



La récolte des données sur les consommateurs permet aux entreprises de mieux connaître les goûts des clients, de vendre davantage et d'accroître leurs marges.

Les promesses très commerciales du « big data »

TECHNOLOGIE Les entreprises sont capables d'interpréter les informations abondantes récoltées sur le Web pour personnaliser leurs relations avec les clients. Mais attention aux dérives.

Maître de la donnée, maître du monde : les organisateurs du dernier salon *big data* à Paris n'y sont pas allés de main morte pour nommer leur conférence d'ouverture. « *La donnée, c'est le pétrole du XX^e siècle* », affirmait de manière non moins grandiloquente un investisseur américain, il y a quelques années. Avec l'explo-

sion d'Internet, des réseaux sociaux et désormais des objets connectés, les entreprises récoltent une quantité faramineuse d'informations sur les consommateurs, et celle-ci ne cesse de grandir. Des « gisements » très précieux dont elles comptent bien tirer profit. Selon le site IDG Enterprise, les trois quarts des multinationales utiliseront, d'ici à deux ans, des solutions *big data*, autrement

dit, des programmes capables de traiter et d'interpréter rapidement leurs immenses volumes d'informations. Ces technologies et les services associés représenteront en 2018 un marché mondial d'environ 42 milliards de dollars, selon le cabinet d'études IDC.

Pourquoi un tel engouement ? En exploitant correctement leurs montagnes de données (historiques de navigation et d'achat, formulaires clients, *e-mails*, publications sur les réseaux sociaux, etc.), les grands groupes ont pour objectif de mieux connaître les goûts des consomma-

teurs, de vendre davantage et d'accroître leurs marges. Auparavant, les entreprises ne connaissaient vraiment que leurs clients les plus fidèles – ceux et celles qui se rendaient régulièrement en boutiques, et acceptaient de prendre une carte de fidélité. Le Web a changé la donne : en analysant la navigation des internautes, on peut deviner leurs goûts et détecter leurs besoins. « *Tous les sites de e-commerce utilisent des algorithmes de recommandation en fonction des notes, des commentaires et des profils similaires* », rappelle Henri Isaac, chercheur à l'université Paris-Dauphine. Amazon, eBay et autres Pixmania personnalisent ainsi le contenu affiché en proposant les produits susceptibles de plaire à l'utilisateur.

Les smartphones ont par ailleurs rajouté une couche en permettant aux internautes d'intensifier leur usage d'Internet tout en géolocalisant leurs requêtes. Les projets *big data* peuvent s'avérer particulièrement rentables. L'an dernier, la division billetterie de la Fnac a utilisé ces technologies pour mieux cibler ses campagnes d'*e-mailing*. Résultat : le taux de clic* et le taux de conversion* ont doublé.

L'avenir radieux du *big data* re-



- **Taux de clic** : proportion d'*e-mails* qui ont été ouverts.
- **Taux de conversion** : rapport entre le nombre de mails envoyés et le nombre de visites que le site a reçus.
- **Machine learning (apprentissage automatique)** : forme d'intelligence artificielle grâce à laquelle l'ordinateur, au fur et à mesure que sa base de données grandit, dégage des tendances et des corrélations et en déduit des anticipations sur les comportements futurs.
- **Disponibilité (ou propension) maximale à payer** : le prix que le consommateur est prêt à payer pour un bien ou un service. La différence avec le prix effectivement payé constitue le « *surplus* » du consommateur.

ASSURANCE

Un pour tous, tous pour un ?

A terme, le *big data* pourrait faire disparaître un mécanisme fondamental de l'assurance : la mutualisation des risques.

Les compagnies d'assurances ont toujours été friandes de mégadonnées. Elles emploient des bataillons d'actuaires, ces professionnels des probabilités et des statistiques qui classent les individus par profil de risques afin d'ajuster les tarifs et de sécuriser les bénéfices. Un jeune homme ayant déjà eu un accident de voiture après un an de permis paiera ainsi plus cher son assurance auto qu'une quadragénaire sans antécédents en la matière.

Dans ce secteur comme dans bien d'autres, les *big data* ouvrent de nouvelles perspectives, en lien notamment avec les objets connectés qui font office de « mouchards ». Les assureurs proposent des applications smartphones qui analysent votre conduite et certains appliquent déjà des bonus-malus en fonction de celle-ci.

Demain, des bracelets connectés préviendront les mutuelles si vous restez affalé sur votre canapé plutôt que de faire de l'exercice comme le médecin vous l'avait prescrit. « *Une compagnie pourra augmenter les primes, voire refuser d'assurer un individu, à cause de son comportement* », prévient Jean-Pascal Gayant. Une perspective qui pose problème : cela remet en cause le principe de mutualisation qui fonde l'assurance moderne. Et en filigrane, une certaine idée de la solidarité au sein de notre société.

Les objets connectés peuvent faire office de mouchards pour les compagnies d'assurances, notamment dans le domaine de la santé.

Emilie Lereaux - Picturetank



pose en particulier sur ce qu'on appelle le *machine learning**, qui a fait d'énormes progrès ces dernières années. « *Les algorithmes s'améliorent au fur et à mesure qu'on les utilise. Plus vous déployez votre modèle à grande échelle, plus il devient performant* », explique Henri Isaac. D'où l'idée de mettre en commun ses informations avec d'autres entreprises (non concurrentes) afin de constituer d'immenses « lacs de données ». C'est la démarche entreprise par Orange avec BNP Paribas, La Poste et d'autres sociétés. L'opérateur téléphonique s'est d'ailleurs lancé très tôt dans le

big data. Pour chacun de ses clients, il calcule 500 notes qui évaluent, par exemple, la probabilité que ceux-ci détiennent un équipement concurrent ou qu'ils envisagent de changer de fournisseur.

Des prix différenciés

La puissance informatique est également mise à contribution pour modifier le prix des produits en temps réel. La *start-up* Boomerang Commerce propose aux e-commerçants d'adapter en permanence leurs tarifs à ceux d'Amazon. Le géant de Seattle, professionnel de l'exploitation des données, n'hésite pas à faire monter subitement ses prix lorsqu'il anticipe qu'un événement particulier va stimuler la demande. L'association Que Choisir a ainsi révélé qu'à l'approche de Noël, les prix de plusieurs jouets « *Star Wars* » avaient bondi de 30 % les jours où les films étaient diffusés à la télévision (voir graphique page 72).

A l'avenir, les distributeurs pourraient aller encore plus loin. En identifiant, à partir de votre adresse IP, les sites que vous consultez et en rachetant auprès de sites comme Google ou Facebook des informations sur vos intérêts du moment, ils pourraient en déduire votre degré d'addiction à un produit et faire monter les enchères en conséquence. « *Ils proposeront alors un tarif proche de votre disponibilité maximale à payer* », prédit Jean-Pascal Gayant, professeur d'économie à l'université du Mans. *Les big data permettront au final de pratiquer* ▶

17 %

des entreprises françaises sont matures dans l'exploitation de leurs données clients.

► une discrimination tarifaire au niveau de l'individu.» Avec à la clé, de substantiels profits. De quoi remettre en cause la croyance selon laquelle Internet rendrait le consommateur plus puissant face aux marques, voire ouvrir la porte à des abus¹¹.

Mais les *big data* ne sont pas l'apanage du Web. Les magasins physiques peuvent également exploiter avec profit les puissants outils d'analyse disponibles sur le marché. Grâce à eux, l'enseigne britannique Tesco sait par exemple qu'une hausse de température de dix degrés entraîne une augmentation de 300 % de ses ventes de viandes pour barbecue. Ce genre d'indicateurs lui permet d'optimiser le stockage de ses denrées et de limiter les ruptures de stocks et les invendus. La chaîne de supermarchés américaine Target va encore plus loin : en analysant les emplettes de chacun de ses clients,

elle devine les « événements de la vie » susceptibles de modifier leur comportement d'achat. Elle a ainsi envoyé à une jeune cliente près de Minneapolis des bons de réduction pour des produits dédiés aux jeunes mamans. Le père de celle-ci, qui n'était pas (encore) au courant que sa fille était enceinte, s'en est étonné auprès du directeur de magasin !

Enfin, les masses de données peuvent servir de matériau pour proposer de nouveaux services en vue d'améliorer l'expérience client. La société Parkeon, qui gère 200 000 horodateurs dans le monde, a par exemple lancé une application pour orienter ses usagers vers les places disponibles. La SNCF, de son côté, partage des informations comme les horaires théoriques et réels ou les comptages de voyageurs sur une plate-forme *open data*. C'est en puisant dedans que la start-

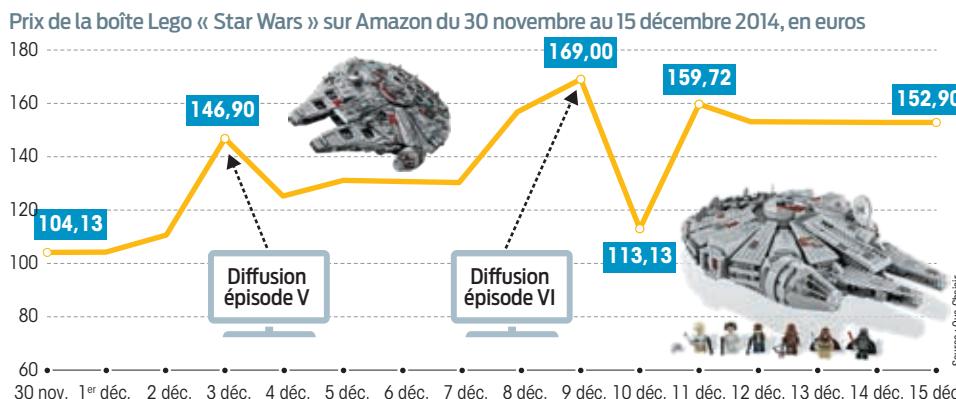
up Snips a conçu l'application bien pratique Tranquiliens, qui prédit le remplissage des RER franciliens.

Multicanal

Dans la ruée vers les données, les géants américains du Net possèdent, sans surprise, une longueur d'avance. « *Les moteurs de recommandation d'Amazon ont déjà une dizaine d'années ; ça montre le décalage avec les sociétés tricolores...* », constate Marc Batty, cofondateur de la start-up Dataiku, qui a mis sur le marché un logiciel d'analyse de données et de création d'applications prédictives. Seulement 17 % des entreprises françaises sont matures dans l'exploitation de leurs données clients, indique une étude récente du cabinet Ernst & Young. « *Les patrons recherchent des profils surspécialisés de type data scientists. Mais avant tout, il est essentiel de partager les données en interne et de favoriser la collaboration entre services : l'informatique, le marketing, le Web et les ventes* », juge Marc Batty. Une démarche indispensable pour mettre en place une stratégie « multicanal », qui croise les informations provenant du digital et celles issues des canaux traditionnels afin d'adresser un message cohérent au client.

De son côté, Google étend déjà le *big data* aux domaines de la domotique (avec le rachat de la société Nest, qui

DES PRIX STIMULÉS PAR LA DIFFUSION TÉLÉ



AUTRES APPLICATIONS

Doper les ventes, mais pas seulement...

Outre la connaissance-client et les ventes, le *big data* peut s'appliquer à bien d'autres domaines de l'entreprise.

Logistique : les industriels bardent de capteurs leurs chaînes de production. L'analyse des données obtenues permet d'anticiper les pannes de machines et de déterminer des corrélations entre la qualité des produits et les conditions de température et de pression (dans le secteur pharmaceutique, par exemple).

Recrutement : pour dénicher la perle rare et éviter les erreurs de casting, de grands

groupes utilisent des algorithmes qui trient les milliers de CV qu'ils reçoivent et analysent les informations diffusées (parfois imprudemment !) par les candidats sur les réseaux sociaux comme LinkedIn, Viadeo, Twitter ou Facebook.

Conception : les fabricants d'éoliennes comme Vestas analysent les paramètres météorologiques sur plusieurs années (voire décennies) et en déduisent la position optimale de leurs installations afin de maximiser le rendement en fonction du vent.

Recherche médicale : la plate-forme

Flatiron Health recueille des données sur les cancers provenant d'hôpitaux et de centres de recherche du monde entier. L'agrégation et le traitement de cette masse d'informations permettent d'améliorer le diagnostic des patients et la prise de décision des médecins.

Agriculture : grâce à des données portant sur des millions d'hectares de terres cultivables aux Etats-Unis, Farmers Business Network évalue la performance des graines et recommande des méthodes agricoles en fonction du lieu et du type de culture.

Data center de Google.
Dans la ruée vers les données, les géants américains du Net possèdent, sans surprise, une longueur d'avance.

conçoit des objets intelligents pour la maison) et de la santé : la firme américaine vient ainsi de signer un partenariat avec le laboratoire Sanofi, auquel il apportera ses compétences en matière de collecte et d'analyse de données pour mettre au point des traitements plus efficaces contre le diabète. La numérisation croissante de l'économie fait ainsi peser le risque de voir une poignée de géants, comme Google, profiter de leur hégémonie pour siphonner la valeur ajoutée de nombreux secteurs.

Big Brother, version 2.0

Derrière ses promesses commerciales, le *big data* soulève cependant aussi de sérieuses questions éthiques. L'institut G9+ souligne les risques potentiels d'une « *dictature de la prévision* » qui ne ferait pas grand cas de notre libre arbitre et enfermerait les individus dans des cases. Prenons le cas (pas si futile qu'il n'y paraît) de la musique en ligne : Deezer, Spotify et iTunes utilisent des algorithmes de recommandation qui vous proposent des morceaux correspondant à vos écoutes précédentes. Du jazz, par exemple. Mais *quid* de la diversité culturelle ? « *Pour contrebalancer la toute-puissance des statistiques, certaines start-up ont mis au point des programmes qui réintroduisent une dose de hasard !* », s'étonne Henri Isaac.

L'autre menace, encore plus sérieuse, concerne évidemment la protection de la vie privée des individus. « *La plupart des citoyens adhèrent avec candeur aux innovations technologiques. Ils racontent leur vie sur Facebook, se laissent géolocaliser par leur smartphone, sans même savoir qui va manipuler ces données* »,

s'inquiète Jean-Pascal Gayant. Big Brother n'est pas loin... Des sociétés ont certes mis en place des chartes *big data* et nommé des *chief privacy officers*, mais les bonnes intentions ne suffiront pas : l'intervention des pouvoirs publics est nécessaire.

Le projet de loi « pour une répu-

“ Les *big data* permettront au final de pratiquer une discrimination tarifaire au niveau de l'individu ”

JEAN-PASCAL GAYANT

Professeur d'économie à l'université du Mans

blique numérique », qui devait être mis en ligne fin septembre par le gouvernement, comporte un volet protection de la vie privée en ligne. Celui-ci affirme notamment les principes d'*« autodétermination informationnelle »*, de *« portabilité des données »* et de *« droit à l'oubli »*, qui visent à rendre aux internautes la maîtrise de leurs données. Outre l'inscription dans le droit de la neutralité du Net, il devait aussi retenir un principe dit « de loyauté » des plates-formes, qui consiste à imposer aux géants comme Google,

Apple ou Amazon une obligation générale de transparence, d'information et de non-discrimination vis-à-vis de leurs utilisateurs. Obligation, dont pourraient s'assurer, ainsi que le propose le Conseil national du numérique (CNNum), des agences de notation spécialisées à créer.

A l'échelon européen, cela fait plusieurs années que les 28 pays de l'Union discutent d'un projet de réforme de la protection des données pour remplacer la directive de 1995, jugée obsolète. Le projet prévoit des amendes pouvant atteindre 2 % du chiffre d'affaires pour les entreprises qui contreviendraient aux règles de manipulation des données.

Le Parlement européen table sur la conclusion d'un accord sur ce texte d'ici à la fin de l'année. Affaire à suivre.

■ Thomas Lestavel

[1] L'Autorité de la concurrence, que nous avons interrogée sur la légalité de ces pratiques, nous a répondu que ces questions « dépassent [son] champ de compétences ».

en savoir plus <

> « **Big data, l'accélérateur d'innovation** », institut G9+ avec Renaissance numérique, décembre 2014, accessible sur <http://cr.g9plus.org/2014-12-16-G9plus-LB-Big-Data.pdf>

> « **(Big) data, où en sont les entreprises françaises ?** », rapport du cabinet EY, 2014, accessible sur [http://www.ey.com/Publication/vwLUAssets/EY-etude-big-data-2014/\\$FILE/EY-etude-big-data-2014.pdf](http://www.ey.com/Publication/vwLUAssets/EY-etude-big-data-2014/$FILE/EY-etude-big-data-2014.pdf)

> **La Révolution Big data. Les données au cœur de la transformation de l'entreprise**, par Jean-Charles Cointot et Yves Eychenne, Dunod, 2014.

