

DATA CATALOGUE & DATA DICTIONARY

Projet MDM - Clinique Privée

1. MDM_PATIENT

DATA CATALOGUE

Nom de la table : mdm.mdm_patient

Description : Référentiel maître unique des patients de la clinique. Consolide les informations patients provenant des systèmes de rendez-vous, ERP (dossiers médicaux, laboratoire) et facturation. Élimine les doublons et garantit une vue 360° fiable de chaque patient pour améliorer la qualité des soins et la sécurité médicale.

Propriétaire métier : Dr. Sophie Laroche - Directrice Médicale

Propriétaire technique : Équipe Data Governance - Service IT Clinique

Fréquence de mise à jour : Temps réel (streaming via Kafka)

Volumétrie :

- Patients actifs : ~80 (après dédoublonnage)
- Enregistrements sources totaux : ~230 (avant consolidation MDM)
- Croissance estimée : +150 nouveaux patients/an

Sensibilité :

- **TRÈS ÉLEVÉE** - Contient des données de santé à caractère personnel (RGPD Art. 9)
- Données Personnelles Identifiables (PII)
- Données sensibles de santé (HADS - Health and Social Care Data)

Sources de données :

1. **Système Rendez-vous** (source_rdv_patients.csv)
 - Colonnes : nom, prénom, sexe, téléphone, adresse
 - Qualité : Doublons présents (~20%), incohérences de format
2. **Système ERP** (source_erp_patients.csv + source_erp_dossiers_medicaux.csv)
 - Colonnes : nom, prénom, date_naissance, email, num_dossier, historique_médical

- Qualité : IDs différents du système RDV, format dates DD/MM/YYYY
- 3. **Système Facturation** (source_fact_patients.csv)
 - Colonnes : nom_famille, prénoms, adresse_facturation
 - Qualité : Encore d'autres IDs, noms de colonnes différents

Services consommateurs :

- **Dossier Patient Partagé** (Application Web Médecins) - Accès lecture/écriture
- **Dashboards Médicaux** (Direction & Qualité) - Accès lecture, analyses statistiques
- **Système d'Alertes Médicales** (Notifications automatiques) - Accès lecture temps réel
- **Outils d'Analyse Statistique** (Études épidémiologiques) - Accès lecture agrégée
- **Module Facturation** (Lien vers factures patients) - Accès lecture
- **Admissions & Secrétariat** (Prise de RDV, coordination soins) - Accès lecture

Relations MDM :

- **Lié à MDM_Médecin** : Un patient peut être suivi par plusieurs médecins (relation gérée dans systèmes opérationnels)
- **Lié à MDM_Service** : Les services (labo, imagerie) produisent des données liées aux patients (relation indirecte)

DATA DICTIONARY - MDM_PATIENT

Colonne	Type	Description métier	Contraintes	Sensibilité	Source(s)	Règle de qualité
Patient_ID	UUID	Identifiant unique universel du patient dans le référentiel MDM. Généré automatiquement et immuable à vie. Utilisé comme clé de référence dans tous les systèmes consommateurs.	PRIMARY KEY, NOT NULL, UNIQUE, DEFAULT gen_random_uuid()	Faible	Généré par PostgreSQL (MDM)	Complétude = 100% (auto-généré)
Nom	VARCHAR(100)	Nom de famille du patient. Normalisé en MAJUSCULES sans accents pour faciliter recherches et matchings. En cas	NOT NULL, Longueur 2-100 caractères	Très élevée (PII)	RDV (nom), ERP (nom), FACT (nom_famille)	Complétude = 100%, Normalisation UPPER + sans accents

		de doublon, prend la valeur la plus fréquente ou la plus récente.				
Prénom	VARCH AR(100)	Prénom du patient. Conserve la casse d'origine (majuscule initiale). En cas de doublon, prend la valeur la plus fréquente .	NOT NULL, Longueur 2-100 caractères	Très Élevée (PII)	RDV (prenom), ERP (prenom), FACT (prenoms)	Complétude = 100%, Capitalisation standardisée
Date_Nais sance	DATE	Date de naissance du patient au format YYYY-MM-DD. Champ critique pour le dédoublement et le calcul d'âge.	NOT NULL, Date < TODAY, Date > '1900-01-01'	Très Élevée (PII)	ERP (date_naissance format DD/MM/YYYY)	Complétude = 100%, Format YYYY-MM-DD , Cohérence temporelle

Sexe	VARCHAR(10)	Sexe biologique du patient. Valeurs standardisées : 'Masculin', 'Féminin'. Utilisé pour statistiques épidémiologiques et protocoles de soins spécifiques.	CHECK (Sexe IN ('Masculin', 'Féminin')), NULL autorisé	Élevée (PII)	RDV (sex)	Complétude cible ≥90%, Valeurs contrôlées
Adresse	VARCHAR(255)	Adresse résidentielle complète du patient. Utilisée pour correspondance postale et localisation géographique (études de santé publique) .	NULL autorisé, Longueur max 255	Élevée (PII)	RDV (adresse), ERP (adresse), FACT (adresse_facturation)	Complétude cible ≥85%, Consolidation meilleure source

Téléphone	VARCHAR(20)	Numéro de téléphone principal du patient. Normalisé à 10 chiffres sans espaces ni caractères spéciaux. Utilisé pour rappels RDV et urgences .	NULL autorisé, Format 10 chiffres, Regex '^[0-9]{10}\$'	Élevée (PII)	RDV (telephone), FACT (tel_contact)	Complétude cible ≥90%, Validation format 10 chiffres
Email	VARCHAR(255)	Adresse email du patient. Validée par regex. Utilisée pour communications non-urgentes et accès portail patient.	NULL autorisé, Format email valide, Regex RFC 5322	Élevée (PII)	ERP (email), FACT (email_contact)	Complétude cible ≥70%, Validation regex email
Num_Dossier	VARCHAR(50)	Identifiant unique du dossier médical patient dans le système	NULL autorisé, UNIQUE si renseigné	Très Élevée (Santé)	ERP (num_dossier format DOSS-XXXXXX)	Complétude cible ≥95%, Unicité = 100%

		ERP. Permet la traçabilité vers l'historique médical complet.				
Historique Médical	TEXT	Antécédents médicaux, allergies et pathologies chroniques du patient. Champ libre enrichi au fil du temps. CRITIQUE pour décisions médicales et prévention erreurs médicamenteuses .	NULL autorisé, Longueur max 10000 caractères	CRITIQUE (Donnée de santé)	ERP (source_erp_dossiers_medicaux : historique_médical + allergies)	Complétude cible ≥70%, Enrichissement continu
Date_Création	TIMESTAMP	Date et heure de création de l'enregistrement MDM. Utilisée pour audit et traçabilité .	NOT NULL, DEFAULT CURRENT_TIMESTAMP	Faible (Métadonnée)	Généré automatiquement par MDM	Complétude = 100% (auto)

2. MDM_MEDECIN

DATA CATALOGUE

Nom de la table : mdm.mdm_medecin

Description : Référentiel maître unique des médecins et praticiens exerçant dans la clinique. Consolide les informations professionnelles provenant du système RH/ERP et du système de rendez-vous. Garantit l'identification unique de chaque praticien et facilite la gestion des plannings, affectations de services et traçabilité des prescriptions.

Propriétaire métier : Marie Dubois - Directrice des Ressources Humaines Médicales

Propriétaire technique : Équipe Data Governance - Service IT Clinique

Fréquence de mise à jour : Quotidienne (batch nocturne à 02h00) + Temps réel pour modifications urgentes

Volumétrie :

- Médecins actifs : 6
- Médecins historiques (retraités/partis) : Conservés pour traçabilité
- Croissance estimée : +2-3 nouveaux praticiens/an

Sensibilité :

- **MOYENNE** - Données professionnelles, non publiques
- Informations de contact professionnel
- Numéro de licence (RPPS) = donnée sensible professionnelle

Sources de données :

1. **Système RH/ERP** (source_erp_medecins.csv) - SOURCE PRINCIPALE
 - Colonnes : nom, prénom, spécialité, num_licence, email_pro, téléphone_pro, service_affecté, disponibilités
 - Qualité : Données complètes et validées RH
2. **Système Rendez-vous** (source_rdv_medecins.csv) - SOURCE SECONDAIRE
 - Colonnes : nom_complet (format "Dr. Prénom Nom"), spécialité, téléphone, email
 - Qualité : Format inconsistante, données partielles

Services consommateurs :

- **Module Planning** (Gestion agenda médecins) - Accès lecture/écriture
- **Système Rendez-vous** (Affectation RDV patients) - Accès lecture
- **Dossier Patient** (Identification prescripteur/médecin traitant) - Accès lecture
- **Facturation** (Lien actes médicaux → médecin) - Accès lecture
- **Direction Médicale** (Tableaux de bord activité) - Accès lecture agrégée

- **Analyses Statistiques** (Études de performance, charge de travail) - Accès lecture

Relations MDM :

- **Lié à MDM_Service** : Un médecin appartient à un service hospitalier (FK Service_ID)
- **Relation inverse** : Un médecin peut être responsable d'un service (lien via MDM_Service.Responsable_ID)

DATA DICTIONARY - MDM_MEDECIN

Colonne	Type	Description métier	Contraintes	Sensibilité	Source(s)	Règle de qualité
Medecin_ID	UUID	Identifiant unique universel du médecin dans le référentiel MDM. Généré automatiquement et immuable. Utilisé comme clé de référence dans tous les systèmes.	PRIMARY KEY, NOT NULL, UNIQUE, DEFAULT gen_random_uuid()	Faible	Généré par PostgreSQL (MDM)	Complétude = 100% (auto-généré)
Nom	VARCHAR (100)	Nom de famille du médecin. Normalisé en MAJUSCULES pour cohérence. Utilisé pour recherche et affichage dans interfaces.	NOT NULL, Longueur 2-100 caractères	Moyenne (Professionnel)	ERP (nom), RDV (parsing de nom_complet)	Complétude = 100%, Normalisation UPPER

Prénom	VARCHAR (100)	Prénom du médecin. Conserve la casse d'origine (majuscule initiale). Utilisé pour communication professionnelle.	NOT NULL, Longueur 2-100 caractères	Moyenne (Professionnel)	ERP (prenom), RDV (parsing de nom_complet)	Complétude = 100%, Capitalisation standard
Spécialité	VARCHAR (100)	Spécialité médicale principale du praticien (Cardiologue, Pédiatrie, etc.). Valeurs contrôlées par table de référence. Utilisée pour affectation RDV et analyses statistiques.	NOT NULL, Valeurs contrôlées (table ref_specialites), ENUM recommandé	Faible	ERP (specialite), RDV (specialite)	Complétude = 100%, Valeurs contrôlées via FK
Service_ID	UUID	Identifiant du service hospitalier d'affectation du médecin. Clé étrangère vers MDM_Service.	FOREIGN KEY (MDM_Service), NULL autorisé	Faible	ERP (service_affecte - matching par nom service)	Complétude cible ≥95%, Intégrité référentielle

		Définit le rattachement organisationnel.				
Num_Licence	VARCHAR (11)	Numéro RPPS (Répertoire Partagé des Professionnels de Santé). Identifiant national unique du praticien. Format : 11 chiffres. CRITIQUE pour validation légale d'exercice.	UNIQUE, NOT NULL, Format '^[0-9]{11}\$', Validation via API RPPS	Élevée (ID professionnel)	ERP (num_licence)	Complétude = 100%, Unicité = 100%, Validation API RPPS
Disponibilité	JSON	Horaires de disponibilité du médecin par jour de la semaine. Format JSON : {"lundi": "08:00-18:00", "mardi": "08:00-18:00", ...}. Utilisé pour planification RDV.	NULL autorisé, Format JSON valide, Schema validation	Faible	ERP (disponibilite_lundi, _mardi, etc. - agrégation)	Complétude cible ≥80%, Validation schema JSON

Email	VARCHAR (255)	Adresse email professionnelle du médecin (domaine @clinique.fr). Utilisée pour communications internes et notifications système.	NULL autorisé, Format email valide, Domaine @clinique.fr	Moyenne (Contact pro)	ERP (email_pro), RDV (email)	Complétude cible ≥95%, Validation domaine + regex
Téléphone	VARCHAR (20)	Numéro de téléphone professionnel du médecin. Normalisé à 10 chiffres. Utilisé pour contact urgences et coordination soins.	NULL autorisé, Format 10 chiffres	Moyenne (Contact pro)	ERP (telephone_pro), RDV (telephone)	Complétude cible ≥90%, Validation format

3. MDM_SERVICE

DATA CATALOGUE

Nom de la table : mdm.mdm_service

Description : Référentiel maître unique des services et départements de la clinique. Consolide la structure organisationnelle incluant services médicaux (Cardiologie, Pédiatrie, etc.), services techniques (Laboratoire, Imagerie) et services administratifs. Facilite la gestion hiérarchique, la planification et la traçabilité des activités par service.

Propriétaire métier : Jean Moreau - Directeur des Opérations

Propriétaire technique : Équipe Data Governance - Service IT Clinique

Fréquence de mise à jour : Hebdomadaire (batch dimanche 23h00) + Ponctuel lors de réorganisations

Volumétrie :

- Services actifs : 6 (Cardiologie, Pédiatrie, Radiologie, Chirurgie, Gynécologie, Urgences)
- Services historiques : Conservés pour traçabilité
- Croissance estimée : +1 nouveau service tous les 2-3 ans

Sensibilité :

- **FAIBLE** - Données organisationnelles publiques en interne
- Informations de contact publiques (standard téléphonique)

Sources de données :

1. **Système ERP** (source_erp_services.csv) - SOURCE UNIQUE
 - Colonnes : service_id_erp, nom_service, description, responsable_nom, localisation, horaires, téléphone, email
 - Qualité : Données complètes et validées Direction

Services consommateurs :

- **Système Rendez-vous** (Affichage services disponibles pour RDV) - Accès lecture
- **Dossier Patient** (Traçabilité services consultés) - Accès lecture
- **Module Planning** (Affectation médecins aux services) - Accès lecture
- **Facturation** (Imputation coûts par service) - Accès lecture
- **Direction** (Tableaux de bord par service, KPIs) - Accès lecture agrégée

- **Patients & Secrétariat** (Informations pratiques : localisation, horaires) - Accès lecture public

Relations MDM :

- **Relation avec MDM_Médecin** : Un service a plusieurs médecins affectés (FK depuis Médecin.Service_ID)
- **Relation hiérarchique** : Un service peut avoir un médecin comme responsable (Responsable_ID → Medecin_ID)
- **Relation indirecte avec MDM_Patient** : Services (labo, imagerie) produisent des données liées aux patients

DATA DICTIONARY - MDM_SERVICE

Colonne	Type	Description métier	Contraintes	Sensibilité	Source(s)	Règle de qualité
Service_ID	UUID	Identifiant unique universel du service dans le référentiel MDM. Généré automatiquement et immuable. Utilisé comme clé de référence dans tous les systèmes.	PRIMARY KEY, NOT NULL, UNIQUE, DEFAULT gen_random_uuid()	Faible	Généré par PostgreSQL (MDM)	Complétude = 100% (auto-générée)
Nom_Service	VARCHAR(100)	Nom officiel du service hospitalier (ex: "Cardiologie", "Laboratoire Biochimie"). Valeur unique servant de référence dans toute la clinique.	NOT NULL, UNIQUE, Longueur 3-100 caractères	Faible	ERP (nom_service)	Complétude = 100%, Unicité = 100%

Description	TEXT	Description détaillée du service : missions, types de soins, équipements disponibles. Utilisée pour documentation interne et information patients.	NULL autorisé, Longueur max 1000 caractères	Faible	ERP (description)	Complétude cible ≥90%
Responsible_ID	UUID	Identifiant du médecin responsable/chef de service. Clé étrangère vers MDM_Médecin. Définit le leadership médical du service.	FOREIGN KEY (MDM_Médecin), NULL autorisé	Faible	ERP (responsable_nom - matching vers MDM_Médecin)	Complétude cible ≥80%, Intégrité référentielle
Localisation	VARCHAR(200)	Localisation physique du service dans la clinique. Format standardisé : "Bâtiment	NULL autorisé, Format "Bâtiment X, Étage Y"	Faible (Public interne)	ERP (localisation)	Complétude cible ≥95%, Validation format

		[A-Z], Étage [1-9]". Utilisée pour orientatio n patients et staff.				
Horaires	JSON	Horaires d'ouvertur e du service par jour de la semaine. Format JSON : {"lundi": "08:00-18 :00", ...}. Utilisé pour informatio n patients et gestion des accès.	NULL autorisé, Format JSON valide, Schema validation	Faible (Public interne)	ERP (horaires_ouver ture_lundi, _mardi, etc. - agrégation)	Compl étude cible ≥90%, Validat ion schem a JSON
Date_Cré ation	TIMESTA MP	Date et heure de création de l'enregistre ment MDM. Utilisée pour audit et traçabilité organisati onnelle.	NOT NULL, DEFAULT CURRENT TI MESTAMP	Faible (Métad onnée)	Généré automatiqueme nt par MDM	Compl étude = 100% (auto)

ANNEXES

A. Glossaire des Termes Métier

MDM (Master Data Management) : Processus de gestion des données de référence permettant d'établir une source unique de vérité (Single Source of Truth - SSOT) pour les données critiques de l'entreprise.

Golden Record : Enregistrement de référence unique consolidant les meilleures données provenant de multiples sources, après dédoublonnage et enrichissement.

PII (Personally Identifiable Information) : Données personnelles permettant d'identifier directement ou indirectement une personne physique (nom, prénom, date de naissance, etc.).

HADS (Health and Social Care Data) : Données de santé à caractère personnel, protégées par réglementation stricte (RGPD Art. 9).

RPPS (Répertoire Partagé des Professionnels de Santé) : Identifiant national unique attribué à chaque professionnel de santé en France (11 chiffres).

Dédoublonnage : Processus d'identification et de fusion des enregistrements en doublon représentant la même entité réelle.

Score de Qualité : Indicateur numérique (0-100) mesurant la fiabilité d'un enregistrement basé sur complétude, exactitude, cohérence et fraîcheur.

UUID (Universally Unique Identifier) : Identifiant unique universel de 128 bits garantissant l'unicité globale sans coordination centrale.

B. Mapping Sources → MDM (Traçabilité)

MDM_PATIENT

Colonne MDM	Source RDV	Source ERP	Source FACT
Patient_ID	-	-	- (généré)
Nom	nom	nom	nom_famille
Prénom	prenom	prenom	prenoms
Date_Naissance	-	date_naissance	-
Sexe	sexe	-	-

Adresse	adresse	adresse	adresse_facturatio n
Téléphone	telephone	-	tel_contact
Email	-	email	email_contact
Num_Dossier	-	num_dossier	-
Historique_Médical	-	historique_medical + allergies	-

MDM_MEDECIN

Colonne MDM	Source RDV	Source ERP
Medecin_ID	-	- (généré)
Nom	nom_complet (parsing)	nom
Prénom	nom_complet (parsing)	prenom
Spécialité	specialite	specialite
Service_ID	-	service_affecte (matching)
Num_Licence	-	num_licence
Disponibilité	-	disponibilite_* (agrégation)
Email	email	email_pro
Téléphone	telephone	telephone_pro

MDM_SERVICE

Colonne MDM	Source ERP
Service_ID	- (généré)
Nom_Service	nom_service
Description	description
Responsable_ID	responsable_nom (matching)
Localisation	localisation
Horaires	horaires_ouverture_* (agrégation)

C. Règles de Dédoublement Patient

Deux enregistrements patients sont considérés comme doublons si **AU MOINS UNE** des conditions suivantes est vraie :

1. **(Nom similaire à 85%+) ET (Prénom similaire à 85%+) ET (Date naissance identique)**
 - Similarité calculée via distance de Levenshtein
2. **Téléphone identique (10 chiffres)**
 - Ignore doublons avec téléphone vide ou invalide
3. **Email identique (format valide)**
 - Comparaison insensible à la casse
4. **Num_Dossier identique**
 - Si renseigné dans les deux enregistrements

Fusion des doublons (Golden Record) :

- Nom : Valeur la plus fréquente ou la plus récente
 - Prénom : Valeur la plus fréquente
 - Date_Naissance : Valeur la plus ancienne (source ERP prioritaire)
 - Autres champs : Prioriser valeurs non vides, sinon source ERP > RDV > FACT
-

D. Calcul du Score de Qualité

Formule globale :

$$\text{Score_Qualité} = (\text{Complétude} \times 40\%) + (\text{Exactitude} \times 30\%) + (\text{Cohérence} \times 20\%) + (\text{Fraîcheur} \times 10\%)$$

Dimension 1 : Complétude (40%)

Complétude = (Nombre de champs obligatoires remplis / Nombre total de champs obligatoires) × 100

Champs obligatoires MDM_Patient : Nom, Prénom, Date_Naissance, Num_Dossier
Champs importants (bonus) : Sexe, Téléphone, Email, Adresse, Historique_Médical

Dimension 2 : Exactitude (30%)

Exactitude = Moyenne des validations réussies

Validations :

- Email : Format regex RFC 5322 valide ? (Oui=100, Non=0)
- Téléphone : Format 10 chiffres ? (Oui=100, Non=0)
- Date_Naissance : Cohérence temporelle (< aujourd'hui, > 1900) ? (Oui=100, Non=0)

- Num_Dossier : Format DOSS-XXXXXX ? (Oui=100, Non=0)

Dimension 3 : Cohérence (20%)

Cohérence = Respect des règles métier

Règles :

- Date_Naissance < Date_Création (Oui=100, Non=0)
- Sexe dans valeurs autorisées (Oui=100, Non=0)
- Num_Dossier unique dans MDM (Oui=100, Non=0)
- Au moins 1 ID source renseigné (RDV, ERP ou FACT) (Oui=100, Non=0)

Dimension 4 : Fraîcheur (10%)

Fraîcheur = Score basé sur ancienneté dernière mise à jour

Si dernière_maj < 7 jours : 100

Si 7 jours ≤ dernière_maj < 30 jours : 80

Si 30 jours ≤ dernière_maj < 180 jours : 60

Si 180 jours ≤ dernière_maj < 365 jours : 40

Si dernière_maj ≥ 365 jours : 20

Seuils d'interprétation :

- **Excellent** : Score \geq 90
 - **Bon** : $70 \leq$ Score < 90
 - **Moyen** : $50 \leq$ Score < 70
 - **Faible** : Score < 50 (nécessite action corrective)
-

E. Standards de Normalisation

Normalisation des Noms

1. Conversion MAJUSCULES : "dupont" → "DUPONT"
2. Suppression accents : "ÉLÉONORE" → "ELEONORE"
3. Suppression espaces multiples : "MARTIN PAUL" → "MARTIN PAUL"
4. Trim espaces début/fin

Table de remplacement des accents :

Caractère	Remplacement
é	é

É, È, Ê, Ë	E
À, Â, Ä	A
Ù, Û, Ü	U
Ô, Ö	O
Î, Ï	I
Ç	C

Normalisation des Dates

Formats sources détectés :

- DD/MM/YYYY (ERP) → Conversion vers YYYY-MM-DD
- YYYY-MM-DD (RDV) → Aucune conversion
- YYYY/MM/DD (FACT) → Remplacement "/" par "-"

Format cible : YYYY-MM-DD (ISO 8601)

Normalisation des Téléphones

1. Suppression caractères non-numériques : "06.12.34.56.78" → "0612345678"
2. Validation longueur = 10
3. Validation préfixe français (01-09)
4. Format stockage : "0612345678" (pas d'espaces ni séparateurs)

Normalisation des Emails

1. Conversion minuscules : "Jean.DUPONT@example.com" → "jean.dupont@example.com"
 2. Trim espaces
 3. Validation regex : `^[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}$`
 4. Vérification domaine @clinique.fr pour emails pro médecins
-

F. Métriques et KPIs à Suivre

Métriques de Qualité Données

Métrique	Calcul	Cible	Fréquence mesure
Score qualité moyen patients	AVG(score_qualite)	≥85	Quotidienne

Taux complétude Historique_Médical	COUNT(historique_medical IS NOT NULL) / COUNT(*) × 100	≥70 %	Hebdomadaire
Taux complétude Email	COUNT(email IS NOT NULL AND email VALID) / COUNT(*) × 100	≥70 %	Hebdomadaire
Taux complétude Téléphone	COUNT(telephone IS NOT NULL AND telephone VALID) / COUNT(*) × 100	≥90 %	Hebdomadaire
Taux doublons éliminés	(Nb enreg sources - Nb enreg MDM) / Nb enreg sources × 100	~25 %	Après chaque ETL
Taux validation RPPS médecins	COUNT(num_licence VALID via API) / COUNT(*) × 100	100 %	Hebdomadaire

Métriques de Performance ETL

Métrique	Cible	Fréquence mesure
Latence temps réel (patients)	<5 secondes	Continue (monitoring)
Taux de succès ETL	≥99%	Après chaque exécution
Durée traitement batch (médecins)	<10 minutes	Quotidienne
Durée traitement batch (services)	<5 minutes	Hebdomadaire
Taux d'erreur transformation	<1%	Après chaque ETL

Métriques d'Adoption

Métrique	Cible	Fréquence mesure
Nb applications consommatoires MDM	≥5	Mensuelle
Nb requêtes MDM/jour	Croissance continue	Quotidienne
Taux satisfaction utilisateurs	≥80%	Trimestrielle (enquête)

G. Plan de Gouvernance des Données

Rôles et Responsabilités

Rôle	Responsable	Responsabilités
Data Owner	Dr. Sophie Laroche (Directrice Médicale)	Validation règles métier, arbitrage qualité, approbation modifications structure
Data Steward	Équipe Data Governance IT	Exécution règles qualité, monitoring métriques, corrections anomalies
Data Custodian	Admin Base de Données	Maintenance technique, backups, sécurité accès
Data Consumer	Médecins, Secrétariat, Direction	Utilisation données, remontée anomalies

Processus de Modification

Ajout d'une nouvelle colonne MDM :

1. Demande formelle au Data Owner (justification métier)
2. Validation comité gouvernance (impact sources + consommateurs)
3. Mise à jour Data Catalogue + Data Dictionary
4. Développement ETL
5. Tests qualité
6. Déploiement + communication utilisateurs

Modification d'une règle métier :

1. Proposition au Data Owner
2. Impact analysis (nb enregistrements affectés)
3. Validation comité
4. Backfill données historiques si nécessaire
5. Documentation + formation

Comité de Gouvernance

Fréquence : Mensuelle (1er lundi du mois, 10h-11h)

Membres :

- Data Owner (présidence)
- Data Stewards (2 personnes IT)
- Représentant Direction
- Représentant Médecins
- DPO (Data Protection Officer) - pour RGPD

Ordre du jour type :

1. Revue métriques qualité du mois
2. Incidents/anomalies

3. Demandes d'évolution
4. Conformité RGPD
5. Actions correctives