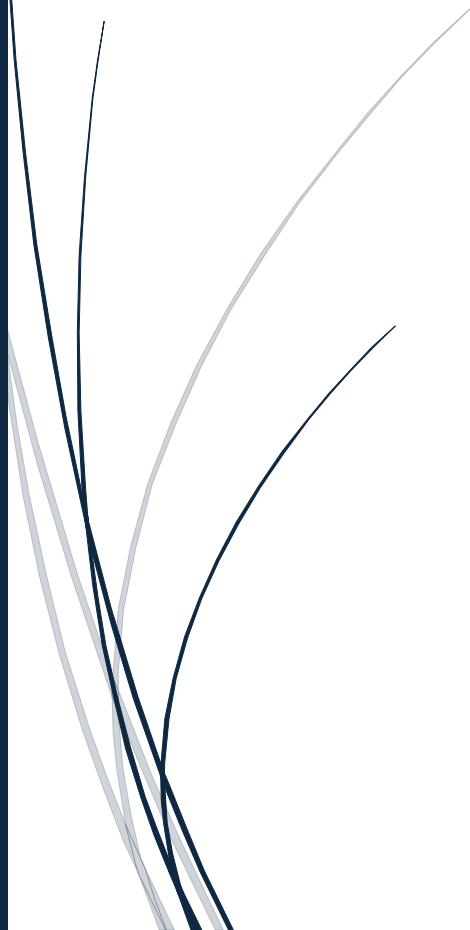


27/01/2026

Data Gouvernance

Documentation des données



Yousra Bouhanna, Abdeljebbar Abid, Elias Ait Hassou
MASTER 2 BUSINESS INTELLIGENCE & ANALYTICS – LYON2

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	4
3.1.1 Table patients	4
3.1.2 Table staff	5
3.1.3 Table staff_schedule	6
3.1.4 Table consultations	6
3.1.5 Table services_weekly	7
3.2 Domaine gestion informatique	8
3.2.1 Table superset_validation_metrics et validation_history	8
4. Apports de la phase OpenMetadata	9

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

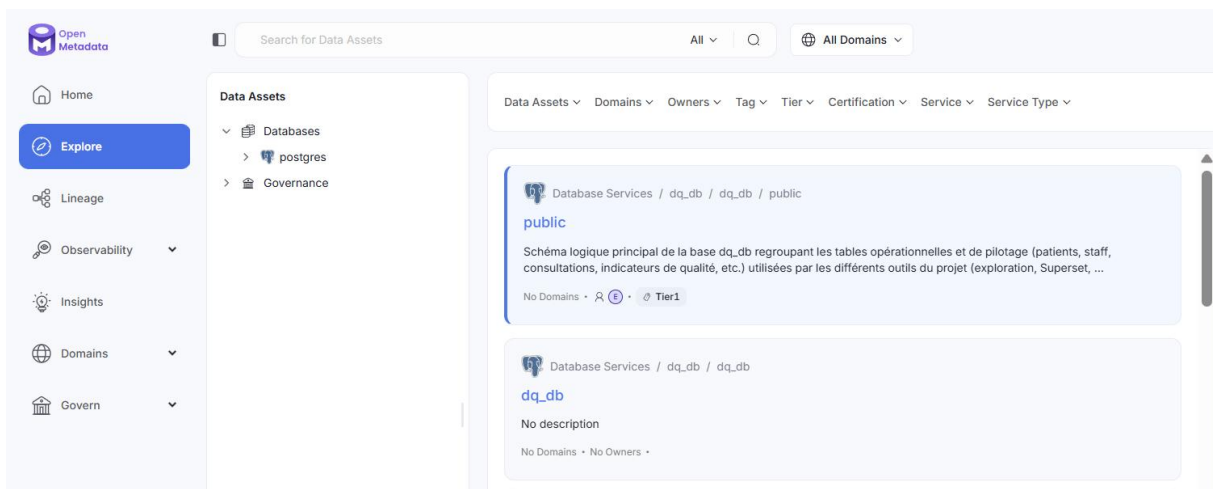
Deux axes de classification ont été mis en place :

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



3. Catalogage des tables


Afin de trouver ce catalogue de données sur l'outil OpenMetadata il faudra se rendre sur l'onglet « Explore » et entrer dans le schéma « public » de la base de données postgresql.





3.1 Domaine gestion hospitalière


3.1.1 Table patients


 **patients_raw**  [Follow](#)

Domains 

Owners 

Tier 



Retention Period 

Certification 

Type


Usage

Gestion hospitalière

  E

Tier1

--

 Bronze

Regular

0th pctile

Columns **14**

Activity Feeds & Tasks 36

Sample Data




Queries 0

Data Observability


Lineage





Contract

Custom Properties

Description   

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.

[Edit](#)  Customize

Name	Type	Description	Tags	Glossary Terms
patient_id	text	Identifiant technique unique du patient		+ Add
name	text	Nom complet du patient	 	+ Add
age	double precision	Âge du patient en années		+ Add

3.1.2 Table staff

staff_raw

Follow

Domains

Gestion hospitalière

Owners

Tier

Tier1

Retention Period

--

Certification

Bronze

Type

Regular

Usage

0th pctile

Columns

Activity Feeds & Tasks

29

Sample Data

Queries

0

Data Observability

Lineage

Contract

Custom Properties

Description

⌵

⌵

⌵

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).

Find in table

Edit

Customize

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

3.1.3 Table staff_schedule

staff_schedule_raw

Follow

Domains

Gestion hospitalière

Owners

Tier

Tier1

Retention Period

--

Certification

Bronze

Type

Regular

Columns

Activity Feeds & Tasks 18

Sample Data

Queries 0

Data Observability

Lineage

Contract

Custom Properties

Description

Cette table contient les informations relatives aux agendas de chaque personnel de l'hôpital. Pour chaque semaine de l'année, nous avons l'information si un personnel de l'hôpital est présent ou absent.

Find in table

Edit

Customize

Name	Type	Description	Tags	Glossary Terms
week	integer	Numéro de la semaine dans une année	None	+ Add
staff_id	text	Identifiant technique unique du personnel	None	+ Add
staff_name	text	Nom complet du professionnel de santé	Sensitive Personal	+ Add
role	text	Fonction exercée	None	+ Add
service	text	Service médical d'affectation	None	+ Add
present	integer	1 si le personnel est présent, sinon 0	None	+ Add

3.1.4 Table consultations

consultations_raw

Follow

Domains

Gestion hospitalière

Owners

Tier

Tier1

Retention Period

--

Certification

Bronze

Type

Regular

Usage

0th pctile

Columns

Activity Feeds & Tasks 14

Sample Data

Queries 0

Data Observability

Lineage

Contract

Custom Properties

Description

Cette table contient l'ensemble des informations relatives aux consultations enregistrées dans le système entre les patients et les professionnels de santé.

Find in table

Edit

Customize

Name	Type	Description	Tags	Glossary Terms
patientid	text	Identifiant technique unique du patient	None	+ Add
staffid	text	Identifiant technique unique du personnel	None	+ Add
consultationdate	date	Date de la consultation	None	+ Add
consultationtime	time without time zone	Horaire de la consultation	None	+ Add
description	text	Description du personnel de santé concernant cette consultation	Sensitive	+ Add

3.1.5 Table services_weekly

services_weekly_raw

Follow

Domains

Gestion hospitalière

Owners

E

Tier

Tier2

Retention Period

--

Certification

Gold

Type

Regular

Usage

0th pctile

Columns

19

Activity Feeds & Tasks

25

Sample Data

Queries

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)

Find in table


Edit

Customize

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

3.2 Domaine gestion informatique

3.2.1 Table superset_validation_metrics et validation_history




superset_validation_metrics

Follow

Domains

Gestion informatique

Owners



Tier

Tier4

Retention Period

--

Certification

Gold

Type

Regular

Usage

Oth pctile

Columns

Activity Feeds & Tasks

Sample Data

Queries

Data Observability

Lineage

Contract

Custom Properties

Description


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

Find in table

Edit

Customize

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




validation_history

Follow

Domains

Gestion informatique

Owners



Tier

Tier4

Retention Period

--

Certification

Gold

Type

Regular

Usage

Oth pctile

Columns

Activity Feeds & Tasks

Sample Data

Queries

Data Observability

Lineage

Contract

Custom Properties

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.

Find in table

Edit

Customize

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

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.