

Les données massives du *big data*, un nouvel **or noir**

PAR MICHEL CRESTIN



partner de LOBI

ET HERVÉ DHELIN

administrateur
du Club des directeurs
marketing des TIC

Comment traiter les masses considérables de données qui apparaissent aujourd'hui ? Un investissement massif et de nouvelles stratégies d'entreprise accompagnent création de valeur et rapidité de service. C'est une tendance lourde qui porte à considérer les données massives, parfois appelées *big data*, comme un nouvel « or noir » pour les entreprises.

■ La croissance des volumes d'informations numériques est telle que les bases de données qui les gèrent ne suffisent plus. Il faut donc imaginer de nouvelles architectures (logicielles et matérielles) pour traiter ces masses considérables de données.

Un investissement massif

L'évolution technologique se traduit par une diminution drastique des coûts de stockage des données. Par exemple, chacun peut acquérir une capacité de stockage d'un *téraoctet* pour seulement quelques centaines d'euros. Le coût d'un gigabit de stockage s'élevait à environ 20 dollars en 2005, à nettement moins d'un dollar aujourd'hui. À cet aspect purement matériel s'ajoute la diminution du coût de création, capture et duplication des données. Les investissements consentis par les entreprises dans l'univers numérique (*cloud computing*, équipement informatique, logiciels,

Ajuster ses prix en temps réel

Une entreprise américaine de grande distribution a constaté une baisse de ses parts de marché. Elle s'est aperçue que son principal concurrent avait lourdement investi dans des outils de collecte, d'intégration et d'analyses de données, de sorte qu'il était capable d'ajuster en temps réel ses prix, d'arbitrer l'approvisionnement entre ses différents magasins et de reconstituer ses stocks grâce à une étroite intégration de son système d'information avec ceux de ses fournisseurs.

services et personnel dédié à la création, gestion, stockage et exploitation des informations) ont augmenté de 50 % entre 2005 et 2011, pour atteindre les 4 000 milliards de dollars au niveau mondial. En 2012, 90 % des sociétés du classement Fortune 500 auront lancé un grand chantier de gestion des données. Les entreprises américaines de plus de mille salariés stockent, en moyenne, l'équivalent de 235 *terabytes* de données.

Une nouvelle stratégie d'entreprise

La valeur des données, et notamment la pertinence de leur exploitation, devient très importante. Une habile exploitation de ces gisements d'information contribue à accroître les avantages compétitifs. Au-delà de ces aspects économiques, la croissance des volumes de données remet en cause le management des entreprises et des organisations. Les dirigeants peuvent ainsi légitimement se poser les questions suivantes.

Sur le plan *managérial* : s'il était possible de tester toutes les décisions, cela changerait-il la stratégie de l'entreprise ? Probablement, ►

**L'évolution
technologique
se traduit par
une diminution
drastique
des coûts
de stockage
des données**

REPÈRES

En 2011, on a créé et répliqué 1,8 *zettaoctet* (c'est-à-dire 1 800 milliards de *gigaoctets*), neuf fois plus qu'en 2005, soit une croissance dépassant les prédictions de la loi de Moore. D'ici la prochaine décennie, la quantité de données gérée par les entreprises pourrait être multipliée par cinquante, et la taille des fichiers par soixante-quinze. Concrètement, un volume de 1,8 *zettaoctet* de données équivaut au volume de plus de 200 milliards de film HD (d'une durée moyenne de deux heures). Il faudrait à un individu pas moins de 47 millions d'années pour en visionner l'intégralité, si tant est qu'il s'y consacre vingt-quatre heures sur vingt-quatre et sept jours sur sept.

L'analyse prédictive

En se fondant sur un historique des informations disponibles sur les clients, l'analyse prédictive établit, avec une analyse statistique des relations entre les données disponibles, si elles sont de nature à prédire, avec la meilleure fiabilité possible, le futur comportement d'un individu. Retenons seulement deux domaines dans lesquels l'analyse prédictive se révèle particulièrement pertinente : la détection de la fraude et le marketing.

- car les managers pourraient mieux identifier les causes d'un phénomène et les conséquences de leurs décisions, en fonction de différents scénarios, au-delà de l'analyse de simples corrélations, et adapter leur stratégie en conséquence.

Sur le plan *opérationnel* : dans quelle mesure les offres d'une entreprise seraient-elles transformées si les énormes volumes de données disponibles étaient utilisés pour personnaliser en temps réel les solutions proposées aux consommateurs et prévoir leur comportement d'achat ? Par exemple, un e-commerçant pourrait traquer en temps réel le comportement de ses prospects et identifier le moment où ils vont s'engager dans l'acte d'achat. Et, ainsi, pousser des offres à plus fortes marges. La grande distribution constitue un secteur privilégié d'expérimentation dans la mesure où foisonnent les informations issues des achats sur le Web, des réseaux sociaux et, aujourd'hui, de plus en plus, les données de géolocalisation des smartphones.

Sur le plan *stratégique* : l'exploitation des gisements de données peut-elle donner lieu à de nouveaux modèles d'affaires ? Probablement pour des entreprises qui se positionneraient

Éviter l'effet domino

Une grande banque européenne a mis en place un système d'analyse prédictive pour détecter les risques de faillites parmi ses entreprises clientes, en particulier en intégrant des informations relatives au contexte économique. Concrètement, la banque est également en mesure d'anticiper les conséquences d'une défaillance, pour elle, d'un client majeur, de manière à éviter les conséquences néfastes d'un « effet domino » sur d'autres clients.

comme intermédiaires dans la chaîne de valeur pour gérer des données transactionnelles de manière exhaustive.

Une création de valeur

La profusion de données génère des opportunités nouvelles pour les entreprises, même s'il faut demeurer conscient des risques induits (intégrité des informations, menaces sur la vie privée, droit à l'oubli, piratage, atteinte à l'image, vol d'identité, fraude, etc.). L'une des révolutions concerne l'analyse prédictive, approche qui étudie les données et les caractéristiques comportementales des individus pour en tirer des modèles prédictifs en vue d'optimiser la relation avec les clients.

Mieux détecter les fraudes

Pour réussir sur un marché dominé par des grands noms de l'assurance, la société américaine Infinity, spécialisée dans l'assurance des conducteurs à risques élevés, avait besoin d'accroître son parc de clients et d'améliorer son efficacité opérationnelle, notamment pour la gestion des sinistres et la réduction de la fraude. Des techniques d'analyse prédictive ont été mises en œuvre. L'objectif était, sur le modèle des crédits bancaires, de « scorer » les sinistres de manière à mieux identifier la probabilité de fraude.

Récupérer ses fonds

Il s'agissait également d'optimiser la collecte dite de subrogation, c'est-à-dire lorsque l'assureur se retourne vers des tiers pour récupérer des fonds, quand son assuré n'est pas responsable. En un mois, la solution mise en œuvre par Infinity a permis d'augmenter les fonds collectés d'un million de dollars et de douze millions de dollars en six mois, uniquement par une meilleure analyse des données. La prochaine étape, elle aussi fondée sur l'analyse de volumes importants de données, consistera à étudier le contenu des documents liés aux règlements des sinistres, par exemple les comptes rendus d'accidents, les rapports médicaux ou les témoignages, autant d'éléments qui sont susceptibles d'améliorer la détection de la fraude le plus en amont possible.

Lutter contre la criminalité

L'analyse de grandes masses de données est également très utile pour lutter contre la criminalité. Par exemple, la police de New York

L'analyse prédictive se révèle particulièrement pertinente dans la détection des fraudes

dispose d'un logiciel d'analyse prédictive de la criminalité. Baptisée *Crush* (*Criminal reduction utilizing statistical history*), cette application regroupe une vingtaine d'années d'archives et de statistiques liées à la délinquance et à la criminalité, avec toutes les caractéristiques des délits : lieu, heure, mode opératoire, etc. L'analyse de ces données est assurée par un programme mathématique conçu par les chercheurs de l'université de Californie. Elle permet de localiser les endroits à risque de la ville, d'identifier les types de délits qui y sont commis, en fonction de critères tels que le moment de la journée ou le jour de la semaine. Concrètement, la police peut alors agir et envoyer des effectifs avant que le crime ou le délit ne soient commis.

L'exemple de Memphis

Déjà déployé au sein de la police de Memphis (Tennessee), le logiciel Crush a permis de réduire la criminalité de 30 % et la criminalité violente de 15 % par rapport à 2006.

En janvier 2010, la police a lancé plusieurs opérations dans un quartier de la ville, en se fondant sur les indications du logiciel Crush. Résultat : plus de cinquante arrestations de trafiquants de drogue et une réduction de près de 40 % de la criminalité.

Améliorer le marketing

L'une des difficultés pour les entreprises qui s'adressent à des millions de clients est de fidéliser ceux-ci et d'en conquérir d'autres. Dans les banques anglo-saxonnes par exemple, le taux de rotation des clients atteint facilement 10 % à 20 % par an, avec un coût par client s'élevant entre 200 et 3500 dollars. Une banque néerlandaise, grâce à l'analyse prédictive, a réussi à accroître l'efficacité de ses campagnes marketing avec un taux de réponse passé de 4 % à 12 %, d'où un retour sur investissement de ses actions amélioré de 10 % à 20 %.

Une tendance lourde

On pourrait multiplier les exemples de création d'avantages compétitifs grâce à une exploitation intelligente de volumes de données. Le phénomène du *big data* va impulser un élan supplémentaire. À moyen ou long terme, on peut discerner ce qui relève de certitudes et ce

qui constituera des incertitudes susceptibles de freiner, mais pas d'annihiler, la tendance lourde à considérer les données, surtout si elles sont massives, comme un nouvel « or noir » pour les entreprises et les organisations.

L'analyse des données

Du côté des certitudes, on retiendra une croissance continue des volumes de données, d'autant que se profilent l'Internet des objets et la civilisation des capteurs, producteurs d'énormes quantités d'informations ; une transformation profonde des modèles d'affaires et, de fait, des stratégies des entreprises ; des investissements massifs en outils d'analyse de données à mesure que les entreprises vont prendre conscience de la valeur de leurs données.

Le rejet des intrusions

Du côté des incertitudes, on retiendra les inquiétudes sur le respect de la vie privée, l'éventuelle montée en puissance de sentiments de rejet face à ce qui peut être considéré comme des intrusions de messages commerciaux dans la sphère privée. De même, nous n'avons pas de certitudes sur les problématiques de sécurité associées au *big data*. Enfin, un frein potentiel réside dans l'évolution de la législation, susceptible de devenir plus restrictive. Entre ces puissants facteurs accélérateurs et les éventuels freins, les entreprises devront trouver un équilibre de manière à concilier création de valeur et éthique. ■

La rapidité du service

Le loueur de véhicules Avis a amélioré le résultat en Europe de ses campagnes de marketing par courriel et divisé ses coûts par deux.

En analysant les données sur ses clients, Avis propose des offres personnalisées dans plus de 18 millions d'envois de courriels chaque année. Avec une meilleure connaissance de l'activité des clients grâce à l'historique des clics dans les courriels et des transactions, Avis personnalise tous les courriels. Par exemple, le loueur dispose d'une base de données de « clients privilégiés » à qui elle promet une prise en charge dans un délai de trois minutes (ils sont assurés de recevoir la clef de leur véhicule dans un délai de trois minutes à partir du moment où ils se présentent au guichet). La rapidité de service étant essentielle pour fidéliser les clients.

Des inquiétudes sur le respect de la vie privée et le droit à l'oubli