



Introduction BI

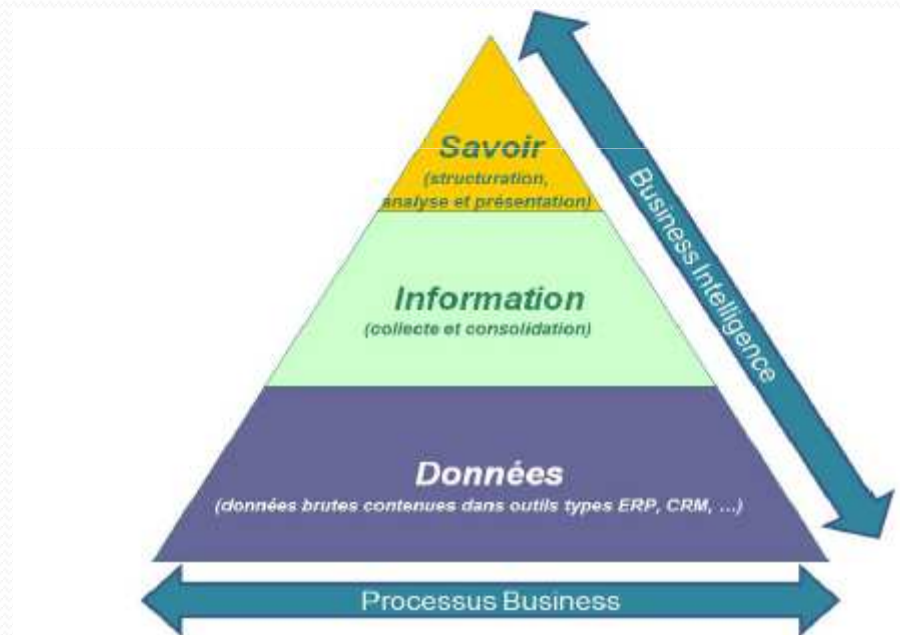
M. BEN SALAH



Que représente le terme BI ?

Business Intelligence (BI) désigne tous les moyens, outils et méthodes qui permettent de collecter, modéliser, consolider et restituer les données d'une entreprise en vue d'offrir une aide à la décision.

- Le processus BI peut se schématiser comme suit :



Les système opérationnel

Rôle :

- automatiser des processus opérationnels (gestion de stocks, gestion de comptes bancaires, ...).
- soutenir le bon fonctionnement de l'activité principale de l'entreprise

Système transactionnel

« Le système transactionnel est généralement une base de données, développée par application, stockant les données courantes d'une organisation, c'est-à-dire qu'il n'y a pas de données d'archives dans les systèmes transactionnels »
(Bédard et al. 1997)

- Généralement les systèmes transactionnels sont implantés selon une structure relationnelle

Système transactionnel

- Les SGBDR sont des systèmes conçus pour de traitement de transactions en ligne ou OLTP (*Online Transaction Processing*).
 - conçus pour assurer de grands débits de transactions, dont les transactions effectuent généralement des modifications aux données opérationnelles de l'organisation, c'est-à-dire les données dont l'organisation a besoin pour gérer ses activités quotidiennes.
- objectifs principaux :
 - insertion, modification, interrogation des données rapidement, suppression des enregistrements.
 - ces opérations doivent pouvoir être effectuées très rapidement et efficacement, par de nombreux utilisateurs simultanément en toute sécurité

Système transactionnel

- Les bases de données transactionnelles sont habituellement normalisées
- A quoi sert la normalisation relationnelle
 - Eviter la redondance
 - Assure l'intégrité des données
 - Simplifie la mise à jour des données

Qui ce que un système décisionnel ?

Un système décisionnel a pour but la transformation des données de l'entreprise en informations pour la compréhension du marché et l'analyse des informations afin de prendre des décisions.

- **Besoin** : décider dans le bon sens
prise de décisions stratégiques et tactiques
- **Pour quoi ?** :
 - anticiper
 - besoin de réactivité
 - reprendre un avantage concurrentiel
 - améliorer les performances
- **Comment ?** :
En répondant aux demandes d'analyse des décideurs

Le processus de prise de décision

Champs d'application des
systèmes décisionnels



Temps de prise d'une décision

- **Comment soutenir la prise de décision ?**

En répondant aux demandes d'analyse des décideurs : se poser un ensemble de questions et trouver des réponses à ces questions en exploitant et explorant les données opérationnelles de l'entreprise : analyse des données



Contexte

- Le soutien de la décision nécessite :
 - une information fiable, précise, actualisée et pertinente
 - des moyens pour effectuer l'analyse
- Données
 - montant total des ventes pour une région données
 - vendeur ayant réalisé le meilleur chiffre du mois
- Information
 - évolution des ventes pour une région au cours des 3 dernière années ?
 - sur quels produits faire des offres promotionnelles ?
 - quelle est la rentabilité d'une activité ?

Contexte

- Comment répondre aux demandes des décideurs?

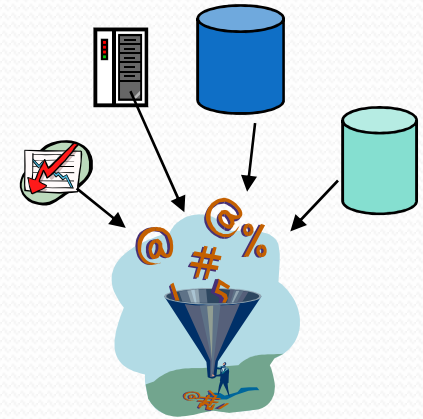
Donner aux décideur une vision synthétique et globale des informations de leurs organisation afin soutenir leur prise de décisions.



- Accéder à toutes les données de l'organisation
- Retrouver une information historique et transversale à l'entreprise
- Rassembler ces données
- Synthétiser/résumer
- Accéder non seulement aux valeurs actuelles présentes dans la base de données mais aussi à des données d'ordre historique afin d'offrir une analyse complète de l'organisation.
- Exploiter ces données en utilisant des outils d'aide à la décision

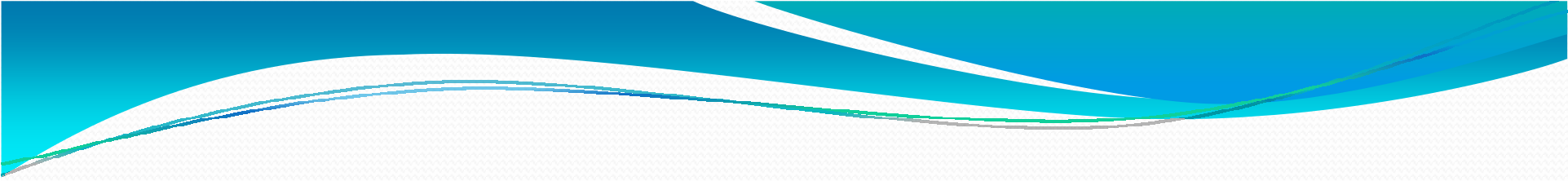
Problématique

- Divers sources de données
 - Divers SGBD (Oracle, SQL Server)
 - Fichiers, web ...
 - Divers services métier (gestion des RH, gestion des commandes...)
- Caractéristiques des données:
 - Distribuées: systèmes éparpillés
 - Hétérogènes: systèmes et structures de données différents
 - Détaillées: organisation des données selon les processus fonctionnels, données surabondantes pour l'analyse
 - Peu (ou pas) adaptées à l'analyse : les requêtes lourdes peuvent bloquer le système transactionnel
 - Volatiles: pas d'historisation systématique
 - Grande masse

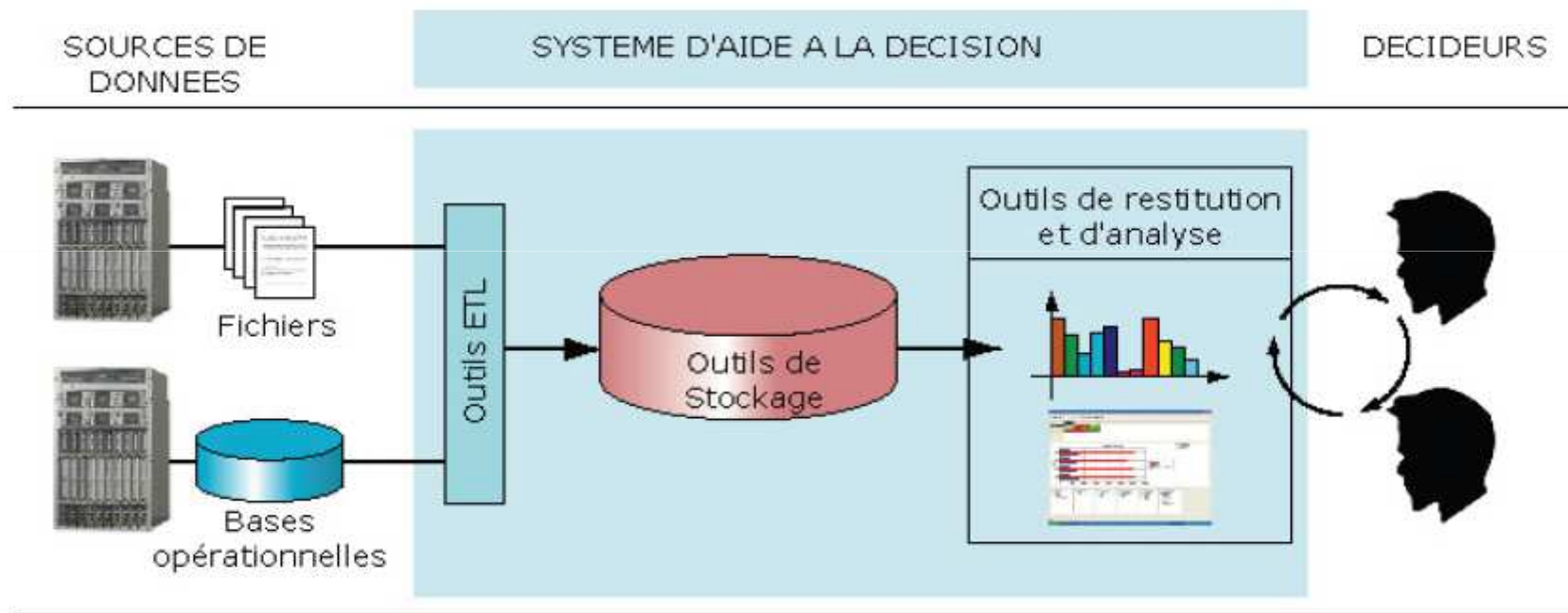


Problématique

- Or pour les décideur on doit retrouver et analyser rapidement les données :
 - Accès rapide et simple à ces données
 - Répondre rapidement aux requêtes souvent lourdes et complexes (agrégation)
 - == > caractère multidimensionnelle : cube
 - Besoin des données historiques

- 
- Comment prendre des décisions sur la base d'informations issues de systèmes hétérogènes n'ayant pas de moyens pour communiquer facilement entre eux

- Solution : Besoin de mettre en place un système décisionnel (dédié aux applications décisionnelles) qui a pour objectif la transformation des données de productions en données stratégique



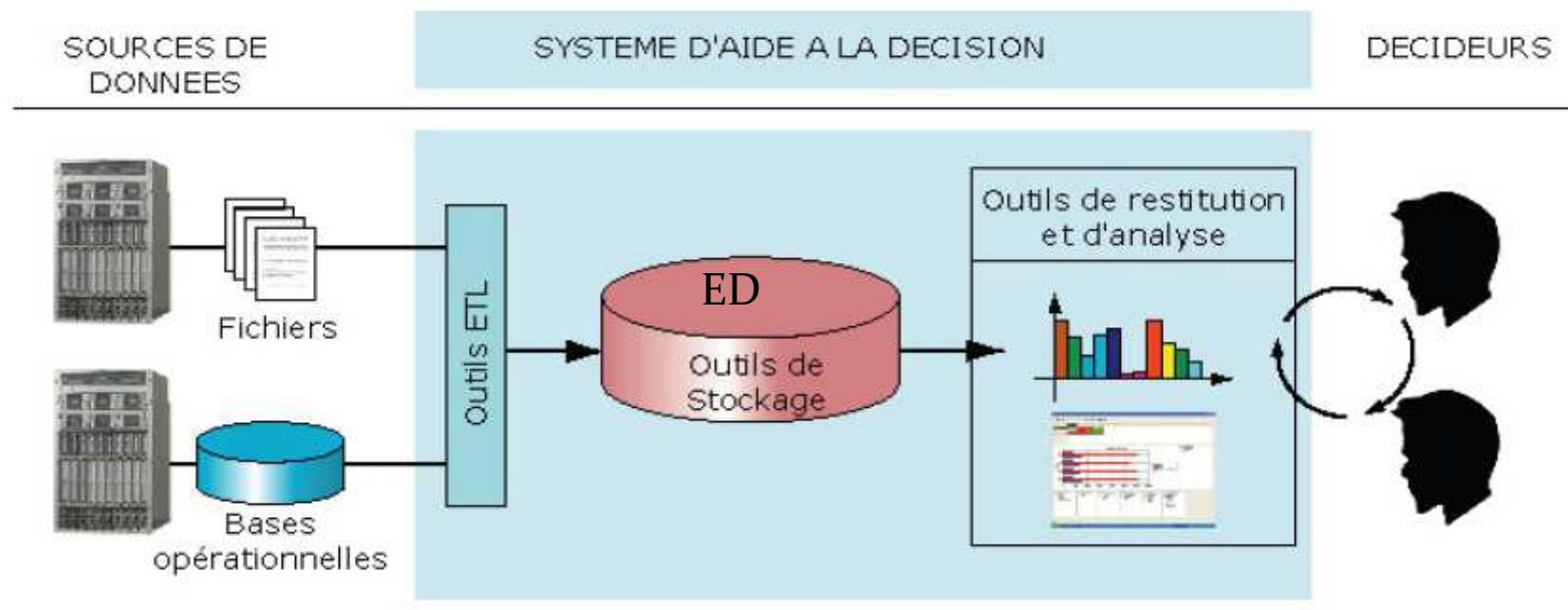
En **Aval** des données
opérationnelles

En **Amont** des
prises de décision

- Pour faciliter ce genre d'analyse, un nouveau concept de base de données a vu le jour :

L'entrepôt de données (*datawarehouse*)

- Besoin aussi de puissants outils d'analyse : Emergence de deux types d'outils d'analyse : les outils de traitement analytique en ligne (OLAP, *Online Analytical Processing*) et d'exploration de données (*datamining*).



Données opérationnelles	Données décisionnelles
Orientées application, détaillées, précises au moment de l'accès.	Orientée activité (thème, sujet), condensées, représente des données historiques.
Mise à jour interactive possible de la part des utilisateurs (insert, delete, update)	Pas de mise à jour interactive, données utilisées pour de l'analyse (reporting), donc accès en lecture principalement.
Accédées de façon unitaire par une personne à la fois (transactions unitaires).	Utilisées par l'ensemble des analystes, gérées par sous-ensemble (en lecture seule, pas de verrou sur les données).
Haute disponibilité en continu. Données critiques qui ne peuvent pas être perdues (sauvegardes/ réplication, ...). L'accès au données doit être rapide.	Exigence différente, haute disponibilité ponctuelle. Les données récentes peuvent être rechargées à partir du SIO en cas de perte des données. L'accès aux données peut être lent (car requêtes complexes et volumétrie élevée).
Uniques (pas de redondance en théorie, schéma relationnel normalisé 3NF)	Peuvent être redondantes (pas de normalisation 3NF), modèle multi-dimensionnel.
Données dédiées au métier de l'entreprise, utilisées potentiellement par tous les salariés.	Données dédiées aux prises de décision, utilisées par les analystes et les décideurs.
Petite quantité de données utilisées par un traitement (en général), requêtes « simples »	Grande quantité de données utilisée par les traitements, requêtes complexes (agrégations, calculs)
Réalisation des opérations au jour le jour	Cycles de vie différents, données collectées sur le long terme.
Volume de données assez faible	Volume de données très grand, car données historisées et redondance.

Problématique à l'analyse dans les systèmes transactionnels

- Les bases de données transactionnelles :
 - mode OLTP
 - normalisation
- Cependant, plus le degré de la normalisation augmente plus l'analyse des données devient complexe :
 - nombre élevé de tables donc nombre élevé de jointures nécessaires entre les tables
 - nécessité de mettre en place des requêtes complexes
 - == > temps de traitement long
 - == > dégradation des performance

Problématique à l'analyse dans les systèmes transactionnels

- Or le processus de décision nécessite :
 - Des données agrégées sur l'ensemble de l'organisation (provenant des différentes sources (BD)
 - Données historiques
 - Réponses rapides (requête complexe)

Les systèmes OLTP sont mal adaptés à l'analyse de données.
L'usage de ces systèmes dans le cadre de la prise de décision n'est pas vraiment une solution aisée

== > Le défi posé : transformer les données opérationnels de l'organisation en une source de connaissance, de manière à présenter à l'utilisateur une seule vue intégrée et consolidée des données de l'organisation.

L'entrepôt de données est considéré comme la solution la plus satisfaisant et la pièce fondamentale d'un système capable d'apporter un soutien à la prise de décision tout en recevant des données de plusieurs sources de données opérationnelles.