

27/01/2026

Data Gouvernance

Documentation des données

Table des matières

| | | |
|-------|---|---|
| 1. | Objectifs de la phase OpenMetadata | 2 |
| 2. | Principes de modélisation dans OpenMetadata..... | 2 |
| 2.1. | Domaines | 2 |
| 2.2. | RGPD et informations personnelles | 2 |
| 2.3. | Tiering et niveaux Bronze / Silver / Gold | 2 |
| 3. | Catalogage des tables | 3 |
| 3.1 | Domaine gestion hospitalière | 3 |
| 3.1.1 | Table patients | 3 |
| 3.1.2 | Table staff | 4 |
| 3.1.3 | Table staff_schedule | 5 |
| 3.1.4 | Table consultations | 5 |
| 3.1.5 | Table services_weekly | 6 |
| 3.2 | Domaine gestion informatique | 7 |
| 3.2.1 | Table superset_validation_metrics et validation_history | 7 |
| 4. | Apports de la phase OpenMetadata | 8 |

1. Objectifs de la phase OpenMetadata

L'objectif de cette phase est de mettre en place un catalogue de données opérationnel pour les principales tables du projet de gouvernance hospitalière, afin de :

- Centraliser la documentation fonctionnelle et technique des tables.
- Organiser les données par domaines métier et domaines informatiques.
- Marquer les données sensibles avec des tags RGPD / PII (informations personnelles identifiables).
- Prioriser les objets via un Tier (importance métier) et un niveau Bronze / Silver / Gold (maturité / qualité).

Au total, 7 tables ont été cataloguées dans OpenMetadata, couvrant à la fois la gestion hospitalière et le pilotage de la qualité des données.

2. Principes de modélisation dans OpenMetadata

2.1. Domaines

Deux grands domaines ont été définis pour structurer la navigation :

- Domaine **Gestion hospitalière** : Regroupe les tables directement liées à l'activité de soins (patients, staff, consultations, indicateurs hebdomadaires).
- Domaine **Gestion informatique** : Regroupe les tables techniques et de pilotage de la qualité (métriques de validation, logs, référentiels techniques).

Ces domaines servent de référentiel : depuis un domaine, on accède directement aux tables correspondantes sans devoir connaître les noms techniques.

2.2. RGPD et informations personnelles

Pour la conformité et la gouvernance :

- Application de tags RGPD / PII sur les tables contenant des données personnelles
- Tagging de colonnes sensibles : identifiants, coordonnées, données de santé.

2.3. Tiering et niveaux Bronze / Silver / Gold

Deux axes de classification ont été mis en place :

- Tier (Tier 1, 2, 3)
 - o Tier 1 : tables critiques pour le métier et les KPIs.
 - o Tier 2 : tables importantes mais non vitales.
 - o Tier 3 : tables de support ou techniques.

- Niveau Bronze / Silver / Gold
 - o Bronze : données brutes ou techniques, peu transformées.
 - o Silver : données intégrées, nettoyées, utilisées pour le pilotage opérationnel.
 - o Gold : tables de référence ou sources de vérité pour les indicateurs stratégiques.

3. Catalogage des tables

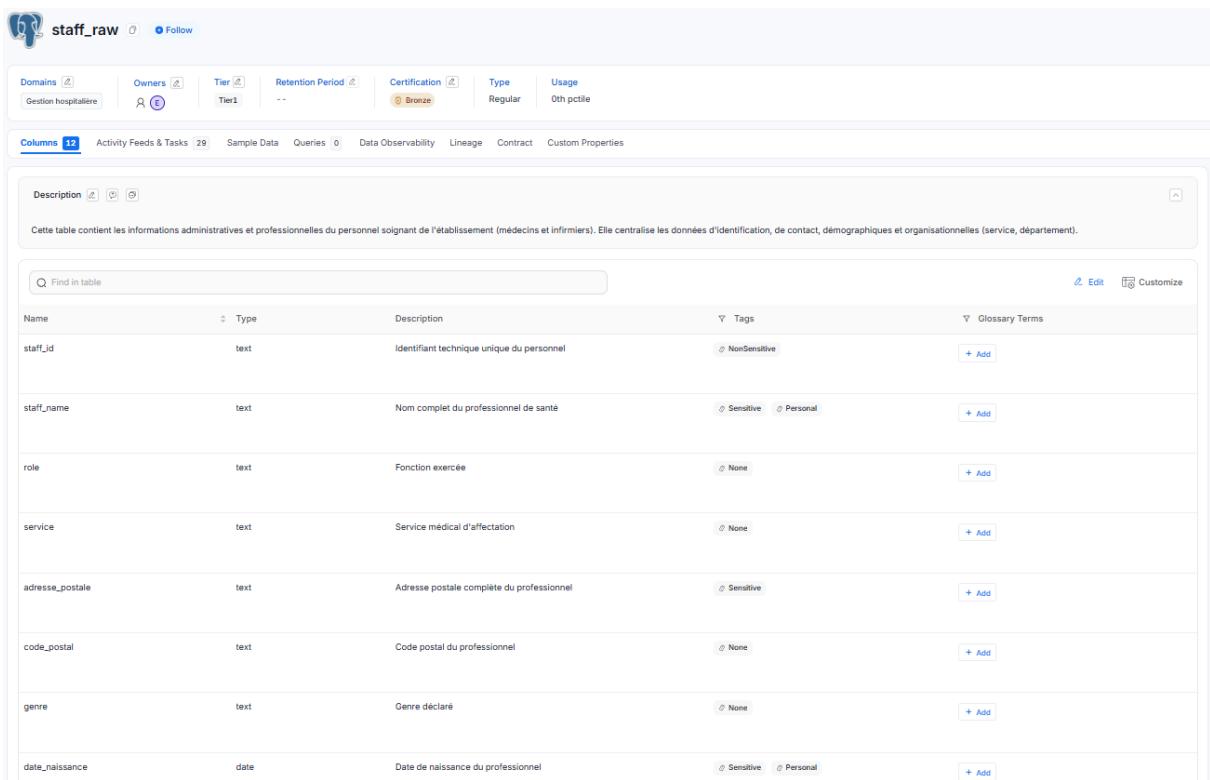
3.1 Domaine gestion hospitalière

3.1.1 Table patients

The screenshot shows a data catalog entry for the 'patients_raw' table. At the top, there are filters for Domains (Gestion hospitalière), Owners (E), Tier (Tier1), Retention Period (--), Certification (Bronze), Type (Regular), and Usage (0th pctl). Below the filters, it shows 14 columns, 36 activity feeds, and 0 queries. The table description states: "Cette table contient les informations administratives et cliniques des patients hospitalisés. Elle centralise les données d'identification, de localisation, de séjour (dates, service) et de contact des patients." The schema table has columns: Name, Type, Description, Tags, and Glossary Terms. The rows are:

| Name | Type | Description | Tags | Glossary Terms |
|------------|------------------|---|--|----------------|
| patient_id | text | Identifiant technique unique du patient | <input checked="" type="checkbox"/> NonSensitive | + Add |
| name | text | Nom complet du patient | <input checked="" type="checkbox"/> Sensitive <input checked="" type="checkbox"/> Personal | + Add |
| age | double precision | Âge du patient en années | <input checked="" type="checkbox"/> None | + Add |

3.1.2 Table staff



The screenshot shows a data catalog interface for a table named "staff_raw". The top navigation bar includes links for Domains, Owners, Tier, Retention Period, Certification, Type, and Usage. The table itself has 12 columns, with the first column being "Description". A note below the table states: "Cette table contient les informations administratives et professionnelles du personnel soignant de l'établissement (médecins et infirmiers). Elle centralise les données d'identification, de contact, démographiques et organisationnelles (service, département)." The table rows are as follows:

| Name | Type | Description | Tags | Glossary Terms |
|-----------------|------|---|--|----------------|
| staff_id | text | Identifiant technique unique du personnel | <input type="radio"/> NonSensitive | + Add |
| staff_name | text | Nom complet du professionnel de santé | <input type="radio"/> Sensitive <input type="radio"/> Personal | + Add |
| role | text | Fonction exercée | <input type="radio"/> None | + Add |
| service | text | Service médical d'affectation | <input type="radio"/> None | + Add |
| adresse_postale | text | Adresse postale complète du professionnel | <input type="radio"/> Sensitive | + Add |
| code_postal | text | Code postal du professionnel | <input type="radio"/> None | + Add |
| genre | text | Genre déclaré | <input type="radio"/> None | + Add |
| date_naissance | date | Date de naissance du professionnel | <input type="radio"/> Sensitive <input type="radio"/> Personal | + Add |

3.1.3 Table staff_schedule

The screenshot shows the details of the `staff_schedule_raw` table. At the top, there are navigation links for Domains (Gestion hospitalière), Owners (with a user icon), Tier (Tier1), Retention Period (--), Certification (Bronze), Type (Regular), and a Follow button. Below this is a header row with columns: Name, Type, Description, Tags, and Glossary Terms. The table contains the following data:

| Name | Type | Description | Tags | Glossary Terms |
|------------|---------|---|--|----------------|
| week | integer | Numeró de la semaine dans une annéé | <input type="radio"/> None | + Add |
| staff_id | text | Identifiant technique unique du personnel | <input type="radio"/> None | + Add |
| staff_name | text | Nom complet du professionnel de santé | <input type="radio"/> Sensitive <input type="radio"/> Personal | + Add |
| role | text | Fonction exercée | <input type="radio"/> None | + Add |
| service | text | Service médical d'affectation | <input type="radio"/> None | + Add |
| present | integer | 1 si le personnel est présent, sinon 0 | <input type="radio"/> None | + Add |

3.1.4 Table consultations

The screenshot shows the details of the `consultations_raw` table. At the top, there are navigation links for Domains (Gestion hospitalière), Owners (with a user icon), Tier (Tier1), Retention Period (--), Certification (Bronze), Type (Regular), and Usage (0th pctl). Below this is a header row with columns: Name, Type, Description, Tags, and Glossary Terms. The table contains the following data:

| Name | Type | Description | Tags | Glossary Terms |
|------------------|------------------------|---|---------------------------------|----------------|
| patientid | text | Identifiant technique unique du patient | <input type="radio"/> None | + Add |
| staffid | text | Identifiant technique unique du personnel | <input type="radio"/> None | + Add |
| consultationdate | date | Date de la consultation | <input type="radio"/> None | + Add |
| consultationtime | time without time zone | Horaire de la consultation | <input type="radio"/> None | + Add |
| description | text | Description du personnel de santé concernant cette consultation | <input type="radio"/> Sensitive | + Add |

3.1.5 Table services_weekly

services_weekly_raw [Follow](#)

Domains [Gestion hospitalière](#) Owners [Gestion hospitalière](#) Tier [Tier2](#) Retention Period [--](#) Certification [Gold](#) Type Regular Usage 0th pctile

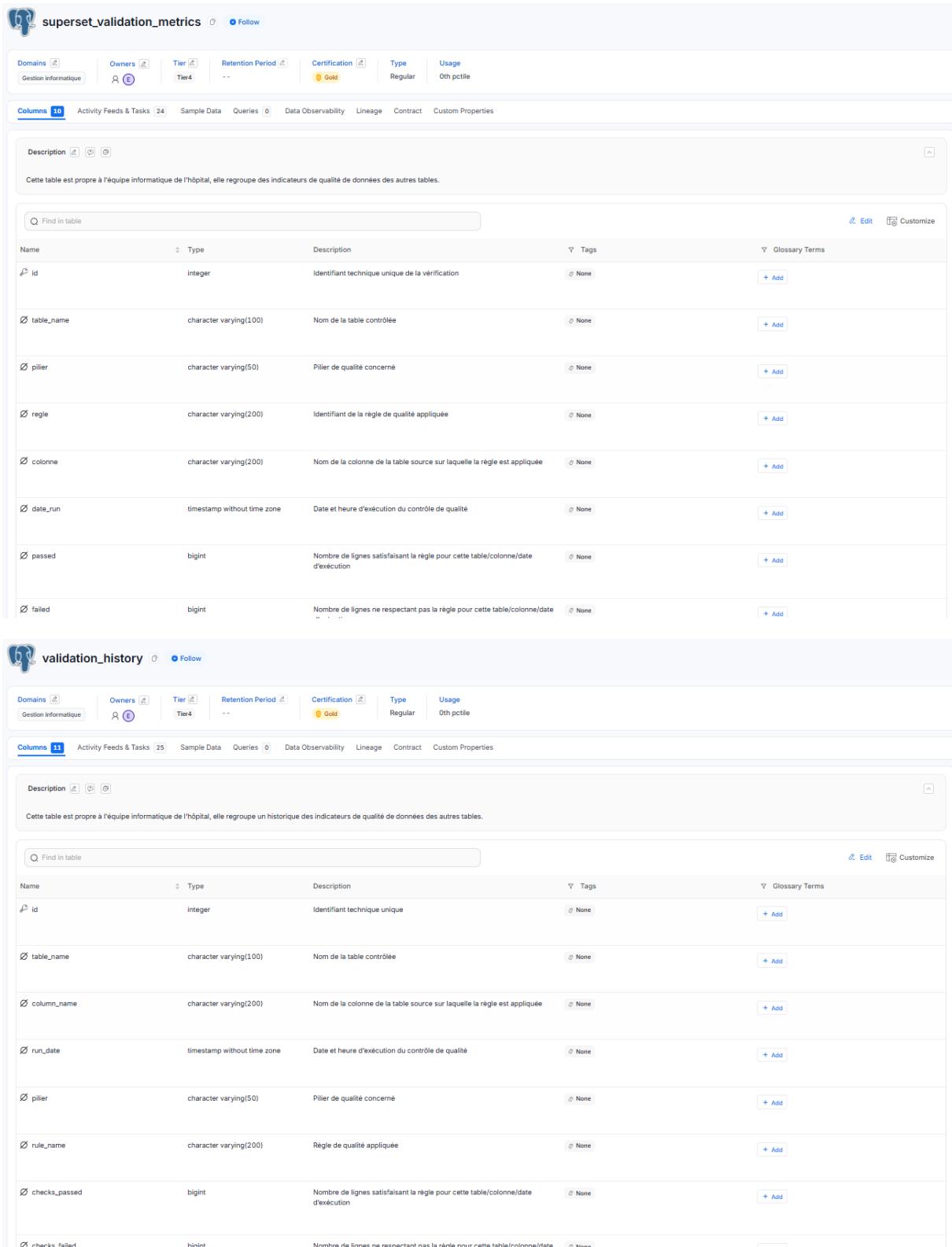
Columns 10 Activity Feeds & Tasks 25 Sample Data Queries 0 Data Observability Lineage Contract Custom Properties

Description [Table de suivi hebdomadaire des capacités et de la demande par service hospitalier \(lits disponibles, demandes, admissions, refus, satisfaction patients et moral du personnel\)](#)

| Name | Type | Description | Tags | Glossary Terms |
|------------------|---------|---|--|----------------|
| week | integer | Numéro de semaine dans l'année (1-52) pour laquelle les indicateurs sont calculés. | None + Add | |
| month | integer | Numéro de mois (1-12) correspondant à la semaine, pour faciliter l'aggrégation mensuelle. | None + Add | |
| service | text | Nom du service hospitalier concerné | None + Add | |
| availablebeds | integer | Nombre de lits disponibles dans le service pour la semaine considérée. | None + Add | |
| patientsrequest | integer | Nombre total de demandes/admissions sollicitées par les patients pour ce service et cette semaine | None + Add | |
| patientsadmitted | integer | Nombre de patients effectivement admis dans le service sur la semaine | None + Add | |
| patientsrefused | integer | Nombre de patients refusés (capacité insuffisante, critères médicaux, etc.) pour la semaine | None + Add | |

3.2 Domaine gestion informatique

3.2.1 Table superset_validation_metrics et validation_history

The screenshot shows two pages from the Superset interface. The top page is for the table 'superset_validation_metrics' (16 columns), which is a Gold-certified, regular table used for hospital IT management. It includes fields like id, table_name, pillar, rule, column, date_run, passed, and failed. The bottom page is for the 'validation_history' table (11 columns), also a Gold-certified, regular table. It includes fields like id, table_name, column_name, run_date, pillar, rule_name, checks_passed, and checks_failed. Both pages show detailed descriptions and a 'Find in table' search bar.

superset_validation_metrics (16 columns)

Description: Cette table est propre à l'équipe informatique de l'hôpital, elle regroupe des indicateurs de qualité de données des autres tables.

| Name | Type | Description | Tags | Glossary Terms |
|------------|-----------------------------|---|--------|----------------|
| id | integer | Identifiant technique unique de la vérification | + None | + Add |
| table_name | character varying(100) | Nom de la table contrôlée | + None | + Add |
| pillar | character varying(50) | Pilier de qualité concerné | + None | + Add |
| rule | character varying(200) | Identifiant de la règle de qualité appliquée | + None | + Add |
| column | character varying(200) | Nom de la colonne de la table source sur laquelle la règle est appliquée | + None | + Add |
| date_run | timestamp without time zone | Date et heure d'exécution du contrôle de qualité | + None | + Add |
| passed | bigint | Nombre de lignes satisfaisant la règle pour cette table/colonne/date d'exécution | + None | + Add |
| failed | bigint | Nombre de lignes ne respectant pas la règle pour cette table/colonne/date d'exécution | + None | + Add |

validation_history (11 columns)

Description: Cette table est propre à l'équipe informatique de l'hôpital, elle regroupe un historique des indicateurs de qualité de données des autres tables.

| Name | Type | Description | Tags | Glossary Terms |
|---------------|-----------------------------|---|--------|----------------|
| id | integer | Identifiant technique unique | + None | + Add |
| table_name | character varying(100) | Nom de la table contrôlée | + None | + Add |
| column_name | character varying(200) | Nom de la colonne de la table source sur laquelle la règle est appliquée | + None | + Add |
| run_date | timestamp without time zone | Date et heure d'exécution du contrôle de qualité | + None | + Add |
| pillar | character varying(50) | Pilier de qualité concerné | + None | + Add |
| rule_name | character varying(200) | Règle de qualité appliquée | + None | + Add |
| checks_passed | bigint | Nombre de lignes satisfaisant la règle pour cette table/colonne/date d'exécution | + None | + Add |
| checks_failed | bigint | Nombre de lignes ne respectant pas la règle pour cette table/colonne/date d'exécution | + None | + Add |

4. Apports de la phase OpenMetadata

La phase OpenMetadata a permis de :

- Rendre les tables clés facilement découvrables grâce aux domaines “Gestion hospitalière” et “Gestion informatique / KPIs”.
- Donner une vision claire de la sensibilité RGPD des données et des champs sensibles.
- Prioriser les efforts de qualité et d’industrialisation via les Tiers et les niveaux Bronze/Silver/Gold.