

Comment construire une gouvernance sur votre data lake ?

<https://training.xebia.fr/formation-data/comment-construire-une-gouvernance-sur-votre-data-lake/>



Agile



Agilité à l'échelle



Management
Agile



DevOps



Data



JVM/Scala

Comment construire une gouvernance sur votre data lake ?

Avec *Julien Buret*

Cette formation propose une **démarche pragmatique et itérative pour construire une gouvernance data** adaptée à vos besoins. Elle est le fruit de retours d'expérience. À la fin de la journée, ces questions ne resteront plus sans réponses :

- Comment puis-je faire confiance aux données de mon data lake ?
- Quelles données sont collectées et transvasées dans mon data lake ?
- À quel rythme et selon quelle stratégie d'intégration ?
- Quels process et traitements sont faits sur mes données ?
- Comment identifier le responsable de tel ou tel jeu de donnée ?
- Comment puis-je y avoir accès ?
- Puis-je avoir confiance en ces données ? Sont-elles de qualité ?
- Comment partager au métier des données compréhensibles et exploitables ?
- Comment limiter l'accès aux données personnelles ?

Vous pensiez insurmontable la mise en place d'**une stratégie de "gouvernance de données"** sur **votre data lake** ? Vous allez changer d'avis !

Programme

Introduction

Changement de paradigme : vers une vision centralisée des données

- L'histoire du Big Data
- Les enjeux des data lakes
- Exemples d'architecture de data lake

Démarrer son projet de gouvernance de data lake

- Bases de données traditionnelles vs. HDFS
- Inventorier la donnée
- Organiser les données

- Documenter les métadonnées

Définir les rôles et compétences clés

- DPO / Responsable de traitements
- Data governor / steward
- etc.

Sécuriser ses données

- Disponibilité
- Confidentialité
- Anonymisation
- Habilitation
- Audit

Mettre en place des workflows

- Mise à disposition de données
- Gérer les habilitations
- Se mettre en conformité avec la régulation

Gérer le cycle de vie des données

- Qualité de la donnée
- Lineage
- Régulation
- Nettoyage

Connecter les systèmes au data lake

- NoSQL
- Kafka
- Dataware
- etc.

Conclusion

À la suite de cette formation, les participants adresseront aisément les sujets suivants :

- Identifier les actions à mettre en place pour construire une gouvernance sur son data lake.
- Savoir démarrer un projet de gouvernance data lake avec les bons réflexes et les bons outils.

Méthodes pédagogiques

Présentation théorique d'une gouvernance sur un data lake. C'est un exposé interactif où les participants sont vivement encouragés à poser des questions durant le déroulé.

Tarif HT

Inter : 900 €

Prochaines dates de formation

30 mars

20 juillet

16 novembre

[Je m'inscris](#)

info@xebia-training.fr

[Programme PDF](#)

[Imprimer le programme](#)

- Chef de projet Big Data
- Architecte Big Data
- Data Governor / Data Steward

Prérequis

Des connaissances générales sur les architectures Big Data sont attendues.

Validation

À la fin de cette formation, les stagiaires recevront une attestation de présence.



SOFTWARE TRAINING **DONE RIGHT**