Systèmes décisionnels, Data Factory
intégrer le Big Data dans le SID pour bâtir la Data Factory

Séminaire de 3 jours - 21h
Réf : SID - Prix 2024 : 2 890€ HT

De l'analyse des données ou analyse décisionnelle, nous sommes passés aujourd'hui à la "science de la donnée", encore appelée data factory, grâce à l'apparition des technologies big data, offrant de nouvelles perspectives de modélisation et de prédiction. Ce séminaire passera en revue l'écosystème d'analyse des données existant, pour faire du gisement des données de l'entreprise un réel levier de croissance.

OBJECTIFS PÉDAGOGIQUES

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

Comprendre la valeur ajoutée, les enjeux et les principes des systèmes décisionnels

Mélanger la modélisation décisionnelle classique et le mode découverte

Mettre en œuvre une démarche pour concevoir un référentiel des données d'entreprise

Intégrer le big data dans le SID pour bâtir la data factory

Suivre pas à pas les étapes pour piloter son projet SID

Faire les bons choix d'outils en ayant un panorama global du marché

EXEMPLE

Un exemple complet de mise en œuvre d'un système d'information décisionnel sera traité.

LE PROGRAMME

dernière mise à jour : 09/2023

1) Finalité et principes des SID

- Positionnement du SID par rapport au big data et au SIO (système d'information opérationnel).
- Tendances et évolutions du décisionnel. Data lake versus data warehouse.
- Nouveaux enjeux : valorisation de l'information, corrélation rapide.
- L'architecture d'un SID.
- Les choix stratégiques d'architecture : intégration du data lake.
- Nouvel équilibre entre modélisation préalable et analyse dynamique.

2) Démarche de conception du SID, impact du mode découverte

- La typologie universelle des demandes sur un SID autour du pilotage et de l'analyse comportementale prédictive.
- Maîtriser le processus de conception des data marts et des data labs.
- Comment passer de data marts en silo à un SI ouvert à la BI découverte et au big data ?
- Cohérence des modèles en étoile : multidimensionnel. Analyse en mémoire contre modèle en étoile. Modèle du data lab.
- Dénormalisation NoSQL versus dénormalisation décisionnelle classique.
- Différence entre analyse multidimensionnelle et prédictive.
- Éviter le foisonnement d'agrégats et indicateurs en réutilisant les développements.

PARTICIPANTS

Responsables informatiques, responsables des études, architectes SI, consultants et chefs de projet décisionnel, autres chefs de projets fonctionnels et techniques.

PRÉREQUIS

Aucune connaissance particulière.

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.

- Concevoir une architecture mixte BI classique-BI découverte-analytique dynamique.

Etude de cas : Proposer une démarche de conception à partir de besoins d'analyse.

3) Construire les référentiels

- Comment faciliter l'accès aux informations ? Réconcilier MOA et MOE autour du référentiel.
- Construire le référentiel d'entreprise. Les dimensions d'analyse et les indicateurs partageables.
- Construire une architecture couvrant toutes les étapes, du pilotage à l'analyse des comportements.
- Construire les dictionnaires pour le SID, utiliser les métadonnées pour gérer la cohérence.

Etude de cas : Déploiement des méthodes proposées sur des exemples.

4) Optimiser l'accès aux données

- Comment répondre aux besoins et définir les réponses informatiques optimales ?
- Le multidimensionnel, ROLAP, MOLAP ou hybride : les critères de choix, les évolutions futures.
- Le multidimensionnel in-memory.
- Organiser son data lake. Construire les data labs.
- Appliquer une normalisation décisionnelle à vos modèles en étoile.

5) Mesurer la valeur du SID

- Faire de son SID un levier de la stratégie d'entreprise.
- Réunir les cas d'usage en processus décisionnels.
- Urbanisation du SID : éviter trop de traitements et des couches sémantiques surchargées.
- Identifier les zones éligibles au cloud computing.
- Cartographier son SID pour lier finalité d'usage et données utilisées pour le RGPD.
- Définir les critères d'efficacité d'un SID.
- Piloter la valeur des données.

6) Big data en mode industriel

- Principaux types de cas d'usage.
- Problème de l'industrialisation des projets big data.
- Check-list des recommandations.
- Analytique – prédictif temps réel et streaming (CEP : complex event processing).

7) État de l'art des outils décisionnels

- Lier ou fusionner son data lake et son data warehouse pour créer la data factory.
- Panorama des suites décisionnelles : SAS, Microsoft, SAP BusinessObjects...
- Degré d'intégration du mode découverte, de l'analytique et de la data visualisation.
- ETL. Outils multidimensionnels. Déploiement web.
- Big data intégré au SID. BD NoSQL. BD NewSQL. Cohabitation entre les différentes bases de données.
- Analyse in-memory. Appliance ou commodity hardware.
- Positionnement de l'écosystème Hadoop dans l'analyse du comportement client.
- Basculer le SID sur une base de données NoSQL, NewSQL ou intégrer les approches ?
- Combiner une solution Agile de data discovery et des capacités d'industrialisation de la BI.

Etude de cas : Déterminer votre trajectoire d'évolution vers une architecture intégrée.

8) Opportunité et création de valeur pour l'entreprise

- Évaluer la valeur ajoutée pour l'entreprise et la conduite du changement utile.
- Gérer son portefeuille de projets, les prioriser. Critères de lotissement.
- Spécificités d'un projet décisionnel et d'un projet big data.
- Techniques d'analyse de besoin en décisionnel : les pièges à éviter.
- Comment évaluer la complexité, la maturité des besoins ?

- Conduire le passage du décisionnel classique au mixte classique-analytique en mode libre.

9) Organisation de la gouvernance : équipes - processus - données

- Différents acteurs et rôles respectifs. Nouvelle relation entre métiers et IT.
- Leviers de la BI découverte, du big data et de la data science.
- Positionner le décisionnel au sein de l'entreprise. Organiser la gouvernance, la cohérence globale des données.
- Créer une organisation cohérente et pluridisciplinaire.
- Préserver l'autonomie des utilisateurs. Gérer la réactivité.
- Intégrer les métiers dans le pilotage de la valeur : données et cas d'usage.
- Organiser la data factory. Administrer les composants du SID.
- Garantir la fiabilité des données et des informations.
- Mettre en place le processus de la gestion de la qualité et de la véracité.
- Définir les contrôles qualité minimum. Définir le phasage des contrôles.
- Impact de la réglementation RGPD sur la sécurité d'accès aux données.

LES DATES

CLASSE À DISTANCE
2024 : 02 juil., 15 oct., 10 déc.

PARIS
2024 : 25 juin, 08 oct., 03 déc.