



Royaume du Maroc
UNIVERSITÉ MOHAMED V - RABAT
ECOLE NATIONALE SUPÉRIEURE D'INFORMATIQUE
ET D'ANALYSE DES SYSTÈMES

DATA Quality Project

Groupe Hotellerie

Filière : Génie de la Data (GD)

Réalisé par :
CHAOUA Wiam
OULHAJ Salwa
BOUGOUTAYA Hiba ALLAH

Encadré par :
PR DAOUDI Najiya

Année Scolaire 2024/2025

Remerciements

Nous tenons à exprimer notre gratitude envers toutes les personnes qui ont contribué à la réalisation de ce projet. Merci à notre encadrante, Pr. Najiya, pour son soutien et ses précieux conseils. Nous remercions également nos familles, amis, et collègues pour leur encouragement constant.

Résumé

Dans un monde où les entreprises reposent de plus en plus sur les données pour prendre des décisions stratégiques, la gestion de la qualité des données est un enjeu essentiel. Une mauvaise qualité des données, caractérisée par des informations incomplètes, incohérentes, obsolètes ou dupliquées, peut entraîner des erreurs coûteuses, des décisions erronées et une perte de compétitivité. Ce rapport se concentre sur l'importance de la gestion de la qualité des données dans l'industrie hôtelière, où des informations provenant de multiples sources (ERP, PMS, CRM, plateformes de réservation en ligne) doivent être centralisées et validées pour garantir leur intégrité. Nous avons exploré les trois parties clés de la gestion des données dans l'hôtellerie : les données client, produit et fournisseur. Nous avons démontré comment une gouvernance efficace des données, soutenue par des solutions de Master Data Management (MDM), peut améliorer la prise de décision, optimiser les processus et améliorer l'expérience client. Le rapport met en évidence l'importance d'une gestion proactive des données, incluant des contrôles automatisés et une documentation rigoureuse pour assurer la qualité, la fiabilité et la traçabilité des informations dans l'ensemble des systèmes et processus de l'organisation.

Abstract

In a world where businesses increasingly rely on data for strategic decision-making, data quality management has become a critical concern. Poor data quality, characterized by incomplete, inconsistent, outdated, or duplicate information, can lead to costly errors, incorrect decisions, and a loss of competitiveness. This report focuses on the importance of data quality management in the hospitality industry, where information from multiple sources (ERP, PMS, CRM, online booking platforms) must be centralized and validated to ensure its integrity. We examined three key areas of data management in hospitality : customer, product, and supplier data. The report demonstrates how effective data governance, supported by Master Data Management (MDM) solutions, can enhance decision-making, optimize processes, and improve the customer experience. It emphasizes the need for a proactive approach to data management, including automated controls and rigorous documentation to ensure data quality, reliability, and traceability across all organizational systems and processes.

Introduction Generale

Dans un monde où les entreprises s'appuient de plus en plus sur les données pour prendre des décisions stratégiques, la qualité des données (Data Quality) est devenue un enjeu fondamental pour garantir la fiabilité et l'efficacité des processus métier. Une mauvaise qualité des données, caractérisée par des informations incomplètes, incohérentes, obsolètes ou dupliquées, peut entraîner des erreurs coûteuses, une prise de décision erronée et une perte de compétitivité. Dans des secteurs tels que l'hôtellerie, où les données sur les réservations, les services, les tarifs et les clients proviennent de multiples sources (ERP, PMS, CRM, plateformes de réservation en ligne), assurer l'intégrité et l'homogénéité de ces informations est crucial pour offrir une expérience client fluide et maximiser les performances opérationnelles. La mise en place d'une gouvernance efficace des données repose sur plusieurs dimensions clés : complétude, exactitude, cohérence, fraîcheur et traçabilité des informations. Sans un cadre rigoureux de gestion des données, les entreprises risquent de se retrouver confrontées à des inefficacités, des pertes financières et une dégradation de la satisfaction client. C'est dans ce contexte que les solutions de Master Data Management (MDM) jouent un rôle central en permettant de centraliser, structurer et valider les données critiques d'une organisation, assurant ainsi leur qualité et leur fiabilité à travers tous les systèmes et processus internes. Une gestion efficace de la qualité des données ne se limite pas à une simple correction d'erreurs, mais implique une approche proactive intégrant des contrôles automatisés, une documentation rigoureuse (data catalogue, data lineage) et une implication continue des parties prenantes. Ainsi, investir dans une stratégie de Data Quality robuste est aujourd'hui un impératif pour les organisations souhaitant exploiter pleinement le potentiel de leurs données et prendre des décisions éclairées basées sur des informations fiables et précises.

Table de matières

Remerciements	2
Résumé	3
Abstract	4
Introduction Générale	5
Table des figures	8
1 Présentation de projet	9
1.1 Contexte de l'étude	9
1.2 Architecture	9
1.3 Conception des tables MDM e création de data catalogue et dictionnaire de données pour chaque table MDM	10
1.3.1 CLIENT	10
Problématique de Gouvernance des Données dans une Chaîne Hôtelière . .	10
Solution Proposée	11
Mise en Œuvre Technique	11
Rôle d'OpenMetadata dans la Gestion des Données Clients	12
Résultats Obtenus	13
1.3.2 PRODUITS	14
A-Objectif du MDM Produits	14
B-Les Services Hôteliers Couverts par le MDM Produits	15
C-Exemple de Flux de Données Sans et Avec MDM Produits	15
D- Provenance de données	16
E- ETL	16
F- Tables $mdm_{produits}$	17
G- Data catalogue	17
H- Dictionnaire de données	17
I- Data profile	18
J- Airflow	18
K- test : nom unifié	19
1.3.3 FOURNISSEURS	19
A- Problématiques Identifiées	19
B- Exemples Réels de Problèmes Rencontrés	20

C- Mise en place de la solution mdm	21
Catalogue de données	21
Dictionnaire de données	22
Data Profiling	23
Tests de Qualité des Données	24
Alertes et Monitoring	27
Conclusion	29

Table des figures

1.1	Solution proposée	13
1.2	Data catalog sur OpenMetadata	13
1.3	Data dictionary	14
1.4	categorie produit	15
1.5	prenance de données	16
1.6	ETL	16
1.7	Tables mdm _{produits}	17
1.8	data-catalog	17
1.9	MDM-produit	17
1.10	data-profile	18
1.11	Airflow	18
1.12	nom unifié	19
1.13	nom unifié	19

Présentation de projet

1.1 Contexte de l'étude

Le projet s'inscrit dans le cadre d'une stratégie de gouvernance des données pour un groupe d'hôtellerie gérant plusieurs établissements. L'industrie hôtelière repose fortement sur l'exploitation des données pour offrir des services personnalisés aux clients, optimiser les opérations, et améliorer la prise de décisions.

Les principales sources de données incluent :

- Les données clients (profils, préférences, historique de réservations).
- Les données des produits (services).
- Les données fournisseurs (contrats, approvisionnements).

Une gestion optimale de ces données est essentielle pour garantir leur qualité, leur disponibilité, et leur utilisation efficace dans un contexte multisites.

1.2 Architecture

Architecture Hybride (Collaboration)

- Principe : Une base MDM centralisée qui synchronise régulièrement les différents systèmes tout en maintenant des références locales.
- Avantages : Assure l'unicité des données tout en permettant une synchronisation entre

différents outils.

- Inconvénients : Nécessite une gouvernance forte pour assurer la cohérence.
- Pourquoi Recommandée ? : Idéale pour la gestion d'un groupe hôtellerie international pour gagner en équilibre entre contrôle et flexibilité.

1.3 Conception des tables MDM e création de data catalogue et dictionnaire de données pour chaque table MDM

1.3.1 CLIENT

Problématique de Gouvernance des Données dans une Chaîne Hôtelière

Dans l'industrie hôtelière, la gestion des données clients est un enjeu majeur. Les hôtels collectent des informations provenant de multiples sources, ce qui entraîne des défis importants en termes de gouvernance des données. Les principales problématiques rencontrées sont les suivantes :

- **Sources multiples de données** : Les données clients proviennent de diverses sources telles que les réservations en ligne, les appels téléphoniques, les interactions en personne, et les plateformes tierces. Cette multiplicité de sources rend difficile la centralisation et l'intégration des informations.
- **Données hétérogènes** : Les données sont souvent stockées dans des formats variés (CSV, JSON, bases de données relationnelles, etc.), ce qui complique leur traitement et leur analyse.
- **Absence de standardisation** : L'absence de règles de normalisation pour les données clients (comme les formats de téléphone, d'e-mails, etc.) entraîne des incohérences et des erreurs, rendant difficile l'exploitation efficace des données.
- **Impact sur les opérations et la prise de décision** : Sans une vue unifiée des données clients, il est difficile pour les hôtels de prendre des décisions stratégiques basées sur des informations fiables. Cela peut entraîner des inefficacités opérationnelles et une

dégradation de l'expérience client.

Solution Proposée

Pour résoudre ces problématiques, une solution de **Master Data Management (MDM)** a été mise en place, centrée sur la gestion des données clients. Cette solution repose sur trois piliers principaux :

- **Centralisation et Intégration des Données :**
 - Utilisation d'un système ETL (Extract, Transform, Load) : Les données provenant des différentes sources sont extraites, transformées et chargées dans un référentiel central.
 - Création d'une base de données maître : Les données clients sont stockées dans une base de données centralisée, appelée **MDM_Clients**.
- **Standardisation des Données :**
 - Mise en place de règles strictes de normalisation pour les champs critiques tels que les numéros de téléphone, les adresses e-mail, et les noms.
 - Validation automatique lors de la saisie des données.
- **Conformité et Sécurité des Données :**
 - Gestion des consentements clients conformément aux réglementations en vigueur (comme le RGPD).
 - Suivi des accès et gestion conforme des données.

Mise en Œuvre Technique

La mise en œuvre de cette solution repose sur une architecture technique robuste, intégrant des outils modernes de gestion des données. Voici les étapes clés de l'implémentation :

1. **Extraction des Données :** Les données clients sont extraites des différentes sources (CRM, site web, systèmes de réservation, etc.) à l'aide de scripts ETL.
2. **Transformation et Nettoyage des Données :** Les données sont transformées et nettoyées pour assurer leur conformité aux règles de normalisation.

3. **Chargement des Données dans le MDM** : Les données nettoyées et standardisées sont chargées dans la base de données maître **MDM_Clients**.
4. **Intégration avec OpenMetadata** : Le système est intégré avec **OpenMetadata** pour assurer la traçabilité et la gouvernance des données.

Rôle d'OpenMetadata dans la Gestion des Données Clients

OpenMetadata est une plateforme open source de gestion des métadonnées qui joue un rôle clé dans la gouvernance des données clients. Voici les principales fonctionnalités d'OpenMetadata utilisées dans ce projet :

- **Visibilité et Traçabilité des Données** : OpenMetadata offre une vue complète sur la provenance, l'utilisation et l'évolution des données clients. Cela permet de suivre les modifications apportées aux données et de comprendre leur impact sur les processus métier.
- **Gouvernance des Données** : La plateforme permet de gérer les accès aux données clients et de s'assurer que les données sont utilisées conformément aux politiques de sécurité et de conformité.
- **Documentation Automatique** : OpenMetadata automatise la documentation des sources de données, des schémas et des transformations appliquées aux données. Cela facilite la compréhension des données par les différentes équipes.
- **Recherche Améliorée** : La plateforme permet une recherche facile des données selon divers critères (nom, description, usage), ce qui accélère l'accès aux informations nécessaires.
- **Intégration avec des Outils Externes** : OpenMetadata se connecte à des bases de données, des outils ETL, des outils de BI, et d'autres systèmes pour centraliser la gestion des métadonnées.
- **Suivi des Changements** : La plateforme permet de suivre les modifications dans les métadonnées au fil du temps, ce qui est essentiel pour maintenir la qualité des données.
- **Qualité des Données** : OpenMetadata inclut des outils de validation et de nettoyage des données pour garantir que les données clients restent propres et fiables.

Résultats Obtenus

Grâce à cette solution, les résultats suivants ont été obtenus :

- **Vue Unifiée des Clients** : Les hôtels disposent désormais d'une vue unifiée et complète des clients.
- **Amélioration de la Qualité des Données** : Les erreurs et les incohérences dans les données clients ont été considérablement réduites.
- **Conformité et Sécurité** : Les données clients sont désormais gérées conformément aux réglementations en vigueur.
- **Optimisation des Opérations** : La centralisation et la standardisation des données clients ont permis d'optimiser les processus opérationnels.

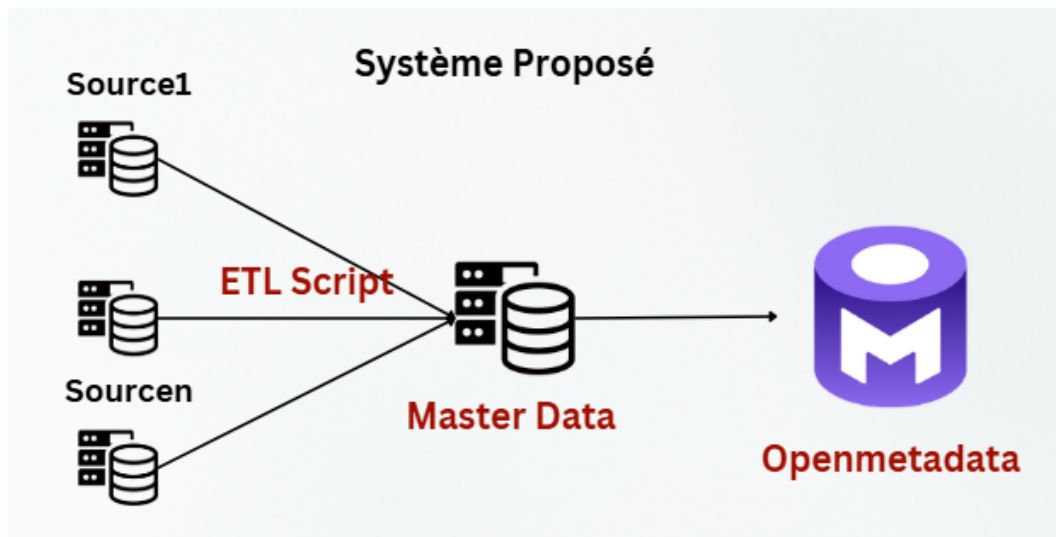


FIGURE 1.1 – Solution proposée

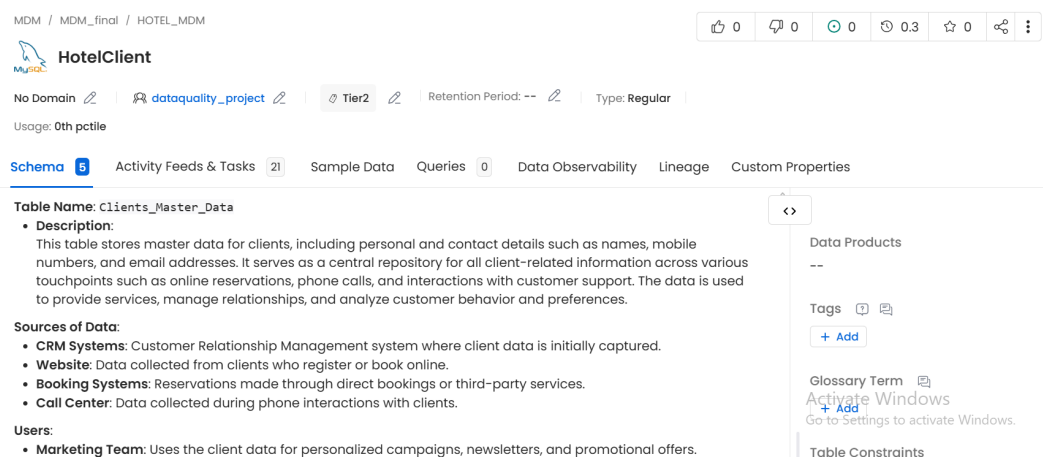


FIGURE 1.2 – Data catalog sur OpenMetadata

NAME	TYPE	DESCRIPTION	TAGS	GLOSSARY TERMS
ClientID	int	Identifiant du client (UNIQUE)	+ Add	+ Add
First_name	varchar(50)	Prénom du client	+ Add	+ Add

FIGURE 1.3 – Data dictionnary

1.3.2 PRODUITS

Dans un groupe hôtelier, chaque hôtel peut proposer une large gamme de services (chambres, restauration, spa, événements, etc.), souvent gérés de manière décentralisée dans des systèmes distincts (ERP, PMS, CRM, plateformes de réservation). Cela crée des problèmes de cohérence et d'exploitation des données, notamment :

- **Incohérences entre établissements** : Un même service (ex. "Massage détente") peut avoir un nom ou un prix différent selon l'hôtel.
- **Difficulté de mise à jour** : Les offres et services évoluent (nouvelles offres, promotions, changements de prix), mais les modifications ne sont pas centralisées.
- **Problèmes d'accès aux données** : Chaque département (opérations, marketing, finance) utilise des sources différentes, rendant la consolidation difficile.

A-Objectif du MDM Produits

Le but du MDM Produits est de :

- Centraliser et standardiser les services proposés pour garantir une information uniforme.
- Faciliter la gestion et la mise à jour des services en temps réel sur tous les canaux (site web, application mobile, agence de voyages en ligne).
- Optimiser l'expérience client en proposant des services adaptés selon l'historique et les préférences.

B-Les Services Hôteliers Couverts par le MDM Produits

Le MDM Produits en hôtellerie ne concerne pas des "produits" physiques mais bien des services proposés aux clients. Ces services peuvent être regroupés dans les catégories suivantes :

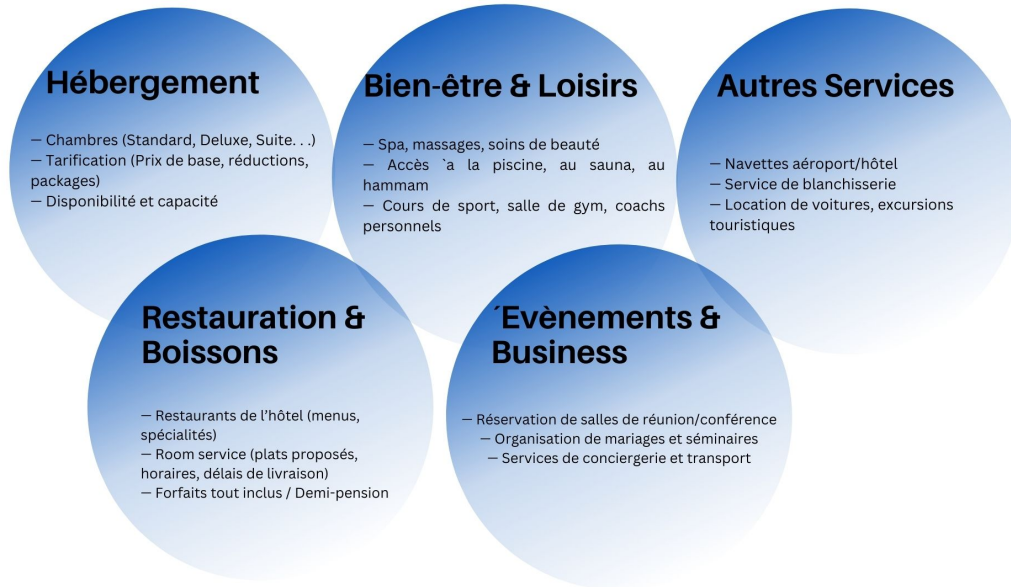


FIGURE 1.4 – categorie produit

Ces services doivent être bien définis et catégorisés dans le MDM Produits pour assurer une gestion efficace.

C-Exemple de Flux de Données Sans et Avec MDM Produits

- **Avant MDM (Système Fragmenté)**
 - Un client consulte les services disponibles sur le site web d'un hôtel.
 - Les informations proviennent du PMS (Property Management System), mais elles ne sont pas synchronisées avec celles du CRM.
 - Le client appelle pour réserver un massage, mais l'accueil n'a pas les bonnes informations sur la disponibilité.
 - Un problème de facturation survient car le prix dans l'ERP est différent de celui affiché en ligne.
- **Après MDM (Système Unifié)**
 - Tous les services sont stockés et gérés dans une base MDM centralisée.

- Les canaux de vente (site web, app mobile, CRM, front desk) récupèrent les informations en temps réel.
- Les modifications (ex. changement de prix d'un massage) sont appliquées immédiatement et partout.
- L'expérience client est améliorée car les offres sont personnalisées et cohérentes sur tous les supports.

D- Provenance de données

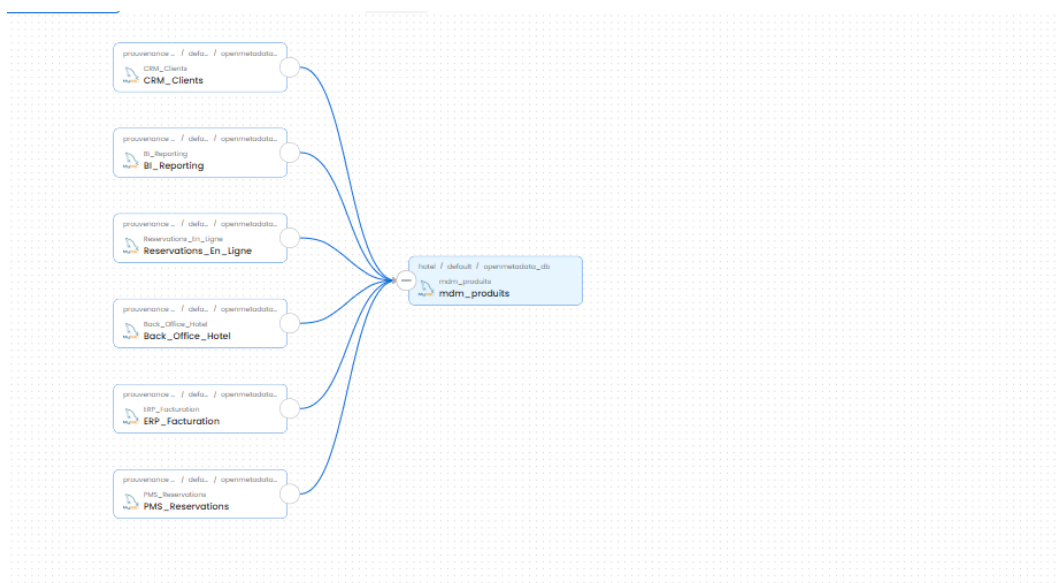


FIGURE 1.5 – prenanace de données

E- ETL

```

20 # * Garder uniquement les colonnes communes
21 df_pms = df_pms[colonnes_communes]
22 df_erp = df_erp[colonnes_communes]
23 df_crm = df_crm[colonnes_communes]
24 df_reservations = df_reservations[colonnes_communes]
25 df_bi = df_bi[colonnes_communes]
26
27 # * Fusionner tous les datasets
28 df_produits = pd.concat([df_pms, df_erp, df_crm, df_reservations, df_bi], ignore_index=True)
29
30 # * Supprimer les doublons et nettoyer les valeurs manquantes
31 df_produits.drop_duplicates(inplace=True)
32 df_produits.fillna("Inconnu", inplace=True)
33
34 # * Sauvegarder les données consolidées
35 df_produits.to_csv("MDM_Produits.csv", index=False)
36 print(df_produits.head())

```

FIGURE 1.6 – ETL

F- Tables $mdm_{produits}$

```
mysql> select*from mdm_produits;
```

Produit_ID	Nom_Produit	Type_Produit	Categorie	Prix	Disponibilite	Description
Date_Ajout	Date_Mise_A_Jour	Fournisseur_ID				
1	Laptop X100	lectronique	Ordinateurs	799.99	1	Laptop performant avec process
eur Intel 17.	2025-02-24 22:16:08	2025-02-24 22:16:08				
2	Smartphone Y200	lectronique	Tlphones	499.99	1	Smartphone Android avec cran A
MOLED.	2025-02-24 22:16:08	2025-02-24 22:16:08				
3	Casque Z500	lectronique	Audio	199.99	1	Casque Bluetooth avec rduction
de bruit.	2025-02-24 22:16:08	2025-02-24 22:16:08				
4	Tablette T300	lectronique	Tablettes	399.99	1	Tablette 10 pouces avec 128 Go
de stockage.	2025-02-24 22:16:08	2025-02-24 22:16:08				
5	Imprimante L400	Bureautique	Imprimantes	149.99	1	Imprimante laser multifonction
.	2025-02-24 22:16:08	2025-02-24 22:16:08				
6	Souris Gamer M600	Accessoires	Priphriques	59.99	1	Souris gaming RGB avec capteur
16K DPI.	2025-02-24 22:16:08	2025-02-24 22:16:08				
7	Clavier Mcanique K800	Accessoires	Priphriques	89.99	1	Clavier mcanique avec switches
Cherry MX.	2025-02-24 22:16:08	2025-02-24 22:16:08				
8	Disque Dur Externe H1TB	Stockage	Disques Durs	99.99	1	Disque dur externe 1 To USB 3.
.	2025-02-24 22:16:08	2025-02-24 22:16:08				

FIGURE 1.7 – Tables $mdm_{produits}$

G- Data catalogue

hotel / default / hotel

mdm_produits

Schema Activity Feeds & Tasks 24 Sample Data Queries 0 Data Observability Lineage Custom Properties

Description

La **référence MDM Produits** centralise et standardise la gestion des services proposés au sein de notre groupe hôtelier. Elle regroupe l'ensemble des prestations disponibles dans nos établissements, telles que l'hébergement, la restauration, les services de bien-être et les activités annexes.

Ce MDM garantit la **qualité**, la **cohérence** et l'**accessibilité** des données produits pour les équipes **opérationnelles**, **marketing** et **financières**, assurant ainsi une gestion efficace et une meilleure expérience client.

Provenance des données :

- PMS_Reservations (Property Management System)
- ERP_Facturation (Enterprise Resource Planning)
- CRM_Clients (Customer Relationship Management)
- Reservations_En_Ligne (Systèmes de Réservation Externes)
- Back_Office_Hotel (Gestion Opérationnelle)
- Administration & BI_Reporting (Analyse de Performance)

Utilisateurs de la Table MDM_Produits :

- Utilisateurs Internes (Employés & Systèmes Internes)
- Utilisateurs Externes (Clients & Partenaires)
- Systèmes & Applications Consommateurs

Niveaux d'Accès pour MDM_Produits

- **Accès complet** : Le **Chef de Projet**, responsable des données, peut créer, modifier et supprimer les services, gérer les permissions et assurer la cohérence des informations.
- **Accès restreint** : Les **Ingénieurs Data** et **Ingénieurs Sécurité** peuvent modifier certaines informations comme les prix, la disponibilité et la description des services, mais ne peuvent ni ajouter ni supprimer un service.
- **Lecture seule** : Les **Clients** et **Fournisseurs** peuvent uniquement consulter les services disponibles, leurs prix et leur disponibilité, sans possibilité de modification.

FIGURE 1.8 – data-catalog

H- Dictionnaire de données

hotel / default / operatordata_db

mdm_produits

no domain Data No Tier Retention period Type Regular Usage 0th pctile

Schema Activity Feeds & Tasks 57 Sample Data Queries 0 Data Observability Lineage Custom Properties

Column	Type	Description	Tags	Glossary Terms	Data Quality
Product_ID	int	Identifiant unique du service	Sensitive		--
Nom_Produit	varchar(255)	Nom du service (Ex: "Massage relax")	Sensitive		--
Type_Produit	varchar(100)	Type du produit (Ex: "Spa", "Restauration")			--
Categorie	varchar(100)	Catégorie du service (Ex: "Bien-être", "Hébergement")			--
Prix	decimal(10,2)	No Description	Sensitive		--
Disponibilite	tinyint	No Description			--
Description	text	No Description			--
Date_Ajout	datetime	No Description	Sensitive		--
Date_Mise_A_Jour	datetime	No Description	Sensitive		--
Fournisseur_ID	int	No Description	Sensitive		--

FIGURE 1.9 – MDM-produit

I- Data profile

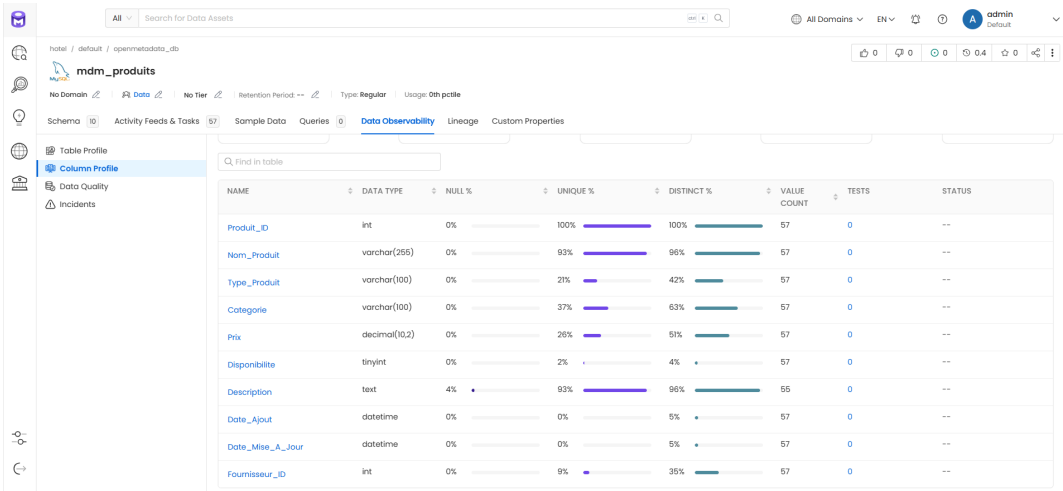


FIGURE 1.10 – data-profile

J- Airflow

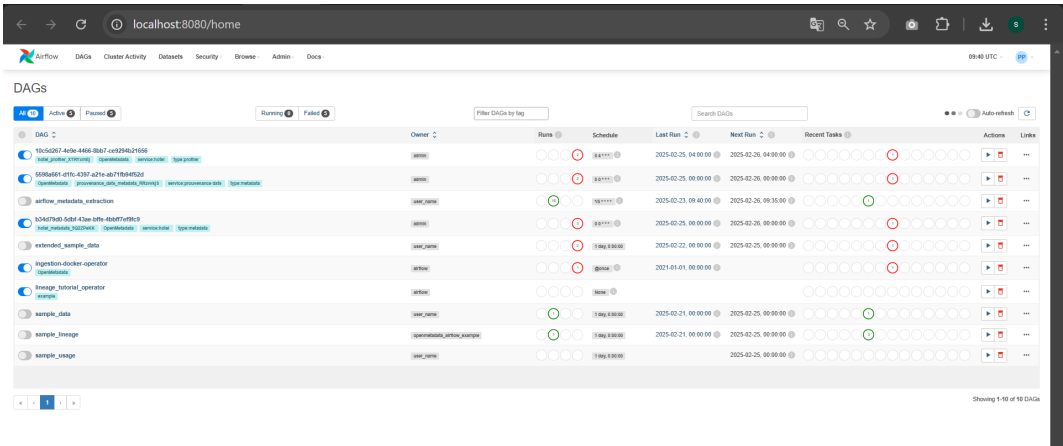


FIGURE 1.11 – Airflow

K- test : nom unifié

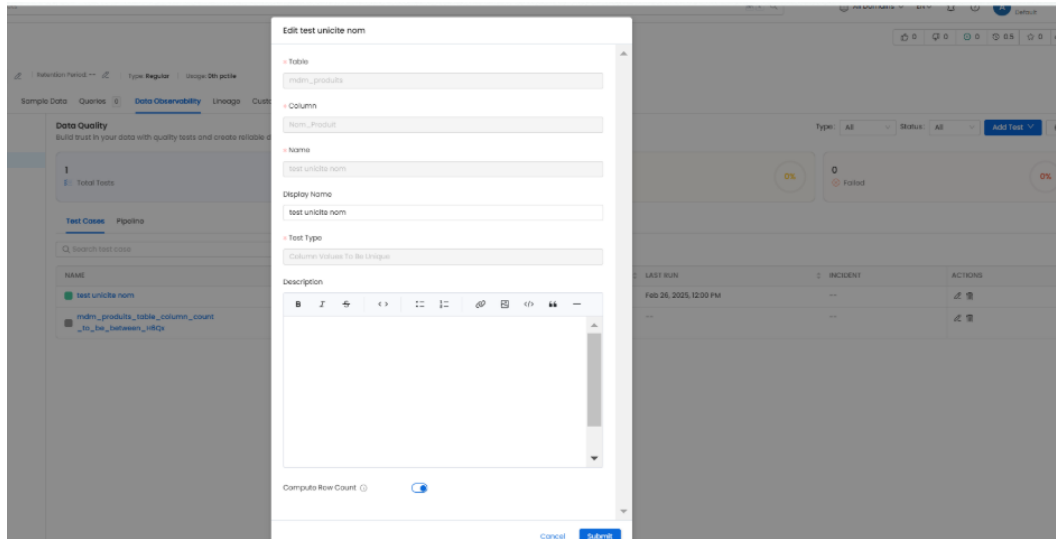


FIGURE 1.12 – nom unifié

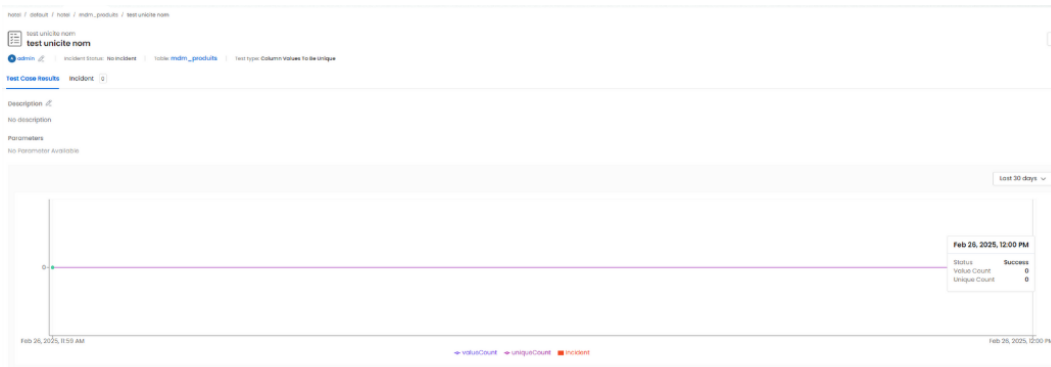


FIGURE 1.13 – nom unifié

1.3.3 FOURNISSEURS

A- Problématiques Identifiées

2.1 Multiples Sources de Données Non Synchronisées

- ERP (Oracle Hospitality, Opera PMS) : Gestion des achats avec des données parfois incohérentes avec d'autres systèmes.
- CRM (Salesforce, HubSpot, etc.) : Gestion relationnelle avec absence de normalisation et doublons.
- Système comptable (Sage X3, SAP Finance, QuickBooks) : Données fournisseurs non harmonisées.

2.2 Risques Financiers et Opérationnels

- Doublons fournisseurs : Paiements en double, fragmentation des achats.
- Absence de validation légale : Risques de fraude et d'amendes.
- Contacts fournisseurs obsolètes : Ruptures d'approvisionnement.
- Mauvaise classification des fournisseurs : Opportunités commerciales perdues.

B- Exemples Réels de Problèmes Rencontrés

3.1 Doublons de Fournisseurs et Erreurs de Paiement

Problème : "fournisseur X" apparaît différemment dans Opera PMS et Sage X3, menant à un paiement en double.

Impact : Perte financière, absence de remises de volume.

Solution MDM : Attribution d'un ID unique et base centralisée.

3.2 Problème de Conformité et Risque de Fraude

Problème : Un fournisseur sans numéro fiscal valide met en péril la conformité fiscale de l'hôtel.

Impact : Risque d'amendes et d'image ternie.

Solution MDM : Vérification automatique des registres CNSS, OMPIC.

3.3 Contact Fournisseur Obsolète et Rupture de Stock

Problème : Une commande de boissons VIP échoue car le contact fournisseur a quitté l'entreprise.

Impact : Achat d'urgence à prix plus élevé.

Solution MDM : Mise à jour automatique des contacts.

3.4 Mauvaise Catégorisation des Fournisseurs

Problème : Un fournisseur de caviar est oublié lors des négociations stratégiques.

Impact : Coût d'approvisionnement plus élevé.

Solution MDM : Catégorisation correcte et segmentation stratégique.

C- Mise en place de la solution mdm

Pour résoudre ces problématiques, une solution MDM a été mise en place en utilisant **Open-Metadata** avec les éléments suivants :

Catalogue de données : Recense les metadonnées business des tables MDM.

Dictionnaire de données : Définit les termes normalisés pour éviter les incohérences et assurer une gouvernance efficace.

Intégration des données : Utilisation d'Ingestion Data pour intégrer nos données à Open-Metadata.

Automatisation et contrôle qualité : vérification périodique de la qualité des données en utilisant des règles définies.

Catalogue de données

The screenshot shows the Open-Metadata interface for the 'mdm_hotel' domain, specifically the 'MDM_fournisseurs' catalog. The interface includes a navigation bar with 'Tables' (3), 'Stored Procedures' (0), 'Activity Feeds' (6), and 'Custom Properties'. Below the navigation bar, there is a description of the catalog: 'les tables master data pour les fournisseurs'. A table with two columns, 'NAME' and 'DESCRIPTION', lists the master data tables. The table is currently not deleted, as indicated by the 'Deleted' toggle switch.

NAME	DESCRIPTION
mdm_categorie_fournisseur categories	Cette table contient la classification des fournisseurs en fonction de leur secteur d'activité. Une bonne catégorisation facilite la gestion des achats et le suivi des relations fournisseurs
mdm_fournisseur fournisseurs	Cette table centralise les informations des fournisseurs du groupe hôtelier, en consolidant les données provenant de différentes sources (ERP, CRM, achats, facturation, etc.).
mdm_fournisseur_contact contacts fournisseurs	Cette table stocke les contacts professionnels des fournisseurs pour faciliter la communication entre les services achats et les fournisseurs

Dictionnaire de données

mdm_fournisseur

fournisseurs

No Domain

wiam

Tier1

Retention Period: --

Type: Regular

Usage: 0th pctlle

Schema 12

Activity Feeds & Tasks 514

Sample Data

Queries 0

Data Observability

Lineage

Custom Properties

Description

Cette table centralise les informations des fournisseurs du groupe hôtelier, en consolidant les données provenant de différentes sources (ERP, CRM, achats, f

Find in table

NAME	TYPE	DESCRIPTION	TAGS
id_fournisseur	int	Identifiant unique du fournisseur dans le MDM.	NonSensitive
nom	varchar(255)	Nom officiel du fournisseur, correspondant à son enregistrement légal.	Sensitive Personal
id_categorie	int	Référence à la catégorie du fournisseur, pour structurer les achats et les relations fournisseurs.	NonSensitive
numero_fiscal	varchar(50)	Numéro d'identification fiscale unique, permettant d'éviter les doublons et assurer la conformité réglementaire	Sensitive Personal
cnss	varchar(50)	Numéro d'affiliation à la CNSS (optionnel pour les fournisseurs soumis à cotisation sociale au Maroc	Sensitive Personal
registre_commerce	varchar(50)	Numéro d'immatriculation au registre du commerce, garantissant la légitimité de l'entreprise	Personal
registre_commerce	varchar(50)	Numéro d'immatriculation au registre du commerce, garantissant la légitimité de l'entreprise	Personal
adresse	text	Adresse complète du siège social ou de l'établissement principal du fournisseur	None
pays	varchar(100)	Pays de localisation du fournisseur, utile pour la gestion fiscale et douanière	None
ville	varchar(100)	Ville où est situé le fournisseur	None
telephone	varchar(20)	Numéro de contact principal du fournisseur	Sensitive Personal
email	varchar(255)	Adresse email principale pour les communications avec le fournisseur	Sensitive Personal
date_creation	date	Date d'ajout du fournisseur dans le système MDM	None

mdm_categorie_fournisseurcategories

No Domain | wiam | Tier2 | Retention Period: -- | Type: Regular | Usage: 0th pctlle

Schema 3 | Activity Feeds & Tasks 28 | Sample Data | Queries 0 | Data Observability | Lineage | Custom Properties

Description

Cette table contient la classification des fournisseurs en fonction de leur secteur d'activité. Une bonne catégorisation facilite la gestion des achats et le suivi des relations fournisseurs

Find in table

NAME	TYPE	DESCRIPTION	TAGS	GLOSSARY TERMS
id_categorie	int	Identifiant unique de la catégorie fournisseur	None	+ Add
nom	varchar(255)	Nom de la catégorie fournisseur (ex: Produits alimentaires, Maintenance, Services, Technologie, etc.)	NonSensitive	+ Add
description	text	Description détaillée des types de fournisseurs appartenant à cette catégorie	NonSensitive	+ Add

mdm_fournisseur_contactcontacts fournisseurs

No Domain | wiam | Tier2 | Retention Period: -- | Type: Regular | Usage: 0th pctlle

Schema 6 | Activity Feeds & Tasks 190 | Sample Data | Queries 0 | Data Observability | Lineage | Custom Properties

Description

Cette table stocke les contacts professionnels des fournisseurs pour faciliter la communication entre les services achats et les fournisseurs

Find in table

NAME	TYPE	DESCRIPTION	TAGS
id_contact	int	Identifiant unique du contact	NonSensitive
id_fournisseur	int	Référence du fournisseur associé au contact	NonSensitive
nom	varchar(255)	Nom du contact au sein du fournisseur	Sensitive, Personal
fonction	varchar(255)	Rôle du contact dans l'entreprise (ex: Responsable achats, Commercial, Comptable, etc.)	Personal
telephone	varchar(20)	Numéro de téléphone du contact.	Sensitive, Personal
email	varchar(255)	Adresse email du contact pour les échanges professionnels	Sensitive, Personal

Data Profiling

On a activé le Data Profiling dans OpenMetadata pour analyser la qualité des données contenues dans la table fournisseur. Et J'ai configuré un pipeline pour l'analyse automatique chaque heure (ingestion chaque heure)

MDM_HOTEL_profiler_jsi2tErRprofiler

Toutes les heures
Tous les jours

SUCCESS

ACTIVE

Pause

Journal

Suite à l'exécution du Data Profiling, nous avons obtenu les métriques suivantes :

MDM_HOTEL / mdm_hotel / MDM_fournisseurs

mdm_fournisseur

fournisseurs

Pas de Domaine | wlam | Tier1 | Retention Period: -- | Type: Regular | Utilisation: 0th pctile

Schéma 12 | Flux d'Activité & Tâches 514 | Échantillon de Données | Requêtes 0 | Data Observability | Lignage | Propriétés Personnalisées

Table Profil

Profil Colonnes

Entité de Données Qualité

Incidents

Profil de Colonne

Surveillez et comprenez la structure de vos colonnes avec le profilage

Décompte de Ligne

27

Décompte Colonnes

12

Échantillon du Profil

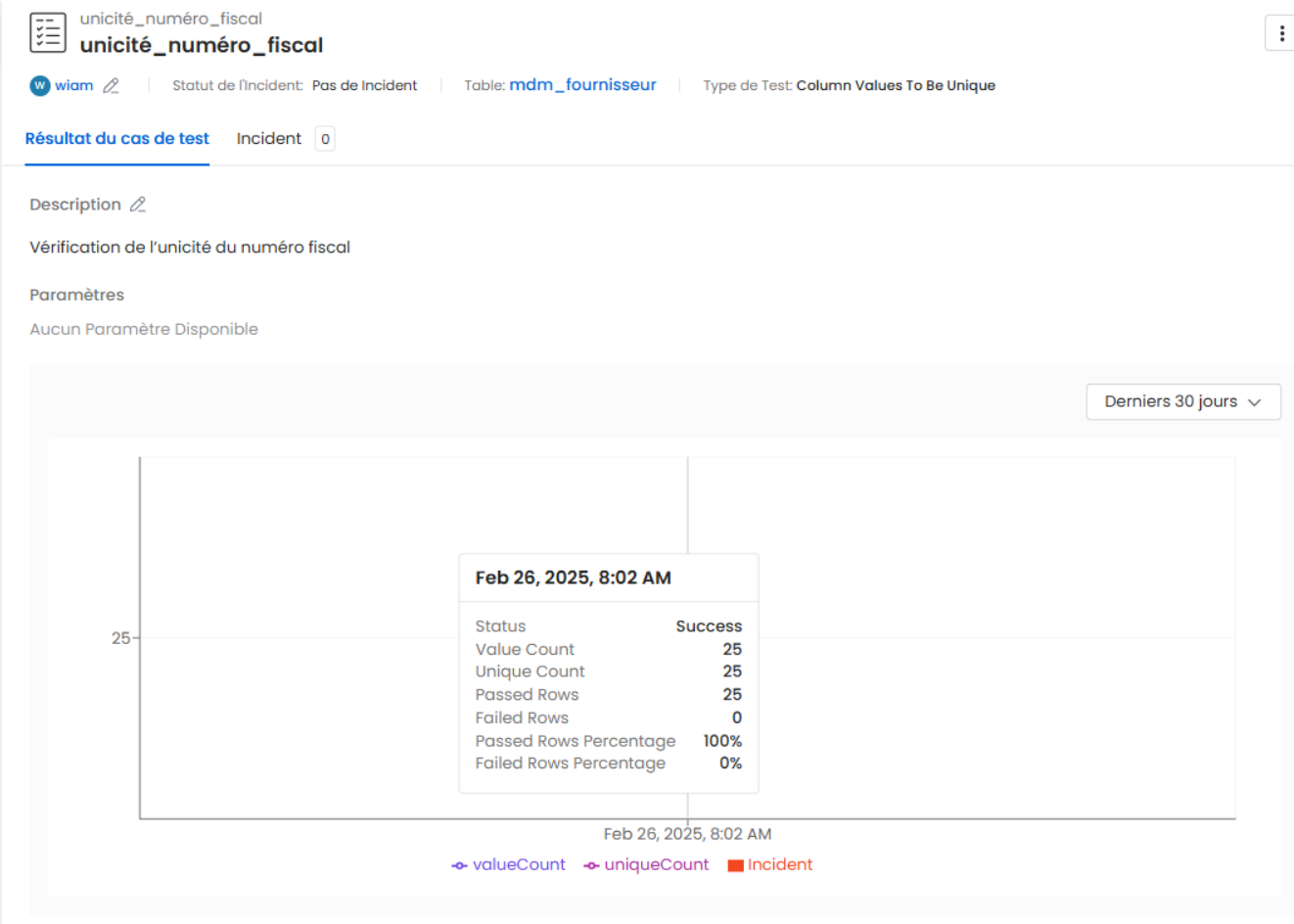
1%

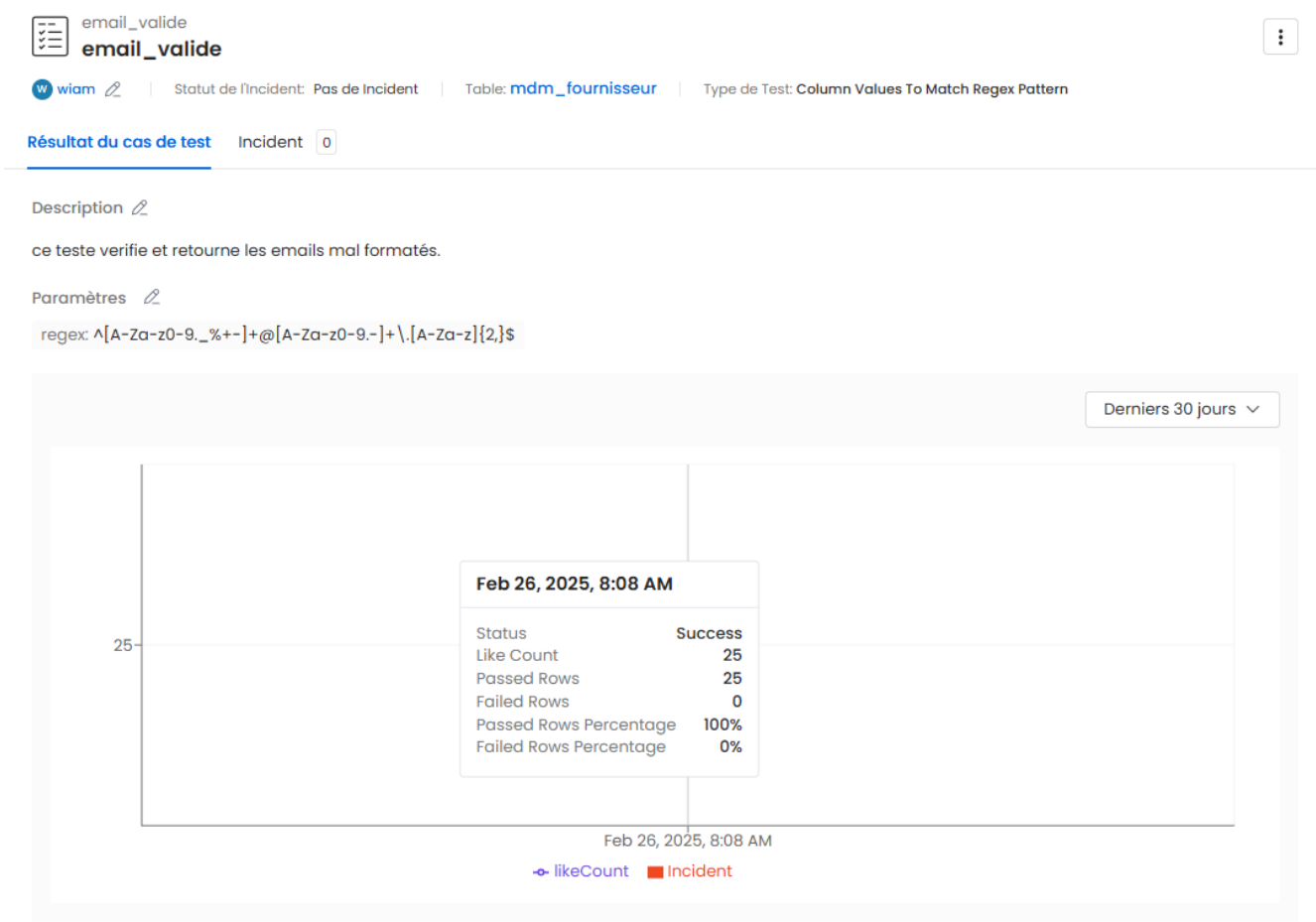
NOM	TYPE DE DONNÉES	NULL %	UNIQUE %	DISTINCT %
id_fournisseur	int	0%	0%	0%
nom	varchar(255)	0%	0%	100%
id_categorie	int	0%	0%	100%
numero_fiscal	varchar(50)	0%	0%	0%
cnss	varchar(50)	0%	0%	0%
registre_commerce	varchar(50)	0%	0%	0%
adresse	text	0%	0%	0%
pays	varchar(100)	0%	0%	0%
ville	varchar(100)	0%	0%	0%
telephone	varchar(20)	0%	0%	0%
email	varchar(255)	0%	0%	0%
date_creation	date	0%	50%	100%

Tests de Qualité des Données

On a défini des règles de validation pour garantir la conformité des données stockées, les Tests de Qualité Implémentés :

test1 : Unicité du numéro fiscal : Vérifie qu'il n'existe pas de doublons pour les numeros fiscales chaque heure :



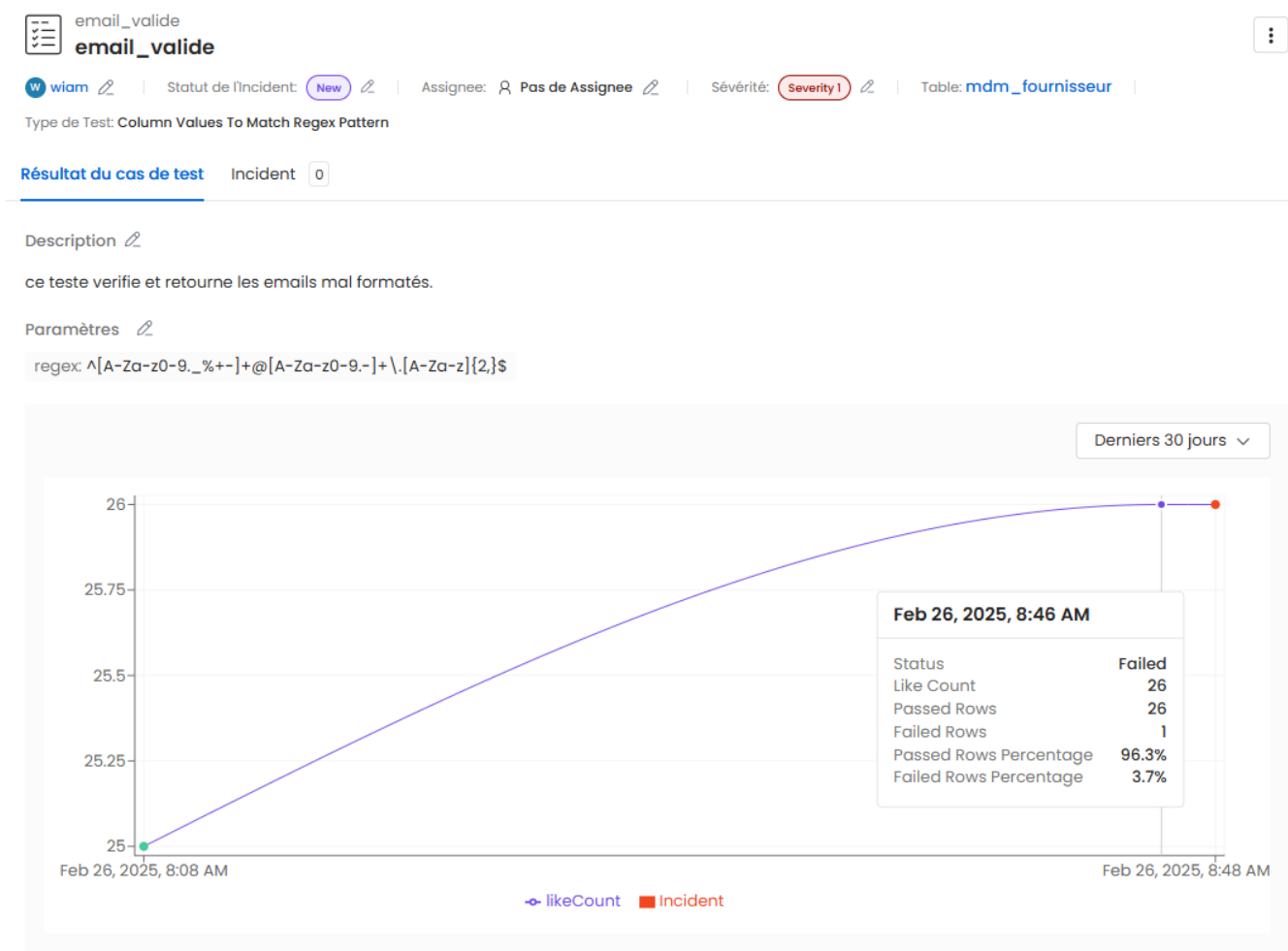


On a ajouté un enregistrement avec un mail non valide pour verifier la fonctionnalite du test :

```
USE mdm_hotel;
INSERT INTO mdm_fournisseur (nom, id_categorie, numero_fiscal, cnss, registre_commerce, adresse, pays, ville, telephone,
VALUES
-- Invalid email format (missing @ and domain)
> ('Fournisseur_Test_Invalid_Email', 1, 'MA123456789', 'CNSS987654321', 'RC123456', '123 Rue Exemple', 'Maroc',
- 'Casablanca', '+212600000000', 'invalid_email.com', '2023-05-10'),

-- Duplicate numero_fiscal (assuming 'MA123456789' already exists)
> ('Fournisseur_Test_Duplicate_Fiscal', 2, 'MA123456755', 'CNSS123123123', 'RC654321', '456 Avenue Exemple', 'Maroc',
- 'Marrakech', '+212611111111', 'test@valid.com', '2023-06-15');

select * from mdm_fournisseur ;
```



Alertes et Monitoring

On a mis en place des alertes automatiques pour signaler les anomalies détectées :

Seuil d’alerte : Dès qu’un test de qualité échoue.

Moyen de notification : Email + Tableau de bord OpenMetadata.

Comme Exemple d’implementation, on a implementé une alerte pour Email non conforme

Un email ne respecte pas le format attendu.

Mettre à Jour Alerte

Restez informé avec des alertes en temps réel à l'aide de webhooks.

Nom

INVALID_MAIL

Description

B I ↺ < > :: ≡ 🔗 📎 </> “ ”

Source

Spécifiez la source qui déclenchera l'alerte.

☐ Cas de Test

Filtres

Spécifiez les événements de modification pour réduire le périmètre de vos alertes.

Test Case Name

Inclure: ☒

Ajouter Filtre

Filtres

Spécifiez les événements de modification pour réduire le périmètre de vos alertes.

Test Case Name

MDM_HOTELmdm_hotelmdm_hotel... X

Inclure: ☒

Ajouter Filtre

Déclencheur

Selectionnez les événements déclencheurs importants comme 'Modification de Schema' ou 'Echec du Test' pour générer une alerte.

Get Test Case Status Updates

Failed X

Inclure: ☒

Ajouter Déclencheur

Destination

Envoyer des notifications à Slack, MS Teams, Email, ou utiliser des Webhooks.

Délai d'Attente de Connexion :

Users

wiamchawa X

Email

Assurez-vous que "Users" ont bien configuré leur "Email" pour recevoir les alertes. X

Ajouter Destination

Test Destinations

Annuler

Enregistrer

Conclusion

Dans le contexte actuel où la qualité des données est devenue un enjeu stratégique majeur pour les entreprises, notamment dans le secteur de l'hôtellerie, assurer une gestion rigoureuse et efficace des informations est indispensable pour maintenir la compétitivité et offrir une expérience client optimale. À travers ce travail, nous avons abordé les trois parties fondamentales de la gestion des données dans l'hôtellerie : le client, le produit et le fournisseur.

Tout d'abord, la gestion des données client s'est avérée cruciale pour comprendre les besoins et attentes des consommateurs, garantir une personnalisation de l'offre, et fidéliser la clientèle. Ensuite, la gestion des données produit a permis de mieux structurer les informations liées aux services proposés, optimisant ainsi les processus internes et la rentabilité des offres. Enfin, la gestion des données fournisseur a assuré la traçabilité et la fiabilité des informations relatives aux partenariats, garantissant des approvisionnements cohérents et efficaces.

L'implémentation d'une gouvernance de données basée sur des principes tels que la complétude, la précision et la traçabilité est essentielle pour que les hôtels puissent maximiser leur performance opérationnelle et offrir une expérience sans faille à leurs clients. En conclusion, l'investissement dans une stratégie robuste de gestion de la qualité des données, appuyée par des solutions de Master Data Management (MDM), est un levier incontournable pour garantir l'intégrité des informations et, par conséquent, la prise de décisions éclairées et l'amélioration continue des services dans l'industrie hôtelière.