

# Formation : Outiller la gouvernance des données

Formation pratique - 3j - 21h00 - Réf. GDP

Prix : 2360 CHF H.T.

★★★★☆ 4,9 / 5

NEW

La gouvernance des données répond à des défis complexes. Pour être efficace, elle a besoin d'être outillée pour apporter des services pertinents aux différents acteurs, notamment par le catalogue de métadonnées, son principal outil. Les métadonnées sont toutes les informations créées automatiquement ou conçues par les professionnels de la data. Ce catalogue gagne à être fortement automatisé puis augmenté par l'IA et en particulier par un LLM pour atteindre une certaine efficacité. Son intégration dans votre écosystème technique est essentielle pour capter automatiquement les métadonnées.

## Objectifs pédagogiques

À l'issue de la formation, le participant sera en mesure de :

- ✓ Comprendre le rôle des métadonnées pour avoir une gouvernance efficace (DAMA)
- ✓ Comprendre l'écosystème des métadonnées dans le système d'informations
- ✓ Connaître les outils qui gèrent les référentiels de métadonnées
- ✓ Connaître les outils d'architecture et référentiels de métadonnées connexes aux outils de la gouvernance
- ✓ Connaître un panorama d'outils de la connaissance du patrimoine et leur extension via l'IA et les LLM
- ✓ Identifier les critères pour construire une grille d'évaluation d'un catalogue adapté à un contexte
- ✓ Identifier l'impact de l'IA sur la gouvernance et le rôle de la gouvernance pour l'IA

## Public concerné

Toutes les personnes intéressées pour outiller la gouvernance des données autour d'un catalogue de données efficace.

## Prérequis

Avoir suivi le cours de la gouvernance des données (ref. GDD).

### PARTICIPANTS

Toutes les personnes intéressées pour outiller la gouvernance des données autour d'un catalogue de données efficace.

### PRÉREQUIS

Avoir suivi le cours de la gouvernance des données (ref. GDD).

### 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.

## Méthodes et moyens pédagogiques

### Méthodes pédagogiques

Active

### 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.

## Programme de la formation

### 1 Introduction : la gouvernance des données et le catalogue de données

- Définition de la gouvernance des données
- Le ROI d'une gouvernance et le TTM des projets
- Gouvernance des données, data management et framework du DAMA
- Rôle et périmètre de la gouvernance : compliance, qualité et connaissance du patrimoine
- Le rôle de la gouvernance pour le développement des cas d'usage d'IA
- Un lien fort entre la gouvernance des données et le catalogue des données

#### Etude de cas

Carte des acteurs de la data et leur rôle dans la gouvernance. Analyse du ROI de la gouvernance sur 2 cas d'usage. Échange.

### 2 Les métadonnées

- Définition des métadonnées
- Quelques exemples sur le cycle de vie de la donnée
- Multiplicité des métadonnées
- Métadonnées créées pour décrire les données fonctionnellement
- Métadonnées créées par des systèmes et collectables
- Gestion de métadonnées : un ou plusieurs référentiel(s) de métadonnées ?

#### Etude de cas

Identifier et modéliser les métadonnées pour respecter les exigences de sécurité et les exigences de la réglementation sur les données personnelles.

### 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 formations pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque formation 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 participant 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

Pour toute question ou besoin relatif à l'accessibilité, vous pouvez joindre notre équipe PSH par e-mail à l'adresse [psh-accueil@orsys.fr](mailto:psh-accueil@orsys.fr).

### 3 La connaissance du patrimoine à travers les métadonnées

- Connaissance du patrimoine et gestion des assets à travers les métadonnées
- Connaissance des sources des données et le lien avec l'architecture
- Connaissance des données (datasets) disponibles sur une plateforme de données
- Identification des données personnelles, anonymisée ou pas sur une plateforme de données
- Descriptions techniques et fonctionnelles (objets métier, processus, définition) des données d'une plateforme data
- Connaissance de la disponibilité et de la fraîcheur des données d'une plateforme data
- Connaissance des caractéristiques de la qualité des données de la plateforme
- Connaissance des caractéristiques de la sécurité et des réglementations des données présentes sur une plateforme

#### Etude de cas

Modéliser une première version du modèle de métadonnées de son futur catalogue. Consulter son RSSI pour certains besoins spécifiques (possibilité d'avoir des échantillons de données ou des éléments de qualité décrits par des données dans son catalogue).

### 4 L'outillage de la connaissance du patrimoine

- Catalogue de métadonnées pour parler des données, les partager et les valoriser
- Catalogue de données et modèles de données, dictionnaire de données, glossaire, thésaurus
- Données structurées et non structurées dans un catalogue
- Fonctionnalités d'un catalogue de données
- Besoins des acteurs et leur grande diversité
- Connaissance tribale
- Mise à jour de ces différents éléments de connaissance des données d'une plateforme
- Catalogue de données ou plateforme de gouvernance des données
- Panorama d'outils de catalogue de données et de plateforme de gouvernance

#### Etude de cas

Élaboration d'une grille d'évaluation des outils, puis choix de fonctionnalités à mettre en œuvre pour l'entreprise de l'étude de cas

## 5 Automatisation et augmentation de la connaissance du patrimoine

- Automatisation et augmentation de la connaissance du patrimoine
- Enjeux de l'automatisation et de l'augmentation
- Présentation du potentiel d'un LLM
- Mise en commun des métadonnées existantes et organisées des outils d'architecture
- Automatisation des échanges entre les outils d'architecture et les catalogues de données
- Automatisation de la mise à jour du référentiel des métadonnées techniques disponibles sur les plateformes (modèle de données physiques, description des champs...)
- Automatisation et augmentation des descriptions des datasets et de leur contenu à partir d'un contexte (modèle de données, dictionnaires, processus métiers, etc.)
- Automatisation et augmentation de l'identification des données personnelles
- Automatisation et augmentation de l'application des règles de confidentialités sur les données à partir d'un contexte
- L'impact de l'IA sur la gouvernance

### Etude de cas

Un standard pour les échanges de données avec les outils d'architecture.  
L'usage d'un LLM et du NLP pour décrire des données et des datasets à partir d'une base de connaissance.

## 6 Les autres outils de métadonnées : data lineage, observabilité de la data et qualité

- Définition et exemple de métadonnées associées
- Enjeux du data lineage
- Besoins de traçabilité quant à la qualité et au réglementaire (BCBS239)
- Automatisation du data lineage : succès et limite
- Qualité des données : un problème complexe
- Description de la qualité des données d'une plateforme de données
- Référentiel de tests de qualité partagé par tous
- Observabilité des données et ses métadonnées
- Lien du catalogue de données et d'un outil d'observabilité des données
- Outils et plateformes disponibles

### Etude de cas

Les liens du catalogue et de l'observabilité de la data : comment suivre la disponibilité des données sur une plateforme de données ? Comment partager ces informations avec les métiers et les data ingénieurs ?

## 7 Lancer son projet de catalogue de données

- Identification des acteurs
- Comprendre un écosystème data pour automatiser et augmenter
- Le prérequis d'une couche sémantique
- One size doesn't fit all : une grille d'évaluation de maturité et des exigences
- Premiers pas dans une organisation

### Échanges

Partir avec les premières actions pour son contexte

