

## COMMUNIQUE DE PRESSE

### Ouvrage « Le Big Data dans l'assurance »

Par Marc Dupuis & Emmanuel Berthelé

Optimind Winter annonce la publication, aux éditions de l'Argus de l'assurance, de son ouvrage « Le Big Data dans l'assurance ». A travers cet ouvrage, les auteurs et contributeurs apportent un éclairage inédit sur les enjeux liés au Big Data dans l'assurance.

De la compréhension des phénomènes sous-jacents aux réflexions déontologiques, le lecteur pourra aisément se forger une opinion sur les opportunités et les risques associés à cet outil de transformation massive qu'est le Big Data, et ainsi comprendre qu'il n'est pas une fin en soi mais bien un nouveau vecteur de création de valeur.

« Le Big Data n'est pas une éventualité, une simple option ou un concept futuriste ! La lecture de cet ouvrage permet de prendre conscience du caractère particulièrement immédiat et révolutionnaire qu'il représente aujourd'hui pour l'industrie assurantielle. » déclare Christophe Eberlé, président d'Optimind Winter.



« Il est désormais important, pour le secteur assurance, de mettre en œuvre des projets concrets, afin d'entrer dans une phase de valorisation des nouveaux flux de données. C'est à cette condition que les promesses pourront devenir réalité et que ces outils seront pleinement associés à la nouvelle ère numérique », ajoute Marc Dupuis, directeur associé en charge du métier Digital.



« Même s'il convient de s'interroger quant à l'impact qu'aura le Big Data sur le secteur de l'assurance, une remise en question profonde des prérogatives des actuaires ne semble pour autant pas de rigueur à court terme. De nouvelles approches et un enrichissement mutuel au travers de collaborations avec les data scientists seront très vraisemblablement au rendez-vous. », précise Emmanuel Berthelé, actuaire senior, responsable de la Practice.

L'extrait de l'ouvrage est disponible sur [ce lien](#). Pour plus d'information, merci de contacter Optimind Winter à l'adresse [edition@optimindwinter.com](mailto:edition@optimindwinter.com).

#### Qui sommes-nous

Leader de l'actuariat conseil et de la gestion des risques en France avec ses 180 collaborateurs, Optimind Winter constitue l'interlocuteur de référence pour les organismes assureurs, banques et grandes entreprises qui souhaitent un partenaire métier de haut niveau les accompagnant dans leurs projets stratégiques en actuariat conseil, protection sociale, risk management, finance & performance, business transformation, digital.



# Big Data dans l'assurance



Marc DUPUIS et Emmanuel BERTHELÉ

optimind winter.∴.

# Sommaire

|   |     |
|---|-----|
| <b>Introduction</b>   | 5   |
| <b>1. Cerner les fondamentaux et les enjeux du Big Data</b>                             | 9   |
| Bien commencer avec quelques définitions et chiffres                                    | 9   |
| L'évolution de l'informatique crée un environnement favorable à l'émergence du Big Data | 11  |
| La révolution digitale en marche  | 16  |
| Écosystème du Big Data : les acteurs du Big Data  | 20  |
| Quelques enjeux pour le Big Data  | 21  |
| <b>2. Placer la donnée au cœur de ce nouveau paradigme</b>                              | 23  |
| La donnée comme point de départ   | 23  |
| Les sources de données se diversifient  | 26  |
| Données non structurées : un des enjeux majeurs du Big Data                             | 31  |
| <b>3. Identifier l'écosystème du Big Data</b>   | 35  |
| Les besoins comme moteur de l'innovation  | 35  |
| Les fondements technologiques des outils Big Data                                       | 36  |
| Architectures et solutions Big Data   | 45  |
| Les algorithmes   | 49  |
| <b>4. L'impact du Big Data sur le secteur de l'assurance</b>                            | 59  |
| L'environnement spécifique de l'assurance   | 59  |
| Les solutions Big Data : quels impacts en assurance                                     | 66  |
| De nouveaux paramètres s'immiscent dans la gestion des risques                          | 69  |
| Modèles et modélisation   | 71  |
| Aller au-delà des données des portefeuilles   | 72  |
| <b>5. Appréhender une nouvelle relation entre assureur et assuré</b>                    | 75  |
| Les impacts du Big Data pour l'assureur   | 75  |
| Une nouvelle approche de l'assurance pour le client                                     | 80  |
| Vers une convergence des intérêts   | 81  |
| <b>6. Passer de la théorie à la pratique</b>  | 83  |
| Comment concevoir un projet Big Data ?  | 84  |
| Une nécessaire transformation de l'entreprise   | 89  |
| Des clés pour le lancement d'un projet Big Data   | 93  |
| <b>7. Cerner les enjeux réglementaires, juridiques et sociétaux</b>                     | 103 |
| Réglementation : contraintes et opportunités  | 103 |
| Comment la certification peut fournir une légitimité                                    | 114 |
| Réflexions sociologiques et sociétales  | 120 |
| <b>Conclusion</b>   | 123 |
| <b>Index alphabétique</b>   | 125 |

## 3.1 Le cloud

### ► Le développement des grilles de calcul et de stockage : un préliminaire au cloud

La décennie 1990-2000 a vu se développer le « grid computing » (grilles de calcul). Une fois parallélisés, il convient de répartir l'exécution des calculs sur des ordinateurs internes et non utilisés (la nuit par exemple) et plus tard à distribuer ces calculs sur des serveurs localisés n'importe où dans le monde : les grilles de calcul étaient nées.

Le principe est simple : si le traitement de  $E$  éléments demande un temps  $T$  de traitement sur une seule machine et si le traitement est parallélisable, passer à  $N \times E$  éléments peut toujours être réalisé dans le même temps  $T$  en utilisant  $N$  machines. En pratique un serveur distribue les calculs sur les différentes machines puis collecte les résultats de chaque calcul réalisé sur chacune des machines pour ensuite les agréger.

D'autres secteurs d'activité tels que les sciences de la vie - séquençage du génome et des protéines, recherches sur le cancer - les industries mécaniques et aéronautiques - simulations et processus - les États - recherche militaire et nucléaire - et bien entendu les services financiers - analyses de produits dérivés, statistiques, analyses des risques sous Bâle 3 et Solvabilité 2 - utilisent les grilles de calculs.

Parallèlement, des grilles dédiées au stockage ont vu le jour. La combinaison des grilles de calcul et des grilles de stockage est l'un des piliers fondateurs des technologies utilisées pour le traitement des Big Data, notamment les technologies Hadoop qui sont décrites dans cet ouvrage.

### ► Le passage au cloud : une évidence économique pour mutualiser la puissance de calcul au service de tous

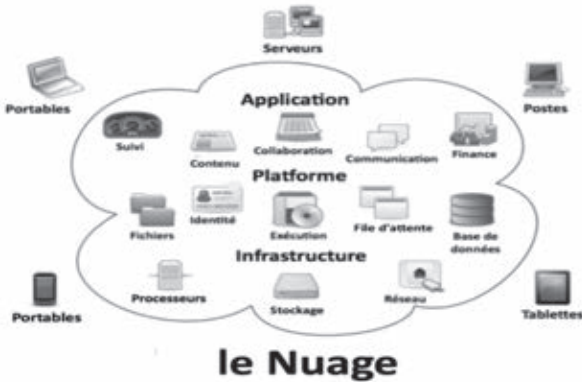
L'utilisation de grilles de calcul est devenue très populaire grâce au lancement en 1999 de SETI@Home dans le cadre du projet SETI visant à détecter des signaux d'origine extraterrestre. SETI@Home utilise les processeurs de milliers d'ordinateurs de particuliers volontaires et connectés à Internet pour analyser les données collectées.

On peut considérer qu'il s'agit là d'un prototype du cloud computing qui consiste à exploiter la puissance de calcul et de stockage à travers un réseau. Si le réseau est internet, on parle de cloud public, s'il s'agit d'un réseau interne à l'entreprise, on parle de cloud privé.

L'intérêt de mutualiser la puissance de calcul et de stockage est d'abord économique en remarquant que tout le monde n'aura pas besoin de puissance de calcul au même moment.

Cela permet aussi de disposer de puissance de calcul à la demande et de ne payer que ce qui est consommé. En effet, les applications en mode SaaS (software as a service) l'exploitent largement en proposant une offre modulaire où seuls les modules nécessaires sont loués par l'utilisateur, tandis que les applications résident chez le fournisseur. D'autres déclinaisons sont proposées : PaaS (platform as a service), IaaS (infrastructure as a service), etc.

Enfin, il faut souligner le véritable changement de paradigme pour les directions informatiques, habituées à disposer de ces machines en interne ou en infogérance. Cela explique certainement la frilosité de certains acteurs à tirer pleinement parti du cloud, notamment dans le monde de l'assurance.



Source : Sam Johnston, Internet.

### 3.2 Social

Les médias sociaux, ou encore le Web 2.0, rendent l'interaction sociale et le partage de contenu si facile qu'ils positionnent chaque individu comme contributeur. L'intelligence en réseau, collaborative et collective, enrichit notablement les informations dont chacun peut disposer. Moins timide derrière un écran, masqué par un pseudo, l'ouverture vers une tranche de la population qui ne se serait pas forcément exprimée en pleine lumière enrichit plus encore la diversité des opinions exprimées et le contenu ainsi construit et rendu accessible à tous.

Le Web 2.0 réalise une transformation de la relation entre individus, pose la question de la perte de liant social, paradoxale sur des sites d'échanges et de discussion, tout en participant à la construction de passerelles entre monde virtuel et monde réel, relations intermédiées par la technologie et relations humaines directes. Dans ce contexte, qualifier la valeur ajoutée et la qualité des informations devient clef : c'est d'ailleurs l'un des nouveaux enjeux de l'éducation. Analyser toute cette masse d'informations présente toutefois un réel intérêt : l'ensemble des avis, critiques, remarques et commentaires constitue des bruits faibles, des tendances de fond en train de prendre forme, des comportements de masse et donne une image de notre société qu'une multitude d'acteurs cherche à comprendre et à analyser. Toutes ces informations à l'échelle du web constituent une base de renseignements d'une grande valeur.

Ces données entrent clairement dans le champ du Big Data et l'intérêt de la puissance des outils d'analyse associés devient évident tant les volumes générés, chaque jour, sont considérables.

### 3.3 Mobilité

#### ► Une double déclinaison de la mobilité

- **L'utilisateur en tant qu'individu connecté au monde par son smartphone, sa tablette ou son ordinateur**

Grâce aux services en temps réel, il est possible de disposer sans démarche des informations utiles pour un utilisateur donné, à un instant précis, et ce de manière contextualisée - à la

location de l'utilisateur notamment (mode push). Il est également possible d'aller chercher l'information (mode pull) où qu'elle soit par rapport à une position à cet instant.

• **L'utilisateur capable de gérer à distance des objets connectés**

- pour la surveillance de ses proches : enfants, parents âgés...
- pour la gestion de ses biens, du fait du développement de la domotique notamment.

► **Des modes de connexion encore en devenir : la technologie tend à être portable**

Devenus le mode de connexion le plus répandu, les smartphones remplissent encore parfois la fonction pour laquelle ils ont été initialement conçus : téléphoner !

D'autres modes émergent au fur et à mesure que la technologie, et surtout la miniaturisation, les rendent suffisamment robustes pour un large déploiement :

- des lunettes à réalité augmentée : superposition à ce que l'on voit en temps réel d'un modèle virtuel en 2D ou 3D ;
- des montres connectées offrant, outre des fonctionnalités d'information de l'individu, différentes capacités de mesure liées à l'individu ;
- des systèmes de combat, par exemple liés au projet français « fantassin à équipements et liaisons intégrés », améliorant les capacités de communication, d'observation, de protection et de mobilité des fantassins grâce, notamment, à une plateforme électronique intégrée aux vêtements et reliant les équipements périphériques entre eux : lunettes de tir, jumelles, écrans, GPS, radio, moyens de vision diurne et nocturne.

### 3.4 Le Big Data : vision disruptive

Tout en étant une voie d'évolution naturelle de la révolution digitale, le Big Data est une révolution en devenir pour certains secteurs économiques dont l'assurance. Elle se trouve déjà intégrée comme une composante essentielle de l'activité pour d'autres, comme la téléphonie mobile ou la grande distribution.

Le Big Data ne casse pas l'existant et ne le remet pas en cause. Il constitue un enrichissement et une évolution par l'ajout d'autres informations.

En revanche, il induit une révolution mentale, une disruption dans la recherche d'informations contenues dans les données à disposition. Tel qu'exposé plus haut, identifier les informations dispersées dans la masse des Big Data, à l'image d'une aiguille dans une botte de foin, implique une approche sans idée préconçue.

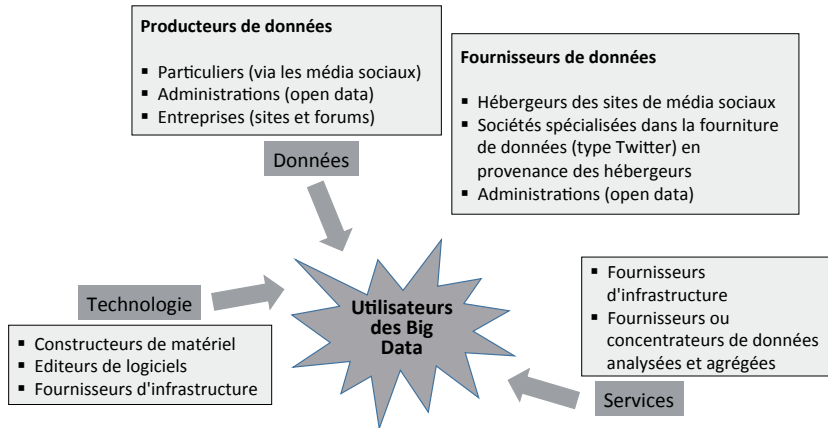
La démarche de formulation d'hypothèses, pour chercher à les vérifier, n'est plus la seule option possible pour se mettre dans un état d'esprit où l'on ne sait pas, à l'avance, ce que l'on recherche. Il en résulte que l'on doit apprendre à se tromper et accepter ses échecs, apprendre à lancer des projets sans certitude de réussite, apprendre à les arrêter si les résultats escomptés ne sont pas au rendez-vous, mais également apprendre à se réinventer sans cesse.

**!** **IMPORTANT**

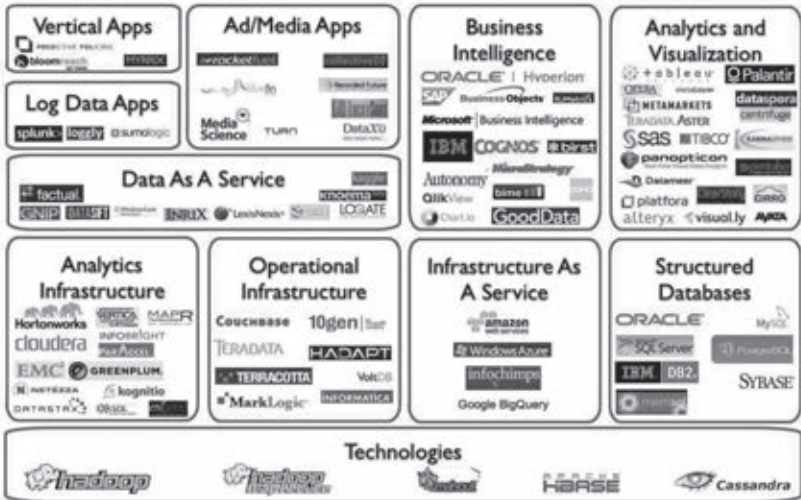
L'expérience utilisateur devient une priorité et prend une position centrale dans tous les systèmes et applications. C'est une nouvelle approche que doivent envisager les entreprises pour répondre à cette exigence de l'utilisateur qui veut une réponse adaptée à ses besoins précis, et pas ceux de son voisin, besoins actuels et futurs donc une réponse flexible et évolutive.

## 4. Écosystème du Big Data : les acteurs du Big Data

L'écosystème du Big Data est constitué de multiples intervenants, des producteurs de données primaires aux utilisateurs des informations extraites de ces données. En voici une première classification par famille d'acteur :



Et une autre par société - sources 2014



Les nombreuses réactions sociétales vis-à-vis de la confidentialité et de la protection des données demeurent et, en même temps, s'estompent, devant l'avantage que chacun peut y trouver. L'individu qui utilise des applications ou services gratuits devient l'objet à vendre à d'autres, selon la célèbre maxime : si c'est gratuit vous êtes le produit. Contrairement à une idée reçue assez largement répandue, nombreux sont ceux qui acceptent ou accepteraient de transmettre des informations confidentielles s'ils y trouvent un avantage.

## 5. Quelques enjeux pour le Big Data

Données massives, complexes et récoltées en temps réel, le Big Data nécessite des formes innovantes de traitement, d'autres modes de pensée et des approches différentes : c'est l'un des enjeux fondamentaux du Big Data.

L'intégration du Big Data dans les applications existantes pour enrichir les données disponibles au sein de l'entreprise, fait évoluer l'architecture technique des outils de « business intelligence », voire de l'ensemble du système d'information.

La mise en œuvre et l'exploitation des technologies Big Data pour traiter les données internes ou réduire les temps de calcul constitue une première étape pour se familiariser avec l'ensemble de cet écosystème.

Enfin, et c'est certainement l'enjeu principal, il faut rendre la donnée intelligente pour la transformer en information.

