Architecture Data Lakehouse — Concepts, design patterns et implémentation

Lien : https://innov-systems.com/formation/architecture-data-lakehouse-concepts-design-patterns-et-implementation

DURÉE
5 jours (35h)

RÉFÉRENCE
BDI168

CATÉGORIE
Big Data : fondamentaux et architecture

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre les principes du Data Lakehouse
- Maîtriser les architectures data modernes unifiées
- Concevoir des pipelines de données scalables
- Implémenter des design patterns Lakehouse
- Optimiser les performances des systèmes data
- Gérer la gouvernance et la qualité des données
- Intégrer des outils de data engineering modernes
- Industrialiser les traitements de données
- Migrer des architectures existantes vers le Lakehouse
- Améliorer la flexibilité des plateformes data
- Réduire les coûts de stockage et de traitement
- Étudier des cas concrets de déploiement Lakehouse

POUR QUI ?

- Data engineers
- Data architects
- Cloud engineers
- BI developers
- Analytics engineers
- DevOps engineers
- SRE (Site Reliability Engineers)
- Consultants data
- Chefs de projet data
- Responsables data platform
- Machine learning engineers
- Toute personne impliquée dans les architectures data modernes

Innov Systems

Programme détaillé

1 / Introduction au Data Lakehouse

- Évolution : Data Warehouse vs Data Lake vs Lakehouse
- Principes fondamentaux du Lakehouse
- Cas d’usage en data moderne

2 / Architecture globale du Lakehouse

- Stockage unifié des données
- Séparation stockage / calcul
- Moteurs analytiques compatibles

3 / Formats de données et stockage

- Parquet, Delta Lake, Iceberg, Hudi
- Optimisation du stockage distribué
- Gestion des métadonnées

4 / Ingestion des données

- Batch vs streaming ingestion
- Pipelines de données modernes
- ETL vs ELT dans le Lakehouse

5 / Design patterns Lakehouse

- Medallion architecture (Bronze, Silver, Gold)
- Data vault adapté au lakehouse
- Patterns de transformation analytique

Innov Systems

6 / Traitement et transformation des données

- Spark et moteurs distribués
- SQL analytics sur data lake
- Optimisation des transformations

7 / Gouvernance et qualité des données

- Data lineage et catalogage
- Contrôle de qualité des données
- Gestion des accès et sécurité

8 / Performance et optimisation

- Partitionnement et indexing
- Caching et optimisation des requêtes
- Gestion des coûts cloud

9 / Orchestration et industrialisation

- Pipelines de données automatisés
- CI/CD pour data engineering
- Monitoring des workflows

10 / Cas d’usage et mise en production

- Architecture Lakehouse en entreprise
- Migration depuis Data Warehouse
- Études de cas réels

Innov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

15 au 19 Fév. 2027	Présentiel - Casablanca
12 au 16 Avr. 2027	Présentiel - Casablanca
07 au 11 Juin 2027	Présentiel - Casablanca
02 au 06 Août 2027	Présentiel - Casablanca
27 Sep. au 01 Oct. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés