

# Comment implanter le Big Data dans l'entreprise

Leleux Laurent

2017 - 2018

# Les modèles de mise en place

- Choix premier :
  - Data-warehouse et BI
  - Big data
- Traitements plus rapides
- Stockage plus important
- Coûts plus faibles

# Modèle disruptif

- Big data comme élément principal du domaine décisionnel
- Hadoop + MapReduce + Machine learning + visualisation
- Approche réactive > < étude à posteriori
- Données structurées ou non (2 in 1)
- Rmq : Yahoo = Hadoop 40 000 Machine, 350Po

# Modèle évolutif

- Traitement en amont d'un data warehouse
- Conservation de l'infrastructure existante
- Enrichissement des données
- Protection du DW
- The Now Factory

# Modèle Hybride

- Flux vers le DW
- Nouveaux flux vers le Big Data
- Intégration des deux
  
- Conservation du système décisionnel existant
  
- Teradata, SAP...

# Compétences nécessaires

- Trois ressources
  - Données (internes ou externes)
  - Extraction de la valeur
  - Méthode de gestion des idées créatives pour la transformation en produits/services

# Gestion des données

- Département informatique
- Nouveaux modèles de stockage/traitement
- Hadoop
- NoSQL
- Intégration, cohérence, budget...

# L'extraction de la valeur

- Data scientist
  - Statisticien et mathématicien
  - Profil de consultant
- 
- Connaissances générales globales
  - Maîtrise des coûts



# Management

- État d'esprit « Big Data »
- Cultiver une image positive et dynamique autour du concept
- Encourager à créer de la richesse
- Manque de diplômés

# Changements organisationnels

- Expert-métier
- Outils de recommandations d'Amazon
- Ligne éditoriale de Google
- Plateformes de trading
- Processus de gestion du risque
- Liseuses numériques

# Centre d'expertise

- Assister au démarrage
- Formations
- Créer une culture du Big Data
- Support des utilisateurs

# Externalisation

- Atout stratégique
- Le partage crée également de la valeur
- Assurances
- Trafic automobile
- Facebook Atlas

# Démarrage

- Sessions de travail (1-2J)
- Expliquer le big-data
  - Pourquoi plus de données ?
  - Pourquoi des corrélations ?
  - Pourquoi du traitement temps-réel ?

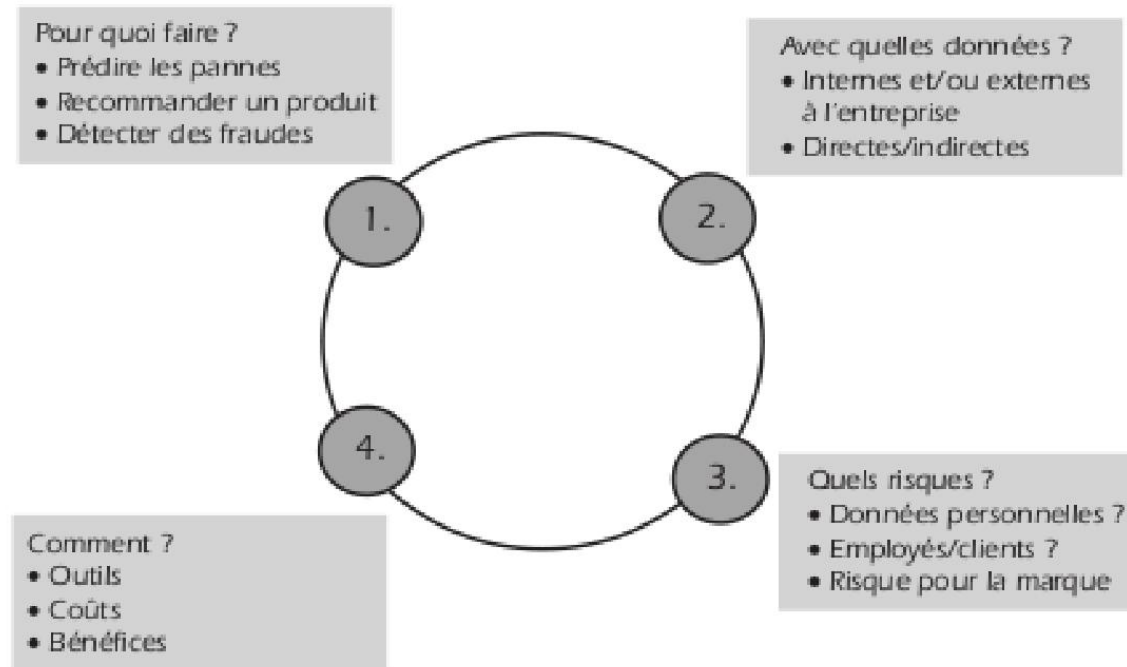
# Démarrage

- Comprendre les sources d'alimentation
  - Internes et externes
  - Lien direct/indirect avec la source
- Comprendre les risques
  - Personnelles (clients, employés...)
  - Sécurité des données
  - Marque, réputation de l'entreprise

# Démarrage

- Idée des coûts – bénéfices (ROI)
  - Exemples
  - Retour d'expériences
  - Outils, externalisation

# Démarrage





# Déploiement - Intégration

- Itératif
- Développement de la culture
- Cohérence et intégration aux systèmes existants
- Prévion de panne de voiture

# Intégration

- Outils marketing intégrés au processus de vente (parcours client)
  - Conflit de canaux
  - Effet d'aubaine