

Formation DataGalaxy : gouvernance et catalogage des données d'entreprise

Réf. SF34592

Maîtrisez DataGalaxy pour structurer, cataloguer et tracer vos données à l'échelle de votre organisation.

2 j Durée	1 900 € HT Inter / participant	3 600 € HT Intra (≤12 pers.)	0 Sessions à venir
--------------	-----------------------------------	---------------------------------	-----------------------

Objectifs pédagogiques

- ✓ Identifier les fonctionnalités et les modules de la plateforme DataGalaxy
- ✓ Mettre en place des politiques de gouvernance et définir les rôles associés
- ✓ Modéliser des ensembles de données et gérer les relations entre entités
- ✓ Tracer les flux de données et analyser l'impact des modifications (data lineage)
- ✓ Créer et exploiter des catalogues de données pour faciliter la découverte d'information
- ✓ Appliquer des règles de validation pour assurer la qualité des données

Public visé & prérequis

Public visé Business analysts, Chefs de projet, Data analysts, Data scientists, Ingénieurs Big Data

Prérequis- Connaissances générales en gestion de données (bases de données, modèles de données) - Notions de gouvernance des données ou de data management

Programme détaillé

Découverte de DataGalaxy et de son écosystème

- Positionnement de DataGalaxy : rôle d'une plateforme de data governance dans l'écosystème data d'entreprise
- Interface et navigation : organisation de l'espace de travail, modules principaux et vues disponibles
- Concepts clés : objets, propriétés, métadonnées et liens au sein de la plateforme
- Cas d'utilisation typiques : scénarios de mise en œuvre selon les besoins métier et IT
- Philosophie collaborative : approche DataGalaxy pour impliquer les différents acteurs de la donnée

Activités pratiques :

- Exploration guidée de l'interface et navigation dans un workspace de démonstration
- Identification des modules et de leur rôle dans un cas d'usage concret

Modélisation des données

- Glossaire et dictionnaire de données : structuration des définitions métier et techniques dans DataGalaxy
- Création de modèles de données : définition des entités, attributs et schémas
- Gestion des relations : liens entre objets, cardinalités et dépendances
- Diagrammes de modélisation : visualisation graphique des modèles pour faciliter la compréhension
- Bonnes pratiques de modélisation : conventions de nommage, granularité et maintenance des modèles

Activités pratiques :

- Création d'un modèle de données à partir d'un cas métier (entités, attributs, relations)
- Construction d'un glossaire métier et liaison avec les objets techniques

Gouvernance des données

- Principes de gouvernance : enjeux, objectifs et bénéfices d'une politique de gouvernance structurée
- Politiques de gouvernance : définition et mise en place de règles adaptées au contexte de l'organisation

- Rôles et responsabilités : attribution des rôles (data owner, data steward) et périmètres associés
- Surveillance de la conformité : indicateurs et tableaux de bord pour suivre l'application des politiques
- Documentation : centralisation de la documentation métier et technique au sein de la plateforme

Activités pratiques :

- Définition d'une politique de gouvernance et attribution des rôles sur un périmètre donné
- Mise en place d'un tableau de bord de suivi de la conformité

Catalogage des données

- Enjeux du catalogage : valeur d'un catalogue de données pour la découverte et la réutilisation de l'information
- Création et gestion de catalogues : alimentation manuelle et par import de fichiers en masse
- Classification et organisation : taxonomies, tags et regroupements pour structurer le catalogue
- Recherche et découverte : fonctionnalités de recherche avancée et filtres pour naviguer dans le catalogue
- Enrichissement par les métadonnées : ajout de descriptions, propriétaires et indicateurs de qualité

Activités pratiques :

- Création d'un catalogue de données et import d'un jeu de données via fichier
- Recherche et découverte de données à partir de critères métier

Data lineage et traçabilité des flux

- Principes du data lineage : traçabilité de bout en bout des données, de la source à la consommation
- Visualisation des parcours : cartographie graphique des flux de données dans DataGalaxy
- Identification des sources et destinations : mapping des systèmes amont et aval
- Analyse d'impact : évaluation des conséquences d'une modification sur les données en aval
- Connecteurs : principe d'utilisation des connecteurs pour alimenter automatiquement le lineage

Activités pratiques :

- Traçage d'un flux de données de bout en bout et visualisation du lineage
- Simulation d'une analyse d'impact suite à la modification d'une source

Qualité des données

- Dimensions de la qualité : complétude, exactitude, cohérence, fraîcheur et unicité des données
- Évaluation de la qualité : identification et mesure des problèmes de qualité dans le catalogue
- Règles de validation : définition et application de contrôles pour maintenir un niveau de qualité cible
- Monitoring : suivi continu de la qualité via indicateurs et alertes

Activités pratiques :

- Définition de règles de validation sur un jeu de données et interprétation des résultats
- Configuration d'un tableau de bord de suivi de la qualité

Sécurité, confidentialité et personnalisation

- Gestion des accès et des permissions : configuration des droits utilisateurs et périmètres de visibilité
- Protection des données sensibles : classification des données sensibles et application de restrictions
- Conformité réglementaire : alignement avec les exigences RGPD et normes sectorielles
- Personnalisation de l'expérience : paramétrage des écrans, vues et tableaux de bord selon les profils

Activités pratiques :

- Configuration des droits d'accès par rôle sur un espace de travail
- Personnalisation d'un tableau de bord adapté à un profil data steward

Pourquoi choisir Sparks Formation ?

20 ans d'expertise IT, des formateurs experts et des sessions garanties contre l'annulation.

20 ans
d'expertise

1 068+
formations

11 178
apprenants formés/an

+67
Net Promoter Score

Avis des participants

Note moyenne des participants de la filière Data Engineering & Big Data : 4,9/5 — 13 avis

Arnaud M. — That was really useful for me, even though I am a PDA. The trainer was very patient with me (I needed more explanation as I was the only...

Thierry R. — Formation idéale pour appréhender la technologie nosql Bon support de formation. Bonne pédagogie

Votre formateur

Isabelle V. — Formateur

Isabelle intervient notamment sur l'analytics et la Business Intelligence, ainsi que sur Microsoft Power Platform. La conception de supports pédagogiques fait également partie de ses activités.

Ils nous font confiance

Cette année notre filière Data Engineering & Big Data a accompagné plus de 73 entreprises dont :

SOPRA STERIA CAPGEMINI TECHNOLOGY SERVICES AMADEUS SAS GROUPAMA SUPPORTS & SERVICES AUBAY FRAIKIN FRANCE LBC FRANCE
ORANGE BUSINESS SERVICES SA ACATUS INFORMATIQUE ACENSI SANA +63 autres

Financement & certifications

Cette formation est éligible aux financements OPCO et FNE-Formation. Nos équipes vous accompagnent dans le montage de votre dossier.



Réservez votre place dès maintenant

20 ans d'expertise · 1 000+ formations · NPS +67 · réponse sous 24h ouvrées

0 805 950 800

demande@sparks-formation.com