

NFE211 - Business Intelligence (1) - Data Warehouses

Présentation

Prérequis

Ce cours s'adresse aux auditeurs préparant le diplôme d'ingénieur informatique option système d'information et/ou aux étudiants suivant le Master STIC mention Informatique Spécialité Systèmes d'Information et de Décision

Prérequis : Bonnes connaissances en bases de données et en systèmes d'information.

Objectifs pédagogiques

Acquérir une bonne connaissance de l'informatique décisionnelle (Business Intelligence), et en particulier de tous les processus liés à une architecture d'entrepôt de données (Data Warehouse).

Cette UE cible tout particulièrement la conception de l'entrepôt de données, la gestion du projet, et sa mise en œuvre, en particulier la mise en œuvre de la partie ETL.

Remarque : Pour une maîtrise complète du processus, tant du point de vue gestion des données que du point de vue analyse des données (data Science) il est conseillé de suivre également le NFE212 (semestre 2) qui fait suite au NFE211 (semestre 1)

Compétences

Capacité d'intégration dans une équipe de développement de système d'information décisionnel. En particulier compétences en conception et exploitation d'entrepôts de données :

- Capacité à concevoir et implémenter la partie ETL d'un Data Warehouse
- Capacité à concevoir et implémenter des cubes décisionnels
- Capacité à concevoir dans sa totalité une chaîne décisionnelle
- Capacité à mettre en œuvre une chaîne décisionnelle à l'aide des outils du marché
- Capacité à exploiter une chaîne décisionnelle
- Capacité à gérer l'intégration des données
- Capacité à gérer un projet décisionnel

Compétences

Capacité d'intégration dans une équipe de développement de système d'information décisionnel. En particulier compétences en conception et exploitation d'entrepôts de données :

- Capacité à concevoir et implémenter la partie ETL d'un Data Warehouse
- Capacité à concevoir et implémenter des cubes décisionnels
- Capacité à concevoir dans sa totalité une chaîne décisionnelle
- Capacité à mettre en œuvre une chaîne décisionnelle à l'aide des outils du marché
- Capacité à exploiter une chaîne décisionnelle
- Capacité à gérer l'intégration des données
- Capacité à gérer un projet décisionnel

Programme

Contenu

Introduction

- La Business Intelligence
- Business Intelligence et Big Data
- Objectifs d'un entrepôt de données
- OLAP versus OLTP

Architecture d'un entrepôt de données

Valide le 11-08-2022



Code : NFE211

Unité d'enseignement de type cours

6 crédits

Volume horaire de référence (+/- 10%) : **50 heures**

Responsabilité nationale :
EPN05 - Informatique / 1

Contact national :

Informatique d'entreprise

2D4P10 , 33 , 2 rue Conté

75003 Paris

01 58 80 84 71

Alexandre LESCAUT

alexandre.lescaut@cnam.fr

- Architecture matérialisée /architecture médiateur
- Dualité Entrepôt / magasins, Architecture de Inmon, Architecture de Kimball
- Data Warehouse / Data Lake
- Les méta-données, gestion, standardisation CWMI

Modélisation multidimensionnelle

- La modélisation multidimensionnelle, faits, dimensions, hiérarchies, indicateurs
- Modèles OLAP, ROLAP, MOLAP
- Modélisation en étoile, en flocon, en constellation
- L'algèbre multidimensionnelle
- Les dimensions à changement lent
- Les différents types de table de faits (récapitulatif / transaction / instantané)

Méthodes de modélisation d'un entrepôt de données

- Méthode par matrice
- Méthode par indicateurs
- Méthode MAP
- Méthode par analyse des requêtes
- Volumétrie, choix de la granularité, choix des Data Marts
- Méthodes de projets BI

Ingénierie d'extraction et d'intégration des données

- L'extraction des données
- Le nettoyage des données
- L'intégration sémantique des données
- Le rafraichissement des données
- Solution par programmation (Embedded SQL, déclencheurs)
- Solution par outils ETL

L'environnement technologique

- Outils ETL, Systèmes de gestion de bases de données OLAP, outils de modélisation multidimensionnelle
- Data Warehouse et Cloud,
- Data Warehouse et Big Data

Modalités de validation

- Projet(s)
- Examen final

Description des modalités de validation

L'UE sera validée par un examen final et par un projet réalisé individuellement ou en binôme.

Le projet - dont le sujet sera choisi par l'élève en accord avec le professeur - permettra de mettre en place une chaîne décisionnelle fondée sur les outils du marché (parmi l'offre industrielle ou celle du logiciel libre). Le projet peut aussi permettre d'approfondir un point théorique (par exemple la prise en compte du RGPD dans le Data Warehouse).

La note finale sera la moyenne de l'examen et du projet (avec la contrainte note examen ≥ 10).

Bibliographie

Titre	Auteur(s)
Fundamentals of datawarehouses (2d édition, Springer, 2003).	M. Jarke , M. Lenzerini, Y. Vassiliou, P. Vassiliadis

Systèmes d'aide à la décision et entrepôts de données (Encyclopedia E. Métais
Universalis) <http://www.universalis.fr/encyclopedie/systemes-informatiques-systemes-d-aide-a-la-decision/>

Building the Datawarehouse W.H. Inmon

Entrepôts de données, guide pratique de modélisation Kimball R, Ross M.
multidimensionnelle, Vuibert 2003