Modélisation en étoile, conception et mise en œuvre

Cours Pratique de 3 jours - 21h
Réf : AMT - Prix 2024 : 2 230€ HT

Ce cours, construit autour de nombreux cas pratiques, vous donnera une idée précise de la démarche de modélisation en étoile dans le cadre du projet Data Warehouse. Vous verrez pourquoi elle est l'expression même du besoin de la maîtrise d'ouvrage et comment elle permet de faire converger la vision des opérationnels, des analystes et des pilotes sur les activités de l'entreprise.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

Conduire des interviews pour recueillir les besoins d'analyse auprès des métiers

Définir les critères de qualité d'un Data Warehouse

A partir d'un cahier des charges d'analyse, identifier les dimensions et les faits à placer dans le modèle

Concevoir et décrire un macro-modèle en étoile

Optimiser le modèle en étoile en vue de la prise en compte de la volumétrie et des problématiques de reporting

Décrire une architecture de chargement des données dans le modèle en étoile décrit

LE PROGRAMME

dernière mise à jour : 09/2022

1) Introduction et rappels

- Qu'est-ce qu'un système d'information décisionnel ?
- Evolution des exigences de décision dans le contexte actuel.
- Infocentres, SIAD, EIS, Data Warehouse, définition et positionnement.
- Comprendre la finalité de l'approche Data Warehouse.

2) Les architectures en réponse aux besoins décisionnels

- Les composants principaux, Data Warehouse, ODS ou "staging area", datamarts.
- Les architectures proposées par Kimball et Inmon. Avantages et inconvénients.
- Positionnement du modèle en étoile dans le Data Warehouse selon l'architecture.
- Les phases du cycle de vie d'un Data Warehouse.
- Les critères de qualité d'un Data Warehouse.
- La notion de métadonnée, de référentiel.

Réflexion collective : Définition des critères de qualité d'un Data Warehouse.

3) Principes et définitions de base sur la modélisation en étoile

- Rappels sur la modélisation des bases de données opérationnelles.
- Différences entre OLTP et OLAP.
- Entités, attributs, cardinalités, formes normales.
- Le principe de la dénormalisation pour concevoir un modèle en étoile.
- Comprendre les notions de fait, dimension et axe d'analyse.
- Les alternatives de modélisation : modèle en flocon, en galaxie.

PARTICIPANTS

Maîtres d'ouvrage et maîtres d'œuvre, responsables des systèmes décisionnels, responsables informatiques, responsables des études, architectes de Systèmes d'Information, chefs de projet.

PRÉREQUIS

Connaissances de base de l'analyse décisionnelle et des SGBD relationnels.

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques…
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

• Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les stages pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
• À l'issue de chaque stage ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
• Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

- Les règles et bonnes pratiques de modélisation en étoile. Proposition alternative de Kortink et Moody.
Etude de cas : A partir d'un cahier des charges d'analyse, identifier les dimensions et faits principaux d'un modèle.

4) Conception du modèle en étoile
- Organisation et synthèse des interviews utilisateur pour le recueil du besoin.
- Compréhension et identification des processus métiers à modéliser.
- Choix des dimensions d'analyse.
- Création de hiérarchies dans les dimensions.
- Identification des mesures et croisements avec les dimensions.
- Définition de la granularité de l'analyse.
- Définition des règles d'agrégation.
- Utilisation d'outils de modélisation.

Exercice : A partir d'objectifs fournis par la MOA, réaliser un macro-modèle, en reliant les dimensions.

5) Optimisation fonctionnelle du modèle en étoile
- Gestion de l'évolution des référentiels et du changement des nomenclatures.
- Gestion des dimensions à évolution lente et rapide.
- Les clés de substitution.
- Gestion de la qualité, fiabilité des données.
- Gestion du contexte non renseigné ou inconnu.
- Les dimensions dégénérées.

Echanges : Décrire l'impact d'un changement donné à partir d'un modèle proposé. Optimisation du modèle pour son évolution.

6) Replacer la modélisation dans le cadre du projet décisionnel
- Présentation de la méthode Kimball et Inmon pour l'organisation du projet.
- Les acteurs et livrables du projet.
- Recueil des besoins métier. Formalisation des exigences techniques et d'organisation.
- Identification des priorités et du périmètre pilote.
- Modélisation des informations.
- Choix de l'infrastructure. Implémentation et recette.
- Déploiement et maintenance du modèle.
- Gestion des historiques.

Jeu de rôle : Conduite d'interview de recueil de besoin pour l'analyse.

7) Optimisation physique du modèle
- Gestion de la performance des requêtes.
- Estimation de l'espace disque requis pour le modèle.
- Limitation de la taille occupée par une dimension.
- Agrégation directe de certains éléments dans les tables.
- Dimensions techniques pour assurer la traçabilité des faits.

Exercice : Estimations de volumétrie moyenne sur quelques cas d'analyse.

8) Alimentation du modèle en étoile
- Contraintes des systèmes opérationnels sources.
- Rôle des ODS dans l'alimentation.
- L'organisation des traitements dans la DSA (Data Staging Area).
- Les différents types d'alimentation (delta, stock, complète).
- Les étapes, les règles et les prérequis de l'alimentation.
- Gestion des rejets.
- Gestion des sources différentes pour l'alimentation d'une dimension ou d'un fait.
- ETL, les solutions d'alimentation disponibles sur le marché.

Exercice : Sur une étude de cas, proposer une architecture de chargement : ODS / Staging Area.

9) Restitution des informations d'un modèle en étoile

- Les différents types d'outils au service de la restitution.
- Le marché des outils de restitution.
- Optimisation du modèle pour l'exploration des données.
- Optimisation des index.
- Utilisation du partitionnement des tables.

Echanges : Présentation de bonnes pratiques pour optimiser le modèle en vue du reporting.

10) Conclusion

- Ce qu'il faut retenir.
- Les pièges à éviter.
- Pour aller plus loin.

LES DATES

CLASSE À DISTANCE
2024 : 17 juin, 16 oct., 16 déc.

PARIS
2024 : 10 juin, 09 oct., 09 déc.