, déclarait que l'Open Graph de Facebook générerait en 2012, à lui seul, près d'un milliard d'éléments de contenus (liens Internet, news, articles de blogs, notes, albums photos) par jour.

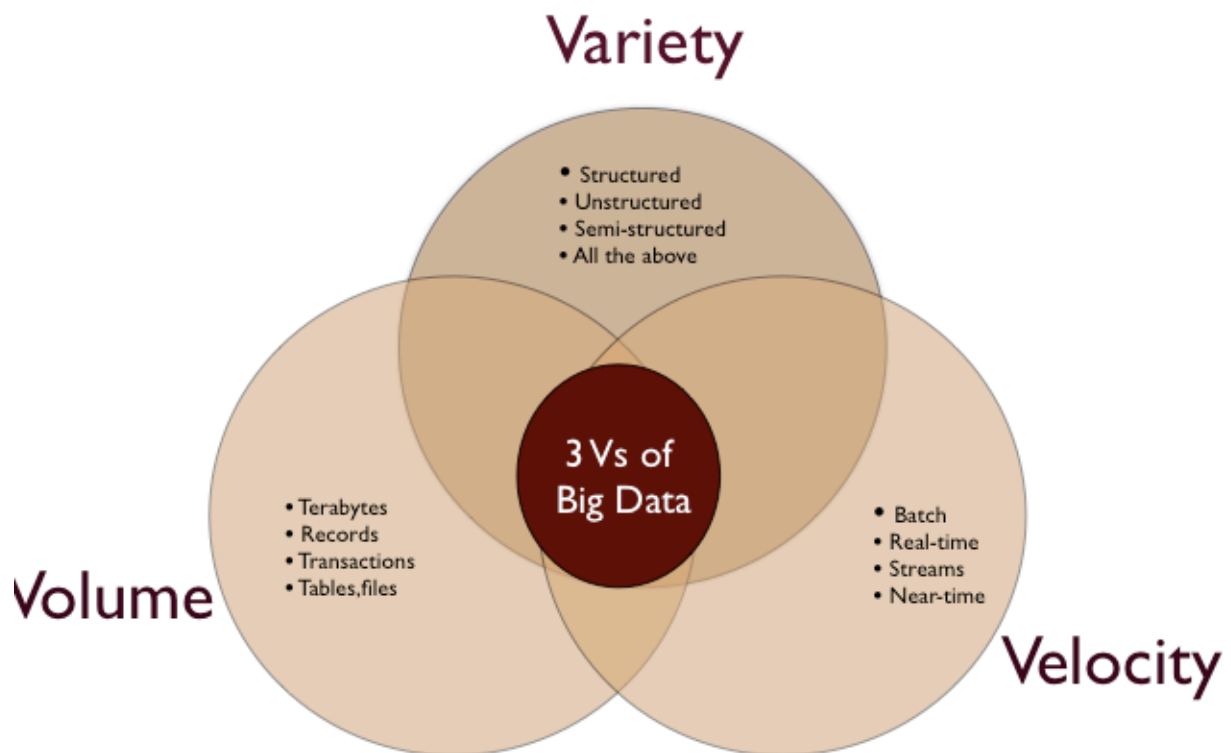
Cette anecdote nous permet d'en tirer une corrélation indubitable entre la production de données toujours plus importantes et ce que les anglo-saxons nomment la « **digitization** ». En effet, force est de constater que la production en masse de données est étroitement liée à la transition qui s'est opérée entre deux mécanismes de transmission de l'information. On parlera ici de signaux de transmission. Les signaux analogiques (ancien mécanisme) et les signaux numériques (l'ère Digital que nous connaissons aujourd'hui, sont utilisés pour transmettre des informations, généralement par le biais de signaux électriques. Dans ces deux technologies, l'information, tels que les fichiers audio et vidéo, est transformée en signaux électriques. La différence entre les technologies analogiques et numériques, c'est que dans la technologie analogique, l'information est traduite en impulsions électriques d'amplitude variable. Dans la technologie numérique, la traduction de l'information est en format binaire (zéro ou un) où chaque bit est représentatif de deux amplitudes distinctes.

De cette manière, ce que nous appelons Digitization, et qui correspond à ce mouvement de transition de l'analogique au digital corrélée aux progrès de la technologie, comme l'augmentation de la vitesse de la mémoire et de traitement ainsi que l'accroissement du stockage pas cher, y compris le stockage en Cloud, ont permis de recueillir et de stocker des données comme jamais auparavant. Selon une étude si en 2000, 75% de l'information dans le monde était générée par des signaux analogiques, en 2007 l'analogique ne représentait plus que 6% de l'information disponible, le Digital ayant pris de plus en plus d'ampleur et qui continue aujourd'hui sa croissance exponentielle.

Sur la période 2011-2016, on estime que le taux de croissance annuel moyen mondial du marché de la technologie et des services du Big Data atteindra de 31,7%. Ce marché devrait ainsi atteindre 23,8 milliards de dollars en 2016 (d'après l'étude de l'IDC, International Data Corporation menée en mars 2013). Une autre source estime que le volume de données créées chaque année par des sociétés américaines à elles seules, suffit à remplir dix mille Bibliothèques du Congrès. Ces chiffres toujours plus impressionnants les uns que les autres donne une vision imagée de l'entendue du phénomène Big Data qui n'en est qu'à ses balbutiement et qui représente déjà aujourd'hui une richesse incontestable pour les entreprises. Toute interaction avec les ordinateurs, Smartphones, e-lecteurs, lecteurs de musique numériques, tablettes, ordinateurs portables, recherches sur Internet, génère des données précieuses sur les potentiels clients des entreprises. Le cabinet de conseil McKinsey, dans un rapport de 2010 consacré au Big Data prédisait une augmentation de 60 % de la marge d'exploitation des distributeurs qui utiliseraient pleinement ces énormes volumes de données. Une richesse oui, mais il ne reste que ces données doivent être très soigneusement traitées et c'est précisément pour cela qu'au delà de la collecte d'informations, les stratégies pour diriger et contrôler ces dernières représentent un défi majeur pour les entreprises.

2. Le Big Data et Les 3V

Tout cela nous amène à présent à analyser la mesure de ces données. Pour cela, il est important de comprendre les 3 V du Big Data. Ce concept a été mis en avant par les chercheurs et cabinets portant un intérêt au phénomène Big Data et qui a permis d'en dégager ses caractéristiques fondamentales. Ces trois V correspondent aux initiales de trois notions fondamentales pour la définition du Big data : «Volume» qui, comme nous avons pu l'aborder correspond à la quantité de données disponibles, "Vitesse" comme la vitesse à laquelle l'information est produite et se dessert dans l'entreprise) et «variété» correspondant au type de données disponibles.



Le volume décrit la quantité de données générées par des entreprises ou des personnes. La quantité d'information traitée par les réseaux informatiques dans le monde chaque jour est mesurée en téraoctets (un trillion d'octets) et pétaoctets (un quadrillion octets). Selon la Salem Press Encyclopedia on estime en 2013, que 90 pour cent de la somme totale de données disponibles dans le monde aujourd'hui, incorporant de l'information générale, les statistiques, les dossiers médicaux, les données scientifiques et de recherche, des enregistrements téléphoniques, documents commerciaux, documents gouvernementaux n'a été générée que depuis 2010, soit en l'espace de trois ans seulement.

La vitesse décrit la fréquence à laquelle les données sont générées, capturées et partagées. Du fait des évolutions technologiques récentes, les consommateurs mais aussi les entreprises génèrent plus de données et ce dans des délais de plus en plus courts. À ce niveau de vitesse, les entreprises ne peuvent capitaliser sur ces données que si elles sont collectées et partagées en temps réel. C'est précisément à ce stade que de nombreux systèmes d'analyse, de CRM, de personnalisation, de point de vente ou autres, échouent. Ils peuvent seulement traiter les données par lots toutes les quelques heures, dans le meilleur des cas. Or, ces données n'ont alors déjà plus aucune valeur puisque le cycle de génération de nouvelles données a déjà commencé.

La Variété correspond elle à la prolifération de types de données provenant de sources comme les médias sociaux, les interactions entre différents postes ordinateurs ou les terminaux mobiles, créant une très grande diversité au-delà des données transactionnelles traditionnelles. C'est ce que l'on a pu aborder précédemment lorsque l'on définissait trois types de données (internes, externes et mixtes). Les données se complexifient, ne sont plus homogènes et ne s'inscrivent plus dans des structures nettes, faciles à consommer. Les nouveaux types de données incluent contenus, données géographiques spatiales, points de données matériels, données de géolocalisation, de connexion, de mesures, données mobiles, points de données physiques, processus, données issues de recherches, données

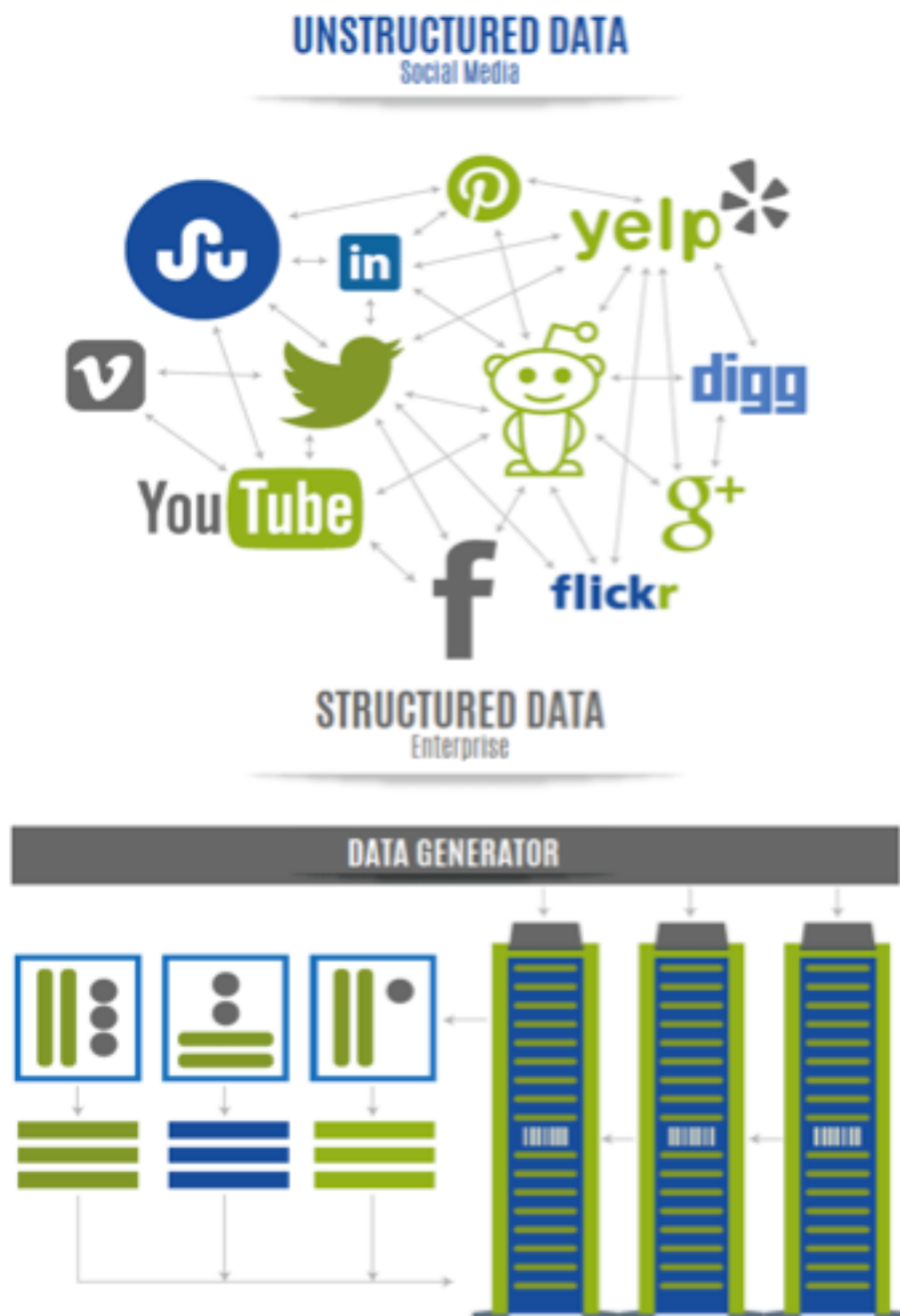
de flux, données issues des médias sociaux, données texte et données issues du Web.

3. Données structurées et données non structurées

A travers les trois caractéristiques du Big Data pré citées, il est également important de comprendre le mélange des données non structurées et multi-structurées que comprennent les volumes d'informations. La data structurée est un semble de données analysées et facilement exploitables car leurs valeurs possibles sont déterminées et connues à l'avance. Prenons l'exemple d'une base de données rassemblant les résultats d'une enquête d'opinion, de Didier Bourigault qui est Directeur de la R&D chez Synomia et chercheur au CNRS : les tranches d'âges ou la liste des catégories socio-professionnelles possibles sont déterminées a priori. C'est alors de la data structurée.

Il en existe d'autres qui sont dites non structurées et qui proviennent de l'information qui n'est pas organisé ou facilement interprété par bases de données ou par les modèles traditionnels. Toujours dans l'exemple de l'enquête d'opinion, les réponses libres aux questions ouvertes sont des données non structurées, car ces réponses sont potentiellement toutes différentes et impossibles à catégoriser a priori. Plus concrètement encore, les *tweets* sur Twitter, et autres messages de médias sociaux sont de bons exemples de données non structurées. De façon générale, les données non structurées sont des données textuelles. Enfin, les données multi-structurées se réfère à une variété de formats et types de données et peut être obtenue à partir des interactions entre les hommes et les machines, telles que les applications Web ou les réseaux sociaux. Un bon exemple serait l'ensemble des données de journal web, qui comprend une combinaison de texte et des images avec des données structurées comme forme ou des informations transactionnelles.

En vingt ans, le Big Data est rapidement devenu un axe stratégique fondamental de réussite de l'entreprise dans tous les secteurs. Il constitue une opportunité d'obtenir des connaissances sur le secteur, sur les consommateurs afin de rendre l'entreprise plus agile, plus réactive aux préférences des clients, aux modifications de comportement de consommation, aux nouveaux besoins. Jusqu'à présent, il n'y avait encore aucun moyen d'exploiter cette opportunité. Néanmoins, le facteur clé de succès pour l'entreprise n'est pas de posséder un maximum de données, mais de savoir traiter ces dernières et les analyser et c'est précisément en cela que le Big Data pose le grand défi aux entreprise et qui concerne le traitement des données, notamment les données non structurées.



De nombreux dirigeants affirment aujourd'hui, ne pas avoir leurs entreprises proprement équipées pour tirer profit du Big Data. Les deux risques majeurs dans ce grand défi auquel fait face l'entreprise aujourd'hui sont : la non compréhension des données et le retard dans le traitement de la data qui ne fera que noyer l'entreprise dans un amas d'informations. Une étude établie par le cabinet de conseil Bain and Company a permis d'interroger des dirigeants de plus de 400 entreprises dans le monde entier dont le chiffre d'affaires excédait le milliard de dollars. Ces dirigeants se sont vus interrogés sur leurs capacités à analyser la Data et la vitesse de prise de décision qui s'ensuivait. Les résultats surprenants ont permis de constater que seulement 4% des entreprises mettent en place les bons outils et les bonnes personnes pour le traitement des données. Face à ce constat, il a également été observé que ce groupe d'élite sont deux fois plus susceptibles d'être dans le premier quartile de la performance financière au sein de leurs industries, trois fois plus susceptibles d'exécuter les décisions comme prévu et cinq fois plus susceptibles de

prendre des décisions plus rapidement. Accompagner l'entreprise dans la transition vers le Big Data et optimiser l'utilisation du Big Data est un processus en trois parties qui requiert une stratégie définie en amont, la mise en place des outils nécessaires à l'analyse des informations en un minimum de temps afin de garantir la pertinence de l'information qui leur est délivrée et enfin une capacité organisationnelle telle que l'entreprise puisse en tirer le meilleur. Et c'est précisément cette problématique de gestion du changement que nous aborderons dans la suite de notre analyse.

C. Le défi de la sécurité de l'information :

1. Le risque industriel et la menace de la cybercriminalité

Le changement concerne aussi bien les enjeux de sécurité de l'information qu'il implique l'exploitation de données et il est nécessaire de nous y attarder. En effet, au vu de la complexité exponentielle des environnements informatiques et de la démocratisation de l'Internet, les menaces de sécurité deviennent de plus en plus graves et complexes. Ces menaces touchent aussi bien l'entreprise elle-même que le consommateur final. En effet dans les deux cas, l'entreprise est confrontée à des problématiques de protection de ses propres données contre les risques industriels d'une part, et d'autre part de respect des libertés individuelles en corrélation avec l'éthique de l'entreprise et ses *best practice*.

En effet, le Big Data a entraîné des volumes très importants de données informatiques qui rendent la sécurisation de la data collectée de plus en plus difficile. Malgré des investissements colossaux pour assurer cette sécurité, les auteurs des attaques et des vols d'information semblent être assez organisés pour prendre le dessus. Plusieurs paramètres permettent d'expliquer cette situation délicate : tout d'abord, les auteurs des attaques, comme mentionnés ci-dessus sont de plus en plus organisés et disposent de plus en plus de moyens. Face à ce phénomène croissant de vol d'information, la défense des entreprises, elle est restée statique et n'a pas été proportionnelle aux menaces. Les auteurs des attaques ont alors misé sur les failles des infrastructures hyper connectées de ces entreprises pour mieux les exploiter. De plus, les entreprises continuent de se complexifier notamment les entreprises liées aux technologies de l'information. Désormais dépendantes de systèmes plus ouverts et plus agiles, elles ont le besoin imminent d'ouvrir les champs des possibles en matière de communication et d'innovation. La combinaison de ces différents facteurs dans les environnements informatiques, la multiplication des interdépendances et l'extension du champ de responsabilités, a rendu la gestion de la sécurité bien plus complexe que jamais. D'après un rapport d'enquête sur les violations de données de Verizon établie en 2012, 91% des violations entraînent des altérations de données en quelques jours ou moins, alors que 79% des violations ont nécessité plusieurs semaines ou plus pour être découvertes. Face à la croissance des opportunités qu'impliquent la collectées, les violations informatiques et ce qu'on pourrait appeler la cybercriminalité n'ont jamais causé autant d'interruptions ni coûté aussi cher aux entreprises. L'entreprise est de plus en plus menacée. La cybercriminalité est un domaine en pleine croissance de la criminalité. De plus en plus de groupes de

personnes exploitent la rapidité, la commodité et l'anonymat de l'Internet pour commettre un large éventail d'activités criminelles à l'encontre des personnes ou des entreprises. De nouvelles tendances dans la cybercriminalité apparaissent tout le temps, avec des coûts pour l'entreprise et pour l'économie mondiale que l'on estime à des milliards de dollars : vol d'information financières, piratage des systèmes d'informations, accès aux données touchant à la recherche et le développement.

Dans le passé, la cybercriminalité a été commis principalement par des individus ou des petits groupes. Aujourd'hui, avec l'importance du phénomène de méga données, l'accès à plus d'éléments connectés, nous assistons à la naissance d'organisations criminelles où des professionnels des technologies commettent les plus grands vols de données.

Business intelligence et protection contre les risques industriels, représentent là tout le défi Big Data, de la gestion de la sécurité pour les entreprises.

L'exemple Wikileaks :

En 2010, le jeune Julian Assange connu en tant que fondateur, rédacteur en chef et porte-parole de Wikileaks fait la une de tous les médias mondiaux. Site lanceur d'alerte publiant des documents et analyses politiques et sociales dont la raison d'être est de donner une audience aux fuites d'informations, tout en protégeant ses sources, *Wikileaks* publie le 5 avril 2010 une vidéo de l'armée américaine montrant deux photographes de Reuters, tués par un hélicoptère Apache, lors du raid aérien du 12 juillet 2007 à Bagdad. Cette publication marque le début de la célébrité mondiale du site Internet, jusqu'alors moins connu du grand public que ses propres fuites.

Le scandale Wikileaks révèle la plus grande violation de données classifiées dans l'histoire. L'organisation On estime que Wikileaks a posté sur son site Internet. 391 832 documents classés en concernant que la guerre en Irak.

Quelles que soient les conséquences que ce scandale a pu avoir sur les forces armées américaines et les dirigeants politiques soumis à une soudaine transparence involontaire. Les révélations Wikileaks (« *leak* » signifiant en anglais « fuite »), ont surtout mis en évidence les dangers liés à la sécurité de l'information et a servi de rappel pour la communauté que la cybercriminalité et la mauvaise gestion de la sécurité est réelle et menacent les données de tous. Le scandale de Wikileaks, qui a mis au grand jour des dizaines de milliers de documents secrets, est resté un traumatisme pour le renseignement américain, contraint de mettre en œuvre des réformes pour cela ne se reproduise plus. Le patron du renseignement, James Clapper, déclarait alors qu'empêcher de nouvelles fuites comme celles du site Internet Wikileaks en 2010 représentait un "*défi*" de tous les instants.

2. La conduite du changement pour une meilleure gestion de la sécurité des données de l'entreprise

Nous comprenons aisément à travers le cas Wikileaks, que la maîtrise de ses données à travers la bonne gestion de la sécurité est le deuxième défi auquel se confronte l'entreprise aujourd'hui. Résoudre efficacement les problèmes représente une vraie stratégie Big Data aujourd'hui dans la gestion de la sécurité. Accompagner le changement dans l'entreprise pour assurer sa sécurité revêt alors trois aspects :



Figure 1. Les piliers du Big Data dans la gestion de la sécurité

L'infrastructure : la gestion de la sécurité doit pouvoir soutenir les nouvelles initiatives de l'entreprise qui ont une incidence sur l'infrastructure informatique, qu'il s'agisse des nouvelles applications ou de la virtualisation ou du Cloud Computing ou de l'outsourcing. L'infrastructure de gestion de la sécurité doit permettre de collecter et gérer les données de sécurité à l'échelle de l'entreprise et d'évoluer pour répondre aux besoins de celle-ci, aussi bien d'un point de vue physique, qu'économique. L'infrastructure doit de surcroît pouvoir s'adapter aux nouveaux environnements et évoluer rapidement et conjointement pour prendre en charge l'analyse de toutes les nouvelles menaces.

Des outils d'analyse prédictive et de visualisation : le Big Data n'est pas uniquement, comme nous avons pu le voir, un amas de gros volumes de données. Il nécessite une analyse prédictive bien plus avancée pour identifier en amont les menaces de sécurité. Ces outils sont alors les éléments nécessaires aux

professionnels de la sécurité pour des analyses spécialisées. Si les responsables n'auront besoin de ces outils que pour avoir un aperçu général des tendances, les analystes eux, les utiliseront pour détecter de fichiers suspects et pour automatiser les tests de ces derniers et détecter les événements qui ont eu lieu sur l'ensemble des réseaux pour ainsi reconstruire l'ensemble de l'information.

Enfin, des renseignements sur les menaces : cet aspect revêt la nécessité d'appliquer des techniques d'analyse de données sur les informations collectées. L'entreprise doit pouvoir visualiser l'environnement des menaces externes afin de pouvoir établir la corrélation avec les informations collectées au sein même de l'entreprise. Cette corrélation est essentielle pour permettre aux analystes de bien connaître les indicateurs de menace actuels et savoir ainsi où investiguer.

3. Entre création de valeur et éthique, l'entreprise doit maintenir ses Best Practice face au risque d'atteinte aux Libertés individuelles

Comme nous le verrons plus tard dans notre analyse, le Big Data a participé de la redéfinition de la stratégie d'entreprise, en créant de nouveaux business modèles. Les entreprises ont en effet intégré l'importance de la création de valeur à travers l'exploitation du Big Data. Créer de la valeur à travers la collecte de données n'est en soi évidemment pas une chose nouvelle. La donnée a toujours servi le marketing et les analyses de marché et plus précisément du consommateur. Néanmoins, ce qu'apporte de nouveau le Big Data est non seulement de fournir de l'information d'une qualité sans précédent et actualisée mais surtout de procéder à la réactualisation continue et automatique de celle-ci.

Les données générées par l'utilisation des Smartphones, les services d'accès à internet sans fil (réseau Wifi ou 3G) a permis aux consommateurs d'avoir accès à des services multiples peu importe où ils se trouvent. Aujourd'hui, selon des données fournies par le cabinet Forrester qui qualifie la croissance des appareils électroniques tactiles de « *spectaculaire* », le nombre de possesseurs de tablettes en Europe va quadrupler d'ici à 2017 tandis que le nombre de Smartphones en activité dans le monde dépassera cette année le milliard d'unités. Les tablettes rencontrent le plus de succès auprès des 18-24 ans : un individu sur quatre dans cette tranche d'âge en possède une. mais ce ne sont pas les seuls. Dans le rapport du cabinet, Thomas Husson, analyste chez Forrester et co-auteur du rapport du cabinet Forrester, affirme que « *les tablettes sont principalement utilisées dans les foyers connectés et permettent des expériences davantage partagées que les Smartphones – par nature plus mobiles et plus personnels* ».

Enfin, selon des estimations, la téléphonie mobile devrait atteindre la barre des quatre milliards de clients en 2018, soit la moitié de la population mondiale ! Ces chiffres toujours plus impressionnants les uns que les autres tendent à rendre compte du fait que, sans s'en rendre compte, des millions de consommateurs laissent des quantités impressionnantes de données personnelles dont la géolocalisation.

Big Data/Big Brother est devenue ironiquement aujourd'hui la nouvelle comparaison utilisée pour mettre en exergue le phénomène de l'exploitation de la donnée personnelle. Mine d'or pour l'entreprise, les informations sur son consommateur final sont précieuses et tous les moyens sont mis en œuvre pour collecter le maximum de données, justes, précises et utiles au business model de l'entreprise pour créer de la valeur. Du moteur de recherche au réseau social en passant par le commerce en ligne, la vie *privée* de l'utilisateur-consommateur s'étale sur un ensemble de bases de données innombrables. Nombre d'informations sont délivrées délibérément par l'internaute, notamment lors d'une inscription, avec au minimum le nom, l'adresse e-mail et un mot de passe, mais aussi souvent l'âge et le sexe. Les sites recueillent également, automatiquement l'adresse IP qui identifie l'internaute, son type d'ordinateur et la version du navigateur web.

En France, la loi Informatique et libertés encadre cette récolte : un site peut "collecter ces informations pour des finalités déterminées, explicites et légitimes" mais elles ne peuvent "pas être traitées ultérieurement de manière incompatible avec ces finalités. [Seul] un traitement à des fins statistiques ou à des fins de recherche scientifique ou historique est considéré comme compatible." Judicaël Phan, juriste au sein de la Commission nationale de l'informatique et des libertés (Cnil) affirme, en substance, que si une société souhaite monétiser sa base de données, elle doit agir en toute transparence, c'est-à-dire informer les personnes. Face aux accusations de livrer la vie privée en pâture aux sociétés de marketing, Facebook fait désormais de principal accusé. Le réseau social aux 955 millions d'utilisateur collecte nombre d'informations personnelles et Julian Assange, lui même dénonçait "la plus terrible machine à espionner jamais inventée". Google se retrouve aussi régulièrement au rang des accusés. Le géant du web collecte, via ses différents services de recherches, d'e-mails ou de vidéos, une montagne d'informations sur les habitudes des internautes. Le plus impressionnant étant "l'historique web" qui compile toutes vos recherches sur Google. Google affirme "ne communiquer des données personnelles à des entreprises, des organisations ou des personnes [le] consentement [de l'utilisateur]. Au-delà de l'intérêt marketing, ces informations "publiques" intéressent de près les autorités. En novembre dernier, l'agence de presse AP a révélé que la CIA, l'agence centrale de renseignement américaine, a créé une cellule qui s'intéresse aux réseaux sociaux. C'est ce qui nous amène alors à aborder la loi américaine faisant suite aux événements du 11 septembre 2001, adoptée le 25 octobre 2001, portant le nom de USA Patriot Act.

Les enjeux du USA Patriot Act

Considéré par de nombreux analystes et défenseurs des droits de l'Homme, comme l'avènement des prédictions de Georges Orwell, auteur du célèbre roman *1984* dont la figure principale, Big Brother, est devenue une figure métaphorique du régime policier et totalitaire, de la société de la surveillance, ainsi que de la réduction des libertés, la loi du Patriot Act érigée par l'administration Bush, renouvelée par le Président actuel des Etats Unis, Barack Obama, s'inscrit dans un véritable contexte de terreur politique et de guerre où le pouvoir administratif a incontestablement empiété sur le judiciaire. Cette loi, au départ considérée comme une loi antiterroriste permet au gouvernement des Etats unis de détenir sans limite et sans inculpation toute personne soupçonnée de projet terroriste. Cette loi dont la structure s'organise en plusieurs « axes » comporte notamment les titres suivants :

Titre I: « Procédures augmentées de surveillance » ;

Titre VII : «Partage accru des informations pour la protection d'infrastructures critiques », car en effet, la loi autorisera alors, quelques temps plus tard, la mise en écoute de tout appareil de communication utilisé par toute personne en rapport de près ou de loin, avec un présumé terroriste.

Le titre V intitulé « Enlèvement des obstacles sur l'investigation dans le terrorisme », réforme en particulier le Foreign Intelligence Surveillance Act de 1978, en permettant la coordination des opérations entre les agences de renseignement chargées de la sécurité extérieure et les agences chargées de la sécurité intérieure. Les « lettres de sécurité nationale » ou NSL qui sont des sortes de mandats délivrés directement par le FBI et sans supervision judiciaire, obligeront alors les groupes privés, et l'occurrence les fournisseurs d'accès à internet, à donner accès au FBI à l'ensemble de leurs bases de données personnelles non plus restreintes aux personnes soupçonnées de terrorisme ou d'espionnage mais à l'ensemble des américains « ordinaires ». Aujourd'hui ces bases de données concernent des dizaines de milliers de citoyens ordinaires et selon les nouvelles dispositions, les agents auraient le droit d'accéder aux bases de données publiques même s'ils ne mènent pas une enquête spécifique mais uniquement dans le but de la pêche à l'information du numérique.

Les Etats-Unis, acteur dominant de l'Internet mondial, se veulent porte-parole de la liberté d'expression sur le Net. Pourtant le 11 septembre 2001, leur législation antiterroriste empiète de plus en plus sur les libertés individuelles des internautes.

Les inquiétudes engendrées par le USA Patriot Act ont été en croissance constante chez les responsables informatiques européens car il donnerait au gouvernement américain un accès sans entrave à leurs données s'ils sont stockés sur les serveurs Cloud d'Américain fournisseurs. L'anxiété s'est renforcée lorsqu'un directeur général de Microsoft Royaume-Uni a admis qu'il ne pouvait pas garantir que les données stockées sur les serveurs de la société, même en dehors des États-Unis, ne seraient pas saisies par le gouvernement américain.

Cette loi met bien en évidence les dérives d'une sur collecte et sur exploitation des données personnelles d'un individu donné. En effet, l'individu est au centre de cette problématique d'atteinte aux libertés individuelles car il en devient l'objet 'expérimentation. Si le Big Data comprend un ensemble de «renseignements personnels» à lors les mêmes règles en vertu de la Loi sur la protection des renseignements personnels 1988 s'appliquent à la collecte de data par l'entreprise, l'utilisation et la divulgation de toute autre "renseignements personnels. Si ces données correspondent à des informations délicates (telles que les informations relatives à la santé) ou les données de télécommunication les mêmes règles s'appliquent également. Par conséquent, il est extrêmement important, lors de l'élaboration et de la mise en œuvre d'une stratégie Big Data de considérer et de comprendre l'ensemble des implications de la détention des données recueillies et déterminer ou sont les limites de cette exploitation lorsqu'elle remet en cause le respect de la vie privée.

Les «Renseignements personnels » correspondent à une information ou une opinion sur une personne identifiée ou dite raisonnablement identifiable. Cela peut inclure, par exemple, des données qui sont rassemblées autour d'un identifiant unique qui se rapporte à un seul individu, mais ne révèle pas le nom de l'individu.

Le champ d'application de la nouvelle Loi sur la protection de l'individu qui restreint les utilisations de Big Data signifie que les entreprises, dans un souci de répondre aux Best Practice, doivent adopter une approche de gestion des risques à leurs stratégies Big Data. Les analystes en marketing se doivent d'envisager de réduire au minimum la quantité de données qu'ils collectent. Dans le meilleur des cas, elles doivent utiliser des données agrégées ou rendues anonymes pour obtenir, à partir d'analyses, des données qui servent la stratégie. Une approche de minimisation des données aidera les entreprises à éviter la collecte de données pour lesquelles ils ne peuvent pas justifier une utilisation, et de réduire le risque de la violation des données. En effet, il est essentiel pour l'entreprise, de placer le client-consommateur au centre de ses préoccupations. Comme le présente l'ADMA, Association for Data-driven Marketing & Advertising, dans son *Best Practice guidelines for Big Data*, un des leviers serait d'envisager la réalisation d'enquêtes «d'attentes de la clientèle» sur les questions de collecte de données et l'utilisation. Cela permettrait de clarifier ce que les clients admettent comme « donnée détenue acceptable ».

Dans un souci de respect des libertés individuelles et d'atténuer le risque que le client s'oppose à l'utilisation de ses données personnelles d'une part, et dans leur propres intérêts d'autre part, les entreprises ont intérêt à peser les risques qui pèsent sur leur réputation, le risque d'amendes civiles ou autres sanctions en vertu de la Loi sur la protection des renseignements personnels. Une approche centrée sur le client est utile pour aider les entreprises à déterminer à agir sur les données acquises grâce à des analyses. Permettant ainsi d'atténuer le risque que le client remette en cause l'atteinte à la vie privée. Par ailleurs il est un impératif de s'assurer que les données sont exactes, et que des idées issues des données ne sont pas utilisées de façon à nuire le consommateur individuel final.

II. Un phénomène qui implique par conséquent de nouveaux enjeux commerciaux entre les différents acteurs du marché et un nouveau modèle organisationnel

A. De nouvelles relations commerciales

1. Le marché du Big Data

Le marché des Big Data représentera en 2016 près de 24 milliards de dollars, il connaît une croissance fulgurante de près de 30% annuel depuis 2011. C'est l'un des marchés les plus porteurs de cette décennie. Qui sont donc les acteurs présents sur ce marché?

- Le marché des fournisseurs :

Il s'agit d'un marché constitué de produits et services en lien avec le Big Data, c'est-à-dire les serveurs, les logiciels et le stockage (qui représente la majorité du marché). Les technologies privilégiées du Big Data sont les bases de données objets ou graphiques, l'indexation de contenus (ou encore les moteurs de recherches avancés) et les technologies en mémoire.

Le marché est composé de pionniers des solutions IT comme par exemple IBM, Oracle, SAP, de spécialistes des solutions data et Big Data tels que MapR, Teradata, des acteurs d'internet comme Google, Amazon ou Facebook (qui prospèrent grâce aux données personnelles) et enfin des intégrateurs comme par exemple Atos, Accenture ou Cap Gemini. Pour répondre aux mieux aux attentes clients, ces fournisseurs créent des partenariats entre eux et leur proposent à la fois l'infrastructure et les applications.

Les réseaux sociaux ont contribué à l'explosion des données informatiques. Facebook, Twitter ou encore LinkedIn génèrent des milliards de données via la mise en ligne de photos ou de commentaires et représentent un nouveau type de communication qui s'est propagé en entreprise. C'est l'internaute qui crée du contenu et qui crée donc de la donnée sur le Web. Ces informations peuvent être vendues aux utilisateurs, aux entreprises qui vont les utiliser en marketing par exemple.



Les premières entreprises qui ont utilisé les technologies du Big Data sont les moteurs de recherches (Google, Amazon, Yahoo) car elles avaient à gérer de gros volumes de données à traiter en temps réel. C'est pour cela qu'elles ont mis à disposition des outils de gestion en Open source qui permettent d'analyser et de restituer des données très rapidement. Cela a été possible grâce au Cloud computing qui permet de stocker de gros volumes de données.

Yahoo a créé Hadoop, qui est le modèle le plus utilisé par les acteurs du Big Data. C'est l'emblème par excellence du Big Data. Hadoop permet de traiter tous types de données en grand nombre, celles-ci ne sont pas triées en fonction de leur relation entre elles. Les données sont collectées depuis diverses sources puis elles sont distribuées, c'est pourquoi on parle d'une librairie de l'information.

Google a produit sa propre technologie Big Data appelé Map Reduce. Il s'agit d'un mode de calcul permettant de traiter de grandes quantités de données en les distribuant dans des clusters. Cette solution est souvent associée à Hadoop et est très appréciée des sociétés qui possèdent de grands centres de traitement de données comme Facebook.

Oracle, leader du domaine des bases de données a développé sa propre solution NoSQL (Not only SQL). Ce terme est utilisé lorsque des requêtes sont effectuées dans un autre langage que SQL. NoSQL permet de traiter un grand nombre de données non structurées, ce que ne permet pas SQL jugé trop stricte. Ce modèle est utilisé par des sites Web de grande audience comme Google ou Amazon qui doivent manipuler des quantités énormes de bases de données.

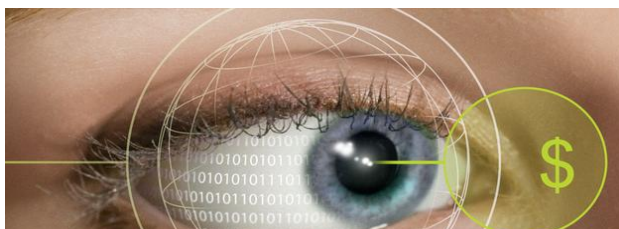
Ces technologies sont aujourd'hui accessibles gratuitement grâce à Hadoop qui propose un écosystème en version libre.

- Le marché des utilisateurs :

Les utilisateurs du Big Data sont assez peu nombreux aujourd'hui en France, environ 10% des entreprises seulement contre 30% en Grande Bretagne et 90% aux Etats-Unis. Ce qui empêche les entreprises de se lancer est le coût de l'investissement en Big Data, le manque de ressources (humaines notamment) et d'organisation. Ce sont les secteurs de la finance, des télécoms (qui connaît un fort succès grâce aux données mobiles de géolocalisation) et de la grande distribution qui utilisent le plus le Big Data.

2. Nouvelle relation B to B : Les données en vente

Ce qui peut sembler inoffensif, des informations qui pourraient paraître sans valeur : des achats, des préférences musicales, des vacances programmées, sont en réalité scrutées par des informaticiens à l'affût de ces informations qu'ils pourront vendre. Un numéro de téléphone portable, une adresse électronique privée. Les



entreprises qui collectent le plus de données comme Facebook, Google ou Amazon disposent de cette quantité énorme de données, qu'ils stockent et

grâce à laquelle ils assurent leur richesse. Et une fois ces données stockées par l'une de ces compagnies, il est impossible de savoir où elles finiront, qui les achètera. La navigation sur le Web et les Smartphones laissent des traces (cookies), qui piste les habitudes des utilisateurs. Les entreprises sont dans leur droit lorsqu'elles nous espionnent, c'est légal et très lucratif d'où l'apparition, ces dernières années, d'un commerce mondial qui se consacre à trouver des nouvelles façons d'extraire et d'analyser ces informations.

Des « courtiers en données » collectent, exploitent cette masse d'information personnelle, les analysent et les vendent ensuite comme des matières premières en ligne ou hors connexion. Acxiom par exemple, l'une des plus grosses sociétés spécialisée en données client, analytique et service marketing, gère les informations de pas moins de 500 millions de consommateurs à travers le monde et réalise un chiffre d'affaire de plus d'un milliard de dollar annuel. Les réseaux sociaux comme Facebook, qui génère des milliards d'information via les photos, noms, mentions j'aime, commentaires, travaillent en étroite collaboration avec ces courtiers en leur renvoyant la mise à jour des statuts contre des codes postaux ou des données de carte de fidélité. Selon une étude du Boston Consulting Group, Facebook tire en moyenne 5 dollars par profil, et la vie d'un Européen vaudrait aujourd'hui plus de 600 euros (services gratuits mais qui ont un impact sur l'économie).

Les entreprises peuvent donc se procurer pour un prix considérable ces tweets et posts Facebook, les faire analyser et réaliser des ventes grâce aux résultats des analyses. Sony a par exemple payé 200 millions à une société spécialisée dans la donnée au titre d'honoraires.

3. Comment La DSI des entreprises choisit-elle Les outils adéquats parmi Les nombreuses offres disponibles ?

De nombreuses offres sont disponibles sur le marché, les fournisseurs proposent des offres variées qui reposent sur le modèle Hadoop mais qui sont plus faciles à intégrer dans l'entreprise, qui sont directement utilisables sans avoir à les programmer. La majorité des DSI optent pour ces solutions externes plutôt que de développer eux-mêmes leur propre infrastructure car l'investissement en temps et en personnel est trop lourd. Celles-ci doivent quand même se renseigner sur l'offre la plus adéquate selon leurs besoins : Faut-il privilégier la rapidité de traitement ou le traitement de gros volumes ? L'intégration des outils doit relever de la stratégie d'utilisation des données définie par l'entreprise.

Pour mettre en place ces nouvelles architectures, les fournisseurs proposent aux utilisateurs des infrastructures légères et qui s'adaptent facilement à leur infrastructure existante pour une utilisation immédiate. En ce qui concerne le stockage de données, les utilisateurs préfèrent privilégier le Cloud pour avoir plus de souplesse dans l'infrastructure et pour regrouper les données.

Le Big Data est un atout pour les entreprises, il leur permet d'analyser un grand nombre des données commerciales (données de navigation et d'achat laissées par les appareils électroniques connectés à internet pour n'en citer qu'un) et

de production. Ces analyses vont aider à la prise de décision et permettre le déclenchement d'actions marketing ciblées pour fidéliser ou adapter les produits et services en fonction des comportements des consommateurs

⇒ Quel est donc l'impact de l'intégration du Big Data sur les entreprises ?

B. Comment accompagner le changement dans l'organisation de l'entreprise en mutation

Comme évoqué précédemment, le Big Data constitue une mine d'or pour les entreprises, encore faut-il arriver à le manier. Les entreprises ont donc tout intérêt à revoir leur organisation. Plus précisément, elles doivent intégrer de nouvelles compétences liées au Big Data, gérer l'apparition de nouveaux métiers, revoir la structure de leur organisation et faire face aux ressources manquantes.

L'intégration du Big Data et toutes les nouvelles technologies qu'il nécessite entraînent de profonds changements en termes de gouvernance et de mode de fonctionnement. Le Big Data devant récupérer des données présentes dans différents services de l'entreprise pour établir des processus d'aide à la décision, la stratégie de l'entreprise sera donc tournée vers cette donnée qui est au cœur des attentions.

En effet, l'entreprise doit faire face à l'arrivée de nouveaux acteurs, à une concurrence plus forte, à la montée en puissance des réseaux sociaux ainsi qu'à une modification de la relation avec le client. Si elle ne s'adapte pas à tous ces nouveaux facteurs, elle risque de perdre en productivité et d'être moins compétitive car elle n'aura pas su proposer une offre adéquate à chaque client et donc pas répondre au mieux à leurs besoins. Elle risque également de dégrader son image auprès de sa clientèle.

1. Un renforcement des relations professionnelles entre informaticiens et « marketers »

On a assisté à une transformation de la relation client/marketing avec l'apparition d'internet et des réseaux sociaux. En effet, les clients peuvent émettre leurs besoins aux entreprises rapidement. De ce fait, on assiste à un renforcement des relations professionnelles entre le service informatique et le département marketing qui sont les acteurs principaux du Big Data au sein de l'entreprise. Désormais ces derniers doivent arriver à faire davantage de synergies afin que les données soient plus facilement traitées et communiquées. L'équilibre entre les créatifs et les ingénieurs est devenu primordial, d'où une nouvelle relation entre le département informatique et marketing.

Le rôle du département informatique a fortement évolué et pris davantage d'importance dans la relation client grâce notamment aux bases de données relationnelles. Il doit collecter les données provenant des Business Units et les traiter. Il est également possible que les informaticiens soient intégrés au sein même des

équipes opérationnelles afin de réduire le cycle de décision. Il est nécessaire pour les équipes informatiques de parvenir à proposer des cycles plus courts et plus orientés marché ainsi que des interfaces et procédures qui leur conviennent et leur sont accessibles. Tout cela en répondant aux besoins et évolutions des équipes marketing et en préservant la fiabilité du système opérationnel.



Le département marketing lui est devenu essentiel dans toutes les grandes entreprises. Il permet de se développer et de faire face à la concurrence en répondant aux mieux aux besoins des clients et en les fidélisant. Les équipes marketing devront assimiler les outils fournis par l'informatique pour faire les requêtes dont elles auront besoin. Pour ce

faire, le service marketing doit compter sur des applications informatiques adaptées qu'il aura demandé à la DSI. De plus le service marketing, et les autres services opérationnels en général, ont tout intérêt à s'impliquer dans le traitement et l'analyse de leurs données car cela leur donne une valeur ajoutée et une double compétence. Il est donc essentiel que ces deux services coopèrent au mieux afin de solidifier la synergie entre ces équipes.

2. Evolution du rôle des DSI

Les Directions des Systèmes d'Information (DSI) sont les piliers du Big Data dans l'entreprise, c'est sur elles que la mise en place du Big Data repose. Et ce sont elles qui favorisent les échanges entre les différents départements. Elles doivent désormais savoir faire preuve de conviction et être force de proposition envers les opérationnels de l'entreprise tout en améliorant leur réseau technologique. Cela dans le but d'améliorer la performance de l'entreprise.

C'est la DSI qui détient la donnée utile au marketing, elle joue un rôle de logisticien de l'information.

3. Plusieurs types de « data-driven organization »

Il existe plusieurs schémas d'organisation au niveau de l'intégration du Big Data en entreprise. Il n'y a pas d'organisation type car les entreprises n'ont pas toutes les mêmes besoins et elles ont des manières différentes d'utiliser la donnée. Avant de choisir un schéma d'organisation elles doivent réfléchir longuement sur les outils Big Data qu'elles souhaitent mettre en place et faire un bilan sur leur besoins et ressources.

- *L'option décentralisée* : le Big Data est mis en place métier par métier. La DSI aide un département (marketing par exemple) à intégrer des outils à ce

département. Dans ce schéma ce sont les opérationnels les analystes de données d'où leur double compétence.

- *L'option centralisée* : une direction est créée au sein de la structure dédiée au Big Data afin de regrouper toutes les données, les traiter et les exploiter pour en fournir des résultats aux opérationnels et à l'informatique. Des spécialistes sont ainsi réunis dans une même direction ce qui représente un avantage certain.
- *L'option de la sous-traitance* : il s'agit pour une entreprise de faire faire effectuer le traitement et l'analyse de données à une entreprise spécialisée dans le Big Data. Cela lui permet d'avoir plus de flexibilité et d'échanger avec des professionnels expérimentés qui disposent de toutes les compétences nécessaires. Cependant, la prise de décision est plus longue car le prestataire est externe à l'entreprise donc il met plus de temps à fournir les données aux opérationnels.

4. Nouvelles compétences liées au Big Data

Les managers et directeurs se doivent d'avoir conscience de la valeur que représentent les données, ils ont donc intérêt à les exploiter en renforçant les compétences métiers.

Le Big Data a profondément modifié la manière d'analyser les données et de ce fait causé de nouveaux besoins en termes de compétences informatiques. Ces dernières font face à la complexité de certains supports tels que Hadoop ou Mapreduce. Il est donc indispensable pour une entreprise que son service informatique ait les compétences techniques nécessaires à la compréhension et à l'analyse des données fournies par ces supports. Faire appel à des prestataires de service spécialisés en informatique ne suffit plus car le Big Data étant déployé à plusieurs départements de l'entreprise les informaticiens doivent être présents pour répondre aux besoins des Business Units.

Le Big Data implique également que les Business Units puissent traiter et exploiter les données avec une vision propre à leur métier. Dans le service marketing par exemple qui est de plus en plus concerné par les data, les « marketers » souhaitent avoir des compétences informatiques pour analyser et se servir des data qu'on leur aura fourni. Ils se forment donc à cette nouvelle compétence qui leur apporte une valeur ajoutée et leur permet d'exercer au mieux leurs fonctions et répondre aux besoins clients.

Avec l'évolution des algorithmes et statistiques qu'entraîne le Big Data, les entreprises se sont vues « obligées » d'embaucher de nouveaux profils comme des statisticiens ou des mathématiciens. Ceux-ci disposent de compétences qui leur permettent de comprendre les algorithmes, de les traduire, de traiter et d'analyser les données.

Il est possible d'ajouter une dernière nouvelle compétence liée au Big Data : la communication. En effet, les décisions qui sont prises grâce aux données reposent

sur l'aptitude pour une entreprise à représenter et à hiérarchiser les informations. La maîtrise de la communication permet cela.

Pour répondre à ces nouvelles exigences, les entreprises peuvent former en interne sur le Big Data, recourir à des organismes spécialisés, créer des partenariats avec des entreprises ou encore acquérir une startup spécialisée dans l'analyse de données. Elles doivent investir dans la donnée, dans le personnel, dans les outils et les process.

5. Les nouveaux métiers du Big Data

Face à la complexité et au volume des données que fournit le Big Data, les entreprises sont contraintes de mettre en place de nouvelles solutions analytiques, de se doter d'outil spécialisés et d'investir en ressources humaines.

Pour lancer un projet Big Data, les entreprises ne peuvent pas compter que sur le personnel déjà présent. Très peu d'entreprises y parviennent. Il faut du



personnel qualifié, qui a les compétences nécessaires et indispensables au Big Data. Il y en a trois fondamentales :

- *Informatique* : aptitude à programmer sur les nouveaux outils
- *Statistiques* : aptitude à innover et à modéliser
- *Business* : capacité d'interprétation et de transformation des indicateurs en langage opérationnel).

Les profils qui disposent de ces trois compétences sont des Data Scientist. Il s'agit d'un nouveau métier crée par le Big Data. Ces profils à la fois très recherchés et polyvalents sont très rares, c'est pourquoi, une multitude d'offre d'emploi sont émises, à des rémunérations les plus qu'alléchantes.

- *Data scientist* : ce métier crée par le Big Data peut être considéré comme une évolution professionnelle du poste de Data Analyst, de consultant en datamining ou de statisticien. Ces profils sont très recherchés à de très haute rémunération. Le Data Scientist ne doit pas seulement collecter les données et les analyser, il doit pouvoir les faire parler dans le but de sortir des indicateurs concrets qui serviront à la direction générale. Il doit également fournir des recommandations sur la manière d'utiliser les résultats de ces analyses de données. Le Data Scientist communique avec les managers et les responsables informatiques c'est pourquoi des qualités relationnelles sont fortement recommandées pour exercer ce métier. Avec la diversité de leurs missions, leur vision transverse de l'entreprise et leur niveau de compétence, les Data Scientist ont tout à fait leur place au comité de direction.

- *Data Analyst* : son rôle est d'aider la direction générale à la prise de décision grâce aux informations qu'il lui fournit. Le Data Analyst utilise des outils informatiques et des techniques statistiques afin d'organiser, de synthétiser et de traduire la masse d'informations qu'il doit traiter.
- *Responsable de la sécurité de l'information* : il a pour mission d'identifier tous les dangers et impacts pour l'entreprise liés aux informations. Il doit sécuriser toutes les données dont la structure dispose sans pour autant freiner son évolution technologique. Il met en œuvre la politique de sécurité de l'entreprise et garantit la disponibilité, la sécurité et l'intégrité des données.

Avec la restructuration de l'entreprise s'est posée la question de la structuration de la prise de décision et des remontées hiérarchiques auprès des dirigeants. Il s'agissait en effet de créer un poste de « porte-parole », une personne qui pourrait aider à la prise de décision en fonction des données traitées.

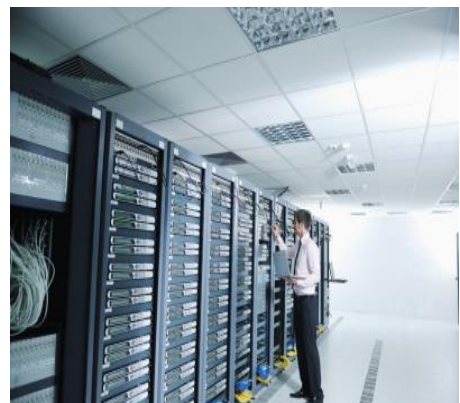
- *Chief Data Officer* : son rôle principal est de faciliter l'accès aux données et de garantir leur fiabilité lors de la prise de décision. Il est au cœur de l'entreprise et doit assumer une fonction transversale pour avoir une parfaite connaissance de celle-ci que ce soit au niveau des business unit ou des process.

Les métiers classiques tels que les administrateurs réseau, les ingénieurs de sécurité informatique et les ingénieurs système sont également concernés par la mise en place d'un projet Big Data. En effet, avec la multiplication du nombre de données se pose le problème de leur stockage, leur acheminement et leur sécurisation.

Les métiers liés aux données ne sont plus cantonnés aux directions informatiques, les nouveaux profils doivent avoir des compétences transversales. Mais faute d'arriver à recruter ces « hybrides », les entreprises se voient dans l'obligation de former le personnel déjà présent.

6. Challenge technologique du Big Data pour l'entreprise

Le Big Data impose aux entreprises d'investir dans des capacités de stockage et d'analyse. Dans des capacités de stockage car avec les anciens systèmes il était impossible de stocker autant de données dans les entrepôts de stockage et les dépenses en électricité nécessaires à leur fonctionnement étaient très coûteuses. La solution a donc été d'investir dans le Cloud computing pour



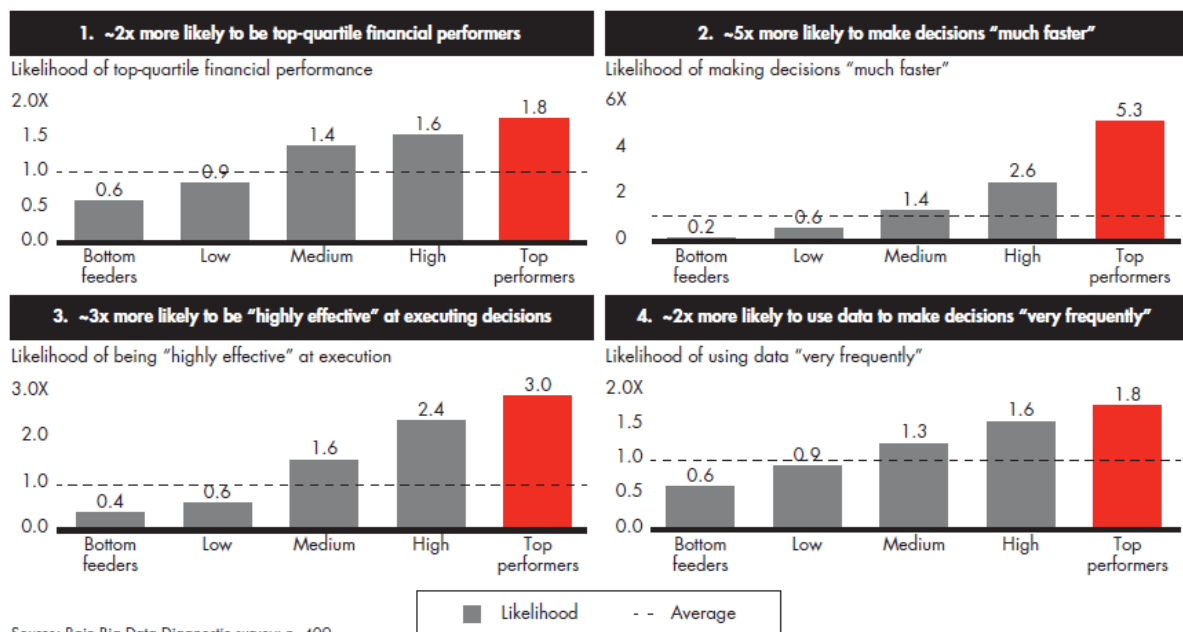
rationaliser les volumes de stockage, solution permise par le Big Data. Les entreprises ont également dû investir dans l'analyse, c'est-à-dire dans les techniques qui permettent d'analyser le croisement entre diverses sources de données. Il y a par exemple le Web Sémantique qui croise les données d'internet avec tout autre donnée du web ou le Datamining qui comporte le crowdsourcing (remontée de données via les réseaux sociaux), l'analyse de sentiments (recherche d'opinion grâce à une analyse textuelle), ou encore l'analyse de réseaux (identification de leaders d'opinions sur les réseaux sociaux).

C. Comment utiliser ces données pour améliorer la prise de décision et l'efficacité organisationnelle

Nous allons voir pourquoi et comment l'exploitation du Big Data peut améliorer la prise de décision et l'efficacité organisationnelle au sein d'une entreprise.

Comme nous l'avons vu précédemment, l'étude menée par Bain & Company sur 400 entreprises au chiffre d'affaires supérieur à 1 milliard de dollars, celles ayant des capacités d'analyse « Big Data » ont globalement de meilleures performances financières, une prise de décision plus rapide et mieux appliquée.

Figure 1: Companies with the best analytic capabilities outperform the competition



1. Des décisions prises en temps réel et plus adaptées

L'exploitation du Big Data permet à l'entreprise de prendre des décisions en temps réel mais aussi plus adaptées. En effet, une organisation et une analyse de pointe des données permet à l'entreprise d'avoir accès à des informations presque instantanées, valorisables dans la prise de décision.

Selon une étude menée par l'IBM Institute for Business Value et la Saïd Business School de l'Université d'Oxford, les organisations engagées dans des projets pilotes de Big Data déclarent mesurer 15 % de bénéfices supplémentaires de leurs informations (dont les Big Data) et de l'analytique par rapport à celles exploitant des fonctions d'analyse traditionnelle. Aussi, plus de 75 % des décisionnaires ont déclaré vouloir disposer des informations nécessaires à leur prise de décision sous 24 heures (contre une semaine auparavant).

Avant le début de l'ère Big Data, la prise de décision se réalisait en prenant en considération l'expérience et les données à disposition, c'est à dire des données historiques, antérieures. Ces données s'utilisaient comme un instrument de la prise de décision. Maintenant, les décisionnaires peuvent utiliser des données en temps réel, fiables et facilement accessibles, celles-ci devenant un élément à part entière de la prise de décision. Ceci permet à l'entreprise d'agir plus rapidement et efficacement. On parle "d'écosystème informationnel" (Lucile Hyon, IBM France).

Donnons l'exemple du service Ressource Humaine d'une entreprise. Il lui est maintenant possible de modéliser les comportements de ses employés à travers l'exploitation du Big Data (via l'utilisation d'algorithmes). Grâce à cette analyse, le service RH peut identifier toute éventuelle volonté de départ, l'anticiper grâce à un dialogue avec l'employé et réduire le turnover.



Autre exemple au sein du département logistique de Procter & Gamble aux Etats-Unis. L'entreprise détient 42 sphères de décision où se trouvent des écrans vidéo et ordinateurs ayant accès à une large base de données. Des équipes inter-fonctionnelles s'y réunissent afin de prendre des décisions transversales et en temps réel. Quand par exemple, le capteur intégré d'une machine Pampers dans une usine du Wisconsin indique que la machine a besoin d'une maintenance, un arrêt de l'usine est programmé. S'il s'avère que la machine doit être arrêtée pour une période longue, alors les équipes anticipent et prennent la décision de livrer Wal-Mart, non pas depuis l'usine en charge (Wisconsin) mais depuis celle d'Albany en Géorgie. Les capacités d'analyse de données qu'a mises en place P&G

permettent de réorganiser efficacement et rapidement les camions de livraisons afin de répondre à tout engagement client, même en cas d'imprévu.

2. Des outils « Big Data » adaptés au projet de l'entreprise

Pour que l'exploitation du Big Data amène à des décisions rapides et adaptées, l'entreprise doit se doter d'outils technologiques performants, capables de stocker, d'organiser et de mettre à disposition des utilisateurs des données pertinentes. En effet, comme nous l'avons vu précédemment, le Big Data ne peut être stocké et exploité dans des entrepôts de stockage traditionnels, c'est pourquoi il faut utiliser des outils plus adaptés tel que MapReduce, Hadoop ou le stockage dans le Cloud. Ces nouveaux outils « Big Data » s'appuient sur des technologies qui permettent de traiter des volumes d'information plus larges, de manière plus rapide et moins coûteuse. Par exemple, la technologie de « Text-Mining » (fouille de texte), permet à un département RH de trier automatiquement les profils correspondant à des critères spécifiques parmi des milliers de CV, économisant ainsi un temps précieux.

Donnons l'exemple de la firme américaine de services financiers Morgan Stanley. Elle s'est mise à utiliser Hadoop en 2010 car les entrepôts de données traditionnels n'étaient plus assez performants pour gérer la masse de données qu'elle générerait. Dans un article au Forbes, Gary Bhattacharjee, le Directeur exécutif des systèmes d'information de Morgan Stanley décrit les bénéfices que l'entreprise a su tirer du Big Data grâce à l'utilisation de Hadoop. Cet outil leur a permis d'améliorer leurs analyses de portefeuille. Ce qui prenait avant plusieurs mois peut maintenant être analysé en temps réel. Quand un événement de marché se produit, ils sont capables de comprendre l'impact en temps réel et d'avoir une trace sur qui, comment, quand et qu'est ce qui a causé le problème.



3. Une équipe de travail « Big Data »

Si les outils Big Data sont un élément essentiel à l'exploitation efficace des données, l'entreprise doit aussi se doter de professionnels compétents, capable

d'utiliser cette technologie, d'analyser les données et de travailler en équipe. Nous avons précédemment développés les particularités de ces nouveaux métiers.

Reprenons l'exemple de Procter & Gamble. L'entreprise s'est mise à embaucher des experts dans le domaine numérique, des développeurs, des experts en réseaux sociaux et des managers confortable avec les outils décisionnels quantitatifs. De cette façon, l'entreprise peut se tourner vers le Big Data de manière globale et efficace.

4. Des professionnels « Big Data » au cœur des décisions

L'entreprise doit s'assurer que ces professionnels « Big Data » aient une place stratégique au sein de l'organisation afin qu'ils prennent part aux décisions majeures.

En effet, pour que l'entreprise bénéficie de ce personnel qualifié, il faut qu'il soit écouté et intégré à la prise de décision. Il s'agit donc de déplacer le pouvoir de décision vers les décisionnaires sachant exploiter la donnée. Ces derniers pourront apporter une nouvelle vision et de nouvelles informations pertinentes à la prise de décision.

Cependant, lorsque les décideurs prennent en compte ces nouveaux paramètres, ils sont amenés à prendre des décisions différentes de celles qui seraient habituellement prises. Ceci risque d'entraîner des clivages au sein des décideurs, entre les conservateurs (réticents à l'exploitation du Big Data ou n'arrivant pas à s'adapter au changement) et ceux mettant l'exploitation du Big Data au cœur de la décision.

C'est pourquoi, un ou des intervenants doivent mener le changement et amener les employés à adopter un état d'esprit où la donnée est une ressource importante de l'entreprise. On retrouve ainsi dans les grandes entreprises le développement du poste de Chief Data Officers qui a autant d'importance que les postes de Directeurs des Ressources Humaines par exemple.

5. Une organisation à repenser dans son ensemble

Lorsque l'entreprise dispose d'outils « Big Data » et d'un personnel capable de les utiliser, elle détient les moyens pour conduire un changement mais il lui reste à mettre en place une stratégie cohérente. Si l'entreprise veut se tourner efficacement vers l'exploitation du Big Data, c'est toute son organisation qui doit être repensée.

Illustrons cet argument avec le "Star model" de Jay R Galbraith, 2010 (figure ci-dessous).

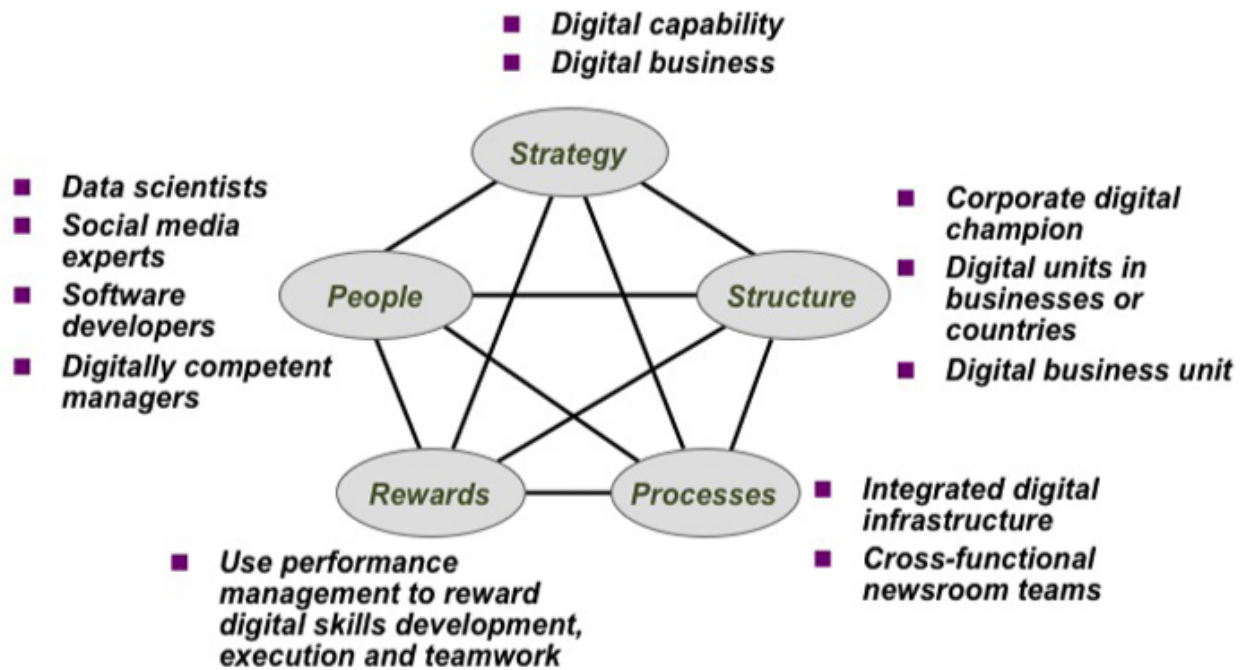


Fig. 4. Big Data's Impact on the Organization

L'exploitation du Big Data impacte les éléments majeurs d'une entreprise que sont: la stratégie, le personnel, la structure, la gestion de la performance et les processus.

Chacun de ces éléments est représenté dans ce modèle.

Au niveau stratégique, l'entreprise peut implémenter une stratégie visant à construire une « digital capability » afin de prendre des décisions rapides et efficaces et/ou implémenter une stratégie « digital business » pour créer une nouvelle source de revenus.

Afin de d'implémenter sa stratégie, l'entreprise doit modifier leurs structures organisationnelles. Il s'agit ici de mettre en place un département tourné vers l'analyse de données mené par un leader (Chief Digital Officer par exemple).

L'entreprise doit mettre en place les infrastructures et processus adéquats. Nous avons précédemment mis en avant l'exemple de P&G qui réunit des équipes inter-fonctionnelles au sein de ses centres d'analyse de données. Ici, P&G a su créer un "écosystème informationnel" qui exploite le Big Data et permet une prise de décision efficace.

Ensuite, comme nous l'avons précédemment vu, l'entreprise doit se doter de personnel compétent avec les outils « Big Data ». Afin de motiver les employés vers l'exploitation du Big Data, l'entreprise doit mettre en place un système de communication, gestion de la performance et récompense en fonctions d'objectifs Big Data.

Pour conduire le changement vers le Big Data, il faut donc repenser chaque élément organisationnel de l'entreprise.

6. Un changement qui doit être organisé et progressif

Après avoir repensé les éléments organisationnels de l'entreprise, la conduite vers le changement Big Data doit s'effectuer de manière organisée et progressive.

Tout d'abord les décideurs doivent s'informer et comprendre les enjeux du Big Data. Il faut ensuite identifier des objectifs atteignables et les données qui s'y rapportent. Pour atteindre ces objectifs, l'entreprise doit établir une stratégie en se posant les questions suivantes:

- Comment investir dans le Big Data?
- Quelles technologies d'informations adopter?
- Quelles formations et outils mettre en place pour que les employés se tournent vers l'exploitation de la masse de données efficacement?

Mais la question que l'entreprise ne doit pas rater est :

- Par où commencer?

Cette question amènera l'entreprise à définir un projet pilote. Il faut commencer sur un périmètre réduit pour maîtriser totalement le déploiement et le retour sur expérience. Cette stratégie doit être développée tout en assurant le choix d'outils et de technologies appropriées, le recrutement et/ou la formation d'une équipe de travail compétente et une coopération entre les acteurs impliqués.

Atteint ce stade, l'entreprise peut commencer la mise en place de son plan d'action à travers son projet pilote (préalablement défini). Lorsque ce dernier arrive à terme, l'entreprise doit tirer les conclusions des réussites et des échecs. Elle peut ensuite déployer son projet « Big Data ». Avec l'expérience, l'entreprise pourra ainsi ses compétences et méthodes d'analyse du Big Data à travers tous les départements de l'entreprise pour créer des synergies.

7. US Xpress, un exemple de « best practices »

Afin d'illustrer cette partie, nous allons mettre en avant l'exemple de US Xpress, entreprise américaine de camionnage qui a su



tirer profit du Big Data pour prendre de meilleures décisions et être plus efficace dans son organisation.

US Xpress a réussi à utiliser ses centaines de milliards de données issus de plusieurs milliers de capteurs placés sur ses camions et donnant par exemple des informations sur les pneus, la consommation d'essence ou la géolocalisation. En exploitant cette masse de données, la firme a pu réaliser une économie de 6 millions de dollars à l'année.

Comment a-t-elle fait?

En 2009, toutes ses informations étaient disséminées à travers 130 différentes bases de données. Les employés avaient même réussi à y écrire Wal-Mart de 178 différentes façons. Les employés devaient vérifier les données sur différents ordinateurs et lancer une requête pouvaient prendre des semaines voir des mois. En d'autres termes, leurs données étaient désorganisées et inutiles. US Xpress se devait d'y mettre un terme.

En premier lieu, la firme a décidé de nettoyer et d'organiser ces données. Pour cela, elle a combiné cette masse de données dans une seule et même interface : le système DriverTech. Ceci a permis de traiter, d'analyser et d'obtenir des données pertinentes en temps réel.

Avec cette donnée disponible en temps réel, US Xpress est devenu capable de comprendre comment les chauffeurs conduisaient, pourquoi certains camions étaient arrêtés alors que le moteur tournait toujours, comment ils pouvaient réduire la consommation d'essence ou même où les incidents avaient lieu. L'analyse géo-spatial qu'a mis en place US Xpress lui permet de surveiller en temps réel ce qu'il se passe et minimiser tout temps d'arrêt. Dès qu'un camion arrive à l'entrepôt pour maintenance ou chargement, la firme a l'information en direct.

US Xpress a réalisé une économie de 20 millions de dollars sur la consommation d'essence la première année en identifiant les chauffeurs conduisant au ralenti et en menant ainsi des actions correctives.

Bien sûr, les données issues des capteurs placés sur les camions ne sont pas les seules données que collecte US Xpress maintenant puisque grâce à l'interface DriverTech, les chauffeurs peuvent discuter sur des blogs directement dans leur camion. Grâce à cette interface, l'entreprise est informée en temps réel de tout problème.

Les managers contrôlant un parc de véhicules ont aussi un iPad qui leur donne accès en direct aux informations dont ils ont besoin.

Ainsi, pour passer d'un système de gestion des données lent et désorganisé à un mode de fonctionnement basé sur l'analyse de données pertinentes et organisées, US Xpress avait mis en place une stratégie de 36 mois. Cette stratégie comprenait 13 différents projets phares qui ont été implémentés les uns après les autres. L'entreprise a utilisé le succès du premier projet pour construire les suivants et déployer l'exploitation du Big Data au sein de l'entreprise, ceci de manière progressive.

La conduite du changement qu'a menée US Xpress pour se tourner vers le Big Data a été un succès puisque l'entreprise a même été récompensée de prix technologiques comme le Ventana Leadership Award in 'Overall IT Leadership'.

III. Un phénomène qui amène les entreprises à repenser leur stratégie

A. Comment utiliser les Big Data pour améliorer l'expertise client ? Efficacité de l'activité des entreprises.

1. Les Big data, des informations précises sur Le client

Nous avons donc pu voir que les Big Data étaient un ensemble de données très complexes et peu structurées. Pourtant, parmi cette masse d'informations se cache une réelle mine d'or pour les entreprises. En effet, en étant correctement exploitées, les Big Data permettent de dresser un portrait très précis des clients, pouvant ainsi aider les entreprises à établir une stratégie marketing qui répond correctement à la demande des consommateurs en l'anticipant.

Tout d'abord il existe les données structurées qui sont des informations recueillies par les entreprises elles-mêmes afin d'établir le comportement des consommateurs. Ainsi dans les magasins de grande distribution, le parcours exercé par les clients pendant leurs achats est étudié pour connaître leurs habitudes de consommation. De la même manière, lorsqu'un consommateur passe commande auprès d'un site en ligne, chacune des informations à propos de celui-ci sont minutieusement enregistrées par l'entreprise pour alimenter leurs fichiers clients. Ainsi, tout est sauvegardé : le prénom, le nom, la date de naissance, l'adresse, le panier moyen et le moment de l'année à laquelle le client a passé commande. Ces informations permettent à l'entreprise d'obtenir un profil très précis des consommateurs.

Cependant, afin d'établir un profil bien plus complet, l'entreprise a également besoin des données contenues dans les Big Data qui vont établir les habitudes et le mode de vie des ménages. Ainsi, dans un monde où le numérique occupe une place toujours plus importante, où l'objet connecté est omniprésent et où le partage d'informations est toujours plus important via la multiplication des réseaux sociaux, il est possible d'étoffer davantage le profil de chaque consommateur. Les Big Data

contiennent donc des données sur les habitudes des consommateurs à travers les recherches menées sur internet. La recherche d'un voyage, d'un appartement, de cadeaux de noëls sont tout autant d'informations qui vont être analysées par les entreprises. De la même manière, Facebook représente une véritable mine d'or sur les goûts des consommateurs. En effet, les Big Data contiennent également les statuts postés, les pages aimées, les liens partagés ainsi que les conversations de chaque personne, soit des données très précises et très personnelles sur les consommateurs. Cela permet aux entreprises de véritablement cerner les goûts et les attentes des ceux-ci.

Avec la popularisation d'internet, il est beaucoup plus facile pour chaque personne d'avoir accès aux avis des autres consommateurs avant d'avoir recourt à l'action d'achat. En effet, on voit apparaître une multiplication de sites de consommateurs pour que chacun puisse donner son impression sur un produit. Ainsi, la marque n'est plus le premier critère de sélection pour une grande partie des ménages. La demande des consommateurs a changé. Pour palier à cette situation et faire prospérer sa clientèle, il convient donc pour les entreprises de se démarquer quant à leur manière d'aborder le client. En effet, il faut non seulement être capable de fidéliser ceux qui ont déjà acheté tout en attirant de nouveaux. Cela passe alors par une offre qui ressemble au client. Ainsi, en combinant leurs données de fichier client avec celles des habitudes des consommateurs, les entreprises peuvent obtenir un profil très précis de leurs clients et leurs proposer une offre personnalisée. Ainsi, en cernant davantage les habitudes et les attentes des consommateurs, les entreprises sont donc capables d'anticiper leurs besoins. Une entreprise, plus à l'aise que les autres sur ce type de fonctionnement pourra donc largement tirer son épingle du jeu en proposant une offre adaptée à sa clientèle.

2. Un marketing transformé

Le domaine de l'entreprise le plus impacté par le Big Data est donc logiquement le marketing qui est au cœur de l'analyse de comportement du consommateur. En effet, les Big Data permettent d'obtenir rapidement des informations très précises sur les habitudes de ces derniers, générant donc un gain de temps en mettant fin aux enquêtes de consommation qui sont relativement longues et dont les données sont limitées. Or, il s'agit là d'informations capitales pour cerner le consommateur (ses goûts, ses envies) afin de tenter de répondre le mieux possible à sa future demande. En connaissant les habitudes du consommateur, les entreprises sont donc plus à mêmes de créer une offre lui ressemblant et répondant à ses envies.

L'Entreprise, consciente des attentes du consommateur va donc lui proposer une offre en accord avec celles-ci. On voit donc se développer une nouvelle stratégie marketing appelée « one to one » qui permet d'offrir aux consommateurs une offre personnalisée. Cette stratégie permet aux entreprises de fidéliser leurs clients actuels tout en attirant de nouveaux. En effet, à travers une offre personnalisée, le consommateur va avoir l'impression que ses attentes ont parfaitement été prises en compte par l'entreprise et va se sentir « important », ce qui va augmenter sa

YVES ROCHER

0€ LIVRAISON OFFERTE SANS MINIMUM D'ACHAT !

La star de l'été, c'est vous !

MES CADEAUX

LA CAPELINE + LES LUNETTES DE SOLEIL

EN CADEAU

OU

LES 3 MINI PRODUITS

EN CADEAU

Je découvre tous mes cadeaux !

LES BEST-SELLERS DE L'ÉTÉ

Cultes !

5€ max

Je craque !

LE SECRÉT DES VAHINÉS

MONOI DE TAHITI

À PARTIR DE 1€60

J'en profite !

20% 30% 50% 10% Voir toutes les PROMOTIONS

f J'aime !

Livraison Relais Colis Paiement sécurisé 25% Frais de port gratuits dès 29 € d'achat Satisfait ou remboursé Livraison en 48 heures

YVES ROCHER RÉCOLTANT, FABRICANT ET DISTRIBUTEUR DEPUIS 1959.

satisfaction. Cela lui permet de se sentir comme une véritable partie-prenante de la société. Ainsi, l'entreprise va alors enregistrer son nouveau comportement d'achat, qui lui permettra par la suite de toujours mieux cibler le consommateur, de répondre à sa demande et ainsi de le fidéliser. Cependant, pour qu'une telle stratégie soit efficace, il convient de connaître parfaitement ses clients, d'où la nécessité également d'entretenir un fichier client régulièrement mis à jour (changement d'adresse du client, situation familiale). Voilà pourquoi, il est maintenant obligatoire pour chaque achat en ligne de créer un compte. Celui-ci permet de fournir des informations essentielles aux entreprises. En effet, les moindres informations sont passées à la loupe : les entreprises observent ainsi les critères de sélection de recherche des clients tels que la taille des vêtements, la fourchette de prix sélectionnée, les couleurs choisies. De plus en plus, il apparaît que les campagnes de communications massives se révèlent inefficaces. Pour qu'une

campagne soit pertinente, il faut que celle-ci s'adresse à la bonne personne et au bon moment. Le moyen devenu incontournable pour différencier les offres se caractérise par l'e-mailing. En effet, le message va directement s'adresser au consommateur et lui proposer une offre adaptée à ses habitudes de consommation et à son panier moyen. On peut ainsi prendre en exemple l'entreprise Yves Rocher qui utilise ses données clients pour construire une offre personnalisée. A travers la carte de fidélité, l'entreprise enregistre et analyse quels sont les produits que la cliente a l'habitude d'utiliser et d'acheter. Par la suite, elle recevra des coupons de réduction à utiliser sur ces produits. On estime qu'une campagne correctement personnalisée peut atteindre un taux de retour avoisinant les 10% contre 1 ou 2% pour une campagne mal ciblée.

En réalité, au vu de la complexité des données et du temps mis en œuvre, il est impossible pour les entreprises de procéder réellement à une offre différenciée pour chaque personne. Les clients sont segmentés par catégorie afin de proposer à un même groupe de personne la même offre qui pourra tous les satisfaire. Ces données recueillies par les entreprises sont également échangées entre elles afin d'obtenir toujours plus d'informations sur de nouveaux clients potentiels et de pouvoir les cibler en leur proposant une offre leur correspondant.

3. Une offre qui répond et anticipe la demande

En combinant les informations de leur fichier client avec celles des Big Data, les entreprises sont parfaitement capables de fournir une offre personnalisée à ses clients. Avec internet et les téléphones portables, ce sont en permanence des informations de géo localisations qui sont envoyées aux entreprises. Celles-ci peuvent donc être utilisées pour proposer une offre commerciale en fonction de l'endroit où se trouve le client. Celui-ci peut ainsi recevoir une publicité avec une réduction pour un restaurant se trouvant à proximité de là où il se trouve actuellement. Par ailleurs, l'analyse des données recueillies par l'entreprise elle-même permettent également de proposer aux clients une offre qui leur ressemble. En effet, à travers les cartes de fidélité et les commandes réalisées par internet, ce sont toute une quantité d'informations qui vont être analysées pour établir le profil du client selon son panier d'achat. Ainsi, à travers une newsletter par exemple, les entreprises vont être capables de savoir quels sont les clients qui l'ont ouverte, ceux qui l'ont placé dans les spams et où les consommateurs ont cliqué. Le constat est le suivant : plus une newsletter contient le nom de l'entreprise, plus les clics seront importants car le nom est perçu comme un renvoi direct au site internet de la marque. A travers l'analyse d'une newsletter, l'entreprise pourra cerner le raisonnement du client et au fur et à mesure lui envoyer une offre de plus en plus personnalisée.

En observant les habitudes des consommateurs, les entreprises ne se limitent plus qu'à présenter une offre répondant à leurs besoins mais anticipent aussi quelle va être leur demande. Ainsi à travers des recherches sur des destinations de voyages ou des billets d'avion, les sociétés y voient un message à travers lequel les consommateurs sont d'humeur à recevoir des offres de produits qui leur seront utiles pour les vacances. C'est un signal pour les entreprises car cela signifie qu'il s'agit du moment propice pour procéder à l'envoi de catalogues contenant des maillots de bain pour l'été ou un équipement de ski pour l'hiver. De la même manière, une étude avait été réalisée par IBM afin de connaître quelles étaient les attentes des consommatrices concernant la taille des talons des chaussures. En analysant un milliard de messages envoyés via les réseaux sociaux, la société est parvenue à démontrer que les clientes souhaitent voir baisser le nombre de centimètres des talons de chaussures.



Par ailleurs, certaines entreprises ne se limitent plus uniquement à anticiper la demande des consommateurs mais cherchent aussi à bousculer leurs habitudes de consommation. L'exemple le plus célèbre est celui de l'entreprise Target qui a voulu capter une clientèle importante de femmes enceintes. En effet, selon les études d'Alan Anderson, les habitudes des consommateurs ont tendance à changer lorsqu'un événement majeur survient dans leur vie. Ainsi, il était donc particulièrement ingénieux de la part de Target de réussir à capter cette nouvelle clientèle qui s'apprêtait à changer son comportement d'achat. Pour ce faire, la société a donc analysé les habitudes des femmes enceintes tout au long de leur grossesse. Target a ainsi réussi à identifier les 25 produits qu'elles consommaient le plus et à établir un cycle de vie des produits achetés pendant leur grossesse. A travers ses analyses, Target est donc devenu capable d'identifier une femme enceinte à partir de son troisième mois de grossesse.



4. Impact sur la publicité

En outre, l'utilisation des Big Data entraîne le développement du real time marketing. Il devient primordial pour les entreprises de pouvoir interagir en direct avec leurs consommateurs. Cette stratégie consiste à analyser en direct l'envie du consommateur afin de proposer le bon produit, au bon endroit et à la bonne personne. Ainsi, lorsqu'une personne réalise des achats en ligne, ceux-ci sont analysés en temps réel afin de pouvoir lui proposer d'autres produits qui pourraient potentiellement répondre à ses goûts. Par ailleurs, à chaque ouverture d'une nouvelle page internet, des cookies (codage informatique) s'installent sur les ordinateurs afin de pouvoir pister et enregistrer l'historique de navigation du consommateur. De cette manière si le consommateur n'a pas eu recours à l'achat, l'entreprise va être capable de faire apparaître une fenêtre publicitaire contenant le produit qui a préalablement attiré l'intérêt du consommateur.

B. Comment les entreprises peuvent-elles exploiter les Big Data pour définir une nouvelle stratégie et un nouveau business model, ceci afin de développer un avantage concurrentiel

1. La pertinence des Big Data

Nous l'avons donc vu, les Big Data permettent de recueillir des informations clés pour cibler et identifier les clients. Les entreprises ne peuvent donc plus se contenter des informations qu'elles arrivent elles-mêmes à collecter à travers l'historique d'achat de ses clients, elle a aussi besoin de connaître leurs goûts et leurs préférences que ces derniers affichent via les réseaux sociaux, leurs recherches sur internet et leur localisation. Or, actuellement, pour obtenir ces informations, il convient de les acheter à une entreprise fournisseuse de données, telles que Facebook, Google ou Amazon qui en génèrent le plus.

Il conviendrait donc, pour les entreprises, d'être capable d'obtenir elles-mêmes ces informations afin de ne plus être dépendantes de ces géants de l'information producteurs de données. Voilà pourquoi il peut paraître ingénieux pour les entreprises de se focaliser davantage sur les objets connectés afin qu'elles puissent obtenir par leurs propres moyens des informations sur leurs consommateurs. C'est donc tout un nouveau business model que l'on voit apparaître.

2. De nouveaux business models

Comme nous l'avons déjà expliqué, la valorisation d'une entreprise dépend maintenant de la quantité d'informations dont elle dispose sur les clients. Voilà pourquoi, on voit également apparaître dans les entreprises une nouvelle source de revenus reposant sur la commercialisation des Big Data. Les entreprises voient donc apparaître une nouvelle clientèle ; des entreprises consommatrices de Big Data; et une manière différente de faire du profit. En effet, chaque information peut être bonne à vendre. On peut distinguer deux types de commercialisation des données.

La première consiste à vendre à une autre entreprise des fichiers clients afin d'obtenir des informations sur de potentiels clients. Une entreprise A va alors vendre à une entreprise B un fichier client contenant le nom, le prénom, l'adresse e-mail et l'âge d'un ensemble de consommateurs. Cette échange de base de données va alors permettre à l'entreprise B de prospecter de nouveaux clients en possédant déjà des informations sur eux et donc de leur envoyer une offre ciblée avant même que ceux-ci ne soient consommateurs de la marque.

L'autre type de commercialisation de données concernent les entreprises, qui à travers leur canal de distribution collectent des données sur les habitudes des

consommateurs. Prenons en exemple la société TOMTOM spécialisée dans la vente de GPS afin d'aider les usagers à trouver leur itinéraire. Cette dernière reçoit en réalité énormément d'informations de la part des utilisateurs : quels sont les itinéraires réalisés ? A quelle vitesse ceux-ci roulent-ils par rapport à la vitesse effectivement autorisée ? Combien de temps faut-il pour aller d'un point A à un point B ? Autant d'informations qui ont été revendues aux collectivités territoriales afin d'accroître leurs connaissances sur les voies de circulation. Par ailleurs, Tomtom a également utilisé toutes ces données pour les revendre à Apple et ainsi devenir l'un des fournisseurs officiel de données pour l'application Map sur Iphone et Ipad. De la même manière, SFR revend des informations sur la géolocalisation de ses clients à des enseignes de grande consommation afin de connaître la provenance de ceux-ci. Dans ces deux cas, l'entreprise ne change pas de secteur mais acquiert de nouveaux clients totalement différents à travers la vente de Big Data. Cependant, certaines entreprises voudraient pouvoir obtenir directement cette source d'information, sans avoir à passer par une tiers personne.

Pour avoir son propre réseau de Big Data, il convient de rappeler que celles-ci sont principalement obtenues par l'utilisation d'un objet connecté. C'est le Smartphone qui va permettre de géolocaliser l'individu. Ce sont ses recherches sur internet et les échanges d'e-mails qui vont permettre de connaître ses goûts et ses habitudes. Pour obtenir leur propre source de d'information, les entreprises peuvent avoir tendance à se positionner davantage sur les objets connectés et des services liés à ceux-ci. On peut ainsi prendre en exemple la société Apple qui multiplie les ventes de nouveaux objets connectés (le dernier en date est la Apple Watch) et qui se développe sur les activités liées à ces objets. Ainsi, chacune des activités proposées par Apple permettent d'analyser toujours plus les consommations des consommateurs : l'Appstore permet d'analyser les applications utilisées par les consommateurs, Itunes permet d'analyser le type de musique achetée et écoutée, l'Icloud permet d'analyser les données personnelles stockées par les utilisateurs.



C. Comment procéder à la conduite du changement « Big Data » ?

Nous l'avons vu, le Big Data peut devenir un puissant moteur d'amélioration de performance à condition que l'entreprise sache faire face aux enjeux organisationnels, commerciaux et stratégiques. Pour cela, elle doit repenser son fonctionnement global et transformer tous les éléments de son organisation et de sa stratégie.

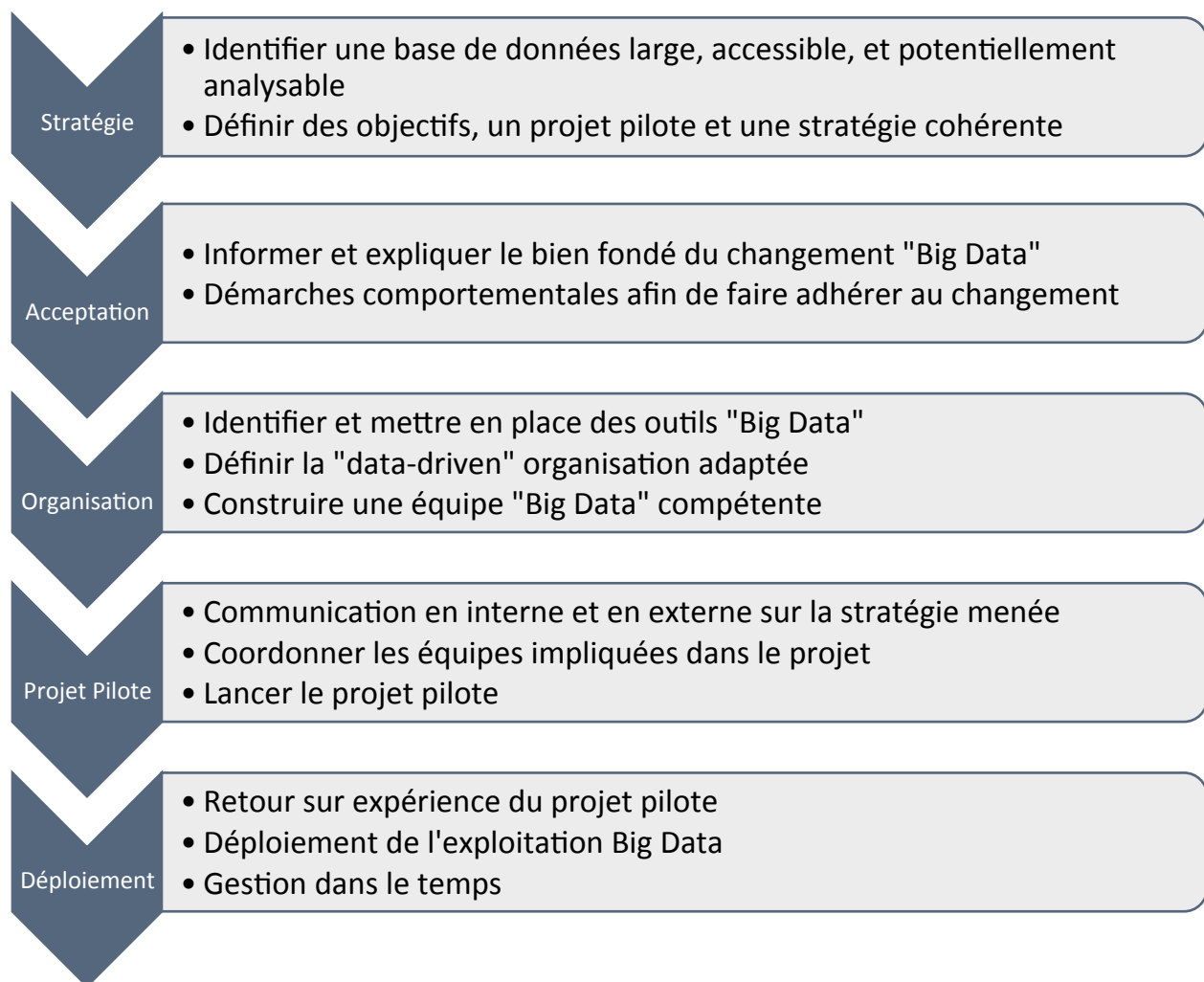
Cependant nous avons aussi vu que le Big Data est un phénomène venant d'émerger mais évoluant à vitesse grand « V ». Ce phénomène est encore peu étudié et difficile à appréhender, surtout pour les entreprises souhaitant en tirer profit au sein même de leur fonctionnement. Il n'y a donc pas vraiment de « solution prêt à l'emploi » ou de réelle formation au Big Data.

Les exemples de « Best Practices » que nous avons précédemment mis en valeur nous enseignent que le « Big Data » n'est pas simplement une initiative technologique mais une stratégie à part entière qui s'appuie sur une expertise technique et technologique. Pour réussir, les entreprises doivent donc définir une stratégie propre alliant outils, équipe de travail compétente et « data-driven » organisation adaptée. Elles doivent ensuite mener le changement de manière organisée et progressive afin d'obtenir les bénéfices du Big Data et créer des synergies.

Nous avons démontré que les aspects techniques d'intégration du Big Data au fonctionnement de l'entreprise sont les principaux défis qu'elle doit relever. Néanmoins il ne faut pas oublier les préoccupations humaines. Car tout changement peut s'avérer être un échec si l'entreprise ne réussit pas à obtenir de la part de son personnel l'acceptation de la nouvelle organisation et des nouvelles missions associées. Il faut donc que les entreprises mettent en place des méthodes efficaces pour soutenir ses collaborateurs lors des transformations et gérer les résistances en interne.

Enfin, il faut que l'entreprise réussisse à faire accepter son changement au niveau de toutes ses parties prenantes : actionnaires, clients, fournisseurs et Etat. Ceci passe notamment par une communication en externe.

Nous avons élaboré un schéma explicatif afin d'éclaircir une conduite du changement « type » (à adapter cependant en fonction des entreprises) relevant les défis techniques et humains qu'engendre les transformations du Big Data dans le fonctionnement de l'entreprise.



1. Stratégie

Avant de se lancer dans un changement « Big Data », l'entreprise doit s'assurer qu'elle détient une large quantité de données facilement accessible et analysable.

Après avoir identifié cette base de données, la Direction doit définir son ambition car adopter le Big Data, c'est adopter une nouvelle manière de fonctionner et mettre l'analyse de données au centre des décisions stratégiques.

Les supérieurs doivent alors répondre aux questions fondamentales :

- Dans quel but?
- Comment le Big Data va améliorer les performances de l'entreprise?

L'entreprise dispose de quatre options pour bénéficier de l'exploitation du Big Data :

- Améliorer l'offre de produits et service existante

- Améliorer le fonctionnement interne de l'entreprise
- Créer une nouvelle offre de biens ou services
- Transformer le business model

Ces quatre options peuvent néanmoins se chevaucher et être menés simultanément ou successivement. Ainsi, l'entreprise doit définir comment elle va investir dans le Big Data mais surtout par où elle va commencer. Pour cela, elle doit définir des objectifs atteignables et un projet pilote sur un périmètre réduit.

Enfin, de cette définition découlera la stratégie que souhaitera mener l'entreprise.

2. Acceptation

Obtenir l'adhésion des salariés ne s'obtient pas sans les informer et leur expliquer le bien-fondé du changement « Big Data ». Les supérieurs doivent convaincre le personnel de mobiliser leur énergie sur un objectif phare : la transformation de l'entreprise vers l'exploitation du Big Data.

L'ennemi que doit combattre la Direction : la résistance au changement. Cette résistance est différente d'un employé à un autre et dépend de nombreux facteurs comme le niveau hiérarchique, le degré de transformation du métier ou bien le service dans lequel travaille l'employé.

Ces préoccupations humaines peuvent se révéler être la tâche la plus difficile dans la conduite du changement. En effet, les habitudes ne se changent pas facilement et le potentiel "retour sur investissement" de l'exploitation du Big Data n'est pas évident à intégrer pour tous les salariés.

Pour répondre à cette résistance, la Direction peut définir des acteurs « moteur du changement » (managers, responsables de département) qui peuvent s'engager dans une communication top-down. Il faut en effet être à l'écoute de chaque personne, l'informer en adaptant son discours et amener chacun d'eux à prendre en charge sa propre évolution.

Il faut aussi faire attention à ce que ces acteurs « porteurs du changement » s'intègrent pleinement au cœur de l'entreprise et soient en interaction avec l'ensemble des collaborateurs pour faire passer le message efficacement. Ils doivent donc sensibiliser et communiquer de manière répétitive sur le sujet. L'adhésion des salariés s'obtient uniquement de l'intérieur.

Des démarches comportementales (brainstorming, action collective) peuvent être mises en place afin de faciliter l'adhésion et l'implication du personnel au changement. Il faut aussi s'appuyer dans un premier temps sur les salariés favorables au changement afin qu'ils influencent positivement les réticents.

3. Organisation

Comme nous l'avons développé précédemment, l'entreprise doit identifier et mettre en place des outils "Big Data" (Hadoop, NoSQL, cloud...) pertinents.

Elle doit ensuite définir la « data-driven » organisation adaptée à la stratégie qu'elle veut mener : organisation décentralisée, centralisée ou sous-traitance.

Si l'entreprise décide de ne pas sous-traiter, le dernier élément nécessaire à une organisation efficace est la construction d'une équipe « Big Data » experte, ayant des compétences statistiques, informatiques et business, capable d'analyser le Big Data (Data scientist, Data analyst...).

4. Projet Pilote

Avant de lancer le projet, la Direction et les acteurs « moteurs du changement » doivent communiquer en interne sur le déroulement du projet pilote et toutes ses implications. Toute conduite du changement se décompose en une multitude d'initiatives à mener au sein des différents services de l'entreprise. Aussi, pour éviter de semer la confusion en diffusant des messages contradictoires, il est indispensable de coordonner les différentes équipes impliquées dans le projet et d'assurer une transversalité dans la composition des groupes de travail.

En externe, afin que toutes les parties prenantes externes à l'entreprise acceptent le changement que l'entreprise s'apprête à mener, cette dernière doit aussi mettre en place une communication externe claire, convaincante et rassurante.

L'entreprise est dès à présent prête à lancer son projet pilote.

5. Déploiement

Le projet pilote achevé ou presque, l'entreprise peut évaluer les réussites et les échecs de ce « test ». Ce retour sur expérience lui permet d'appliquer ses bonnes pratiques sur les projets suivants, de ne pas refaire les mêmes erreurs et d'améliorer les méthodes utilisées. L'entreprise peut maintenant déployer l'exploitation du Big Data sur un périmètre plus large en lançant de nouveaux projets par exemple.

Par ailleurs la conduite de changement ne se limite pas à la seule gestion de projet mais doit s'inscrire dans le temps. L'entreprise doit avoir pour ambition de rendre les transformations permanentes et de veiller à ce que le personnel ne revienne pas aux habitudes passées mais s'inscrive dans la nouvelle vision « Big Data ».

Bibliographie

AIGRAIN Maud (2003), *One to one : coller au plus près aux desiderata des clients*, Action Commerciale, janvier 2003, No. 227

ANTVOICE (2013), *L'efficacité de la recommandation produit chez Amazon* [en ligne], disponible sur <http://www.antvoice.com/e-commerce-2/lefficacite-de-la-recommandation-produit-chez-amazon/>

ARTHUR Lisa (2013), *What Is Big Data?* [en ligne], disponible sur <http://www.forbes.com/sites/lisaarthur/2013/08/15/what-is-big-data/>

AUTHEMAYOU Céline (2013), *Big Data : de nouveaux métiers pour une nouvelle économie* [en ligne], disponible sur <http://www.letudiant.fr/metiers/big-data-quand-les-donnees-deviennent-le-nouveau-petrole-de-l-emploi/big-data-de-nouveaux-metiers-pour-une-nouvelle-economie.html>

BARTLETT Jamie (2013), *iSPY: How the internet buys and sells your secrets* [en ligne], disponible sur <http://www.spectator.co.uk/features/9093961/little-brothers-are-watching-you/>

BELOUEZZANE Sarah, DUCOURTIEUX Cécile (2012), *Vertigineux "big data"* [en ligne], disponible sur http://www.lemonde.fr/technologies/article/2012/12/26/vertigineux-big-data_1810213_651865.html

BIESDORF Stefan, COURT David, WILLMOTT Paul (2013), *Big Data: What's your plan?* [en ligne], disponible sur http://www.mckinsey.com/insights/business_technology/big_data_whats_your_plan

BIG DATA-STARTUP (2014), *Big Data Best Practices* [en ligne], disponible sur <http://www.bigdata-startups.com/best-practices/>

BOUCHER FERGUSON Renee (2013), *Big Data and Big Change Management: A Path Forward* [en ligne], disponible sur <http://sloanreview.mit.edu/article/big-data-and-big-change-management-a-path-forward/>

BOUCHER Simon (2013), *Les Big Data et les réseaux sociaux* [en ligne], In Actes de la conférence AIIM 2013, Montréal disponible sur <http://www.aiim.org/~media/Files/Chapters/FirstCanadian/aiim-MTL-%20Quest-ce%20que%20le%20Big%20Data%20-%20AIIM%20-%20Simon%20Boucher.ashx>

BREWSTER Tom (2014), *Facebook, Google and personal data: What's your worth?* [en ligne], disponible sur <http://www.bbc.com/future/story/20140509-how-much-is-your-facebook-worth>

BROWN Brad, COURT David, WILLMOTT Paul (2013), *Mobilizing your C-suite for big-data analytics* [en ligne], disponible sur http://www.mckinsey.com/insights/business_technology/mobilizing_your_c_suite_for_big_data_analytics

CHATTERJEE Pratap (2013), *How Private Tech Companies are collecting Data on you and sellong them to the feds for huge profits* [en ligne], disponible sur <http://www.alternet.org/civil-liberties/how-private-tech-companies-are-collecting-data-you-and-selling-them-feds-huge>

CHAUMAIS Romain (2013), *Modification organisationnelle liée au big data : le défi* [en ligne], disponible sur <http://pro.01net.com/editorial/593249/modification-organisationnelle-liee-au-big-data-le-big-defi/>

COTTON James (2014), *Le Big data et le consommateur* [en ligne], disponible sur <http://www.lesechos.fr/idees-debats/cercle/cercle-107117-le-big-data-et-le-consommateur-1031238.php>

DAVENPORT Thomas, DYCHE Jill (2013), *Big Data in Big Companies* [en ligne], disponible sur <http://www.sas.com/resources/asset/Big-Data-in-Big-Companies.pdf>

DE MONTCHEUIL Yves (2014), *Facebook: A Decade of Big Data* [en ligne], disponible sur <http://www.wired.com/2014/03/facebook-decade-big-data/>

DE MONTCHEUIL Yves (2014), *Quand les big data créent de nouveaux business models* [en ligne], disponible sur <http://www.latribune.fr/opinions/tribunes/20140407trib000824061/qua>

[nd-les-big-data- creent-de-nouveaux-business-models.html](#)

DELORT Pierre (2014), Interview [en ligne], disponible sur <http://www.bigdataparis.com/2014/itv-pierre-delort.php>

DI DOMENICO Maria Laura, NUNAN Daniel (2013), *Market research and the ethics of big data*, International Journal of Market Research, novembre 2013, Vol. 55, No. 4, p.505-520

DOLIQUE Alexandre (2014), *Le Big Data se démocratise : impacts et enjeux sur l'organisation des entreprises* [en ligne], disponible sur <http://www.journaldunet.com/solutions/expert/58779/le-big-data-se-democratise---impacts-et-enjeux-sur-l-organisation-des-entreprises.shtml>

DUHIGG Charles (2012), *How Companies Learn Your Secrets* [en ligne], disponible sur http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&_r=2&

DUPONT-CALBO Julien, PIQUARD Alexandre (2013), *Les "cookies" divisent les acteurs de la publicité en ligne* [en ligne], disponible sur http://www.lemonde.fr/technologies/article/2013/11/25/les-cookies-divisent-les-acteurs-de-la-publicite-en-ligne_3519753_651865.html

EDWARDS Luke (2014), *How do companies make money from your Data* [en ligne], disponible sur <http://www.pocket-lint.com/news/130366-how-do-companies-make-money-from-your-data>

ELLIOTT Timeo (2013), *L'Oréal: Turning Data Into A Customer Experience* [en ligne], disponible sur <http://timoelliott.com/blog/2013/04/loreal-turning-data-into-a-customer-experience.html>

FEINLEIB Dave (2014), *Guide du Big Data pour les nuls* [en ligne], disponible sur <http://www.data-business.fr/guide-du-big-data-pour-les-nuls/#T%C3%A9l%C3%A9chargement>

FULGONI Gian (2013), *Friend or Foe of Digital Advertising? Five Ways Marketers Should Use Digital Big Data to Their Advantage*, Journal of Advertising Research, December 2013, Volume 53, Issue 4, p. 372-36

GABEL Tim, TOKARSKI Cathy (2014), *Big Data and Organization*

Design: Key Challenges Await the Survey Research Firm, *Journal of Organization Design*, 2014, Vol. 3, No.1

GALBRAITH Jay R. (2014), *Organization Design Challenges Resulting From Big Data*, *Journal of Organization Design*, 2014, Vol.3, Issue 1, p. 2

GASNIER Lise (2013), *Aux origines du Big Data* [en ligne], disponible sur <http://www.solucominsight.fr/2013/08/auw-origines-du-big-data>

GOLLION Anne-Laurence (2014), *Instituts d'études et sondage, l'effet big data* [en ligne], disponible sur <http://www.lenouveleconomiste.fr/lesdossiers/instituts-detudes-et-sondage-leffet-big-data-24104/>

HAMAMI Oren (2014), *Big Data Security: Understanding the Risks*, *Business Intelligence Journal*, juin 2014, Vol. 19 Issue 2, p20

HYON-LE GOURRIEREC Lucile (2013), *Quels usages concrets du Big Data ?* [en ligne], disponible sur <http://www.journaldunet.com/solutions/expert/54903/quels-usages-concrets-du-big-data.shtml>

JORDAN John (2013), *The Risks of Big Data for Companies* [en ligne], disponible sur <http://online.wsj.com/news/articles/SB10001424052702304526204579102941708296708>

JUHAN Virgile (2013), *Ces métiers créés par le Big Data* [en ligne], disponible sur <http://www.journaldunet.com/solutions/analytics/metier-big-data-data-scientist.shtml>

LANEY Doug (2012), *To Facebook You're Worth \$80.95* [en ligne], disponible sur <http://blogs.wsj.com/cio/2012/05/03/to-facebook-youre-worth-80-95/>

LEMBERGER Pirmin (2012), *Quelles compétences pour le Big Data* [en ligne], disponible sur <http://www.journaldunet.com/solutions/expert/51837/quelles-competences-pour-le-big-data.shtml>

MEUNIER Nelly (2014), *Big Data : les applications au sein du parcours*

client [en ligne], disponible sur

<http://www.marquesetreseaux.com/2014/01/big-data-parcours-client/>

NORA Dominique, PENICAULT Nicole, SOULA Claude (2014), *Pourquoi la BIG data devient un très gros business* [en ligne], disponible sur

<http://tempsreel.nouvelobs.com/economie/20140417.OBS4360/pourquoi-i-la-big-data-devient-un-tres-gros-business.html>

PEARSON Travis, WENEGER Rasmus (2013), *Big Data: The organizational challenge* [en ligne], disponible sur

http://www.bain.com/publications/articles/big_data_the_organizational_challenge.aspx

POGGI Nora (2014), *Comment le Big Data bouleverse l'organisation des entreprises* [en ligne], disponible sur <http://www.usine-digitale.fr/article/comment-le-big-data-bouleverse-l-organisation-des-entreprises.N253137>

PRESS Gil (2013), *A Very Short History Of Big Data* [en ligne], disponible sur <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/>

PRESS Gil (2014), *12 Big Data Definitions: What's Yours?* [en ligne], disponible sur <http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/>

RAZA Christian (2014), *Big data : vers une industrialisation d'Hadoop ?* [en ligne], disponible sur

<http://www.journaldunet.com/solutions/expert/56701/big-data---vers-une-industrialisation-d-hadoop.shtml>

REDSSEN CONSULTING (2013), *Qu'est-ce que le Big Data ?* [en ligne], disponible sur <http://www.redsen-consulting.com/2013/06/big-data/>

SINHA Velu, WENEGER Rasmus (2013), *The value of Big Data: How analytics differentiates winners* [en ligne], disponible sur

<http://www.bain.com/publications/articles/the-value-of-big-data.aspx>

STANISLAS Emmanuel (2014), *Big Data, Big Compétences* [en ligne], disponible sur

<http://www.latribune.fr/opinions/tribunes/20140507trib000828825/big-data-big-competences.html>

VINCENT Claude (2013), *La ruée vers l'or des données personnelles* [en ligne], disponible sur

<http://www.journaldunet.com/solutions/expert/56701/big-data---vers-une-industrialisation-d-hadoop.shtml>