

Informatique Décisionnelle (BI)

ABDELTIF EL BYED

Professeur à l'université Hassan II – Faculté
des sciences Ain Chok (FSAC)

Aelbyed@gmail.com /
Abdeltif.elbyed@univh2c.ac.ma

Site: <https://sites.google.com/site/aelbyed>

Plan général du Module

Partie1: Informatique Décisionnelle

1. Introduction à l'Informatique Décisionnelle (BI)
2. Entrepôt de données (DW)
3. OLAP: On-Line Analytical Processing
4. Les système OLAP: ROLAP, MOLAP & HOLAP

Partie 2: IBM Cognos Report Studio (En Anglais)

1. Introduction to the Reporting Application
2. Create List Reports
3. Focus Reports Using Filters
4. Create Crosstab Reports
5. Present Data Graphically
6. Focus Reports Using Prompts

Plan général du Module

Partie 1: Informatique Décisionnelle

1. Introduction à l'Informatique Décisionnelle (BI)
2. Entrepôt de données (DW)
3. OLAP: On-Line Analytical Processing
4. Les système OLAP: ROLAP, MOLAP & HOLAP

Partie 2: IBM Cognos Report Studio (En Anglais)

1. Introduction to the Reporting Application
2. Create List Reports
3. Focus Reports Using Filters
4. Create Crosstab Reports
5. Present Data Graphically
6. Focus Reports Using Prompts

Partie 3: Data Mining

1. Objectifs de la fouille de données
2. Les algorithmes de data mining
3. Le DW et le data mining
4. IA et l'apprentissage

A.EL BYED: Introduction au BI

3

CH1

Informatique Décisionnelle Introduction

Abdeltif EL BYED

Professeur à l'université Hassan II

Aelbyed@gmail.com

Abdeltif.elbyed@univh2c.ac.ma

Plan

1. Emergence de l'Informatique Décisionnelle (ID) / Business Intelligence (BI)
2. Des SI opérationnels aux SI décisionnels
3. Les outils BI
4. Un exemple du BI dans l'entreprise : la relation client (CRM)

1. Emergence de l'Informatique Décisionnelle

1. Contextes économique et informationnel actuel
2. Pilotage de l'entreprise et le BI
3. Historique de le BI
4. Problématique de le BI
5. Domaines d'application de le BI

Contextes économique et informationnel

Contexte économique :

- Mondialisation de l'économie, ouverture de nouveaux marchés
- Concurrence toujours plus accrue
- Besoin d'informations pour prises de décisions de plus en plus rapides

Contexte informationnel :

- Décentralisation des données vers les utilisateurs
- Difficulté d'accès à l'information qui est en trop grande quantité
- Un enjeu stratégique d'entreprise
- Les informations, une source de revenu et de compétitivité

Contexte informatique :

- Puissance de calcul croissante
- Capacité de stockage croissante
- Bases de données de plus en plus importantes
- SGBD de plus en plus performants (parallélisme, ...)
- Ouverture sur le Web, ...

Pilotage de l'entreprise et BI

- Le pilotage d'une entreprise dépend de ses objectifs stratégiques
- Ce pilotage doit prendre en considération :
 - Une organisation de plus en plus orientée clients
 - Des cycles conception/fabrication de plus en plus courts
 - De nouveaux canaux de distribution notamment les ventes en ligne sur le Web
 - L'exigence d'internationalisation
 - ...
- Par conséquent, l'entreprise se doit :
 - Anticiper les besoins des clients,
 - Contrôler l'intégrité et la qualité des flux de gestion
 - Évaluer la performance des différentes entités et composants

Pilotage de l'entreprise: Outils nécessaires

Outils d'analyse :

- **Constituer** et **mettre à jour** à partir de diverses sources des réservoirs de grande quantités de données historiques et multidimensionnelles, ...
- **Extraire** selon divers critères des sous ensembles de données de tels réservoirs
- **Analyser** ces données selon différents axes (OLAP), d'identifier des tendances, des corrélations, faire de la prévision (Data Mining)

Outils de veille stratégique, de recherche d'information (RI) :

- Issus de « l'intelligence économique » (Competitive Intelligence), ces outils ont pour finalité de recueillir des informations sur le marché et sur la concurrence
- **Collecter** une importante quantité de données via le Web, les **filtrer** et **extraire** les informations pertinentes (Web Mining) pour les analyser ensuite

=> **Outils relevant de l'Informatique Décisionnelle (ID) ou Business Intelligence (BI),**

L'Informatique Décisionnelle (ID) /Business Intelligence (BI)

- L'Informatique Décisionnelle (ID), en anglais Business Intelligence (BI), est l'informatique à l'usage des décideurs et des dirigeants des entreprises
 - Les systèmes de ID/BI sont utilisés par les décideurs
- BI permet:
 - D'obtenir une connaissance approfondie de l'entreprise
 - De définir et de soutenir leurs stratégies d'affaires
- Par exemple :
 - Acquérir un avantage concurrentiel,
 - Améliorer la performance de l'entreprise,
 - Répondre plus rapidement aux changements,
 - Augmenter la rentabilité,
 - Création de valeur ajoutée de l'entreprise.

Historique du BI (1/2)

Années 70-90 : débuts de l'Informatique Décisionnelle :

- Développement d'outils d'édition de rapports, de statistiques, exploitant les BD opérationnelles
- Développement de petits systèmes d'aide à la décision à base de tableurs (simulation budgétaire, ...)
- Développement de systèmes experts (IA) systèmes à base de règles, conçus par extraction de la connaissance d'un ou plusieurs experts
 - intérêts et résultats limités
- Développement de systèmes spécifiques d'aide à la décision :
 - basés sur des techniques de Recherche Opérationnelle (RO), la simulation, l'optimisation, ...

=> **Systèmes en général mal intégrés au SI opérationnel**

Historique du BI (2/2)

Années 90 - 2000 : essor de l'informatique décisionnelle:

- Technologie informatique permettant le développement d'Entrepôts de Données (Data Warehouse)
- Nouveaux algorithmes :
 - souvent issus des statistiques et de l'IA, permettant l'extraction des informations à partir de données brutes
 - regroupés dans des logiciels de Fouille de données (Data Mining)
- Données issues du Web :
 - Recherche d'Information (RI) et Fouille de données sur le Web (« Web Mining »)

BI d'aujourd'hui

- L'ID/BI est un ensemble de processus, d'outils et de technologies pour transformer les données de l'entreprise en informations exactes et à jour pour soutenir des processus décisionnels:
 - Data Warehousing (DW)
 - On-Line Analytical Processing (OLAP)
 - Data Mining (DM) et visualisation de données (VIS)
 - Analyse des décisions (what-if)
 - Customer Relationship Management (CRM)
- L'ID/BI peut être considérée comme l'opposé de l'Intelligence Artificielle (IA) :
 - Les systèmes d'IA prennent des décisions pour les utilisateurs
 - Les systèmes de BI permettent aux utilisateurs de prendre les bonnes décisions sur la base des données disponibles
 - Mais de nombreuses techniques de BI ont leurs racines dans l'IA.

Exemples de requêtes BI

- Q1: pour la **date** du 11 Octobre 2018, trouver les 5 **produits** les plus vendus pour chaque **sous-catégorie** de produits qui représente plus que 20% des **ventes** dans sa **catégorie** de produits
- Q2: En **date** du 15 Décembre 2018, déterminer la priorité d'expédition (**shipping priority**) et de **revenu** brut potentiel (potential gross revenue) des **commandes** qui ont les 10 plus **grandes recettes** brutes (largest gross revenues) parmi les commandes qui n'avaient pas encore été **expédiées**.
 - On considère uniquement les commandes du segment du **marché du livre**

BI en pleine croissance (1/2)

- Selon Meta Group : DW seul = 15 billion\$ en 2000
- Selon Palo Alto Management Group : BI = 113 billion\$ en 2002
- Le Web rend le BI encore plus nécessaire :
 - Les clients ne sont pas «physiquement» dans le magasin
 - Les clients peuvent changer à d'autres magasins plus facilement
 - Comment connaître ses clients :
 - Analyser les « Web log » pour analyser le comportement des clients sur le site
 - Combiner ces données Web avec les données traditionnelles des clients
- « Internet sans fil » ajoute à cela :
 - Les clients sont toujours "en ligne"
 - La position de la clientèle est bien connue
 - Combiner la position et la connaissance sur le client => très utile

BI en pleine croissance (2/2)

- Selon une étude du Gartner Group en 2009 :
 - Les entreprises rechercheront de plus en plus des cadres qui auront en charge l'ID/BI et la gestion de la performance pour contribuer à la transformation et l'amélioration de leurs affaires
- En 2009, la prise de décision collaborative va s'imposer et nécessiter la combinaison d'outils de BI et de logiciels sociaux et d'autres..
- En 2010, 20% des entreprises auront une application analytique spécifique de BI
- Jusqu'en 2012, plus de 35% des 5000 principales entreprises mondiales ont régulièrement échoué dans la prise de décisions cruciales concernant leurs marchés et leurs affaires
- « Aujourd'hui, il est difficile de trouver une entreprise prospère qui n'a pas d'effet de levier technologie BI pour leurs affaires »

Problématique de la BD opérationnelle

- Modèles de BD opérationnelles complexes et inutilisables :
 - souvent difficiles à comprendre
 - ne concerne pas un objectif unique d'affaire
- Problématique des données opérationnelles :
 - identiques dans différentes BD
 - même concept souvent défini différemment
 - adaptées pour les systèmes opérationnels (comptabilité, facturation, ...),
 - pas pour l'analyse des fonctions d'affaires
 - Mauvaise qualité de données:
 - données manquantes, données imprécises, ...
 - Volatiles :
 - elles sont supprimées périodiquement dans les systèmes opérationnels
 - la modification des données au fil du temps - aucune information historique

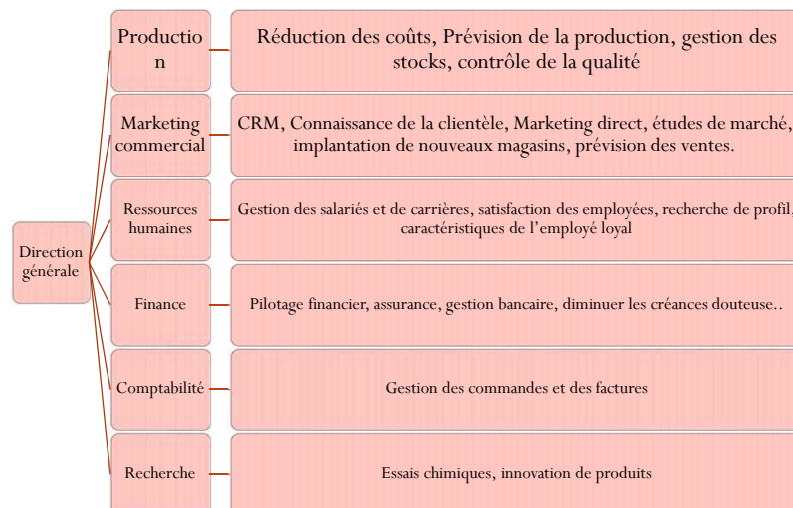
Terminologies

Termes anglais	Termes français
Business Intelligence (BI)	Informatique Décisionnelle (ID)
Decision Support Systems (DSS)	Systèmes d'aide à la décision (SAD)
Competitive Intelligence (CI)	Intelligence Economique (IE)
Data Warehouse (DW)	Entrepôt de données (ED)
On-Line Analytical Processing (OLAP)	Analyse en ligne de données
Knowledge Discovery in databases (KDD)	Extraction de Connaissances dans les bases de données
Data Mining (DM)	fouille de données, orpaillage
Customer Relationship Management (CRM)	Gestion de la Relation Client

BI & les domaines d'application

- Domaines concernés :
 - commercial, financier, transport, télécommunications, santé, services, ...
- Applications :
 - gestion de la relation client
 - gestion de commandes, de stocks
 - prévisions de ventes
 - définition de profil utilisateur
 - analyse de transactions bancaires
 - détection de fraudes
 - ...

BI dans l'entreprise



2. Des SI opérationnels au SI décisionnels

1. Systèmes d'information opérationnels : traitement transactionnel en ligne des données (OLTP)
2. Nouveaux besoins : traitement analytique en ligne des données (OLAP)
3. Des SI Opérationnels aux SI Décisionnels

SI opérationnels: OLTP

- Supportent en général une ou plusieurs grandes fonctions de l'entreprise
 - Ex, production, marketing, commercial, ressources humaines, finance, comptabilité, recherche, ...
- Parfois intégrés dans un ERP,
 - Ils s'appuient sur des SGBD traditionnels (Oracle, DB2, ...) pour gérer des BD « opérationnelles »
 - Permettent des processus de traitement en ligne des données – OLTP (On line Transactionnal Processing) :
 - Interactifs, Concurrents, Nombreux, Répétitifs, Structurés, Simples
- Ces processus OLTP concernent :
 - La mise à jour de données
 - Un nombre restreint d'enregistrements
 - Des données précises et à jour

Par exemple : un supermarché ENREGISTRE ses ventes

Nouveaux besoins (1/2)

- Pour prendre de « bonnes décisions », on doit pouvoir **accéder** en temps réel aux données de l'entreprise, **traiter** et **extraire** l'information pertinente de ces données, par exemple pour savoir :
 - Quels sont les résultats des ventes par gamme de produit et par région pour l'année dernière?
 - Quelle est l'évolution des chiffres d'affaires par type de magasin et par période ?
 - Comment qualifier les acheteurs de mon produit X ?

Nouvelles attentes des SI :

- Considérer des quantités de données historiques de plus en plus importantes (Tera, Penta octets), organisées selon différentes dimensions (temps, espace géographique, gammes de produit, ...)
- Passer du traitement en ligne des données (OLTP) à l'analyse en ligne de ces données (On Line Analysis Processing - OLAP)
 - selon différentes dimensions pour analyser ces données et construire des indicateurs indispensables au pilotage de l'entreprise

Nouveaux besoins (2/2)

L'informatique opérationnelle ne peut satisfaire ces besoins :

- BD opérationnelles trop complexes pour être appréhendées facilement par un décideur.
- SI opérationnel ne peut être interrompu pour répondre à des questions nécessitant des calculs importants.
- Processus d'entreposage des données mal adapté
- Analyse en ligne des données très limitée

=> Recours à la BI pour l'élaboration de SI Décisionnels

3. Les outils BI

1. Entrepôt de données (DW)
2. Analyse en ligne (OLAP)
3. Fouille de données (DM)

Technologies BI

Entrepôt de données (Data Warehouse) :

- Il récolte, stocke et gère efficacement des gros volumes de données pour la prise de décision
- Il assure un regroupement homogène et exploitable de données hétérogènes, très nombreuses et distribuées

Analyse en ligne des données OLAP (On Line Analytical processing) :

- Données organisées selon plusieurs axes d'analyse et selon différents niveaux de détail
- Traitements permettant interactivement de changer de points de vue, de niveau de détail, et effectuer les opérations OLTP classiques

Fouille de données (Data Mining):

- Extraction automatique de connaissances (propriétés cachées) dans de grands volumes de données :
 - Connaissances valides, nouvelles, compréhensibles, pertinentes et implicites

Entrepôt de données



Définition de Bill Inmon (1992) :

- « une collection de données **thématiques, intégrées, non volatiles et historisées**, organisées pour le support d'un processus d'aide à la décision»
- Données :
 - **thématiques** : données pertinentes pour un sujet ou thème et nécessaire aux besoins d'analyse
 - **intégrées** : données résultant de l'intégration de données provenant de différentes sources pouvant être hétérogènes
 - **historisées** : données représentent l'activité d'une entreprise durant une certaine période (plusieurs années)
 - **non-volatiles** : données essentiellement utilisées en interrogation (consultation) et ne peuvent pas être modifiées

Analyse en ligne des données OLAP

- Caractéristiques des processus d'analyse en ligne OLAP :
 - Interactifs
 - Concurrents
 - Peu nombreux
 - Non prévisibles
 - Complexes
 - Les processus OLAP concernent :
 - L'exploitation des données
 - Un nombre d'enregistrement très important
 - Des données consolidées et synthétiques
 - Exploration et analyse de données historiques
- Par exemple : un supermarché ANALYSE l'ensemble de ses ventes**

OLTP versus OLAP

	Caractéristiques	OLTP	OLAP
Conception	Conception	Transaction	Analyse
	Orientation	Entité-Relation	Star / snowflake
Données	Granularité	Détail	Résumées, agrégées
	Nature	Relationnelle	Multidimensionnelle
	Actualisation	Actualisées, mises à jour	Historisées recalculées
	Taille	100 Mo/Go	100 Go/To
Traitements	Unité de travail	Transaction simple	Requête complexe
	Accès	Lecture/écriture	Lecture
	Nb de tuples accédés	Dizaines	Millions
	Métrique	Débit de transactions	Temps de réponse
Utilisateurs	Utilisateur	Agent opérationnel	Analyste/décideur
	Nombre d'utilisateurs	Milliers	Centaines

Fouille de données (Data Mining)

- Recherche de la connaissance sous forme de modèles de comportement caché dans les données :
 - Extraction de connaissance à partir de données (Knowledge Discovery in Database - KDD)
- Domaine jeune à l'intersection de l'Intelligence artificielle (IA) , les statistique, l'analyse de données, et les BD
- Techniques de fouille traditionnelles :
 - issues des statistiques et de l'analyse de données
- Nouvelles techniques de fouille :
 - l'induction d'arbre de décision, les algorithmes génétiques, les réseaux de neurones, ...
- Tendance à une intégration croissante de ces techniques de fouille dans les entrepôts de données (notamment dans l'offre commerciale)

Architecture multi-niveaux

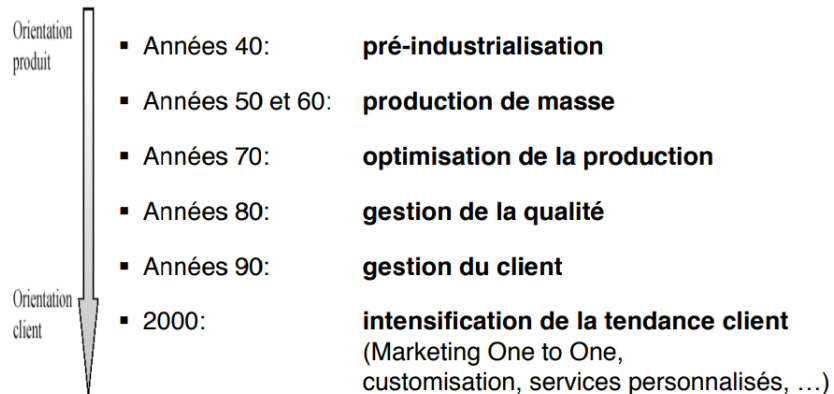
- *Figure chaine décisionnelle à ajouter*

4. Un exemple de la BI dans l'entreprise : la relation client

1. Contexte économique général
2. La Gestion de la Relation Client
3. BI pour la Gestion de la Relation Client

Contexte économique général (1/2)

Evolution historique générale :



33

A.EL BYED: Introduction au BI

Contexte économique général (2/2)

Dans un climat de concurrence mondiale tendu :

- Conquérir un nouveau client coûte 5 fois plus cher que de fidéliser un client existant
- 5% d'amélioration de la fidélité des clients entraîne une augmentation des profits de 10 à 15%
- Tous les clients ne sont pas égaux :
 - 30% des clients génèrent 70% du CA
- Le client attend un service personnalisé, sur-mesure
- La personnalisation est une source de profit

34

A.EL BYED: Introduction au BI

Gestion de la relation client (CRM)

- Gestion de la relation client (Customer Relationship Management – CRM) :
 - Démarche émergente dans les domaines du marketing et du commercial
- Afin de répondre aux questions :
 - Quels sont les besoins et les attentes des clients? Comment y répondre?
 - Quels sont les clients prêts à acheter de nouveaux produits?
 - Quels sont les clients les plus profitables, fidèles et pourquoi?
 - Quels sont les clients mécontents, et pourquoi?
- Besoin de mieux connaître et comprendre ses clients pour :
 - Réduire les coûts (clients infidèles)
 - Comment augmenter les profits ?
 - Comment fidéliser une clientèle?
 - Comment identifier les nouvelles opportunités?
 - Les concurrents s'attaquent à certains de nos clients?

=> Une réponse par l'informatique décisionnelle

Gestion de la relation client

Définition du CRM :

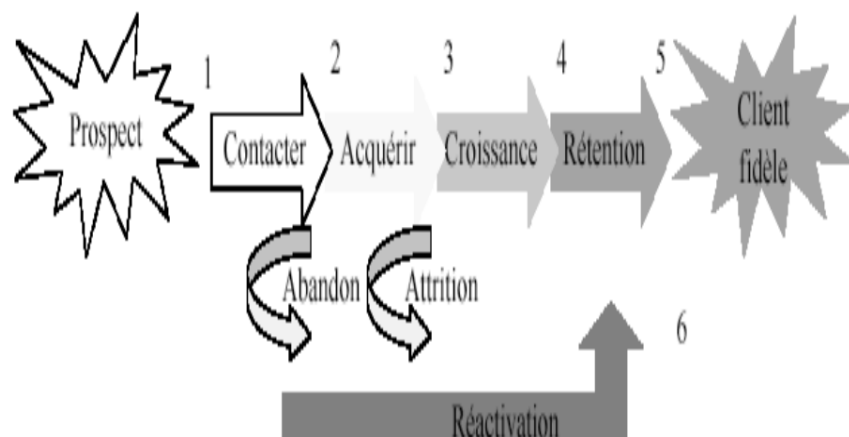
- Capacité à identifier, à acquérir et à fidéliser les meilleurs clients
 - dans le but d'augmenter le chiffre d'affaires et les bénéfices.
- Capacité à bâtir une relation profitable sur le long terme avec les meilleurs clients
 - en capitalisant sur l'ensemble des points de contacts
- Principales dimensions de la Gestion de la Relation Client :
 - stratégique,
 - temporelle et
 - géographique

Gestion de la relation client

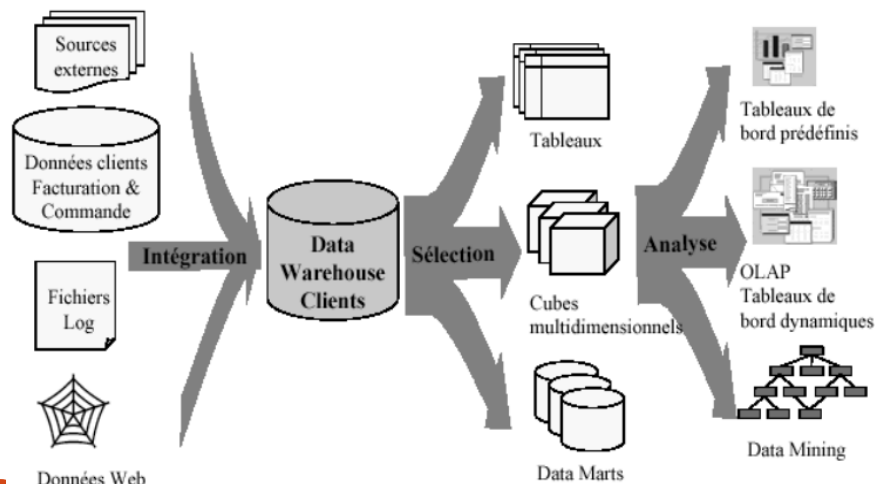
Enjeux du CRM :

- Connaître ses clients rentables et se concentrer sur les clients à fort potentiel
- Augmenter la fidélité des clients
- Comprendre le comportement de ses clients
- Augmenter le panier moyen
- Diminuer les coûts des programmes marketing avec un meilleur ciblage des meilleurs clients
- Adapter les produits au marché
- Aider les commerciaux à vendre
- Rentabiliser la BD Clients
- Mesurer les résultats des actions CRM
- ...

Gestion de la relation client



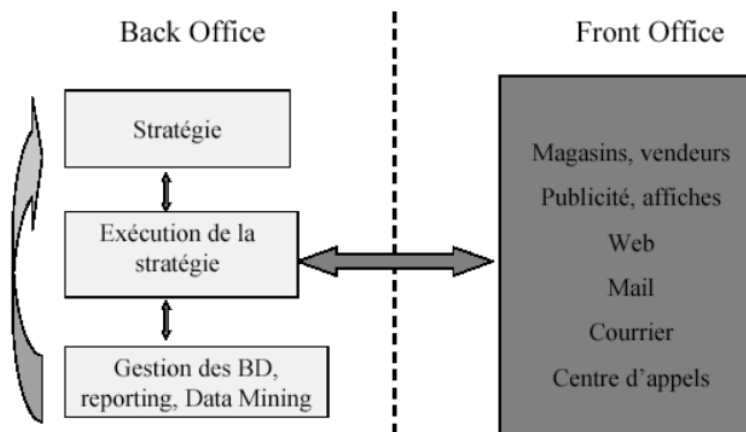
BI et la Gestion de la Relation Client



39

A.EL BYED: Introduction au BI

BI et la Gestion de la Relation Client



40

A.EL BYED: Introduction au BI

BI et la Gestion de la Relation Client



Gestion de la Relation Client et Web

- Coût du contact Client sur le Web très faible :
 - Web = un point de contact important
 - En 2001, 1/3 des contacts clients provenaient d'Internet
 - En 2003, Internet et les applications de commerce électronique génèrent 60% du volume de contact des centres d'appels
- Intégration du Web dans le CRM :
 - e-CRM
 - Web Mining