

Semestre : 6 Parcours ISIL

Unité d'enseignement Méthodologique

Matière : Business intelligence (BI)

Crédits : 2

Coefficient : 1

- **Objectifs de l'enseignement** : l'objectif de cette matière est d'apporter des connaissances à l'étudiant pour : Comprendre le business intelligence
- Déterminer les différentes étapes dans la mise en place d'un projet B.I
- Examiner le marché de la BI et les outils leaders

Connaissances préalables recommandées : maîtrise de base de données relationnelles

Contenu de la matière :

- I. BI vs ERP
- II. Architecture classique d'une solution BI
- III. ETL et Data
- IV. Modèles multi-dimensionnels et OLAP
- V. Data Warehouse Modeling
- VI. Data Warehouse architecture
- VII. BI et Data Mining
- VIII. KPI, Dashboard, Scorecards et Cockpit
- IX. Real-time BI

Mode d'évaluation : examen final (100%)

Références:

- Alphonse Carlier *Business Intelligence et Management*, AfnorÉditions, 2013, (ISBN 978-2-12-465429-1); EAN : 9782124654291
- Jean-Marie Gouarné, *Le Projet décisionnel - Enjeux, Modèles, Architectures du Data Warehouse* [archive], Eyrolles, 1997, (ISBN 978-2-212-05012-7)
- Alain Garnier, *L'Information non structurée dans l'entreprise - Usages et Outils*, Hermes - Lavoisier, 2007, (ISBN 978-2-7462-1605-1)
- R. Kimball, L. Reeves, M. Ross, W. Thornthwaite, *Le Data Warehouse : Guide de conduite de projet*, Eyrolles, 2005, (ISBN 978-2-212-11600-7)
- Alain Fernandez, *Les Nouveaux Tableaux de bord des managers, Le Projet Business Intelligence clés en main*, Eyrolles, 6^e édition, 2013. (ISBN 978-2-212-55647-6) présentation éditeur [archive]
- Roland et Patrick Mosimann, Meg Dussault, *The Performance Manager Faire de la performance le quotidien de chacun* [archive], CognosPress, 2007, (ISBN 978-0-9730124-4-6)
- James Taylor, *Decision Management System* [archive], IBM Press, Pearson Education

I. BI vs ERP

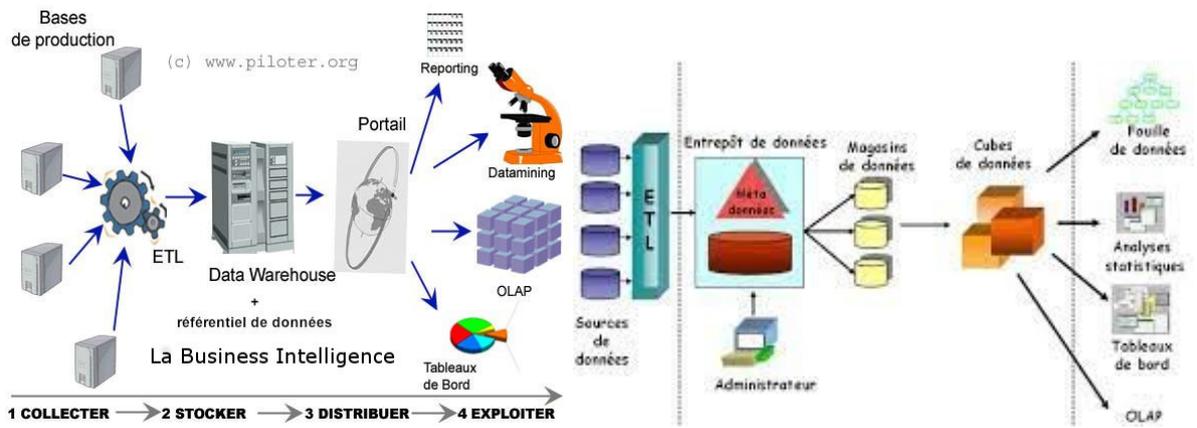
1. Définition ERP. un « **Enterprise Resource Planning (ERP)** » c'est-à-dire, en français, un Progiciel de Gestion Intégrée (PGI). Cette catégorie de logiciel a pour vocation de couvrir tout ou partie des domaines d'une entreprise au sein d'une solution intégrée unique. L'ERP est un logiciel de gestion qui permet à l'entreprise d'intégrer différentes fonctionnalités telles que :

- La gestion comptable et financière (ERP de gestion comptable, ERP de gestion de facturation),
- La gestion des stocks (logiciel ERP gestion des stocks),
- La gestion des ressources humaines,
- La gestion des fournisseurs (ERP fournisseurs grande distribution) ainsi que,
- La gestion de la vente,
- La gestion de la distribution ou encore,
- La gestion de l'e-commerce (ERP commerce, ERP de commerce détails spécialisé).

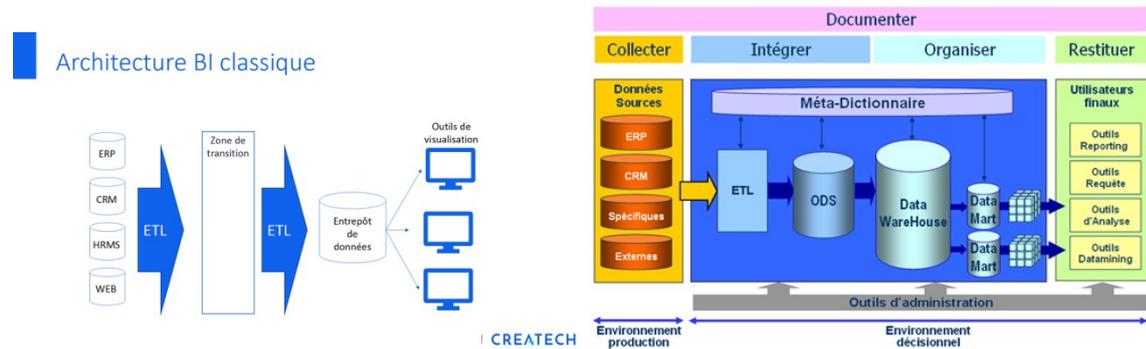
2. Définition BI. La **Business Intelligence (informatique décisionnelle)** a le but est de permettre aux managers de prendre des décisions opérationnelles et stratégiques, sur la base d'informations **structurées, visuelles et pertinentes**. Elle désigne les moyens, les outils et les méthodes qui permettent de collecter, consolider, modéliser et restituer les données, matérielles ou immatérielles, d'une entreprise en vue d'offrir une aide à la décision et de permettre aux responsables de la stratégie d'entreprise d'avoir une vue d'ensemble de l'activité traitée. Ce type d'application utilise en règle générale un datawarehouse (ou entrepôt de données) pour stocker des données transverses provenant de plusieurs sources hétérogènes et fait appel à des traitements lourds de type "batch" pour la collecte de ces informations. L'informatique décisionnelle s'insère dans l'architecture plus large d'un système d'information.

II. Architecture classique d'une solution BI

- 1- *Collecter* : Les outils d'ETL (Extract Transform and Load) nettoier et consolider les données Extraire les données des systèmes de production et les adapter à un usage décisionnel.
- 2- *Stocker* : Le Data warehouse et le datamart les données structurées et traitées afin qu'elles soient disponibles pour un usage décisionnel.
- 3- *Distribuer* Ou plutôt faciliter l'accessibilité des informations selon les fonctions et les types d'utilisation.
- 4- *Exploiter* ou comment assister du mieux possible l'utilisateur afin qu'il puisse extraire la substance de l'information des données stockées à cet usage.



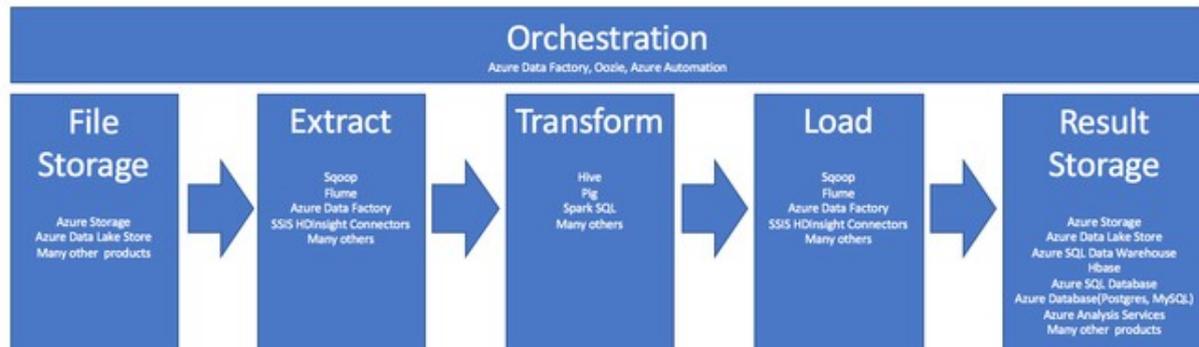
Exemple de source de données hétérogènes



Exemple : CV, dépenses missionnaires.

III. ETL et Data

Extract-transform-load ou ETL, ou extracto-chargeur, (ou parfois : datapumping). Il s'agit d'une technologie informatique intergicielle (comprendre *middleware*) permettant d'effectuer des synchronisations massives d'information d'une source de données (le plus souvent une base de données) vers une autre. Selon le contexte, on est amené à exploiter différentes fonctions, souvent combinées entre elles : « extraction », « transformation », « constitution » ou « conversion », « alimentation ».



Exemples d'Etapes :

- Extraction :
 - Compiler des données provenant de sources pertinentes.
 - Organisez les données pour les rendre cohérentes.
- Transformation :
 - Convertir les données en fonction des besoins de l'entreprise.
 - Reformatez les données converties dans un format standard pour assurer la compatibilité.
 - Nettoyer les données non pertinentes des jeux de données.
 - Trier et filtrer les données.
 - Effacer les informations en double.
 - Traduisez si nécessaire.
- Chargement
 - Chargez des ensembles de données bien transformés via un chargement en bloc.

- Chargez des ensembles de données douteux via des insertions SQL par exemple.

IV. Architecture technique du Data Warehouse

1. Définition du Data Warehouse : ou entrepôt de données, est une base de données dédiée au stockage de l'ensemble des données utilisées dans le cadre de la prise de décision et de l'analyse décisionnelle.

2. Définition du Datamart : est une version plus réduite du data warehouse. Le data mart est orienté sujet ou thème

3. les caractéristiques du Data Warehouse

Le Data Warehouse n'est pas une simple copie des données de production. Le data warehouse est organisé et structuré.

Orienté sujet : Au coeur du Data warehouse, les données sont organisées par thème. Les données propres à un thème, les ventes par exemple, seront rapatriées des différentes bases OLTP de production et regroupées.

Intégré : Les données proviennent de sources hétérogènes utilisant chacune un type de format. Elles sont intégrées avant d'être proposées à utilisation

Non volatile : Les données ne disparaissent pas et ne changent pas au fil des traitements, au fil du temps (Read-Only).

Historisé : Les données non volatiles sont aussi horodatées. On peut ainsi visualiser l'évolution dans le temps d'une valeur donnée. Le degré de détail de l'archivage est bien entendu relatif à la nature des données. Toutes les données ne méritent pas d'être archivées.

3. Les architectures :

3 architectures technologiques typiques et classiques pour le stockage de grandes quantités de données pour des fins décisionnelles : SMP, MMP et Cluster.

SMP (Symetric Multi-Processing)

Principe : le modèle d'architecture de type "SMP" est fondé sur l'exploitation de plusieurs processeurs identiques oeuvrant en parallèle et partageant une mémoire commune. *Inconvénients :* la mémoire est unique, la synchronisation de l'accès à la mémoire par les différents processeurs constitue le principal inconvénient de ce type d'architecture.

MMP (Massively Parallel Processing)

- *Principe :* le modèle d'architecture de type "MPP" est fondé sur l'exploitation d'un nombre important de processeurs. Chaque processeur dispose de sa propre mémoire.
- *Inconvénient :* il nécessite des développements spécifiques. Les traitements doivent être prévus dès la conception pour une exécution sur ce type d'architecture.

Cluster

- *Principe :* avec l'architecture de type "Cluster", les ordinateurs sont organisés en "grappes". Ils sont interconnectés par des liaisons rapides Ethernet. Sur le plan du principe, le fonctionnement est assez proche de l'architecture MMP.
- *Inconvénient :* le programme à exécuter doit impérativement être développé pour ce type d'architecture.

V. BI et datamining