

SYSTÈMES D'INFORMATION DÉCISIONNELS

BUSINESS INTELLIGENCE- BI

Département d'informatique

Module: Systèmes d'information décisionnels

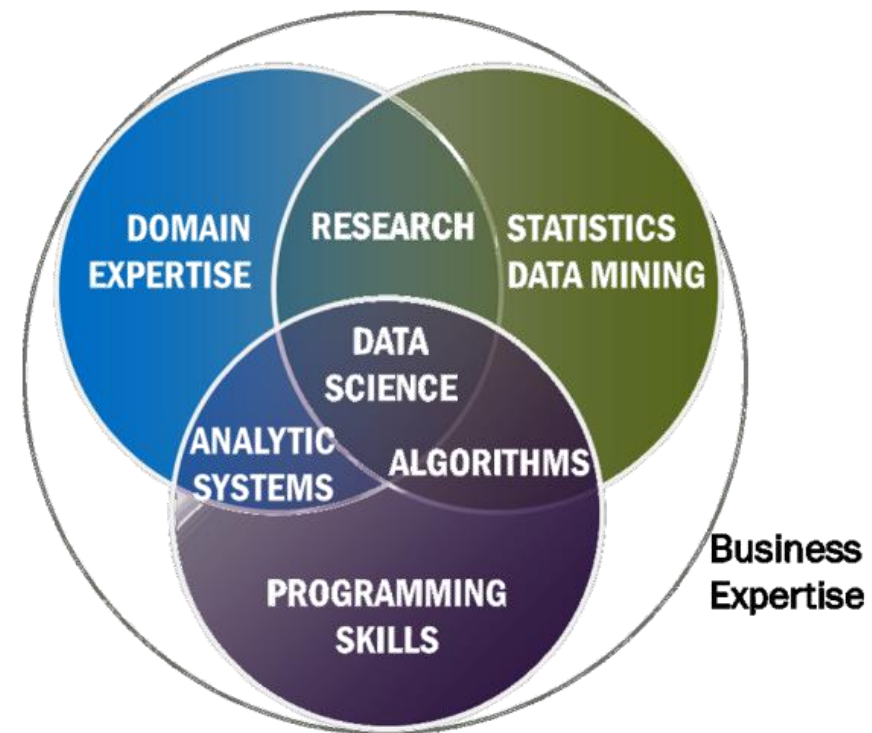
Chargé du module: Mokeddem, S

Email: sidahmed.mokadem@univ-mosta.dz



Année universitaire: 2018/2019

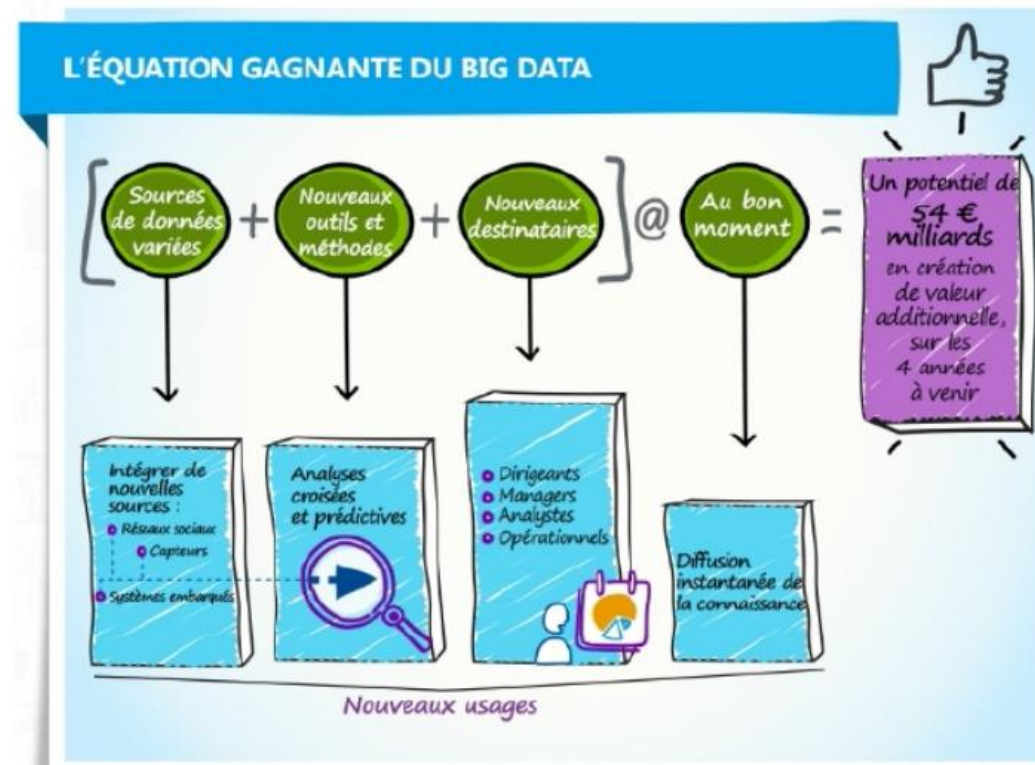
CONTEXTE



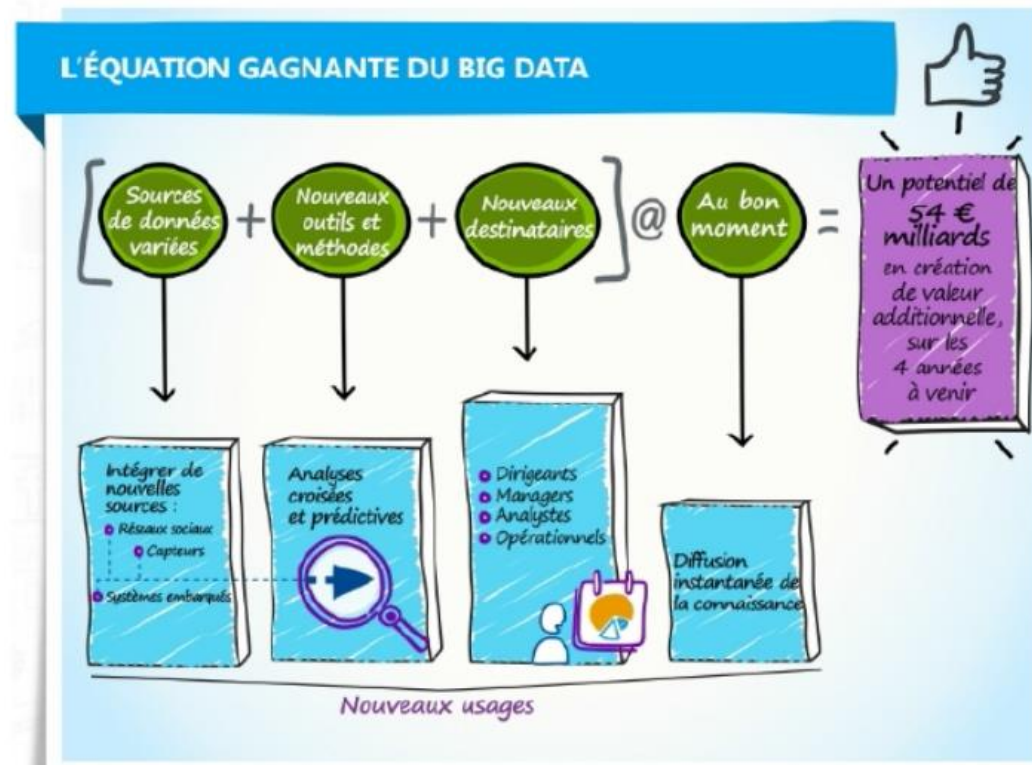
DEFINITION

- Business Intelligence (BI), est l'informatique à l'usage des décideurs et des dirigeants des entreprises.
- Les systèmes de BI sont utilisés par les décideurs pour obtenir une connaissance approfondie de l'entreprise et de définir et de soutenir leurs stratégies d'affaires,
 - par exemple : d'acquérir un avantage concurrentiel,
 - d'améliorer la performance de l'entreprise,
 - de répondre plus rapidement aux changements,
 - d'augmenter la rentabilité, et d'une façon générale la création de valeur ajoutée de l'entreprise.

ETUDE - MICROSOFT 2014



ETUDE - MICROSOFT 2014



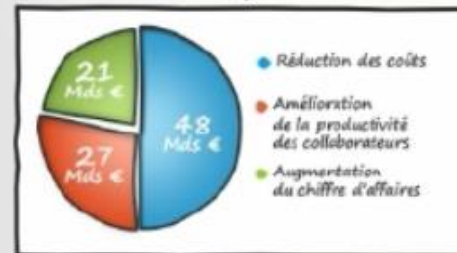
ETUDE - MICROSOFT 2014

COMPÉTITIVITÉ DES ENTREPRISES ET DONNÉES

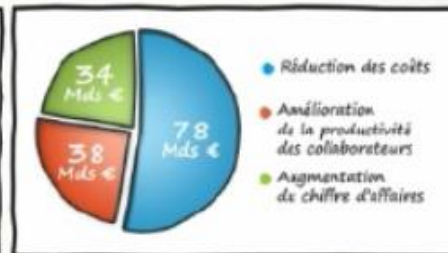


Une meilleure exploitation des données pourrait améliorer la compétitivité des entreprises à plusieurs niveaux

Entreprises aux usages avancés



Entreprises aux usages standards



LES MÉTIERS DU BI

4 métiers

- Data Integrator
- Data Analyst
- Data Scientist
- Data Steward (Responsable des données)

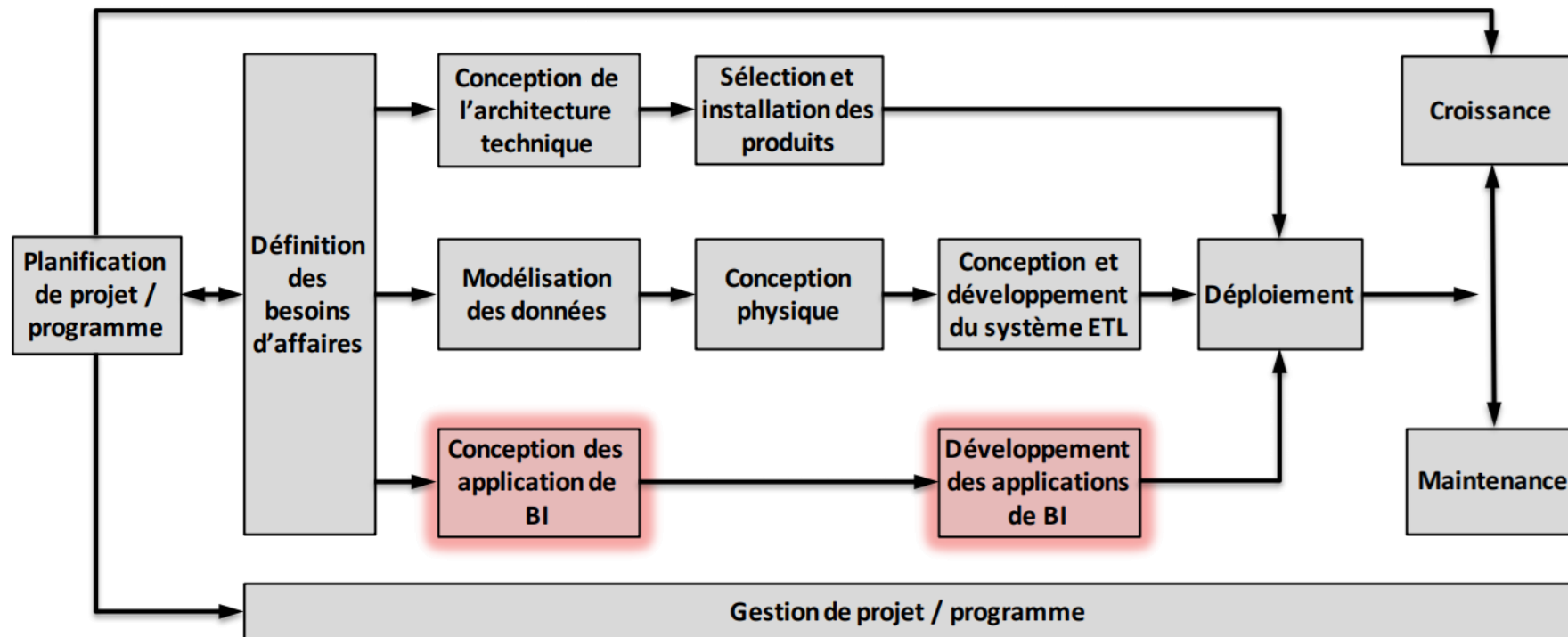
LES MÉTIERS DU BI: BAS NIVEAU

- Data Integration
 - Combiner des informations hétérogènes venants de sources différentes
- Data Analysis
 - Inspection, nettoyage, transformation et modélisation des données.
 - Data Mining, Data Visualisation
 - Rendre la donnée compréhensible
 - Communiquer à partir de la donnée

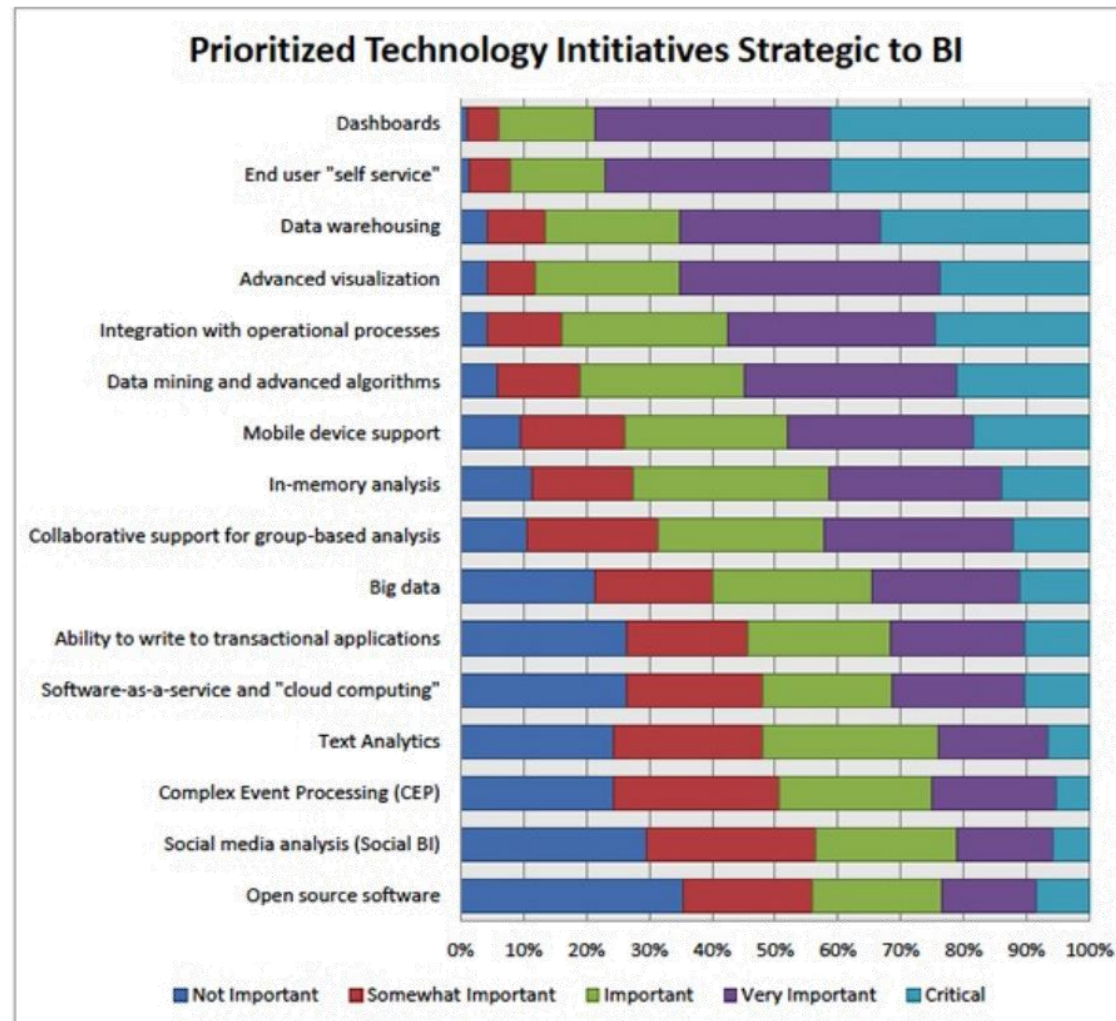
LES MÉTIERS DU BI: HAUT NIVEAU

- Data Scientist
 - Il s'agit de disposer de compétences de haut niveau en matière d'analyse de données, en combinant à la fois les méthodes statistiques, mais aussi d'autres connaissances telles que la linguistique, la sémantique, utiles notamment pour travailler sur des données non structurées, sans oublier la bonne compréhension du métier sur lequel on travaille, et de mettre en œuvre une démarche d'analyse itérative
- Data Steward - Responsable des Données
 - susceptibles sur un périmètre métier sur lequel ils détiennent une expertise reconnue, de spécifier les exigences sur les données et d'en contrôler la qualité. Ces responsables de données peuvent être positionnés à différents niveaux dans l'organisation, et peuvent être pilotés par des coordinateurs au niveau d'un métier

LE CYCLE DE VIE D'UN PROJET EN BI



LE CYCLE DE VIE D'UN PROJET EN BI



Source: Dresner Advisory Services (DAS),
2012 Wisdom of Crowds Business Intelligence
Market Study

TYPES D'ANALYSES EN BI

- BI stratégique:
 - Analyse les tendances d'affaires selon une métrique (ex: ventes);
 - Concerne les objectifs à long-terme de l'entreprise;
 - S'applique surtout aux analystes d'affaires de l'entreprise;

TYPES D'ANALYSES EN BI

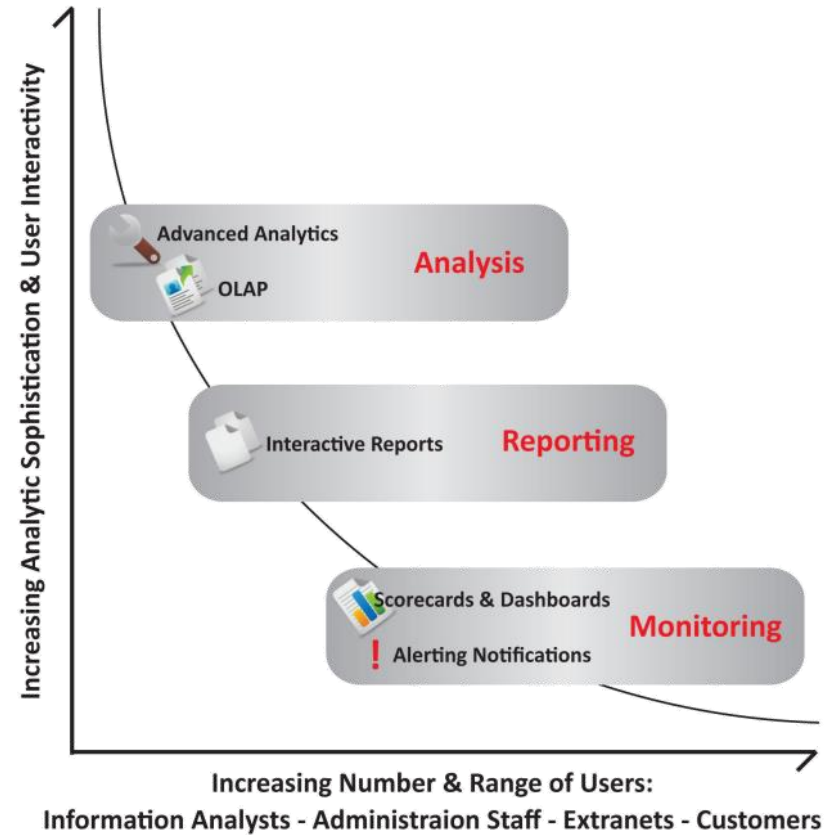
- BI tactique:
 - Fournit des métriques de performance (ex: scorecards);
 - Concerne les objectifs à court-terme de l'entreprise;
 - S'applique surtout aux gestionnaires et cadres supérieurs;

TYPES D'ANALYSES EN BI

- BI opérationnel:
 - Fournit du support informationnel aux points d'affaires de l'entreprise (ex: support à la clientèle);
 - Concerne l'état opérationnel de l'entreprise.

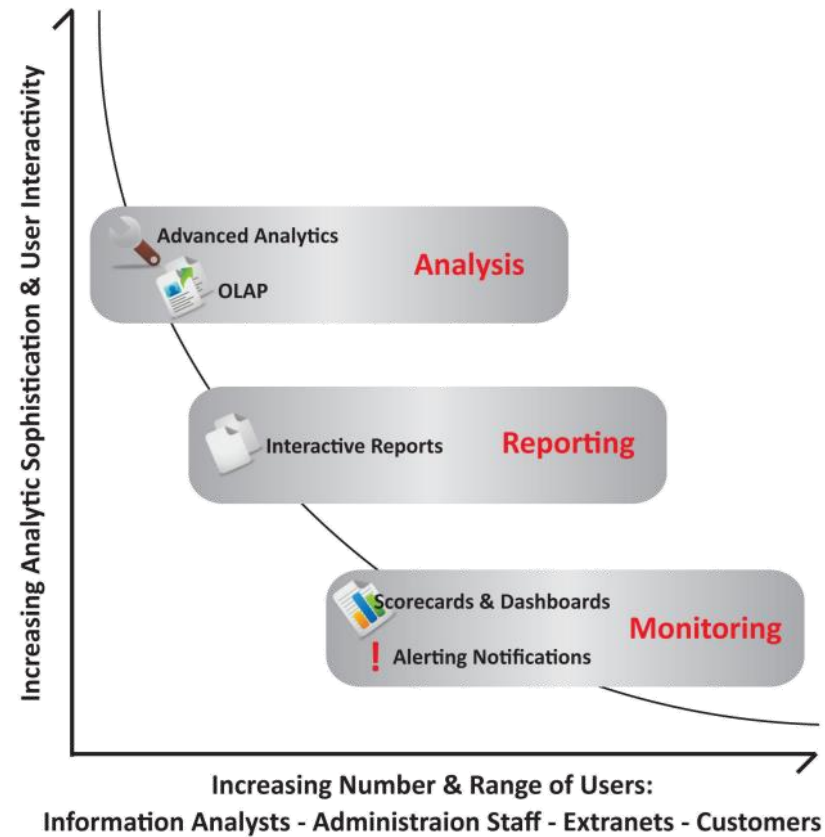
COMPLEXITÉ DES APPLICATIONS BI

- Complexité et utilisation:



COMPLEXITÉ DES APPLICATIONS BI

- Complexité et utilisation:



CONCEPTION D'APPLICATIONS BI

- **Importance**

- La plupart des utilisateurs finaux ne voit que cette partie;

- **Critères de conception:**

- Validité: les applications doivent fournir de l'information correcte et précise aux utilisateurs;
 - Performance: le délai de réponse aux requêtes des utilisateurs doit être court (ex: moins de 5 secondes);
 - Convivialité: les utilisateurs moins expérimentés doivent pouvoir utiliser le système (ex: rapport en 10 clics ou moins);
 - Apparence: l'interface des outils et rapports doit être attirant;
 - Maintenabilité: les applications doivent être suffisamment documentées et un plan d'amélioration doit être en place.

OUTILS BI: PAYANTS

- InfoSphere DataStage et DB2 de IBM:
 - IBM WebSphere DataStage est un ETL et fait parti de la plateforme d'IBM WebSphere Data Integration Suite.
 - Cette suite se base sur une interface graphique pour construire une solution d'intégration de données
 - Server ou Entreprise.
 - Nombre important de sources/cibles.
 - Intégration avec Oracle applications et SAP

OUTILS BI: PAYANTS

- **Oracle Business Intelligence Enterprise Edition Plus**
 - L'outil ETL se nomme Oracle Warehouse Builder (OWB)
- **Sage ERP X3**
- **SAP Netweaver BI (Ancien SAP BW)**
- **SAS**

OUTILS BI: OPENSOURCE

- Pentaho BI Suite
 - Pentaho est un projet global de la chaine décisionnelle.
 - Il comporte les briques suivantes : reporting, analyse, tableau de bord, data mining et workflow.
 - Il a été lancé par des vétérans du décisionnel : des ex de Business Objects, Cognos, Hyperion, IBM, Oracle, et SAS.
 - Selon les propos des fondateurs, leur objectif n'est pas uniquement de proposer une alternative Open Source mais bien de dépasser en termes de fonctionnalités les offres du marché.

OUTILS BI: OPENSOURCE

- Jaspersoft
 - La suite JasperSoft Business Intelligence (BI) fournit un outil intégré d'intégration de données, d'analyse et de reporting.
 - JasperSoft repose sur une architecture ouverte et sur un support multi-langage, notamment Perl, PHP, Python, et Java (sur lequel il repose à 100%),
 - la possibilité de connexion via les services web, notamment via SOAP.
 - Il est possible d'intégrer les notions de groupes d'utilisateurs multi-niveau et d'authentification à la génération de rapports (au format PDF, Excel, Word et HTML).

SYSTÈMES D'INFORMATION DÉCISIONNELS

NOSQL

Département d'informatique

Module: Systèmes d'information décisionnels

Chargé du module: Mokeddem, S

Email: sidahmed.mokadem@univ-mosta.dz



Année universitaire: 2017/2018

NOT ONLY SQL

- Données non-structurées
- Pas de schéma
- Pouvoir s'adapter à de nouvelles données
- Stocker au fur et à mesure et traiter plus tard
- Langage de requête non-standardisé
- Spécifique à chaque système

NOSQL

- Langage spécifique
 - Similaire dans l'idée au SQL
 - ALLER CHERCHER tel truc À L'ENDROIT machin AVEC LES CONTRAINTES
 - Parfois plus puissant, parfois mieux connu du développeur
- Contraintes inutiles
 - Éviter les redondances : stockage de masse pas cher
 - Optimiser les requêtes : requêtes souvent simples
- Localité
 - Données logiquement proches physiquement proches

ORIENTÉ GRAPHE

- Données et relations
 - Stockage de données avec beaucoup de relations complexes
 - En évitant les jointures
- Avantages et inconvénients
 - Beaucoup de données ressemblent à des graphes
 - Approprié pour parcourir les relations
 - Pas pour filtrer selon des contraintes

ORIENTÉ CLÉ VALEUR

- Opérations CRUD
 - Create : créer un objet
 - Read : lire à partir de la clé
 - Update : mettre à jour à partir de la clé
 - Delete : suppression à partir de la clé
- Indexation
 - Seulement la clé
 - On ne regarde pas le contenu de la valeur
- Avantages et inconvénients
 - Facile à utiliser, performances élevées
 - Pas de structure, pas de requêtes complexes

ORIENTÉ CLÉ VALEUR

- Opérations CRUD
 - Create : créer un objet
 - Read : lire à partir de la clé
 - Update : mettre à jour à partir de la clé
 - Delete : suppression à partir de la clé
- Indexation
 - Seulement la clé
 - On ne regarde pas le contenu de la valeur
- Avantages et inconvénients
 - Facile à utiliser, performances élevées
 - Pas de structure, pas de requêtes complexes

ORIENTÉ DOCUMENT

- Collection de documents
 - Clé/Valeur
 - On regarde dans la valeur Indexation
 - Identifiant unique par document
- Avantages et inconvénients
 - Modèle simple mais puissant
 - Requêtes complexes possibles
 - Mauvais passage à l'échelle pour des requêtes complexes

ORIENTÉ COLONNE

- Stockage colonne par colonne
 - Et pas ligne par ligne
- Indexation
 - Par colonne
- Avantages et inconvénients
 - Passe à l'échelle
 - Lecture difficile pour des données complexes

STRUCTURES

- Bases de données
 - Collections
- Collections
 - Documents
 - Chaque document a une clé primaire unique
- Document
 - Dictionnaire de valeurs
 - Aucune structure imposée

DOCUMENTS

```
{ "address":  
  { "building": "1007",  
    "coord": [ -73.856077, 40.848447 ],  
    "street": "Morris Park Ave",  
    "zipcode": "10462" },  
  "borough": "Bronx",  
  "cuisine": "Bakery",  
  "grades": [ { "date": { "$date": 1393804800000 }, "grade": "A", "score": 2 }, { "date": { "$date": 1378857600000 }, "grade":  
    "A", "score": 6 }, { "date": { "$date": 1358985600000 }, "grade": "A", "score": 10 }, { "date": { "$date": 1322006400000 },  
    "grade": "A", "score": 9 }, { "date": { "$date": 1299715200000 }, "grade": "B", "score": 14 } ],  
}
```

BASE ET COLLECTION

- Sélection de la base
 - `use NOM_DE_LA_BASE`
- Sélection de la collection
 - `db.NOM_DE_LA_COLLECTION.xxxx`
 - xxxx est l'instruction à effectuer

INSERTION

```
db.restaurants.insert(  
  { "address" : { "street" : "2 Avenue",  
                  "zipcode" : "10075",  
                  "building" : "1480",  
                  "coord" : [ -73.9557413, 40.7720266 ], },  
  "name" : "Vella", } )
```

RECHERCHE

- Conditions

- `db.restaurants.find({"borough": "Manhattan" })`
- `db.restaurants.find({"address.zipcode": "10075" })`
- `db.restaurants.find({"grades.grade": "B" })`
- `db.restaurants.find({"grades.score": { $gt: 30 } })`

- Logique

- ET : `db.restaurants.find({"cuisine": "Italian", "address.zipcode": "10075"})`
- OR : `db.restaurants.find({ $or: [{ "cuisine": "Italian" }, { "address.zipcode": "10075" }] }) 1`

MISE À JOUR

- `db.restaurants.update(
 { "name" : "Juni" },
 { $set: { "cuisine": "American (New)" },
 $currentDate: { "lastModified": true } })`
- `$currentDate` est une fonction qui met à jour le champ "lastModified"

SUPPRESSION

- Tous les documents
 - `db.restaurants.remove({ "borough": "Manhattan" })`
- Seulement un
 - `db.restaurants.remove({ "borough": "Queens" }, { justOne: true })`

AGGRÉGATION

- `db.restaurants.aggregate(
 [{ $match: { "borough": "Queens", "cuisine": "Brazilian" } },
 { $group: { "_id": "$borough", "count": { $sum: 1 } } }]);`
- Opérateurs
 - `$sum`, `$avg`, `$min`, `$max`, `$first`, `$last`, `$avg`

CALCULS

- `db.collection.aggregate([{ $group: { _id: "$item", avgAmount: { $avg: { $multiply: ["$price", "$quantity"]} } } }])`