

**Facebook regarde dans quels  
magasins vous faites vos  
courses**

	<b>Facebook regarde dans quels magasins vous faites vos courses</b>
---	---

---

**Facebook va désormais traquer les données de ses utilisateurs pour savoir dans quels magasins ils se rendent. Le but est de permettre aux annonceurs de savoir si leurs publicités attirent des consommateurs sur leurs points de vente.**



Facebook ne cesse de renforcer son service de publicités. Le réseau social veut proposer une offre plus précise et pertinente pour ses clients. Pour cela, il se servira désormais des données de localisation de ses utilisateurs pour savoir dans quels magasins ils se rendent. Le but ? Permettre aux entreprises de savoir si leurs annonces sur Facebook attirent du monde dans leurs magasins.

Ainsi, les annonceurs pourront comparer le nombre de personnes qui ont vu leurs annonces au taux de fréquentations de leurs points de vente. Ils peuvent également intégrer une carte interactive à leur publicité – sous la forme d’un carrousel – pour indiquer à l’internaute le chemin qui le mènera au magasin le plus proche.

Ces nouvelles fonctionnalités s’inscrivent dans une volonté de Facebook de proposer des services plus personnalisés – et donc plus efficaces – à ses clients. En 2014, la boîte de Mark Zuckerberg avait déjà lancé une plateforme qui permet d’afficher de la publicité aux utilisateurs du réseau social qui se trouvent à proximité du magasin afin de les inciter à s’y rendre rapidement.

Selon Facebook, plusieurs entreprises ont déjà eu l’occasion de tester, en avant-première, ces nouvelles fonctionnalités. Parmi eux, se trouve E.Leclerc. La chaîne de distribution française « a pu atteindre 1,5 millions de personnes dans un rayon de dix kilomètres autour de ses supermarché et a observé qu’environ 12 % des clics sur leur publicité ont entraîné une visite en magasin dans les sept jours qui suivaient », indique Facebook dans son annonce.

Grâce à ces jeux de données très précis, Facebook fournit des outils pertinents pour les entreprises car, grâce à cela, elles peuvent ajuster leur stratégie de communication en fonction de chaque point de vente et de chaque région. Le réseau social prouve encore plus à quel point il représente un atout bien plus puissant que les modes de diffusion traditionnels.

Quant aux utilisateurs de Facebook, si cette information a de quoi énerver, elle n’a rien de vraiment surprenant. Il est de notoriété publique que la publicité ciblée représente le fonds de commerce principal du réseau social. Celui-ci n’est d’ailleurs pas le seul à traquer les internautes pour savoir dans quels magasins ils vont. Google le fait depuis quelques temps déjà, comme le rappelle, dans un tweet, Jason Spero, responsable de la stratégie et des ventes mobiles chez la firme de Mountain View.

Google dispose de données encore plus importantes destinées aux annonceurs et adapte les publicités en fonction, entre autres, des recherches de l’utilisateur et de sa géolocalisation.

Article original de Omar Belkaab



Réagissez à cet article

**Original de l’article mis en page : Facebook regarde dans quels magasins vous faites vos courses – Business – Numerama**

**Pourquoi vous ne devriez  
jamais publier de photos de  
vos jeunes enfants sur  
Facebook**

	<p><b>Pourquoi vous ne devriez jamais publier de photos de vos jeunes enfants sur Facebook</b></p>
---	--

---

**Atlantico : Poster une photo de son enfant sur Facebook peut-il lui porter préjudice ? Si oui, quand ? Et pourquoi ?**



Publier une photo de ses enfants sur Facebook – qui est de loin le leader des réseaux sociaux dans le monde – est un acte compréhensible mais qui fait surtout plaisir sur le moment aux parents. Les parents façonnent l'identité numérique de leurs enfants à l'insu de leur plein gré alors même que le droit à l'oubli n'existe pas sur Internet. Plus tard, certaines traces numériques (photos ou vidéos postées avec les commentaires et tags associés) peuvent être utilisées contre eux surtout si les paramètres de confidentialités sont mal utilisés.

Et même en postant une photo accessible aux seuls amis, celle-ci peut ensuite être partagée plus largement. En outre les personnes qui vont réagir à la photo permettent de révéler l'écosystème relationnel de la personne. Il est facile d'établir des corrélations entre les personnes. Et en fonction du profil des personnes réagissant de déterminer quel est le profil potentiel de l'enfant sur la photo. Pour les préjudices, on pense avant tout à l'attitude d'un recruteur mais ce peut être aussi des amis potentiels de l'enfant qui le jugeront avec un autre regard. Déjà on google une personne avant de la rencontrer ce qui induit un prisme dans la première rencontre. Le préjudice peut intervenir à des périodes charnières de la vie : adolescence où l'individu se construit et est sensible au regard des autres, entrée dans la vie active, rencontre amoureuse, etc.

**Comment fonctionne le système de tag ? Quelle est sa fonction ? Pourquoi l'utilise-t-on ?**

Il s'agit d'un système mis en place par Facebook qui permet à un utilisateur de Facebook d'indiquer qu'une personne figure sur une photo. En quelque sorte, un traitement manuel du facebooknaute lui-même vient en complément de l'algorithme mis en place par Facebook pour collecter des données personnelles (en l'occurrence les photos des visages des personnes) de nature à faire grandir la base d'information relative à une personne. Facebook peut avec l'expérience lui-même déterminer les personnes reconnues sur les photos, ce qui est parfois bluffant. Facebook peut ensuite, en fonction des références à d'autres posts, déterminer le cercle probable de personnes autour de celle qui a été taguée. Ceci lui permet de faire des suggestions (par exemple amis que l'on pourrait connaître, voire produits ou services que l'on est susceptible d'aimer car les goûts de ses amis sont souvent plus proches des siens que ceux d'inconnus) avec des taux de retour plus pertinents.

L'objectif de Facebook est d'exploiter le *big data* constitué par les photos et leurs tags pour sans cesse améliorer les résultats pour les marques partenaires et qui paient ses services. Par ailleurs, les algorithmes qui permettent de reconnaître les visages et les techniques de bio-identification ne sont qu'à leur début. Demain, à partir d'une simple photo, il sera, avec des outils idoines, possible de dresser le portrait robot d'une personne en allant fouiller sur l'ensemble de la webosphère (pas seulement sur Facebook mais sur l'ensemble des réseaux sociaux et des sites) pour collecter les numéros de téléphone, les adresses mails et d'autres détails personnels associés. Ceci peut présenter des opportunités réelles pour mieux connaître rapidement une personne, mais présente des risques. Des garde-fous et une éthique sont à construire pour éviter que le numérique ne soit un facteur d'exclusion ou un moyen d'ostraciser les internautes. Alors que les États-Unis sont dans le mécanisme d'*opt-out* (utilisation *a priori* des données personnelles sans autorisation préalable), l'Europe préfère l'*opt-in* qui constitue un principe de précaution quant à l'exploitation des données personnelles. Mais force est de constater que les outils majoritairement utilisés en Europe sont Américains et que nous sommes GAFA-dépendant (*Ndlr* : GAFA = Google, Apple, Facebook, Amazon) et qu'en contrepartie de la gratuité d'utilisation d'un service, nous fournissons et souvent avec beaucoup de zèle des données personnelles que ces outils utilisent à la fois avec un traitement automatique et un traitement humain qui le perfectionne comme celui des tags.

Article original de David Fayon Lire la suite...



Réagissez à cet article

Original de l'article mis en page : Pourquoi vous ne devriez jamais publier de photos de vos jeunes enfants sur Facebook | Atlantico.fr

---

# Mieux connaître le consommateur avec ses données

Denis JACOPINI



vous informe



Mieux  
connaître le  
consommateur  
avec  
l'analyse  
prédictive  
et le Big  
Data

Grâce aux nouvelles technologies et particulièrement aux réseaux sociaux, il est désormais possible d'étudier tout ce que font vos clients.

Grâce aux nouvelles technologies et particulièrement aux réseaux sociaux, il est désormais possible d'étudier tout ce que font vos clients. Habitudes d'achat, fréquence et lieux des visites, horaires... Toutes ces informations forment une base de données gigantesque et sans cesse en mouvement. C'est ce que l'on nomme « Big Data » et il s'agit d'une véritable mine d'or pour les professionnels du marketing. Fini les suppositions logiques et autres préjugés, l'analyse prédictive permet maintenant de dégager des statistiques et schémas de consommation concrets.

**N'où viennent les informations qui composent le Big Data ?**  
 Chaque fois que vous activez votre géolocalisation en consultant un site internet ou une application, cela laisse une trace. Les données du Big Data sont également composées par vos habitudes de navigation sur le net, les endroits où vous vous rendez, combien de temps vous restez, ce que vous regardez. Bien sûr toutes ces informations sont rendues anonymes, mais vos terminaux, dont votre smartphone, sont de véritables capteurs dans votre poche. Un data scientist, tel que sont nommés les experts du Big Data, s'intéresse aux patterns et trie les données avec celles de milliers d'autres personnes. Il s'agit par exemple de créer des algorithmes adaptés aux habitudes de navigation des utilisateurs d'un moteur de recherche. L'idée est d'aller chercher dans les données des tendances, et d'identifier des comportements. Analyser, comprendre, puis prédire les actions futures. Cela est désormais possible et relativement simple avec les outils dont disposent les analystes.








**Le Big Data, un outil d'analyse prédictive qu'il faut savoir exploiter**  
 Si le Big Data peut servir à améliorer l'expérience des utilisateurs d'un produit, il révèle surtout son potentiel dans le secteur du marketing. Grâce à l'analyse de flux de données, il est possible d'établir des segments toujours plus pertinents. Finalement la publicité « à destination de la ménagère de 40 ans ». Vous êtes désormais en mesure de savoir qui est réellement susceptible d'utiliser vos produits, et avec quel argument mettre en avant votre offre. Bien sûr, cela demande un réel travail d'analyse et ce n'est pas un hasard si vous voyez fleurir les offres d'emploi de data scientist ou de data mining. Le marketing et l'analyse prédictive deviennent des travaux de statisticien. Cela demande également de disposer des bons outils. Il s'agit d'un investissement en plusieurs étapes :

1. Vous collectez les données transmises par toutes les sources pertinentes ;
2. Vous analysez les données et isolez les schémas de consommation qui vous intéressent. L'étude de leurs occurrences sera la base de vos analyses prédictives ;
3. Enfin, vous établissez une stratégie de marketing ciblée en fonction des résultats obtenus.

Pour une efficacité maximale, la majeure partie de ce processus sera automatisée. Pour gagner en efficacité mais aussi en efficience grâce à des outils de traitement des données en temps réel, il est possible de créer des processus semi-automatisés. L'intervention humaine n'est plus utile ? C'est le contraire. Elle est essentielle. L'œil humain est là pour aller chercher dans les données, fouiner et faire émerger des signaux faibles. La technologie libère le potentiel des données, mais il faut une intervention humaine pour bien utiliser ces outils, et en tirer des décisions actionnables.

**Comment se servir de l'analyse prédictive pour optimiser son ROI ?**  
 S'il peut être intéressant d'analyser le Big Data pour de multiples raisons, en matière de marketing l'objectif est avant tout d'améliorer votre ROI (Return On Investment). Pour cela, votre démarche analytique doit s'inscrire dans un plan d'action concret. Que vous soyez spécialisé dans le e-commerce ou que vous réalisiez toutes vos ventes dans des magasins physiques, utilisez les données pour améliorer votre marketing digital.

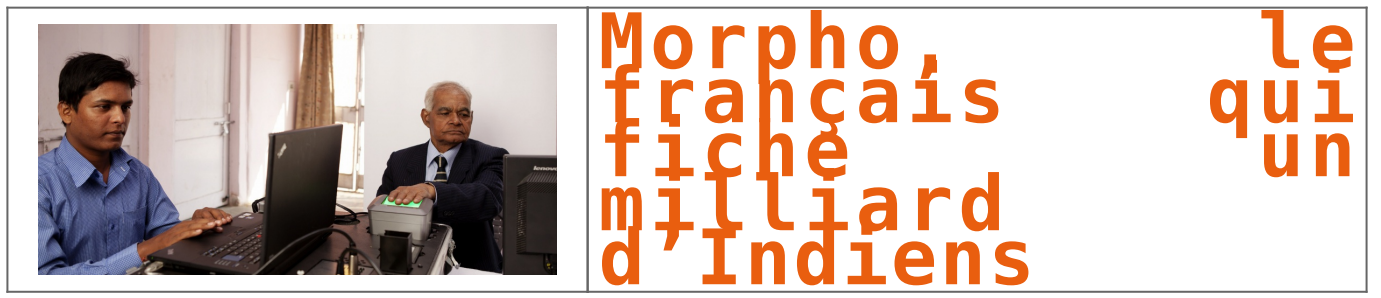
**Lancez des campagnes de marketing ciblées :**  
 Démarez-vous de flux de publicités, et adaptez votre proposition aux envies réellement exprimées de vos clients. Mais l'analyse prédictive ne sert pas qu'à générer des ventes. Elle trouve aussi son utilité dans le maintien de la relation client. Il est par exemple possible de déterminer quand un client est sur le point de résilier un abonnement, quand celui-ci est sur le point de basculer chez un concurrent, pour pouvoir le retenir ! À l'aide de ces informations contenues dans votre Big Data, vous pouvez améliorer votre taux de fidélité en adaptant vos offres au bon moment. Un exemple ? La chaîne d'hôtel Hyatt utilise désormais l'analyse prédictive pour donner à son personnel d'accueil des informations supplémentaires sur les clients. En analysant la recherche menée par ces derniers sur le site et les applications du groupe, Hyatt précise si un client peut être intéressé par une chambre avec vue (car il a regardé plusieurs fois la page) ou s'il désire peut-être une chambre avec des oreillers allergiques, car il a tapé ce mot-clé dans le moteur de recherche interne. Un bel exemple de personnalisation de la relation client, grâce aux données. [Lire la suite](#)

Share nous sur       

Magistrez à cet article

Source : *Analyse prédictive et Big Data : mieux connaître le consommateur avec ses données*

# Morpho, le français qui fiche un milliard d'Indiens – Challenges.fr



La filiale de Safran est en train de fournir une identité numérique à 1,2 milliard d'Indiens. Une base de données biométrique unique au monde, qui effraie certains.



Une base de données biométrique rassemblant 1,3 milliard d'individus, soit 18% de la population mondiale... C'est le défi incroyable que le français Morpho, filiale de Safran, est en train de relever en Inde.

Concrètement, le programme, baptisé Aadhaar (socle, en hindi), consiste à offrir un numéro d'identification unique à 12 chiffres à chaque citoyen. Cette identité numérique est sécurisée par la prise des données biométriques de son propriétaire: les 10 empreintes digitales, les 2 iris, et une photo du visage. Quatre ans après le début de l'opération, la base de données vient d'atteindre la barre symbolique du milliard d'individus fichés. « Chaque jour, jusqu'à 1 million de personnes peuvent être « enrôlées » dans le système », souligne Jessica Westerouen van Meeteren, directrice de la division Government Identity chez Morpho.

Pourquoi cette base de données géante? L'idée de départ du programme, lancé en 2009 par New Delhi, était d'offrir une existence officielle à des centaines de millions d'Indiens qui, faute de carte d'identité, restaient invisibles à l'administration, et donc exclus des programmes d'aide sociale. Dans un pays à l'administration pléthorique où la corruption reste importante, l'argent attristait souvent dans les mauvaises poches. Le numéro d'identification doit permettre de corriger le problème des fraudes à l'identité, mais aussi d'ouvrir un compte en banque simplifié ou d'obtenir un passeport plus facilement.

### La complexité d'un programme spatial

Pour mener à bien ce projet colossal, le gouvernement indien a créé une agence d'Etat, la Unique Identification Authority of India (UIDAI).

Morpho est l'un des fournisseurs retenus par l'agence, avec le japonais NEC et l'américain L1 (autre filiale de Safran). Le groupe français fournit les scanners biométriques destinés à l'enregistrement des données, mais aussi la technologie de « dédoublement » qui permet de vérifier qu'un individu n'est pas déjà enregistré sous un autre numéro. Le système est capable de répondre à un million de requêtes par jour. « C'est un programme d'une complexité inédite dans le secteur, qu'on peut comparer à celle d'un programme spatial », assure Jean-Pierre Pellestor, directeur de programme chez Morpho.

Si le projet est en train d'arriver à bon port, c'est en grande partie grâce à l'action d'un homme: Nandan Nikelani, le cofondateur du géant de l'informatique indien Infosys. Le puissant homme d'affaires, qui fut le premier président de l'UIDAI, a pesé de tout son poids pour passer outre les légendaires pesanteurs de l'administration indienne. Au point que la loi avalisant le programme n'a été votée à la Lok Sabha, la chambre basse du parlement indien, que le 16 mars dernier... soit six ans après le début des opérations d'enregistrement. Nikelani avait même réussi à convaincre le premier ministre Narendra Modi, très critique contre Aadhaar durant la campagne électorale de 2014, de poursuivre le projet. « Modi l'a finalement accéléré », se félicite-t-on chez Morpho.

### Risque de Big Brother?

Le programme ne fait pourtant toujours pas l'unanimité en Inde. Si plus d'un milliard de personnes ont accepté de s'enregistrer dans la base de données, d'aucuns y voient un Big Brother potentiel, qui pourrait être détourné au détriment de la vie privée des citoyens. « Le gouvernement peut-il nous assurer que Aadhaar et les données collectées ne vont pas être détournées comme ce qui a été fait par la NSA aux Etats-Unis? », s'interrogeait auprès de Reuters Tathagata Satpathy, une avocate basée dans l'Odisha (est de l'Inde). L'accès au fichier pour un usage lié à la « sécurité nationale » fait notamment débat. « Le projet apporte une protection de la vie privée d'une grande robustesse, au-delà de tout ce qu'ont apporté les autres lois en Inde », répondait mi-mars Nandan Nikelani à l'Indian Express.

En tout cas, Morpho espère bien surfer sur le contrat indien pour vendre d'autres systèmes similaires. « Nous avons des campagnes commerciales en cours dans d'autres pays sur des programmes comparables, mais la taille du projet indien restera probablement unique », détaille Jessica Westerouen van Meeteren. Mais la bonne santé de Morpho (1,9 milliard d'euros de chiffre d'affaires en 2015, en croissance organique de 11%) n'empêche pas le directeur général de Safran Philippe Petitcolin de réfléchir à son avenir, la division n'ayant pas vraiment de synergie avec le reste du groupe, ni le poids suffisant pour équilibrer les activités aéronautiques. Après avoir mis en vente l'activité de détection d'explosifs (Morpho Detection), le groupe pourrait annoncer la cession de toute la division dans le courant de l'année 2016... [Lire la suite]

Denis JACOPINI est Expert Judiciaire en Informatique spécialisé en « Sécurité » « Cybercriminalité » et en RGPD (Protection des Données à Caractère Personnel).



- Mises en conformité RGPD ;
- Accompagnement à la mise en place de DPO ;
- Formations (et sensibilisations) à la cybercriminalité (Autorisation n°93 84 03041 84) ;
- Audits Sécurité (ISO 27005) ;
- Expertises techniques et judiciaires ;
- Recherche de preuves téléphones, disques durs, e-mails, contentieux, détournements de clientèle... ;
- Expertises de systèmes de vote électronique ;



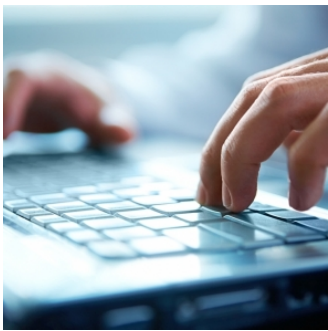
[Contactez-nous](#)

Réagissez à cet article

Source : *Morpho, le français qui fiche un milliard d'Indiens – Challenges.fr*

---

# Big Data : gare à vos données personnelles !



**Big Data : gare à vos données personnelles !**

---



Les marques, enseignes et sites Web cherchent à capter un maximum d'informations sur leurs clients. Leur objectif ? Vous vendre plus en vous soumettant des offres et promotions personnalisées. À la clé, des bons plans ou du harcèlement ? Débat.



Renseigner votre mail pour participer à un jeu concours, indiquer votre numéro de portable pour ne rater aucune vente privée, ou encore remplir un formulaire pour intégrer un programme de fidélité : il ne se passe pas une semaine où un magasin ou une marque ne vous sollicitent pour obtenir des informations personnelles sur vous, votre famille et vos habitudes de consommation. Et sur le Web, même combat. Les sites Internet ont même un avantage puisqu'ils peuvent obtenir des informations via les fameux cookies et ainsi savoir quel site vous avez visité ou encore quels sont les produits que vous avez regardés sur la Toile.

Peut-être avez-vous déjà remarqué que le fameux ordinateur ou la paire de bottes que vous convoitez se retrouve dans un carré de pub les heures ou les jours suivants votre session de surf ?

Globalement, les Français communiquent facilement et de façon importante des informations sur eux. C'est ce que révèle l'enquête 2015 « Les Français et leurs données personnelles » réalisée par Ipsos pour Elia Consulting. Plus exactement, les résultats démontrent qu'il y a un décalage entre la méfiance grandissante des utilisateurs sans pour autant qu'ils changent leur comportement. Ainsi, paradoxalement, les Français renseignent fréquemment et en quantité leurs données personnelles, malgré leur opposition de principe à leur utilisation par les entreprises et la conscience des risques qui y sont associés. 74% déclarent partager régulièrement des données personnelles, ce qui s'explique souvent par la contrainte de fournir ces informations pour terminer un acte d'achat ou bénéficier d'un service (73% des Français renseignent leurs données personnelles pour terminer un acte d'achat). D'ailleurs, seul un internaute sur trois prend le temps de lire les conditions générales de vente ou de modifier les paramètres de sécurité de leurs réseaux sociaux et smartphones (33% dans les deux cas).

### **Des données pour quoi faire ?**

Les Big Data, ou mégadonnées en français, désignent l'ensemble des informations que l'on peut capter et représentent le nouvel or noir des professionnels. En effet, l'objectif pour ces derniers consiste à bien cerner un client et lui glisser ainsi la bonne offre ou le bon service, au bon moment, et avec le bon prix. Amazon teste même un service aux États-Unis où le client reçoit un produit sans l'avoir commandé, en fonction de son historique d'achats. Libre à lui de le garder – dans ce cas, il sera débité dans les jours suivants – ou de le renvoyer si le produit ne lui convient pas.

En France, nous n'en sommes pas encore là. L'usage des mégadonnées reste encore peu développé, faute de moyens techniques et financiers à allouer à ce sujet. L'idée serait plutôt, dans le secteur du commerce, d'utiliser la « smart data », autrement dit celle qui fera la différence pour déclencher une intention d'achat chez le consommateur. Pour illustrer l'usage des mégadonnées, prenons un cas pratique avec Monoprix par exemple. L'enseigne envoie à tous les porteurs de la carte de fidélité des bons de réduction. Or, ces fameux coupons nominatifs ont été définis en fonction des informations fournies initialement (genre, âge, situation matrimoniale, etc.) et des derniers achats effectués par le consommateur. Ainsi, un homme n'aura, en théorie, pas de réduction pour du maquillage. Et s'il consomme surtout des plats cuisinés, ces derniers se retrouveront dans ses bons de réductions.

### **Manque de transparence ?**

Toujours selon l'enquête Ipsos, les Français sont parfaitement conscients que leurs données personnelles peuvent être utilisées. D'ailleurs, 9 sondés sur 10 (92%) pensent que les informations qu'ils renseignent peuvent être utilisées ou conservées pour un usage futur par le fournisseur de services. Mais dans la majorité des cas, les Français se sentent mal informés sur l'utilisation qui en est faite.

L'utilisation des données par les professionnels reste en réalité un jeu à double tranchant. En effet, si votre magasin d'électronique préféré vous envoie des promotions qui ne vous intéressent absolument pas, il y a fort à parier que vous allez vite couper toute relation avec lui. Pas forcément ne plus acheter chez lui, mais il ne pourra plus vous contacter pour vous inciter à venir en point de vente ou à vous rendre sur son site Web. Par ailleurs, même si les professionnels ont en leur possession de nombreuses données sur vous, ils font attention à ne pas devenir trop intrusifs. Imaginons que vous êtes en train de chercher un nouveau parfum et que la vendeuse vous lance : « Vous avez acheté depuis deux ans uniquement des fragrances sucrées, celui-ci est très différent ». D'un côté, vous bénéficieriez d'un conseil super personnalisé mais, d'un autre côté, vous réaliseriez que la dame en face de vous que vous ne connaissez pas sait beaucoup de choses sur vous...

### **Vigilance de mise**

Si dans les faits, il devient presque impossible de ne fournir aucune donnée personnelle, il convient, néanmoins, de réfléchir à qui vous les donnez, de vérifier dans les petites lignes à quoi elles serviront, et d'identifier les gains que cela vous apportera. En effet, comme les professionnels veulent un maximum d'informations pour mieux vous cerner, ils proposent... [Lire la suite]



Réagissez à cet article

Source : *Big Data : gare à vos données personnelles !*

---

# Les principales tendances en 2016 du Big data



**La nouvelle année verra le Big Data prendre de l'ampleur et de la vitesse. L'évolution des usages des données va progresser. Les entreprises ne pourront pas ignorer la généralisation des analyses en libre-service et l'adoption à grand échelle du cloud et de Hadoop, ainsi que les nouvelles technologies venant compléter ce framework, entraînant de nombreux changements.**

## **1. Montée en puissance du NoSQL**

Dans nos prévisions des tendances du Big Data de l'année dernière, nous avons noté une accentuation de l'adoption des technologies NoSQL, qui sont généralement associées aux données non structurées. Dorénavant, les bases de données NoSQL commenceront à occuper une place centrale dans le paysage IT des entreprises, tandis que les avantages des bases de données sans schéma seront de plus en plus notables. Pour s'en convaincre, il suffit de consulter le Magic Quadrant de Gartner consacré aux systèmes de gestion de bases de données opérationnelles. Par le passé, Oracle, IBM, Microsoft et SAP dominaient le classement. Aujourd'hui, ce Magic Quadrant fait la part belle aux prestataires de solutions NoSQL, comme MongoDB, DataStax, Redis Labs, MarkLogic et Amazon Web Services (avec DynamoDB), qui viennent supplanter les fournisseurs traditionnels dans la catégorie des leaders.

## **2. Apache Spark révolutionne le Big Data**

De simple composant de l'écosystème Hadoop, Apache Spark est devenu une référence en matière de plate-forme Big Data pour de nombreuses entreprises. Spark offre une rapidité de traitement bien supérieure à Hadoop et constitue désormais le plus important projet Big Data open source, selon son créateur Matei Zaharia, également cofondateur de Databricks. Les exemples convaincants de mises en œuvre de Spark en milieu professionnel sont de plus en plus nombreux, à l'instar de Goldman Sachs qui en a fait sa solution de choix pour ses analyses du Big Data.

## **3. Les projets Hadoop arrivent à maturité et les entreprises passent à l'environnement de production**

Selon une enquête récente menée auprès de 2 200 clients Hadoop, seuls 3 % d'entre eux envisagent de moins s'appuyer sur ce framework au cours des 12 prochains mois. 76 % des sondés prévoient de recourir davantage à Hadoop au cours des 3 prochains mois, et enfin près de 50 % des entreprises qui n'ont pas déployé Hadoop déclarent avoir l'intention de le faire dans les 12 prochains mois. Cette enquête révèle également que Tableau est le principal outil d'aide à la décision pour les entreprises utilisant Hadoop ou prévoyant de l'utiliser, ainsi que pour celles qui ont déjà parfaitement intégré ce framework à leurs projets.

## **4. Le Big Data prend de l'ampleur : Hadoop s'ajoute aux normes de l'entreprise**

Hadoop occupe une place de plus en plus importante dans le paysage de l'IT. L'augmentation des investissements dans les éléments gravitant autour des systèmes professionnels, comme la sécurité, viendra corroborer cette tendance. Apache Sentry procure un système granulaire et basé sur des rôles pour gérer les autorisations d'accès aux données et métadonnées stockées sur un cluster Hadoop. Il s'agit là d'exemples de ce que les clients attendent de leur plate-forme de gestion de bases de données relationnelles. Ces fonctionnalités se retrouvent désormais à l'avant-garde des nouvelles technologies en matière de Big Data, ce qui simplifie d'autant plus l'adoption de telles solutions en entreprise.

## **5. Le Big Data prend de la vitesse : Hadoop aussi**

À mesure que Hadoop gagne en importance dans les entreprises, les utilisateurs attendent de plus en plus des fonctions d'exploration de données aussi rapides que celles proposées par les entrepôts de données traditionnels. En réponse à cette demande grandissante, nous assistons à une adoption croissante de technologies telles que Cloudera Impala, AtScale, Actian Vector et Jethro Data, qui favorisent la compatibilité de Hadoop avec les cubes OLAP et simplifient d'autant plus le rapprochement entre solutions traditionnelles d'aide à la décision et Big Data.

## **6. Le nombre croissant d'outils de préparation des données favorise la découverte d'informations**

Les outils de préparation de données en libre-service gagnent en popularité. Cette explosion est due en partie à l'adoption d'outils de découverte de données générées par les utilisateurs métiers, tels que Tableau, qui permettent d'accélérer les analyses. Désormais, ces utilisateurs souhaitent également réduire le temps nécessaire à la préparation des données et la complexité d'une telle opération, ce qui revêt une importance toute particulière pour le Big Data qui implique une multiplicité de types et de formats de données. Nous avons pu assister à de nombreuses innovations en matière de préparation de données pour le Big Data de la part de prestataires comme Alteryx, Trifacta, Paxata ou Lavastorm. Les leaders traditionnels en matière de solutions ETL, comme Informatica, qui développe Rev, ne sont pas en reste et réalisent d'importants investissements en la matière.

## **7. Les entrepôts de données à traitement MPP se tournent vers le cloud**

Même si la croissance du segment des entrepôts de données ralentit, ces solutions ne sont pas pour autant en passe de disparaître. Cette technologie opère actuellement sa transition vers le cloud, avec en tête Amazon Redshift, un entrepôt de données à la demande dans le cloud. Redshift est le service AWS qui a connu la croissance la plus rapide, mais il doit désormais faire face à la concurrence de Google et sa solution BigQuery, d'autres acteurs bien établis sur le marché comme Microsoft (avec Azure SQL Data Warehouse) ou Teradata, ou encore de nouveaux prestataires tels que Snowflake, lauréat du Strata + Hadoop World 2015 Startup Showcase, qui connaissent un succès grandissant dans le secteur. Les analystes estiment que 90 % des entreprises qui ont adopté Hadoop conserveront également leurs entrepôts de données. De plus, grâce aux nouvelles offres cloud, ces entreprises peuvent augmenter ou diminuer la capacité de stockage et la puissance de calcul de leurs entrepôts en fonction du volume d'informations stockées dans leur lac de données Hadoop.

## **8. Convergence de l'IoT, du cloud et du Big Data**

Même si l'IoT n'en est encore qu'à ses balbutiements, les pétaoctets de données générées par les différents objets connectés vont favoriser l'explosion des solutions cloud. Dans cette optique, les leaders en la matière, comme Google, Amazon Web Services et Microsoft proposent désormais des services d'IoT pour transférer en toute transparence ces données vers leurs moteurs analytiques dans le cloud.

Si ces tendances et ces évolutions peuvent sembler disparates, elles sont toutes liées par une même nécessité : exploiter les données rapidement et confortablement. Alors que le Big Data continue d'évoluer et que de nouvelles manières d'exploiter ces données voient le jour, une seule chose ne change pas : l'analyse de données est désormais à la portée de tous, et nous avons hâte de nous y mettre.



# Algorithmes prédictifs et Big Data



**Savoir ce que sera demain, ce mythe philosophique personnalisé par Cassandre dans l'antiquité Grecque est-il aujourd'hui en passe de devenir une réalité scientifique ?**

Nombreuses sont les applications qui aujourd'hui s'appuient sur des algorithmes prédictifs, que ce soient dans les domaines du marketing, de la finance ou encore de la santé. Cet engouement multidisciplinaire est le résultat d'un phénomène, qui naquit il y a plus de 10 ans mais qui a aujourd'hui trouvé un terreau favorable de croissance : le Big Data.

Socle de cette émergence, la donnée constitue un or noir disponible, aisée d'exploitation de prime abord mais aussi potentiellement facteur décisionnel dans certaines disciplines. Au cours des dernières années est apparu le terme de «données massives» pour englober la mine de renseignements collectée à partir de nos activités quotidiennes, des articles publiés, de nos interactions sociales et des objets de plus en plus connectés, c'est à dire eux même générateurs de données.

La donnée offre la possibilité d'envisager des évolutions à court, moyen et long termes comme des comportements à venir parfois même en temps réel. Elle permet de faire le grand écart entre la globalisation et l'individualisation. En effet, il est, à la fois, possible de suivre et d'anticiper les grandes épidémies à l'échelle mondiale tout en analysant le comportement individuel de monsieur X par rapport à ses déplacements ou ses achats.

La lutte contre la criminalité n'échappe pas à l'intérêt que présente la donnée dans la capacité à prévoir les évolutions et pourquoi pas le comportement criminel. En effet, les données massives constituent la source de l'analyse prédictive en ce qu'elles alimentent des applications à vocation opérationnelle.

Derrière la notion simplificatrice et vulgarisée d'algorithme prédictif se dissimule un processus analytique complexe et polyvalent. Loin d'être le fruit d'une génération spontanée, l'analyse prédictive repose sur des préceptes mathématiques et statistiques à des fins d'extraction de connaissances et de formes criminelles particulières. Il n'existe pas de logiciels ou d'algorithmes miracles pour lutter contre la délinquance. Il s'agit de développer des méthodes, de les tester, de les évaluer préalablement à une quelconque utilisation. Les méthodes reposent sur des techniques d'apprentissage capable d'exploiter les données dans leurs multiples dimensions. Loin de toute notion de préemption, l'analyse prédictive a une vocation de prévention, c'est à dire non pas d'agir préalablement à toute commission d'infraction mais plutôt d'interrompre l'évolution d'un processus en cours. A l'opposé des clichés véhiculés par la fiction (Minority Report, Person of Interest), l'analyse prédictive constitue une aide à la décision pour un chef opérationnel qui la complète par une approche prospective. En effet, l'anticipation criminelle nécessite de prendre en compte l'héritage des événements du passé, c'est la prédiction. Mais elle intègre aussi, en élaborant les scénarii les plus probables, des événements ponctuels impactant le futur, c'est le domaine de la prospective. Dès lors, en matière de lutte contre la criminalité, le pilotage ne peut s'effectuer par la donnée comme cela peut être le cas dans d'autres disciplines. En effet, l'analyse prédictive apporte des éléments objectifs de compréhension mais quoiqu'il arrive et en dépit de la masse de données disponible, incomplets. Elle est, par exemple, utilisée dans la prise en compte des perspectives d'évolution d'infractions de masse telles que les cambriolages. Elle peut aussi être utilisée dans la détection préalable de fraudes sociales ou bancaires ou encore dans la compréhension à des fins d'anticipation des variables pesant sur certaines formes d'infractions. En effet, la politique d'ouverture et de partage des données publiques permet de tirer profit de variables liées à un contexte social ou économique ou encore à l'évolution météorologique, voire à la création de nouvelles infrastructures.

En dépit de l'intérêt manifeste que présentent les méthodes prédictives, il convient de se garder de tout risque d'atteintes aux libertés individuelles. Ce point est un préalable obligatoire à un quelconque déploiement en matière de sécurité publique où la donnée à caractère personnel est exclue du champ d'analyse. Outre cette question essentielle, le risque de l'analyse prédictive est aussi une mauvaise interprétation de ce qu'est réellement l'apport de la prédiction. Gardons nous du joug infligé par Apollon à Cassandre.



Réagissez à cet article

**Source : *L'algorithme prédictif [par Patrick PERROT, Chef de la Division Analyse et Investigations Criminelles, Service Central de Renseignement Criminel de la Gendarmerie Nationale] | Observatoire FIC***

# Big Data Paris 2016



**Le congrès Big Data Paris 2016 se tiendra à Paris (Palais des Congrès de la Porte Maillot) le 7 et 8 mars 2016. Organisé par Corp Agency.**

Le Congrès Big Data Paris vous invite, pour cette 5e année, à plonger dans l'univers passionnant du prédictif ! Sommet du Big Data en France, le congrès a réuni plus de 6 500 participants en 2015, animés par un seul et unique but : participer à la construction et au développement d'une filière française d'excellence !

Véritable laboratoire d'innovation et de disruption, le Congrès Big Data Paris 2016 valorisera les acteurs les plus avant-gardistes de la scène internationale : retours d'expériences, conférences stratégiques et prospectives, parcours immersif ...

Plus d'informations sur le site de l'événement : <http://www.bigdataparis.com>



Réagissez à cet article

Source : *BIG DATA PARIS*

# Mégadonnées, petits secrets? – Pas facile d'assurer la confidentialité des données personnelles à l'ère des réseaux sociaux et de l'open data | Le Net Expert Informatique



## Mégadonnées, petits secrets? – Pas facile d'assurer la confidentialité des données personnelles à l'ère des réseaux sociaux et de l'open data

Vous croyez qu'il suffit d'enlever les informations nominatives qui figurent dans une banque de données pour assurer la confidentialité? Détrompez-vous. Non seulement est-ce insuffisant, mais même avec des mesures beaucoup plus serrées, la confidentialité d'une banque de données peut être facilement mise à mal par recoupement avec d'autres sources d'informations, notamment celles que les individus disséminent à tout vent sur le Net. Voilà la leçon de prudence livrée par Anne-Sophie Charest, professeure au Département de mathématiques et de statistique, aux participants du colloque «Big Data, le défi du traitement des données», présenté le 29 octobre sur le campus.

La professeure Charest compte au nombre de la dizaine de spécialistes que le Centre de recherche en données massives de l'Université Laval et l'Institut technologies de l'information et sociétés avaient réunis pour discuter du potentiel et des défis des mégadonnées, ces banques d'information si volumineuses et si complexes qu'elles exigent des méthodes de traitement particulières. D'entrée de jeu, Anne-Sophie Charest a rappelé les termes du contrat qui lie les chercheurs et les gens qui acceptent de participer à des enquêtes ou à des études. «On assure aux participants que les données resteront confidentielles et qu'elles ne seront utilisées qu'à des fins statistiques. Par contre, les chercheurs rendent publics des études ou des rapports à partir de ces informations et ils partagent même les données avec d'autres chercheurs ou avec la population. Le défi, qui existait même avant l'avènement du Big Data, est de concilier ces deux objectifs contradictoires.»

La solution intuitive, qui consiste à supprimer les informations nominatives, ne suffit pas à blinder une banque de données. À preuve, la professeure Charest a cité le cas du Massachusetts qui avait accepté, à la fin des années 1990, que les dossiers médicaux anonymisés des 135 000 employés de l'État soient mis à la disposition des chercheurs. Le gouverneur William Weld avait alors assuré que la confidentialité était garantie étant donné que les noms, les adresses et les numéros d'assurance sociale des employés avaient été supprimés. Une étudiante-chercheuse du MIT, Latanya Sweeney, aujourd'hui professeure à l'Université Harvard, avait toutefois trouvé une brèche de taille. En recoupant cette banque de données avec la liste électorale, elle a démontré qu'elle pouvait associer une bonne partie des dossiers médicaux à la personne correspondante. «Elle a même fait imprimer le dossier médical du gouverneur et elle l'a fait livrer à son bureau», raconte la professeure Charest.

Des organismes comme Statistique Canada et l'Institut de la statistique du Québec travaillent fort à assurer le respect de leur promesse de confidentialité, souligne Anne-Sophie Charest. La chercheuse explore, elle aussi, de nouvelles façons de compliquer le travail des personnes mal intentionnées qui tentent d'extraire des informations personnelles des bases de données publiques ou des publications qui en découlent. Ces différentes méthodes présentent toutefois un inconvénient important: si on les applique trop rigoureusement, on réduit l'accès aux données, ce qui n'est guère dans le ton en cette ère de l'open data, et on limite l'information utile qu'on peut en tirer.

Autre problème, il ne suffit plus que la confidentialité d'une base de données soit intrinsèquement protégée, il faut qu'elle le soit en tenant compte des recoupements possibles avec les autres sources d'informations. «Il est très difficile de prédire quelle information pourrait causer du tort au répondant si elle était rendue publique. L'approche de la confidentialité différentielle offre toutefois un compromis intéressant, estime-t-elle. Elle promet aux répondants qu'une tierce personne ne pourra rien apprendre de plus sur eux qu'ils acceptent ou non de participer à l'enquête. Cette nuance est importante considérant toute l'information que chaque personne diffuse maintenant sur elle-même. On ne peut plus fonctionner en vase clos.»

Par Jean Hamann

Source: Université Laval

Pour plus d'informations:  
Organisation: Université Laval  
Adresse: 1160 Université Laval  
Québec, Québec  
Canada, G1V 0A6  
www.ulaval.ca

Denis JACOPINI est Expert Informatique assermenté, consultant et formateur en sécurité informatique, en mise en conformité de vos déclarations à la CNIL et en cybercriminalité.  
Nos domaines de compétence :

- **Expertises et avis techniques** en concurrence déloyale, litige commercial, piratages, arnaques Internet... ;
  - **Consultant** en sécurité informatique, cybercriminalité, en accompagnement aux mises en conformité et déclarations à la CNIL ;
  - **Formateur et chargé de cours** en sécurité informatique, cybercriminalité et déclarations à la CNIL et accompagnement de Correspondant Informatique et Libertés.
- Contactez-nous

Cet article vous plaît ? Partagez !  
Un avis ? Laissez-nous un commentaire !

Source :

[http://www.rimq.qc.ca/detail\\_news.php?ID=551819&titre=Pas+facile+d%27assurer+la+confidentialit%C3%A9+des+donn%C3%A9es+personnelles+%C3%A0+l%27%C3%A8re+des+r%C3%A9seaux+sociaux+et+de+l%27open+data&cat=71;21](http://www.rimq.qc.ca/detail_news.php?ID=551819&titre=Pas+facile+d%27assurer+la+confidentialit%C3%A9+des+donn%C3%A9es+personnelles+%C3%A0+l%27%C3%A8re+des+r%C3%A9seaux+sociaux+et+de+l%27open+data&cat=71;21)

# Big Data et IoT bousculent le droit des individus | Le Net Expert Informatique



Big Data et IoT  
bousculent le  
droit des  
individus



**Gaspard Koenig, philosophe, apporte un éclairage original sur les relations entre Big Data, objets connectés et individus. Les paradigmes sont bousculés et l'intellectuel milite pour la création d'un droit de propriété sur les données personnelles.**

Les Assises de la Sécurité ont pris comme habitude « d'aérer » les esprits des RSSI et des partenaires avec des contributions de personnes non issues du sérail. Luc Ferry, Jean-Christophe Ruffin, Laurent Alexandre se sont pliés à l'exercice sur des sujets aussi variés que les propos sécuritaires, les mouvements de rébellion ou le hacking du cerveau. C'était au tour de Gaspard Koenig de se plier à l'exercice pour la 15ème édition des Assises. Normalien, philosophe et ancienne plume de Christine Lagarde, il avait choisi un thème elliptique : « l'utopie numérique est-elle dangereuse pour l'individu ? ».

#### **Basculer de la déduction à la corrélation**

Pour sa démonstration, l'utopie numérique se focalise sur le Big Data et l'Internet des objets, la partie individu se concentre sur la transparence ou la fin de la confidentialité. Pour lui, il s'agit tout d'abord d'une question épistémologique. « Le Big Data est la revanche de la sensibilité sur la généralité », explique-t-il en se basant sur la seconde méditation de Descartes et l'épisode du morceau de cire, qui pose comme postulat que l'appréhension d'un objet dans sa globalité est dictée par la raison. Or le Big Data remet en cause cette façon de penser en appliquant des singularités au morceau de cire. Cela signifie que les objets ne sont plus classés dans des catégories, mais selon plusieurs subtilités, finalités. « Le Big Data fait évoluer la science de la déduction à la corrélation. On est dans le quantitatif, dans la convergence de données », constate le philosophe. Et ce changement de paradigme a déjà des applications concrètes. Dans les assurances, les objets connectés vont « redéfinir l'individu » (façon de conduire, habitude de vie et de consommation) alors qu'il était catégorisé selon certaines données (âges, sexes, etc).

#### **Bonheur collectif ou individuel ?**

Pendant de cette course à « la connaissance universelle », il y a les problématiques de vie privée et de transparence des données individuelles. « Le Big Data a une finalité qui est de gagner du bien-être, des gains de productivité, de confort, etc. Il s'agit d'une vision utilitariste pensée par Jeremy Bentham pour qui le but de toute politique est d'aboutir au bonheur des gens », remarque Gaspard Koenig. Une analyse qui ne lui convient pas et qui a été critiquée par « John Stuart Mill qui substituait le bonheur général au bonheur individuel, c'est-à-dire le droit à la diversité à la faillibilité ». Mais dans ce cas-là, « pourra-t-on refuser le trajet dicté par une voiture autonome ou ne pas vouloir partager les informations d'un compteur intelligent sans impacter le bonheur général ? ». Des exemples triviaux, mais qui pose la question des missions de l'Etat, « apporter le bonheur collectif ou protéger les droits individuels ».

#### **Un droit de propriété des données personnelles**

La réponse apportée par le philosophe est de « donner un droit de propriété sur les données personnelles ». Aujourd'hui, on part du principe que dans le Big Data, nos données sont gratuites. « Elles sont en réalité pillées par les GAFAs avec l'aval des utilisateurs via les 'terms and conditions' qui comme dans le cas de Paypal sont plus longs que Hamlet », accuse Gaspard Koenig. Pour lui, il faut rentrer dans une logique patrimoniale des données personnelles. « Nous avons inventé le droit de propriété intellectuelle lors de la révolution industrielle. A l'heure du numérique, il faut créer un droit de propriété des données individuelles ». Une approche qui implique de reconfigurer le business model de l'Internet, « Google devrait rémunérer les gens pour obtenir leur données », conclut le philosophe un brin utopiste, mais qui sait...

Même si remplir un formulaire de déclaration à la CNIL est gratuit et enfantin, il vous engage cependant, par la signature que vous apposez, à respecter scrupuleusement la loi Informatique et Libertés. Cette démarche doit d'abord commencer par un Audit de l'ensemble de vos systèmes de traitements de données. Nous organisons régulièrement des **actions de sensibilisation ou de formation** au risque informatique, à l'hygiène informatique, à la cybercriminalité et à la mise en conformité auprès de la CNIL. Nos actions peuvent aussi être personnalisées et organisées dans votre établissement.

Besoin d'informations complémentaires ?

Contactez-nous

Denis JACOPINI

Tel : 06 19 71 79 12

formateur n°93 84 03041 84

Denis JACOPINI est Expert Judiciaire en Informatique, consultant, formateur et chargé de cours.

Nos domaines de compétence :

- **Expertises et avis techniques** en concurrence déloyale, litige commercial, piratages, arnaques Internet... ;
- **Consultant** en sécurité informatique, cybercriminalité et mises en conformité et déclarations à la CNIL ;
- **Formateur et chargé de cours** en sécurité informatique, cybercriminalité et déclarations à la CNIL.

Contactez-nous

Cet article vous plait ? Partagez !  
Un avis ? Laissez-nous un commentaire !

Source : <http://www.silicon.fr/assises-de-securite-2015-big-data-iot-bousculent-droit-individus-127948.html>  
Par Jacques Cheminat