



Plan du cours

- 2-3 Exemples
- Introduction
- Le phénomène
- C'est quoi?
- Pour qui, pour quoi
- **Le Big Data pour l'industrie**
- **Les enjeux**
- Conclusion



Big Data : détection des fraudes dans les services financiers

Fraude bancaire = un réel problème

➤ **600- 800 millions d'euros dérobés en 202X**

« Le Big Data améliore considérablement l'expérience de lutte anti-fraude »

Intérêt du Big Data

- Traitement en temps réel des données ➔ **Réactivité inégalée**
- Réduction des délais de traitement des opérations de détection
3 semaines à 5 minutes
- Augmentation des taux de détection

Comment ?

- Améliorations des solutions existantes grâce aux algorithmes
- Les comportements anormaux détectés sont affinés
- Détection des fraudes non avérées en temps réel. (Grâce à un large volume de données analysées)
- Comportements à risques plus facile à détecter grâce au **machine learning**

Un exemple de fraude bancaire

- 202X : Jems Groupe intervient sur des problèmes de blanchiment d'argent.

Limite de la démarche

- Analyse limitée aux transactions supérieures à 10 000 euros
- Masse de données trop importantes pour étudier les autres cas
- Analyse longue qui s'étale sur la journée



Mise en place d'un système de Big Data

« Une aide considérable en matière de lutte anti-fraude. »

- Possibilité de traiter toutes les transactions
- Analyse en temps réel

Avancées considérable en matière de détection mais toujours insuffisant contre les attaques

Exemple: Big DATA et SANTÉ

2003

10 ans pour un
séquençage total de
l'ADN humain.

2018-2X

Moins de 24h pour un
séquençage total de
l'ADN humain.

Il existe des centaines de bases de données publique en épidémiologie, médico-économique ou encore sur des études en cours.

Augmentation du nombre d'objets connectés pour la santé.
Les compagnies du GAFA s'en mêlent.

Le résultat est une augmentation massive du nombre de données à traiter



Introduction

- Le traitement de l'information, un dilemme permanent pour les entreprises.
- Le volume d'information est passé de peu abondant à surabondant en quelques années.



Introduction



- Ce nouveau paradigme peut se résumer en une phrase : **une abondance de données sans réelle explication et sans contexte rend difficile la transformation de ces données en informations actionnables.**



Introduction

- **Faits**
- Selon les prévisions d'ici à 202X, environ 3,0 MB de nouvelles informations **seront créés chaque seconde pour chaque être humain** de la planète.
- De la grande distribution à la santé, du voyage à l'automobile, chaque secteur devra rapidement adopter les services numériques adaptés.



Introduction

- **Sources multiples: sites, base de données, téléphone, serveurs:**



(IOT)

- Détecter les sentiments et les réactions des clients
- Détecter les conditions critiques ou potentiellement mortelles dans les hôpitaux, et à temps pour intervenir
- Prédire des modèles météorologiques pour planifier l'usage optimal des éoliennes
- Prendre des décisions risquées basées sur des données transactionnels en temps réel
- Identifier les criminels et les menaces à partir de vidéo, sons et flux de données (téléphone (GPS))
-

Introduction

- **Sources multiples: sites, base de données, téléphone, serveurs:**
 - Les données massives sont le résultat de la rencontre de trois éléments essentiels qui sont:
 - Internet
 - Les réseaux sociaux
 - Les appareils intelligents: les ordinateurs, les tablettes, smartphones, les objets connectés.
 - L'internet permet la transmission de l'information quelle que soit sa forme sur les appareils intelligents
 - Appareils intelligents ; création de données
 - Utilisateur des réseaux sociaux : consommateur
 - Internet: vecteur de transmission.

Introduction

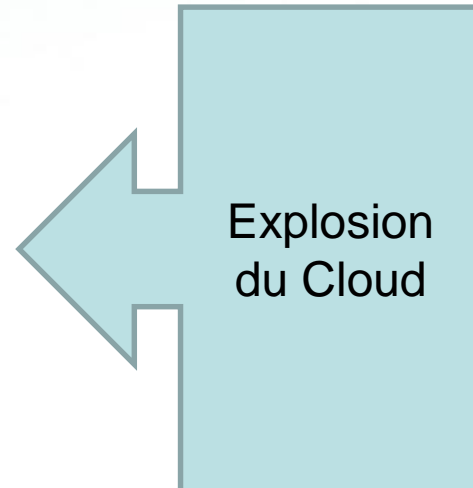
- Il devient donc important de savoir comment **traiter cette information pour en tirer des tendances en termes de nouveaux business** dans des perspectives particulières telles que;
 - Réorganiser les organisations
 - Parfaire la connaissance client
 - Innover plus vite
 - Favoriser l'économie collaborative
 - etc.

Introduction



- **Défis**

- Réunir un grand volume de données variées pour trouver de nouvelles idées.
- Capturer des données créées rapidement.
- Sauvegarder toutes ces données.
- Traiter ces données et les utiliser.



Introduction



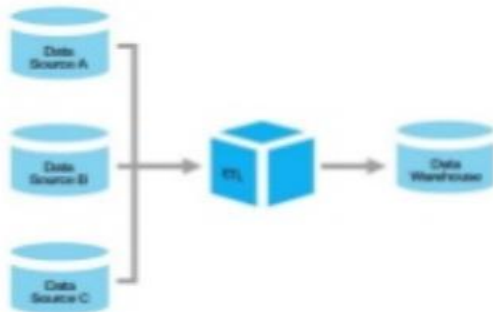
- **Rappel des fondamentaux : Business Intelligence versus Big Data.**
 - La BI consiste en un ensemble d'outils et de techniques permettant de collecter, de nettoyer et d'enrichir des données structurées ou semi structurées pour les stocker dans différentes formes de base de données multidimensionnel.
 - Les données vont donc être gérées dans des formats normalisés pour faciliter l'accès à l'information et les vitesses de traitement.

Introduction

- **Bi traditionnel**



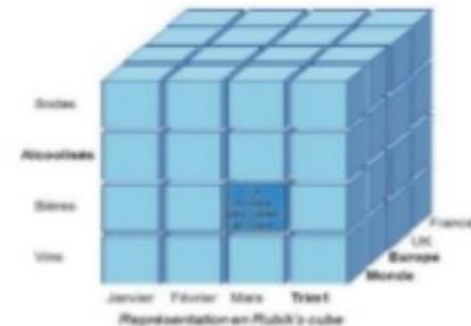
Mécanismes d'extraction,
de transfert et de consolidation



Centralisation des données



L'analyse multidimensionnelle
des données



Reporting



Introduction

- Les technologies du Big Data permettent de:
 - stocker les mêmes données
 - mais dans des contextes différents
 - en appliquant des traitements distincts
 - **et des séries d'algorithmes différenciées**
 - **et ceci pour traiter plusieurs problématiques simultanément.**

Introduction

- Méthodologies d'analyse différentes



Introduction

- **Big Data et décisionnel**



Approche Big Data versus Traditionnelle



- La question n'est pas :
 - De choisir entre l'approche classique et l'approche Big Data?
 - Mais plutôt comment le faire?
 - Comment les faire fonctionner ensemble?



Le Big Data, c'est quoi ?

Le Big Data vise à ;

- Tirer un avantage concurrentiel au travers de méthodes de collecte
- D'analyser et d'exploiter des données qu'on ne pouvait utiliser jusqu'à présent du fait des contraintes économiques, fonctionnelles et techniques liées aux volumétries, à la vitesse de traitement et à la variété des données à considérer.



Big Data: pour qui, pour quoi?



Industrie

- Produit comme un service
- Qualité, innovation R&D
- Maintenance préventive

Distribution



- Offres temps réel et service personnalisés
- Optimisation de l'expérience magasin
- Pricing dynamique

Banques



- Parcours clients multi-canaux
- Fraude, anti blanchiment
- Partage des données consommateurs pour personnalisation

Assurance



- Fraudes et risques
- Recommandation client
- Tarification à l'usage, personnalisation

Santé



- Gestion des effets indésirables
- Traitements personnalisés.
- Amélioration des diagnostics

Transports, loisirs



- Planification et gestion des evts liés à la logistique
- Service client temps réel
- Economie d'énergie
- Pricing dynamique

Secteur public



- Services informationnels
- Fraudes, abus
- Sécurité publique
- Personnalisation de la relation cito

Telecom



- Parcours clients multi-canaux
- Partage de données de géo localisation
- Fraudes et analyse du comportement client

Produits gde conso.



- Analyse de sentiments et retour produits
- Relation personnalisée avec le consommateur
- Produit comme un service

Exemple application TD Conduite auto

Exemple Pandemie Alerte Covid

Des applications qui se déclinent dans tous les secteurs d'activité

[Score et détails](#)

Vous

Les autres

Vous vous dirigez vers un rabais de **9%**

Trajets récents

[Voir plus](#)

2 juin à 14:23



Ville de Saguenay, Québec

Ville de Saguenay, Québec



32 km - 34 min



Bilan du trajet



02 juin

32 km | 34 min



Ville de Saguenay, Québec

14:23



Ville de Saguenay, Québec

14:57



Vitesse



Fluidité



Vigilance

Vous avez eu un accident?

Allez à un endroit sûr. Appelez-nous immédiatement pour obtenir de l'aide (service de remorquage ou d'un atelier de réparation).

 Appeler pour de l'aide

Conseils de conduite



Des épisodes récents d'accélération ont abaissé votre score global d'accélération.



Des frein scor

[Voir tous les conseils de conduite](#)

Détails sur votre dernier trajet

lundi, mars 18, 2019 à 15:06



Boul Talbot



Université du Québec à Chic...



1.72 km

 4 min

97



Fermer



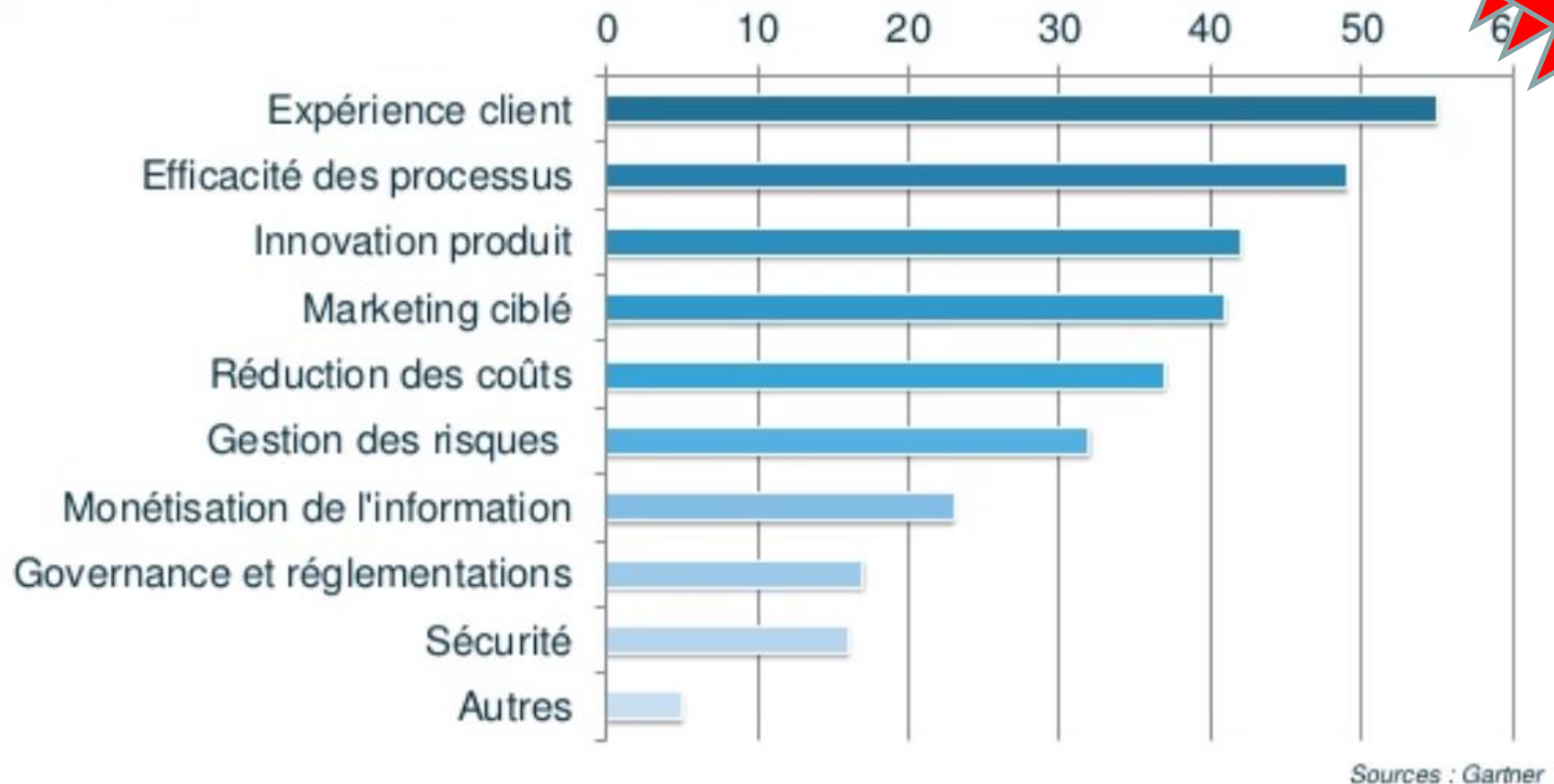
Fonctionnement d'Alerte COVID

L'application s'exécute en arrière-plan sans interrompre vos activités.

Quand vous serez près d'une personne ayant Alerte COVID, vos téléphones s'échangeront des codes aléatoires toutes les 5 minutes.

Ces codes changent souvent et ne peuvent pas servir à vous identifier.

Le Big data: Pour quoi faire?



La relation client avant tout, puis l'efficacité des processus et l'innovation

Le Big Data pour l'industrie. De quoi s'agit-il?



- Réduire les arrêts non programmés
- Améliorer la fiabilité des équipements
- Réduire la perte de production
- Optimiser les coûts de maintenance
-

Les enjeux du Big Data



Les enjeux du Big Data



1. Garantir la qualité des informations

- La qualité des données doit être une priorité pour ne pas fausser les stratégies découlant de leur traitement.

2. Optimiser le traitement des données

Les informations arrivent en masse et se présentent sous divers formats. L'entreprise doit donc investir dans des outils de gestion et de traitement permettant de transformer les données recueillies en informations intelligibles.

Les enjeux du Big Data



3. Mettre en relation tous les métiers

- Faire travailler ensemble différents professionnels dans le but d'atteindre des objectifs précis.

4. Assurer la sécurité

- La sécurité de l'information a une importance particulière, car elle engage la responsabilité et la réputation de l'enseigne.

Les enjeux du Big Data



5. Humaniser les données

- L'entreprise ne doit pas oublier l'humain derrière les données recueillies. De toute façon, l'intérêt du Big Data est de placer les clients au centre du processus décisionnel.

Conclusion



- Au final, avec toutes ces mesures (organisationnelles, métiers, techniques...), le positionnement des données dans l'entreprise est en passe de changer radicalement.
- Elles prennent une place centrale dans le SI et dans les préoccupations des métiers.
- Avec cela, Business Intelligence et Big Data sont très complémentaires et vont faire bon ménage dans les années qui viennent.

Conclusion

Big Data = Du traitement organisé et structuré de gros volumes de données, au rêve de la prédictibilité...

