

Analytique : De l'usage concret des Big Data

Ou comment les entreprises innovantes exploitent le potentiel de données de valeur incertaine



IBM Institute for Business Value

IBM Global Business Services, via l'IBM Institute for Business Value, développe à l'intention des cadres dirigeants des études stratégiques fondées sur des faits, autour des problématiques propres au secteur privé et au secteur public. Ce rapport de synthèse se base sur une étude approfondie conduite par l'équipe de recherche de l'institut. Il a été réalisé dans le cadre de l'engagement permanent d'IBM Global Business Services pour fournir des analyses et une réflexion qui aident les entreprises à produire une forte valeur ajoutée. D'autres études publiées par l'IBM Institute for Business Value sont disponibles sur ibm.com/iibv

Saïd Business School de l'Université d'Oxford

La Saïd Business School est l'une des plus grandes écoles de commerce du Royaume-Uni. Cette école s'emploie à établir un nouveau modèle d'enseignement commercial en collaborant étroitement avec l'Université d'Oxford, université de réputation mondiale, et tente de relever un certain nombre de défis du monde actuel. Pour plus d'informations, vous pouvez contacter les auteurs ou visiter le site suivant : www.sbs.ox.ac.uk

Par Michael Schroeck, Rebecca Shockley, le Dr. Janet Smart, le Professeur Dolores Romero-Morales et le Professeur Peter Tufano

Le terme « **Big Data** », qui revêt une multitude de significations selon les personnes, dépasse désormais la sphère de la technologie. Ce concept représente aujourd'hui un enjeu commercial prioritaire, étant donné sa capacité à impacter le commerce en profondeur dans l'économie mondiale intégrée. En plus d'apporter des solutions aux enjeux commerciaux durables, les Big Data inspirent de nouvelles manières de transformer les processus, les organisations, les secteurs d'activités entiers, voire même toute la société. Mais du fait de son importante couverture médiatique, il est difficile de distinguer le mythe de la réalité. Qu'en est-il vraiment ? Notre toute dernière recherche montre que les organisations utilisent les Big Data pour obtenir des résultats orientés client, consulter des données internes et développer un meilleur écosystème de l'information.

Alors que l'expression « Big Data » se généralise, cette notion continue de créer de la confusion. Ce terme est utilisé pour véhiculer toutes sortes de concepts : on évoque notamment de gigantesques volumes de données, l'analytique des réseaux sociaux, les fonctions de gestion de données de nouvelle génération, les données en temps réel, et bien d'autres choses encore. Quelle que soit la formule utilisée, les organisations commencent à comprendre et à explorer de nouvelles manières de traiter et d'analyser un large éventail d'informations. C'est ainsi qu'un groupe restreint mais croissant de pionniers a déjà atteint des résultats commerciaux exceptionnels.

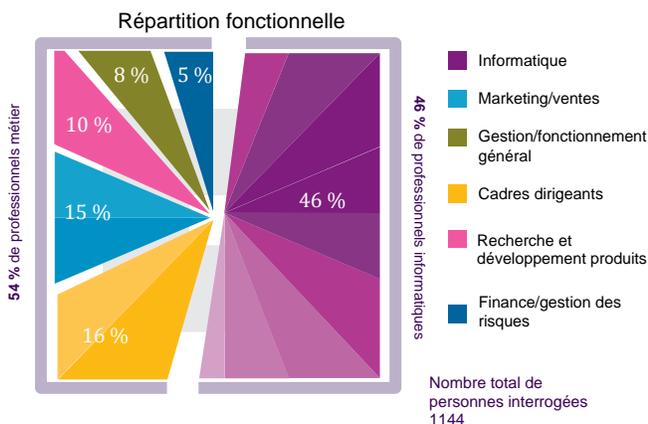
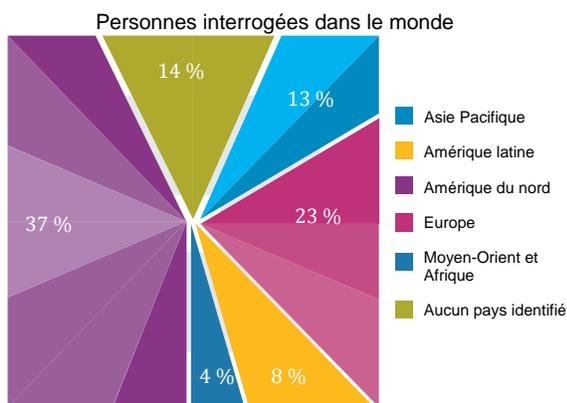
Des dirigeants d'entreprises du monde entier reconnaissent la nécessité d'apprendre à exploiter les Big Data. Mais malgré une attention sans relâche de la part des médias, il semble difficile de trouver des informations précises sur ce que font réellement les entreprises.

C'est pourquoi nous avons cherché à mieux comprendre la manière dont les entreprises envisagent les Big Data et dans quelle mesure elles l'utilisent aujourd'hui pour profiter à leurs activités. L'IBM Institute for Business Value, en collaboration avec la Saïd Business School de l'Université d'Oxford, a mené une étude sur l'utilisation des Big Data en 2012 (Big Data @ Work) auprès de 1 144 entreprises et professionnels de l'informatique dans 95 pays, en interrogeant plus d'une vingtaine d'universitaires, des experts en la matière et des responsables d'entreprises.

À propos de l'étude

L'IBM Institute for Business Value et la Saïd Business School de l'Université d'Oxford ont collaboré à la publication de ce rapport. Il est basé sur l'étude « Big Data @ Work » réalisée par IBM courant 2012 auprès de 1144 professionnels de 26 secteurs industriels dans 95 pays différents. Les personnes interrogées sont issues de métiers variés répartis entre fonctionnels (54 % du total des personnes) et informatiques (46 %). Les participants se sont inscrits eux-mêmes pour prendre part à l'étude en ligne.

Les résultats de l'étude sont basés sur l'analyse des données de l'enquête et des discussions avec des universitaires de l'Université d'Oxford, des experts en la matière et des responsables d'entreprises. IBM est la source principale des recommandations de l'étude.



L'étude révèle que 63 % des personnes interrogées, près des deux tiers, déclarent que l'utilisation des informations (dont les Big Data) et de l'analytique crée un avantage concurrentiel pour leur organisation. Ce chiffre est à comparer avec les 37 % de réponse comparable obtenue pour l'étude l'Entreprise intelligente 2010 (New Intelligent Enterprise Global Executive Study and Research Collaboration) conduite par IBM ; une augmentation de 70 % en tout juste deux ans.

En tant que segment toujours plus important du marché de l'information et de l'analytique, les Big Data ont un réel impact. Les personnes interrogées faisant partie d'organisations engagées dans des projets ou des déploiements pilotes de Big Data ont 15 % plus de chances de tirer un maximum de bénéfices des informations (dont les Big Data) et de l'analytique par rapport à celles exploitant des fonctions d'analyse traditionnelles.

L'un des résultats surprenants de cette étude est l'impact relativement faible des réseaux sociaux sur le marché actuel des Big Data. Étant donné l'importance de la couverture médiatique concernant l'impact des données sociales sur l'expérience client, on pourrait facilement croire que les Big Data seraient associées aux données des réseaux sociaux mais seulement 7 % des personnes interrogées ont défini les Big Data de cette manière. Et moins de la moitié des personnes interrogées impliquées dans des projets de Big Data ont déclaré recueillir et analyser les données des réseaux sociaux. Les personnes interrogées en revanche ont affirmé utiliser les sources de données internes pour développer les Big Data.

Qu'est-ce qui rend les activités de type « Big Data » différentes aujourd'hui ? Certaines organisations gèrent d'importants volumes de données de puis de longues années. Par exemple, une entreprise internationale de télécommunications collecte des milliards d'enregistrements détaillés d'appels provenant de 120 systèmes différents et les stocke tous pendant neuf mois au minimum. Une société d'exploration pétrolière analyse des téraoctets de données géologiques et échange des millions de transactions en bourse chaque minute. Pour ces entreprises, le concept de Big Data n'est pas nouveau.

Toutefois, deux tendances importantes ont fait évoluer les Big Data récemment :

- La numérisation quasi-systématique crée désormais de nouveaux types de groupes de données volumineux en temps réel pour un grand nombre de secteurs industriels. Il s'agit pour la plupart de données non structurées : par exemple les données de flux, géo spatiales ou générées par capteur qui ne peuvent pas être stockées dans des entrepôts de données traditionnels, structurés et relationnels.
- Les technologies et techniques d'analyse avancée actuelles aident les organisations à extraire des connaissances grâce aux données avec des niveaux de sophistication, de vitesse et de précision impensables jusqu'à aujourd'hui.

Pour l'ensemble des secteurs industriels et des zones géographiques, notre étude a révélé que les organisations ont une approche pragmatique des Big Data. Les solutions Big Data les plus efficaces commencent par identifier les besoins métier puis personnalisent l'infrastructure, les sources de données et l'analytique nécessaires à cette initiative. Ces organisations peuvent extraire de nouvelles connaissances grâce aux sources d'informations existantes et à ces nouvelles sources disponibles, définir une stratégie de technologie de Big Data puis réaliser en conséquence des mises à niveau incrémentielles de leurs infrastructures dans le temps.

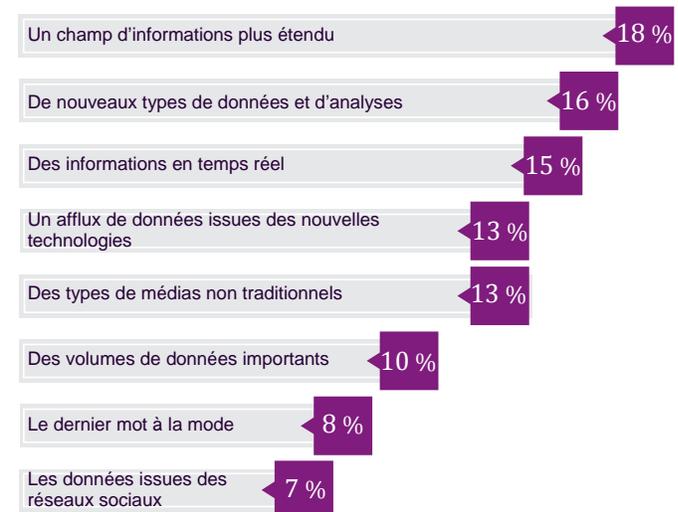
Les résultats de notre étude ont permis de mettre en évidence cinq recommandations clés permettant aux organisations de faire évoluer leurs Big Data et de les valoriser au maximum :

- Concentrer ses efforts initiaux sur les résultats orientés client
- Développer un projet Big Data à l'échelle de l'entreprise
- Partir des données existantes pour obtenir des résultats quasiment en temps réel
- Bâtir des fonctions d'analyse basées sur les priorités métier
- Créer une analyse de faisabilité basée sur des résultats mesurables.

Vers une définition des Big Data

La plus grande part de la confusion autour des Big Data est due à sa définition elle-même. Afin de comprendre les définitions données par les personnes interrogées au cours de l'étude, nous avons demandé à chacune de sélectionner jusqu'à deux caractéristiques des Big Data. Plutôt que de voir une caractéristique unique dominer nettement toutes les autres, les personnes interrogées se sont montrées divisées quant à savoir si les Big Data se définissent davantage par l'augmentation actuelle des volumes de données, les nouveaux types de données et d'analyse ou les besoins émergents pour l'analyse des informations en temps réel (voir la Figure 1).

Vers une définition des Big Data



Les personnes interrogées étaient invitées à choisir jusqu'à deux réponses parmi les choix proposés ci-dessus pour décrire la façon dont leur organisation envisage les Big Data. Les descriptions ont été abrégées et les réponses normalisées pour atteindre 100 %. Nb total de personnes interrogées = 1 144.

Figure 1 : Les personnes interrogées sont divisées sur la manière d'envisager les Big Data.

Ces résultats illustrent une manière utile de définir les trois dimensions des Big Data, appelée « les trois V » : volume, variété et vélocité. Et bien que ces trois dimensions couvrent les attributs clés des Big Data, les organisations devraient prendre en compte une quatrième dimension : la véracité. La prise en compte de la véracité comme quatrième attribut des Big Data souligne l'importance de traiter et de gérer l'incertitude inhérente à certains types de données (voir la Figure 2).

La convergence de ces quatre dimensions permet de définir et de distinguer les Big Data :

Volume : quantité de données. Sans doute la caractéristique la plus étroitement liée aux Big Data, le volume se réfère aux grandes quantités de données que les organisations tentent de mettre à profit pour améliorer la prise de décision dans l'ensemble de l'entreprise. Les volumes de données continuent d'augmenter à un rythme sans précédent. Toutefois, ce qui constitue véritablement d'importants volumes de données varie par secteur et même par zone géographique et reste inférieur aux pétaoctets et zétaoctets souvent évoqués. Un peu plus de la moitié des personnes interrogées seulement considèrent les volumes de données compris entre un téraoctet et un pétaoctet comme d'importants volumes de données, et 30 % des participants ne savent simplement pas à partir de quel stade les volumes de données sont considérés comme « copnséquent » par leur

organisation. Nous sommes pourtant tous d'accord pour dire que ce que l'on considère comme conséquent aujourd'hui le sera bien plus demain.

Variété : différents types de données et de sources de données. Il s'agit de gérer la complexité d'une multitude de types de données y compris des données structurées, semi structurées et non structurées. Les organisations doivent intégrer et analyser des données provenant d'un ensemble complexe de sources d'informations traditionnelles et non traditionnelles, internes et externes à l'entreprise. Face à l'explosion des capteurs, des périphériques intelligents et des technologies de communautés sociales, les données peuvent être générées dans un nombre incalculable de formats : texte, données Web, tweets, données générées par capteur, audio, vidéo, flux de clics, fichiers journaux, etc.

Vélocité : données en mouvement. La vitesse à laquelle les données sont créées, traitées et analysées continue d'augmenter. Contribuer à augmenter la vélocité constitue la nature même de la création de données en temps réel, ainsi que le besoin d'incorporer des flux dans des processus métier et de prise de décision. La vélocité impacte la latence, c'est-à-dire le délai compris entre le moment où les données sont créées ou capturées et celui où elles deviennent accessibles. Aujourd'hui, les données sont créées de manière continue à un rythme bien trop élevé pour que les

Les dimensions des Big Data

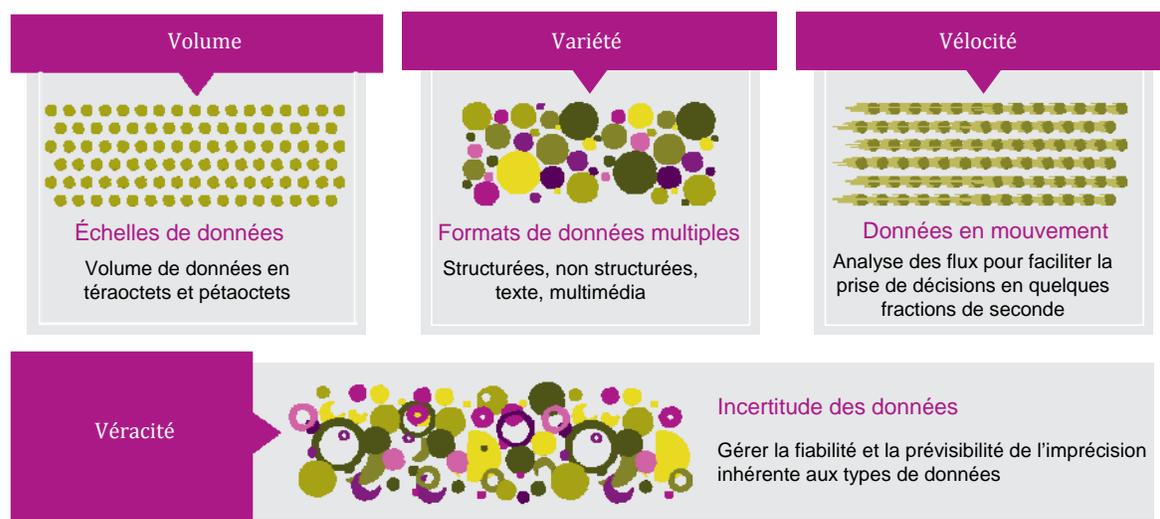


Figure 2 : Les quatre dimensions des Big Data.

systèmes traditionnels puissent les capturer, les sauvegarder et les analyser. Dans le cas des processus sensibles au facteur temps comme la détection des fraudes en temps réel ou le marketing « instantané » multi-canal, certains types de données doivent être analysés en temps réel pour apporter une valeur ajoutée à l'entreprise.

Véracité : incertitude des données. La véracité fait référence au niveau de fiabilité associé à certains types de données. Chercher un niveau de qualité élevé est une exigence clé et un défi majeur des Big Data mais même la meilleure méthode de nettoyage des données ne peut pas supprimer le caractère aléatoire inhérent à certaines données comme le temps, l'économie ou les décisions d'achat futures d'un client. Le besoin de reconnaître et de planifier cette incertitude est une dimension des Big Data qui est apparue lorsque les dirigeants ont cherché à mieux comprendre la nature incertaine de ces données (voir l'encadré « Véracité, le quatrième 'V.' »).²

En fin de compte, les Big Data sont une combinaison de ces différentes caractéristiques qui donne aux entreprises un avantage concurrentiel sur le marché actuel du numérique. Cet avantage permet aux entreprises de transformer la façon dont elles communiquent et servent leurs clients et donnent aux organisations, voire à des secteurs de marché entiers, la possibilité de se transformer eux-mêmes. Toutes les organisations n'auront pas la même approche pour exploiter et développer les performances des Big Data. Mais les possibilités d'utiliser les nouvelles technologies Big Data et l'analytique pour améliorer la prise de décisions et les performances existent dans tous les secteurs.

Les organisations ont une vision pratique des Big Data

Malgré un certain « battage » médiatique, on s'accorde à dire que les entreprises ne sont qu'aux balbutiements de l'adoption de l'environnement Big Data. Nous avons utilisé dans cette étude l'expression « adoption des Big Data » pour définir la progression naturelle des données, sources, technologies et compétences nécessaires pour créer un avantage concurrentiel sur le marché mondial intégré.

Véracité, le quatrième « V »

Certaines données sont incertaines par nature ; par exemple : les sentiments et la sincérité pour les humains ; les capteurs GPS qui rebondissent sur les gratte-ciels de Manhattan ; les conditions météorologiques ; les facteurs économiques, ou encore l'avenir. Concernant ces types de données, aucun nettoyage, aussi important soit-il, ne peut corriger cette incertitude. Pourtant, malgré ce caractère aléatoire, les données contiennent des informations de valeur. Le besoin de reconnaître et d'intégrer cette incertitude est l'une des caractéristiques des Big Data.

Le caractère incertain des Big Data se manifeste de différentes façons. Il réside dans le scepticisme qui entoure les données créées dans un environnement humain comme les réseaux sociaux, l'imprévisibilité quant à l'avenir et à la manière dont les individus, la nature et les forces du marché invisibles réagiront à la variabilité du monde alentour.

Un exemple de cette incertitude concerne la production d'énergie : Le temps est incertain mais une compagnie d'électricité doit être capable de prévoir la production. Dans de nombreux pays, les régulateurs exigent qu'un pourcentage de la production provienne de sources renouvelables, mais ni le vent ni les nuages ne peuvent se prévoir avec précision. Alors comment pouvez-vous planifier ?

Pour gérer cette incertitude, les analystes doivent créer un contexte autour des données. Une façon de créer ce contexte consiste à fusionner les données afin que la combinaison de sources multiples peu fiables produise un point de données plus précis et plus utile, comme les commentaires sociaux ajoutés aux informations de localisation géographique. Une autre manière de gérer ce caractère aléatoire peut être l'utilisation de mathématiques avancées pour couvrir cette incertitude, comme les techniques d'optimisation efficaces et les approches de logique floue.

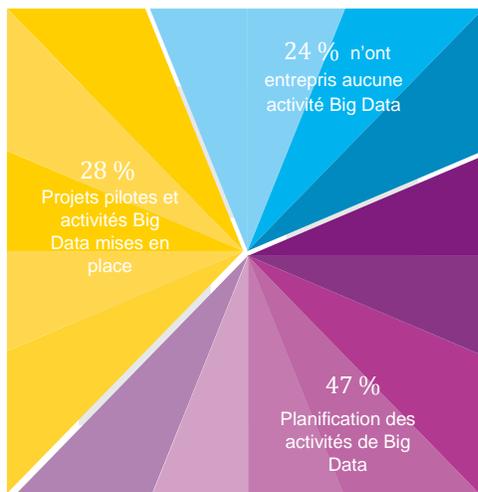
De part leur nature, les êtres humains n'aiment pas l'incertitude mais le fait de l'ignorer risque de créer encore davantage de difficultés que l'incertitude elle-même. À l'ère des Big Data, les dirigeants devront adopter une approche différente de la dimension d'incertitude. Ils devront la reconnaître, l'intégrer et déterminer comment en tirer profit. La seule certitude à propos de l'incertitude est qu'elle n'est pas prête de disparaître.

Notre étude Big Data @ Work a confirmé que la plupart des organisations ne sont qu'à l'aube de leur développement Big Data et une majorité d'entre elles axent leurs efforts soit sur la compréhension des concepts (24 %) soit sur la définition d'une feuille de route associée aux Big Data (47 %). Cependant, 28 % des personnes interrogées travaillent dans des organisations à la pointe dans lesquelles des études de faisabilité ont été réalisées ou des solutions de Big Data sur mesure ont déjà été implémentées (voir figure 3).

En analysant les différentes réponses, cinq résultats majeurs permettent de révéler quelques tendances et approches communes intéressantes :

- Pour l'ensemble des segments du marché, l'analyse des cas de Big Data révèle une approche fortement axée sur la compréhension des objectifs clients.
- La mise en place d'une gestion de l'information évolutive et extensible est un pré-requis essentiel au développement Big Data.

Activités de Big Data



Les personnes interrogées devaient identifier l'état actuel des activités Big Data au sein de leurs organisations. Le total n'atteint pas 100 % car les résultats ont été arrondis. Nb total de personnes interrogées = 1061

Figure 3 : La plus part des organisations ne sont qu'à l'aube de leur développement Big Data.

- Les organisations ont lancé des projets pilotes et des implémentations en utilisant les sources de données existantes et nouvelles sources internes disponibles.
- Des fonctions avancées d'analytique sont nécessaires, bien que souvent manquantes, pour que les organisations puissent exploiter les Big Data au maximum.
- Alors que la compréhension et l'implication pour les Big Data augmentent, quatre étapes dans l'adoption des Big Data émergent nettement.

L'analytique des clients favorise les initiatives autour des Big Data

Lorsqu'elles ont été invitées à déterminer les trois objectifs principaux des Big Data, près de la moitié des personnes interrogées ont identifié les objectifs orientés client comme étant la priorité de leur organisation (voir figure 4). Les organisations cherchent à améliorer l'expérience client et à mieux comprendre les préférences et les comportements de leurs clients. Comprendre comment satisfaire ses clients aujourd'hui était également identifié comme l'une des priorités majeures dans les deux études IBM Global Chief Marketing Officer Study de 2011 et IBM Global Chief Executive Officer Study de 2012.³

Les entreprises considèrent les Big Data comme un moyen évident de mieux comprendre et de prévoir les comportements des clients et par conséquent d'améliorer l'expérience client. Les transactions, les interactions multi-canal, les réseaux sociaux, les données transmises à travers des sources comme les cartes de fidélité et d'autres informations relatives aux clients ont permis aux organisations d'avoir un aperçu plus complet et plus clair des préférences et des exigences des clients, l'un des objectifs majeurs des services de marketing, des ventes et de la clientèle depuis plusieurs dizaines d'années.

Grâce à cette connaissance approfondie, les organisations de tout type cherchent de nouvelles façons d'améliorer leurs relations avec leurs clients et prospects. Ce principe s'applique largement dans les commerces, mais également dans les télécommunications, la santé, le gouvernement, la banque et les finances, ainsi que les produits de consommation pour lesquels les consommateurs finaux et les citoyens s'impliquent et les interactions business-to-business entre les partenaires et fournisseurs.

On peut considérer les Big Data comme une voie à double sens entre les clients et les organisations. Par exemple, le modèle de voiture La Ford Focus électrique génère d'importants volumes de données qu'elle soit en circulation ou garée. En conduisant, le conducteur reçoit constamment de nouvelles informations relatives à l'accélération, le freinage, la charge de la batterie et la localisation de la voiture. « Ces données sont utiles au conducteur mais sont également transmises aux ingénieurs Ford et leur révèlent les habitudes de conduite des clients y compris comment, quand et où ils rechargent leur voiture. » Et lorsque le véhicule est à l'arrêt, les données relatives à la pression des pneus et à l'état de la batterie sont transmises au smartphone le plus proche.⁶

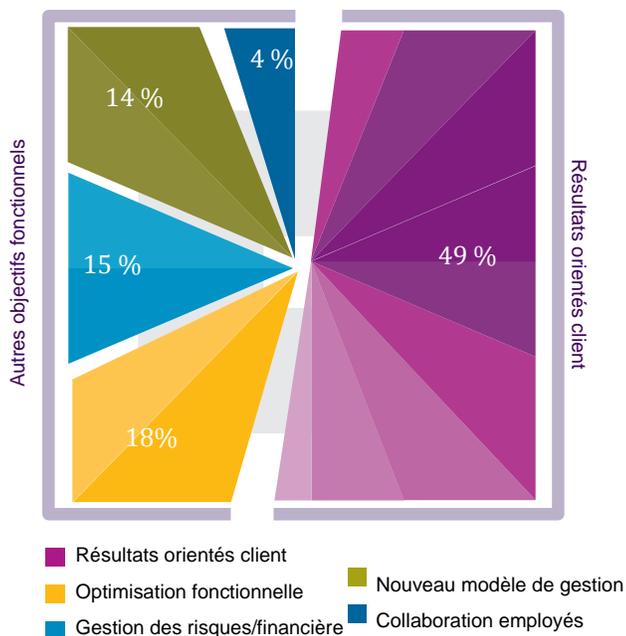
Les Big Data permettent d'avoir un aperçu plus complet des préférences et des exigences des clients. Grâce à cette connaissance approfondie, les organisations de tout type cherchent de nouvelles façons d'améliorer leurs relations avec leurs clients et prospects.

De multiples avantages découlent de ce scénario orienté clients grâce aux Big Data qui favorisent de nouveaux types de collaboration de valeur. Les conducteurs reçoivent des informations immédiates pendant que les ingénieurs à Detroit collectent les informations sur les comportements de conduite pour améliorer leurs connaissances des clients et prévoir les améliorations produits. De plus, les compagnies d'énergie et les autres fournisseurs tiers analysent des données de conduite collectées sur des millions de kilomètres afin de décider où installer de nouvelles stations de recharge et comment protéger les réseaux d'électricité fragiles contre la surcharge.⁷ Grâce aux Big Data, les entreprises du monde entier perfectionnent leurs engagements au service de leurs clients et améliorent leurs fonctionnements. Des entreprises telles que McLeod Russel India Limited ont réussi à éliminer totalement les temps d'arrêt des systèmes dans le commerce du thé grâce à un meilleur suivi des récoltes, de la production et de la vente de plus de 100 million de kilos de thé chaque année.⁸ Premier Healthcare Alliance a utilisé le partage de données et l'analytique avancée afin d'améliorer les résultats pour les patients tout en réduisant les dépenses de 2,85 milliards de dollars.⁹

La compagnie Santam quant à elle a pu améliorer l'expérience client en mettant en place l'analytique prédictive pour réduire la fraude (voir l'encadré, « Santam : l'analytique prédictive améliore la détection des fraudes et accélère le traitement des déclarations »).

En plus des objectifs orientés client, d'autres objectifs fonctionnels sont également identifiés dans les premières initiatives associées aux Big Data. L'optimisation fonctionnelle, par exemple, a été citée par 18% des personnes interrogées même s'il s'agit principalement de projets pilotes. D'autres applications du modèle Big Data fréquemment citées incluent : la gestion des risques/gestion financière, la collaboration entre employés et la mise en place de nouveaux modèles de gestion.

Résultats par objectifs fonctionnels



Les participants ont été invités à classer les principaux objectifs fonctionnels associés aux Big Data dans leurs organisations. Les réponses ont été pondérées et cumulées. Nb total de personnes interrogées = 1067

Figure 4 : Près de la moitié des personnes interrogées déclarent axer leurs efforts relatifs aux Big Data sur les résultats orientés client.

Santam : l'analytique prédictive améliore la détection des fraudes et accélère le traitement des déclarations¹⁰

La lutte contre la fraude est l'un des enjeux réels majeurs des compagnies d'assurance dans le monde. Qu'il s'agisse d'une infraction à grande échelle comme un incendie criminel ou une fraude moins grave comme une facture de réparation automobile gonflée, le paiement des déclarations frauduleuses coûte des millions de dollars aux entreprises chaque année, et ce montant se répercute sur les clients sous la forme d'une augmentation de la prime d'assurance. Les compagnies d'assurance combattent les fraudes mais les techniques traditionnelles comme les procédures judiciaires et les enquêtes privées demandent du temps et ont un coût inabordable.

Santam, plus grande compagnie d'assurance à court terme d'Afrique du sud, a évidemment ressenti les effets de la fraude à l'assurance. Les pertes causées par la fraude représentent 6 à 10 % du coût des primes d'assurance annuelles pour les clients de Santam. La fraude a une autre conséquence : l'inefficacité opérationnelle. Alors que les agents doivent traiter des déclarations à haut risque comme des déclarations à faible risque, il faut maintenant trois jours au minimum pour gérer toutes ces déclarations ; la réputation de Santam pour bons services à la clientèle s'est ainsi dégradée face à des clients dans l'attente de résultats toujours plus rapides.

Santam est désormais capable de déceler la fraude de manière anticipée grâce à la solution d'analytique avancée permettant de recueillir les données des déclarations, d'évaluer chaque déclaration par rapport aux facteurs de risque identifiés et de segmenter les sinistres en cinq catégories afin de distinguer les déclarations probablement frauduleuses et à risque élevé des cas à faible risque. Ce nouveau système a permis à la compagnie non seulement de réduire la fraude à l'assurance de plusieurs millions de dollars mais également de diminuer considérablement les temps de traitement des déclarations à faible risque ; certains dossiers peuvent ainsi être résolus en moins d'une heure. Au cours des premiers mois qui ont suivi la mise en place du système, Santam a découvert un groupe de fraudeurs à l'assurance automobile. Les Big Data, l'analytique prédictive et la segmentation des risques ont contribué à l'identification de schémas permettant à la compagnie de déceler ce système de fraude.

Les Big Data sont fondées sur une gestion de l'information évolutive et extensible

La promesse de créer une véritable valeur ajoutée mesurable grâce aux Big Data ne peut être tenue que si les organisations mettent en œuvre une gestion de l'information permettant de prendre en charge un volume, une variété et une vélocité croissants de données. Les personnes interrogées ont été invitées à définir l'état actuel de leurs infrastructures de Big Data. La moitié d'entre elles environ ont signalé avoir démarré leur développement Big Data en se basant sur une gestion de l'information intégrée, évolutive, extensible et sécurisée. Quatre composantes de gestion de l'information ont été le plus souvent citées par les personnes interrogées pour faire état de leurs initiatives de Big Data (voir figure 5).

Un système d'information intégré est un composant clé de tout développement analytique et son rôle est encore plus important dans le cas des Big Data. Selon l'étude de l'IBM Institute for Business Value réalisée en 2011 et portant sur l'analytique avancée, les données d'une organisation doivent être rapidement disponibles et facilement accessibles aux personnes et systèmes qui en ont besoin.¹¹

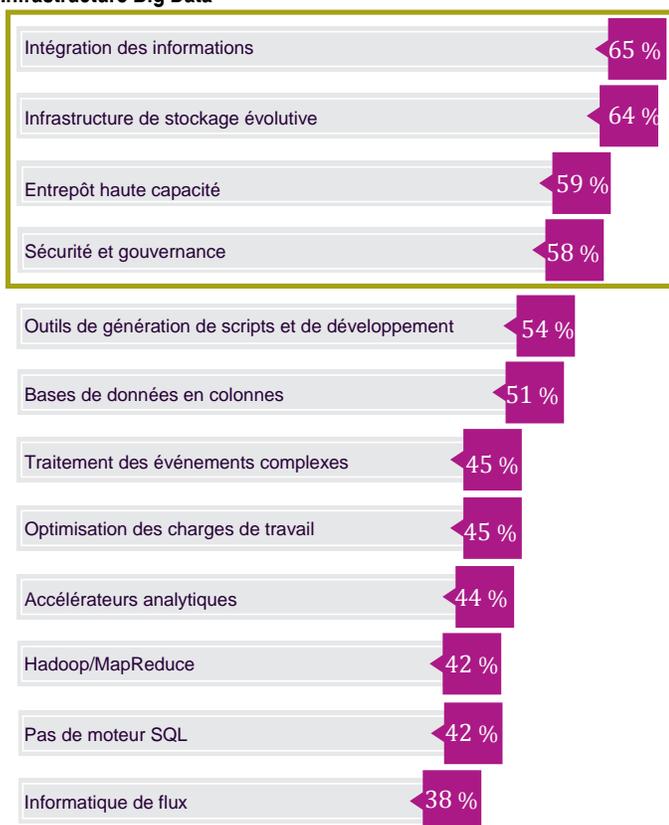
La gestion des données de référence et l'intégration des types de données clés (client, produit, fournisseur, employé et autres données connexes) exigent des données multisectorielles régies par une norme d'entreprise unique. L'incapacité à connecter des données entre plusieurs organisations et départements a constitué pendant de longues années l'un des défis majeurs de la veille stratégique. Cette intégration est encore plus importante, et d'autant plus complexe, avec les Big Data. Parmi les organisations ayant déclaré avoir mis en œuvre des initiatives de Big Data, 65 % considèrent leur capacité de gestion intégrée des informations comme suffisante pour prendre en charge les Big Data.

Après la gestion intégrée d'informations, les éléments de gestion de l'information les plus importants dans le cadre du développement Big Data sont une infrastructure de stockage évolutive et un entrepôt à capacité élevée. Ces deux éléments permettent en effet de prendre en charge le flux croissant de données actuelles et futures entrant dans l'organisation.

Au premier abord, l'association d'un stockage supplémentaire et d'un ou plusieurs serveurs plus importants permettent de prendre en charge l'infrastructure croissante de gestion de l'information. Toutefois, il est important de comprendre que l'anticipation et l'architecture de l'infrastructure sont des éléments clés qui permettront de créer la valeur ajoutée et d'atteindre ainsi les objectifs de l'organisation. Les organisations doivent trouver la meilleure façon de gérer les fluctuations des données afin de permettre aux utilisateurs d'accéder aux données quand ils en ont besoin et d'analyser ces données dans les contraintes de temps imposées par l'organisation (en jours, heures, secondes ou millisecondes). L'équilibre de la configuration et du déploiement de serveurs et de systèmes de stockage favorisent la création d'une infrastructure optimisée.

Ces technologies gèrent également la vélocité croissante des données entrantes (et stockées) en permettant un flux cohérent et automatisé des données au sein de l'entreprise, car de plus en plus de personnes doivent accéder à des types d'informations supplémentaires et différents. Les technologies émergentes, telles que le tiering et la compression des données, ou encore les systèmes de fichiers étendus et les bases de données en mémoire, permettent la gestion de charges de travail plus importantes que les entrepôts conventionnels. Pour de nombreuses organisations, l'optimisation de la gestion de volumes croissants est la première grande priorité Big Data, suivie de près par la capacité à prendre en charge une variété croissante de données. (Voir l'encart, « Vestas : L'amélioration des capacités d'analyse des données permet de réduire les coûts et d'optimiser l'efficacité. »)

Infrastructure Big Data



On a demandé aux participants ayant des initiatives Big Data actives d'indiquer les composants de plateforme actuellement en phase pilote ou intégrés dans l'architecture. Chaque point de données a été collecté indépendamment. Le nombre total de participants pour chaque point de données est compris entre 297 et 351

Des processus de sécurité et de gouvernance solides sont en place dans 58 % des organisations qui indiquent avoir des initiatives actives Big Data en cours. Si la sécurité et la gouvernance sont depuis longtemps une partie inhérente des activités de business intelligence, de nouvelles considérations légales, éthiques et réglementaires liées aux Big Data introduisent de nouveaux risques, notamment en termes de malades publiques, comme nous l'avons déjà vu dans certaines sociétés ayant perdu le contrôle des données ou les ayant utilisées de manière discutable.

En conséquence, la sécurité des données, et surtout la confidentialité des données, est une partie critique de la gestion des informations, selon plusieurs experts en la matière et cadres interrogés. La sécurité et la gouvernance deviennent encore plus importantes et préoccupantes lorsque les organisations intègrent de nouvelles sources d'informations, surtout des données issues de médias sociaux. Par ailleurs, les réglementations sur la confidentialité évoluent encore et peuvent varier grandement en fonction du pays.

« La confidentialité et la sécurité sont perçues comme des tâches faciles, mais elles sont très réglementées, et surveillées très étroitement », explique un cadre d'une société de télécommunications. Et la surveillance n'est pas seulement effectuée par les agences gouvernementales, mais aussi par les clients eux-mêmes. Ce cadre a ajouté : "« Il y a un certain nombre de nouveaux secteurs - tels que les données de navigation sur le Web - où une zone d'ombre existe entre ce qui est légal et ce qui ne l'est pas.

Nous avons choisi comme approche de prendre en compte toute action en fonction de ce que le client penserait si (la manière dont nous utilisons les données) faisait la une de la presse. »

Figure 5 : Composants des infrastructures Big Data des participants.

Vestas : L'amélioration des capacités d'analyse des données permet de réduire les coûts et d'optimiser l'efficacité.¹²

Les éoliennes sont un investissement de plusieurs millions de dollars avec une durée de vie moyenne de 20 à 30 ans. Pour déterminer le positionnement optimal d'une éolienne, un grand nombre de facteurs liés à l'emplacement doivent être pris en compte, y compris la température, les précipitations, la vitesse du vent, l'humidité et la pression atmosphérique.

Pour Vestas Wind Systems A/S (Vestas), un fabricant d'éoliennes danois, le processus d'analyse des données utilisé pour créer les modèles de localisation d'éolienne pour ses clients devenait de plus en plus insatisfaisant ; le processus prenait plusieurs semaines pour s'exécuter et ne pouvait pas prendre en charge l'intégralité de l'analyse des données que la société estimait nécessaire pour un positionnement précis des éoliennes et pour la prévision de la production d'électricité. Les ingénieurs de Vestas ont voulu développer leurs propres modèles de prévision basés sur des données réelles enregistrées, au lieu des modèles industriels ; le défi consistait à augmenter la capacité de gestion des données pour atteindre les six pétaoctets anticipés.

A l'aide d'une solution Big Data exécutée sur un superordinateur parmi les plus performants existant à ce jour, et d'une solution de modélisation conçue pour recueillir des informations correspondant à un très grand nombre de facteurs (données structurées et non structurées), la société peut maintenant aider ses clients à optimiser le positionnement des éoliennes et, en conséquence, leurs performances.

Ce nouvel environnement d'information permet à la société de gérer et d'analyser les données météorologiques et de positionnement de manière inédite, afin d'obtenir des informations permettant d'optimiser les décisions quant au positionnement et à l'exploitation des éoliennes, ainsi que des prévisions plus précises concernant la production d'électricité. Des modèles détaillés se traduisent par des études de faisabilité plus fiables, des résultats plus rapides et une prévisibilité et une fiabilité plus importantes, qui diminuent les coûts par kilowatt produit pour les clients et augmentent la précision des estimations de retour sur investissement. Ces technologies réduisent d'environ 97 % le temps de réponse pour les requêtes des utilisateurs métier (réponses obtenues en quelques heures au lieu de plusieurs jours) et améliorent grandement l'efficacité du positionnement des éoliennes.

Les coûts associés à la mise à niveau des infrastructures ont été signalés comme un problème par plusieurs cadres interrogés. La Direction, ont-ils indiqué, demande une étude de faisabilité solide et quantifiable, qui prévoit des investissements progressifs et indique des opportunités de rationalisation et d'optimisation des coûts pour l'environnement de gestion des informations. Des architectures à coûts plus faibles (notamment le cloud computing, l'externalisation stratégique et la tarification en fonction de la valeur) ont été citées comme des tactiques en cours de déploiement. Cependant, d'autres sociétés ont fait le choix d'investir dans leur plateforme d'informations car elles considèrent que les opportunités commerciales qui en découlent valent largement les investissements consentis.

En matière de Big Data, les efforts initiaux se concentrent sur l'obtention d'informations à partir de sources de données internes existantes et nouvelles.

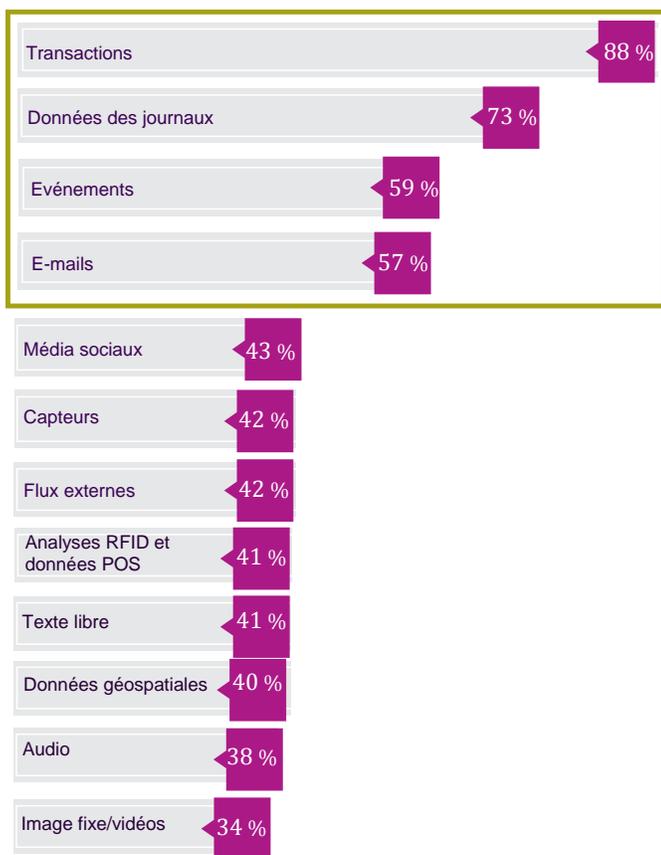
Au départ, les efforts en termes de Big Data sont ciblés sur la collecte et l'analyse des données internes. Selon notre enquête, plus de la moitié des participants ont indiqué que les données internes sont la principale source de Big Data dans leur organisation. Cela laisse entendre que les sociétés adoptent une approche pragmatique pour la mise en œuvre des initiatives Big Data et que ces systèmes internes renferment encore un grand potentiel inexploité (voir la Figure 6).

Comme prévu, les données internes sont les données les plus matures et les mieux comprises pour les organisations. Elles sont collectées, intégrées, structurées et normalisées depuis des années, grâce à la planification des ressources d'entreprise, à la gestion des données de référence, aux initiatives de business intelligence et à d'autres tâches associées. A l'aide des analyses, les données internes extraites des transactions clients, des interactions, des événements et des e-mails peuvent fournir des informations de grande valeur (voir l'encart « Automercados Plaza's : Des revenus accrus grâce à des informations plus complètes »). Cependant, dans de nombreuses organisations, le volume et la portée de ces données internes, telles que les transactions détaillées et les données des journaux d'exploitation, sont devenus trop importants ou hétérogènes pour être gérés dans les systèmes traditionnels.

Presque trois participants sur quatre avec un projet Big Data actif analysent les données issues de journaux. Ce sont des données « générées par une machine / un capteur » visant à enregistrer les détails des fonctions automatisées exécutées dans les systèmes métier ou d'informations, c'est-à-dire des données trop volumineuses pour les capacités de stockage et d'analyse de la plupart des systèmes traditionnels. Par conséquent, la plupart de ces données sont collectées, mais pas analysées.

Les cadres interrogés ont confirmé que la plupart des Directeurs des technologies de l'information chargés des initiatives Big Data de leur société commencent par ces sources inexploitées d'informations internes, à l'aide de la puissance de traitement supplémentaire fournie par une infrastructure plus évolutive.

Sources Big Data



On a demandé aux participants avec des initiatives Big Data actives d'indiquer les sources de données utilisées actuellement pour la collecte et l'analyse des données. Chaque point de données a été collecté indépendamment. Le nombre total de participants pour chaque point de données est compris entre 557 et 867

Figure 6 : Les organisations utilisent principalement des sources de données internes pour les initiatives Big Data.

Automercados Plaza's : Des revenus accrus grâce à des informations plus complètes¹³

Automercados Plaza's, une chaîne familiale de supermarchés au Venezuela, s'est retrouvée avec plus de six téraoctets de données produit et client sur différents systèmes et bases de données. Par conséquent, la société ne pouvait pas évaluer aisément les opérations de chaque magasin et les cadres savaient que des informations de valeur restaient inexploitées.

« Nous étions confronté à une grande pagaille en ce qui concerne la tarification, les stocks, les ventes, la distribution et le merchandising », déclare Jesus Romero, Directeur des technologies de l'information chez Automercados Plaza's. « Nous avons environ 20 millions de dollars en stock et nous recherchons les informations correspondantes dans différents systèmes et les compilons manuellement. Nous avons donc besoin d'une vue intégrée pour comprendre exactement ce que nous avons ».

En intégrant les informations à l'échelle de l'entreprise, la chaîne de supermarchés a réalisé près de 30 % d'augmentation de revenus et une augmentation de 7 millions de dollars en rentabilité annuelle. M. Romero attribue ces excellents résultats à une meilleure gestion des stocks et à la capacité à s'adapter plus rapidement aux évolutions du marché. Par exemple, la société a empêché la perte d'environ 35 % de ses produits, car elle peut maintenant planifier des réductions de prix pour vendre les denrées périssables avant leur date d'expiration.

Un projet Big Data nécessite des capacités d'analyse puissantes

Un environnement Big Data ne crée pas de valeur, jusqu'à son utilisation pour résoudre d'importants défis commerciaux. Cela nécessite un accès à plusieurs types de données différents, ainsi que des capacités d'analyse puissantes qui incluent des outils logiciels et les compétences requises pour les utiliser.

Il s'avère que les organisations s'engageant dans des activités Big Data démarrent avec un noyau puissant de capacités d'analyse conçues pour traiter les données structurées. Ensuite, elles ajoutent des fonctions pour tirer parti des très nombreuses données semi-structurées entrant dans l'organisation (données pouvant être converties en formulaires de données standard) et des données non structurées (données non standardisées).

Plus de 75 % des participants avec des initiatives Big Data actives ont indiqué qu'ils utilisent des fonctions d'analyse de base, telles que les requêtes et la génération de rapports, et le data mining pour analyser les Big Data, tandis que plus de 67 % indiquent utiliser une modélisation prédictive. L'utilisation de ces fonctions d'analyse de base est une manière pragmatique de commencer à interpréter et à analyser les Big Data, surtout lorsqu'elles sont stockées dans une base de données relationnelle. (voir la Figure 7).

Le besoin en fonctions de visualisation de données plus avancées augmente avec l'introduction des Big Data. Les volumes de données sont souvent trop importants pour que les analystes métier et de données puissent les consulter et les analyser avec les outils de génération de rapports et de data mining conventionnels. Dans notre étude, les participants ont déclaré que 71 % des efforts en matière de Big Data reposent sur des compétences de visualisation des données.

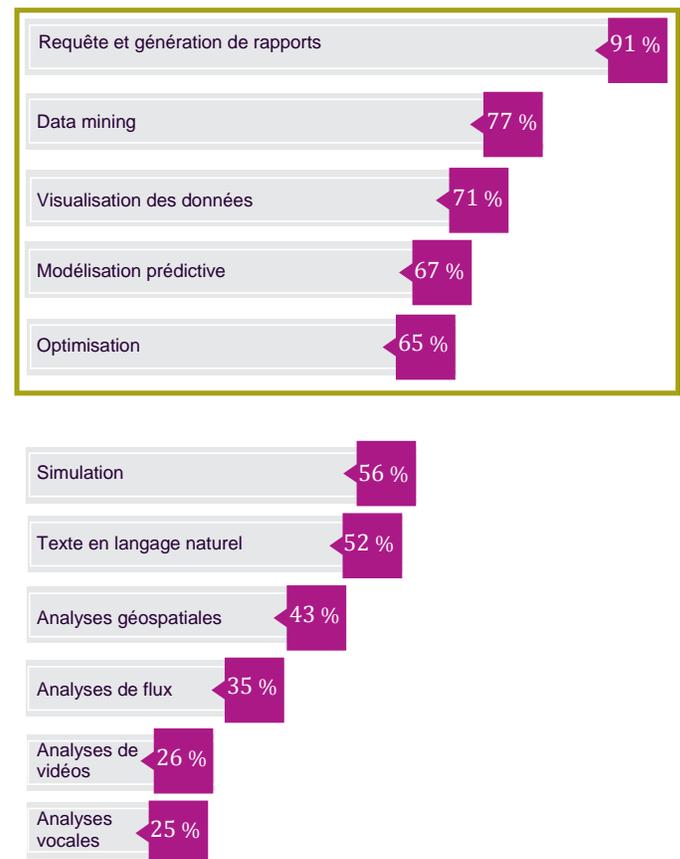
Les organisations engagées dans une initiative Big Data ont de plus en plus besoin de fonctions avancées pour trouver des schémas logiques dans la complexité inhérente. Pour ce faire, les participants appliquent des modèles d'optimisation et des analyses avancées pour mieux comprendre comment transformer les processus métier clés. Ils utilisent des fonctions de simulation pour analyser la myriade de variables présentes dans un environnement Big Data. Notre enquête a indiqué que plus de 50 % des initiatives Big Data actives utilisent ces fonctions de modélisation avancées.

Aujourd'hui, la plupart des sociétés orientent leurs efforts en matière de Big Data sur l'analyse des données structurées, dans un premier temps. Mais un environnement Big Data nécessite également l'analyse de plusieurs types de données, qui peuvent être entièrement nouveaux pour de nombreuses organisations. Dans plus de la moitié des initiatives Big Data, les participants ont indiqué utiliser des fonctions avancées conçues pour analyser le texte dans son état naturel, notamment les transcriptions de toutes les conversations des centres d'appel. Ces analyses comprennent la capacité à interpréter et à comprendre les nuances de la langue, telles que les sentiments, l'argot et les intentions.

Mettre en place la capacité à analyser des données non structurées (par exemple, données de positionnement géospatial, vocales et vidéos) ou en flux continu reste un défi pour la plupart des organisations. Si le matériel et les logiciels sont maintenant au point dans ces domaines, les compétences restent rares. Moins de 25 % des participants avec des initiatives Big Data actives ont indiqué avoir les capacités requises pour analyser des données non structurées telles que les données vocales et les vidéos.

L'acquisition ou le développement de ces compétences analytiques et techniques plus avancées requises pour la progression d'une initiative Big Data devient le défi principal dans de nombreuses organisations avec des initiatives Big Data actives. Parmi ces organisations, le manque de compétences analytiques avancées est un frein majeur à une exploitation plus approfondie des Big Data.

Fonctions d'analyse des Big Data



On a demandé aux participants avec des initiatives Big Data actives d'indiquer les fonctions d'analyse actuellement disponibles dans leur organisation. Chaque point de données a été collecté indépendamment. Le nombre total de participants pour chaque point de données est compris entre 508 et 870.

Figure 7 : Les participants appliquent une grande variété d'analyses avancées.

Le modèle émergent pour l'adoption de l'environnement Big Data se concentre sur la fourniture de valeur métier mesurable

Pour mieux comprendre le paysage du Big Data, nous avons demandé aux participants de décrire le niveau des activités Big Data dans leur organisation aujourd'hui. Les résultats suggèrent quatre niveaux principaux d'adoption du Big Data et une progression le long d'un continuum que nous avons appelé Prise de conscience, Exploration, Engagement et Exécution (voir la Figure 8).

Prise de conscience : développement d'une base de connaissances (24 % des participants)

À l'étape de Prise de conscience, l'objectif principal se concentre sur la découverte et le développement des connaissances. Près de 25 % des participants ont signalé qu'ils n'utilisent pas encore l'approche Big Data dans leur organisation. Si certaines personnes en savent relativement peu au sujet des Big Data, nos entretiens suggèrent que la plupart des organisations à cette étape étudient les avantages potentiels des technologies et des analyses Big Data et essaient de mieux comprendre comment les Big Data peuvent aider à traiter des opportunités commerciales importantes sur le secteur ou le marché concerné. Au sein de ces organisations, ce sont principalement des individus qui rassemblent les connaissances, et non des groupes de travail formels, et leurs conclusions ne sont pas encore utilisées par l'organisation. En conséquence, le potentiel des Big Data n'a pas encore été totalement compris et intégré par les cadres métier.

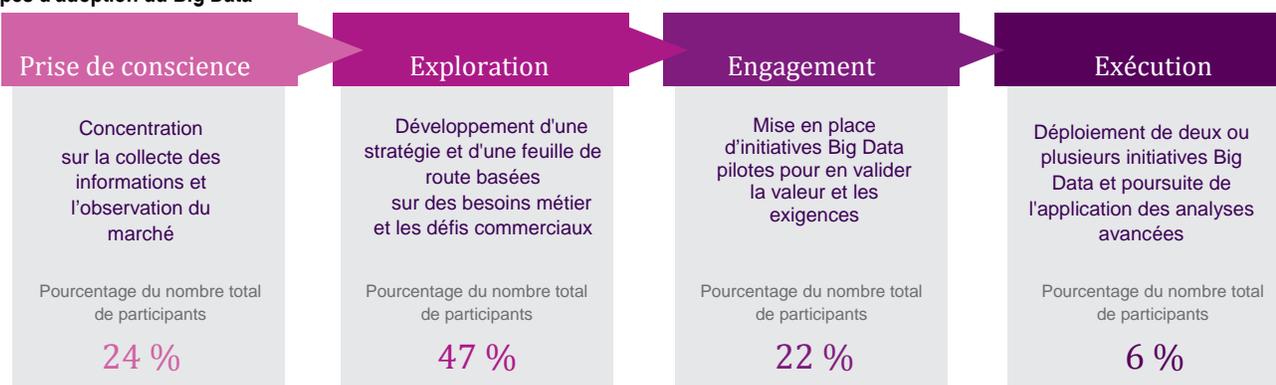
Exploration : Définition de l'étude de faisabilité et de la feuille de route (47 %)

L'objectif de la phase d'exploration est de développer une feuille de route pour le projet Big Data au sein de l'organisation. Presque la moitié des participants ont déclaré tenir des discussions formelles et régulières au sein de leur organisation sur la manière d'utiliser les Big Data pour résoudre d'importants défis commerciaux. Les objectifs clés de ces organisations comprennent le développement d'une étude de rentabilité quantifiable et la création d'un plan directeur pour les Big Data. Cette stratégie et cette feuille de route prennent en considération les données, la technologie et les compétences existantes, puis indiquent par où commencer et comment développer un plan cohérent avec la stratégie commerciale de l'organisation.

Engagement : intégration des Big Data (22 %)

À l'étape d'Engagement, les organisations commencent à prouver la valeur métier des Big Data, et à effectuer une évaluation de leurs technologies et compétences. Parmi les organisations interrogées, plus d'une sur cinq est actuellement en train de développer des études de faisabilité afin de valider les exigences associées à l'implémentation des initiatives Big Data, et d'évaluer les retours attendus. Les organisations de ce groupe travaillent dans un périmètre bien défini, afin comprendre et tester les technologies et les compétences requises pour capitaliser sur les nouvelles sources de données.

Étapes d'adoption du Big Data



On a demandé aux participants d'identifier l'état actuel des activités Big Data au sein de leur organisation. Le pourcentage n'est pas égal à 100 % car les chiffres sont arrondis. Nb total de participants = 1 061

Figure 8 : Quatre étapes dans le modèle d'adoption Big Data émergent.

Exécution : Implémentation de la solution Big Data à l'échelle de l'entreprise (6 %)

A l'étape Exécution, les capacités Big Data et analytiques sont rendues opérationnelles et implémentées à l'échelle de l'organisation. Cependant, seuls 6 % des participants ont signalé que leur organisation a déjà implémenté deux ou plusieurs solutions Big Data à l'échelle de l'entreprise, seuil correspondant au début de cette étape. Le nombre restreint d'organisations parvenues à l'étape d'exécution est cohérent avec les implémentations que nous voyons sur le marché. Ces organisations pionnières optimisent les Big Data afin de transformer leurs activités et exploitent ainsi largement le potentiel des informations dont elles disposent. La vitesse d'adoption des stratégies Big Data par les entreprises augmentant très rapidement, comme le montrent les 22 % de participants à l'étape d'Engagement, avec leurs études de faisabilité ou projets pilotes en cours, nous pensons que le pourcentage des organisations à cette étape doublera dans l'année à venir.

Plus de changement par étape au fur et à mesure de l'évolution des capacités Big Data

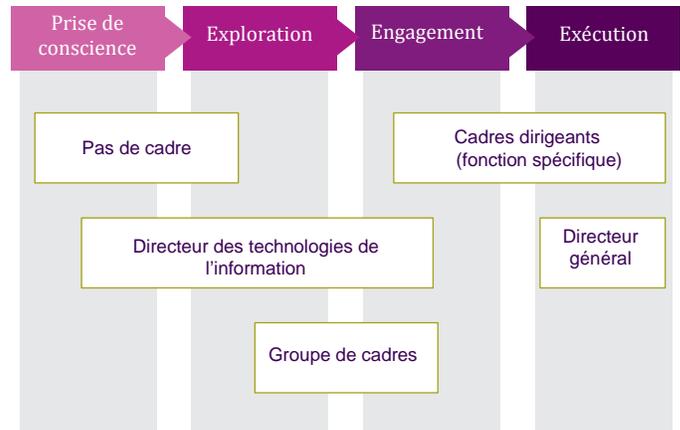
En plus des cinq résultats clés précédents, notre analyse a également révélé plusieurs autres schémas relatifs aux différentes étapes d'adoption : soutien actif des cadres, exigences de disponibilité des données et principaux obstacles. Ces schémas révèlent les étapes interconnectées que les organisations suivent pour faire évoluer leurs infrastructures Big Data et leurs fonctions d'analyse.

Soutien actif d'un projet Big Data

Un examen plus approfondi de l'adoption des Big Data montre une évolution intéressante du soutien apporté au projet (voir la Figure 9). Au sein de notre échantillon, les participants indiquent que plus d'un tiers des efforts Big Data sont menés par le Directeur des technologies de l'information. Cependant, la majorité de ces efforts se situe en tout début de projet car l'organisation investit dans la technologie et commence à identifier les exigences et les opportunités commerciales.

Au fur et à mesure de la progression du projet, le soutien actif est fourni par un cadre métier (Responsable financier ou Responsable marketing par exemple), voire par le Directeur général lui-même. Ce schéma de soutien actif avec un seul cadre commercial est considéré comme essentiel au succès de l'implémentation Big Data.

Soutien actif d'un projet Big Data



On a demandé aux participants d'identifier la personne la plus directement associée au projet d'utilisation des informations issues des Big Data et des analyses. Le positionnement dans le schéma indique à quel niveau un cadre est dominant à chaque étape. Nb total de participants =1028

Figure 9 : Les changements de soutien actif sont évidents au fur et à mesure de l'évolution du projet Big Data dans une organisation.

Ce schéma suggère que les organisations peuvent se concentrer initialement sur la technologie et la construction de leurs infrastructures Big Data, mais au fur et à mesure qu'elles commencent à développer leurs études de faisabilité et leurs feuilles de route, le soutien actif glisse vers un ou plusieurs cadres métier.

Toutefois, le Responsable des technologies de l'information et le service informatique doivent continuer à jouer un rôle important dans l'exécution de la feuille de route définie.

Disponibilité des données

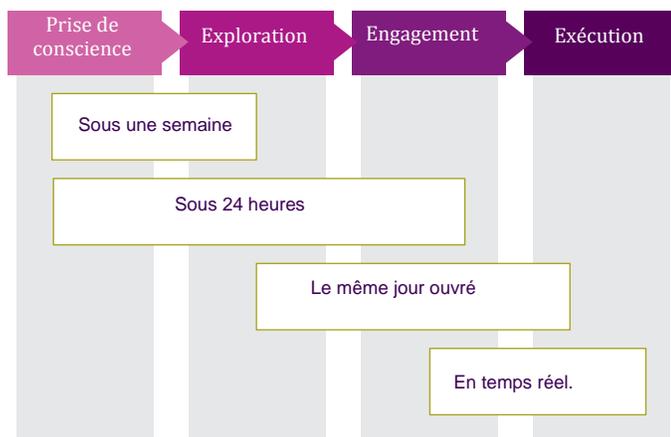
La Figure 10 montre que les besoins en matière de disponibilité des données évoluent considérablement au fur et à mesure que les entreprises font progresser leurs initiatives de Big Data. L'analyse des réponses a révélé que quel que soit le stade d'adoption des Big Data, les organisations sont de plus en plus confrontées au besoin de réduire la latence entre la capture des données et l'action. Il semble que les dirigeants prennent de plus en plus en compte la valeur de données récentes pour prendre leurs décisions métier quotidiennes et stratégiques. Les données ne servent plus seulement à soutenir une décision, elles représentent un élément à part entière de cette prise de décision.

Nous pensons que la demande pour un accès en temps réel devrait continuer d'augmenter alors que les modèles métier évoluent et que les organisations investissent dans des solutions de streaming des données, d'analyse en mémoire, de traitement de machine à machine et autres innovations.

Les obstacles aux Big Data

Les obstacles à l'adoption des Big Data varient au fur et à mesure que les organisations avancent dans l'adoption du concept. Une difficulté perdure toutefois, quel que soit le stade du processus d'adoption, à savoir la capacité d'articuler une analyse de rentabilité attrayante (voir Figure 11). À chaque étape, les initiatives de Big Data font l'objet d'un examen budgétaire. La conjoncture économique mondiale actuelle laisse aux entreprises peu d'appétit pour de nouveaux investissements aux avantages non mesurables, un phénomène non limité au domaine des Big Data. Une fois l'étude de faisabilité réalisée, la plus grande difficulté consiste à trouver les compétences nécessaires pour implémenter les Big Data : compétences techniques, analytiques et en gouvernance.

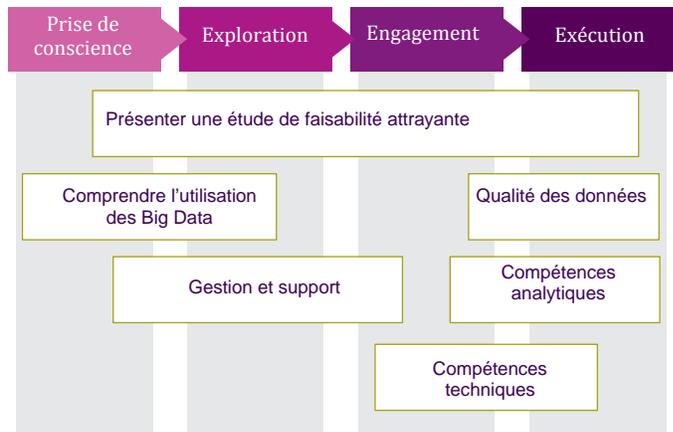
Disponibilité requise des données



Les participants ont été invités à indiquer dans quels délais les données doivent être disponibles aux utilisateurs et processus. La position des encadrés reflète la fréquence de cette exigence dans une étape. Nb total de participants = 973

Figure 10 : Comment obtenir des données en temps réel.

Principaux obstacles



Les participants ont été invités à classer les principales difficultés associées aux Big Data dans leur organisation. La position des encadrés reflète la fréquence de cette difficulté à chaque étape. Les réponses ont été pondérées et cumulées. Nb total de participants = 1062

Figure 11 : Principaux obstacles à l'adoption des Big Data.

Recommandations pour favoriser l'adoption des Big Data

L'analyse par IBM de son étude Big Data @ Work Study a apporté de nouvelles connaissances sur la façon dont les entreprises font évoluer leurs initiatives de Big Data à chaque étape. Face à la nécessité de résoudre leurs problèmes métier, à l'évolution des technologies et à la nature changeante des données, les organisations commencent à examiner de plus près les avantages des Big Data. Pour exploiter au maximum le potentiel des Big Data, nous proposons aux organisations de nombreuses recommandations adaptées à chaque étape du processus d'adoption.

Commencer par centrer les efforts sur le client

Il est impératif que les organisations centrent leurs initiatives de Big Data sur les domaines susceptibles d'avoir le plus d'impact sur leurs résultats. Dans de nombreux secteurs, ceci impose de commencer par une analyse de la clientèle en vue d'un service de meilleure qualité grâce à une meilleure compréhension des besoins des clients et à l'anticipation de leurs comportements futurs.

La numérisation en masse, l'un des facteurs qui a favorisé l'essor des Big Data, a également modifié l'équilibre du pouvoir entre l'individu et l'institution. Pour comprendre les clients et les citoyens, et leur apporter de la valeur ajoutée, les organisations doivent apprendre à connaître leurs clients en tant qu'individus. Elles devront également investir dans de nouvelles solutions et capacités analytiques avancées permettant d'avoir une meilleure vision des interactions et préférences de chacun de leurs clients.

Mais aujourd'hui, qu'il soit un particulier ou une entreprise, le client ne veut pas seulement être compris. Pour qu'il s'intéresse à une relation et qu'elle soit fructueuse, la communication doit se faire en utilisant des méthodes qu'il percevra comme avantageuses.

La valeur ajoutée peut se traduire par des interactions plus opportunes, mieux informées ou plus pertinentes, ou par des opérations sous-jacentes qui améliorent la qualité de ces interactions. D'une façon ou d'une autre, l'analyse nourrit les connaissances issues des Big Data, de plus en plus importantes dans l'approfondissement de ces relations.

Développer un projet de Big Data à l'échelle de l'entreprise

Un avant-projet englobe la vision, la stratégie et les exigences en matière de Big Data dans une organisation. Il est essentiel afin de mettre en correspondance les besoins des utilisateurs métier et le calendrier de mise en œuvre informatique. Cet avant-projet permet de faire comprendre à toute l'entreprise l'utilisation prévue des Big Data pour améliorer les objectifs métier.

Un avant-projet efficace définit la portée des Big Data dans l'organisation en identifiant les problèmes métier clés auxquels il sera appliqué, les processus métier qui déterminent le mode d'utilisation des Big Data, et l'architecture utilisée avec les données, les outils et le matériel nécessaires. Il sert de base à la définition d'un calendrier visant à guider l'organisation dans le développement et la mise en œuvre de ses solutions de Big Data, par le biais d'une approche pragmatique, en vue de créer une valeur métier durable.

Partir des données disponibles pour obtenir des résultats quasiment en temps réel

Pour produire des résultats quasi temps-réel tout en créant la dynamique et l'expertise nécessaires à la conduite d'un programme de Big Data, il faut adopter une approche pragmatique. Comme les participants l'ont confirmé, l'entreprise est l'endroit le plus logique et le plus efficace pour commencer à chercher de nouvelles connaissances.

L'examen des ressources internes permet d'exploiter ses propres données, logiciels et compétences afin d'obtenir une valeur métier quasi temps-réel et d'acquérir une expérience qui sera utile pour envisager d'étendre les capacités en place afin d'utiliser des sources et des types de données plus complexes. La plupart des organisations voudront procéder ainsi pour tirer parti des informations conservées dans leurs référentiels tout en dimensionnant leurs entrepôts de données afin de gérer des données plus nombreuses et plus variées.

Bâtir des fonctions d'analyse basées sur les priorités métier

Partout dans le monde, les entreprises se voient proposer des outils d'analyse toujours plus variés alors qu'elles sont aussi confrontées à une pénurie de compétences analytiques. L'efficacité des Big Data devrait combler ce manque. En bref, les organisations devront investir dans les outils et dans les compétences. Dans le cadre de ce processus devraient apparaître de nouveaux rôles et de nouveaux modèles de carrière pour les personnes proposant le juste équilibre entre les compétences analytiques, fonctionnelles et informatiques.

Les dirigeants doivent accorder une attention particulière à la formation professionnelle et au plan de carrière de leurs analystes internes, qui connaissent déjà les processus métier et les problèmes de l'organisation. Parallèlement, les universités et les individus, quels que soient leur profil ou leur spécialisation, doivent acquérir de solides compétences analytiques.

Créer une étude de faisabilité basée sur des résultats mesurables

La mise en place d'une stratégie de Big Data viable et complète, ainsi que de son calendrier, nécessite une étude de faisabilité solide et quantifiable. Il faut donc s'assurer du soutien et de l'implication active d'un ou plusieurs dirigeants tout au long du processus.

La réussite à long terme du projet dépend également d'une collaboration étroite continue entre l'informatique et les services métier.

Un grand nombre d'organisations justifient leurs études de faisabilité Big Data par les avantages suivants :

- *Des décisions plus intelligentes*, en exploitant de nouvelles sources de données pour améliorer la qualité de la prise de décision.
- *Des décisions plus rapides*, grâce à une capture et une analyse des données davantage en temps réel, au service d'une prise de décision au « point d'impact », par exemple au moment où le client consulte le site Web ou est au téléphone avec un agent du service clients.
- *Des décisions qui font la différence*, en ciblant les domaines dans lesquels les Big Data peuvent faire une réelle différence.

Toutes ces recommandations s'appuient sur un principe important : l'aspect métier et le service informatique doivent travailler ensemble tout au long du projet de Big Data. Les solutions de Big Data les plus efficaces commencent par identifier les besoins métier puis personnalisent l'infrastructure, les sources de données et l'analyse quantitative nécessaires à cette initiative.

Recommandations complémentaires étape par étape : partir de la situation actuelle

Certaines activités clés sont caractéristiques de chaque étape du cycle de vie de l'adoption des Big Data. Les recommandations suivantes, étape par étape, fournissent une approche pratique, éprouvée et progressive.

De la prise de conscience à l'exploration : partir d'une base solide pour agir

- Continuez à développer vos connaissances en ciblant les cas pour lesquels les Big Data apportent un avantage concurrentiel, dans et hors de votre secteur d'activité.
- Travaillez avec différentes divisions et fonctions de l'entreprise pour identifier les opportunités métier et les difficultés les plus critiques, susceptibles d'être résolues grâce à un accès plus efficace et plus opportun aux informations. De nombreuses organisations commencent par les données et l'analyse de la clientèle afin d'améliorer l'efficacité de leur programme de front-office.
- Centrez vos efforts sur le renforcement de votre environnement et de votre infrastructure de gestion des informations, en développant notamment votre avant-projet de Big Data.

Ces avant-projets sont souvent basés sur des normes du secteur, des architectures de référence et d'autres cadres et ressources techniques.

De l'exploration à l'engagement : transformer les plans en action

- Vérifiez que vous bénéficiez du soutien actif de la direction métier lorsque vous mettez au point votre stratégie et votre calendrier de Big Data.
- Développez l'analyse de rentabilité d'une ou plusieurs opportunités métier clés auxquelles vous envisagez de vous attaquer, par le biais de preuves de concept ou d'un ou plusieurs projets pilotes.
- Tout en commençant à prévoir les besoins à plus long terme, vérifiez régulièrement que votre base de gestion des informations et votre infrastructure informatique sont capables de fournir les solutions et les fonctions de Big Data nécessaires à la preuve de concept ou au projet pilote.
- Évaluez vos processus actuels de gouvernance des informations et leur adéquation aux nouveaux aspects des Big Data.
- Analysez les ensembles de compétences disponibles en interne et commencez à analyser les lacunes pour prévoir de développer des compétences ou d'embaucher du personnel supplémentaire.

De l'engagement à l'exécution : comprendre les opportunités et les difficultés futures

- Communiquez largement les bons résultats du projet pilote afin de maintenir la dynamique tout en commençant à impliquer d'autres parties de l'entreprise.
- Finalisez l'analyse de rentabilité en validant et en quantifiant les retours sur investissements et les bénéfices prévus, y compris les critères de réussite et les indicateurs définis.
- Identifiez les modifications et les améliorations des processus métier attendues grâce à l'accès à des informations plus opportunes et de meilleure qualité (marketing, ventes, service clients et sites de réseaux sociaux).
- Mettez au point un plan de compétences pour vous assurer de disposer de compétences techniques adéquates en nombre suffisant afin d'atteindre les objectifs à court et long terme.
- Documentez en détail le programme de migration des projets pilotes en phase de production. Ce programme doit notamment confirmer la valeur métier, les coûts, les ressources et les délais prévus.

Étape d'exécution : exploiter l'innovation associée aux Big Data

- Documentez les résultats quantifiables des premiers succès pour encourager les efforts futurs.
- Mettez en place des communications formelles sur les Big Data dans toute l'organisation pour continuer de soutenir la dynamique et l'aide reçue.
- Élargissez les solutions et les compétences requises pour utiliser les Big Data afin de relever de nouveaux défis dans les divisions, les fonctions et les régions.
- Restez vigilant sur la gouvernance des informations (notamment la gestion du cycle de vie des informations), le respect de la vie privée et la sécurité.
- Continuez d'évaluer les outils et les solutions de Big Data qui évoluent rapidement. Équilibrez l'infrastructure en place avec des solutions plus récentes capables d'améliorer l'évolutivité, l'optimisation et la fiabilité.

Lancez-vous et entrez dans l'ère Big Data

Pour rester compétitif dans l'économie mondiale, il est de plus en plus clair que vous devez comprendre parfaitement les marchés, les clients, les produits, les réglementations, les concurrents, les fournisseurs, les employés, etc. Ceci impose d'utiliser efficacement les informations et les données d'analyse. En fait, de nombreuses entreprises considèrent les informations comme leur ressource la plus précieuse et la meilleure source de différenciation, après leurs employés..

Aujourd'hui, face à l'essor des Big Data, les organisations du monde entier découvrent de nouveaux moyens d'assurer leur compétitivité et leurs ventes. Elles acceptent de se transformer pour tirer parti du vaste ensemble d'informations disponibles afin d'améliorer leur prise de décision et leurs performances à tous les niveaux. C'est ce que fait déjà un groupe relativement restreint d'entreprises innovantes, en dotant tous les employés (des dirigeants jusqu'aux employés du marketing et des ateliers) des informations, des compétences et des outils nécessaires à une prise de décision plus pertinente et opportune au moment du « point d'impact »

Toutes les organisations n'auront pas à gérer tout l'éventail des capacités des Big Data. Cependant, tous les secteurs peuvent utiliser, à un niveau ou à un autre, les nouvelles données, solutions et fonctions d'analyse. Les organisations génèrent de la valeur ajoutée en analysant la quantité et la rapidité des données, anciennes et nouvelles, en se dotant des compétences et des outils adaptés pour mieux comprendre leurs activités, leurs clients et le marché.

Quel que soit leur point de départ, les organisations du monde entier continueront de développer l'utilisation des Big Data pour développer leur activité et renforcer leur avantage concurrentiel dans le contexte de l'économie mondiale.

Pour plus d'informations sur cette étude IBM Institute for Business Value, contactez-nous à iibv@us.ibm.com. Pour obtenir un catalogue complet de nos recherches, consultez le site : ibm.com/iibv

Abonnez-vous à IdeaWatch, notre newsletter mensuelle en ligne qui présente les rapports les plus récents basés sur les recherches de l'IBM Institute for Business Value : ibm.com/gbs/ideawatch/subscribe

Accédez aux rapports IBM Institute for Business Value depuis votre tablette en téléchargeant l'application gratuite « IBM IBV » pour iPad ou Android.

Pour plus d'informations sur la Saïd Business School de l'université d'Oxford, consultez le site www.sbs.ox.ac.uk

Publications connexes

Kiron, David, Rebecca Shockley, Nina Kruschwitz, Glenn Finch et Dr. Michael Haydock, "Analytics: The widening divide: How companies are achieving competitive advantage through analytics" IBM Institute for Business Value in collaboration with MIT Sloan Management Review. Octobre 2011. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-analytics-widening-divide.html> © 2011 Massachusetts Institute for Technology.

LaValle, Steve, Michael Hopkins, Eric Lesser, Rebecca Shockley et Nina Kruschwitz. "Analytics: The new path to value: How the smartest organizations are embedding analytics to transform insights into action." IBM Institute for Business Value in collaboration with MIT Sloan Management Review. Octobre 2010. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-embedding-analytics.html> © 2010 Massachusetts Institute for Technology.

Teerlink, Dr. Marc et Dr. Michael Haycock. "Customer analytics pay off: Driving top-line growth by bringing science to the art of marketing." IBM Institute for Business Value. Septembre 2011. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-customer-analytics.html>

Auteurs

Michael Schroeck est partenaire et vice-président d'IBM Global Business Services, dont il est Global Information Management Foundation Leader. M. Schroeck est également IBM Distinguished Engineer. Contact : mike.schroeck@us.ibm.com

Rebecca Shockley est Business Analytics and Optimization Global Research Leader d'IBM Institute for Business Value. Elle dirige une étude basée sur des faits portant sur l'analyse métier dans l'objectif de renforcer le leadership éclairé des dirigeants. Contact : rshock@us.ibm.com

Le Dr Janet Smart est membre et co-directrice du groupe de recherche Complex Agent-Based Dynamic Networks de la Saïd Business School. Elle est experte dans les systèmes complexes, l'ingénierie des systèmes, les Big Data et la gestion de projet. Elle collabore également au projet ATLAS du CERN au niveau de la gestion de projet et de l'ingénierie des systèmes des projets Big Science. Elle enseigne l'ingénierie des systèmes dans le cadre du MSc de Major Programme Management et du programme britannique Major Projects Leadership Academy. Contact :

janet.Smart@sbs.ox.ac.uk

Le professeur Dolores Romero-Morales enseigne la Recherche opérationnelle à la Saïd Business School. Elle dispense le cours principal sur la Décision et l'analyse des données dans le cadre des programmes MBA et EMBA, et a publié des dizaines d'articles de recherche dans d'importantes revues dans les domaines de l'optimisation de la chaîne d'approvisionnement, d'exploration des données et de gestion du chiffre d'affaires. Elle a également travaillé sur ces thèmes en collaboration avec des acteurs de divers secteurs. Contact : dolores.Romero-Morales@sbs.ox.ac.uk

Le professeur Peter Tufano est doyen de la Fondation Peter Moores de la Saïd Business School. M. Tufano a effectué des recherches spécialisées dans l'innovation, et sur la façon dont elles peuvent améliorer les services financiers proposés aux familles. Avant de venir à Oxford, le professeur Tufano a passé 33 ans à Harvard et a fondé D2Dfund :

claire.eggleton@sbs.ox.ac.uk – Cette adresse ne correspond pas au nom cité

Contributeurs

Fred Balboni, Global Leader Business Analytics and Optimization, IBM Global Business Services

Dr Stephen Buckley, BAO Applied Research Leader, IBM Research

Wendy Olivier, Information Management Foundation Program Director, IBM Global Business Services

Katharyn White, vice-présidente du marketing, IBM Global Business Services

Les auteurs remercient également les personnes suivantes d'IBM et de l'université d'Oxford pour l'aide apportée à la rédaction de ce rapport : Dr John Bell, Min Chen, Michael Coleman, Richard Cuthbertson, Tom Deutsch, Angela Finley, Mark Graham, Larry Gosselin, Tina Groves, Bambi Grundweg, Pamela Hartigan, Bernie Hogan, Matin Jouzdani, Jim Kocis, Eric Lesser, Monica Logan, Eduardo Lopez, Dr. Robin Lougee, Piyush Malik, Helen Margetts, Joni McDonald, Brian Morris, Richard Perret, Nancy Puccinelli, Tarun Ramadorai, Dean Ranalli, Eric Sall, Ralph Schroeder, Stephanie Schneider, Ronald Shelby, Neil Shephard, James Taylor, Owen Tebbutt et Andy Twigg.

Références

- 1 LaValle Steve, Michael Hopkins, Eric Lesser, Rebecca Shockley et Nina Kruschwitz. « *Analytics: The new path to value: How the smartest organizations are embedding analytics to transform insights into action.* » IBM Institute for Business Value en collaboration avec MIT Sloan Management Review. octobre 2010. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-embedding-analytics.html>
© 2010 Massachusetts Institute for Technology.
- 2 « *2012 IBM Global Technology Outlook* » IBM Research. mars 2012. http://www.research.ibm.com/files/pdfs/gto_booklet_executive_review_march_12.pdf
- 3 « *From Stretched to Strengthened: Insights from the IBM Chief Marketing Officer Study* » IBM Institute for Business Value. mai 2011. www.ibm.com/cmstudy; « *Leading Through Connections: Insights from the IBM Chief Executive Officer Study* » IBM Institute for Business Value. mai 2011. www.ibm.com/ceostudy
- 4 Woody, Todd. "Automakers, Tech Companies Mining Electric Car Big Data to Plot Industry's Future." Forbes. 18 Juin 2012. <http://www.forbes.com/sites/toddwoody/2012/06/18/automakers-tech-companies-mining-electric-car-big-data-to-plot-industrys-future/>
- 5 Ibid.
- 6 Ibid.
- 7 Ibid.
- 8 Étude de cas IBM. « *McLEOD RUSSEL INDIA LIMITED: Eliminating downtime in the tea trade with IBM Informix* » 9 septembre 2011. http://www-01.ibm.com/software/success/cssdb.nsf/CS/JHUN-8LFLWH?OpenDocument&Site=dmain&cty=en_us
- 9 Étude de cas IBM. « *Premier Healthcare Alliance* » 30 avril 2012. http://www-01.ibm.com/software/success/cssdb.nsf/CS/JHUD-8TS39R?OpenDocument&Site=wp&cty=en_us
- 10 Communiqué de presse IBM. « *Using IBM Analytics, Santam Saves \$2.4 Million in Fraudulent Claims.* » 9 mai 2012. <http://www-03.ibm.com/press/us/en/pressrelease/37653.wss>
- 11 Kiron David, Rebecca Shockley, Nina Kruschwitz, Glenn Finch et Dr. Michael Haydock. « *Analytics: The widening divide: How companies are achieving competitive advantage through analytics* » IBM Institute for Business Value en collaboration avec MIT Sloan Management Review. octobre 2010. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-analytics-widening-divide.html> © 2011 Massachusetts Institute for Technology.
- 12 Pittman David. « *Lords of the Data Storm: Vestas and IBM Win Big Data Award* » The Big Data Hub: Understanding Big Data for the enterprise. 28 septembre 2012. <http://www.ibmbigdatahub.com/blog/lords-data-storm-vestas-and-ibm-win-big-data-award>
- 13 Étude de cas IBM. « *Automercados Plaza's increases revenue by 30 percent with greater insight into operations* » 15 juillet 2011. http://www-01.ibm.com/software/success/cssdb.nsf/CS/JHUN-8JPSK?OpenDocument&Site=de_fault&cty=en_us



© Copyright IBM Corporation 2012

IBM France
17 avenue de l'Europe
92275 Bois-Colombes
France

Produit en France
Novembre
Tous droits réservés

IBM, le logo IBM et ibm.com sont des marques ou marques déposées d'International Business Machines Corporation aux Etats-Unis et/ou dans d'autres pays. Si ces marques et d'autres marques d'IBM sont accompagnées d'un symbole de marque (® ou ™), ces symboles signalent des marques d'IBM aux Etats-Unis à la date de publication de ce document. Ces marques peuvent également exister et éventuellement avoir été enregistrées dans d'autres pays. La liste actualisée de toutes les marques d'IBM est disponible sur la page Web « Copyright and trademark information » à : ibm.com/legal/copytrade.shtml.

Les autres noms de sociétés, produits ou services peuvent être des marques commerciales ou marques de services appartenant à leurs propriétaires respectifs. Toute référence à des produits et services IBM dans cette publication n'implique pas que la société prévoit de les proposer dans tous les pays où elle est présente.

Certaines parties de ce rapport sont utilisées avec l'autorisation de la Saïd Business School de l'université d'Oxford. © 2012 Saïd Business School de l'université d'Oxford. Tous droits réservés.



Merci de recycler

GBE03519-FRFR00