

Blog Customer Matrix

Big Data : Faites danser l'éléphant !

Posté le 3 mai 2012

Révéler l'information pertinente à chaque utilisateur issue du tsunami de données incessant, voici la promesse des technologies Big Data.

Suite à l'ouverture de la boîte de Pandore par le cabinet McKinsey, de nombreuses discussions ont suivi, certains craignant de percevoir le Big Data comme « une stratégie marketing pour exciter les foules IT pensant la crise ».

Car au juste, le datamining – qui, selon le mythe, permit à Wal-Mart de comprendre qu'il fallait placer les Pampers à côté des Budweiser pour les pères de famille effectuant leurs courses le samedi après-midi - le data-mining donc ne sonnerait pas un peu comme ce que l'on définit comme Big Data ?

Datamining – Business analytics – Business Intelligence – Pourquoi le Big Data aujourd'hui ?

Parce que les données ont changé

En 2008, Chris Anderson distinguait déjà l'obsolescence des méthodes traditionnelles face à l'explosion de données hétérogènes et multidirectionnelles : « The Data Deluge Makes the Scientific Method Obsolete ».

En effet, l'avènement bien heureux du web conversationnel et du contenu généré par les utilisateurs ont profité à l'explosion des **données non-structurées** jusqu'alors marginales.

Le numérique grâce à la multiplicité des devices, ne lâche plus l'utilisateur qui fournit et consomme des données tout au long de sa journée (voir l'infographie ci-dessous). L'expansion du volume de ces données fines, comportementales et non pré-formatées au traitement de données traditionnel est également propulsée par les médias sociaux et autres outils collaboratifs.

Bref, les données non structurées vont augmenter de 80% au cours des cinq prochaines années, selon Gartner.

Les avantages concurrentiels et les opportunités des entreprises naîtront en se forgeant des intuitions, fruit du croisement de données de types variés; relationnelles et non relationnelles, structurées et non structurées, internes et externes à l'entreprise.

Parce que les technologies d'analyse de données ne suivent plus l'amélioration de la vitesse de stockage et d'indexation

Si le stockage des données est bien meilleur marché qu'hier, il est également plus rapide voire quasi-instantané.

Les serveurs de plus en plus rapides, robustes et abordables ont permis aux entreprises de stocker de monstrueux volumes de données dans nos datawarehouses qui ne peuvent plus être explorés avec les outils traditionnels. Pour jouir de l'investissement en stockage supplémentaire, les entreprises doivent pouvoir analyser leurs données.

Par ailleurs, les acteurs du marché du stockage ont permis de réduire le temps de stockage des données. Mais il s'agit d'un gain marginal si les outils d'analyse et de recherche prennent toujours autant de temps.

Ce sont précisément ces éléments qui ont donné naissance à une nouvelle génération de processus et d'outils nommés **Big Data**.

Création, partage et stockage de données se rapprochent du temps réel, l'imposant datawarehouse doit pouvoir garder le rythme.

En une métaphore, nous dirions, que le Big Data permet de faire danser l'éléphant ou de le rendre agile.

Ici, on aurait préféré l'appeler Poly Data car la data est avant tout polysources, polyformat et polyusers.