Point de vue

Exploiter les bénéfices du big data fragmenté

Comment les infomédiaires créeront-ils un nouvel écosystème de données ?

Auteurs
Bill Gerhardt
Kate Griffin
Roland Klemann

Juin 2012

Cisco Internet Business Solutions Group (IBSG)

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Exploiter les bénéfices du big data fragmenté

Comment les infomédiaires créeront-ils un nouvel écosystème de données ?

Les données sont « le nouvel or noir »

Déjà en 2006, l'analyste Clive Humby avait décrit les données comme étant « le nouvel or noir ». [1]Aujourd'hui, alors que les données enflent sans discontinuer pour créer un royaume du big data sans limite, les experts commencent à y croire. » Ils voient dans le big data un potentiel de transformation et d'enrichissement identique à celui que les visionnaires du 19e siècle avaient perçu dans la substance noire et suintante qui s'échappait du sol.

Il y a un siècle ou plus, l'or noir alimentait les rêves. Aujourd'hui, ce sont les données qui ouvrent la voie aux innovations économiques et techniques. Si nous parvenons à extraire, à raffiner et à acheminer les données « brutes » de sorte qu'elles influencent les décisions en temps réel, elles seront de véritables mines d'or. Lorsque nous pourrons y puiser des informations essentielles grâce à une analyse qui révélera des comportements, des tendances et des événements complexes au moment où ils se produiront (si ce n'est *avant*), alors les données concrétiseront tout leur potentiel. Et si nous parvenons à partager convenablement les données sur un écosystème complet et à les rendre disponibles lorsque les analyses sont essentielles, elles transformeront véritablement tout : notre mode de vie et notre façon de travailler, de jouer, d'apprendre, d'acheter, de vendre, de voyager…

Cependant, le big data s'accompagne de défis importants. Le problème à l'heure actuelle, c'est que le déluge de données qui déferle et qui ne cesse de croître (des téraoctets, des pétaoctets, puis des exaoctets) menace de nous submerger d'informations non filtrées, non structurées, non traitées et apparemment impossibles à gérer. Dans l'état actuel des choses, l'écosystème de données est extrêmement fragmenté. Les entités qui génèrent les données et celles qui pourraient en tirer profit se heurtent à un labyrinthe de complexités, de disparités et de problèmes de communication. Si l'analyse doit jouer le rôle de la « raffinerie », il faudra réduire cette fragmentation en renforçant la connectivité, la confiance et l'efficacité.

Qui sera capable d'assumer cette fonction ? Cisco® Internet Business Solutions Group (IBSG) estime que ce rôle de premier plan reviendra à une nouvelle catégorie d'intervenants : les infomédiaires de données. Ces entités entreront en jeu et établiront le lien entre les auteurs des données et ceux qui pourront les exploiter, en valorisant les premiers et en simplifiant chaque maillon de la chaîne de valeur. Dans ce cadre, ils permettront à tous d'harmoniser, de réaliser et de monétiser le vaste potentiel du big data.

[1]http://ana.blogs.com/maestros/2006/11/data_is_the_new.html

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Les infrastructures de transport et de stockage du big bata joueront un rôle essentiel dans ce processus. Les intervenants capables de voir grand et loin dans le domaine des réseaux et des data centers et qui sauront collecter et contrôler les données en mouvement seront les mieux placés pour rassembler les acteurs disparates du big data et former un tout parfaitement rodé.

Dans ce document, nous allons observer les répercussions au sens large du big data et les transformations qu'il promet, et aborder les problèmes rendant sa maîtrise difficile. Nous allons surtout explorer la dynamique potentielle du big data qui se pose en véritable force de transformation pour les infomédiaires de données. Le big data est sans conteste « le nouvel or noir » de notre ère.

La véritable nature du big data

Dans le but de permettre des décisions plus avisées, de nombreuses entreprises utilisent des entrepôts de données et des méthodes traditionnelles d'analyse des données pour réduire les désabonnements des clients, renforcer l'efficacité des campagnes et lutter contre la fraude, pour ne citer que quelques exemples. Le big data représente une étape révolutionnaire faisant suite à l'analyse traditionnelle des données. Il repose sur trois piliers : la variété, le volume et la rapidité.

- Il n'existe que deux **variétés** de données : les données structurées et les données non structurées. Les données structurées débouchent dans un entrepôt de données déjà organisé et sont facilement triées. Pourtant, la plupart des données actuelles ne sont pas structurées. Elles proviennent de sources telles que Facebook, Twitter et de vidéos. Elles sont aléatoires, difficiles à analyser et *envahissantes*.
- Le **volume** brut du big data submerge l'entrepôt de données classique. Par exemple, Facebook indique que ses utilisateurs génèrent 2,7 milliards de « likes » et de commentaires par jour. L'abondance de ces données est impressionnante : les entreprises ne parviennent pas à y faire face et encore moins à les trier, à les analyser et à les exploiter.
- La gestion d'un tel volume de données peut poser problème, surtout lorsqu'elles arrivent à une **vitesse** qui dépasse largement les capacités de traitement de la plupart des entreprises. Pour que le big data puisse transformer le monde, l'analyse doit être effectuée à la vitesse fulgurante à laquelle les informations parviennent dans les entrepôts de données. La décision de collecter ou non un bit spécifique de données doit être prise en quelques microsecondes. Sa pertinence lorsqu'il est combiné à d'autres données doit également être déterminée presque instantanément. Par exemple, le visage d'un délinquant connu parmi des milliers d'images peut déclencher un signal « stop » ; un comportement potentiellement frauduleux peut générer des avertissements et les données indiquant le désabonnement croissant des clients peuvent donner lieu à un coupon de réduction. Dès qu'une anomalie importante apparaît dans les données, elle doit être signalée *avant* le stockage des données afin de permettre l'application de mesures en temps réel.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

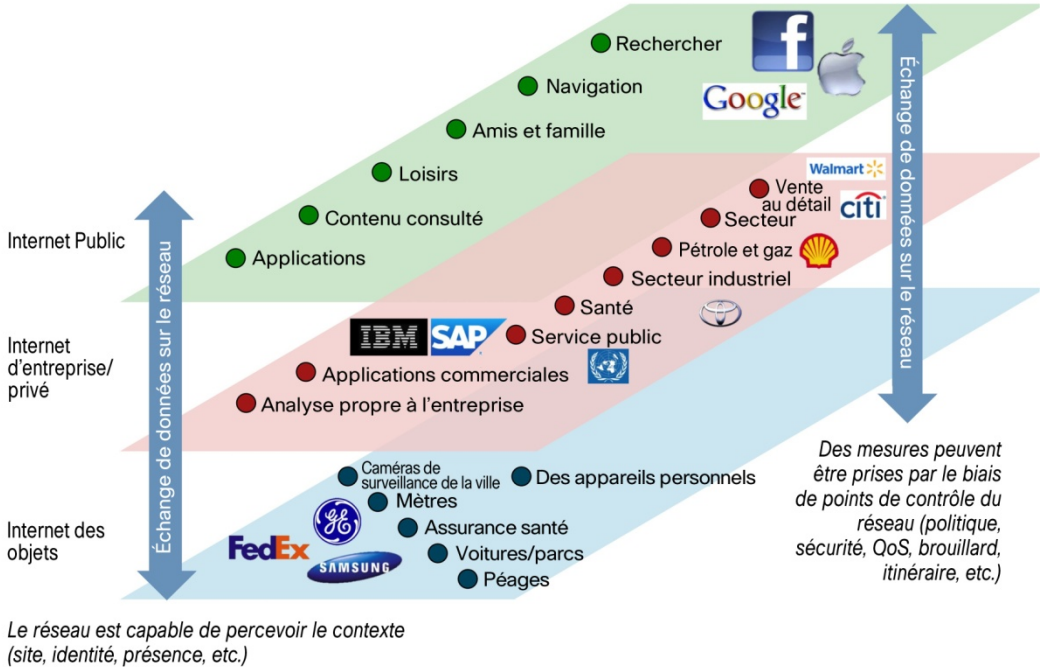
Les promesses du réseau

Par sa nature même, le trafic réseau constitue le big data. Dans la seule portion du réseau mobile, 6 milliards d'abonnements de téléphonie mobile ont été souscrits dans le monde et, chaque jour, 10 milliards de SMS sont envoyés.[2] International Data Corp., la société de recherche internationale sur le marché de l'information, a estimé qu'en 2011, toutes les données créées dans le monde représentaient 1,6 trillion de gigaoctets.

La croissance exponentielle du nombre de données transitant sur les réseaux est loin d'être terminée. D'ici à 2020, 50 milliards d'appareils seront connectés à des réseaux et à Internet.[3] Mais le réseau est bien plus qu'un « simple tuyau » acheminant ces torrents de données. Outre le transport du volume brut, le réseau peut jouer deux rôles importants contribuant à augmenter le potentiel du big data pour les entreprises (voir la figure 1).

Tout d'abord, le réseau peut collecter des données et fournir le contexte très rapidement. Il peut rassembler des données provenant d'appareils, de réseaux privés et de réseaux publics étendus, souvent en « heure réseau », c'est-à-dire que la valeur est extraite et exploitée au moment où les événements surviennent, en temps réel ou presque. Grâce aux données qu'il collecte, le réseau peut déterminer le contexte. Cela peut consister à les recouper avec des informations de géolocalisation, d'identification ou de présence. En outre, le contenu des données collectées peut également être intéressant. Le fait de connaître non seulement les sites visités par les internautes, mais également leurs actions sur ces sites, peut fournir des informations essentielles.

Figure 1. Le réseau connecte les domaines de données.



Source : Cisco IBSG, 2012

[2] Union internationale des télécommunications, 2012.

[3] « L'Internet des objets : Comment l'évolution actuelle d'Internet transforme-t-elle le monde ? », Cisco IBSG, 2011.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Ensuite, le réseau permet aux entreprises de réagir immédiatement aux informations qui leur parviennent. Le réseau et ses points de contrôle en temps réel peuvent avoir un impact sur le client ou influencer l'expérience client. Ces points de contrôle entraînent des modifications des politiques, de la confidentialité, des préférences, de la sécurité, de la qualité de service et du réacheminement ou du traitement local.

Les entreprises maîtrisant le réseau de bout en bout (contexte et contrôle) voient grand et à long terme.

Comment le big data va-t-il transformer votre vie ?

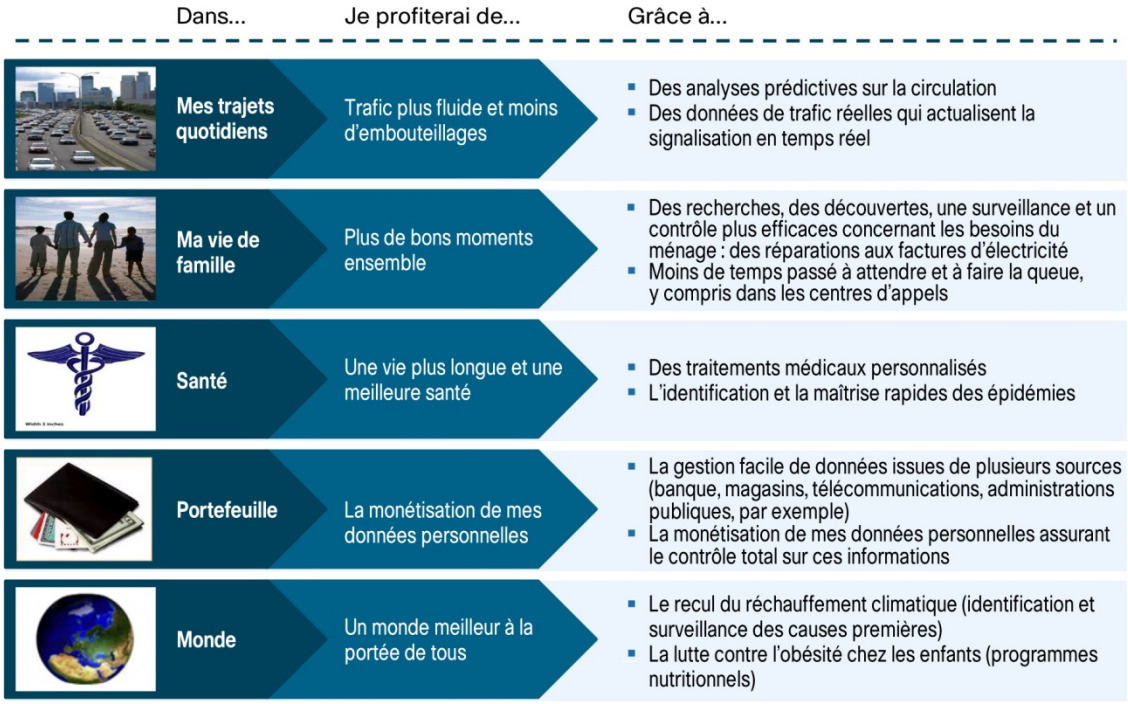
Si le big data réalise tout son potentiel, il transformera une multitude de tâches et d'activités quotidiennes de routine ou plus complexes (voir la figure 2).

Prenons l'exemple des embouteillages. On peut penser que c'est un mal nécessaire et inévitable, mais imaginez que l'on puisse les réduire considérablement, voire les éliminer. Ce serait une amélioration énorme sur le plan du stress et des pertes de temps (sans parler des économies *d'argent)*. La qualité de vie serait meilleure, puisque les temps de trajet seraient réduits et la vie de famille y gagnerait. La solution n'est pas de proposer des voitures volantes, mais plutôt des analyses prévisionnelles et en temps réel de la circulation. Celles-ci entraîneront des modifications immédiates des feux tricolores, des panneaux numériques et des itinéraires, avant même le début des embouteillages.

Et ces tickets de caisse et reçus bancaires qui envahissent votre portefeuille ? Imaginez qu'ils soient remplacés par des justificatifs électroniques. Les entreprises pourraient compléter ces documents par des informations contextuelles et comparatives. Cela permettrait aux particuliers de gérer, de partager, de monétiser et d'utiliser les données, par exemple à l'aide d'applications de gestion du budget et de conseil sur la santé.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Figure 2. La pertinence du big data.



Source : Cisco IBSG, 2012

Les progrès médicaux découlant du big data pourraient notamment inclure la personnalisation des traitements en fonction du dossier de chaque patient. Les épidémies ou les contaminations alimentaires pourraient être anticipées, suivies et écourtées. L'application Google Suivi de la grippe a déjà rassemblé des données sur les recherches d'informations relatives à la santé afin de prévoir les épidémies sur l'ensemble de la planète.

À cette échelle (voire au-delà), le big data permet de surveiller et d'analyser de grandes quantités d'informations sur tous les phénomènes, du changement climatique à l'astrophysique, en passant par la consommation énergétique et les tendances socioéconomiques. Dans tous les cas, l'analyse du big data peut apporter une meilleure connaissance des principales causes des événements majeurs et permettre d'anticiper les tendances.

Mettre le big data au service des entreprises

Le big data peut également transformer nos interactions avec les entreprises. Petit à petit, les interactions des consommateurs avec le monde qui les entoure ont un effet qui passe souvent inaperçu : elles produisent de vastes quantités de données personnelles. Il s'agit notamment de recherches sur la météo, de comparaisons de prix, d'achats et de milliers d'autres opérations et choix quotidiens. Grâce à l'analyse des données générées par toute cette activité, le big data permet de révolutionner les interactions entre les consommateurs et les vendeurs.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Certaines entreprises commencent déjà à exploiter ces possibilités. En voici quelques exemples :

- La société américaine Harrah's a fait progresser son chiffre d'affaires de 8 à 10 pour cent en analysant les données de segmentation des clients,[4] tandis qu'Amazon déclarait il y a quelques années que 30 pour cent de son chiffre d'affaires provenaient de son moteur de recommandations analytiques.[5]
- La chaîne Wal-Mart a intégré la technologie RFID (identification de la fréquence radio) à son système de gestion des stocks afin de connecter en temps réel les informations entre les fournisseurs et son entrepôt de données Retail Link. Cette opération a permis de réduire d'environ 16 pour cent les ruptures de stock. [6]
- FedEx, qui collecte des données sur les expéditions et les clients sur plus de 46 000 sites répartis sur sa chaîne de distribution et d'approvisionnement, a obtenu une visibilité en temps réel sur ses livraisons.
- La société Kaiser Permanente a réduit ses coûts d'exploitation de 40 pour cent grâce à l'intégration de ses dossiers médicaux électroniques et de son système de gestion des patients hospitalisés et traités en ambulatoire. L'intégration de ces données cliniques et financières a permis de faire apparaître les effets indésirables du Vioxx, ce qui a entraîné le retrait du marché de ce médicament.[7]

Par conséquent, le big data est déjà devenu la priorité des directeurs informatiques.[8] Pourtant, Cisco IBSG imagine un niveau encore plus poussé d'analyse des données et d'interaction tout au long de la chaîne de valeur des données. Nous avons analysé quelques centaines d'applications potentielles de big data sur plusieurs marchés verticaux et identifié huit catégories d'exemples d'utilisation :

1. **Garantie de chiffre d'affaires.** L'intégration de données internes à l'entreprise pourrait affiner l'identification et prévenir les fraudes. Les secteurs particulièrement touchés par les abus en seraient les premiers bénéficiaires.
2. **Réduction des risques.** Chaque jour, les réseaux transportent des pétaoctets d'informations critiques pour les entreprises, les organismes publics et les particuliers, ce qui augmente sans cesse les risques d'intrusions et d'atteintes à la sécurité. La fédération des données sur des zones géographiques et des réseaux plus étendus permettrait d'identifier les comportements suspects tout en signalant la nécessité d'une intervention immédiate.
3. **Cycle de vie du client.** Les entreprises peuvent se concentrer sur les désagréments subis par les clients et y répondre immédiatement, afin d'améliorer leur satisfaction et de limiter le taux de désabonnements. Les bénéfices de cette approche seraient à la portée de tous les secteurs tertiaires dont la relation au client pourrait bénéficier.

---

[4] Gartner, 2005.
[5] McKinsey & Company, 2011.
[6] « Does RFID Reduce Out-of-Stocks? », University of Arkansas, 2005 ; « Opening Up the Big Box: Measuring the Wal-Mart Effect », The Economist, 23 février 2006.
[7] McKinsey & Company, 2011.
[8] D'après une enquête menée en 2011 par IBM, 83 pour cent des dirigeants plaçaient le big data au premier rang de leurs priorités, avant les solutions de mobilité, la virtualisation et le cloud computing.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

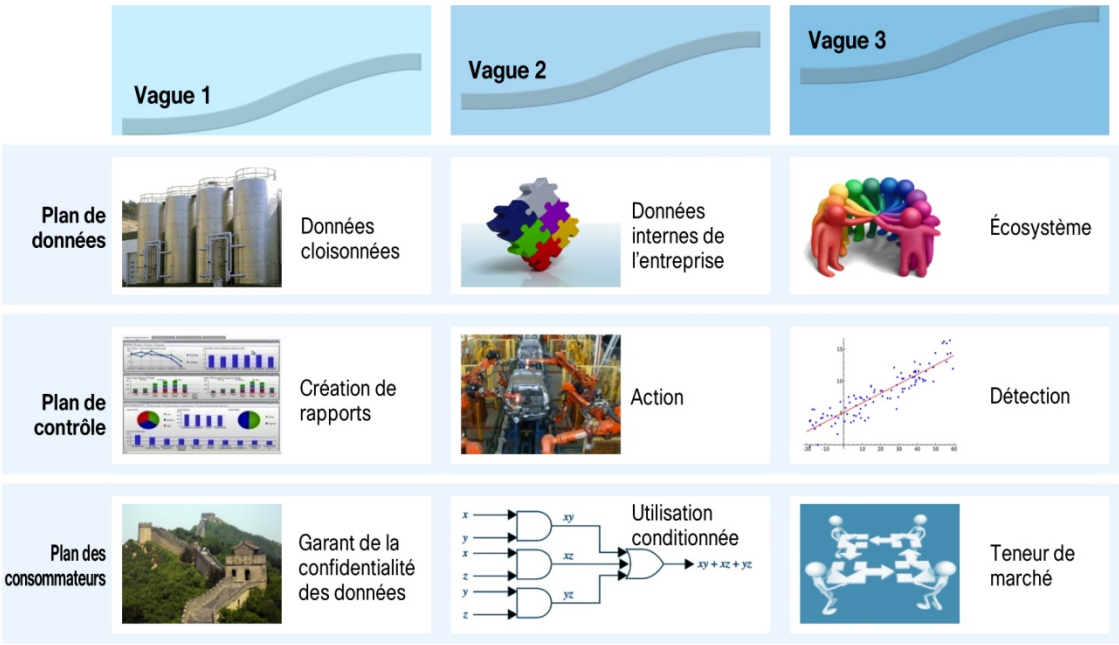
4. **Exécution des services.** Grâce à l'analyse, le big data permet de meilleures prestations de services sur le marché et améliore les possibilités de ventes croisées et de ventes additionnelles. Les établissements bancaires et les commerces électroniques apparaissent comme les bénéficiaires potentiels les plus importants.
5. **Innovation produit.** L'avis des consommateurs est essentiel au développement des produits et actuellement, de nombreuses entreprises souhaitent véritablement en savoir plus sur les appréciations de leurs clients. L'intégration de sources de données externes à l'entreprise, notamment les flux de réseaux sociaux, renseignerait mieux sur l'impression des clients vis-à-vis d'un produit et pourrait même révéler la nécessité d'un nouveau produit *avant* qu'il ne soit inventé ou conçu.
6. **Modèles commerciaux.** De nouveaux algorithmes tarifaires créeront de nouvelles possibilités de monétisation et des interactions plus efficaces avec les clients. Les commerces de détail, en particulier, tireraient d'importants bénéfices en mettant en place de nouveaux modèles de tarification associés au lieu et au comportement des consommateurs.
7. **Gestion des opérations.** Le big data a le potentiel d'aider presque tous les types d'entreprise à améliorer leurs processus et leur efficacité. Les prestataires de services pourraient améliorer les opérations quotidiennes de leurs réseaux. Les détaillants pourraient créer des interactions plus efficaces et plus lucratives dans leurs points de vente. Enfin, toute chaîne d'approvisionnement gagnerait en fluidité. Globalement, une infrastructure commune d'informations améliorerait l'efficacité des processus et fournirait un aperçu complet des ressources.
8. **Publicité avancée.** Plus les entreprises détiennent d'informations sur les comportements des clients, plus il leur est facile de se servir de ces données à des fins publicitaires. La société de recherche eMarketer a déterminé que le ciblage avancé reposant sur une base de données double l'efficacité de la publicité. Il est possible d'augmenter de plus de 50 pour cent le coût par mille. Pour exploiter davantage ces bénéfices à l'aide de l'analyse du big data, des mécanismes de confiance devront réduire les contraintes liées à la confidentialité et garantir l'adéquation des publicités ciblées aux informations sur le client qui transitent sur le réseau.

Tirer parti des phases de l'évolution du big data

Tout le potentiel de ces cas d'utilisation sera concrétisé au cours des trois phases de l'évolution sectorielle (voir la figure 3). À chaque phase, l'analyse peut être décrite sur le plan des données, du contrôle et des consommateurs. Le plan des données concerne les sources de données utilisées. Sur le plan du contrôle, nous nous intéressons à la manière dont les données sont utilisées, à la façon dont les connaissances sont générées et aux processus mis en place lorsqu'un constat a été établi. Sur le plan des consommateurs, nous étudions la relation du consommateur au processus d'analyse des données, sachant que ces personnes génèrent souvent ces données, mais en sont rarement les propriétaires.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Figure 3. L'évolution du secteur du big bata.



Source : Cisco IBSG, 2012

- Actuellement, la plupart des secteurs sont enlisés dans la **première phase.** Les données sont généralement cloisonnées dans certaines divisions de l'entreprise. Elles servent principalement à générer des rapports ex post. Après l'événement, les données sont compilées, mais les rapports ainsi générés n'ont qu'un impact indirect sur les décisions. Les consommateurs ont, quant à eux, peu de contrôle sur leurs données et ne sont même pas toujours conscients qu'elles existent. La sécurité et l'intégrité de leurs données les inquiètent et une vague crainte d'utilisation frauduleuse prédomine.
- Dans la **deuxième phase,** les silos de données deviennent accessibles dans l'ensemble de l'entreprise. Les données peuvent être déplacées entre les différents silos de l'entreprise et vers un référentiel commun. Celui-ci permet l'analyse et le partage d'un ensemble bien plus vaste de données, et permet donc des prises de décision bien plus stratégiques. Sur le plan des données, les nouvelles informations déclenchent des actions. Auparavant, les entreprises disposaient de simples informations sur une seule variable. À présent, plusieurs variables et plusieurs facteurs peuvent générer une alerte qui déclenche une action (la commande d'un produit en stock ou l'envoi d'une offre spéciale à un client, par exemple). Les consommateurs se rendent de plus en plus compte qu'ils peuvent tirer profit de la transmission de leurs données personnelles. Ils peuvent choisir explicitement, parmi les options proposées, comment leurs données seront utilisées.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

- Dans la **troisième phase,** les données sont collectées, partagées et traitées sur l'ensemble de l'écosystème. Les entreprises peuvent non seulement exploiter leurs propres données, mais également une profusion de données provenant de tiers, notamment des informations météorologiques, des indicateurs économiques, des ventes au détail anonymisées, des flux de médias sociaux, etc. Lors de cette phase, les entreprises utilisent les données dans des processus de plus en plus complexes afin d'effectuer des analyses prédictives, en *anticipant* les événements. Une fois ces informations obtenues, l'entreprise peut prendre des mesures automatisées en temps réel, sans aucune intervention humaine (modifier les itinéraires routiers ou réassortir le stock, par exemple). Lors de la 3e phase, les consommateurs commencent à apprécier le fait que leurs données personnelles peuvent leur être bénéfiques. Ils définissent des politiques d'utilisation de leurs données et savent que ces règles sont appliquées à l'ensemble de l'écosystème. Plus celui-ci est transparent, plus les consommateurs ont confiance, car ils peuvent observer la façon dont leurs données sont utilisées. Ils contribuent souvent volontairement à ce marché où leurs données sont valorisées et leur apportent des bénéfices concrets.

Toutes ces idées font leur chemin et, d'après Cisco IBSG, elles ont déjà un impact. Entre la 1re et la 2e phase, les bénéfices se multiplient, mais les obstacles peuvent également s'accentuer.

Les obstacles à l'exploitation universelle du big data et les moyens de les surmonter

L'objectif final du big data est de collecter, de stocker, d'exploiter les données et d'en tirer profit. Actuellement, l'écosystème du big data est fragmenté. Il est pourtant possible d'en orchestrer tous les aspects pour permettre au créateur et au bénéficiaire des données de travailler de concert dans un contexte de confiance.

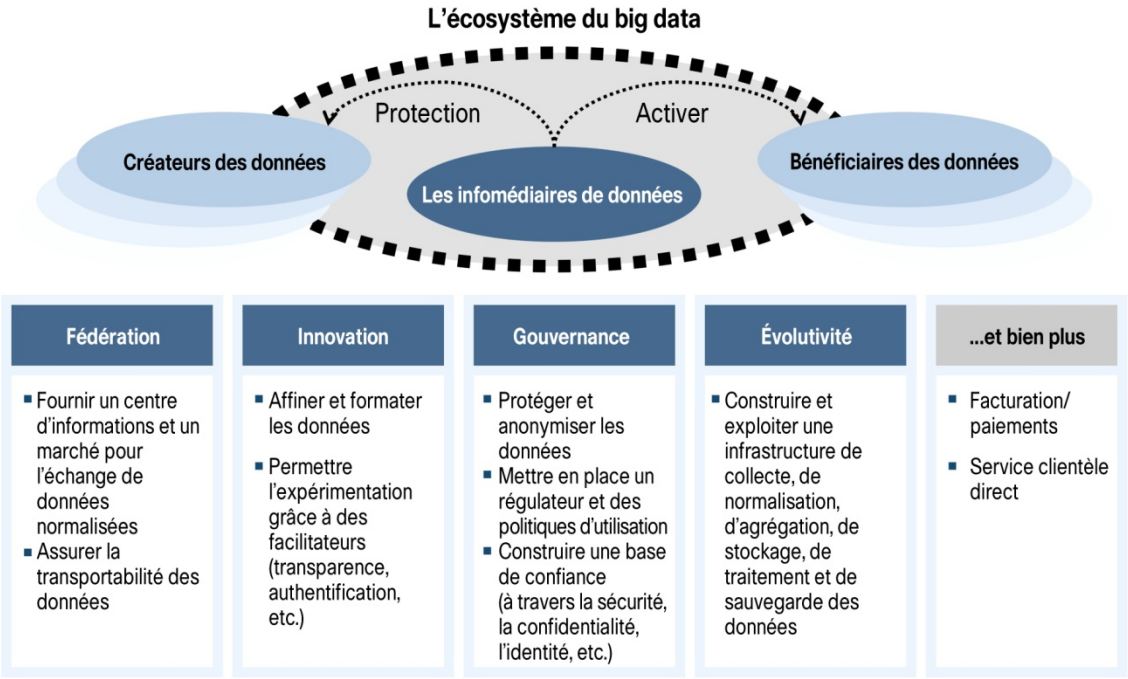
Intéressons-nous aux facteurs de succès d'autres écosystèmes complexes tels que la compensation des transactions financières, l'itinérance mobile, les échanges de données électroniques, les péages routiers automatisés et les dossiers médicaux électroniques. D'après nos recherches, les écosystèmes complexes ont besoin d'entités externes pour exécuter des fonctions à valeur ajoutée permettant la fédération, l'innovation, la gouvernance et l'évolutivité.

Dans certains écosystèmes, des intermédiaires sont apparus. Ils remplissent ces fonctions et contribuent à la prospérité de ces secteurs : TransUnion dans le domaine bancaire, Syniverse Technologies pour l'itinérance mobile et RosettaNet dans la gestion de la chaîne d'approvisionnement. En revanche, d'autres secteurs tels que les péages automatiques en Europe et les dossiers médicaux électroniques aux États-Unis rencontrent des difficultés. En l'occurrence, aucune entité n'a souhaité ou n'a été en mesure d'assumer le rôle d'infomédiaire, ce qui a freiné la progression.

Cisco IBSG pense que les intermédiaires ou infomédiaires seront indispensables dans l'écosystème d'analyse du big data pour que celui-ci réalise tout son potentiel et tienne toutes ses promesses. Nous pensons que les infomédiaires de données doivent jouer quatre rôles (voir la figure 4 et les descriptions ci-dessous) :

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Figure 4. Le rôle des infomédiaires de données.



Source : Cisco IBSG, 2012

1. **La fédération : viabiliser l'écosystème.** Les données doivent être transférées de l'endroit où elles ont été créées vers l'endroit où elles sont exploitées à l'aide d'un mécanisme qui limite les problèmes de confidentialité et de sécurité. Pourtant, il n'existe aujourd'hui aucun marché viable sur lequel partager différentes sources de données et les exploiter. Toutefois, les pionniers du secteur commencent à découvrir les avantages de cette diversité. La chaîne commerciale américaine Target, par exemple, a étudié de près les habitudes de consommation des acheteuses enceintes, car la fidélisation des clients à une nouvelle marque est forte du début de la grossesse à la période suivant l'accouchement. Comme l'indiquait le magazine *Forbes*, Target pourrait éventuellement partager cette base de données d'informations avec d'autres entreprises dans des circonstances appropriées et au sein du marché adéquat.[9] De même, différentes compagnies d'assurance complètent leurs propres données relatives aux demandes d'indemnisation par des bases de données tierces comprenant les médias sociaux et les ressources de l'ensemble du secteur afin de détecter les anomalies et d'identifier ainsi les nouveaux comportements frauduleux.

[9] « How Companies Learn Your Secrets », Charles Duhigg, *The New York Times,* 16 février 2012 ; « How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did », Kashmir Hill, *Forbes,* 16 février 2012.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Cependant, dans tous les cas de figure actuels, les entreprises doivent encore développer les moyens leur permettant d'utiliser elles-mêmes les données externes. Les infomédiaires de données seraient, quant à eux, capables d'identifier la valeur et de localiser la source de données, de régler les problèmes légaux/commerciaux, de filtrer les données et de normaliser le format de données pour le rendre compatible avec les systèmes internes. Les infomédiaires de données faciliteraient le partage de données dans l'ensemble de l'écosystème et créeraient un centre d'informations et un marché d'échange des données inter et intrasectorielles. Ils simplifieraient également largement la complexité à laquelle sont confrontées les entreprises lorsqu'elles étudient les bénéfices du big data, en particulier les petites entreprises qui ne font pas partie des rares sociétés investissant déjà massivement dans le big data. Pour ce faire, les infomédiaires de données pourraient créer des normes communes et identifier les acheteurs potentiels de données, ainsi que les possibilités d'exploitation de ces données qui s'offrent à eux.

2. **Innovation : valoriser un océan de données.** Pour l'utilisateur final, retirer de la valeur d'un tsunami de données reviendrait presque à « chercher une aiguille dans une botte de foin ». De nombreuses entreprises éprouvent des difficultés à expérimenter cette méthode sur de vastes quantités de données et parviennent encore moins à glaner de la valeur à partir de particules d'informations isolées. En assumant les rôles de gestion et d'authentification des identités, les infomédiaires de données facilitent cette expérimentation. Une fois authentifiées, ces données seraient transportables et prêtes à être partagées avec tous. Du point de vue du consommateur, les données stockées sur Facebook, par exemple, pourraient être transférées sur MySpace ou Amazon sans que l'utilisateur doive à chaque fois retaper ses informations personnelles.

3. **Gouvernance : alignement interne et externe des ressources.** L'écosystème actuellement fragmenté du big data doit être unifié. Les données transitent par différentes routes, à la fois internes et externes, et ce manque d'harmonisation et de gouvernance est un obstacle à la valorisation du big data. Pour y remédier, les infomédiaires de données aligneraient les ressources, de manière interne et externe, publique et privée, pour harmoniser le tout en une structure cloud commune. Cette approche protégerait les données contre le sabotage et garantirait aux organismes publics que celles-ci sont utilisées correctement. À en croire les récentes études approfondies sur Facebook, par exemple, les problèmes de confidentialité pèsent lourd dans la balance. Les infomédiaires pourraient contribuer à définir des normes applicables à l'ensemble du secteur et garantir que les données de Facebook ne soient pas perçues de manière négative, mais comme un bien précieux à utiliser conformément au souhait des utilisateurs. En gagnant la confiance des consommateurs et des acteurs du big data, les infomédiaires de données augmenteraient encore le nombre de données disponibles, ce qui permettrait de renforcer les actions en temps réel.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

4. **Évolutivité : faire face aux besoins informatiques.** La gestion du big data sera trop onéreuse pour la plupart des entreprises, qui devront externaliser le gros œuvre. En l'absence de capacités à grande échelle (en interne ou non), la gestion des données ne peut être fluide. Et même si le dispositif de stockage interne est capable de gérer plusieurs pétaoctets, les requêtes simultanées d'analyse de données peuvent tout de même mettre à mal le système. Les infomédiaires de données joueraient un rôle crucial dans la gestion d'énormes volumes de données. Ils proposeraient des méthodes de stockage plus évolutives et plus économiques, et aideraient les entreprises à procéder à l'incontournable traitement de leurs données. Les entreprises mondiales telles qu'Apple, Amazon et Facebook investissent déjà dans ces ressources sans intermédiaires et sont de plus en plus en mesure d'analyser les données en temps réel, au fur et à mesure qu'elles sont créées.

Différents acteurs du secteur peuvent assumer les rôles d'infomédiaires de données : par exemple, les prestataires de services, les entreprises informatiques, les banques, les grandes entreprises, les entreprises Web 2.0, les organismes gouvernementaux ou les nouvelles start-ups. La capacité de ces intervenants à faire tomber ensemble les obstacles décrits ici déterminera le succès à venir du big data.

L'aventure du big data : un appel aux décideurs

Les dirigeants d'entreprise commencent à se demander comment rejoindre l'aventure du big data et avec quel degré d'implication. Voici quelques options que les décideurs devront étudier :

- **Expérimenter lors de la première phase.** L'entreprise peut continuer à collecter des données de manière sélective pour des applications individuelles cloisonnées, telles que les applications de tarification, et en extraire de la valeur le cas échéant.
- **Passer à la deuxième phase.** À ce stade, une entreprise peut choisir de passer à l'échelon supérieur, de la réflexion tactique à l'exploitation stratégique des données. La stratégie peut rester cantonnée aux opérations internes, mais l'engagement et l'investissement permettent de différencier l'expérience des services grâce à une analyse des données et à une réaction proactive.
- **S'ouvrir aux autres.** À ce stade, une entreprise se rend compte que cette profusion d'informations traitées bien plus efficacement et plus intelligemment pourrait servir à d'autres. Ainsi, l'acteur d'un écosystème pourrait vendre ses données de géolocalisation à un autre intervenant et obtenir ainsi une nouvelle source de chiffre d'affaires. La chaîne américaine Wal-Mart, par exemple, pourrait choisir de vendre des données sur le commerce de détail à l'entreprise de télécommunications Virgin Media ; Amazon pourrait louer une partie de sa capacité de cloud computing excédentaire à des PME qui n'auraient pas d'autres moyens d'accéder à une telle infrastructure.
- **Devenir un infomédiaire de données.** Certaines entreprises peuvent souhaiter devenir cet acteur de confiance du secteur qu'est l'infomédiaire de données. À ce stade, elles peuvent contribuer à faire tomber les barrières du secteur, en reliant les domaines, en proposant de nouveaux services et en assurant la protection et la sécurité des données.

Cisco IBSG © 2012 Cisco et/ou ses filiales. Tous droits réservés.

Le chemin à parcourir : comment exploiter le big data ?

Au fur et à mesure que Cisco IBSG dessine l'évolution future du big data, il définit les compromis à envisager par les entreprises pour optimiser leurs bénéfices potentiels. Cisco IBSG pense en particulier que l'importance grandissante du big data s'accompagne d'opportunités importantes pour les infomédiaires potentiels. Et alors que le marché se dirige vers un point d'inflexion clé, en passant d'analyses centrées sur l'entreprise à un modèle fédéré/contextuel, il est temps d'agir.

Les entreprises peuvent toujours transformer leurs analyses de données, même si elles choisissent de maintenir une stratégie à orientation interne. Toutefois, la plupart d'entre elles retireront davantage de bénéfices d'une stratégie et de perspectives externes faisant appel au partage de données et d'informations dans l'ensemble de l'écosystème. D'autres entreprises, notamment les prestataires de services (dotés d'un accès de bout en bout au réseau), peuvent endosser le rôle d'infomédiaire de données, en répandant l'harmonie, la confiance, l'innovation et l'efficacité dans l'ensemble du big data hétéroclite.

Du point de vue de Cisco IBSG, la question n'est pas de savoir si la plupart des entreprises devraient investir dans l'avenir du big data, mais plutôt jusqu'où elles doivent aller. À l'image de ces magnats qui ont un jour évalué les coûts et les bénéfices potentiels de l'exploitation du pétrole, les décideurs actuels doivent optimiser les retombées de l'ère du big data qui s'annonce, indépendamment du niveau qu'ils choisissent.

Comme à la fin du 19e siècle, l'ampleur des retombées potentielles est telle qu'il est difficile de ne pas y penser.

Remerciements

Les auteurs souhaitent remercier Neeraj Arora, Marc Latouche, Thomas Renger, Rob Rowello et Evgenia Ryabchikova pour leur précieuse contribution à la rédaction de ce document.

Informations supplémentaires

Cisco IBSG (Internet Business Solutions Group) favorise la création de valeur marchande pour les clients en établissant un leadership éclairé qui façonne le secteur, des services de conseil pour la direction et la conception et l'incubation de solutions innovantes. En alliant stratégie, processus et technologie, Cisco IBSG agit en qualité de conseiller de confiance pour aider ses clients à prendre des décisions transformatrices pour créer de la valeur à partir de leurs idées.

Pour plus d'informations sur IBSG, visitez le site http://www.cisco.com/ibsg.

Siège social aux États-Unis
Cisco Systems, Inc.
San Jose, Californie

Siège social Asie-Pacifique
Cisco Systems (États-Unis) Pte, Ltd.
Singapour

Siège social en Europe
Cisco Systems International BV Amsterdam.
Pays-Bas

monde. Les adresses, numéros de téléphone et numéros de fax de nos bureaux sont indiqués sur le site Web Cisco, à l'adresse suivante : **www.cisco.com/go/offices**.

Cisco et le logo Cisco sont des marques de commerce ou déposées de Cisco et/ou de ses filiales aux États-Unis et dans d'autres pays. Pour voir la liste des marques commerciales Cisco, rendez-vous à l'adresse suivante : www.cisco.com/go/trademarks. Les marques de commerce de sociétés tierces mentionnées sont la propriété de leurs détenteurs respectifs. L'utilisation du terme « partenaire » n'implique nullement une relation de partenariat entre Cisco et toute autre entreprise. (1110R)