

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université A/Mira de Béjaïa
Faculté des Sciences Exactes
Département d'Informatique



MÉMOIRE DE MASTER
Option : Génie Logiciel

Thème

Mise en place d'un référentiel de données
produit
Cas d'étude : CEVITAL

Réalisé par:

M. *HACHEMAOUI* Salim
M. *HAMMADACHE* Hichem

Devant le jury composé de :

Examineur 1	M. AMROUN Kamal
Examineur 2	M. ALLEM Khaled
Encadrant	M. ACHROUFENE Achour

2020/2021

Résumé

La Gestion des Données de Référence ou Master Data Management (MDM) vise à gérer un inventaire unique de données de référence dans une entreprise à l'aide d'un ensemble standard d'outils et de processus. Les données de référence étant très critiques pour une entreprise. Parallèlement, le MDM facilite le traitement informatique dans les environnements comportant plusieurs services, plateformes et applications, permettant ainsi à une entreprise d'associer toutes ses données critiques à un seul et unique emplacement.

Ce projet consiste à réorganiser le système d'information de l'entreprise CEVITAL, en y intégrant un référentiel unique nommé Master Data services(MDS) qui permet de gérer les bases de données de chaque service de manière uniforme, centralisée et synchronisée, afin de garantir la manipulation des mêmes données dans chaque service en temps réel, en améliorant la qualité de ses données ainsi que la communication entre les services. La solution proposée est réalisée à l'aide de l'outil SQL Server 2019 qui alimente et exploite les données avec le module SSIS (Microsoft SQL Server Intégration Services) en utilisant un ETL (Extract Transform Load).

Mots clés : *MDM, MDS , SQL Server,ETL, DQS*

Abstract

Master Data Management (MDM) aims to manage a single inventory of master data across an enterprise using a standard set of tools and processes. Because master data is so critical to an organisation, MDM facilitates IT processing in environments with multiple services, platforms and applications, allowing an organisation to associate all of its critical data to a single location. This project consists of reorganising CEVITAL's information system, integrating a single repository called Master Data Services (MDS), which allows the databases of each service to be managed in a uniform, centralised and synchronised manner, guaranteeing the handling of the same data in each service in real time, improving the quality of its data and communication between services. The proposed solution is implemented using the SQL Server 2019 tool which feeds and exploits the data with the SSIS (SQL Server Integration Services) module using an ETL (Extract Transform Load).

Keywords : *MDM, MDS, SQL Server, ETL, DQS*

Remerciements

Nos remerciement vont en premier lieu à nos parents qui ont su nous donner courage, volonté et patience afin d'achever ce projet ainsi tous les projets de la vie.

*Nous tenons à remercier également **Mr.ACHROUFENE Achour** notre promoteur, pour son aide, ses conseils, ses encouragements et ses orientations qui nous ont guidés jusqu'à l'aboutissement de ce travail.*

Merci aux membres du jury d'avoir accepté de juger notre travail et de nous faire part de leurs remarques pertinentes.

Nous exprimons notre plus grande gratitude à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Dédicaces

*Merci Allah de m'avoir donné la force et la patience afin d'atteindre mon objectif.
Je dédie ce mémoire :*

À mes parents ;

À ma mère, qui a œuvré pour ma réussite, de par son amour, son soutien, ces sacrifices consentis et ses précieux conseils, pour son assistance et sa présence dans ma vie ; Reçu à travers ce travail également modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.

À mon père, qui par l'achèvement de ce travail fruits de longues années de sacrifices et de privations afin de m'aider à avancer dans la vie, à l'éducation reçue et aux valeurs nobles inculquées.

*À mon oncle **Zizi Mouhamed** qui nous a quitté récemment paix a son âme.*

À toute ma grande famille

À mes amis que je remercie pour leurs soutiens pendant toutes ces années.

HACHEMAOUI Salim

Dédicaces

Je dédie ce modeste travail :

À ma petite famille, aucune dédicace ne saura exprimer ma reconnaissance pour tout ce qu'ils ont fait pour moi, pour les valeurs qu'ils m'ont inculqué, je dédie particulièrement ce modeste travail à ma très chers maman sans elle je ne serais pas là, pour son soutien et ses conseils judicieux qui m'ont éclairés le chemin, que dieu les protège tous pour moi.

À mon père, qui peut être fier de trouver ici le résultat de longues années de sacrifices et de privations pour m'aider à avancer dans la vie. Puisse Dieu faire en sorte que ce travail porte son fruit ; Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de toi.

À toute ma famille et tous mes cousins et cousines.

À tous mes amis(es) sans exception.

Tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste travail.

HAMMADACHE Hichem

Table des matières

Table des matières	iv
Liste des figures	v
Liste des tableaux	vii
Liste des abréviations	viii
Introduction générale	1
1 Étude Préalable	3
1.1 Introduction	3
1.2 Présentation de la société d'accueil CEVITAL	4
1.3 Historique de CEVITAL	4
1.4 Activités de CEVITAL	5
1.5 Missions de CEVITAL	6
1.6 Objectifs de CEVITAL	6
1.7 Organigramme du groupe CEVITAL	7
1.8 Présentation de l'organisme d'accueil	7
1.8.1 Organigramme de la direction système d'information	8
1.8.2 Département système d'information	8
1.8.3 Département information technologie	9
1.8.4 Département application métiers	9
1.8.5 Département transformation digitale	9
1.9 Gestion des données au niveau de CEVITAL	9
1.10 Problématique	12
1.11 Solution proposée	13
1.12 Conclusion	13

2	Master Data Management	14
2.1	Introduction	14
2.2	Système d'information et son importance au sein d'une entreprise	14
2.2.1	Définitions	15
2.2.1.1	Donnée	15
2.2.1.2	Information	15
2.2.1.3	Système d'Information	15
2.2.2	Problématiques du Système d'Information	15
2.2.2.1	Spécialisation des données échangées	15
2.2.2.2	Contraintes du Système d'Information	16
2.3	Données de référence	17
2.3.1	Typologie des données de référence	18
2.3.2	Notion de famille de données	18
2.4	Présentation du concept MDM	19
2.4.1	Définition	19
2.4.2	Architectures des MDMs	20
2.4.2.1	Mode Centralisation	20
2.4.2.2	Mode Consolidation	20
2.4.2.3	Mode Coopération (mode mixte)	21
2.4.3	Comparaison des trois architectures MDMs	22
2.4.4	Solutions MDM présentes sur le marché	22
2.4.5	Choix d'une solution	24
2.5	Gestion de la Qualité de données (DQM)	27
2.5.1	Qualité des données	27
2.5.1.1	Critères intrinsèques	27
2.5.1.2	Critères de service	27
2.5.1.3	Les critères de sécurité	28
2.6	Systèmes décisionnels	28
2.6.1	Définition d'un système d'aide à la decision	28
2.6.2	Objectif d'un système décisionnel	29
2.7	Entrepôt de données	29
2.7.1	Définition d'un entrepôt de données	29
2.7.2	Objectifs d'un entrepôt de données	30
2.8	Processus ETL (Extraction, Transformation, Load)	30
2.8.1	Définition	30
2.8.2	Phases constructives d'un ETL	30

2.9	Conclusion	31
3	Conception du MDS	32
3.1	Introduction	32
3.2	Environnement de l'entreprise	32
3.3	Identification des acteurs	33
3.4	Diagramme de cas d'utilisation	34
3.5	Description textuelle des cas d'utilisation	36
3.6	Choix de l'architecture	36
3.7	Diagramme de classes	36
3.7.1	Service Production	37
3.7.2	Service Achat	38
3.7.3	Service Commercial	39
3.8	Dictionnaire de données	40
3.9	Conception du référentiel Produit	42
3.9.1	Diagramme de classes	42
3.9.2	Modèle relationnel	44
3.10	Conclusion	44
4	Réalisation	45
4.1	Introduction	45
4.2	Présentation des outils utilisés	45
4.2.1	SQL Server 2019	45
4.2.2	SQL Server Management Studio	46
4.2.3	Visual studio 2019	46
4.2.4	Excel	46
4.2.5	SSIS (SQL Server Integration Services)	46
4.2.6	DQS (Data Quality Services)	46
4.2.7	Draw io	47
4.2.8	Navicat Premium	47
4.3	Installation de SQL server 2019 management studio	47
4.4	Configuration de Master Data Services	49
4.5	Déploiement et exploitation de la solution MDS	52
4.6	Alimentation du référentiel Produit	54
4.6.1	Contrôle de la qualité de données	54
4.6.2	Partie intégration de donnée SSIS (SQL Server Integration Services)	55

4.6.3 Attribution des rôles	59
4.7 Conclusion	64
Conclusion générale	65
Bibliographie	67

Liste des figures

1.1	Logo de CEVITAL	4
1.2	Organigramme de l'entreprise CEVITAL [source interne]	7
1.3	L'organigramme de la DSI de CEVITAL [source interne].	8
1.4	Diagramme de flux de processus	11
2.1	Gartner Group – Syndrome Spaghetti	16
2.2	Quelques données de référence dans un système d'information [1]	17
2.3	quelques exemples de regroupement en famille [1].	19
2.4	Mode Centralisation	20
2.5	Mode Consolidation	21
2.6	Mode Coopération	21
2.7	Architecture de MDS [2]	26
3.1	Diagramme de cas d'utilisation du système	35
3.2	Diagramme de classes du service de production	37
3.3	Diagramme de classes du service d'achat	38
3.4	Diagramme de classes du service Commercial	39
3.5	Diagramme de classes du Modèle Produit	43
4.1	Centre d'installation SQL Server	48
4.2	Sélection de fonctionnalités	48
4.3	Configuration d'analyses Services	49
4.4	Gestionnaire de configuration Master Data Services.	50
4.5	Fonctionnalités d'IIS	50
4.6	L'assistant de création de base de données	51
4.7	Interface principale de MDS	52
4.8	Interface gestion des modèles	53
4.9	Interface gestion des entités	53

4.10	interface gestion attribut pour l'entité Produit	54
4.11	Interface principale de DQS	55
4.12	Sources des données du projet.	56
4.13	Sources des données du projet.	57
4.14	Exécution de chargement et transformation des données.	58
4.15	Entité Produit	59
4.16	Interface du super administrateur	60
4.17	Interface d'attribution des rôles	61
4.18	Interface du superviseur de données	61
4.19	Interface de l'utilisateur du SI	62
4.20	Interface Excel pour l'utilisateur du SI	62
4.21	Interface utilisateur du SI du service commercial	63
4.22	Interface utilisateur du SI de production-maintenance	64

Liste des tableaux

2.1	Comparaison entre les trois architectures [3]	22
3.1	Identification les rôles des acteurs	33
3.2	Dictionnaire de données de la classe Produit	41
3.3	Dictionnaire de données de la classe souscategorieProduit	42
3.4	Dictionnaire de données de la classe CategorieProduit	42

Liste des abréviations

- 2TUP** : Two Track Unified Process.
- BDD** : Base de données.
- BI** : Business Intelligence.
- DSI** : Direction de Système d'information.
- DQM** : Data Quality Management.
- ETL** : Extract Transform and Load.
- IDE** : Integrated Development Environment.
- IIS** : Internet Information Services.
- MDM** : Master Data Management.
- MDS** : Master Data Services.
- OLAP** : OnLine Analytical Processing.
- OLTP** : OnLine Transactional Processing.
- SDI** : Système d'infomation décisionnelle.
- SGBD** : Système de Gestion de Bases de Donnée.
- SI** : Système D'information.
- SQL** : Standard Query language.
- SSAS** : SQL Server Analysis Services.
- SSIS** : Sql Server Integration Services.
- SSMS** : SQL Server Management Studio.
- UML** : Unified Modeling Language.

Introduction générale

Depuis plusieurs années, les entreprises sont confrontées à des problèmes importants dus à la qualité des données. Du secteur de la grande distribution au secteur de la Vente en passant par le secteur Achat, personne ne peut ignorer cette réalité. La maîtrise de l'information est non seulement un facteur de réussite mais surtout un besoin vital pour l'avenir des entreprises. Les données de mauvaise qualité peuvent avoir un effet néfaste et conséquent en matière d'efficacité organisationnelle interne de l'entreprise qui peuvent être catastrophiques tel que les pertes de clients, perte de temps et coût lié au retraitement des données erronées, documents financiers légaux erronés, erreurs stratégiques, etc.

Parmi les entreprises algériennes qui sont confrontées à ce type de problèmes, nous trouvons CEVITAL. Dans le cadre de notre projet de fin de cycle, nous avons effectué un stage au niveau de cette entreprise qui comporte plusieurs services (service Production-Maintenance, service Achat et service Commercial). Nous avons constaté que ces services traitent les informations d'une manière séparée, ce qui amène à des problèmes de redondances des données et l'incohérence des résultats qui deviennent inexploitable, provoquant une défaillance dans la qualité de données.

C'est dans ce contexte que se situe l'objectif de notre projet qui consiste à la mise en place d'un référentiel unique et centralisé, de l'ensemble des données du SI, accessible par tous les services de l'entreprise CEVITAL, en utilisant un Master Data Management (MDM). Ce travail s'inscrit dans une logique de gestion du Système d'Information qui poussera à réfléchir aux échanges des données et à restructurer, parfois sur plusieurs années, tout le Système d'Information de l'entreprise.

Notre mémoire se présente en un document structuré en quatre chapitres comme suit :

- **Chapitre 1 " Étude Préalable " :** Dans ce présent chapitre, nous allons introduire l'organisme d'accueil, en décrivant l'entreprise CEVITAL, ses différents départements et son mode de gestion des données. Nous allons présenter les différents problèmes qu'engendre sa gestion et proposer une solution adéquate pour les traiter.
- **Chapitre 2 " Master Data Management " :** Dans le deuxième chapitre, nous décortiquons le Master Data management par la présentation des généralités reliées à ce dernier, à savoir son architecture, son mode de gestion de la qualité de données. Nous terminons par présenter le processus ETL (Extract Transform Load) servant d'outil d'extraction des données.
- **Chapitre 3 " Conception du MDM " :** Dans ce chapitre, nous allons recenser les besoins de l'organisation, en présentant la démarche suivie durant notre projet. Nous allons définir la conception des différents services ainsi que la conception de référentiel produit.
- **Chapitre 4 " Réalisation" :** Ce chapitre va décrire les outils utilisés, les étapes de réalisation du projet MDS sous SQL server 2019 et les résultats fournis à chaque étape.

Enfin, nous clôturons ce mémoire par une conclusion dans laquelle nous résumons notre mémoire et nous exposons quelques perspectives futures.

Étude Préalable

1.1 Introduction

CEVITAL est une société par action (SPA), elle fait partie des entreprises algériennes qui ont vu le jour dès l'entrée du pays en économie de marché. C'est le premier groupe privé algérien, présent également à l'international et la troisième entreprise algérienne par le chiffre d'affaires.

CEVITAL est composée de plusieurs départements contribuant à son bon fonctionnement. Cependant, le traitement des informations d'une manière séparée posent certains problèmes tels que la redondance des données et les mises à jour.

Dans ce chapitre, nous allons représenter l'état actuel de l'entreprise CEVITAL qui nous a accueillis tout au long de notre stage. Nous présenterons ses multiples activités et objectifs, en particulier le département de transformation digitale afin de comprendre le système d'informations reliant les différents services. Ensuite, nous allons exposer les problèmes de gestion rencontrés par CEVITAL pour établir le lien entre ses services et nous proposerons une solution basée sur le principe de Master Data Management pour résoudre ces problèmes.

1.2 Présentation de la société d'accueil CEVITAL



FIGURE 1.1 – Logo de CEVITAL

Le groupe CEVITAL est un conglomérat algérien de l'industrie agroalimentaire, créé par l'entrepreneur ISSAD RABRAB en 1998. C'est un groupe de vingt-cinq sociétés, réparties dans cinq secteurs d'activités : l'industrie métallurgique, l'information et la communication, la distribution automobile, le transport terrestre et maritime et l'industrie agroalimentaire, la figure 1.1 représente logo de groupe CEVITAL.

Son complexe de production est implanté au niveau du nouveau quai du port de Béjaïa, à 3 Km du sud-ouest de cette ville, à proximité de la RN 26. Limité par le centre-ville de Béjaïa au nord, Oued Ghir et Sonatrach au sud, NAFTAL et ECOTEX au sud-ouest et la méditerranée à l'est.

Sa situation géographique stratégique lui confère l'avantage de proximité économique. De plus, elle se trouve près du port et de l'aéroport, ce qui lui permet de bien distribuer ses produits et de dominer, d'une grande part, le marché national.

CEVITAL Agro-industrie est composée de plusieurs unités de production telles que : une raffinerie d'huile, deux raffineries de sucre, une margarinerie, une unité de conditionnement d'eau minérale qui se situe à Tizi Ouzou, une unité de fabrication et de conditionnement de boisson rafraîchissante (site EL-KSEUR), une conserverie, silos portuaires ainsi qu'un terminal de déchargement portuaire[4].

1.3 Historique de CEVITAL

CEVITAL est un groupe familial qui s'est bâti sur une histoire, un parcours et des valeurs qui ont fait sa réussite et sa renommée, c'est la première entreprise privée algérienne à avoir investi dans des secteurs d'activités diversifiés.

ISSAD Rebrab crée son cabinet d'expert-comptable en 1968, puis se lance dès 1971 dans l'entrepreneuriat, en créant des sociétés dans la métallurgie et en 1991 dans la sidérurgie. À la tête du Groupe CEVITAL, il n'a cessé de le faire grandir en diversifiant ses activités,

pour atteindre sa taille et sa notoriété actuelle l'entreprise a traversé d'importantes étapes historiques qui sont [5] :

- **1971** : Lancement de la construction métallique.
- **1988** : Création de META SIDER (SIDERURGIE).
- **1991** : Reprise des activités I.B.M en Algérie /Création du quotidien Liberté.
- **1997** : Création de Hyundai Motorsen Algerie.
- **1998** : Création de CEVITAL SPA industrie agroalimentaires
- **2006** : Création de numidis et immobis ; acquisition de COJEK.
- **2007** : SAMHA- Production et distribution SAMSUNG /Création MFG (verre plat).
- **2008** : Nolis-Transport maritime/commercialisation du verre plat en Europe/Création de Numilog.
- **2009** : Augmentation de la production de sucre.
- **2013** : OXXO (Italie/ ALAS (Italie)).
- **2014** : BRANDT (Italie)/AFFERPI(Italie) ex LUCCHINI PIOMBINO.
- **2015** → **2020** : Le groupe a su marquer sa présence sur les trois continents (Afrique, Europe et Amérique latine) avec un volume d'export parmi les plus élevés en Algérie et en se positionnant parmi les fournisseurs majeurs des marchés européens, de l'Amérique latine ainsi que les pays du Moyen Orient.

Le groupe CEVITAL réalise l'année 2020 un chiffre d'affaire de 3,5 Milliards de dollars et vise à atteindre 25 Milliards de dollars à l'horizon 2025. Cette évolution est le résultat d'une vision moderne, ambitieuse et stratégiquement cohérente avec l'économie algérienne et internationale.

1.4 Activités de CEVITAL

CEVITAL a débuté son activité pour le conditionnement en décembre 1998. En février 1999, les travaux de génie civil de la raffinerie ont débuté. Cette dernière est devenue fonctionnelle en août 1999 [4].

L'ensemble des activités de CEVITAL sont concentrées sur la production et la commercialisation des huiles végétales, margarine et sucre, elles se présentent comme suit [source interne] :

- Raffinage d'huile 1600T/J pouvant passer après extension à 1800 T/J.

- Production de margarines avec une capacité de 6000T/J.
- Fabrication d'emballage en PET (9600 unités/h).
- Raffinage de sucre avec capacité de production de 1600T/J.
- Sucre liquide.
- Silos portuaires.
- Minoterie semoulerie.
- Hydroélectrique d'huile.
- Production de lait et produits laitiers.
- Trituration de graines oléagineuses.
- Stockage et négoce de céréales.

La possession de plusieurs silos portuaires ainsi que celle d'un terminal de déchargement d'une capacité de 2000 tonnes/heure fait de CEVITAL le premier terminal de déchargement portuaire en Méditerranée.

1.5 Missions de CEVITAL

L'entreprise a pour mission principale de développer la production et d'assurer la qualité de conditionnement des huiles des margarines et du sucre a des prix nettement plus compétitifs et cela dans le but de satisfaire le client et de le fidéliser [4].

1.6 Objectifs de CEVITAL

Les objectifs visés par le groupe CEVITAL se présentent comme suit [source interne] :

- L'extension de ses produits sur tout le territoire national.
- L'importation des graines oléagineuses pour l'extraction directe des huiles brutes.
- L'encouragement des agriculteurs par des aides financières pour la production de ces graines localement.
- L'optimisation de ses offres d'emploi sur le marché du travail.
- La modernisation des installations en termes de marché étranger et leur exportation.
- La réponse aux besoins du client.
- L'accroissement du chiffre d'affaire.

- L'amélioration de ses relations de partenariat étrangers.
- Le positionnement de ses produits sur le marché étranger par leurs exportations.
- Amélioration sans cesse de ces performances à tous niveaux.

1.7 Organigramme du groupe CEVITAL

La figure 1.2 montre l'organisation générale de l'entreprise CEVITAL et notre domaine d'étude la direction du système d'information (DSI), en particulier dans le département transformation digitale [source interne] :

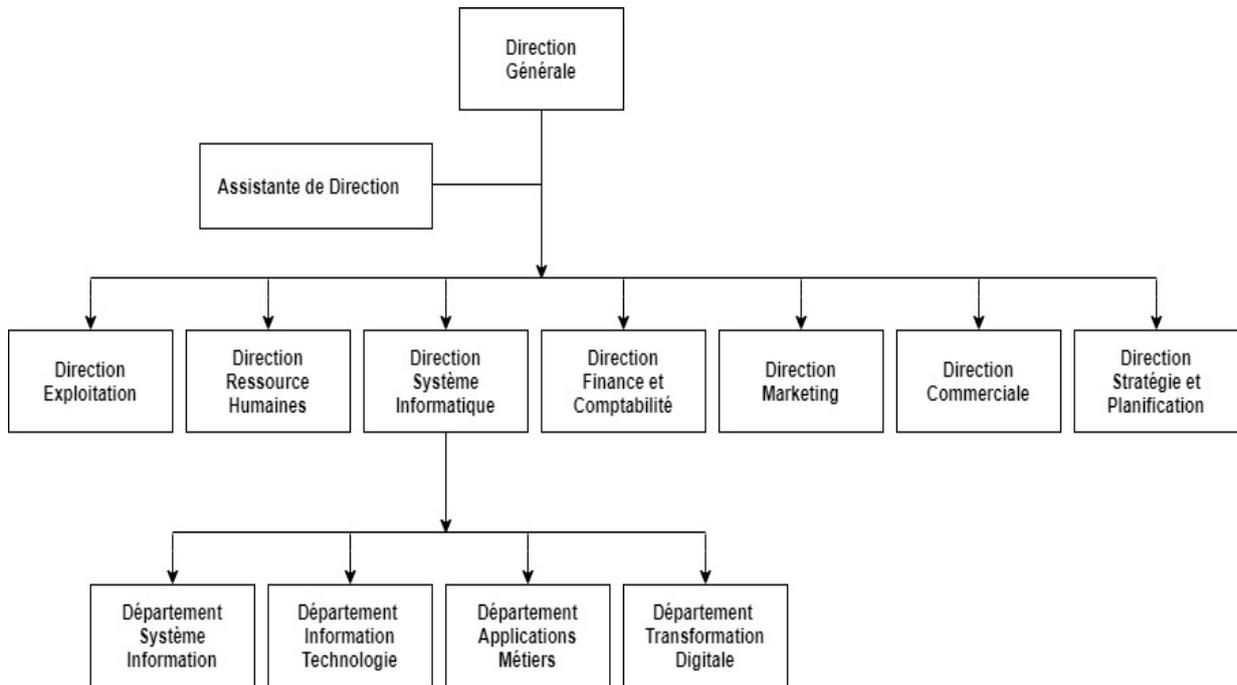


FIGURE 1.2 – Organigramme de l'entreprise CEVITAL [source interne]

1.8 Présentation de l'organisme d'accueil

CEVITAL donne une grande importance au domaine de l'informatique qui possède un système d'information et une direction informatique.

Notre étude est focalisée au niveau du groupe CEVITAL de Bejaïa où nous avons effectué notre stage, dans la direction du système d'information (DSI), en particulier dans le département transformation digitale.

1.8.1 Organigramme de la direction système d'information

La direction système d'information de CEVITAL est composée de deux départements [source interne] :
 département Information Technologie, département Application Métiers, département Transformation Digitale et département Système Information.

L'organigramme de la direction système d'information est montré dans la figure 1.3 :

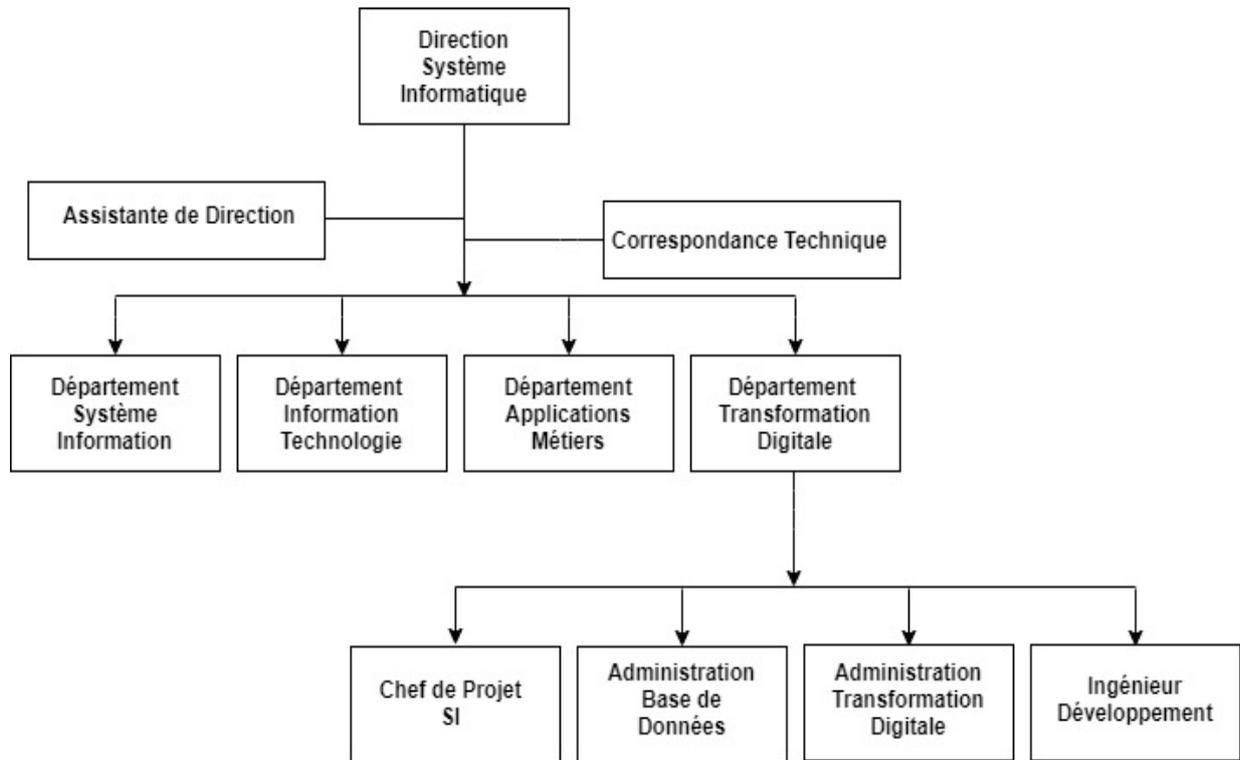


FIGURE 1.3 – L'organigramme de la DSI de CEVITAL [source interne].

1.8.2 Département système d'information

Ce département a pour mission de veiller à la performance et à l'optimisation du Système d'information, d'anticiper les évolutions métiers et de proposer des projets de développement en tenant compte de la stratégie de l'entreprise.

1.8.3 Département information technologie

Le département de la technologie de l'information est chargé de la compréhension approfondie de la gestion et la résolution des problèmes d'affaires à l'aide de solutions technologiques innovatrices.

1.8.4 Département application métiers

Il assure et facilite la gestion de l'activité de l'entreprise (la gestion des salariés, collaborateurs...), d'informatiser et d'automatiser les processus de gestion (développement d'application, administration de bases de données...).

1.8.5 Département transformation digitale

Son objectif est d'adapter l'entreprise à la concurrence au sein d'une économie numérique, savoir choisir les nouvelles technologies et les utiliser de manière à générer de l'innovation et un avantage concurrentiel durable.

Finalement, chaque département a pour but d'améliorer le niveau de l'information au sein de l'entreprise, ils collaborent afin de garantir le développement et la progression des services du groupe CEVITAL.

1.9 Gestion des données au niveau de CEVITAL

Les activités principales de CEVITAL passent essentiellement par quatre processus [source interne] :

- **Achat :**

Il prend en charge l'achat de toutes les gammes de produits : pièces de rechange, matières premières, packaging, etc.

- **Commercial :**

Il a en charge de commercialiser toutes les gammes de produits et le développement de Fichier clients de l'entreprise, au moyen d'actions de détection ou de promotion de projets à base de hautes technologies. Elle est aussi en relation directe avec la clientèle, et possède des qualités relationnelles pour susciter l'intérêt des prospects.

- **Production :**

Le service production est la colonne vertébrale de CEVITAL. Tous les autres services lui sont en quelque sorte des structures de soutien. Il est chargé du raffinage des huiles brutes, production des bouteilles en plastique d'un litre, de deux litres et de cinq litres, et du conditionnement de ces produits.

• **Maintenance :**

Il met en place et intègre de nouveaux équipements industriels et procédés, planifie et assure la maintenance pour l'ensemble des installations. Il gère et déploie aussi avec le Directeur Industriel et les Directeurs des Pôles, les projets d'investissement relatifs aux lignes de production, et rédige les cahiers de charges en interne.

Depuis l'ouverture de CEVITAL, la gestion de la pièce de rechange se fait par un magasin central (service maintenance). Avec le développement des capacités de production et la réalisation de nouvelles unités (service production), les besoins en pièces de rechange ont augmenté, pour cela les deux services sont regroupés en un seul.

Afin d'assurer la gestion, ces processus sont regroupés en trois services qui sont production-maintenance, achat et commercial.

Les trois services communiquent entre eux assez fréquemment, chacun d'entre eux a besoin des services de l'autre afin de compléter ses propres tâches. Des exemples de communication sont cités dans les points suivants :

- le service production demande au service d'achat d'acheter un produit (matière première), il peut aussi demander au service maintenance une opération de maintenance (maintenance industrielle)
- le service maintenance demande au service d'achat d'acheter un produit (une pièce de rechange par exemple)
- le service commercial (vente) demande au service de production de lui fournir des produits.

La figure 1.4 représente l'ensemble des transferts et des échanges de données entre les différents services.

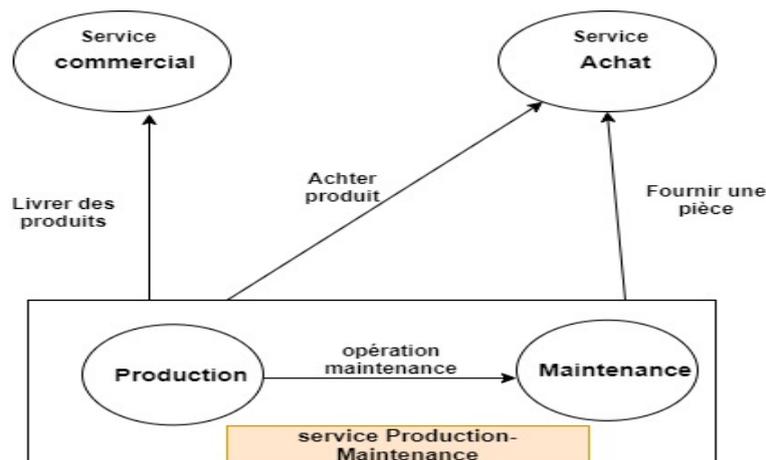


FIGURE 1.4 – Diagramme de flux de processus

1.10 Problématique

Parmi les majeurs problèmes de CEVITAL on compte la difficulté d'établir un lien entre les différents services, par exemple, un produit peut changer de prix de vente ce dernier peut être mis à jour dans une base de données de service Achat mais pas dans un autre (par exemple : celle de service Production). Ces services utilisent un logiciel mono-poste pour gérer ces données, cependant il peut y avoir des champs gérés dans un service qui ne sont jamais utilisés dans les autres services. Aucun des services ne fournit individuellement une vue complète des produits, c'est-à-dire, un service n'a pas la possibilité de voir ce que l'autre sauvegarde ni ce qu'il saisit.

l'entreprise dispose d'un entrepôt de données utilisé comme moyen de stockage performant mais vu la quantité de données stockées dans plusieurs services, devient difficile de déterminer quelle est la donnée la plus fiable ou la plus actualisée au sein de l'entreprise, les utilisateurs ne peuvent pas savoir avec certitude lequel des nombreux services disponibles possède les informations correctes. Ce qui conduira à une forte redondance qui engendre par la suite une grande difficulté lors de génération des rapports et une grande perte de temps en cherchant l'information souhaitée.

L'enjeu réside dans le fait de pouvoir réaliser une solution complète pour l'entreprise, qui puisse résoudre les problèmes suivants :

- Problème de cohérence des données entre les services qui mène à l'incohérence des résultats.
- Problème de données incomplètes, données NULL ou vides pour un attribut d'entité donné.
- Problème de duplication de données.
- Grand volume de données stockées dans plusieurs bases de données, l'exploitation de ces données est donc très difficile.
- Qualité de données douteuse et changeante dans le temps.
- Données dans un format difficilement interprétable ou ambigu.
- Difficulté de consommer la donnée et d'en avoir un aperçu entre services.

Afin de résoudre ces problèmes, source de perte d'efficacité pour l'entreprise, la mise en place d'un référentiel unique regroupant l'ensemble des données des services est nécessaire. Ainsi, chaque utilisateur pourra voir ce que son collègue a déjà traité ce qui permet d'utiliser la même version de la donnée.

1.11 Solution proposée

Notre travail consiste à mettre en place une solution basée sur la technologie de Master Data Management (MDM) et Data Quality Management (DQM) pour pouvoir résoudre les problèmes de l'entreprise mentionnés précédemment. Cette solution doit permettre :

- La centralisation, la synchronisation et l'uniformisation des données.
- D'améliorer la performance et soutenir la montée en charge.
- De fournir une meilleure vue sur les données de l'entreprise.
- De définir les accès aux données pour les services et personnes autorisées.
- D'améliorer la communication entre les différentes parties du Système d'information.
- La désignation d'un coordinateur de données (Data Steward ou responsable du MDM).
- De faciliter l'exploitation des informations stockées (données organisées selon différentes dimensions).
- D'offrir des informations fiables, cohérentes et pertinentes, contenant la logique business souhaitée.
- La mise en place d'un Dashboard pour la visualisation de l'état des données (Power BI).

1.12 Conclusion

Dans ce premier chapitre, nous avons présenté en premier lieu l'organisme d'accueil CEVITAL, puis nous nous sommes intéressés aux services qui sont en relation directe avec la suite de notre travail. Par la suite, nous nous sommes focalisés sur la problématique decelée ainsi que la solution proposée. Les concepts fondamentaux relatifs à la solution à proposer seront aborder dans le chapitre suivant.

Master Data Management

2.1 Introduction

La gestion des données de référence (GDR), ou Master Data Management en anglais (MDM), est une discipline des technologies de l'information qui concerne les données de référence. Elle est le garant de la cohérence entre diverses architectures de systèmes et fonctions métiers.

Dans un premier temps, nous décrirons les points importants à connaître tels que la donnée et le système d'information afin de bien comprendre les concepts avancés tels que le MDM et ses architectures, la qualité de la donnée, système décisionnel et entrepôt de données. Dans un deuxième temps, nous présentons les architectures MDM, les solutions MDM présentes sur le marché et celle retenue dans ce travail.

2.2 Système d'information et son importance au sein d'une entreprise

Il est très important de maîtriser l'information au sein d'une entreprise afin d'assurer sa continuité. Le travail effectué et les actions prises par chaque acteur dépendent des données disponibles. Pour cela, il faut s'assurer que tous les acteurs partagent les mêmes données, car des données incorrectes influencent le bon fonctionnement opérationnel au sein de l'entreprise et peuvent conduire à des fails.

2.2.1 Définitions

2.2.1.1 Donnée

Dans les technologies de l'information, une donnée est une description élémentaire, d'un objet, d'une transaction d'affaire, d'un événement... [1]. La donnée est représentée sous forme codée, en vue d'être collectée, traitée, conservée, échangée [1]. Elle peut être valorisée et a un cycle de vie [6].

2.2.1.2 Information

Une information est un ensemble d'éléments porteurs de sens [7], ou un ensemble de données agrégées dans le but d'une utilisation par l'homme (par exemple, le résultat d'une requête décisionnelle qui somme des données individuelles est une information) [6].

2.2.1.3 Système d'Information

Le système d'information est un ensemble organisé de ressources, matériels, logiciels, personnels, données, procédures permettant d'acquérir, de traiter, stocker, communiquer des informations sous plusieurs formes (textes, images, sons...) dans des organisations [8].

2.2.2 Problématiques du Système d'Information

Les systèmes d'informations ont beaucoup contribué au développement des entreprises, cependant plusieurs problèmes sont apparus.

2.2.2.1 Spécialisation des données échangées

Si, à l'origine, les données étaient cloisonnées dans le SI de chaque service pour des raisons de commodité et de réactivité, aujourd'hui, les différents services ont besoin de communiquer entre eux des données qu'ils partagent. Cependant, le format des données qu'ils échangent pouvant être différent, des interfaces sont généralement mises en place par les entreprises pour pouvoir communiquer. Le souci lié à ce genre de pratique est qu'à chaque nouvel échange, une nouvelle interface doit être créée, ce qui entraîne ce que Gartner Group appelle le « syndrome spaghetti » illustré dans la figure 2.1 [1].

Aussi, ces données échangées entre les différentes parties du Système d'Information deviennent de plus en plus hétéroclites, éparses, incohérentes et obsolètes. Elles peuvent donc être source d'erreurs, et influencer le bon fonctionnement de l'entreprise.

Comme nous pouvons le constater sur la figure 2.1, il sera très difficile de comprendre et maintenir ce système d'information. Les liens entre les applications ne sont pas indépendants, ils forment des enchaînements correspondant à une logique métier, mais sans aucun contrôle central.

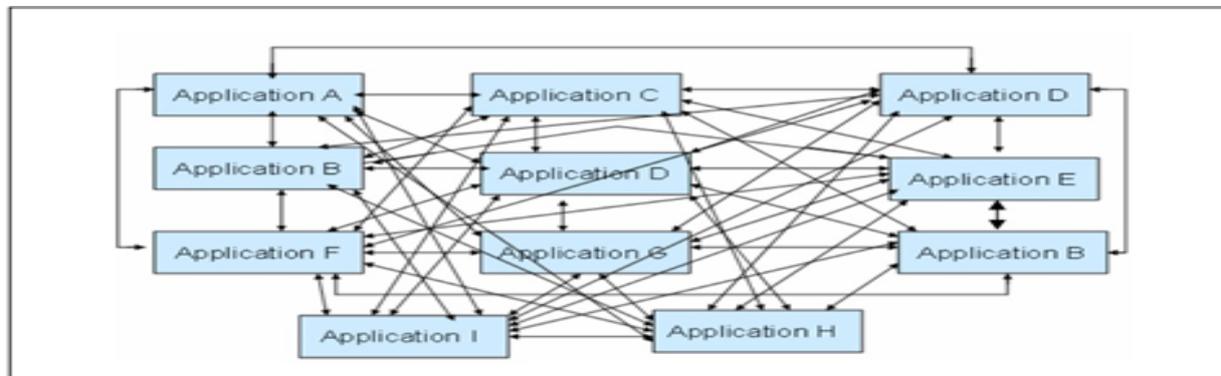


FIGURE 2.1 – Gartner Group – Syndrome Spaghetti

2.2.2.2 Contraintes du Système d'Information

Actuellement, les données échangées, souvent éparses et hétéroclites, représentent un volume extrêmement important qui tend à grossir au fil des années. Cela entraîne de nombreux problèmes qui poussent les entreprises à adopter des solutions afin d'améliorer la qualité de leur Système d'Information, décision motivée par plusieurs contraintes.

- L'apparition de nouveaux impératifs métiers nécessite d'avoir de plus en plus l'information en temps réel (gestion des stocks, état de livraison, etc).
- Le phénomène de multiplication des systèmes et applications, les données opérationnelles de chaque système se retrouvent éparpillées, dupliquées, désynchronisées, incohérentes, ce qui a des répercussions sur la performance de l'entreprise :
 - désaccords internes sur la valeur à attribuer à telle ou telle donnée.
 - faible contrôle des données.
 - coût financier pour réparer les erreurs liées aux données erronées localement.
 - altération des données partagées au fil des traitements par la chaîne d'applications intermédiaires.
 - augmentation des contraintes métier pour garantir leur cohérence, entraînant une explosion de leur complexité.

- dysfonctionnement opérationnel dans les processus métiers critiques.

En plus de ces contraintes il existe d'autres contraintes dites externes résultent des phénomènes économiques et sociologiques.

2.3 Données de référence

Toute entreprise possède ses propres activités ou processus métiers tels que la production, la maintenance, l'achat et la vente. Tous ces derniers s'appuient et utilisent des données comme éléments de base : produits, clients, fournisseurs, employés, pays, catégories, codes, etc.

La figure 2.2 présente quelques données de référence dans un système d'information. On voit que ces données, comme d'ailleurs l'ensemble des données du système d'information, sont dispersées et souvent dupliquées entre différents processus et applications.

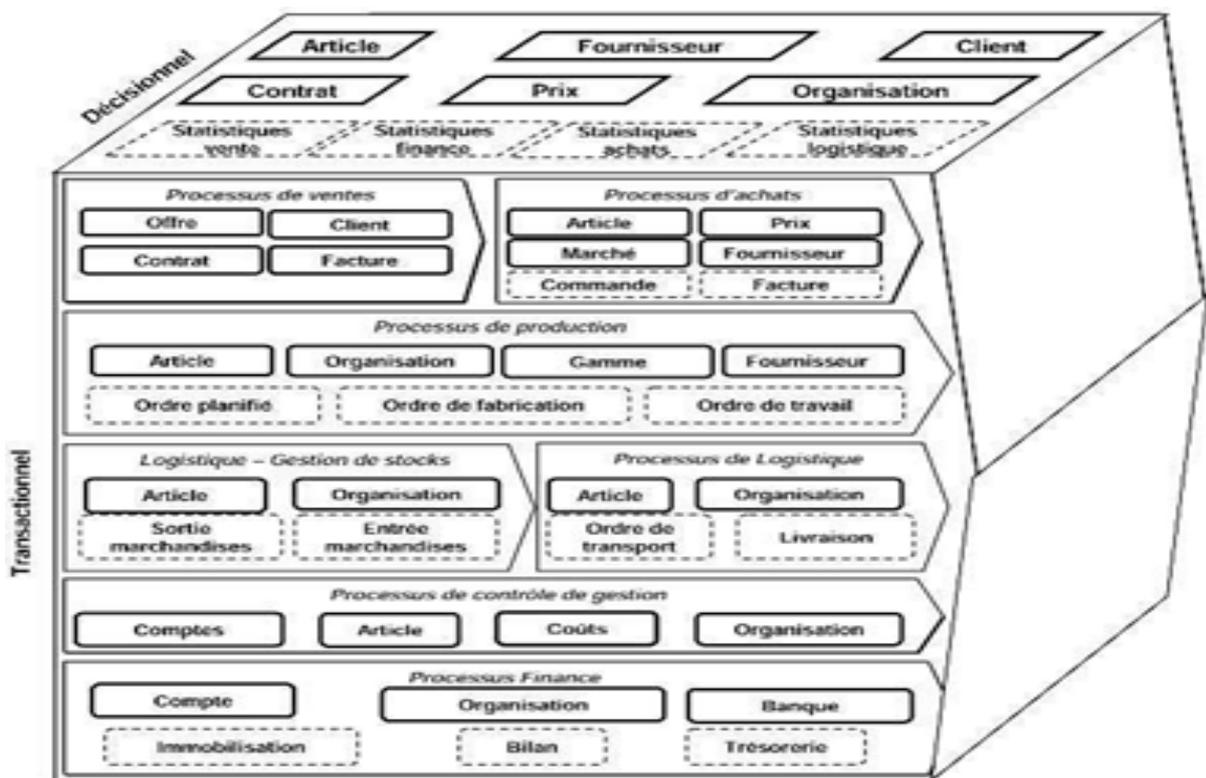


FIGURE 2.2 – Quelques données de référence dans un système d'information [1]

2.3.1 Typologie des données de référence

Il existe trois grands types de données de référence [1] :

1. **Les données maître** : ces données se manifestent dans les objets métiers principaux d'un domaine fonctionnel tel que les données client, produit et fournisseur. Elles sont au coeur d'un système d'information et structurent les principales applications.
2. **Les données constitutives** : sont des données constituées elles-mêmes d'attributs (par exemple, une adresse peut-être composé de plusieurs attributs tels que la ville et la rue). Ces attributs caractérisent en général des données maître, mais aussi d'autres objets métier (adresse peut représenter l'adresse de livraison de l'objet métier **bon de livraison**).
3. **Les données paramètre** : sont des tables de valeurs ou un glossaire de termes tels que les codes postaux et devises. Ce sont essentiellement les données les plus partagées du système d'information, une attention particulière doit donc être portée à ces données.

2.3.2 Notion de famille de données

Les données de référence sont liées entre elles et forment ce qu'on appelle famille de données, quoiqu'elles sont classifiées et semblent pouvoir être étudiées isolément.

Chaque entreprise forme ses propres familles de données selon les activités qu'elle exerce. Ceci aura notamment une incidence sur les préconisations d'organisation et de gouvernance des données par la suite [1]. La figure 2.3 montre des exemples de regroupement des données en famille.

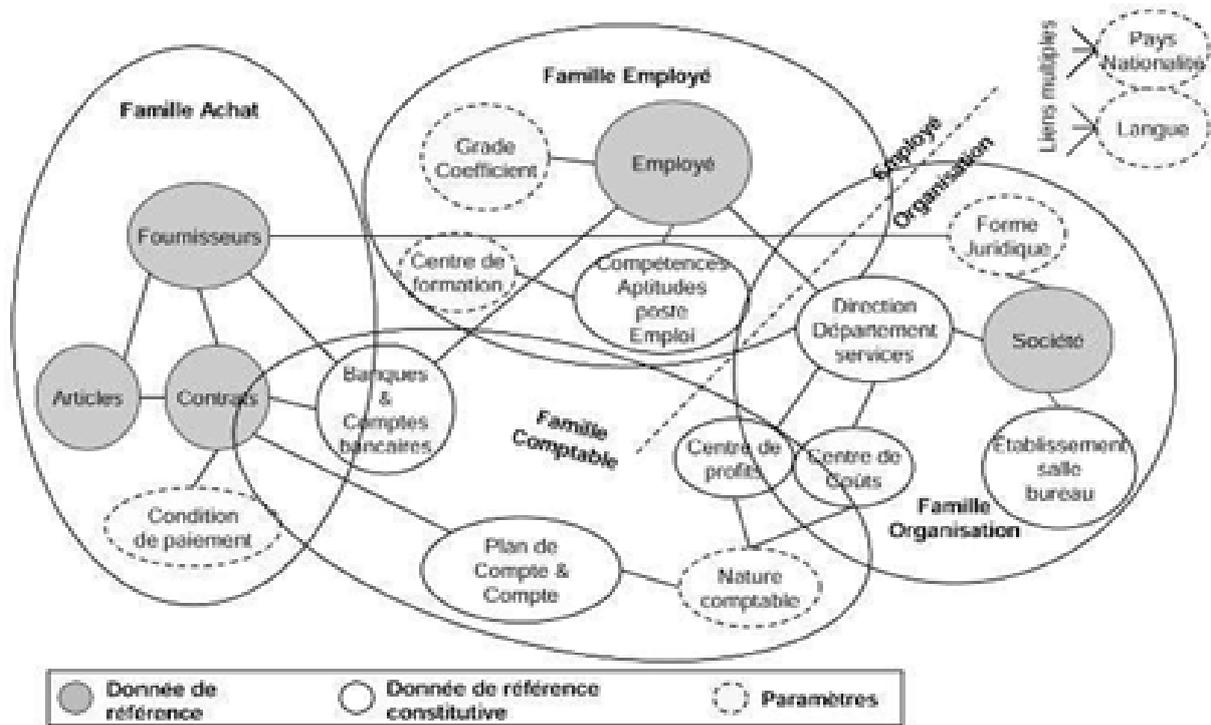


FIGURE 2.3 – quelques exemples de regroupement en famille [1].

Il est important d'assurer la gestion de ces données de référence. Ces dernières sont des données partagées et utilisées dans l'ensemble du système d'information de l'entreprise. Nous présentons dans la section suivante l'approche MDM qui répond à cette problématique.

2.4 Présentation du concept MDM

Le concept MDM est une démarche qui met en œuvre des procédures durables par la gouvernance des données. Cette gouvernance est assurée par une organisation de circonstances, composée d'individus aux tâches précises, et assistée par des outils dédiés afin d'améliorer la qualité et le partage des données transversalement à l'organisation.

2.4.1 Définition

Il n'existe pas de définition de l'approche MDM qui soit communément acceptée par l'ensemble de la communauté MDM. Différentes définitions existent, en voici deux habituellement citées. La première définition stipule que :

Le MDM est la base fiable et faisant autorité pour les données utilisées à travers de nombreuses applications et circonscriptions dans le but de fournir une vue unique de l'entreprise. L'objectif est de fournir une vue unique de la vérité, où qu'elle se trouve [3].

La deuxième a défini le MDM comme une discipline des technologies de l'information qui concerne les données de référence et qui inclut d'autres disciplines tel que le DQM (Data Quality Management) afin d'atteindre ses objectifs [6].

2.4.2 Architectures des MDMs

Il existe trois architectures principales qui peuvent être distinguées en ce qui concerne les MDMs [9].

2.4.2.1 Mode Centralisation

Dans ce mode, le Master Data devient le seul et unique système maître des données de référence et tous les applicatifs de l'entreprise sont simples consommateurs de ces données.

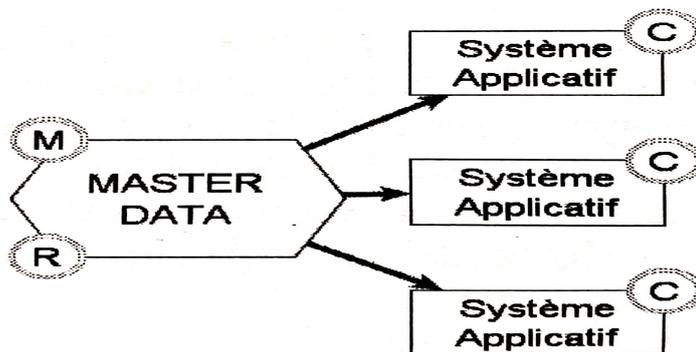


FIGURE 2.4 – Mode Centralisation

M : Système maître de la donnée de référence, **R** : Point de référence, **C** : Point de consommation de la donnée.

2.4.2.2 Mode Consolidation

Dans ce mode, le Master Data est le seul et unique point de référence des données de référence, qui sont gérées dans plusieurs systèmes sources, puis intégrées au master data

pour être mises à disposition des systèmes cibles.

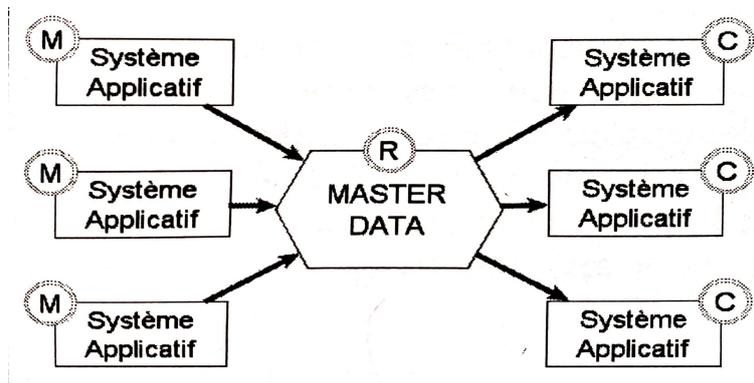


FIGURE 2.5 – Mode Consolidation

M : Système maître de la donnée de référence, **R** : Point de référence, **C** : Point de consommation de la donnée.

2.4.2.3 Mode Coopération (mode mixte)

Suivant la donnée de référence, il peut être soit simplement un point de référence, soit disposer aussi d'un rôle de maître.

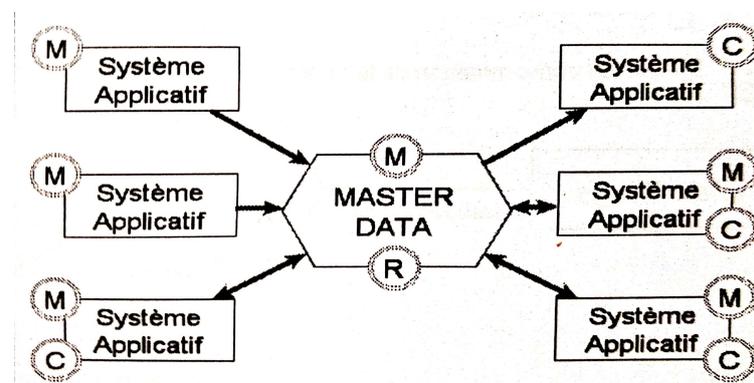


FIGURE 2.6 – Mode Coopération

M : Système maître de la donnée de référence, **R** : Point de référence, **C** : Point de consommation de la donnée.

2.4.3 Comparaison des trois architectures MDMs

Chacune des trois architectures MDM que nous avons introduit possède des avantages et des inconvénients. L'architecture de centralisation est la plus efficace opérationnellement car les systèmes applicatifs ont toujours accès à une seule source de données de référence consistante et à jour [3]. Cependant, le coût de la mise en place d'une telle architecture est élevé par rapport aux autres, et la charge sur les systèmes fournisseurs des données est souvent trop importante.

Dans les deux autres architectures, la duplication des données est présente, ce qui nécessite plus de temps pour la mise à jour des données de références, et des outils de synchronisation afin d'assurer la cohérence. Le tableau 2.1 montre les points de comparaison entre ces trois architectures.

	Consolidation	Coopération	Centralisation
Contrôle	-	+	++
Disponibilité	-	+	++
Cohérence	-	+	+++
Enrichissement	-	++	++
Indépendance	++	-	—
Mise en place	++	-	—
Coûts de gestion	-	-	+++

TABLE 2.1 – Comparaison entre les trois architectures [3]

- : médiocre, + : bon, ++ : très bon, +++ : excellent

2.4.4 Solutions MDM présentes sur le marché

Le marché MDM a évolué rapidement au cours des dernières années, et la raison en est le large éventail d'exigences qu'un système MDM devrait être capable de gérer. Étant donné que la nature des données, les types de technologies utilisées, les différentes exigences frontales liées à la qualité des données et à la gérance des données varient fortement d'une entreprise à l'autre, le marché du MDM est très large. Nous citons les plus connues :

— **IBM [10] :**

IBM InfoSphere Master Data Management gère tous les aspects des données critiques de l'entreprise, indépendamment du système ou du modèle, et les transmet aux uti-

lisateurs d'applications via une vue unifiée. L'outil assure la conformité aux règles et politiques de gouvernance des données, et présente un cadre configurable qui prend en charge les environnements de cloud hybride. InfoSphere MDM est disponible en deux éditions (standard et avancée), et toutes les deux sont disponibles sur site ou en tant qu'offres de cloud entièrement gérées.

- **Avantages :**

- Le produit reste à jour avec les besoins en matière de gestion des données.
- Très bon ETL (Extract, Transform and Load tool) et très bon en gouvernance de données

- **Inconvénients :**

- Une meilleure intégration du cloud est souhaitée, afin de pouvoir faire des choses dans le cloud si nécessaire.
- Le produit n'est vraiment créé que pour les grandes organisations.

- **Informatica [10] :**

Informatica propose une solution MDM modulaire qui offre une vue unique des données. Le produit permet aux utilisateurs de créer une vue faisant autorité des données critiques de l'entreprise à partir de sources disparates, dupliquées et conflictuelles. Informatica MDM est également doté de l'IA et de l'apprentissage automatique, et comprend des fonctionnalités de qualité des données, d'intégration des données, de gestion des processus métier et de sécurité des données qui permettent d'enrichir facilement les enregistrements de données de référence avec des données provenant de fournisseurs externes. Les fonctionnalités MDM d'Informatica peuvent être déployées sur site ou dans le cloud.

- **Avantages :**

- La composante qualité des données est très bonne
- Doté de grandes capacités de nettoyage des données

- **Inconvénients :**

- La solution ne se connecte à aucun des outils d'analyse d'une façon directe.
- Manque de supports techniques

- **Microsoft [9] :**

Microsoft a étendu sa plate-forme SQL Server avec Master Data Services (MDS) à partir de la version 2008r2, mais à partir de la version 2012, elle est devenue mature.

Il possède son propre magasin de données et est très flexible dans son déploiement, ce qui lui permet d'être utilisé pour une variété d'applications. Comme il s'agit d'un composant de SQL Server, le coût de la licence est très acceptable et pour certaines versions de SQL Server (SQL Server Business Intelligence et SQL Server Enterprise).

- **Avantages :**

- L'assistance technique est très bonne.
- La console d'administration est la fonction la plus précieuse.
- Sa disponibilité pour un serveur sécurisé est l'une de ses caractéristiques les plus précieuses. De même, ses répliques, où nous pouvons gérer huit répliques. Son clustering et ses groupes de disponibilité sont également précieux
- Possibilité d'intégration avec Azure Active Directory.
- Il dispose d'une très bonne fonctionnalité appelée plug-in Excel. Il suffit d'installer le plug-in, puis il se connecte directement à l'instance MDS où nous pouvons maintenir et gérer nos données et les publier directement via Excel. Nous n'avons pas besoin d'aller à l'interface du MDM pour faire des changements.

- **Inconvénients :**

- C'est un outil distinct à installer séparément de Microsoft SQL Server, en conséquence des erreurs d'installation sont souvent rencontrées
- La version on-cloud n'est pas encore disponible.

2.4.5 Choix d'une solution

L'un des facteurs majeurs qui nous pousse à la mise en place d'un MDM est le désir de cohérence et d'exactitude des données de l'entreprise. Ces traitements relatifs à la qualité des données sont assurés par les fonctions de DQM (Data Quality Management), ces dernières doivent répondre et compléter les objectifs du MDM et permettre à l'entreprise CEVITAL l'utilisation des technologies proposées par Microsoft pour la réalisation de ses projets informatiques.

afin de réaliser notre projet, nous allons travailler avec les outils MDS (Master Data Services) et DQS (Data Quality Services) qui sont, respectivement, le MDM et le DQM de Microsoft, pour les raisons suivantes :

- ▷ Il s'intègre facilement à SQL Server, et nous pouvons utiliser SSIS (SQL Server Integration Services) pour effectuer les ETL.
- ▷ MDS est intégré à SQL Server, alors, nous pouvons profiter des autres services fournis en standard (par exemple le moteur OLAP et serveur de rapports).
- ▷ L'alimentation du référentiel peut être aussi effectuée à partir de multiples sources hétérogènes grâce à la suite Office proposée par Microsoft.
- ▷ MDS offre une interface web riche, ergonomique, simple et facile à utiliser. Elle adresse toutes les fonctionnalités d'administration et de la gestion de la sécurité.
- ▷ MDS propose l'outil Add-in Excel qui permet de se connecter au service web du référentiel afin de manipuler les données et effectuer des traitements sur les entités (consommation, modification, etc.).
- ▷ MDS permet à travers des vues d'abonnements d'exposer les entités et les hiérarchies sous forme de vue SQL que les applications et les ETL et les cubes vont consommer.
- ▷ Possibilité de compression de données ce qui nous permet de gagner en performances lors de la consommation des données.
- ▷ MDS nous donne la possibilité de mettre en oeuvre une sécurité plus granulaire, où nous pouvons rajouter des utilisateurs et des super-administrateurs au niveau des référentiels.
- ▷ MDS de Microsoft offre en termes d'architecture une possibilité de séparer la partie web et la partie base de données.

Brièvement le MDS se caractérise par une architecture à trois niveaux comprenant respectivement la couche de base de données, la couche de service et la couche d'interface utilisateur [2] :

- **La base de données MDS** est une base de données SQL Server, elle est responsable de la sauvegarde des données de référence. Cette base de données est livrée avec de nombreuses procédures stockées, tables de base de données et fonctions qui prennent collectivement en charge le traitement back-end.

Une BDD MDS nous permet de gérer des **modèles, entités, attributs, membres,...**etc :

- Un modèle représente l'objet de plus haut niveau au sein des services de données de référence tel que la donnée Produit.
- Les entités sont l'équivalence des tables dans une BDD relationnelle.
- L'attribut est l'équivalence d'une colonne dans une BDD relationnelle.

- Un membre est l'équivalence d'un tuple dans une BDD relationnelle.
- **Service** est le niveau intermédiaire de l'architecture, elle propose des services web à l'aide de la technologie de communication Windows WCF, et qui permet une liaison entre les deux autres niveaux de l'architecture.
- **L'interface utilisateur** peut être présente sous forme d'interface web ou bien en tant que Excel Add-in qui permet aux utilisateurs la manipulation des différentes parties du MDS.

La figure 2.7 montre l'architecture de MDS et ses interactions avec les différents systèmes externes possibles.

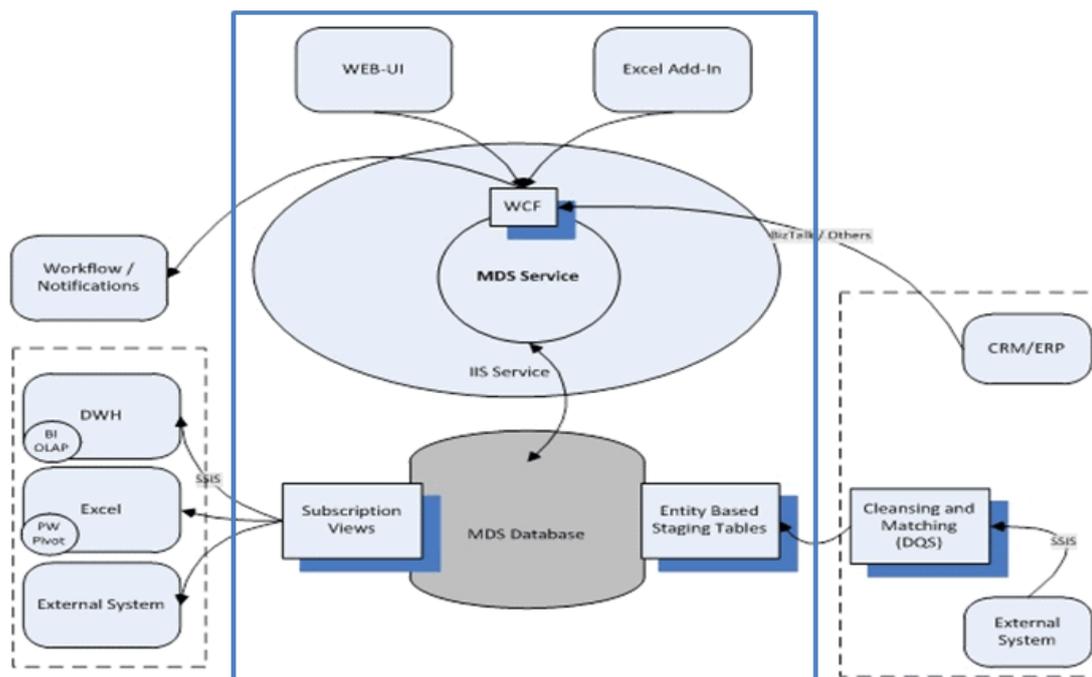


FIGURE 2.7 – Architecture de MDS [2]

L'objectif du MDM est d'assurer la fiabilité et la qualité des données de l'entreprise, cela grâce à l'approche DQM qui est une brique indispensable pour l'application des méthodologies MDM et l'obtention de meilleurs résultats, nous verrons plus sur ce concept dans la prochaine section.

2.5 Gestion de la Qualité de données (DQM)

La gestion de la qualité de données est une exigence fondamentale des bons systèmes d'information. Elle regroupe toutes les actions et procédures destinées à garantir et maintenir la qualité des données au sein d'une entreprise.

2.5.1 Qualité des données

les données sont de qualité si elles satisfont les exigences de leur utilisation prévue et elles manquent de qualité dans la mesure où elles ne satisfont pas à l'exigence. La qualité des données dépend autant des utilisations prévues que des données elles-mêmes mais elles doivent caractérisées par [11] [6] :

2.5.1.1 Critères intrinsèques

- **l'unicité** : la donnée doit apparaître qu'une et une seule fois en base de données. Elle possède un identifiant (sa clef primaire, par exemple, ou une clef composite comme le triplet nom/prénom/adresse) et ne doit pas être redondante.
- **la conformité** : se caractérise par le respect d'un ensemble de contraintes (le genre d'une personne doit être soit mâle soit femelle).
- **la complétude** : concerne les schémas et l'ensemble des données attendues (présence ou non de tous les attributs requis par exemple).
- **l'exactitude** : une donnée est considérée exacte si elle est correcte, valide et précise.
- **la cohérence** : une donnée peut être cohérente par rapport à une valeur de référence (par exemple le prix d'achat d'un produit est forcément inférieur au prix de vente).
- **l'intégrité** : se définit par les relations obligatoires entre les données (par exemple, une commande doit être liée à un client. Si elle ne l'est pas, la donnée « commande » a un problème d'intégrité).

2.5.1.2 Critères de service

- **l'actualité** : si la donnée n'est pas à jour à un instant donné, on dit qu'elle est obsolète (puisque modifiée mais ne reflétant pas son nouvel état) ou que sa représentation a changé (elle peut changer de représentation d'un instant à un autre).
- **l'accessibilité** : elle concerne la facilité d'accès aux données, c'est-à-dire que la donnée est à la fois accessible en mode requête à un service (service web, par exemple),

en mode batch pour des synchronisations de masse (pour le décisionnel, par exemple), et en mode événement (présente à chaque mise à jour).

- **la pertinence** : une donnée doit être accessible selon le niveau de granularité adéquat à son usage.
- **la compréhensibilité** : la bonne information doit être retrouvable facilement parmi les attributs disponibles d'un objet peu importe le biais (processus, utilisateur, application).

2.5.1.3 Les critères de sécurité

- **la disponibilité** : capacité du système à fournir une fonction pour un service précis.
- **l'intégrité** : concerne le droit à toute modification de données.
- **la confidentialité** : accès limité à qui de droit pour les besoins de service.
- **la traçabilité** : conservation des opérations et événements effectuées avec leurs auteurs.

2.6 Systèmes décisionnels

Du moment que la mesure de la performance de l'activité d'une entreprise est devenue un élément indispensable du plan d'action des entreprises pour l'atteinte de leurs objectifs, la conservation de leurs parts de marché et l'optimisation de leurs processus métiers. Un système d'information d'aide à la décision est l'un des outils permettant de répondre d'une façon parfaite à ce besoin.

2.6.1 Définition d'un système d'aide à la décision

Les systèmes décisionnels appartiennent à la famille des systèmes d'information, ils sont souvent appelés SID (système d'information décisionnelle) en anglais : DSS pour Décision Support System ou encore BI pour Business Intelligence.

Un SDI est un système permettant aux décideurs de l'entreprise de disposer des informations pertinentes et des outils d'analyse puissants pour aider à prendre les bonnes décisions aux bons moments [11].

Un SDI est défini aussi comme un regroupement de données orientées vers certains sujets,

intégrées, dépendantes du temps, non volatiles, ayant pour but d'aider les gestionnaires dans leurs prises de décision [12].

Nous avons constaté que les systèmes décisionnels sont définis de plusieurs façons, mais chaque nouvelle définition repose sur la même idée qu'un SDI a pour but d'aider à la prise de décision.

2.6.2 Objectif d'un système décisionnel

- Établir des prévisions et analyser la performance actuelle et prédire le futur.
- Le système décisionnel exige une flexibilité de la part des entreprises afin de répondre avec agilité aux besoins d'évolution, dans un monde en perpétuel mouvement où les variables de succès sont beaucoup et ne cessent d'évoluer.
- Il offre une vision globale de l'entreprise, en intégrant toutes les dimensions.
- Grâce aux outils d'analyse, nous pouvons avoir des estimations qui aident à la prise de décision.

L'entrepôt de données est un élément important dans l'architecture d'un SID car il représente le lieu de stockage, alors, sa mise en place est une étape primordiale pour la réalisation d'un SID.

2.7 Entrepôt de données

Le concept d'entrepôt de données est apparu dans les années 1980, alors que l'intérêt des gens pour les systèmes d'aide à la décision grandissait.

2.7.1 Définition d'un entrepôt de données

Un entrepôt de données est une base de données particulière destinée à l'aide à la décision. Il prend les données de diverses bases de données opérationnelles et d'autres sources de données telles que les fichiers et les transforme en de nouvelles structures mieux adaptées à la tâche d'analyse. Les entrepôts de données sont basés sur un modèle multidimensionnel, où les données sont représentées sous forme d'hypercubes, les dimensions correspondant aux différentes perspectives de l'entreprise et les cellules des cubes contenant les mesures à analyser [13].

2.7.2 Objectifs d'un entrepôt de données

La mise en place d'un entrepôt de données permet de :

- L'accès facile à l'utilisateur final et avec rapidité ou le résultat de sa requête doit être retourné dans un temps d'attente optimale
- Malgré les informations provenant d'un service de l'entreprise, elles peuvent être mises en corrélation avec celles d'un autre service, les données de l'entrepôt doivent être justes, assemblées d'une manière correcte à partir d'une variété de sources, nettoyées, et libérées afin d'être prêtes à la consommation.
- L'offre d'une vision globale de l'entreprise, en intégrant toutes les dimensions.
- Grâce aux outils d'analyse, nous pouvons avoir des estimations qui aident à la prise de décision.

2.8 Processus ETL (Extraction, Transformation, Load)

Pour alimenter une base de données d'une façon régulière, les données d'un ou de plusieurs systèmes opérationnels doivent être extraites et copiées dans une autre source, d'où la nécessité d'établir des flux de données entre la destination et les bases de données spécialisées où elles ont été créées. ‘

2.8.1 Définition

L'ETL est un type de logiciel permettant de collecter des données en provenance de sources multiples pour ensuite les convertir dans un format adapté à un Data Warehouse et les y transférer. ‘

2.8.2 Phases constructives d'un ETL

Le processus ETL se déroule en 3 phases qui sont [14] :

- **Extraction des données** : découvrir, analyser et extraire les données à partir de sources hétérogènes.
- **Transformation des données** : nettoyer et standardiser les données selon les règles d'affaires établies par l'entreprise.
- **Chargement des données** : charger les données dans un entrepôt de données et/ou les propager vers les datamarts.

2.9 Conclusion

Dans ce chapitre, nous avons étudié l'importance des données au sein des entreprises et les problématiques liées au système d'information. Les approches théoriques du MDM et DQM ont été présentées en introduisant tout d'abord les concepts de base tels que les données de référence. Nous avons aussi présenté les architectures du MDM, ces différentes architectures dépendent des besoins et des contraintes posés par l'entreprise. Enfin, nous avons défini quelques généralités sur les entrepôts de données et les systèmes décisionnels qui sont des notions importantes pour la réalisation de notre projet.

La plus grande difficulté rencontrée dans un projet décisionnel de n'importe quelle entreprise est de faire correspondre ses données de référence aux données de l'entrepôt de données, pour cela, dans le chapitre suivant nous appliquerons l'approche MDM sur le référentiel "Produit" de l'entreprise Cevital.

Conception du MDS

3.1 Introduction

Notre projet consiste à proposer une solution informatique pour la gestion des données de références, afin de créer une seule source de données centralisée et synchronisée. Le but est de réduire les redondances et éliminer la désynchronisation entre les différents services.

Dans ce chapitre, nous allons définir ce que notre système doit accomplir pour procurer les fonctionnalités souhaitées et recommandées par notre organisation en prenant en compte les notions théoriques. Pour mener à bien notre projet, nous avons suivi une démarche agile proposant un cycle de développement 2TUP « 2 Track Unified Process » avec un minimum de diagrammes UML. Nous allons entamer le processus de développement par une phase analyse qui va mettre en évidence les différents acteurs intervenant dans le système cible ainsi que leurs besoins. Puis une phase conception du MDM. Ce choix est justifié par la démarche de réalisation de notre application.

3.2 Environnement de l'entreprise

Chacun des trois services production-maintenance, achat et commercial dispose d'un système opérationnel (OLTP) pour organiser ses fonctions, et participer dans les tâches de gestion quotidiennes, pour les organiser dans une BDD propre à lui. L'entreprise dispose aussi d'une plate-forme BI pour la réalisation des analyses prédictives et l'automatisation de prise de décision. Ces tâches permettent à l'entreprise de prendre en charge :

- La surveillance de la qualité des produits qu'elle achète.
- La surveillance des prix et les quantités des produits en entrée et en sortie.

- La livraison qui doit être dans les temps.
- Le suivie de l'inventaire.
- La traçabilité de toute activité effectuée, par exemple la raison de la vente.
- Le suivie détaillé des ordres de production.

3.3 Identification des acteurs

Un acteur représente un rôle joué par une entité externe qui interagit directement avec le système étudié. Dans notre cas d'étude, nous avons mis en place un MDS, où nous sommes intéressés aux comportements de trois acteurs principaux qui sont : Super Administrateur, Superviseur des données et Utilisateur du SI.

Acteur	Rôle
Super Administrateur	<ul style="list-style-type: none"> - Construit le modèle référentiel - Définit les règles métiers de l'entreprise - Met en place la sécurité. - Gère les versions. - Gère les membres. - Gère les utilisateurs.
Superviseur de données	<ul style="list-style-type: none"> - Valide les transactions des utilisateurs des trois services - Reçoit des notification des anomalies dans les données - Gère la qualité des données
Utilisateur du SI (Services)	<ul style="list-style-type: none"> - Met à jour le référentiel de données

TABLE 3.1 – Identification les rôles des acteurs

3.4 Diagramme de cas d'utilisation

Un cas d'utilisation (use case) représente un ensemble de séquences d'actions réalisées par le système, et produisant un résultat observable intéressant pour un acteur particulier. Un cas d'utilisation modélise un service rendu par le système. Il exprime les interactions acteurs/système et apporte une valeur ajoutée « notable » à l'acteur concerné. L'objectif principal est que l'ensemble des cas d'utilisation doit décrire exhaustivement les exigences fonctionnelles du système. Nous présentons dans cette section le diagramme de cas d'utilisation (Figure 3.1) puis la description de chacun d'eux.

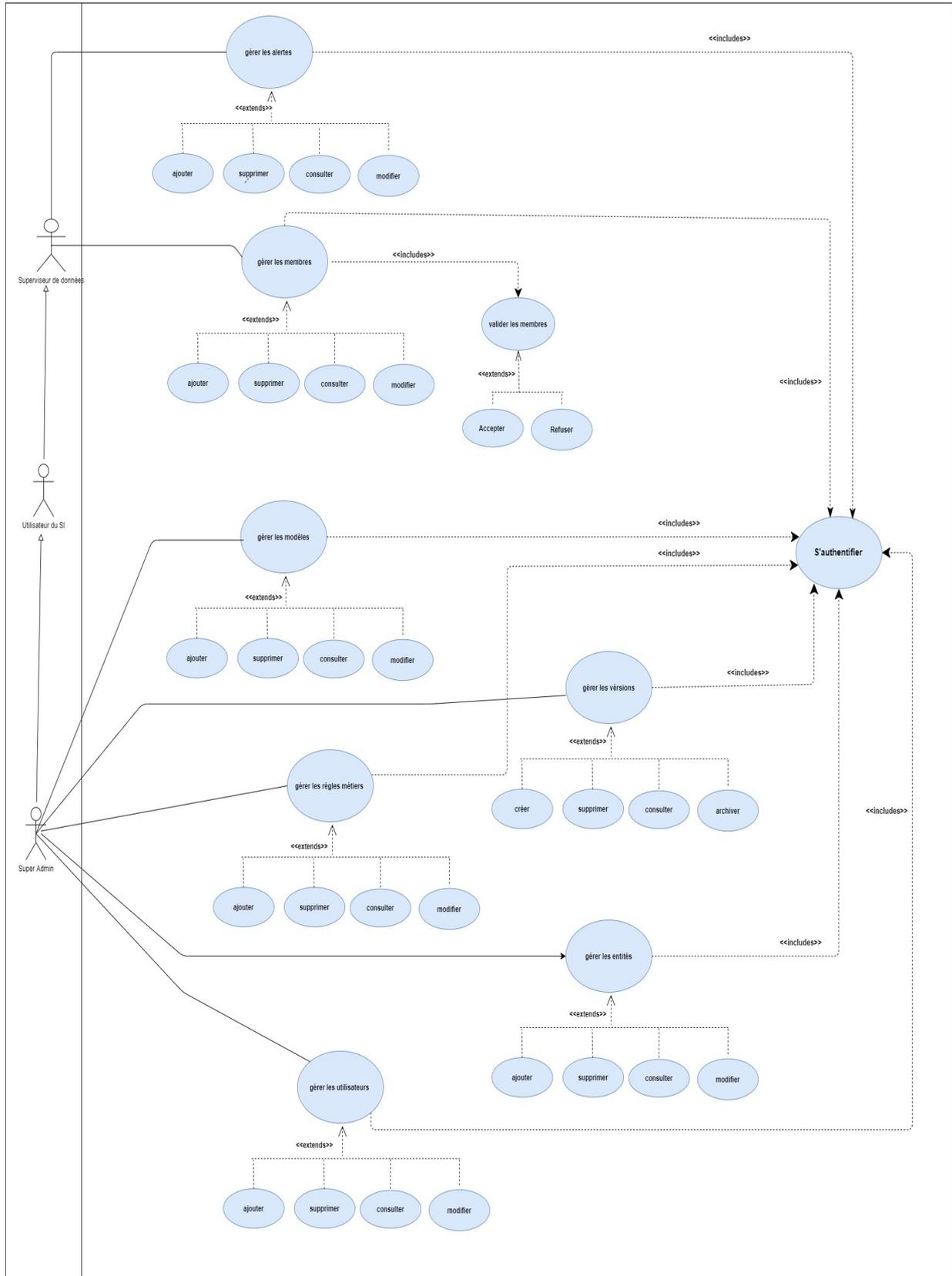


FIGURE 3.1 – Diagramme de cas d'utilisation du système

3.5 Description textuelle des cas d'utilisation

Dans cette section, nous présentons une description détaillée de quelques cas d'utilisation.

- ▷ **Gérer les Modèles** : Le Super Administrateur doit s'authentifier d'abord, pour pouvoir construire des modèles ou bien effectuer des mises à jour sur des modèles déjà existants.
- ▷ **Gérer les Membres** : Tous les utilisateurs doivent s'authentifier d'abord, pour pouvoir manipuler les données existants dans la BDD MDS, en ajoutant, supprimant, et modifiant des membres appartenant à un entité spécifique.
- ▷ **Valider les membres** : Les membres ajoutés ou bien manipulés par un utilisateur du SI doivent passer par le superviseur de données, ce dernier se charge de la validation ou non de la tâche.

3.6 Choix de l'architecture

Selon les besoins de l'entreprise, nous avons opter pour l'architecture coopération, dite aussi mixte ou bien hybride, et cela pour les raisons suivantes :

- La plus réaliste à choisir par rapport aux besoins de l'entreprise par exemple Avoir un référentiel unique regroupant l'ensemble des données.
- Une base de données unique et centrale qui conserve les entités de données de référence.
- Les autres systèmes tel que l'entrepôt de données peuvent être alimenté à partir de cette base de données centrale.
- Les données sont correctes et synchronisées vers tout les services.

3.7 Diagramme de classes

Le diagramme de classes est le point central dans un développement orienté objet, en analyse, il a pour objectif de décrire la structure des entités manipulées par les utilisateurs. En conception, le diagramme de classes représente la structure d'un code orienté objet ou, à un niveau de détail important, les modules du langage de développement.

En se basant sur les schémas des bases de données existants dans les trois services, nous avons pu modéliser leurs diagrammes de classes comme suit :

3.7.1 Service Production

La figure 3.2 représente le diagramme de classes de la solution logicielle utilisée dans le service Production.

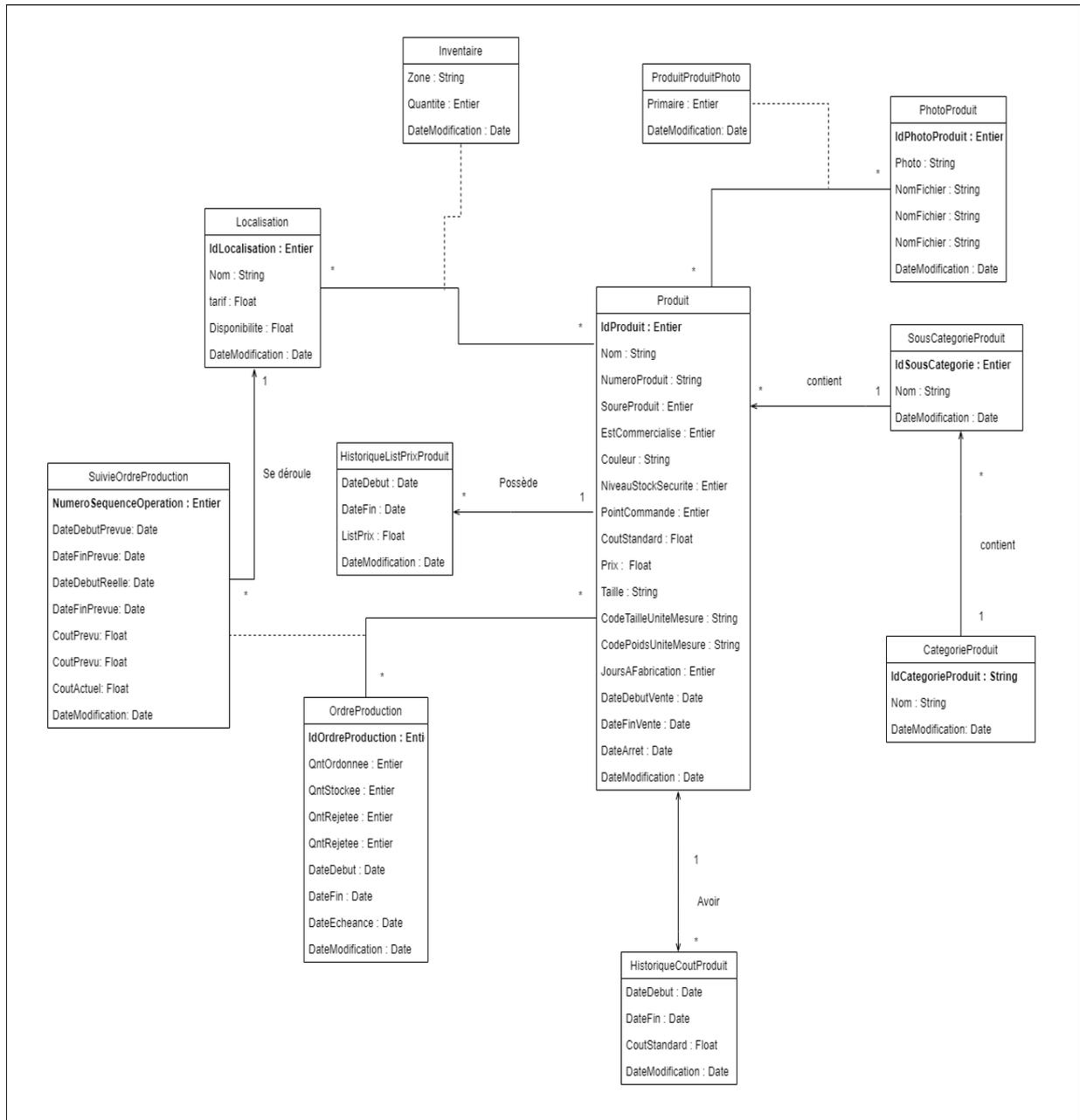


FIGURE 3.2 – Diagramme de classes du service de production

3.7.2 Service Achat

La figure 3.3 représente le diagramme de classes de la solution logicielle utilisée dans le service Achat.

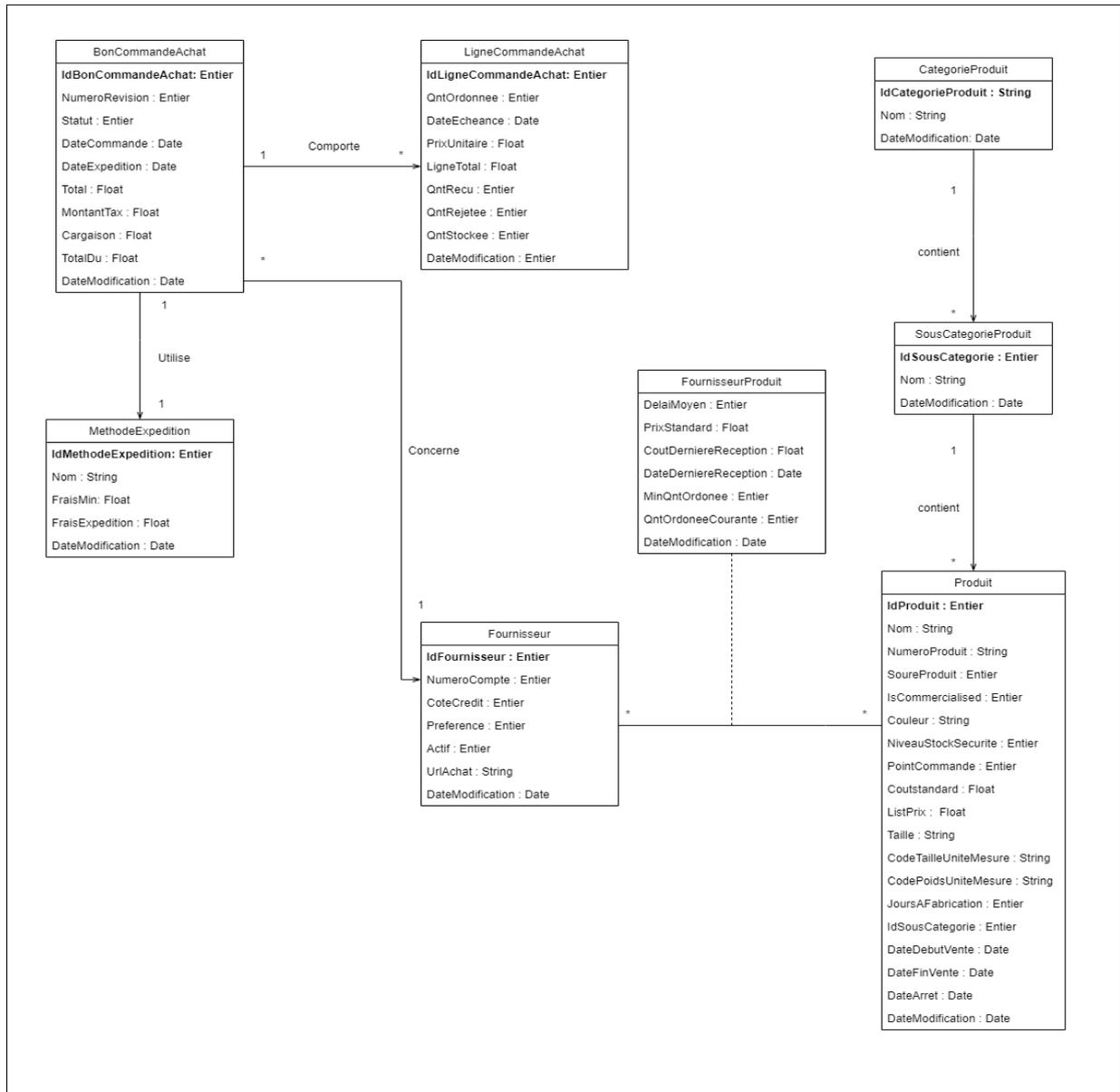


FIGURE 3.3 – Diagramme de classes du service d’achat

3.8 Dictionnaire de données

Le dictionnaire de données représenté par les tableaux 3.2, 3.3 et 3.4 regroupe toutes les données recensées, spécifiées et recueillies par rapport aux systèmes existants (et qui figureront donc dans les diagrammes de classes des trois services), et que nous aurons à conserver dans notre BDD MDS afin de réaliser le référentiel **Produit**.

Les tableaux 3.2, 3.3 et 3.4 définissent les caractéristiques des classes Produit ,Catégorie-Produit et SousCatégorieProduit et leurs types respectivement :Alphanumérique (AN) et Numérique (N) et Date (D).

Classe Produit :

Attribut	Type	Taille	Description de l'attribut
IDProduit	N	10	Clé primaire pour les enregistrements de produit.
Nom	AN	25	Nom du produit
Numeroproduit	AN	25	Numéro d'identification unique du produit.
SourceProduit	N	2	0 = Le produit est acheté, 1 = Le produit est fabriqué en interne.
IsCommercialised	N	2	0 = Le produit n'est pas un article vendable. 1 = Le produit est vendable.
Couleur	AN	10	Couleur de produit
NiveauStockSecurite	N	10	Quantité minimale d'inventaire.
PointCommande	N	10	Niveau de stock qui déclenche un bon de commande ou un ordre de travail.
Coutstandard	N	10	Coût standard du produit.
Prix	N	10	Prix de vente
Taille	AN	15	Taille du produit
CodeTailleUniteMesure	AN	15	Unité de mesure pour la colonne Taille.
CodePoidsUniteMesure	AN	15	Unité de mesure pour la colonne Poids.
Poids	N	10	Poids du produit
JoursAFabrication	N	10	Nombre de jours nécessaires à la fabrication du produit.
DateDebutVente	D	/	Date à laquelle le produit était disponible à la vente.
DateFinVente	D	/	Date à laquelle le produit n'était plus disponible à la vente.
DateArret	D	/	Date à laquelle le produit a été arrêté.
DateModification	D	/	Date de la dernière mise à jour de l'enregistrement

TABLE 3.2 – Dictionnaire de données de la classe Produit

Classe SousCatégorieProduit

Attribut	Type	Taille	Description de l'attribut
IdSousCatégorie	N	15	Clé primaire pour les enregistrements sous catégorie produit
Nom	AN	25	Description de la sous-catégorie.
DateModification	D	/	Date de la dernière mise à jour de l'enregistrement.

TABLE 3.3 – Dictionnaire de données de la classe souscategorieProduit

Classe CatégorieProduit

Attribut	Type	Taille	Description de l'attribut
IdCatégorieProduit	N	15	Clé primaire pour les enregistrements catégorie produit
Nom	AN	25	Description de la catégorie de produit.
DateModification	D	/	Date de la dernière mise à jour de l'enregistrement.

TABLE 3.4 – Dictionnaire de données de la classe CategorieProduit

3.9 Conception du référentiel Produit

Après étude des diagrammes de classes des différents services, nous avons constaté que pour la réalisation du référentiel Produit nous avons besoin de :

- Un seul Modèle.
- Cinq entités et leurs attributs.

3.9.1 Diagramme de classes

La figure 3.5 représente le diagramme de classes du Modèle Produit.

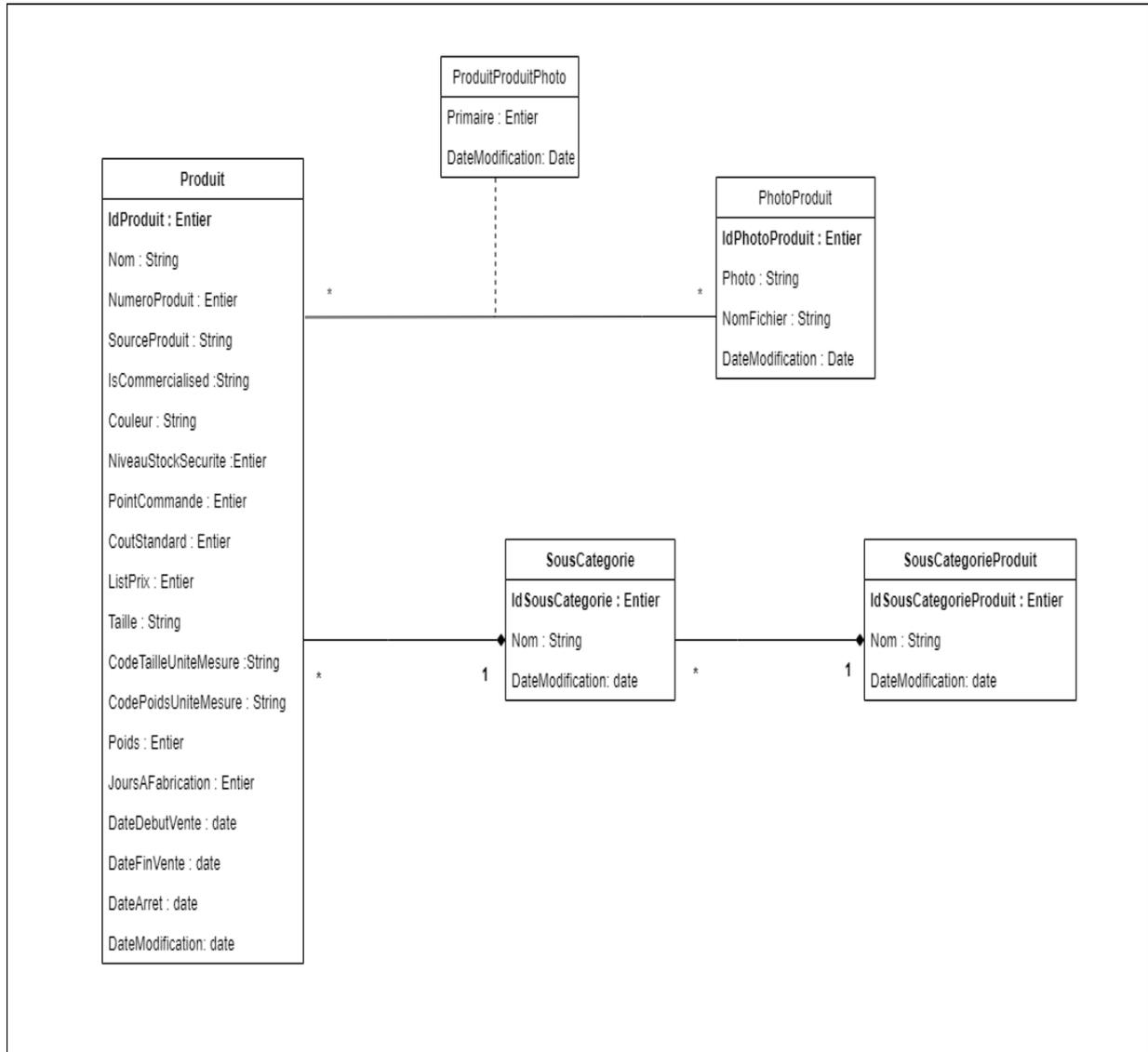


FIGURE 3.5 – Diagramme de classes du Modèle Produit

3.9.2 Modèle relationnel

Le modèle relationnel correspondant au diagramme de classes de la figure 3.5 est donné par les tables suivantes :

Produit (IdProduit, Nom, NumeroProduit, SourceProduit, isCommercialised, Couleur, NiveauStockSecurite, PointCommande, CoutStandard, ListPrix, Taille, CodeTailleUniteMesure, CodePoidsUniteMesure, Poids, JoursAFabrication, DateDebutVente, DateFinVente, DateArret, DateModification, #IdSousCategorie).

SousCategorie(IdSousCategorie, Nom, , DateModification, #IdProduitCategorie)

ProduitCategorie(IdProduitCategorie, Nom, , DateModification)

PhotoProduit (IdPhotoProduit, Photo, NomFichier, DateModification)

ProduitPhotoProduit (#IdProduit, #IdPhotoProduit, Primaire, DateModification)

3.10 Conclusion

Au cours de ce chapitre, nous avons présenté les modèles conceptuels de l'existant sur lesquels nous nous sommes basé pour mener une conception détaillée du travail à réaliser. Les résultats de cette partie sont un fondement pour l'étape de la réalisation qui sera expliquée dans le prochain chapitre.

Réalisation

4.1 Introduction

Dans ce chapitre, nous nous intéressons à la description de la phase de réalisation de notre solution. Nous commençons par la spécification des différents environnements de développement, matériels et logiciels. Nous présentons par la suite les différentes étapes d'installation et configuration des outils utilisés pour la mise en place de notre solution MDS. Nous terminons par quelques interfaces et explications de fonctionnement de l'application.

4.2 Présentation des outils utilisés

En premier lieu, on a utilisé le framework .NET [15] en vu de pouvoir utiliser le logiciel SQL Server 2019.

4.2.1 SQL Server 2019

Microsoft SQL Server 2019 [16] est à la fois un SGBD et une plate forme décisionnelle complète (ETL, OLAP, DATAMINING et Reporting). Il est développé et commercialisé par Microsoft .

SQL Server 2019 permet une efficacité informatique et développement en déployant des applications plus sûres, plus puissantes et plus fiables, et en obtenant une analyse décisionnelle très évolutive. Il s'articule autour d'un moteur de base de données qui assure la gestion et le stockage de données, en utilisant comme langage le T-SQL .



4.2.2 SQL Server Management Studio

Ce paragraphe concerne l'utilisation d'un outil indispensable à SQL Server 2019 qui est le SQL Server Management Studio [17] que l'on appellera SSMS .Il constitue l'outil majeur de SQL Server 2019. C'est grâce à lui qu'on peut créer des bases de données, travailler sur celles-ci, manipuler des données... SSMS est donc l'outil incontournable de SQL Server 2019 par son utilité.

4.2.3 Visual studio 2019

IDE de Microsoft [18], employé pour développer des solutions de gestion incluant des projets Analysis Services (SSAS), Integration Services (SSIS) et Reporting Services (SSRS). Chaque type de projet fournit des modèles permettant de créer les objets requis pour les solutions décisionnelles ainsi qu'une variété de concepteurs, d'outils et d'assistants destinés à manipuler ces objets.



4.2.4 Excel

Excel est un logiciel de la suite bureautique Office de Microsoft [19], il permet la création de tableaux, de calculs automatisés, de plannings, de graphiques et de bases de données interrogeables. Les fichiers de base de données Excel facilitent la saisie la recherche d'informations spécifiques.



4.2.5 SSIS (SQL Server Integration Services)

Cet outil [20] permet d'intégrer des données provenant de différentes sources dans une source cible, tout en donnant la possibilité d'effectuer des traitements sur les données manipulées.

4.2.6 DQS (Data Quality Services)

Cet outil DQS se compose de Data Quality Server et de Data Quality Client, qui sont tous deux installés en tant que partie intégrante de SQL Server [20].

DQS est un outil dédié à la gestion de qualité des données , il est axé sur les connaissances. DQS permet de créer une base de connaissances et de l'utiliser pour effectuer diverses tâches critiques de qualité de données, notamment la correction, l'enrichissement, la normalisation et la déduplication de vos données.

4.2.7 Draw io

Draw.io [21] est une application gratuite en ligne, accessible via son navigateur (protocole https) qui nous permet de dessiner des diagrammes ou des organigrammes. Cet outil nous propose de concevoir toutes sortes de diagrammes, de dessins vectoriels, de les enregistrer au format XML puis de les exporter.



4.2.8 Navicat Premium

Navicat Premium [22] est un outil de développement de bases de données qui nous permet de nous connecter simultanément à des bases de données MySQL, MariaDB, MongoDB, SQL Server, Oracle, PostgreSQL et SQLite à partir d'une seule application. Il est compatible avec des bases de données du Cloud telles Amazon RDS, Amazon Aurora, Amazon Redshift, Microsoft Azure, Oracle Cloud, Google Cloud et MongoDB Atlas. nous pouvons rapidement et facilement construire, gérer et maintenir nos bases de données.



4.3 Installation de SQL server 2019 management studio

SQL Server Management Studio (SSMS) est la console principale de SQL Server. Il permet de créer et de gérer les bases de données ainsi qu'administrer la configuration SQL Server. SSMS inclus les tâches suivantes :

1. Créer, modifier et supprimer des bases de données et des objets de base de données tels que des tables, des procédures stockées, etc.
2. Créer et maintenir des sauvegardes (manuelles ou planifiées).
3. Importer et exporter des données depuis / vers d'autres bases de données.

La première étape se fait par la sélection de "Nouvelle installation autonome SQL Server..."

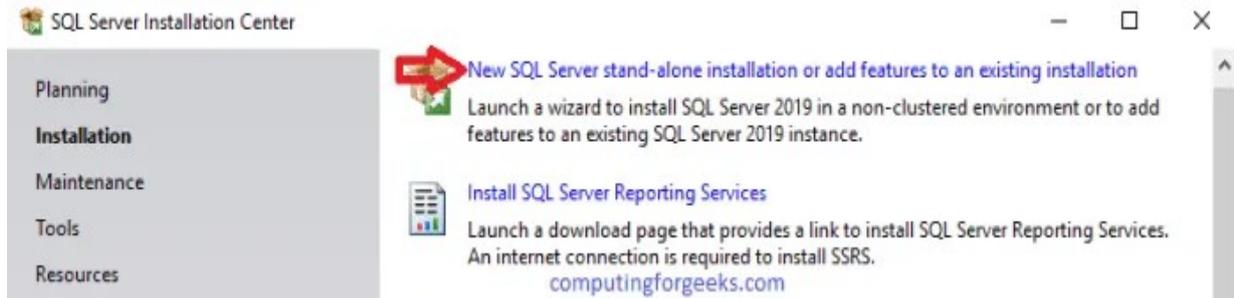


FIGURE 4.1 – Centre d’installation SQL Server

la deuxième étape consiste à confirmer notre accord avec les termes du contrat de licence. Maintenant, sur la page sélection des fonctionnalités, nous allons sélectionner les fonctionnalités indiquées dans la figure 4.2 pour continuer l’installation.

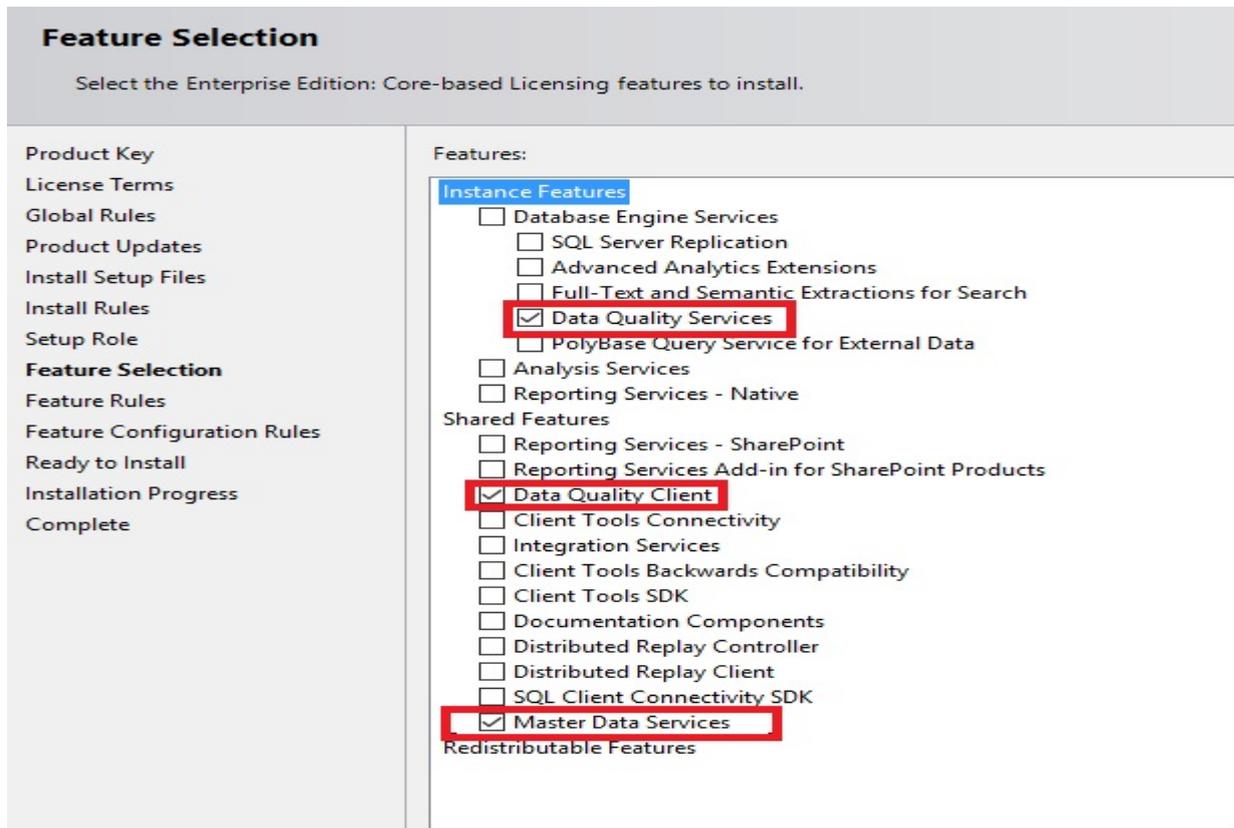


FIGURE 4.2 – Sélection de fonctionnalités

la troisième étape nous allons configurer le serveur , en mettant les paramètres par défaut.

la quatrième étape, nous choisissons un administrateur et définissons le mode de serveur pour la configuration d'analyses Services.

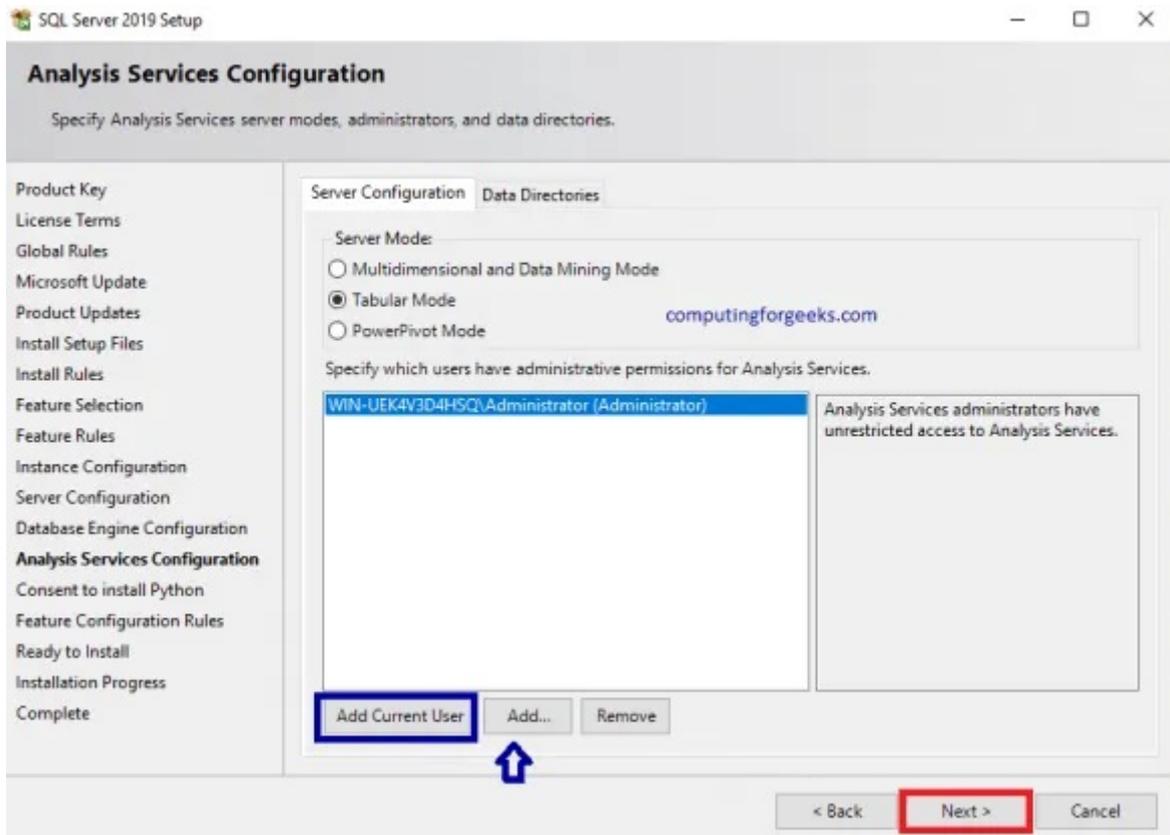


FIGURE 4.3 – Configuration d'analyses Services

La dernière étape est la fin de l'installation, nous recevons la confirmation que l'installation s'est terminée correctement.

4.4 Configuration de Master Data Services

L'installation de MDS effectuée séparément de l'installation de SQL Server à l'aide d'une interface d'assistant. L'assistant installe Master Data Services Configuration Manager, installe les fichiers nécessaires à l'exécution du service Web Master Data Services. Une fois l'installation terminée, nous ouvrons le gestionnaire de configuration Master Data Services.



FIGURE 4.4 – Gestionnaire de configuration Master Data Services.

Dans cette étape, nous allons configurer et installer d'IIS. Dans la fenêtre Panneau de configuration, nous allons cliquer sur l'option activer ou désactiver les fonctionnalités Windows du panneau latéral gauche.

Dans la page Fonctionnalités, nous sélectionnons que les fonctionnalités sélectionnées dans la figure 4.5 :

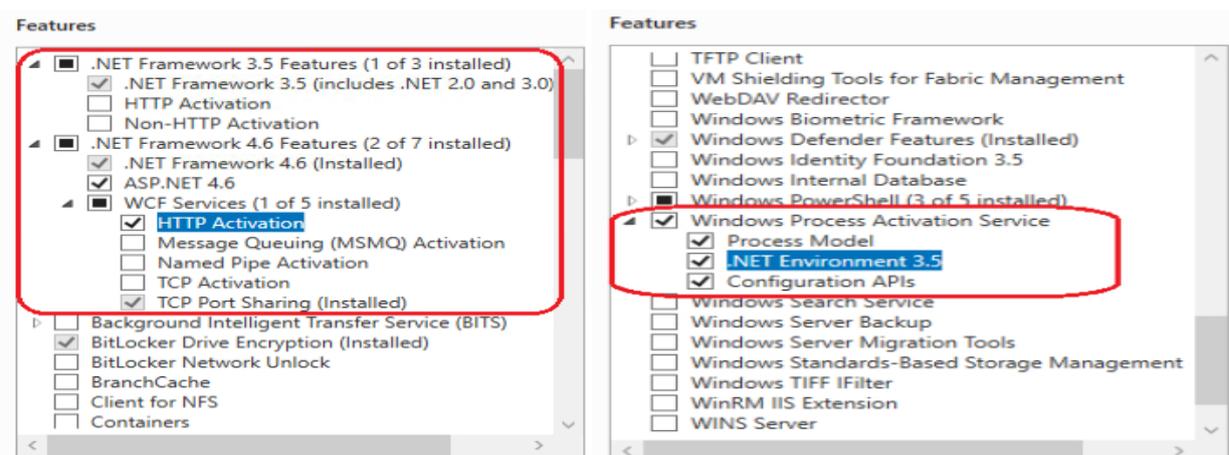


FIGURE 4.5 – Fonctionnalités d'IIS

une fois les fonctionnalités IIS installées, nous lançons le gestionnaire de configuration Master Data Services. Nous continuons à cliquer sur l'option Configuration de la base de données du panneau latéral gauche. Dans l'Assistant Création d'une base de données, nous cliquons sur créer une base de données.

Sur l'écran suivant de l'assistant, nous allons fournir des informations d'identification pour nous connecter à l'instance de base de données avec suffisamment d'autorisations pour créer la base de données MDS, nous allons créer une base de donnée Master Data Services (MDS) nommée **ProduitMDS**.

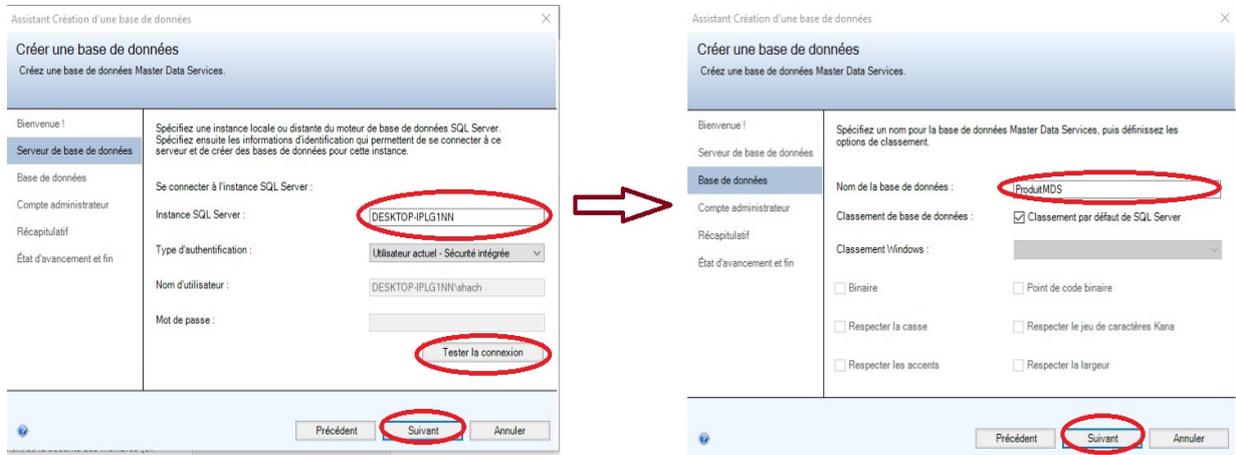


FIGURE 4.6 – L'assistant de création de base de données

Maintenant, sur la page suivante de l'assistant, nous allons fournir le compte windows qui aura accès à tous les modèles, la base de données est créée et configurée. Une fois que la base de données MDS a été créée, nous allons créer un nouveau site Web et une nouvelle application pour MDS et attribuer un port TCP pour accéder au site Web MDS.

Maintenant, que nous avons terminé l'installation et la configuration de MDS, nous ouvrons l'application Web MDS en utilisant le nom de l'ordinateur et le port que nous avons attribué à l'application.

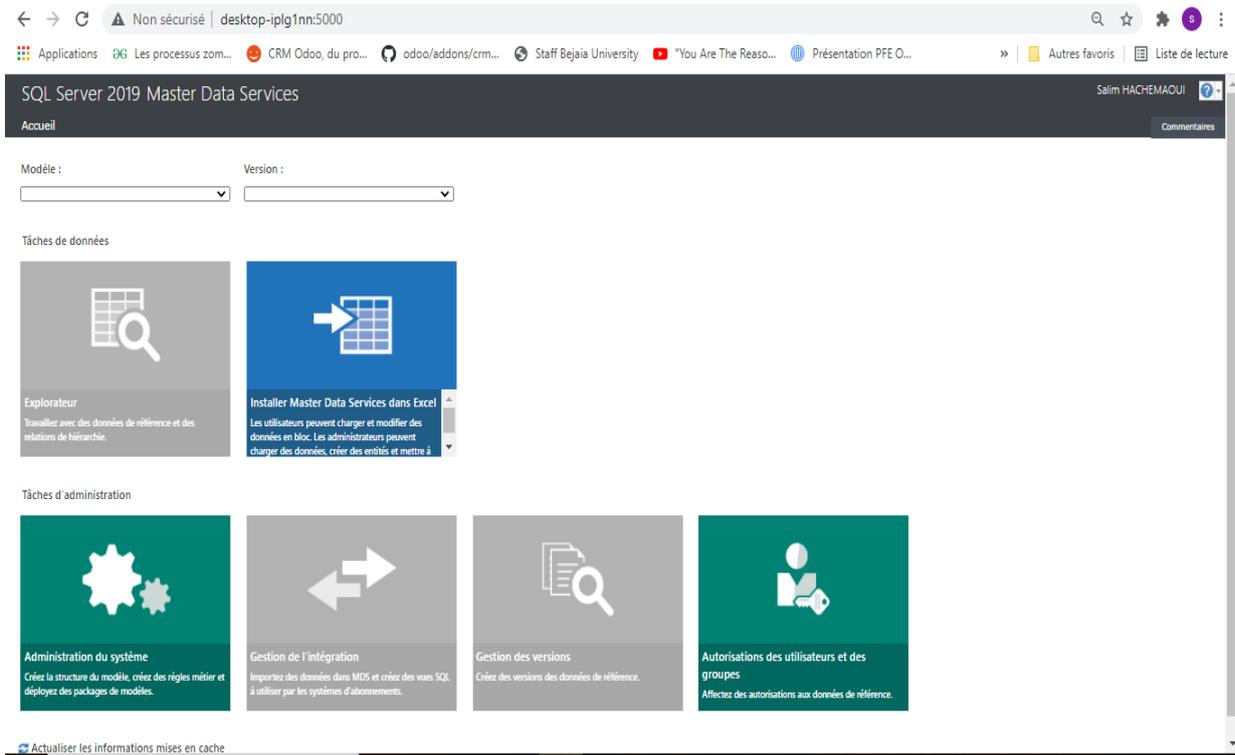


FIGURE 4.7 – Interface principale de MDS

Une fois le MDS est installé et configuré, nous allons déployer et exploiter la solution au sein de l'entreprise CEVITAL.

4.5 Déploiement et exploitation de la solution MDS

Une fois l'interface MDS chargée, dans le volet Administration système, nous pouvons créer les modèles et les entités ainsi que les attributs et les règles métiers de l'entreprise.

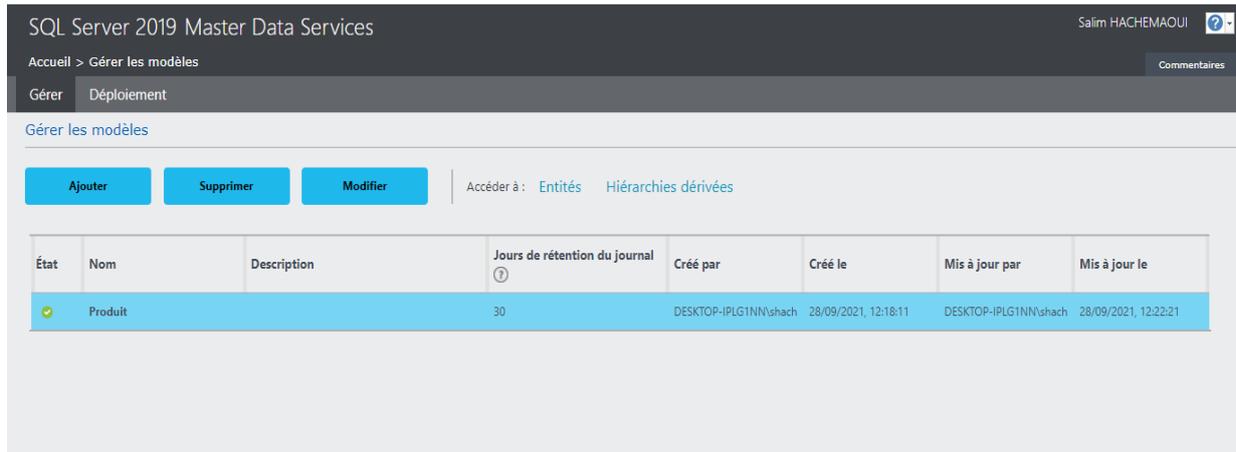


FIGURE 4.8 – Interface gestion des modèles

Le modèle est le niveau le plus élevé de l'organisation de données, après avoir créer le modèle, nous allons créer ses entités. Dans Master Data Services, nous créons une entité pour contenir des membres et leurs attributs.

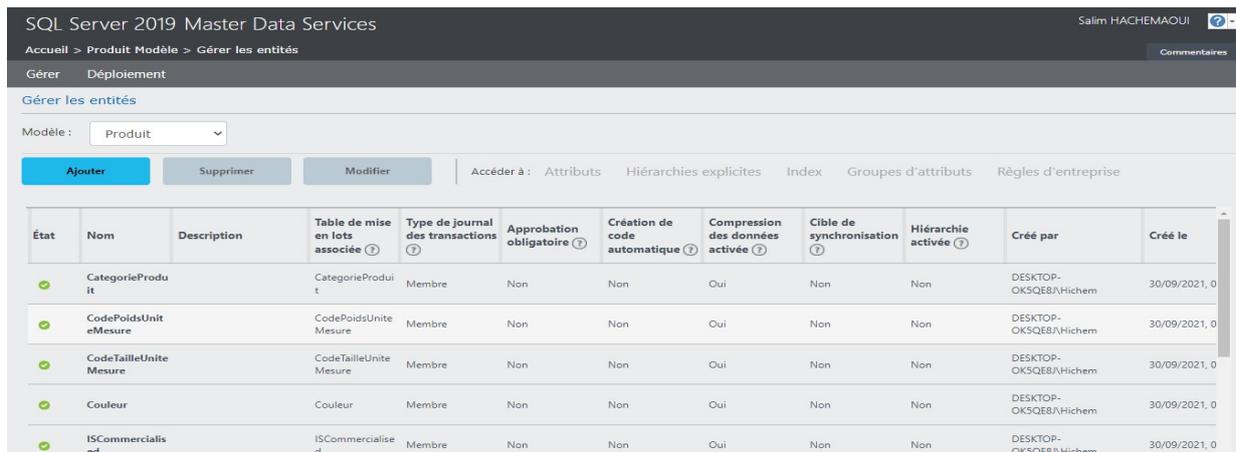


FIGURE 4.9 – Interface gestion des entités

Pour ajouter des attributs à l'entité créée , nous allons cliquer sur attributs.

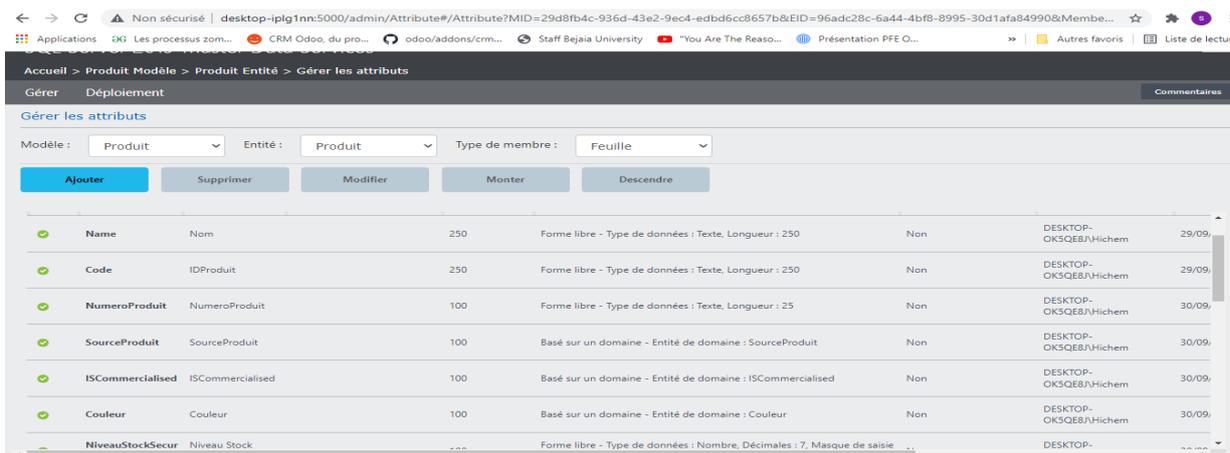


FIGURE 4.10 – interface gestion attribut pour l’entité Produit

4.6 Alimentation du référentiel Produit

L’alimentation de la BDD MDS est une étape très importante dans la réalisation de notre projet. Le but de cette étape est d’offrir une source cohérente et unique de la donnée référentielle Produit, pour réussir cette étape nous allons passer par plusieurs phases qui sont le contrôle de la qualité, le chargement des Données et l’attribution des rôles.

4.6.1 Contrôle de la qualité de données

En fonction des besoins et des règles métiers de l’entreprise CEVITAL, nous avons pu déterminer les points nécessaires à prendre en compte lors de la phase du contrôle de la qualité de données telle que la codification que les attributs doivent respecter.

- **Création de la base de connaissances**

La base de connaissances est un référentiel de métadonnées qui a pour but d’améliorer la qualité des données par le nettoyage et le rapprochement des ces données. Pour la création de cette base de connaissance, nous avons opter pour l’outil DQS que nous avons défini précédemment.

Une fois le DQS est installé et configuré, nous allons procéder à la création d’une nouvelle base de connaissances que nous avons appelé **ProduitBaseConnaissance**.

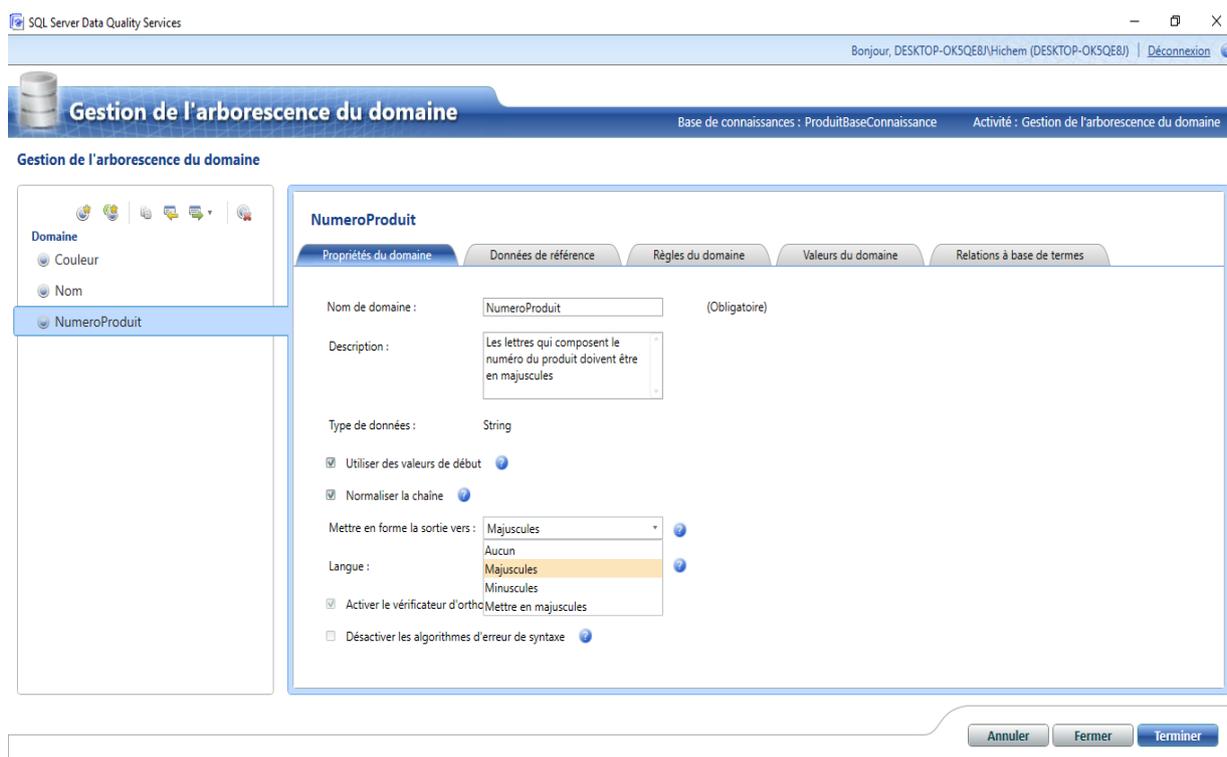


FIGURE 4.11 – Interface principale de DQS

Comme le montre la figure 4.11, la base de connaissances comporte trois règles de domaine qui s'appliqueront sur les attributs Couleur, Nom, NumeroProduit. Les lettres de l'attribut **NumeroProduit** doivent être en majuscule, ainsi que l'attribut **Nom** doit commencer par une majuscule. Cette base de connaissances peut être modifiée à n'importe quel moment par l'administrateur, cela en fonction des besoins de l'entreprise, où il pourra ajouter, modifier et supprimer des règles métiers.

Ces règles seront appliquées durant la phase du chargement des données. À l'aide de l'outil DSQ et cette base de connaissances, les données seront nettoyées avant leur chargement à la BDD MDS.

4.6.2 Partie intégration de donnée SSIS (SQL Server Integration Services)

Avant d'entamer la partie intégration, on a conçu nos données par nous-même, en modifiant table par table et champs par champs. Voici un aperçu des données sources.

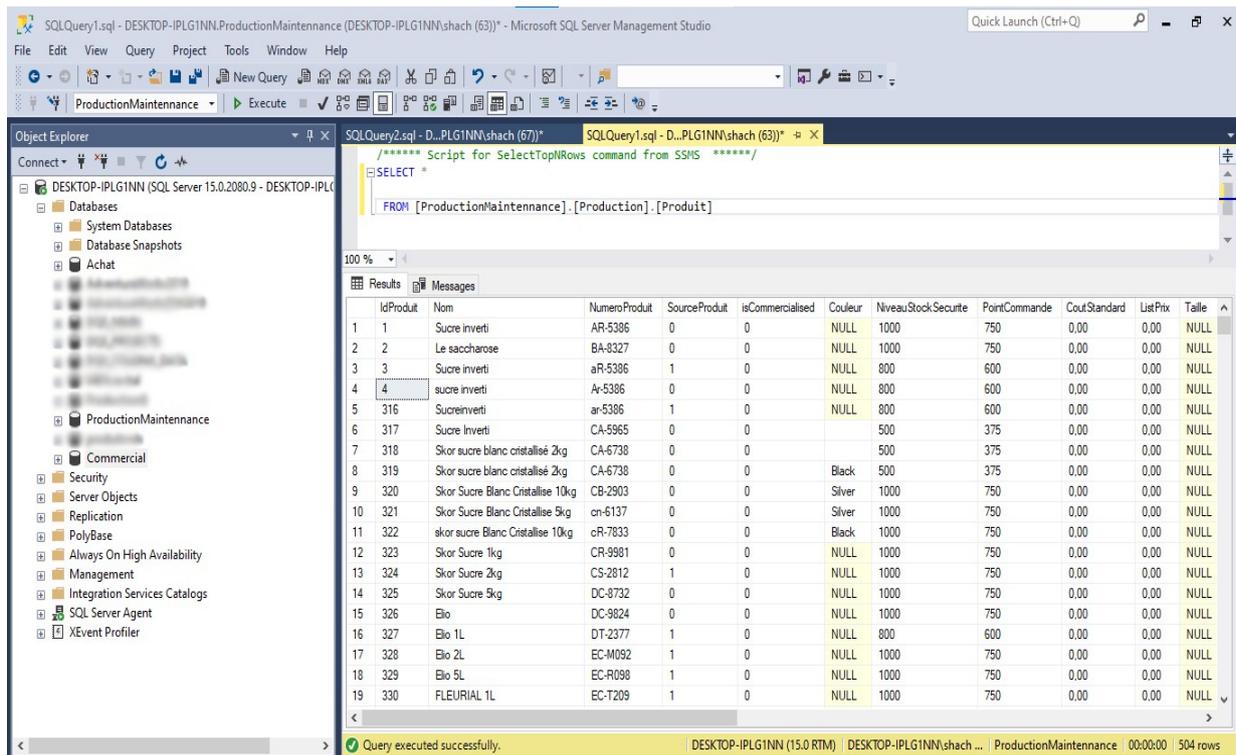


FIGURE 4.12 – Sources des données du projet.

L'étape suivante est l'extraction des données qui se trouve au niveau des 3 services (Production-Maintenance, Commercial, Achat), toute en créant des tables sur SSIS et les alimenter, puis les charger dans la base de donnée MDSCe vital.

Pour la réalisation de ce flux de données, nous avons poursuivi les étapes suivantes :

1. Créer trois connexions vers les bases de données des trois services en configurant trois composants **Source OLE DB**. Nous avons renommé ces trois composants respectivement BDD Achats, BDD Ventes, BDD Production-maintenance.
2. Appliquer la requête montrée dans la figure 4.13 pour chaque source de données OLE DB.
3. Relier chaque source de données à un composant **Nettoyage DQS** afin d'effectuer le nettoyage des données avant de les regrouper à l'aide du composant **Regroupement**. Le composant **Nettoyage DQS** est connecté à la base de connaissances créée précédemment.
4. Effectuer un trie avec élimination des doublants, cela juste après le regroupement des données.

5. Diffuser le résultat du trie grâce au composant **MultiDiffusion**, cela permet l'utilisation du résultat dans de plusieurs tâches.

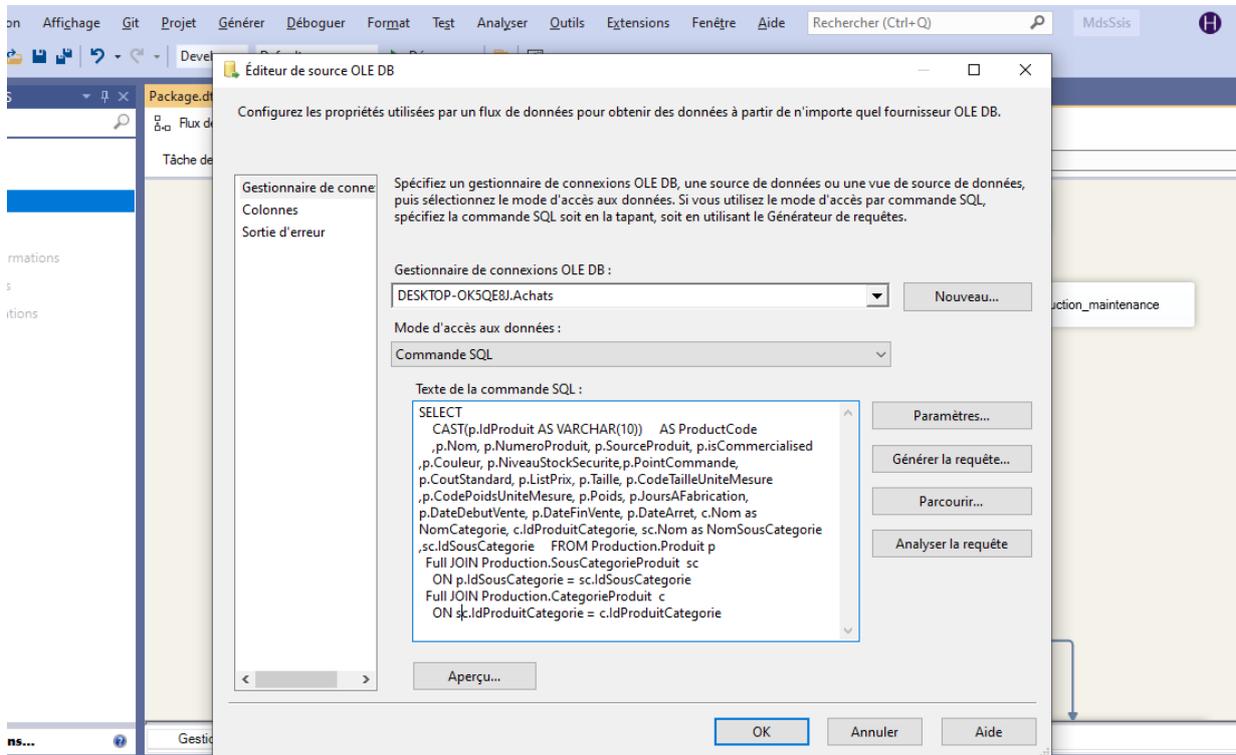


FIGURE 4.13 – Sources des données du projet.

À ce moment tous les produits avec leurs sous-catégorie et catégorie sont nettoyés et regroupés dans un seul emplacement (Multidifution). Nous procédons maintenant à l'insertion du résultat dans les entités cibles *Categorie*, *SousCategorie* et *Produit*. Nous continuons la construction de flux de données comme suit :

1. Créer trois connexions vers les entités *Categorie*, *SousCategorie*, *Produit* de la BDD *MDSceVital*, cela en configurant trois composants **Destination OLE DB**. Nous avons renommé ces trois composants respectivement **BDD mds-produit - Table Categorie**, **BDD mds-produit - Table SousCategorie**, **BDD mds-produit - Table Produit**.
2. Appliquer une sélection et une fonction d'agrégation du type **Group By** sur les attributs ***IdProduitCategorie***, ***NomCategorie*** afin d'insérer le résultat dans l'entité *Categorie* de la BDD *MDS*.

3. Appliquer la même chose sur les attributs **IdProduitCategorie**, **IdSousCategorie**, **NomSousCategorie** de l'entité SousCategorieProduit.
4. Effectuer un Mapping directement avec le résultat du trie et les attributs de l'entité Categorie.

MDSCEvital : est une base de donnée conçue pour centraliser les données issues de sources hétérogènes dans le but de faciliter les opérations d'analyse. Dans la figure 4.14, nous montrons le chargement des données à la BDD MDSCEvital.

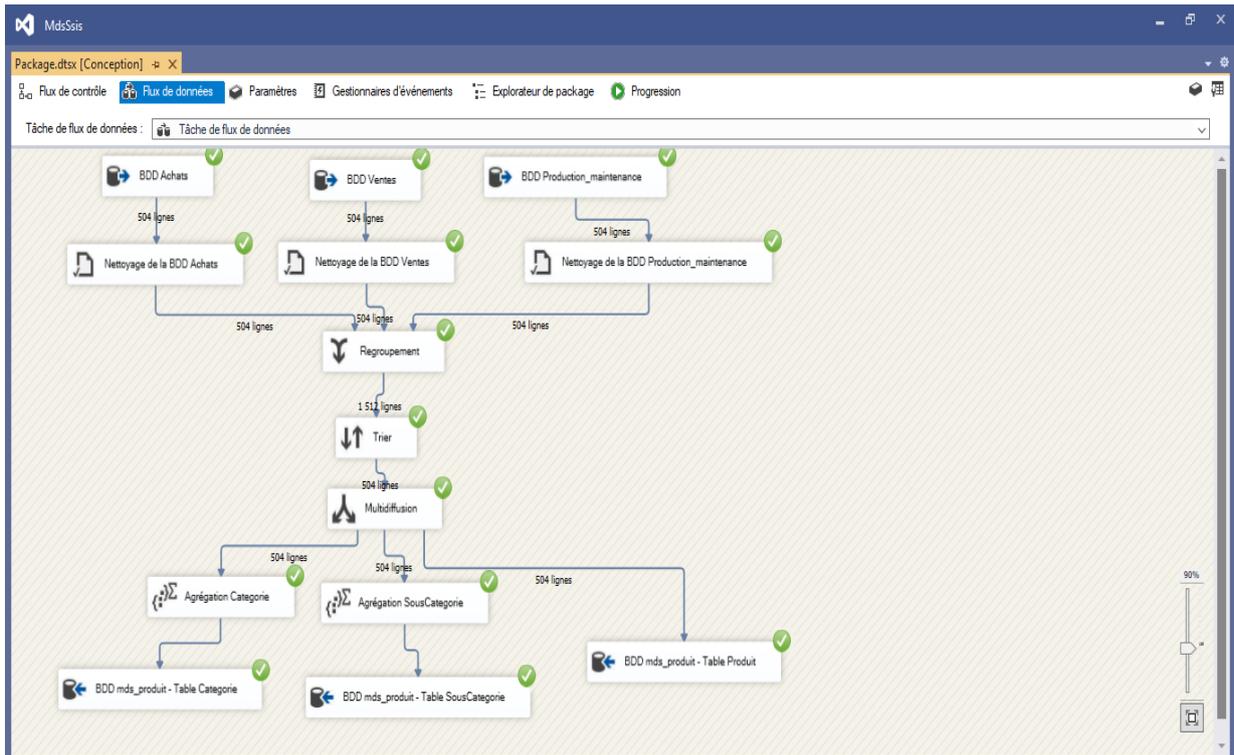


FIGURE 4.14 – Exécution de chargement et transformation des données.

On remarque que la couleur tourne en vert donc toutes les données sont extraites en bon état et on remarque que le nombre de ligne transportées est indiqué.

Le résultat de toutes ces transformations se retrouve dans la base de donnée MDSCEvital sous SQL server 2019.

La figure 4.15 présente l'entité produit après le chargement et la suppression des doublants.

SQL Server 2019 Master Data Services
 Accueil > Explorer (Modèle : Produit | Version : VERSION_1)

Entités Dépendances de l'entité Hiérarchies Collections Ensembles de modifications

Entité Produit

Ajouter un membre | Supprimer des membres | Appliquer des règles | Afficher l'historique | Filtrer | Paramètres | Excel

Service Production Service Achat Service Commercial Emballage [Tous les attributs]

	Name	Code	IdProduit	SourceProduit	isCommercialis
▶ ?	Le saccharose	BA-8327	2	False {F}	False {False}
▶ ?	LL Bottom Bracket	BB-7421	994	True {T}	True {True}
?	ML Bottom Bracket	BB-8107	995	True {T}	True {True}
?	HL Bottom Bracket	BB-9108	996	True {T}	True {True}
✓	Mountain Bottle Cage	BC-M005	871	False {F}	True {True}
?	Road Bottle Cage	BC-R205	872	False {F}	True {True}
?	Mountain-500 Black, 40	BK-M18B-40	989	True {T}	True {True}
?	Mountain-500 Black, 42	BK-M18B-42	990	True {T}	True {True}
?	Mountain-500 Black, 44	BK-M18B-44	991	True {T}	True {True}
?	Mountain-500 Black, 48	BK-M18B-48	992	True {T}	True {True}

1 - 50 of 500 membres Page 1 sur 10

FIGURE 4.15 – Entité Produit

4.6.3 Attribution des rôles

La spécification et l'attribution des rôles et des fonctionnalités aux utilisateurs de l'application est aussi une étape primordiale pour assurer la sécurité des données de l'entreprise. La solution MDS propose en général une sécurité à plusieurs niveaux, à savoir les fonctionnalités systèmes, les modèles, les entités et la manipulation des membres.

En suivant la conception faite dans le chapitre précédent, nous avons trois types d'acteurs dont le super Administrateur, le superviseur de données et l'utilisateur du SI. Nous allons appliquer l'architecture MDS Hybride afin de satisfaire le nombre et le type d'acteur de notre application.

La première configuration du MDS est considérée comme l'interface du **Super Administrateur** où il aura accès à toutes les fonctionnalités de l'entreprise, il est aussi considéré comme l'administrateur du système, et c'est le niveau le plus haut des autorisations. Alors, il va pouvoir :

- Il possède toutes les fonctionnalités d'administration système.
- Accéder à toutes les autres fonctionnalités de l'application.

La figure 4.16 présente l'interface principale du Super Administrateur, où il a accès à toute l'application.

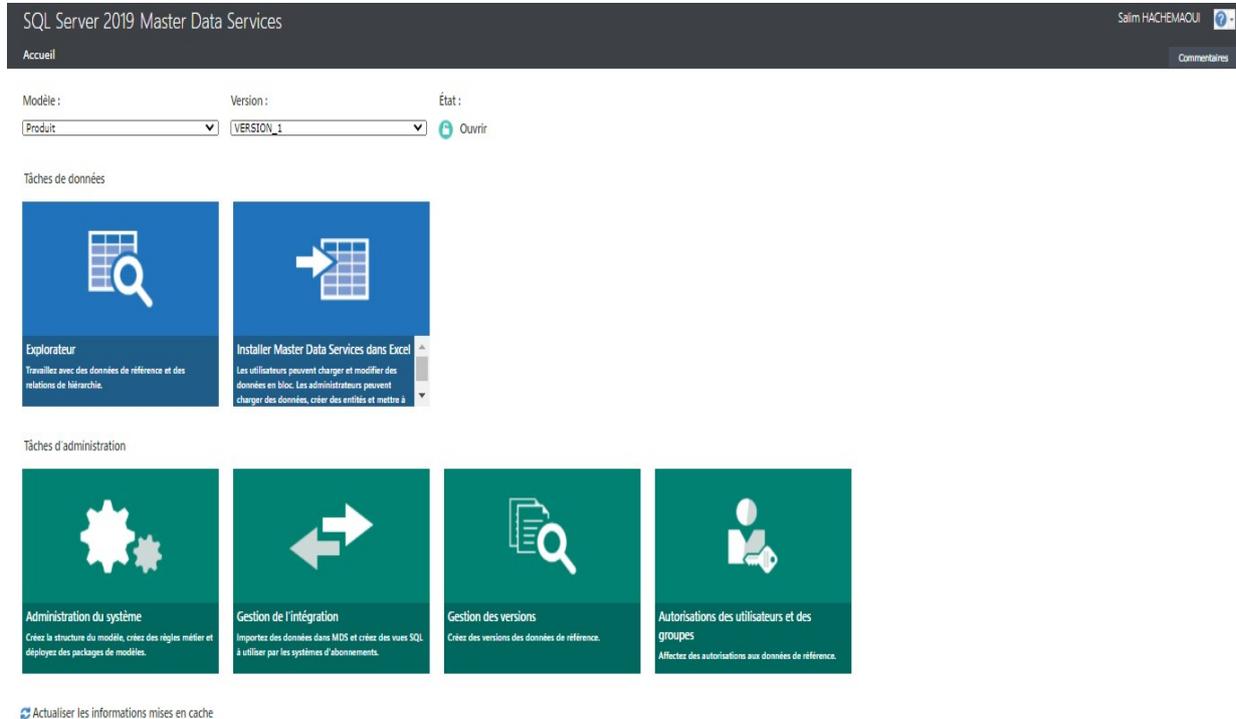


FIGURE 4.16 – Interface du super administrateur

Pour l'instant un seul utilisateur qui est présent dans l'application. En accédant avec le compte du Super administrateur nous pourrions créer le reste des utilisateurs avec les privilèges nécessaires pour chacun. Pour cela, nous accédons à l'interface dédiée à la gestion des utilisateur en cliquant sur **Autorisation des utilisateurs et des groupes**. La figure 4.17 représente les utilisateurs créés par le super administrateur, ou il pourra ajouter, modifier, supprimer les utilisateurs ou leurs privilèges.

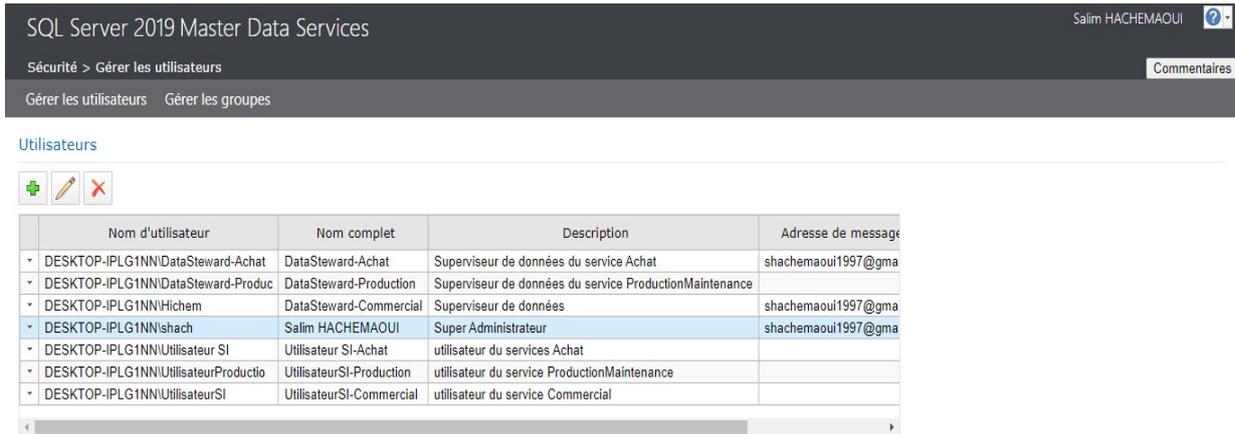


FIGURE 4.17 – Interface d’attribution des rôles

Pour satisfaire les besoins actuels de l’application, nous ajoutons l’ensemble des utilisateurs ainsi que leurs privilèges. La figure 4.17 montre aussi un tableau contenant tous les utilisateurs de l’application, et qui sont gérés par le Super Administrateur. La colonne description indique le type d’utilisateur ajouté.

La figure 4.18 représente l’interface principale du Superviseur de données, cette interface montre aussi toutes les fonctionnalités qui lui sont accessibles.

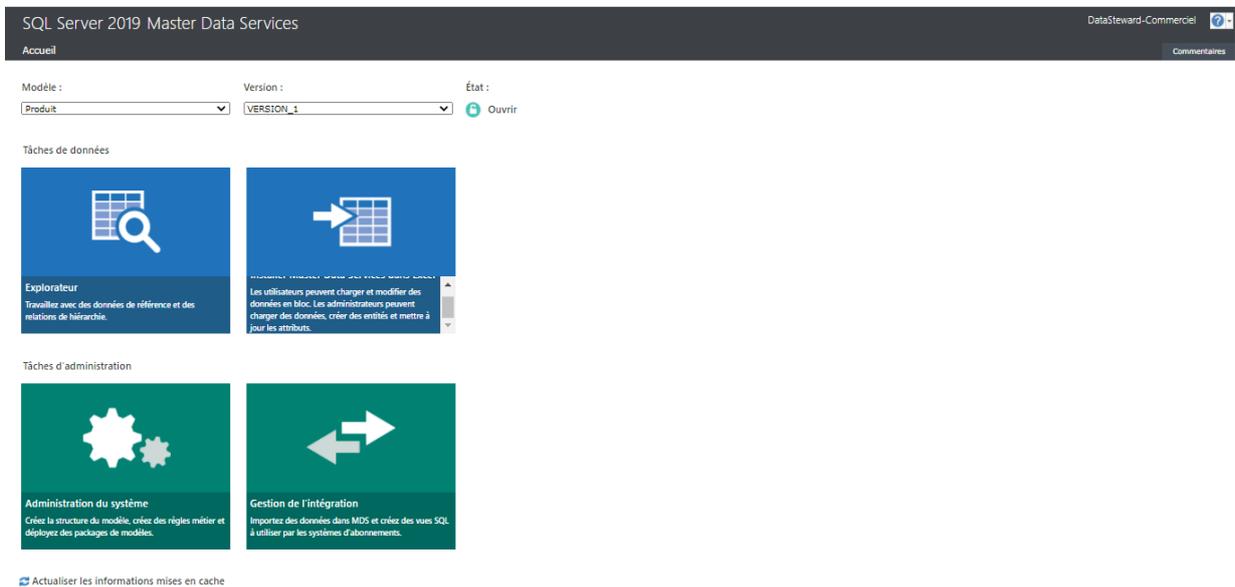


FIGURE 4.18 – Interface du superviseur de données

La figure 4.19 représente l'interface principale de l'utilisateur du SI, cette interface montre aussi toutes les fonctionnalités qui lui sont accessibles.

Notant que ce dernier utilisateur pourra accéder aux données afin d'effectuer les manipulations via l'interface Excel compatible avec MDS comme le montre la figure 4.20.

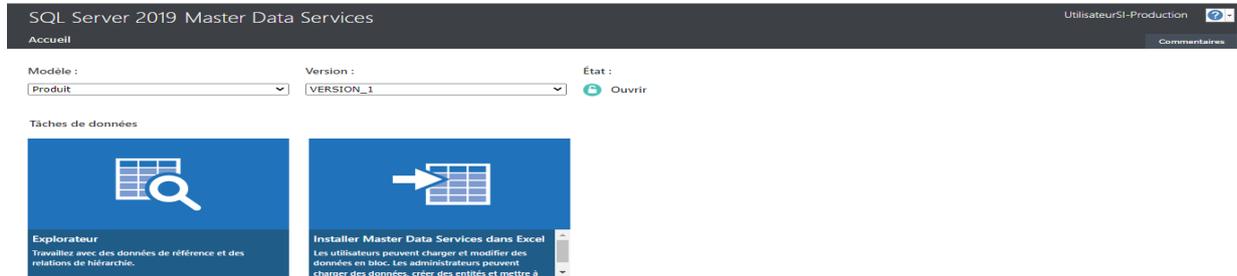


FIGURE 4.19 – Interface de l'utilisateur du SI

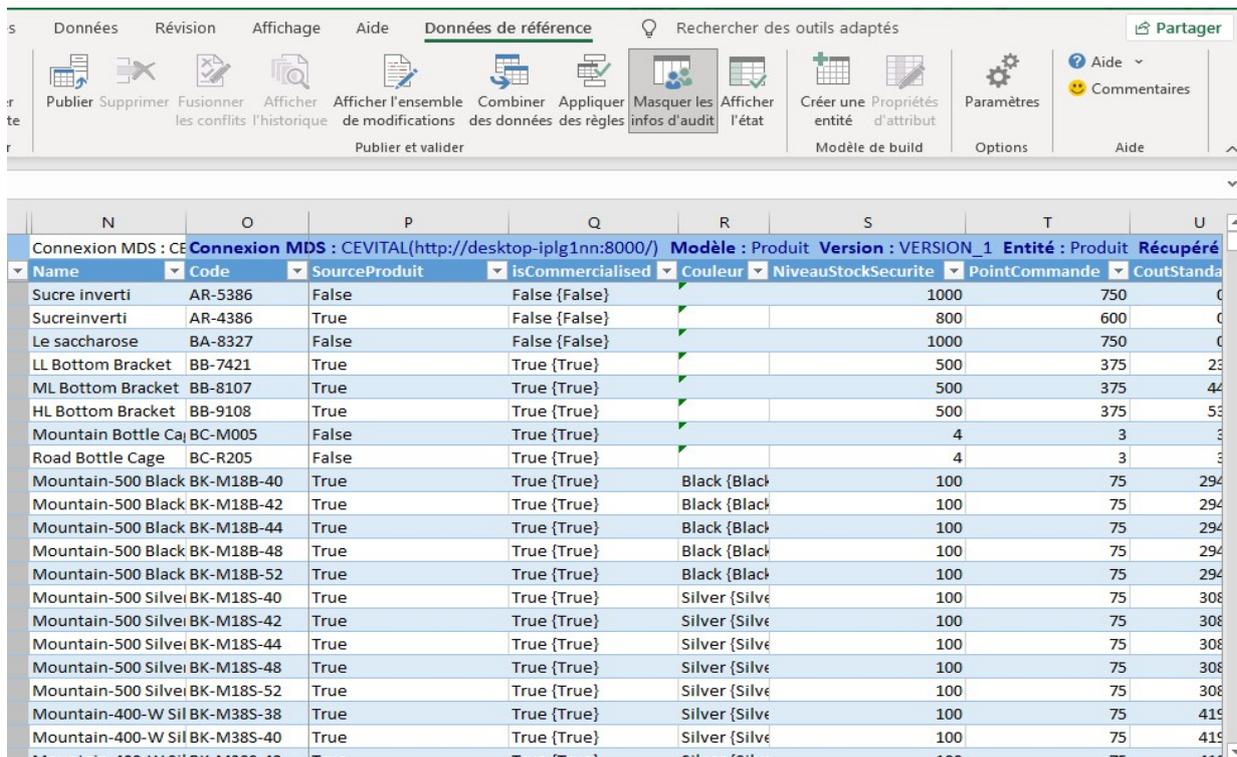


FIGURE 4.20 – Interface Excel pour l'utilisateur du SI

La visualisation des données diffère d'un service à un autre, cela en fonction des besoins

de leurs activités. Par exemple l'utilisateur du service production-maintenance n'a aucun droit d'effectuer des manipulations sur la donnée prix de vente d'un produit. Les figures 4.21 et 4.22 montrent les groupes d'attributs autorisés aux utilisateurs du service commercial.

SQL Server 2019 Master Data Services

Accueil > Explorer (Modèle : Produit | Version : VERSION_1)

Entités Dépendances de l'entité Hiérarchies Collections Ensembles de modifications

Entité Produit

Ajouter un membre | Supprimer des membres | Appliquer des règles | Afficher l'historique | Filtrer | Paramètres | Excel

Service Commercial

	Name	Code	isCommercialised	NiveauStockSecurite	CoûtStandard	
?	Le saccharose	BA-8327	False {False}	1000	0	0
?	LL Bottom Bracket	BB-7421	True {True}	500	23,9716	53,99
?	ML Bottom Bracket	BB-8107	True {True}	500	44,9506	101,24
▶ ?	HL Bottom Bracket	BB-9108	True {True}	500	53,9416	121,49
✓	Mountain Bottle Cage	BC-M005	True {True}	4	3,7363	9,99
?	Road Bottle Cage	BC-R205	True {True}	4	3,3623	8,99
?	Mountain-500 Black, 40	BK-M18B-40	True {True}	100	294,5797	539,99
?	Mountain-500 Black, 42	BK-M18B-42	True {True}	100	294,5797	539,99
?	Mountain-500 Black, 44	BK-M18B-44	True {True}	100	294,5797	539,99
?	Mountain-500 Black, 48	BK-M18B-48	True {True}	100	294,5797	539,99
?	Mountain-500 Black, 52	BK-M18B-52	True {True}	100	294,5797	539,99

1 - 50 of 500 membres Page 1 sur 10

FIGURE 4.21 – Interface utilisateur du SI du service commercial

La figure 4.22 représente l'entité de produit du service Production-Maintenance.

SQL Server 2019 Master Data Services
 Accueil > Explorer (Modèle : Produit | Version : VERSION_1)

Entités Dépendances de l'entité Hiérarchies Collections Ensembles de modifications

Entité Produit

|
 |
 |
 |
 |
 |

Service Production Appliquer des règles d'entreprise aux membres visibles uniquement

	Name	Code	SourceProduit	PointCommande	Taille	CodeT:
▶ ?	Le saccharose	BA-8327	False {F}	750		
?	LL Bottom Bracket	BB-7421	True {T}	375		
?	ML Bottom Bracket	BB-8107	True {T}	375		
?	HL Bottom Bracket	BB-9108	True {T}	375		
✓	Mountain Bottle Cage	BC-M005	False {F}	3		
?	Road Bottle Cage	BC-R205	False {F}	3		
?	Mountain-500 Black, 40	BK-M18B-40	True {T}	75	40	CM {CM}
?	Mountain-500 Black, 42	BK-M18B-42	True {T}	75	42	CM {CM}
?	Mountain-500 Black, 44	BK-M18B-44	True {T}	75	44	CM {CM}
?	Mountain-500 Black, 48	BK-M18B-48	True {T}	75	48	CM {CM}
?	Mountain-500 Black, 52	BK-M18B-52	True {T}	75	52	CM {CM}
?	Mountain-500 Silver, 40	BK-M18S-40	True {T}	75	40	CM {CM}
?	Mountain-500 Silver, 42	BK-M18S-42	True {T}	75	42	CM {CM}

FIGURE 4.22 – Interface utilisateur du SI de production-maintenance

4.7 Conclusion

Dans ce chapitre, nous avons présenté les outils utilisés dans notre projet, par la suite, nous avons présenté les étapes suivies pour l'installation et la configuration de SQL Server 2019 Master data services. Ensuite, nous avons procédé à une démarche de déploiement et l'exploitation de Master Data Service, tout en utilisant des figures qui montrent chaque étape de la mise en œuvre de la solution.

Conclusion générale

Dans le cadre de notre projet de fin de cycle, nous avons conçu et réalisé et mis en place un référentiel de données par le biais du Master Data Service (MDS) au sein de l'entreprise CEVITAL, cette solution est proposée pour répondre à la gestion séparée des services (service Production-Maintenance, service Achat, Service Commercial) qui causait des problèmes de redondance et de duplication des données, l'incohérence des données et l'absence de synchronisation entre les services impliquant un manque de qualité de données, par la suite des résultats imprécis lors de l'établissement des statistiques et des rapports pour des raisons de décisions.

Dans le but de remédier aux problèmes cités auparavant, nous avons, à l'aide du département de la transformation digital, initié le projet de la mise en place d'un référentiel Produit unique, regroupant l'ensemble de données des services, qui caractérisent tous les produits de l'entreprise. C'est-à-dire, que chaque service utilisera les mêmes données, ce qui permet à chaque acteur de manipuler la même donnée en temps réel. De plus, le MDS met en place une gestion de sécurité à plusieurs niveaux sur les données, où chaque acteur peut uniquement visualiser ou manipuler les données auxquelles il doit avoir accès. Et cela afin de fournir aux responsables de l'entreprise CEVITAL un accès à des informations fiables et pertinentes pour prendre plus intelligemment les décisions de gestion stratégiques.

Notre projet s'est déroulé en trois grandes parties, nous avons procédé par l'identification des données de référence, la modélisation et la création de la base de données MDS. Ensuite, vient la partie gestion de la qualité de données qui s'est faite avec l'outil DQS, qui nous a permis de concevoir et de créer une base de connaissances, avec laquelle nous avons pu supprimer l'ensemble des données incohérentes. La phase de la suppression de données est effectuée pendant le chargement des données au master data services en utilisant l'outil ETL de Microsoft SSIS. Le chargement des données correctes au MDS a permis d'avoir une vue unique et cohérente de données.

Au final, nous sommes passé à la réalisation et la configuration du MDS dont la création

des différents acteurs ainsi que leurs privilèges, cela a permis les insertions, les mises à jour et les suppressions appliquées sur tous les services en même temps.

La réalisation de ce projet, nous a permis de découvrir plus en détails les outils d'intégration des applications d'entreprise, d'approfondir nos connaissances dans le domaine de la gestion des données de référence, la gestion de la qualité des données et d'explorer plus en détail l'outil ETL (Extract Transform Load).

La réalisation de notre projet a été limité par le manque de temps, l'absence de complément de la donnée utilisée, fournie par l'organisme d'accueil, la durée de stage insuffisante afin de nous permettre de recueillir toutes les données dont nous avons besoin. Par ailleurs, notre projet peut s'étendre vers les perspectives suivantes :

- Création des rapports et des cubes de données d'autres activités afin de répondre au mieux aux attentes de l'entreprise entière.
- Étendre le champs de travail vers d'autres données de référence avec plusieurs règles métiers.

Bibliographie

- [1] Axel KAMALAK. «les apports de la méthode mdm dans la performance du si des entreprises ,mémoire de fin d'études. École privée des sciences informatiques tunisie», 2008. [en ligne ;visité :20/03/2021].
- [2] Hari Yadav. Services de données de référence – les bases. <https://www.red-gate.com/simple-talk/devops/database-devops/master-data-services-basics/>. en ligne ; visité :25/04/2021.
- [3] Jean-Christophe TTRIGAUX. «master data management mise en place d'un référentiel de données», 2009. en ligne, consulté :20/03/2021.
- [4] CEVITAL. <https://www.cevital.com>. en ligne ; visité :26/03/2021.
- [5] Ouardia TIRCHI. «la gestion de rémunération, un outil de motivation au service de la performance au travail» ,mémoire de magister,université mouloud mammeri de tizi-ouzou, 2012. en ligne ; visité :20/03/2021.
- [6] Jacques Finet Franck Régner-Pécastaing, Michel Gabassi. «mdm enjeux et méthodes de la gestion des données», novembre 2008.
- [7] Intd CNAM MARIKO Dominique. Mdm & dqm. http://www.dsi.cnrs.fr/description_referentiel/donnees/Documents/Synthese_CXP.pdf. en ligne ; visité : 15/03/2021.
- [8] Morley CHANTAL. « management d'un projet systÈme d'information », mai 2016.
- [9] Franck GAVAND Sébastien FANTINI. « business intelligence avec sql server 2014».
- [10] «the 16 best master data management tools (mdm solutions) for 2021». <https://solutionsreview.com/data-management/the-best-master-data-management-tools/>. en ligne ; visité :6/04/2021.

- [11] Jack E.Olson. « data quality, the accuracy dimension».
- [12] A.Devisy. « livre blanc e-business intelligence, l'effet internet sur le décisionnel»,aubay.
- [13] Esteban ZIMANYI Alejandro VAISMAN. E. (2014).« data warehouse systems, design and implementation » data-centric systems and applications.
- [14] Fatima Zohra YOUNSI. «mise en place d'un système d'information décisionnel pour le suivi et la prévention des epidémies» mémoire de fin d'études. université lyon 2, 2016. en ligne ;visité :20/05/2021.
- [15] Microsoft. Télécharger framwork .net. <https://dotnet.microsoft.com/download/dotnet-framework>". en ligne,visité :23/08/2021.
- [16] Microsoft. Télécharger sql server. <https://www.microsoft.com/en-us/sql-server/sql-server-downloads>. en ligne, visité :25/05/2021.
- [17] Microsoft. <https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver15>. en ligne, visité :29/05/2021.
- [18] Microsoft. <https://visualstudio.microsoft.com/fr/downloads/>. en ligne, visité :16/06/2021.
- [19] Microsoft. <https://microsoft-excel.fr.malavida.com/#gref>. en ligne ; visité :25/02/2021.
- [20] Microsoft. Télécharger ssis. <https://www.microsoft.com/en-us/sql-server/sql-server-downloads?rtc=1>. en ligne, visité :23/08/2021.
- [21] Gaudenz Alder. Outil pour dessiner des diagrammes en ligne. <https://app.diagrams.net/>. en ligne, visité :22/08/2021.
- [22] Navicat. Navicat premium. <https://www.navicat.com/fr/products/navicat-premium>. en ligne, visité :20/03/2021.