

MDM MASTER DATA MANAGEMENT MÉTHODOLOGIE

— Public	Responsables SI, chefs de projet data et acteurs de la gouvernance des données
— Durée	2 jours - 14 heures
— Pré-requis	Connaître les principaux processus et applications de son organisation. Une compréhension générale des bases de données et des échanges entre systèmes facilite la participation aux ateliers.
— Objectifs	Distinguer inventaire, dictionnaire, glossaire, cartographie et lignage des données. Construire un inventaire intégrant les métadonnées métier, techniques et de sécurité. Cartographier les échanges entre applications et identifier les zones à risque. Définir les responsabilités de gouvernance. Élaborer une feuille de route adaptée à un premier périmètre.
— Méthodes pédagogiques	Pour bien préparer la formation, le stagiaire remplit une évaluation de positionnement et fixe ses objectifs à travers un questionnaire. La formation est délivrée en présentiel ou distanciel (e-learning, classe virtuelle, présentiel et à distance). Le formateur alterne entre méthodes démonstratives, interrogatives et actives (via des travaux pratiques et/ou des mises en situation). La validation des acquis peut se faire via des études de cas, des quiz et/ou une certification. Cette formation est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par Audit Conseil Formation.
— Moyens techniques	1 poste de travail complet par personne De nombreux exercices d'application Mise en place d'ateliers pratiques Remise d'un support de cours Passage de certification(s) dans le cadre du CPF Remise d'une attestation de stage
— Modalité d'évaluation des acquis	Evaluation des besoins et objectifs en pré et post formation Evaluation technique des connaissances en pré et post formation Evaluation générale du stage
— Délai d'accès	L'inscription à cette formation est possible jusqu'à 5 jours ouvrés avant le début de la session
— Accessibilité handicapés	Au centre d'affaires ELITE partenaire d'ACF à 20 m. Guide d'accessibilité à l'accueil.

JOUR 1 – MATIN : CADRAGE STRATÉGIQUE, TYPOLOGIE DES DONNÉES ET FONDATIONS THÉORIQUES

1. ENJEUX, DÉFINITIONS ET CONCEPTS FONDAMENTAUX

- Pourquoi cartographier ?
- Éliminer les silos d'information
- Réduire la dette technique
- Accélérer les projets d'analytique/IA, et répondre aux impératifs réglementaires et de sécurité (RGPD, NIS 2, DORA, secret des affaires)
- Déconstruire le jargon (distinctions clés) :
- Inventaire de données : liste statique des actifs (« ce que l'on possède »)
- Dictionnaire de données : définition technique et règles de formatage au niveau champ/colonne (types, contraintes, nullabilité)
- Glossaire métier: sémantique partagée des termes de l'entreprise (« qu'est-ce qu'un client actif ? »)
- Cartographie des flux : dynamiques d'échanges et dépendances inter-systèmes
- Lignage des données: traçabilité de bout en bout de l'origine jusqu'au tableau de bord (horizontal) et décomposition système-table-colonne (vertical)
- Niveaux de granularité :
- Macro (système d'information)
- Méso (schémas, tables, API)
- Micro (champs, attributs spécifiques)

2. TYPOLOGIE ET CLASSIFICATION DES DONNÉES

- Données de référence vs données transactionnelles :
- Concepts du MDM
- Sources de vérité
- Classification par criticité :
- Disponibilité
- Intégrité
- Confidentialité (critères DICP)
- Classification réglementaire :
- identification des DCP
- Données sensibles (santé, biométrie)
- Données d'importance vitale

3. DÉFINITION DU MODÈLE DE GOUVERNANCE (RÔLES & RESPONSABILITÉS)

- **Matrice RACI appliquée à la donnée :**
- **Data Owner :** responsabilité légale, validation des accès, validation sémantique
- **Data Steward :** documentation continue, résolution des anomalies de qualité
- **Data Custodian / DSI :** hébergement, sauvegardes, sécurité technique et performance
- **DPO / RSSI :** contrôle de la conformité, analyse d'impact (AIPD/EIVP) et politique de chiffrement

JOUR 1 – APRÈS-MIDI : MÉTHODOLOGIE D'INVENTAIRE, MÉTADONNÉES ET CAS PRATIQUE DE COLLECTE

1. ARCHITECTURE DES MÉTADONNÉES : QUOI COLLECTER CONCRÈTEMENT ?

- **Métadonnées techniques :**
- Nom de table/base
- Nom du champ
- Type (VARCHAR, INT...)
- Contraintes
- Volumétrie, fréquence de rafraîchissement
- **Métadonnées métier :**
- Définition vulgarisée
- Domaine d'activité
- Cas d'usage autorisés
- Règles de gestion métier
- **Métadonnées opérationnelles :**
- Date de création, dernier accès
- Statut d'archivage, durée de rétention légale (purge)
- **Métadonnées de sécurité/conformité :**
- Niveau de sensibilité (Public, Interne, Confidentiel, Secret)
- Base légale de traitement
- Statut de chiffrement (au repos/en transit)

2. MÉTHODE D'INVESTIGATION POUR LE CADRAGE DU PILOTE

- **Sélection du périmètre :**
- Choisir un périmètre à forte valeur ajoutée (ex. Processus Lead-to-Cash, Onboarding client ou Gestion des réclamations)
- **Techniques de collecte terrain :**
- Entretiens qualitatifs avec les équipes métier (questionnaires types)
- Extraction des schémas de bases de données (Reverse engineering SQL/NoSQL)
- Inventaire des échanges occultes : les « bases Excel » et fichiers partagés non répertoriés (Shadow IT)

3. ATELIER PRATIQUE : CONSTITUTION DU PREMIER CATALOGUE/INVENTAIRE

- **Cas pratique :** À partir d'un scénario d'entreprise (ex. chaîne de commande omnicanale avec e-commerce, CRM, ERP, passerelle de paiement et transporteur).
- **Exercice :**
- Définir le glossaire métier commun
- Remplir la grille d'inventaire standardisée (modèle matriciel)
- Identifier les Data Owners et qualifier le niveau de sensibilité de chaque attribut
- **Débriefing :**
- Analyse des écarts de définitions entre équipes et arbitrage des ambiguïtés

JOUR 2 – MATIN : MODÉLISATION DES FLUX, CARTOGRAPHIE DYNAMIQUE ET LIGNAGE

1. MODÉLISER LA CIRCULATION DE LA DONNÉE (DATA FLOW MAPPING)

- **Cycle de vie de la donnée :**
- Collecte/Création → Stockage → Transformation/Enrichissement → Exposition/Partage → Archivage → Destruction
- **Typologie des flux et canaux d'échange :**
- Synchrones (API REST/GraphQL, Webhooks, requêtes directes)
- Asynchrones (fichiers plats SFTP, batchs ETL/ELT nocturnes, files de messages Kafka/RabbitMQ)
- **Formalismes de représentation :**
- Notations standards : BPMN (avec artefacts Data Objects et Data Stores), DFD (Data Flow Diagram), ou diagrammes d'architecture (C4 Model, ArchiMate)
- Identification des règles de transformation : comment un champ A d'une source S1 devient un champ agrégé B dans un Data Warehouse

2. ANALYSE CRITIQUE ET DIAGNOSTIC DE VULNÉRABILITÉ

- **Détection du Shadow IT :**
- Flux non sécurisés (exports CSV manuels par e-mail, synchronisations sur des clouds personnels)
- **Points de rupture et dette technique :**
- Points de défaillance uniques (SPOF)
- Transformations non documentées dans du code propriétaire
- Doublons non synchronisés
- **Conformité des flux :**
- Détection des transferts de données hors UE
- Zones de chiffrement manquant entre deux briques applicatives

3. ATELIER PRATIQUE : CARTOGRAPHIE VISUELLE D'UN FLUX COMPLEXE

- **Cas pratique :**
- Modéliser le schéma de flux complet du cas pratique (des systèmes sources jusqu'au reporting décisionnel / BI)
- Matérialiser visuellement les transformations, les protocoles (API vs Batch) et les zones à risque (RGPD / failles de sécurité)
- Restitution croisée : Comparaison des diagrammes, critique constructive de la clarté visuelle et de la complétude technique

JOUR 2 – APRÈS-MIDI : OUTILLAGE, PÉRENNISATION ET FEUILLE DE ROUTE OPÉRATIONNELLE

1. PANORAMA ET BENCHMARK DES OUTILS DU MARCHÉ

- **Niveau 1 – Démarrage & Prototypage (low-cost / agile) :**
- Inventaire : Tableurs structurés (Excel, Google Sheets), bases relationnelles légères ou no-code (Airtable, Notion)
- Diagrammes de flux : Outils visuels (Draw.io/diagrams.net, Miro, Lucidchart, PlantUML)
- **Niveau 2 – Documentation collaborative & Wiki :**
- Documentation semi-structurée (Confluence, BookStack) couplée à des modèles de pages strictes
- **Niveau 3 – Catalogues de données d'entreprise (Data Catalogs) :**
- Open-source : DataHub, Apache Atlas, OpenMetadata (focalisés sur le scanning automatique des métadonnées et le lignage)
- Solutions Editeurs / Entreprise : Collibra, Alation, Microsoft Purview, Atlan, Zeenea (gestion des politiques, workflow de validation, intégration multi-cloud)
- **Outils spécialisés conformité / Privacy :**
- OneTrust, Dastra (focus registre de traitement RGPD art. 30)
- **Critères de choix d'un outil :**
- Capacité d'automatisation (scanners de bases)
- Ergonomie pour les métiers
- Coût des licences, hébergement (SaaS vs On-Premise)

2. BONNES PRATIQUES DE PÉRENNISATION

- **Automatiser vs déclaratif :**
- Quand privilégier les collecteurs automatiques d'API/catalogues et quand maintenir la saisie humaine manuelle
- **Intégration dans le cycle de vie projet (Data by Design) :**
- Inclure l'étape « mise à jour de la cartographie » dans la définition de projet (DoD / Definition of Done en méthode Agile).
- **Gestion des versions et cycle de vie :**
- Politique de revue périodique

3. ATELIER PRATIQUE : FEUILLE DE ROUTE ET KIT DE DÉMARRAGE

- **Exercice :**
- Choisir le premier cas d'usage réel de l'entreprise d'un stagiaire
- Rédiger la charte de démarrage : objectifs mesurables, sponsors à convaincre, choix de l'outillage de lancement (Phase 1 vs Phase 2)
- Identifier les 3 risques majeurs du projet (résistance au changement, absence de documentation technique, manque de temps des métiers) et plan de remédiation

NOUS CONTACTER

Siège social

16, ALLÉE FRANÇOIS VILLON
38130 ÉCHIROLLES

Téléphone

04 76 23 20 50 - 06 81 73 19 35

Suivez-nous sur les réseaux sociaux, rejoignez la communauté !



ACF Audit Conseil Formation



@ACF_Formation

Dernière mise à jour : 14/10/2026

PROFIL Formateur : Les formateurs sont recrutés selon plusieurs critères :
Expérience, pédagogie, dynamisme et prévoyance.