



MEMOIRE DE FIN D'ÉTUDES

En vue de l'obtention du diplôme de Master
En Management Stratégique et Système d'Information

**L'intégration de la virtualisation de données dans les Systèmes
d'Information Décisionnels (SID) : proposition d'une solution cible.**

Cas: Ministère des Finances

Élaboré par :

BADRI Badria

BOUSSAHA Abdelkader

Encadré par :

Dr. LADJOUZI Soumiya

MCA à l'ENSM Koléa

Co encadrante :

Mme SEGHIRI Narimane

Année universitaire 2024 – 2025

Résumé

Ce travail de recherche explore l'intégration de la virtualisation de données dans les Systèmes d'Information Décisionnels (SID) pour résoudre les défis des architectures traditionnelles, telles que les Data Warehouses et les processus ETL. Ces derniers étant rigides, lents et coûteux en duplication de données. La data virtualisation offre une vue unifiée et un accès en temps réel aux données hétérogènes sans les déplacer physiquement. Elle permet une agilité accrue, une réduction des coûts et une gouvernance des données améliorée. L'étude de cas au Ministère des Finances algérien, dont le SID existant est fragmenté, a démontré que cette technologie est une solution sécurisée et performante. L'implémentation d'un prototype avec la plateforme Denodo et son intégration à Power BI a validé la capacité à fournir des informations stratégiques, sécurisées et pertinentes, optimisant la prise de décision.

Mots clés : Virtualisation de données, SID, entrepôt de données, processus ETL, plateforme Denodo, Power BI, Ministère des finances.

Abstract

This brief explores the integration of data virtualization into Business Intelligence Systems (BIS) to solve the challenges of traditional architectures, such as data warehouses and ETL processes, which are rigid, slow and costly in terms of data duplication. Data virtualization offers a unified view and real-time access to heterogeneous data without physically moving it. It delivers greater agility, lower costs and improved data governance. The case study of the Algerian Ministry of Finance, whose existing SID is fragmented, demonstrated that this technology is a secure, high-performance solution. The implementation of a prototype with the Denodo platform and its integration with Power BI validated the ability to provide strategic, secure and relevant information, optimizing decision-making.

Key words: Data virtualization, Business Intelligence Systems (BIS), Data warehousing, ETL (Extraction, Transformation, Loading) processes, Denodo platform, Power BI, Ministry of Finance.

ملخص

يتناول هذا البحث موضوع دمج تقنية افتراضية البيانات في أنظمة المعلومات القرارية (SID) بهدف معالجة التحديات التي تواجهها البنى التقليدية مثل مستودعات البيانات (Data Warehouse) وعمليات الاستخراج والتحويل والتحميل (ETL) ، والتي غالبًا ما تكون بطيئة وتسبب تكاليف عالية نتيجة تكرار البيانات. توفر افتراضية البيانات رؤية موحدة وإمكانية وصول لحظية إلى البيانات المتنوعة دون الحاجة إلى نقلها فعليًا، مما يمنح المؤسسات مرونة أكبر، ويقلل التكاليف، ويحسن حوكمة البيانات. تم تطبيق هذه الدراسة على وزارة المالية الجزائرية، حيث أظهرت النتائج أن هذه التقنية تمثل حلاً آمناً وفعالاً لتحسين نظام المعلومات القرارية لديها. كما أن تنفيذ نموذج أولي باستخدام منصة Denodo وربطه مع أداة Power BI أكد قدرة هذه التقنية على تزويد صناع القرار بمعلومات استراتيجية، موثوقة وذات صلة، مما يعزز جودة عملية اتخاذ القرار.

الكلمات المفتاحية: افتراضية البيانات، أنظمة المعلومات القرارية، مستودع البيانات، عمليات الاستخراج والتحويل والتحميل ، منصة Denodo ، أداة Power BI ، وزارة المالية.

Remerciements

Tous d’abord, grâce à Allah le tout puissant de nous avoir donné la santé et la volonté d'entamer et de terminer ce mémoire.

Nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de ce mémoire de fin d’études.

Nos remerciements les plus sincères vont en premier lieu à **Mme LADJOUZI Soumiya**, notre encadrante, pour son accompagnement précieux, ses conseils avisés et sa disponibilité tout au long de ce travail. Son expertise et son soutien ont été d’une aide inestimable. Aussi, a notre co-encadrante **Mme SEGHIRI Narimane** pour son expertise et son soutien qui ont grandement contribué à l’aboutissement de ce projet.

Nous adressons également nos vifs remerciements à **M. NOUR Moncef**, notre encadrant de stage au sein du **ministère des Finances**, pour son encadrement bienveillant, ses orientations éclairées et l’opportunité qu’il nous a offerte de mettre en pratique nos connaissances dans un environnement professionnel enrichissant.

À nos **parents**, nous exprimons toute notre reconnaissance pour leur amour inconditionnel, leurs sacrifices et leurs encouragements constants qui nous ont permis de poursuivre nos études et d’atteindre ce jalon important.

À nos **épouses et époux**, merci pour votre patience, votre compréhension et votre soutien quotidien durant cette période exigeante. Vos encouragements ont été notre moteur pour aller jusqu’au bout de ce projet.

Enfin, une pensée spéciale à notre **petite famille** (enfants, frères, sœurs, etc.) pour leur affection et leur réconfort, qui ont illuminé nos journées même les plus chargées.

Ce mémoire est le fruit d’un travail d’équipe, tant sur le plan académique que familial, et nous en sommes fiers.

Avec toute notre gratitude,

BADRI Badria ;

BOUSSAHA Abdelkader.

Liste des abréviations

BI : Business Intelligence

BIS : Business Intelligence Systems

DWH : Data Warehouse (Entrepôt de données)

ETL : Extraction, Transformation, Loading (Extraction, Transformation, Chargement)

MdF : Ministère des Finances

DGNDSIE : Direction Générale de la Prévision, de la Numérisation et des Systèmes d'Information Economiques

IGF: Inspection Générale des Finances

SID : Systèmes d'Information Décisionnels

API : Application Programming Interface (Interface de programmation d'application)

ORM : Object Relational Mapping

ESB : Enterprise Service Bus (Bus de service d'entreprise)

EII : Enterprise Information Integration

VDP : Virtual DataPort (Composant principal de la Plateforme Denodo)

Power BI : Outil Microsoft de visualisation et d'analyse de données

CSV : Comma Separated Values

DBMS : DataBase Management System

RDBMS : Relational DataBase Management System

GDPR : General Data Protection Regulation

GOUV. : Gouvernance

IHM : Interface Homme-Machine

IT : Information Technology

OLAP : Online Analytical Processing

OLTP : Online Transaction Processing

ODS: Operational Data Store (Magasin de données opérationnel)

RNI: Référentiel National d'Interopérabilité

KPI: Key Performance Indicator (Indicateur clé de performance)

REST : Representational State Transfer

SGBD : Système de Gestion de Base de Données

SQL : Structured Query Language

VM : Virtual Machine

VPC: Virtual Private Cloud (Nuage privé virtuel)

XML : eXtensible Markup Language

Table des matières

Résumé.....	I
Abstract.....	II
ملخص.....	III
Remerciements.....	IV
Liste des abréviations.....	V
Table des matières.....	VII
Liste des tableaux.....	X
Liste des figures.....	XI
I. Introduction.....	3
I.1. Contexte algérien.....	4
I.2. Problématique.....	5
I.3. Objectifs de la recherche.....	6
I.4. Méthodologie de recherche.....	6
I.5. Plan de la recherche.....	6
Chapitre 1 : Revue de littérature et cadre conceptuel.....	9
1.1.Revue de Littérature.....	9
1.1.1.Contexte mondial.....	11
1.1.2.Situation en Afrique.....	11
1.1.3.Cas de l'Algérie.....	12
1.2. Cadre conceptuel.....	12
1.2.1.Les Systèmes d'Information Décisionnels (SID).....	12
1.2.1.1.Définition et évolution des SID.....	13
1.2.1.2.Architecture classique d'un SID.....	14
1.2.1.3.Les Niveaux de l'Architecture SID.....	15
1.2.1.4.Processus de Traitement des Données.....	16
1.2.1.4.1.Extraction, Transformation, Chargement (ETL).....	16
1.2.1.4.2.Analyse et Reporting.....	17
1.2.1.5.Architecture Technique.....	17
1.2.1.6.Enjeux et limites des approches traditionnelles.....	17
1.2.2.La Virtualisation de données.....	20
1.2.2.1.Définition et principes fondamentaux.....	20
1.2.2.1.1.C'est quoi la virtualisation des données?.....	20
1.2.2.1.2.Définition de la data virtualisation.....	21
1.2.2.2.Principales caractéristiques.....	22
1.2.2.3.Les concepts relatifs de la data virtualisation.....	22
1.2.2.4.Avantages.....	24
1.2.2.5.Limites et défis.....	25

1.2.2.6. Différence entre data virtualisation, ETL et data fédération.....	27
1.2.2.7. Différentes implémentations de la virtualisation des données.....	30
1.2.3. Intégration de la Data Virtualisation dans les Système d'Information Décisionnel.....	32
1.2.3.1. Rôle de la data virtualisation dans l'architecture décisionnelle moderne.....	32
1.2.3.2. Impact de l'intégration d'un serveur de virtualisation de données sur la prise de décision stratégique.....	32
1.2.3.3. Stratégies de mise en place de la virtualisation des données.....	33
Chapitre 2 : Cadre méthodologique et contexte organisationnel.....	39
2.1. Contexte organisationnel.....	40
2.1.1. Présentation de l'organisme d'accueil (Ministère des Finances)	40
2.1.2. Missions du Ministère.....	42
2.1.3. Ressources Humaines Informatiques.....	43
2.1.4. Compétences de la RH Informatique.....	43
2.1.5. Analyse du système décisionnel existant.....	43
2.1.5.1. Caractéristiques du SID existant.....	44
2.1.5.2. Forces et Faiblesses du SID existant.....	45
2.2. Cadre méthodologique.....	46
2.2.1. Le choix du sujet	46
2.2.2. La position épistémologique et méthode de recherche	50
2.2.2.1. La position épistémologique	50
2.2.2.2. La méthode de recherche.....	51
2.2.2.2.1. Les méthodes utilisées dans le terrain d'enquête.....	51
2.2.2.2.2. Guide d'entretien pour les utilisateurs opérationnels	54
2.2.2.2.3. Guide d'entretien pour les décideurs stratégiques	55
2.2.2.2.4. Guide d'entretien pour les experts techniques	55
Chapitre 3 : Modélisation de la Solution de virtualisation de données.....	58
3.1. L'Architecture cible du système décisionnel virtualisé.....	58
3.2. Vues de données virtuelles pour les décideurs.....	59
3.2.1. Définition d'une vue de données.....	59
3.2.2. Modélisation logique avec un moteur de virtualisation.....	60
3.2.3. Exemples des vues métiers :	60
3.2.4. Illustration des vues virtuelles Métier en Sql:.....	62
3.3. Mise en place d'un prototype fonctionnel.....	63
3.3.1. Les quatre (04) étapes principales du modèle fonctionnel.....	63
3.3.2. Les Résultats attendus du prototype:	64
3.4. Les règles de transformation et d'intégration des données.....	65
3.5. Élaboration des maquettes des tableaux de bord et des rapports décisionnels.....	65
3.6. Implémentation et validation de La solution proposée.....	68

3.6.1. Installation et Configuration de la plateforme de virtualisation choisie dans un environnement de test.....	69
3.6.1.1.La plateforme Denodo.....	69
3.6.1.2.Types de connexion de la plateforme Denodo.....	70
3.6.1.3.Installation et Configuration de la plateforme Denodo 9 Express sur Windows.....	70
3.6.2.Intégration des sources de données dans la Plateforme.....	79
3.6.3Création de quelques vues d'agrégation.....	87
3.6.4.Création de quelques vues métier (avec précalculs).....	91
3.6.5.Connexion de la plateforme Denodo avec un outil de visualisation (PowerBI)	101
3.6.5.1.Étapes Générales de Connexion.....	103
3.6.5.2.Connexion de PowerBI avec la BDD virtuelle (BDD_ministere_finances).....	103
Chapitre 4 : Discussion des résultats.....	114
4.1. Synthèse analytique des réponses des utilisateurs opérationnels.....	114
4.1.1.Difficultés majeures avec le système actuel.....	114
4.1.2.Attentes pour la solution future.....	114
4.1.3.Perception des bénéfices de la virtualisation.....	114
4.2.Analyse critique des réponses des décideurs stratégiques.....	115
4.2.1.Forces observées.....	115
4.2.2.Limites et défis majeurs.....	115
4.2.3.Perspective critique.....	116
4.3.Analyse critique des réponses des experts techniques.....	116
4.3.1.Points forts relevés.....	117
4.3.2.Défis identifiés.....	117
4.4.Synthèse analytique du déploiement du prototype de la plateforme de virtualisation des données.....	118
Conclusion.....	120
Bibliographie.....	122
Annexes.....	126

Liste des tableaux

Tableau 1	comparatif entre les notions suivantes : data virtualisation, ETL et data fédération.....	28
Tableau 2:	Liste des interviewés.....	53
Tableau 3	Architecture cible du système décisionnel virtualisé.....	58

Liste des figures

Figure 1: Les 4 Phases Du Processus De Business Intelligence, De La Donnee A L'information...	14
Figure 2 : Lorsque La Virtualisation Des Donnees Est Appliquee, Toutes Les Sources De Donnees Sont Presentees Comme Une Source De Donnees Integree.....	21
Figure 3 La Virtualisation Des Donnees.....	22
Figure 4 Organisation Du Ministère Des Finances.....	41
Figure 5 Organisation De La Dgndsie.....	42
Figure 6: Schéma Du Système Décisionnel Actuel Du Mdf.....	48
Figure 7: Schema Du Sid Cible Avec Data Virtualisation.....	49
Figure 8: Architecture De Haut Niveau De Gaiandb.....	59
Figure 9 : Schema Conceptuelle D'une Vue Virtuelle.....	61
Figure 10 : Rapport Decisionnel : Situation Macroeconomique Trimestrielle – Dgpp/Dgp.....	66
Figure 11: Rapport Trimestriel De Gouvernance Des Finances Publiques -- Dgndsie.....	67
Figure 12: Authentification Pour Acceder Au Site Officiel Denodo.....	71
Figure 13: Lancement Du Telechargement Sur Site.....	71
Figure 14: Fenetre D'installation De La Plateforme Denodo.....	72
Figure 15: Choix Du Repertoire D'installation.....	72
Figure 16: L' Icône Denodo Platform 9.....	73
Figure 17: Configuration De Denodo Platform Control Center.....	74
Figure 18: Demarrage Du Serveur Vdp Et Le Web Design Studio.....	75
Figure 19: Page De Connexion Denodo Design Studio.....	76
Figure 20: Page D'accueil Denodo Design Studio.....	77
Figure 21: Lancement De Vql Shell.....	78
Figure 22: Requette Sql Pour La Creation D'une Bdd Virtuelle.....	79
Figure 23: Creation D'une Table Sous Postgresql 16.....	80
Figure 24: Donnees De La Table Declarations_Douaniers.....	80
Figure 25: Creation Des Fichiers Plats De Type .Csv.....	81
Figure 26: Creation Des Sources De Type .Json.....	81
Figure 27: Integration D'une Nouvelle Source De Donnees.....	82
Figure 28: Interface Pour Le Choix De Type De La Source.....	83
Figure 29: Connexion A Une Bdd Sql.....	83
Figure 30: Creation De La Vue Virtuelle Declaration_Douaniers.....	84
Figure 31: Resultats De L'execution De La Vue Virtuelle Declarations_Douaniers.....	84
Figure 32: Integration Du Fichier Plat Loi_Finances.Csv.....	85
Figure 33: Execution De La Vue Virtuelle Loi_Finances.....	85
Figure 34 : Intégration D'une Source Api Json.....	86
Figure 35: Résultats De L'exécution De La Vue Virtuelle Indicateurs_Performance.....	86
Figure 36: Sources De Donnees Heterogenes Sur Une Seule Base De Donnees Virtuelle.....	87
Figure 37: Creation De La Jointure Pour La Vue Declarations_Enrichies.....	88
Figure 38: Condition De Jointure Pour La Vue Virtuelle Declarations_Enrichies.....	88
Figure 39: Resultats De L'execution De La Vue Virtuelle Declarations_Enrichies.....	89
Figure 40: Resultats De L'execution De La Vue Virtuelle Budget_Execution.....	90
Figure 41: Resultats De L'execution De La Vue Virtuelle Impact_Ajustements.....	91
Figure 42: Utilisation Group By Pour Le Filtrage.....	92
Figure 43: Calcul De La Somme Des Montants D'engagement.....	93
Figure 44: Resultats De L'execution De La Vue Metier Engagement_Paiement.....	93
Figure 45: Resultats De L'execution De La Vue Metier Engagement_Paiement.....	94
Figure 46 : Details Sur L'axe Publicite (Bdd Ministere Des Finances).....	96
Figure 47: Nombre D'actes Par Type Et Par Annee.....	97
Figure 48: Nombre D'actes Par Region.....	98
Figure 49: Filtrage De Structures Selon Le Nombre D'actes.....	99

Figure 50: Les Structures Les Plus Actives.....	100
Figure 51: <i>Top Type Acte Par Region</i>	101
Figure 52: Connexion Avec La Plateforme Denodo.....	104
Figure 53: Configuration De La Connexion Avec La Plateforme.....	105
Figure 54: Chargement De Donnees Sur Powerbi.....	105
Figure 55: Visualisation Du Meilleur Type D'acte Par Region Avec Powerbi.....	106
Figure 56: Chargement De Donnees De La Vue Impact Des Ajustements.....	108
Figure 57: Tableau De Bord Impact Des Ajustements.....	109

Introduction

I. Introduction

Au cours des deux dernières décennies, les besoins décisionnels des organisations ont radicalement évolué sous l'effet de la digitalisation, de l'explosion du volume des données et de la demande croissante d'analyses en temps réel. Initialement, les systèmes décisionnels s'appuyaient sur des Data Warehouses (**DWH**) centralisés, conçus pour stocker des données historiques structurées et permettre des requêtes complexes via des outils de Business Intelligence. Ces architectures, basées sur des processus **ETL** (Extract, Transform, Load), étaient adaptées à des environnements où les données étaient relativement stables et où les analyses pouvaient être réalisées par batch. Cependant, avec l'avènement du Big Data, de l'IoT et des applications cloud, les entreprises font face à des données hétérogènes, décentralisées et souvent non structurées, nécessitant une agilité que les DWH peinent à offrir.

Les limites des Data Warehouses traditionnels sont devenues de plus en plus apparentes. D'une part, le processus ETL, bien que robuste, est rigide et lent : l'extraction, la transformation et le chargement des données peuvent prendre des heures, voire des jours, ce qui retarde considérablement la disponibilité des informations pour les décideurs. D'autre part, ces architectures nécessitent une duplication des données, augmentant les coûts de stockage et créant des problèmes de cohérence entre différentes sources. Par exemple, une entreprise utilisant à la fois un CRM, un ERP et des logs IoT doit constamment synchroniser ces données dans son DWH, ce qui génère des goulots d'étranglement et des risques d'erreurs. Enfin, les DWH traditionnels ne sont pas conçus pour gérer des requêtes ad hoc ou des analyses en temps réel, ce qui limite leur capacité à soutenir une prise de décision agile dans des secteurs comme la finance, la santé ou le e-commerce.

Face à ces contraintes, les organisations ont exploré des alternatives comme les Data Lakes, capables de stocker des données brutes dans leur format natif. Cependant, sans gouvernance stricte, ces solutions peuvent rapidement devenir des "**marécages de données**", où l'information est difficile à exploiter. De plus, bien que les Data Lakes permettent une plus grande flexibilité, ils ne résolvent pas entièrement le problème de l'intégration en temps réel ni celui de la complexité des requêtes distribuées. C'est dans ce contexte que la virtualisation de données émerge comme une solution prometteuse, permettant d'abstraire l'accès aux données sans nécessiter de duplication, tout en offrant une vue unifiée et à jour pour les utilisateurs métiers.

Cette évolution reflète un changement de paradigme : les entreprises ne cherchent plus seulement à stocker des données, mais à y accéder de manière dynamique, souple et évolutive. Les architectures traditionnelles, bien qu'encore utiles dans certains cas, montrent leurs limites face à des besoins décisionnels toujours plus exigeants en rapidité, en granularité et en interactivité. La virtualisation de données, combinée à des approches modernes comme le cloud computing et l'analytique embarquée, pourrait ainsi représenter l'étape suivante dans l'évolution des systèmes décisionnels, en permettant une intelligence économique plus réactive et moins coûteuse.

I.1. Contexte algérien

La virtualisation de données émerge en Algérie comme une solution prometteuse pour répondre aux enjeux de la transformation numérique, bien que son adoption demeure encore limitée (APS, 2023) (webservices by ebs, s.d.). Le pays affiche une volonté marquée de renforcer sa souveraineté numérique, comme en témoignent les projets structurants tels que le Data Center gouvernemental développé avec Huawei et le Haut-commissariat à la numérisation (APS, 2023) (webservices by ebs, s.d.). La future loi sur la numérisation pourrait d'ailleurs constituer un levier important pour l'adoption de technologies innovantes comme la Virtualisation de données (APS, 2023). Cependant, plusieurs défis persistent : les organisations algériennes restent majoritairement attachées aux architectures traditionnelles (Data Warehouses physiques et processus ETL), malgré leurs coûts élevés (Boumediene, 2024), tandis que le manque d'expertise locale en solutions virtualisées freine le développement de ces nouvelles approches (Boumediene, 2024).

Pourtant, les opportunités offertes par la Virtualisation de données apparaissent particulièrement adaptées au contexte algérien. Sa capacité à réduire les coûts en évitant la duplication des données représente un avantage décisif (Boumediene, 2024). De plus, sa flexibilité et son déploiement rapide en font une solution idéale pour accélérer la modernisation des services publics et privés (APS, 2023) (webservices by ebs, s.d.). Si la virtualisation d'infrastructure (serveurs, stockage) s'est déjà implantée via des solutions comme VMware (Boumediene, 2024), la virtualisation des données spécifiquement dédiée à l'abstraction logique des sources reste encore marginale. Son essor dépendra notamment du développement de formations spécialisées pour les ingénieurs et décideurs, ainsi que de son intégration dans les stratégies nationales (cloud souverain, gouvernance des données) (webservices by ebs, s.d.). En définitive, la Virtualisation de données constitue une piste

sérieuse pour accompagner la transition numérique algérienne tout en respectant les impératifs de souveraineté nationale (APS, 2023) (webservices by ebs, s.d.).

I.2. Problématique

Le Ministère des Finances algérien est confronté à une multiplication rapide et une complexité croissante des sources de données utilisées dans son Système d'Information Décisionnel (SID). Face à ce contexte hétérogène et volumineux, il devient impératif d'intégrer une solution performante de virtualisation de données.

La question centrale qui guide cette démarche est :

Comment intégrer une solution de virtualisation de données afin d'améliorer l'exploitation des données massives dans le SID du Ministère des Finances ?

Plusieurs sous-questions s'en dégagent naturellement pour orienter la réflexion et le développement du projet :

1. Quel est l'état actuel du SID au ministère et quelles sont les limites qui entravent sa performance face à la diversité des données ?
2. En quoi la virtualisation des données peut-elle améliorer l'efficacité et la réactivité des systèmes d'information décisionnels, en particulier dans un environnement technique aussi complexe que celui du Ministère des finances ?
3. Quels sont les défis majeurs, aussi bien techniques qu'organisationnels, liés à l'intégration et à l'exploitation des données au sein du SID actuel ?
4. Comment concevoir et modéliser une architecture cible intégrant la plateforme Denodo, afin de répondre aux exigences fonctionnelles et techniques spécifiques du ministère ?
5. Quelles étapes méthodologiques et opérationnelles doivent être suivies pour une implémentation réussie de cette solution, depuis la conception jusqu'à son déploiement et son exploitation durable ?

Cette problématique pose ainsi un cadre précis pour guider la recherche et soutenir le processus de transformation numérique ministériel, où la virtualisation des données représente une réponse stratégique adaptée aux enjeux actuels de gestion et d'analyse d'une masse croissante d'informations.

I.3. Objectifs de la recherche

L'objectif de cette recherche est de comprendre comment l'intégration d'une solution de virtualisation de données dans un système d'aide à la décision influence le fonctionnement et l'efficacité dudit système. Il s'agit de présenter les avantages et défis de cette solution innovante par rapport à la solution traditionnelle qui est le data Warehouse, ainsi que les conditions favorables à cette intégration, en prenant en compte les points de vue des utilisateurs opérationnels, des experts techniques et décideurs. Enfin, cette étude vise aussi à proposer et tester un prototype d'une solution de virtualisation de données concrète au ministère des Finances, en tenant compte des aspects techniques (architecture, interopérabilité), organisationnels (gouvernance des données) et stratégiques (alignement avec les objectifs métiers) pour garantir une mise en œuvre harmonieuse et efficace.

I.4. Méthodologie de recherche

Notre étude adopte deux approches complémentaires. D'une part, une approche qualitative qui permet de mieux comprendre les aspects humains et organisationnels liés à l'intégration de la virtualisation des données : comment elle influence le fonctionnement et l'efficacité du système de décision, quels sont les bénéfices, les difficultés rencontrées, et les conditions nécessaires pour réussir cette intégration, vus du point de vue des utilisateurs, des experts techniques et des décideurs. D'autre part, une approche technique accompagnée d'une validation pratique, pour proposer une solution de virtualisation de données concrète que le ministère des Finances pourra appliquer facilement. Ainsi, la solution proposée prendra en compte non seulement les aspects techniques (architecture, interopérabilité), mais aussi les dimensions organisationnelles (gestion et gouvernance des données) et stratégiques (alignement avec les objectifs du ministère).

I.5. Plan de la recherche

Ce travail de recherche est scindé en trois chapitres complémentaires. Le premier chapitre, intitulé « Revue de littérature et cadre conceptuel », présente les principaux travaux ayant servi de fondement à notre étude, et établit le cadre théorique et conceptuel en présentant l'évolution des systèmes d'information décisionnels (SID) et les principes de la virtualisation de données, en suite l'intégration de la virtualisation de données dans les systèmes décisionnels.

Le deuxième chapitre, intitulé « Cadre méthodologique et contexte organisationnel », présente l'organisme d'accueil où nous avons effectué notre stage, expose les choix méthodologiques retenus pour la collecte et l'analyse des données ainsi que l'étude du système décisionnel existant au niveau du ministère des finances, aussi la conception d'une solution proposée et enfin la mise en place du prototype fonctionnel de la solution proposée (installation et configuration de la plateforme de virtualisation de données).

Enfin, le troisième chapitre est consacré à l'analyse des résultats issus des entretiens et des évaluations menées, afin d'appréhender le concept innovant de la virtualisation des données et son intégration aux systèmes d'aide à la décision, et aussi une synthèse analytique du déploiement du prototype de la solution proposée au ministère des finances.

Chapitre 1 : Revue de littérature et cadre conceptuel

Chapitre 1 : Revue de littérature et cadre conceptuel

1.1. Revue de Littérature

Durant les dernières années, plusieurs travaux de recherche ont exploré l'intégration de la virtualisation de données et son impact dans les systèmes décisionnels, voici une synthèse de quelques articles clés :

(Van der Lans, 2012) met en évidence dans son les défis liés à la création et la maintenance des DWH rationnels, quand les données proviennent de multiples sources hétérogènes, l'auteur explique comment la virtualisation des données résout ces problèmes, en basant sur la création des tables virtuelles à partir des divers types des sources (BDD relationnelles, service web, des fichiers plats..etc), selon l'auteur la virtualisation de données est présentée comme une solution qui offre une grande flexibilité et des capacités d'optimisation des requêtes supérieures à celles d'une base de données unique.

(Albadri, 2016) propose avec son ouvrage un modèle de conception pour la virtualisation des données afin de permettre la prise de décision en temps quasi-réel dans les environnements de BI. Il souligne que la virtualisation des données est une solution pour surmonter les limitations des processus ETL traditionnels, qui ne peuvent pas fournir d'informations en temps réel pour la prise de décision.

(Serano, M., 2018) avec son article du groupe IBM Software. Cet article explore le rôle croissant de la virtualisation des données dans les architectures décisionnelles hybrides (cloud et on-premise). L'auteur y montre que la virtualisation permet d'abstraire les silos de données tout en garantissant des performances proches des systèmes centralisés. IBM y présente une architecture cible basée sur une couche de virtualisation interagissant avec des sources hétérogènes en temps réel. L'article insiste sur la gouvernance des données et l'importance des métadonnées partagées dans les environnements complexes. Des cas d'usage sont présentés, notamment dans les services publics et la finance, où l'agilité et la conformité réglementaire sont prioritaires. L'article conclut que la virtualisation est aujourd'hui un pilier incontournable des systèmes décisionnels modernes.

(Barrera, J. et al., 2021) a publié un article sur "Journal of Big Data". Une étude menée par les auteurs de cet article explore comment la virtualisation de données permet d'améliorer les

performances décisionnelles en temps réel dans des environnements hétérogènes. Elle démontre que la data virtualisation réduit significativement le temps d'accès à l'information.

(Aaliyah, J., 2023) avec son article démontre comment la virtualisation des données transforme l'analyse du Big Data. Elle explique que les entreprises n'ont plus le temps d'attendre des traitements par lots avec les approches classiques telles qu'ETL : elles ont besoin d'analyses rapides, dynamiques et connectées. La virtualisation est présentée comme une réponse efficace aux défis du volume, de la variété et de la vitesse. L'auteure insiste sur la capacité de la virtualisation à unifier des données structurées, semi-structurées et non structurées issues de bases SQL, fichiers plats, APIs, etc. Elle discute aussi des risques (latence, sécurité) et propose des bonnes pratiques d'implémentation, comme le masquage des données sensibles. Son article est très utile pour comprendre les apports spécifiques de la virtualisation dans les projets Big Data.

(N. Huda & S. Roy, 2023) ont défini la virtualisation des données comme une couche logique qui fournit un accès en temps réel aux données métiers, en se connectant à des sources de données hétérogènes sans les dupliquer, cette solution permet d'éliminer les silos de données. Les auteurs expliquent que cette approche réduit les coûts, augmente l'agilité et simplifie l'accès à toutes les données.

(Boumediene, H., 2024) examine avec son article les obstacles et opportunités liés à la modernisation numérique des institutions publiques algériennes, en mettant l'accent sur la gestion des données. L'auteur critique la dépendance excessive aux entrepôts de données physiques et aux solutions ETL, coûteuses et rigides. Il présente la virtualisation des données comme une solution pertinente pour le contexte algérien, notamment grâce à sa flexibilité et son faible coût d'infrastructure. Boumediene évoque aussi le manque d'expertise locale, qui freine l'adoption de solutions comme Denodo, malgré les efforts institutionnels (data center national, cloud souverain). L'article recommande d'intégrer la virtualisation dans les politiques publiques de numérisation, tout en développant les compétences locales en ingénierie des données.

(AtScale, 2025) met en lumière les raisons pour lesquelles les entreprises adoptent massivement la data virtualisation dans les projets analytiques modernes. AtScale y insiste sur la capacité de cette technologie à offrir un accès immédiat aux données sans les déplacer, ce qui réduit les coûts, les délais et les efforts liés à l'intégration. L'article analyse les impacts de

la virtualisation sur les performances analytiques, notamment pour les outils comme Tableau, Power BI ou Excel. Il illustre que la virtualisation, combinée à des modèles sémantiques, améliore la cohérence et la gouvernance des données dans les grandes entreprises. L'article met en garde contre les erreurs de conception (comme la surcharge réseau) et recommande l'utilisation de plateformes spécialisées comme Denodo. Il conclut que la virtualisation est désormais centrale dans l'architecture data agile

1.1.1. Contexte mondial

La virtualisation des données dans les systèmes décisionnels est reconnue comme une avancée majeure qui permet d'accéder, manipuler et analyser des données provenant de multiples sources en temps réel, sans réplication ni déplacement physique des données. Cette technologie assure une meilleure gouvernance, une sécurité renforcée et une réduction des coûts opérationnels, tout en facilitant une prise de décision plus rapide et plus souple à l'échelle globale. (Thangaraj, 2024)

En Europe, le projet Gaia-X, soutenu par plusieurs États membres de l'Union européenne, vise à créer des espaces de données sécurisés et interopérables utilisant largement la virtualisation des données dans divers secteurs publics et privés. (Christophe Auffray, avr 2025)

Aux États-Unis, des entreprises de la finance et de l'assurance investissent fortement dans les data rooms virtuelles pour sécuriser et analyser de gros volumes de données en temps réel, facilitant les opérations de fusion-acquisition et de conformité. (M&A, Dec 27, 2024)

1.1.2. Situation en Afrique

En Afrique, la virtualisation et autres technologies de l'informatique dématérialisée, telles que le Cloud Computing, commencent à émerger comme solutions clés face aux défis d'équipement informatique limité et de connectivité. Plusieurs pays investissent dans la mise en place de centres de données et dans l'adoption progressive du Cloud Computing, bien que des enjeux réglementaires et de sécurisation des données subsistent. Les expériences montrent un fort potentiel d'amélioration des systèmes décisionnels grâce à ces innovations, surtout dans les secteurs bancaires et éducatifs. (ITU, 2012)

Plusieurs pays africains, sous l'impulsion de programmes comme ceux de l'UIT (Union Internationale des Télécommunications) et de partenariats internationaux, adoptent progressivement la virtualisation et le Cloud Computing pour renforcer leurs systèmes

décisionnels, notamment dans les secteurs de la finance, de la santé et de l'éducation. (ITU, 2012)

Des entreprises privées dans des pays comme le Sénégal explorent des systèmes décisionnels intégrant la virtualisation pour mieux gérer les données éducatives et réseaux. (Janaushek, 2024)

1.1.3. Cas de l'Algérie

En Algérie, des initiatives similaires visent à intégrer ces technologies dans les systèmes d'information, notamment pour améliorer l'efficacité des systèmes décisionnels publics et privés. Toutefois, la mise en œuvre doit tenir compte des particularités locales, notamment en matière de gouvernance des données, d'interopérabilité des architectures et de cadre réglementaire. Des travaux universitaires et mémoires ont également étudié ces intégrations innovantes, soulignant les défis spécifiques d'adaptation et recommandant des approches hybrides combinant technologies et stratégies organisationnelles.

En Algérie, des établissements universitaires et certaines administrations publiques expérimentent des solutions hybrides combinant virtualisation et gouvernance des données pour améliorer les performances des systèmes décisionnels.

Certaines initiatives visent à adapter ces technologies étrangères au contexte local, en tenant compte des spécificités techniques et réglementaires nationales, notamment dans les ministères clés comme celui des Finances. (Radio Algérienne, 2023)

Ces exemples montrent que la virtualisation des données est adoptée à différentes échelles et secteurs, avec des adaptations spécifiques à chaque région et contexte organisationnel.

1.2.Cadre conceptuel

1.2.1. Les Systèmes d'Information Décisionnels (SID)

Les Systèmes d'Information Décisionnels (SID) ont transformé complètement la manière dont les entreprises collectent, organisent et exploitent les données pour orienter leurs prises de décisions stratégiques. En se basant sur une variété de technologies et des méthodes, les SID permettent de convertir des données brutes en informations exploitables, facilitant ainsi l'analyse des tendances, la prévision des résultats et l'optimisation des opérations. L'évolution de ces systèmes, marquée par l'émergence d'architectures comme les entrepôts de données et les data lakes, a répondu à un besoin croissant d'intégration et de consolidation des données en

provenance de sources multiples. Toutefois, ces systèmes présentent aussi certains défis, notamment en termes de latence, de coût et de complexité qui peuvent parfois entraver leur efficacité. Dans cette section, nous examinerons la définition des SID, leur architecture classique ainsi que les enjeux qui y sont associés, afin d'établir une base solide pour la compréhension des évolutions technologiques à venir.

1.2.1.1. Définition et évolution des SID

Les Systèmes d'Information Décisionnels (SID) sont des outils essentiels permettant aux organisations de prendre des décisions basées sur des données consolidées. Ils se caractérisent par leur capacité à collecter, intégrer, analyser et présenter des données pour soutenir les processus décisionnels. L'évolution des SID a été marquée par l'émergence et l'adoption d'architectures telles que les entrepôts de données et les data lakes (Inmon, 2005).

Au départ, dans les années 1960, les SID étaient essentiellement basés sur des fichiers maîtres et des rapports statiques. L'apparition des systèmes de gestion de bases de données (SGBD) dans les années 1970 a constitué un tournant majeur, offrant une architecture centralisée et permettant un accès direct aux données. L'entrepôt de données ou bien le data warehouse a été conceptualisé comme une source unique de données intégrée, facilitant l'analyse et la génération de rapports. Un data warehouse (entrepôt de données) s'agit d'une architecture de base de données conçue pour agréger, stocker et gérer des données provenant de sources variées, afin de faciliter les analyses et la prise de décision. Il intègre des données d'application, qui sont structurées et intégrées pour être exploitables dans des contextes analytiques. Les données dans un Data warehouse sont généralement historiques, normalisées, et orientées vers la synthèse pour répondre aux besoins des utilisateurs finaux dans un format pertinent pour l'analyse (Inmon, 2005).

L'avènement des data lakes représente une autre étape significative, permettant de stocker des données non structurées et semi-structurées à grande échelle, ouvrant ainsi la voie à des analyses avancées et à l'intelligence artificielle. D'autre part, un data lake est un référentiel de stockage qui permet de conserver des données dans leur format brut et natif, qu'il s'agisse de données structurées, semi-structurées ou non structurées. Contrairement à un data warehouse, un data lake offre plus de flexibilité quant à la façon dont les données peuvent être stockées et analysées. Bien qu'il ne soit pas défini dans les extraits fournis, le concept de data lake se distingue par sa capacité à gérer de grandes quantités de données diverses sans nécessiter une

structure prédéfinie ou une intégration immédiate, ce qui en fait un choix pertinent pour des analyses ad hoc et des explorations de données (Inmon, 2005).

Sommairement, le data warehouse est destiné à l'analyse de données structurées avec des nécessités d'intégration et de normalisation, tandis que le data lake permet le stockage de données dans leur forme originale, favorisant l'exploration et l'analyse variées.

1.2.1.2. Architecture classique d'un SID

Une architecture classique d'un SID est généralement composée de plusieurs composantes clés, notamment l'Extraction, la Transformation et le Chargement (ETL), le stockage de données, le reporting et l'analytique. Le processus ETL joue un rôle principal en assurant que les données provenant de sources diverses sont extraites, nettoyées et transformées avant d'être chargées dans l'entrepôt de données. Cette étape garantit l'intégrité et la qualité des données analysées. (Inmon, 2005)

Le stockage de données se déroule typiquement dans un entrepôt de données, où les données sont organisées pour optimiser la requête et l'analyse. À partir de cette base de données centralisée, les utilisateurs peuvent générer des rapports traditionnels ou entreprendre des analyses plus poussées à l'aide d'outils d'analyse de données. L'analyse peut inclure à la fois des analyses descriptives et prédictives, permettant aux entreprises de tirer des insights significatifs et d'anticiper les tendances futures. (Inmon, 2005)

Figure 1: Les 4 phases du processus de Business Intelligence, de la donnée à l'information



Source : (Fernandez, 2020).

1.2.1.3. Les Niveaux de l'Architecture SID

Comme nous l'avons déjà mentionné, l'architecture classique d'un SID est un cadre structuré qui intègre divers composants et processus. Cette architecture peut être décomposée en plusieurs niveaux, chacun jouant un rôle spécifique dans la transformation des données brutes en informations significatives. Voici les principaux niveaux qui composent cette architecture (Inmon, 2005):

- Niveau opérationnel

Le niveau opérationnel est la base de l'architecture, composé de systèmes transactionnels qui enregistrent les données en temps réel. Ces systèmes incluent des bases de données opérationnelles et des applications de gestion (telles que les systèmes de gestion des ressources humaines, des ventes et des stocks). Dans ce niveau, les données sont souvent fragmentées et enregistrées sous forme de données brutes, orientées vers l'application, répondant directement aux besoins quotidiens de l'entreprise.

- Entrepôt de données (Data Warehouse)

L'entrepôt de données constitue le cœur du SID, où les données provenant de diverses sources sont intégrées, normalisées et centralisées. C'est un système de stockage permanent qui contient des données historiques et non volatiles, qui sont accessibles pour des analyses à grande échelle. Les caractéristiques de ce niveau de l'entrepôt de données comprennent :

- **Intégration des Données** : Les données sont collectées à partir de sources opérationnelles variées, consolidées et transformées afin d'offrir une vue unifiée.
- **Données Historiques** : Contrairement au niveau opérationnel, l'entrepôt de données stocke des données sur des périodes étendues, ce qui permet des analyses temporelles et des comparaisons historiques.
- **Orienté Vers le Sujet** : Les données sont organisées selon des sujets d'analyse, facilitant les rapports et les requêtes complexes

- Niveau départemental (Data Mart)

Les niveaux départementaux, ou data Marts, représentent des segments de l'entrepôt de données, spécialisés pour répondre aux besoins spécifiques d'un département donné (comme le marketing, la finance, ou les opérations). Ces niveaux permettent une gestion plus granulaire des informations :

- **Vieillessement spécifique** : Contrairement aux entrepôts de données qui intègrent de larges volumes de données, les data marts se concentrent sur des ensembles de données pertinents pour des utilisateurs spécifiques.
- **Accès rapide** : Étant moins volumineux, les data marts rendent l'accès aux données plus rapide et l'analyse plus efficace pour les utilisateurs finaux.
- **Niveau individuel (Desktop)**

Le niveau individuel est l'interface utilisateur où les citoyens des données interagissent avec le SID pour accéder aux informations et pour réaliser des analyses. Ce niveau comprend des outils de visualisation et des fonctionnalités de reporting qui permettent aux utilisateurs finaux de tirer des conclusions à partir des données :

- **Outils BI (Business Intelligence)** : Les solutions BI offrent des fonctionnalités de tableaux de bord, tableaux de bord prospectif, de rapports dynamiques et d'analyses automatisées qui facilitent la prise de décisions.
- **Interface Graphique Utilisateur** : Ce niveau doit être intuitif et convivial pour permettre aux utilisateurs non techniques d'exploiter efficacement les données, sans nécessiter de compétences en développement ou en programmation.

1.2.1.4. Processus de Traitement des Données

Une fois les données collectées, l'architecture du SID doit comprendre les processus nécessaires pour transformer ces données en informations décisionnelles (Inmon, 2005) :

1.2.1.4.1. Extraction, Transformation, Chargement (ETL)

Le processus ETL est indispensable pour le bon fonctionnement de l'architecture SID. Ce processus se déroule en trois étapes :

- **Extraction** : Les données sont extraites des systèmes opérationnels, souvent disparates, et peuvent inclure des formats variés.
- **Transformation** : Les données brut doivent être nettoyées (par exemple, en éliminant les doublons) et transformées (par exemple, en changeant les formats), en conformité avec les besoins d'analyse.
- **Chargement** : Une fois transformées, les données sont chargées dans l'entrepôt de données où elles sont prêtes à être analysées.

1.2.1.4.2. Analyse et Reporting

Les outils d'analyse sont appliqués sur les données chargées dans l'entrepôt de données pour produire des rapports et des aperçus significatifs. Ces rapports peuvent inclure des analyses complexes, des recherches de tendances et des calculs agrégés qui, à leur tour, influencent les décisions stratégiques.

1.2.1.5. Architecture Technique

L'architecture technique d'un SID comprend également des composants matériels et logiciels (Inmon, 2005):

- **Matériel**

Le matériel nécessaire pour un SID peut inclure des serveurs dédiés, des dispositifs de stockage, et des infrastructures réseau qui garantissent la disponibilité et l'accès aux données.

- **Logiciel**

Les systèmes SID reposent sur des logiciels spécifiques pour le traitement des données, y compris des bases de données relationnelles, des systèmes de gestion de base de données (SGBD), et des outils BI adaptés pour l'analyse et le reporting.

On conclue que l'architecture classique d'un Système d'Information Décisionnel est une structure sophistiquée conçue pour gérer le cycle complet de la donnée, depuis la collecte jusqu'à l'analyse. Chaque niveau de cette architecture, du niveau opérationnel à l'interface utilisateur, joue un rôle clé dans le soutien des prises de décisions éclairées au sein de l'entreprise.

1.2.1.6. Enjeux et limites des approches traditionnelles

Les approches traditionnelles de gestion et d'analyse des données présentent plusieurs enjeux et limites, qui soulignent la nécessité d'une réévaluation des systèmes existants. Ces défis affectent la productivité, la qualité des informations, et l'efficacité des processus de prise de décision au sein des organisations (Inmon, 2005).

- **Enjeux**

- **Complexité d'accès aux données :**

L'une des principales préoccupations des systèmes traditionnels est la difficulté d'accès aux données. Alors que les organisations possèdent souvent une multitude de fichiers et

d'applications disparates, localiser les données pertinentes nécessite l'examen minutieux d'une grande variété de fichiers, chacun utilisant potentiellement des méthodes d'accès distinctes, comme le système de gestion d'informations (IMS) ou le système de gestion de bases de données intégrées (IDMS). Cette complexité diminue la productivité des analystes, qui doivent naviguer à travers ces différentes technologies.

- **Long délai de production de rapports :**

Les approches traditionnelles se caractérisent par des délais significatifs dans la création de rapports. La création du premier rapport peut prendre de 9 à 12 mois, alors que les rapports subséquents peuvent prendre jusqu'à 3 à 5 ans. Cette hiérarchie des délais souligne un manque de flexibilité et de réactivité aux standards de reporting des utilisateurs. Les exigences pour les rapports suivants ne sont souvent pas connues lors de la rédaction du premier, ajoutant une couche d'incertitude.

- **Difficulté de transformation des données en informations :**

Un autre enjeu majeur est l'incapacité des systèmes d'information traditionnels à transformer efficacement les données en informations exploitables. Les analystes rencontrent souvent des difficultés en raison de l'existence d'applications non intégrées, ce qui complique le recours aux données nécessaires pour répondre aux exigences d'analyse. Cela conduit à une situation où même les demandes d'information simples peuvent prendre des mois de travail pour être satisfaites.

- **Manque d'intégration des données:**

Les systèmes traditionnels sont souvent marqués par un manque d'intégration, rendant difficile la prise en compte de multiples sources de données. Les données sont souvent fragmentées en différentes applications, rendant leur collecte et leur consolidation laborieuses. En conséquence, la vision d'ensemble des données sur l'entreprise est souvent compromise, ce qui peut entraîner des décisions basées sur des informations incomplètes ou obsolètes.

- **Mauvaise gestion des données historiques :**

Les données historiques sont essentielles pour l'analyse de tendance et la prise de décisions stratégiques. Cependant, les systèmes traditionnels n'ont souvent pas été conçus pour archiver et consolider les données historiques efficacement. Cela entraîne un manque

de contextualisation des nouvelles données, empêchant une analyse historique significative.

➤ **Limites**

- **Rigidité des systèmes:**

Les systèmes traditionnels sont souvent rigides et peu adaptables aux changements rapides des besoins des utilisateurs. Les modifications des exigences du marché ou des besoins utilisateurs peuvent nécessiter des changements longs et coûteux, entravant ainsi l'innovation et l'efficacité.

- **Surcharge de redondance des données :**

Dans de nombreux cas, les différentes applications non intégrées conduisent à une redondance massive des données. Cette situation crée des incohérences, puisque les mêmes données peuvent être interprétées différemment à travers des systèmes variés, rendant l'audit et la traçabilité des données particulièrement difficiles.

- **Coûts cachés:**

La complexité inhérente au maintien des systèmes traditionnels engendre des coûts cachés, non seulement en termes d'infrastructure, mais aussi d'efforts humains. Les équipes doivent consacrer un temps considérable à la gestion quotidienne et à l'intégration des données, ce qui détourne les ressources des initiatives d'innovation.

- **Non-utilisation efficace des données :**

Les données accumulées dans les environnements traditionnels ne sont pas toujours utilisées de manière efficace, souvent en raison d'un manque d'outils adéquats pour leur analyse et leur exploitation. Cette situation fluctue le coût de l'entreposage et la gestion des données, sans bénéfices proportionnels en faveur de la prise de décision.

En somme, les approches traditionnelles de gestion des données, bien que largement utilisées, montrent de sérieux signaux de faiblesse face à un environnement technologique de plus en plus dynamique et exigeant. Les enjeux liés aux délais de production de rapports, à la transformation des données en informations, et à l'intégration efficace des données nécessitent une transition vers des systèmes plus modernes, tels que les entrepôts de données et les data lakes, qui répondent mieux aux besoins contemporains des entreprises.

En conclusion, dans cette section nous avons présenté les bases des Systèmes d'Information Décisionnels, leur évolution vers des solutions plus complexes comme les entrepôts de données, ainsi que les principaux défis qu'ils rencontrent, notamment la lenteur, le coût élevé et la complexité de gestion. Ces difficultés montrent qu'il est nécessaire de chercher de nouvelles approches capables de mieux répondre aux besoins actuels en matière de gestion et d'analyse des données. C'est pourquoi le chapitre suivant sera consacré à la data virtualisation, une technologie innovante qui promet de dépasser les limites des entrepôts de données traditionnels. Elle permet d'intégrer les données de façon plus flexible, en offrant plus d'agilité et une meilleure réactivité face aux besoins analytiques d'aujourd'hui.

1.2.2. La Virtualisation de données

À l'ère du big data, les entreprises sont confrontées au défi de gérer, d'intégrer et d'analyser efficacement des volumes massifs de données diverses provenant de différentes sources. La virtualisation des données apparaît comme une solution transformatrice, révolutionnant la manière dont les données sont accessibles et exploitées pour l'analyse (Aaliyah Jordon, 2023). Ce chapitre explore la manière dont la virtualisation de données change la donnée dans le domaine de l'analyse des big data. En commençant par répondre à la question suivante : C'est quoi la virtualisation des données ? Ensuite, en examinant les concepts relatifs à la virtualisation des données, en définissant le concept de la data virtualisation. Ensuite, en la comparant à des techniques traditionnelles comme l'ETL et la fédération des données, et en examinant ses avantages et défis. Et enfin, nous présentant les différentes alternatives de l'implémentation d'une couche de virtualisation de données. Au final, ce chapitre révèle à quel point cette technologie peut révolutionner le domaine des systèmes d'information décisionnels (SID).

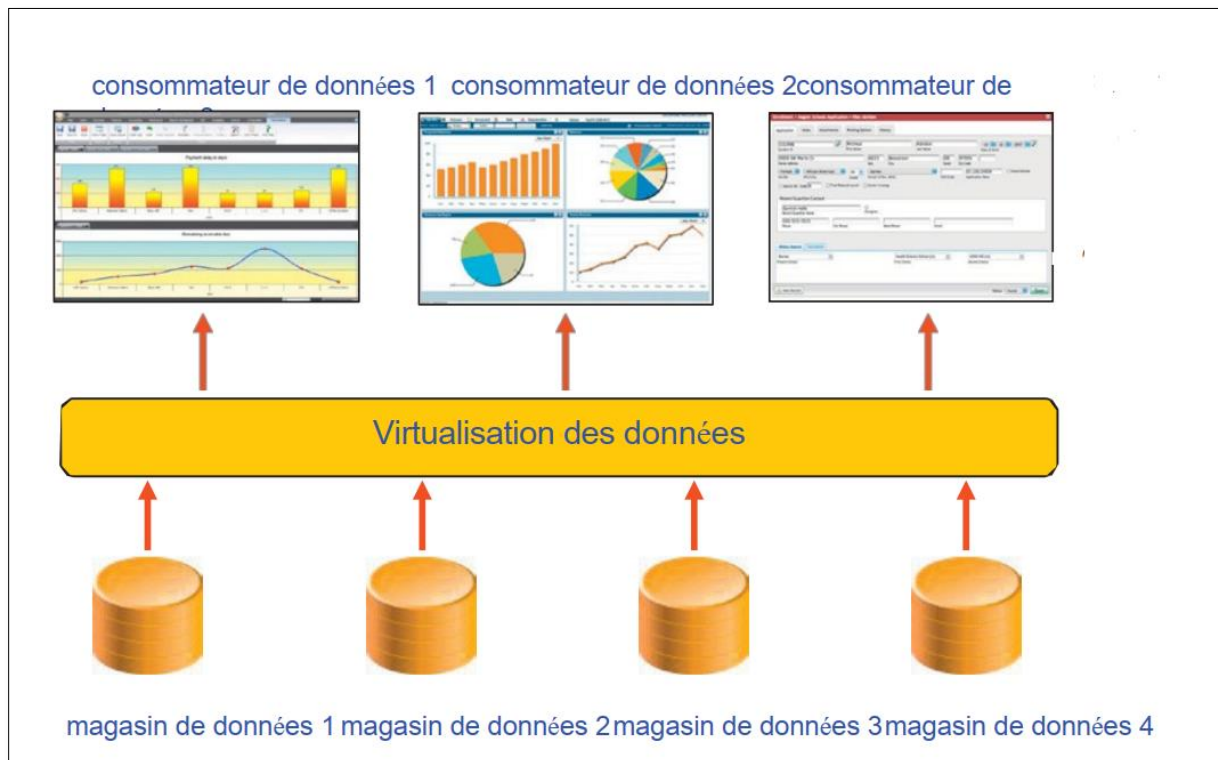
1.2.2.1. Définition et principes fondamentaux

1.2.2.1.1. C'est quoi la virtualisation des données?

Tout au long de ce mémoire, nous allons utiliser les termes "**consommateur de données**" et "**magasin de données**". Le terme "consommateur de données" désigne toute application qui récupère, saisit ou manipule des données. Par exemple, un consommateur de données peut être une application de saisie de données en ligne, une application de reporting, un modèle statistique, une application Internet. De même, le terme "magasin de données" est utilisé pour désigner toute source de données : une table dans une base de données SQL, un simple fichier texte, un document XML, une feuille de calcul, un service web, un fichier séquentiel, une

page HTML, etc. Dans certains cas, un magasin de données n'est qu'un fichier passif et dans d'autres, il s'agit d'une source de données comprenant le logiciel permettant d'accéder aux données, tel qu'un serveur de base de données et un service web (Lans, 2012).

Figure 2 Lorsque la virtualisation des données est appliquée, toutes les sources de données sont présentées comme une source de données intégrée



.source : (Aaliyah Jordon, 2023)

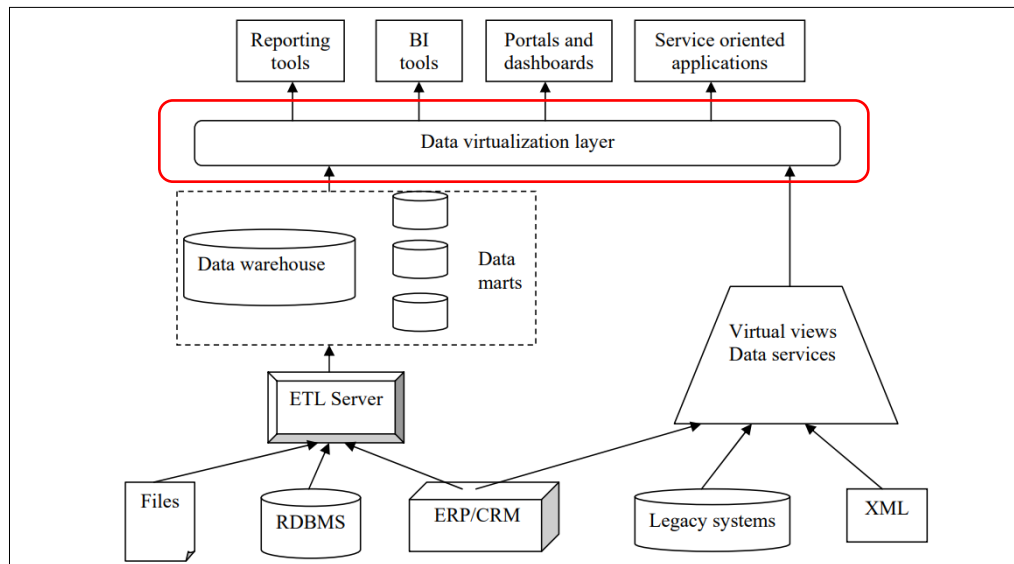
Lorsqu'une solution de virtualisation des données est utilisée, elle réside entre les consommateurs de données et les magasins de données. Les consommateurs de données accèdent aux données par l'intermédiaire de la couche de virtualisation des données et cette dernière cache les magasins de données. (Aaliyah Jordon, 2023)

1.2.2.1.2. Définition de la data virtualisation

La virtualisation des données est une technologie qui offre aux consommateurs de données une vue unifiée, abstraite et encapsulée pour interroger et manipuler des données stockées dans un ensemble hétérogène de magasins de données (Lans, 2012). Autrement dit, Elle sert de couche intermédiaire qui fournit une vue unifiée et virtuelle des données, indépendamment de leur emplacement ou de leur format. Cette technologie a le potentiel de révolutionner

l'analyse des big data en simplifiant l'intégration des données, en améliorant l'agilité des données et en facilitant l'obtention d'informations plus rapides (Aaliyah Jordon, 2023).

Figure 3 La virtualisation des données



source : (Bologa, 2011).

1.2.2.2. Principales caractéristiques :

- Vue Unifiée : Les utilisateurs ont une perspective consolidée des données, quelles que soient leurs sources physiques.
- Accès en Temps Réel : Les données sont accessibles et intégrées à la demande, sans délai de traitement, permettant une actualisation en temps réel.
- Masquage de la Complexité : Les utilisateurs n'ont pas besoin de connaître les détails techniques des sources de données ; les aspects techniques tels que les types de bases de données, les formats de fichiers, etc., sont abstraits.
- Transformations à la Volée : La capacité de nettoyer et de transformer les données pendant l'interrogation, sans prétraitement, est une caractéristique clé.
- Flexibilité et Agilité : Les changements dans les sources de données ou l'ajout de nouvelles sources sont gérés de manière agile sans impacts significatifs sur les utilisateurs finaux.

1.2.2.3. Les concepts relatifs de la data virtualisation

La virtualisation des données représente une approche novatrice pour accéder et manipuler des données de manière efficace en masquant la complexité des sources de données sous-

jacentes. Elle est souvent mise en relation avec plusieurs concepts clés de l'informatique, qui enrichissent sa compréhension et son application. Voici une définition de chaque concept, ainsi que son rapport avec la virtualisation des données (Lans, 2012) :

- **Encapsulation :**

L'encapsulation est un principe fondamental de la programmation qui consiste à cacher les détails internes d'une structure de données ou d'un système. Dans le contexte de la virtualisation des données, l'encapsulation permet aux consommateurs de données d'accéder à des informations sans avoir à se préoccuper des méthodes de stockage ou des interfaces techniques sous-jacentes. Par conséquent, la virtualisation des données offre une encapsulation qui simplifie le développement d'applications en masquant les complexités des différentes sources de données.

- **Masquage des informations :**

Le masquage des informations est le processus de dissimulation des détails qui ne devraient pas influencer d'autres parties d'un système. Cela favorise l'indépendance de l'application par rapport à la structure des données. La virtualisation des données applique ce principe en cachant les détails techniques des diverses sources de données, permettant ainsi aux développeurs de se concentrer sur la logique de l'application sans être affectés par les modifications des structures sous-jacentes.

- **Abstraction :**

L'abstraction est un processus qui permet de se concentrer sur les aspects essentiels d'un phénomène tout en ignorant ses détails superflus. Dans le cadre de la virtualisation des données, l'abstraction signifie que les consommateurs de données interagissent avec des vues simplifiées des données, ne présentant que les tables, colonnes et lignes pertinentes pour leurs besoins. Cela facilite l'accessibilité et améliore l'expérience utilisateur, tout en garantissant que des détails non pertinents sont masqués.

- **Intégration des données :**

L'intégration des données concerne la combinaison de données provenant de diverses sources pour fournir une vue unifiée. La virtualisation des données facilite cette intégration en permettant aux applications d'accéder à des données dispersées à partir de différentes plateformes sans avoir à les copier ou à les déplacer. Cela favorise une exploitation plus efficace des données en temps réel et réduit le besoin de stockage supplémentaire.

- **Intégration des informations d'entreprise (EII) :**

Ce terme fait référence à l'intégration des données critiques pour la prise de décision à l'échelle de l'entreprise. La virtualisation des données est un outil clé dans ce domaine, car elle permet de centraliser l'accès à des données provenant de divers systèmes, tout en conservant leur origine. Cela permet aux entreprises d'obtenir des informations cohérentes et pertinentes sans les inconvénients des systèmes d'intégration traditionnels, qui peuvent être coûteux et complexes à gérer.

En somme, la virtualisation des données s'appuie sur ces concepts pour offrir une approche agile, intégrée et simplifiée de l'accès aux données, facilitant ainsi le développement et la gestion des applications tout en améliorant la prise de décision au sein des organisations.

1.2.2.4. Avantages

La virtualisation des données présente de nombreux atouts notables, à savoir (Lans, 2012)

➤ **Systèmes plus agiles :**

Elle permet une modification aisée des systèmes, ce qui favorise un développement rapide de nouveaux rapports et l'ajustement de ceux déjà en place.

➤ **Simplification et diminution des dépenses :**

Ceci facilite les structures des systèmes d'information décisionnel, ce qui aide à diminuer les coûts.

➤ **Intégration fluide des données :**

Elle offre un accès unifié aux informations, indépendamment de leur localisation, format ou méthode d'accès.

➤ **Facilitation de l'intégration de nouvelles technologies :**

Elle rend plus aisé le renouvellement des technologies actuelles de stockage de données par des alternatives innovantes, sans affecter les rapports.

➤ **Augmentation de la productivité des développeurs :**

En masquant les complexités techniques d'accès aux données, la virtualisation permet aux développeurs de se concentrer sur la définition des mappings pour les tables virtuelles.

En plus des avantages bien connus de la data virtualisation, elle présente des bénéfices particulièrement stratégiques dans un contexte de transformation numérique rapide. En effet, dans un environnement en phase de transformation numérique où les compétences techniques nécessaires pour gérer une architecture classique d'intégration de données sont insuffisantes et manque d'expertise, la data virtualisation offre une solution agile qui réduit considérablement la complexité technique et les ressources de développement requises. Ce modèle permet ainsi de livrer des données intégrées et gouvernées beaucoup plus rapidement, ce qui est important lorsque le temps est un facteur clé et que la transformation doit être achevée dans les meilleurs délais. Par ailleurs, en évitant les lourdeurs liées aux répliquions et aux infrastructures traditionnelles, elle facilite l'adaptation aux évolutions rapides des besoins métiers, tout en permettant aux équipes de se concentrer sur la valeur métier plutôt que sur la gestion technique des données.

1.2.2.5. Limites et défis

Malgré ses avantages, la virtualisation des données présente également des limites et des défis. La sécurité des données est une préoccupation majeure, car l'accès en temps réel aux données provenant de multiples sources peut augmenter le risque d'exposition. De plus, la performance peut être affectée si les systèmes sous-jacents sont mal optimisés ou si la latence des réseaux d'accès aux données devient un facteur limitant. Enfin, la gouvernance des données reste un défi, car assurer la qualité et la conformité des données à travers plusieurs systèmes disparates nécessite une gestion rigoureuse.

Les restrictions et les enjeux liés à la virtualisation des données sont variés et concernent différents aspects techniques, de performance, de sécurité et de gouvernance, voici les principaux éléments (Lans, 2012), (Serano, 2018), (IBM Software Hub, 2024) (GDH, 2024):

- Gestion complexe et intégration :

La virtualisation des données peut rendre l'intégration avec d'autres systèmes et technologies de données plus difficile, comme les entrepôts de données, les data lakes, les pipelines de données, les catalogues de données et les outils de visualisation. Elle peut aussi engendrer des problèmes de compatibilité et d'interopérabilité en raison des divers formats de données, normes, protocoles et API utilisés. Il est donc crucial d'assurer une couche de virtualisation des données performante, en adoptant des techniques telles que la fédération de données, l'orchestration des données, l'ingestion de données, l'extraction de données et les API, afin de faciliter l'intégration et l'interopérabilité de la virtualisation des données.

-Gouvernance et gestion des métadonnées :

La virtualisation des données peut compliquer la gouvernance des données et la gestion des métadonnées, en ajoutant une autre couche d'abstraction et de transformation des données qui doit être documentée, surveillée et contrôlée. Elle peut également accroître la complexité et la diversité des données, en créant davantage de sources de données, de vues et de consommateurs qui doivent être identifiés, classés et catalogués.

-Sécurité et vulnérabilités :

L'hyperviseur, élément central de la virtualisation, est une cible privilégiée des attaques. Des attaques telles que l'hyper-jacking, où un hyperviseur malveillant prend le contrôle total d'un serveur, ou le guest-hopping, où des vulnérabilités permettent à des logiciels malveillants de se propager entre les machines virtuelles, représentent des menaces sérieuses. Des mesures proactives, comme le durcissement de l'environnement et l'isolement de chaque machine virtuelle, sont essentielles pour contrer ces attaques.

Disponibilité et tolérance aux pannes :

La recherche de la disponibilité dans la virtualisation est complexe et peut être perturbée par de nombreux événements, notamment sur le système hôte lorsqu'il est défaillant. Un système virtualisé interconnecté avec un grand nombre de réseaux hétérogènes peut augmenter les possibilités d'accès malveillants, rendant la prévention des intrusions plus difficile et fragilisant la tolérance aux pannes.

Performances ET surcharge :

Une surcharge de performance due à la couche hyperviseur peut être entraînée lors de l'exécution de plusieurs machines virtuelles sur un seul serveur physique, ce qui peut avoir une mauvaise influence sur la performance globale des applications exécutées sur les machines virtuelles.

Coûts et licences :

Les couts sont élevés et les accords de licence pour les logiciels dans les environnements virtualisés peuvent être complexes. Les organisations doivent veiller à respecter les accords de licence tout en optimisant l'utilisation des ressources

1.2.2.6. Différence entre data virtualisation, ETL et data fédération

Dans le domaine de la gestion des données, il est essentiel de comprendre les différences entre plusieurs concepts clés, notamment la virtualisation des données, l'ETL (Extract, Transform, Load) et la fédération de données. Chacune de ces approches a ses propres caractéristiques, avantages et applications.

Comme dans les sections précédentes, nous avons déjà définie et expliqué en détail ses caractéristiques clés. En tant qu'approche innovante, la virtualisation des données permet d'accéder et de manipuler les informations provenant de diverses sources tout en masquant leur complexité technique. Les avantages, tels que la vue unifiée, l'accès en temps réel et la flexibilité, en font un élément essentiel dans la gestion moderne des données. Cette compréhension préalable nous permettra de mieux aborder les distinctions entre la virtualisation des données, l'ETL et la fédération de données, que nous examinerons par la suite.

➤ ETL (Extract, Transform, Load) (Caserta, 2004)

-Définition :

L'ETL est un processus d'intégration de données en trois étapes. Il extrait les données de diverses sources, les transforme pour les adapter à des besoins spécifiques, puis les charge dans un entrepôt de données ou un autre système cible pour une analyse ultérieure.

-Principales caractéristiques :

Processus par Lots : L'intégration des données se fait généralement sur une base horaire, quotidienne, ou selon un autre calendrier fixe, rendant le processus moins réactif que la virtualisation.

Transformation Préalable : Les données doivent être transformées avant leur chargement, ce qui peut inclure le nettoyage, la normalisation et l'enrichissement, permettant une meilleure qualité des données dans l'entrepôt cible.

Stockage Physique : Les données intégrées sont stockées dans un entrepôt de données, ce qui peut nécessiter une infrastructure et des coûts de stockage additionnels.

Contrôle des Données : Grâce à un processus contrôlé, l'ETL garantit que seules les données de haute qualité sont chargées et disponibles pour l'analyse.

Utilisation Fréquente dans les Entrepôts de Données : L'approche ETL est couramment utilisée pour alimenter les entrepôts de données et pour les besoins de reporting.

➤ **Data Fédération (Inmon, 2005)**

-Définition :

La fédération de données désigne l'intégration de plusieurs sources de données autonomes pour présenter une vue unifiée, permettant aux utilisateurs d'accéder à des données disparates comme si elles provenaient d'une seule source.

-Principales caractéristiques :

Sources Autonomes : Chaque magasin de données d'origine peut fonctionner indépendamment en dehors du système de fédération.

Intégration à la Demande : Les données sont intégrées et accessibles sur demande, sans nécessiter de processus de chargement séparés comme dans l'ETL.

Intégration Hétérogène : La fédération permet de rassembler des données de différentes sources, même si elles utilisent des structures, des langages d'accès ou des API variés.

Vue Intégrée : Les utilisateurs obtiennent une présentation intégrée des données, masquant la distribution physique entre différentes sources de données.

Utilisation de l'Intégration en Temps Réel : Contrairement à l'ETL, qui opère sur des lots, la fédération de données permet une consultation en temps réel des données réparties.

La distinction entre ces trois technologies peut être synthétisée dans le tableau ci-dessous, selon les articles (Lans, 2012), (AtScale, 2025) , (What is a data federation?, n.d.) , (tenorsolutions, 2025) , (Work, 2023) et (webservices by ebs, s.d.) :

Tableau 1 comparatif entre les notions suivantes : data virtualisation, ETL et data fédération

Caractéristique	Data Virtualisation	ETL (Extract, Transform, Load)	Data Federation
Objectif Principal	Fournir une vue unifiée et abstraite des données en temps réel, sans déplacer les	Déplacer, transformer et charger les données dans un système cible (souvent un data	Fournir un accès unifié aux données provenant de sources multiples, en laissant les données à

	données sources.	warehouse).	leur emplacement d'origine.
Processus	Création de vues virtuelles des données, requêtes exécutées à la demande sur les sources.	Extraction des données, transformation selon les besoins, chargement dans le système cible.	Fédération des sources de données, les requêtes sont distribuées et les résultats combinés.
Transformation	Transformation à la demande lors de l'accès aux données.	Transformation avant le chargement dans le système cible.	Transformation à la demande lors de l'accès aux données.
Mouvement des Données	Aucun déplacement physique des données. Les données restent à leur source.	Déplacement physique des données vers le système cible.	Aucun déplacement physique des données. Les données restent à leur source.
Temps Réel / Batch	Principalement temps réel ou quasi-temps réel.	Principalement batch (traitement par lots).	Principalement temps réel ou quasi-temps réel.
Stockage	Les données restent dans leurs systèmes sources .	Nécessite un système de stockage cible (data warehouse, data lake, etc.).	Les données restent dans leurs systèmes sources .
Complexité des Données	Gère une grande variété de sources de données (structurées, non structurées, cloud, on-premise).	Souvent optimisé pour les données structurées .	Peut être limité à certains types de sources de données (traditionnellement relationnelles).
Cas d'Usage Typiques	BI en temps réel, reporting agile, intégration d'applications, accès aux données distribuées.	Construction de data warehouses, reporting historique, migration de données, qualité des données.	Reporting distribué, accès à des données hétérogènes sans réplication, requêtes fédérées.
Couplage aux Sources	Faible. Les applications accèdent à des vues virtuelles, indépendantes des	Fort. Les changements dans les sources peuvent nécessiter des modifications importantes	Faible. Les applications accèdent à une vue fédérée, potentiellement abstraite des détails de

	changements sous-jacents.	des processus ETL.	chaque source.
Maturité de la Technologie	Relativement plus récente et en évolution.	Technologie bien établie et mature.	Technologie établie , souvent considérée comme un sous-ensemble de la virtualisation des données.

Source : élaboré par nos soins.

1.2.2.7. Différentes implémentations de la virtualisation des données

Techniquement, il existe plusieurs alternatives pour la mise en œuvre d'une couche de virtualisation des données. Voici quelques exemples (Lans, 2012) :

- Avec un serveur de virtualisation des données dédié, plusieurs magasins de données semblent n'en former qu'un seul. Les applications voient un seul grand magasin de données, alors qu'en réalité les données sont stockées dans plusieurs magasins de données.
- Un bus de service d'entreprise (ESB) peut être utilisé pour développer une couche de services permettant un accès standardisé aux données. Les consommateurs de données qui invoquent ces services n'ont pas besoin de comprendre où et comment les données sont stockées, leur structure de stockage, leur interface avec la source d'origine ou d'autres détails techniques. Ils ne verront, par exemple, qu'une interface SOAP ou ReST (Representational State Transfer). Dans ce cas, l'ESB est la couche de virtualisation des données.
- Le fait de placer des magasins de données dans le Cloud est également une forme de virtualisation des données. Pour accéder à un magasin de données, les consommateurs de données travaillent avec l'API du Cloud et n'ont aucune idée de l'endroit où résident les données elles-mêmes. Le fait que les données soient stockées et gérées localement ou à distance est totalement transparent.
- D'une autre manière, la création d'une base de données virtuelle en mémoire avec des données chargées à partir de plusieurs bases de données physiques peut également être considérée comme une virtualisation des données. La structure de stockage, l'API et l'emplacement des données réelles sont transparents pour l'application qui accède à la

base de données en mémoire. Cette solution est parfois appelée « in-memory analytics».

- Les mappers objet-relationnel (ORM) sont des outils qui convertissent les structures de données des magasins de données aux concepts utilisés dans un modèle de programmation orienté objet, tel que Java et C#. Par exemple, un ORM peut être en mesure de convertir les structures de table plate d'une base de données SQL en concepts orientés objet utilisés en Java. Ainsi, les programmeurs Java n'ont pas à comprendre et à traiter les caractéristiques des concepts SQL, mais uniquement les concepts Java. Exemples d'ORM sont Hibernate et NHibernate.
- Aussi Les organisations peuvent également développer leur propre couche de virtualisation des données qui cache l'endroit et la manière dont les données sont stockées.

Il existe de nombreuses autres façons de développer une couche de virtualisation des données, mais comme notre étude est pointue sur le système d'information décisionnel, l'accent est mis sur la première alternative : les serveurs de virtualisation des données. Un serveur de virtualisation des données est un produit dédié conçu pour prendre en charge la virtualisation des données, ce qui signifie qu'il peut présenter aux consommateurs de données plusieurs magasins de données hétérogènes sous la forme d'un magasin de données logique. L'accès à un serveur de virtualisation des données est très similaire à la connexion à un serveur de base de données. Sans que les consommateurs de données le sachent, les données provenant de différents magasins de données sont réunies (même à partir de magasins de données utilisant des modèles de stockage différents), et les données sont transformées, nettoyées, agrégées, etc.

En résumé, la virtualisation des données se présente comme une avancée prometteuse dans la gestion des informations de manière agile et efficace. Cependant, des défis subsistent en matière de sécurité, de performance et de gouvernance. La prochaine section abordera l'intégration de la virtualisation des données dans les systèmes d'information décisionnels (SID), explorant comment cette technologie peut être mise en œuvre pour répondre aux besoins stratégiques des entreprises dans un environnement de plus en plus dynamique.

1.2.3. Intégration de la Data Virtualisation dans les Système d'Information Décisionnel

Au cours des sections précédentes, nous avons exploré deux notions fondamentales pour la compréhension de l'évolution des systèmes d'information modernes : la data virtualisation et les systèmes d'information décisionnels. Ces concepts jouent un rôle essentiel dans la manière dont les organisations collectent, intègrent et analysent les données pour soutenir leurs processus de décision. Dans ce chapitre, nous allons examiner comment la data virtualisation s'intègre dans les systèmes d'information décisionnels, en abordant son rôle dans l'architecture décisionnelle moderne, son impact sur la prise de décision stratégique et en présentant les différentes stratégies de déploiement de la data virtualisation dans les SID.

1.2.3.1. Rôle de la data virtualisation dans l'architecture décisionnelle moderne

La data virtualisation permet de créer une architecture d'information plus souple et réactive. Elle réduit la complexité en intégrant des sources de données hétérogènes sans nécessiter de stockage physique des données. Dans cette architecture, la data virtualisation représente une couche qui interagit avec différents systèmes de stockage, en offrant aux utilisateurs une vue unifiée des données. Cela facilite l'accès à des données en temps réel, réduisant ainsi le délai nécessaire pour obtenir des indications stratégiques (Lans, 2012).

L'un des principaux avantages de cette approche est qu'elle favorise l'agilité. Les entreprises peuvent rapidement adapter leur architecture d'information pour réagir aux changements du marché ou aux nouvelles exigences réglementaires. Avec des rapports et analyses pouvant être réalisés en temps réel, la prise de décision devient plus efficace (Lans, 2012).

1.2.3.2. Impact de l'intégration d'un serveur de virtualisation de données sur la prise de décision stratégique

D'après (Lans, 2012), le déploiement d'un serveur de virtualisation de données dans les systèmes d'information décisionnels a un impact significatif sur la prise de décision stratégique. Avec un accès rapide et intégré aux données actualisées, les décideurs peuvent baser leurs choix sur des informations pertinentes et précises, réduisant ainsi les risques associés à ces décisions.

L'introduction d'un serveur de virtualisation des données a un impact sur divers aspects d'un système d'information décisionnel. Premièrement, les développeurs de rapports ne sauront pas à quel magasin de données on accède. Ils ne voient qu'une seule grande base de données dans laquelle toutes les données sont intégrées. Ils développent leurs rapports à partir des

tables virtuelles, mais ils n'ont pas besoin de savoir à quoi ressemblent les correspondances, ni où et comment les données sont stockées ?

Deuxièmement, lorsqu'un serveur de virtualisation des données est déployé, les spécifications relatives aux données qui figureraient normalement dans les rapports figurent désormais dans les définitions et les mappages des tables virtuelles et sont gérées par ce serveur de virtualisation des données. Dans laquelle le concept de spécifications de métadonnées partagées est introduit, en laissant ces spécifications être gérées par un serveur de virtualisation des données, tous les outils de création de rapports (même s'ils proviennent de plusieurs fournisseurs) qui accèdent aux données via un serveur de virtualisation des données les utilisent. Tous les rapports partagent donc le même ensemble de spécifications.

Troisièmement, dans un système d'information décisionnel classique, il faut développer une logique de transformation et d'intégration pour obtenir les données sous une forme adaptée aux outils de reporting. En principe, la partie importante de cette logique se retrouve dans des scripts ETL. Lorsqu'un serveur de virtualisation des données est utilisé, cette logique est mise en œuvre dans les mappings. Néanmoins, l'ETL ne disparaîtra pas complètement. Par exemple, les données doivent toujours être copiées périodiquement des bases de données de production vers la zone de transit, et éventuellement de là vers l'entrepôt de données.

Ce qui ne changera pas, c'est ce que les utilisateurs voient et expérimentent. Dans cette nouvelle architecture, les utilisateurs ne voient pas que l'accès aux données est soudainement géré par un serveur de virtualisation des données. Ils voient leurs rapports habituels et ne voient pas si le magasin de données auquel ils accèdent est un data Mart chargé de données à l'aide de scripts ETL ou s'il s'agit de données accédées par un serveur de virtualisation des données. Il faut bien comprendre qu'il n'est pas important pour les utilisateurs de savoir à partir de quel magasin de données les données sont extraites. Tant que les données répondent à toutes leurs exigences, ils sont satisfaits.

1.2.3.3. Stratégies de mise en place de la virtualisation des données

Dans cette section, nous allons présenter trois stratégies distinctes (Lans, 2012) pour la mise en place d'une couche de virtualisation des données dans un système d'information décisionnel. La première stratégie, que nous examinerons en détail, consiste à étendre un système décisionnel existant en intégrant une couche de virtualisation des données. Cette approche permet de tirer parti des infrastructures existantes tout en améliorant l'agilité et l'efficacité des processus décisionnels. Les deux autres stratégies, qui concernent le

développement d'un nouveau système décisionnel, seront simplement mentionnées pour donner un aperçu complet. Ces stratégies alternatives se focalisent sur la création de solutions dédiées dès le départ, maximisant ainsi les potentialités de la virtualisation des données dès l'origine du projet.

➤ **Stratégie 1 : Introduire la virtualisation des données dans un système d'information décisionnel existant.**

La première stratégie pour mettre en place la couche de virtualisation des données dans un système d'information décisionnel existant est divisée en plusieurs étapes, comme suit :

Étape 1: Installation du serveur de virtualisation des données

Dans cette première étape, un serveur de virtualisation des données est installé. Cela implique la configuration matérielle et logicielle nécessaire pour héberger le serveur.

Les tables sources sont importées depuis les différents magasins de données (entrepôt de données, data Marts, etc.). Aussi, les Wrappers et tables virtuelles sont définis pour chacune des tables sources pertinentes. Ces tables virtuelles permettent d'accéder aux données de manière abstraite sans avoir à toucher directement aux données sous-jacentes.

Étape 2: Accès initial aux données

Pendant cette phase, les rapports existants continuent à accéder directement aux magasins de données. Les tables virtuelles développées correspondent généralement aux tables sources, permettant une transition en douceur. Aucun mappage avec des transformations complexes n'est requis à ce stade, ce qui permet de garder la méthode simple et sans perturbations pour les utilisateurs finaux.

Étape 3: Évolution des tables virtuelles

Une fois que le serveur de virtualisation est en place et que les premières tables virtuelles sont opérationnelles, il est temps de les modifier pour accéder à d'autres tables sources.

Cette étape permet d'enrichir l'environnement de données en intégrant plus de sources, ce qui augmente la richesse des informations disponibles.

Étape 4: Nettoyage des données

À ce stade, il est important de commencer à améliorer la qualité des données disponibles. Cela inclut la suppression des données en double et la correction des anomalies.

Les transformations nécessaires peuvent être intégrées dans les tables virtuelles pour garantir que les données présentées aux utilisateurs sont précises et exploitables.

Étape 5: Création de datamarts virtuels

Les tables virtuelles peuvent finalement être structurées pour répondre aux besoins spécifiques des consommateurs de données. Ce processus peut inclure la définition de datamarts virtuels.

Ces datamarts virtuels sont établis en tenant compte des usages envisagés par les utilisateurs finaux, tout en permettant une flexibilité dans l'accès aux différentes sources de données.

Étape 6: Développement d'un prototype

Pour faciliter le processus de validation avec les utilisateurs, un prototype peut être développé. Ce prototype, qui ne nécessite pas de scripts ETL ni de magasins de données supplémentaires, permet de visualiser rapidement les résultats.

Les tables virtuelles et leurs correspondances peuvent être ajustées sur place en fonction des commentaires reçus, ce qui réduit le temps de développement global et améliore l'adaptabilité.

Cette première stratégie permet de mettre en place d'un serveur de virtualisation des données dans un système d'information décisionnel existant de manière évolutive et contrôlée, minimisant les perturbations tout en augmentant la flexibilité et la réactivité du système global face aux besoins des utilisateurs. La virtualisation des données est ainsi adoptée progressivement, permettant une intégration harmonieuse avec les systèmes existants sans nécessiter de restructuration majeure.

➤ Stratégie 2: Créer un nouveau système d'information décisionnel avec la virtualisation des données.

Cette stratégie consiste à développer un nouveau système d'information décisionnel en intégrant dès le départ la virtualisation des données. Plutôt que de partir d'une base de données complexe, on commence par établir une architecture simple, incluant une zone de stockage et un entrepôt de données.

Dans un premier temps, on charge les données directement à partir des systèmes de production, sans transformation. L'idée ici est de conserver les données dans leur état brut afin que les utilisateurs puissent y accéder facilement. Ensuite, on utilise un serveur de virtualisation pour créer des tables virtuelles, ce qui simplifie l'accès aux données tout en permettant des analyses ultérieures. Cela permet aux utilisateurs de bénéficier d'une grande flexibilité tout en ayant accès à des données non filtrées.

➤ **Stratégie 3: Développer un système d'information décisionnel intégrant données brutes et propres.**

Dans cette stratégie, on se concentre sur la création d'un système qui permet aux utilisateurs d'accéder à la fois aux données originales et aux données nettoyées. On commence par établir une structure de base simple comprenant une zone de stockage et un entrepôt. Les données sont d'abord chargées sous leur forme brute, ce qui est précieux pour ceux qui souhaitent avoir une vue complète et authentique des informations.

Ensuite, on met en place un serveur de virtualisation qui crée des représentations virtuelles de ces données. Cela permet de transformer et de nettoyer les données pour répondre aux besoins d'analyse, tout en offrant un accès facile aux versions originales. Avec cette approche, le système devient réactif aux retours des utilisateurs, leur permettant de naviguer facilement entre les données historiques et les analyses modernes.

Ces méthodes sont conçues pour tirer parti de la flexibilité de la virtualisation, tout en s'assurant que les utilisateurs obtiennent exactement ce qu'ils recherchent.

Cette section a exploré l'intégration de la virtualisation des données dans les systèmes d'information décisionnels, en soulignant son rôle essentiel dans l'architecture décisionnelle moderne. La virtualisation des données s'avère être un levier stratégique puissant, permettant d'accélérer l'accès aux informations et d'améliorer la qualité des données disponibles pour la prise de décision.

Nous avons examiné comment l'intégration d'un serveur de virtualisation de données transforme fondamentalement les processus de décision stratégique. En offrant une vue unifiée et actualisée des données provenant de sources disparates, cette technologie facilite des analyses plus rapides et pertinentes, permettant aux organisations de naviguer efficacement dans un environnement commercial complexe.

Par ailleurs, nous avons présenté les stratégies de mise en place de la virtualisation des données, allant de l'introduction dans des systèmes d'information décisionnels existants à la création de nouvelles solutions, démontrant la flexibilité qu'offre cette approche. Que ce soit par l'évolution progressive d'un système déjà en place ou par le développement d'infrastructures entièrement nouvelles, les organisations peuvent adapter la virtualisation des données à leurs besoins spécifiques. De plus, la possibilité de combiner des données brutes

avec des versions transformées répond aux exigences variées des utilisateurs, permettant une analyse approfondie tout en préservant l'intégrité des données originales.

Dans le prochain chapitre, nous entamerons la partie empirique de notre étude, où nous explorerons le système d'information décisionnel au niveau du ministère des finances.

Chapitre 2: Cadre méthodologique et contexte organisationnel

Chapitre 2 : Cadre méthodologique et contexte organisationnel

Le Ministère des Finances (MDF) est le principal acteur dans le développement économique. Il joue un rôle central au sein du gouvernement dans les domaines budgétaire, fiscal, financier et comptable, il est ainsi le conseiller du gouvernement en matière financière, dans sa quête continue d'excellence et d'efficacité, il a élaboré un **Schéma Directeur Stratégique des Systèmes d'Information (SDSSI)** pour la période **2024-2028**. Ce schéma directeur, piloté par la **Direction Générale de la Numérisation, de la Digitalisation et des Systèmes d'Information Economiques (DGNDISIE)**, constitue une feuille de route pour la transformation numérique du ministère, en accord avec les orientations stratégiques nationales.

L'un des axes majeurs de ce schéma directeur est la consolidation d'un système d'aide à la décision performant, visant à soutenir la mise en place d'une gouvernance des finances publiques axée sur la performance et la programmation budgétaire pluriannuelle. Pour atteindre cet objectif, il est crucial d'optimiser l'accès, l'intégration et l'analyse des données financières, souvent dispersées au sein des différentes entités et systèmes du ministère. La data virtualisation émerge comme une approche technologique prometteuse pour répondre à ce besoin.

Le présent projet de mémoire s'inscrit dans cette vision stratégique du Ministère des Finances. Il se propose d'étudier l'application de la data virtualisation pour améliorer le système d'information décisionnel du ministère. Afin de mener à bien cette étude, nous suivrons une démarche méthodologique rigoureuse dans cette partie, qui comprendra : d'abord, une présentation détaillée du contexte organisationnel du Ministère des Finances, permettant de comprendre son fonctionnement et ses enjeux. Ensuite, une analyse approfondie du système décisionnel existant, qui mettra en lumière ses forces, ses faiblesses et les opportunités d'amélioration. Puis, la conception et la modélisation d'une solution de data virtualisation adaptée aux besoins spécifiques du ministère. Cette solution sera ensuite soumise à une phase d'implémentation et de validation, afin d'évaluer sa pertinence et son efficacité. Enfin, nous examinerons les modalités de mise en œuvre et de généralisation de la solution proposée, et nous conclurons par la présentation des résultats escomptés en termes d'amélioration de la prise de décision et d'efficacité globale du Ministère des Finances.

2.1. Contexte organisationnel

2.1.1. Présentation de l'organisme d'accueil (Ministère des Finances)

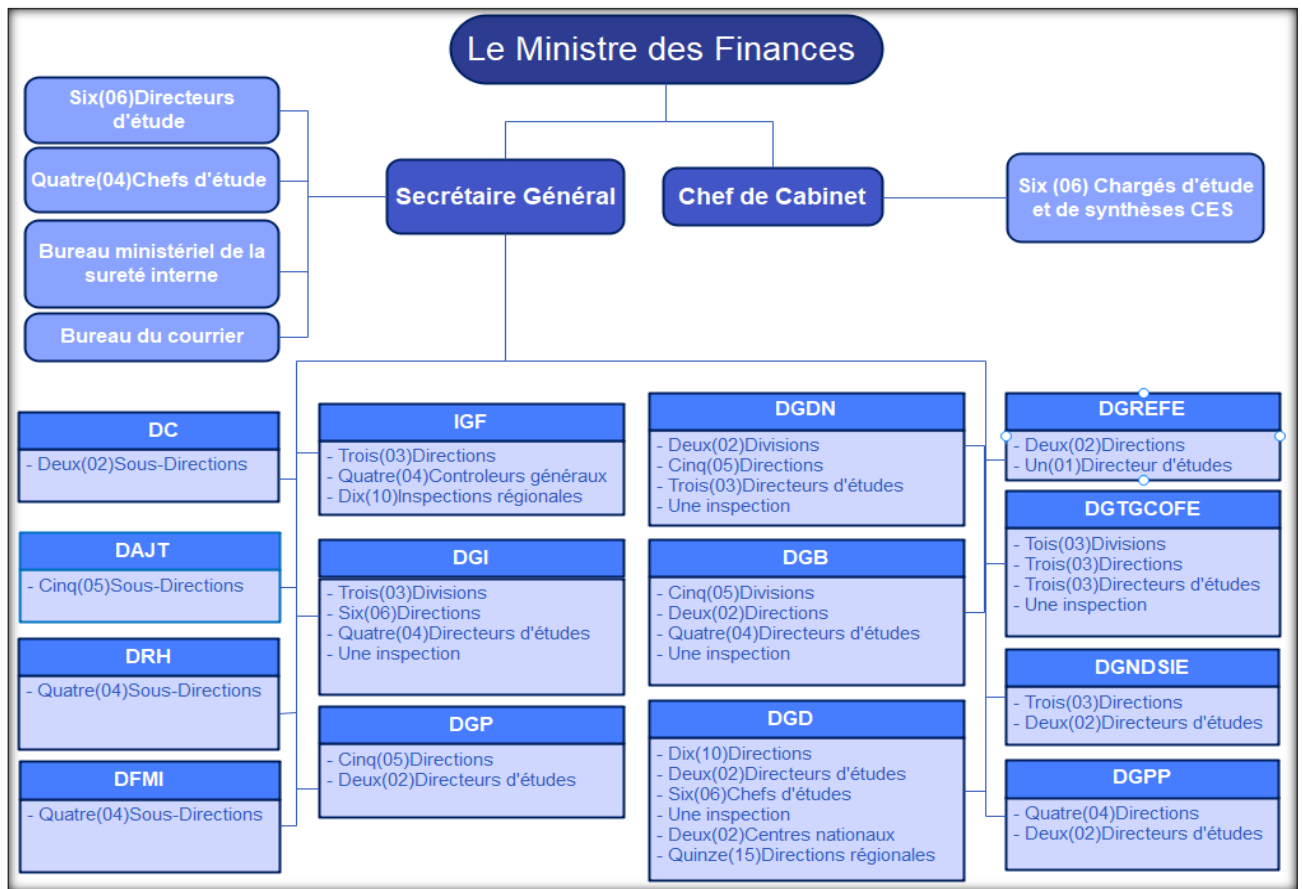
Le Ministère des Finances (MdF) dans sa nouvelle organisation conformément au Décret exécutif N°21-252 du 25 Chaoual 1442 correspondant au 6 juin 2021 portant organisation de l'administration centrale du Ministère des Finances, compte (Ministère des Finances / DGNDISIE, 2024-2028.):

- L'Inspection Générale des Finances (IGF) ;
- Neuf (09) Directions Générales :
 - Direction Générale du Budget (DGB) ;
 - Direction Générale du Trésor et de la Gestion Comptable des Opérations Financières de l'Etat (DGTGCOFE) ;
 - Direction Générale des Impôts (DGI) ;
 - Direction Générale du Domaine National (DGDN) ;
 - Direction Générale des Douanes (DGD) ;
 - Direction Générale de la Prévision et des Politiques (DGPP) ;
 - Direction Générale de la Prospective (DGP) ;
 - Direction Générale de Numérisation, de Digitalisation et des Systèmes d'Information Economiques (DGNDISIE) ;
 - Direction Générale des Relations Economiques et Financières Extérieures (DGREFE);
- Quatre (04) Directions :
 - Direction des Finances, des Moyens et des Infrastructures (DFMI) ;
 - Direction des Ressources Humaines (DRH) ;
 - Direction de la Communication (DC) ;
 - Direction de l'Agence Judiciaire du Trésor (DAJT) ;

En plus du Secrétariat Général et du cabinet de Monsieur le Ministre.

L'organigramme du Ministère des Finances ainsi que l'organisation macro de chacune de ses structures se présentent comme suit :

Figure 4 Organisation du Ministère des Finances

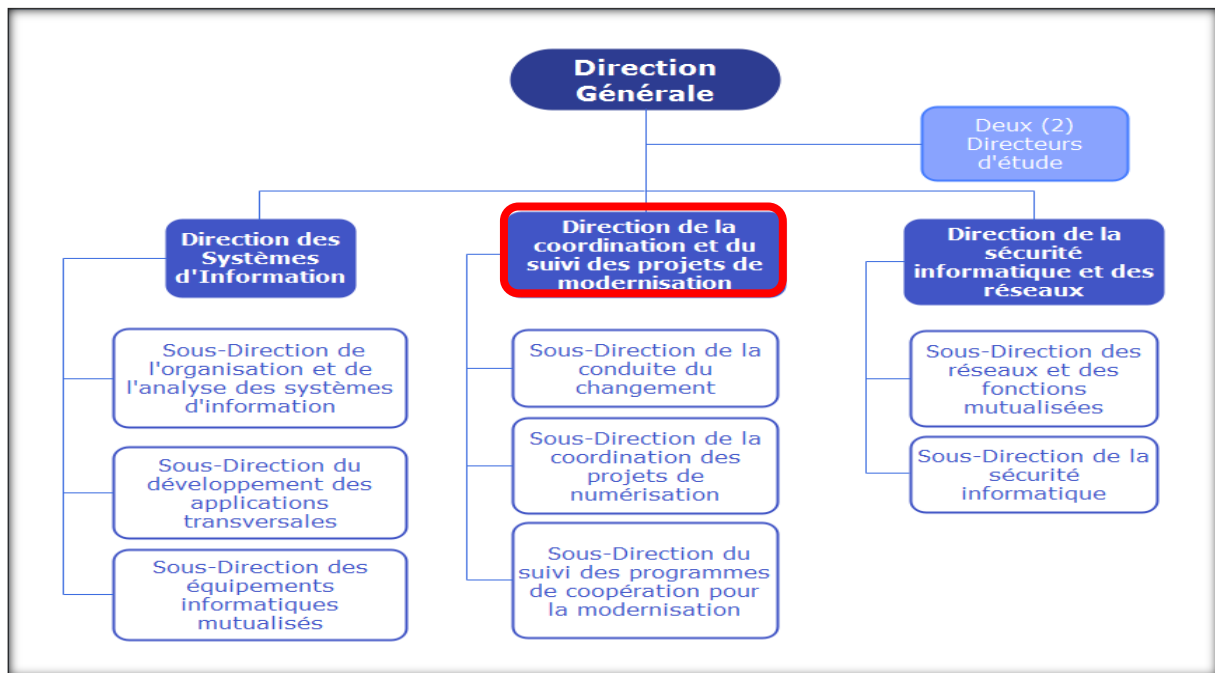


source : (Ministère des Finances / DGNDSIE, 2024-2028.)

Comme le montre l'organigramme ci-dessus, le Ministère des Finances exerce ses activités à travers plusieurs structures de différents types avec des niveaux hiérarchiques différents, qui varient entre : **Divisions, Directions Générales, Directions Centrales, Directions Régionales, Inspections, ...réparties** sur le territoire national. Et cela, en sus des autres entités régies par le MdF, telles que : les Collectivités Territoriales, les Établissements Publics à caractère Économique, les Établissements Publics à caractère Administratif, ... Les activités métiers du MdF sont assurées par les structures suivantes : la DGI, la DGDN, la DGD, la DGB et la DGTGCOFE. Ces structures sont soutenues par les structures supports du MdF, à savoir : la DGPP, la DGP, la DGNDSIE, la DGREFE, la DRH, la DFMI et la DC en plus de la DAJT et l'IGF.

La Direction Générale de la Numérisation, de la Digitalisation et des Systèmes d'Information Économiques (DGNDSIE) est organisée comme suit:

Figure 5 Organisation de la DGNDSIE



source : (Ministère des Finances / DGNDSIE, 2024-2028.)

Remarque:

Le présent mémoire explore la création d'un Système Décisionnel Economique (SDE) au sein de la Direction de la coordination et du suivi des projets de modernisation/DGNDSIE.

2.1.2. Missions du Ministère

Le Ministère des Finances est le principal acteur dans le développement économique. Il joue un rôle central au sein du gouvernement dans les domaines budgétaire, fiscal, financier et comptable, il est ainsi le conseiller du gouvernement en matière financière.

Le Ministère des Finances Algérien a pour mission principale, conformément au décret exécutif N° 95-54 du 15 Ramadhan 1415 correspondant au 15 février 1995 fixant les attributions du Ministre des Finances, la proposition des éléments de la politique nationale en matière financière et la mise en œuvre de cette politique conformément aux lois et règlements en vigueur.

Cette mission générale se traduit par différents métiers, à savoir (Ministère des Finances / DGNDSIE, 2024-2028.):

- Les finances publiques ;
- La fiscalité ;
- La douane ;
- Le domaine national et les affaires foncières ;
- Les dépenses publiques, le budget et la comptabilité publique ;

- La monnaie ;
- L'épargne, le crédit et les assurances économiques ;
- Les ressources du Trésor Public ;
- Les interventions financières de l'État ;
- La politique nationale en matière d'endettement extérieur ;
- Le contrôle des changes ;
- Le contrôle financier relatif aux utilisations des crédits du budget de l'État et des ressources du Trésor Public ;
- Les relations économiques et financières extérieures.

2.1.3. Ressources Humaines Informatiques

Selon (Ministère des Finances / DGNDSE, 2024-2028.), les effectifs informatiques du Ministère, représentent environ 4000 agents tous diplômes et grades confondus.

Ce nombre paraît important, mais en faisant un zoom sur leur répartition, profils et compétences, évoqués plus bas, il s'avère que la quasi-totalité des structures du MdF souffre d'un manque de la ressource humaine qualifiée dans le domaine informatique.

2.1.4. Compétences de la RH Informatique

Selon (Ministère des Finances / DGNDSE, 2024-2028.), les effectifs informatiques du MdF ne sont pas adaptés en compétences aux exigences actuelles de l'informatique.

En effet, il a été constaté que les applications développées, même récemment, majoritairement ne répondent pas aux exigences technologiques actuelles.

Ce constat est le même sur tous les plans (conception, développement, déploiement, sécurisation, ...) et il est dû principalement au fait que la ressource actuelle n'a pas été mise à niveau pour disposer des nouvelles technologies dans le domaine IT.

2.1.5. Analyse du système décisionnel existant

Le Ministère des Finances est en train de mettre en place un système d'information décisionnel, mais qu'il n'est pas encore pleinement opérationnel selon (Ministère des Finances / DGNDSE, 2024-2028.):

- Mr Le Ministre des Finances parle de la "consolidation d'un **système d'aide à la décision**" comme un objectif du schéma directeur. On peut considérer un "système d'aide à la décision" comme une composante d'un système d'information décisionnel.

- Le Directeur Général mentionne la "consolidation d'un **système d'information économique décisionnel**" comme un résultat attendu de la mise en œuvre du schéma directeur.
- La mise en place d'un "Système d'Information des Finances Publiques (SIFP), qui représente le **système décisionnel** du MdF". ce système va consolider les données de pilotage ministériel.
- Le schéma directeur indique une volonté claire de développer et de consolider un système d'information décisionnel au sein du Ministère des Finances, mais il semble que ce soit un projet en cours plutôt qu'une réalité pleinement établie.

2.1.5.1. Caractéristiques du SID existant

Dans cette section, on s'appuie sur les entretiens menés avec les différents acteurs du système décisionnel (les utilisateurs, les experts techniques et les décideurs stratégiques), ainsi que sur une analyse approfondie de plusieurs documents, dont un **document confidentiel relatif à la stratégie nationale de numérisation du Ministère des Finances pour la période 2024-2028**. Cette étude combinée nous a permis de dresser un portrait précis du Système d'Information Décisionnel (SID) actuel, en mettant en avant ses caractéristiques principales.

Le système décisionnel actuel du ministère des Finances repose sur **une architecture de systèmes hétérogènes**, alimentés par différentes directions et entités, et **une absence d'un système d'information décisionnel intégré et performant** (Ministère des Finances / DGNDISIE, 2024-2028.). Il est caractérisé par :

- Inexistence d'un système d'information des finances publiques structuré et consolidé pour la prise de décision;
- Cloisonnement fonctionnel important entre les différentes directions métiers, Ce qui limite l'échange fluide et la consolidation des données nécessaires à une aide à la décision efficace;
- Hétérogénéité des applications et des bases de données utilisées, sans référentiel commun ni gouvernance centralisée des données;
- Manque de référentiel de processus métier et de référentiel de données, ce qui rend difficile la consolidation et l'analyse des informations à des fins décisionnelles ;
- Des outils de reporting traditionnels (type Datawarehouse, OLAP), souvent développés en interne ou via des solutions bureautiques, ce qui limite la capacité

d'analyse multidimensionnelle et la génération de rapports stratégiques et ne permettant pas une analyse dynamique en temps réel;

- Une multitude de bases de données sectorielles détenues par diverses entités (Trésor, Douanes, Impôts, etc.), souvent cloisonnées et peu intégrées;
- Des flux de données non automatisés, limitant la réactivité dans la prise de décision stratégique;
- Risques liés aux systèmes obsolètes : La présence de systèmes obsolètes, utilisant des technologies dépassées, expose le ministère à des risques de sécurité, de compatibilité et de performance, ce qui peut compromettre la fiabilité des informations et la capacité à prendre des décisions rapides et efficaces.

2.1.5.2. Forces et Faiblesses du SID existant

Dans cette section, nous proposons une analyse des points forts et des limites du Système d'Information Décisionnel (SID) actuellement en place au Ministère des Finances. Les informations présentées proviennent d'un document confidentiel qui nous a été remis, retraçant la stratégie nationale de numérisation des systèmes économiques pour la période 2024-2028, élaborée par le Ministère des Finances. Ce document officiel offre une vision stratégique claire des capacités existantes ainsi que des priorités à suivre pour optimiser et renforcer la performance du SID.

➤ Forces :

- Existence d'un socle de données riche, réparti sur les différentes directions.
- Capacités techniques internes (informaticiens, analystes) capables de développer des tableaux de bord.
- Prise en compte croissante de la donnée comme actif stratégique dans le SDSSI.

➤ Faiblesses:

- Fragmentation de l'information : données cloisonnées, difficulté d'accès transversal.
- Faible automatisation et intégration des systèmes décisionnels (absence d'architecture datawarehouse centralisé ou de couche OLAP).
- Problèmes de qualité et de traçabilité des données échangées entre les systèmes.
- Sécurité hétérogène : les règles de sécurité varient d'une entité à une autre, ce qui rend difficile une vision unifiée sécurisée.
- Temps d'accès à l'information élevé, ralentissant le processus décisionnel.

En résumé, le système décisionnel existant est fragmenté, peu mature, orienté vers des rapports statiques et peu adapté aux besoins analytiques dynamiques des décideurs ministériels et ne répond pas aux besoins croissants de pilotage stratégique et opérationnel du ministère.

2.2. Cadre méthodologique

2.2.1. Le choix du sujet

Compte tenu de la confidentialité et la complexité d'implémentation d'un système décisionnel classique (Datawarehouse, OLAP, etc.), La **virtualisation de données** est une approche qui permet d'accéder, de fédérer et d'interroger les données en temps réel, sans les déplacer physiquement. Elle présente plusieurs avantages dans le contexte du Mdf, à savoir:

-Adaptée à la confidentialité des données

Les données restent dans leur source d'origine (impôts, douanes, trésor...), ce qui limite le risque lié au transfert ou à la duplication des données sensibles.

Elle permet de **respecter les règles de sécurité propres à chaque entité** tout en fournissant une vue consolidée.

-Réponse à l'hétérogénéité

La virtualisation permet de connecter des sources hétérogènes (bases Oracle, SQL Server, fichiers Excel, etc.) sans imposer un format commun ou un entrepôt centralisé.

Elle permet de créer **des vues virtuelles unifiées** à partir de systèmes disparates, sans coût élevé d'intégration physique.

-Complexité d'un datawarehouse traditionnel

La mise en place d'un datawarehouse nécessite un lourd investissement en infrastructure, ETL, sécurité, et gouvernance.

Les données doivent être copiées et synchronisées régulièrement, ce qui augmente les risques d'obsolescence ou de divergence.

- Agilité** : permet de créer rapidement des tableaux de bord analytiques.
- Performance** : accès en temps réel aux données avec des moteurs d'optimisation.
- Sécurité renforcée** : chaque source conserve son contrôle d'accès ; les droits sont fédérés au niveau de la couche de virtualisation.
- Coût réduit** : pas de duplication massive ni d'infrastructure lourde comme pour les entrepôts de données classiques.

La data virtualisation devient ainsi **un levier stratégique pour un système décisionnel performant, sécurisé et interconnecté**, particulièrement dans les environnements complexes et sensibles comme celui du ministère des Finances.

L'intégration de la métadonnée est essentielle pour :

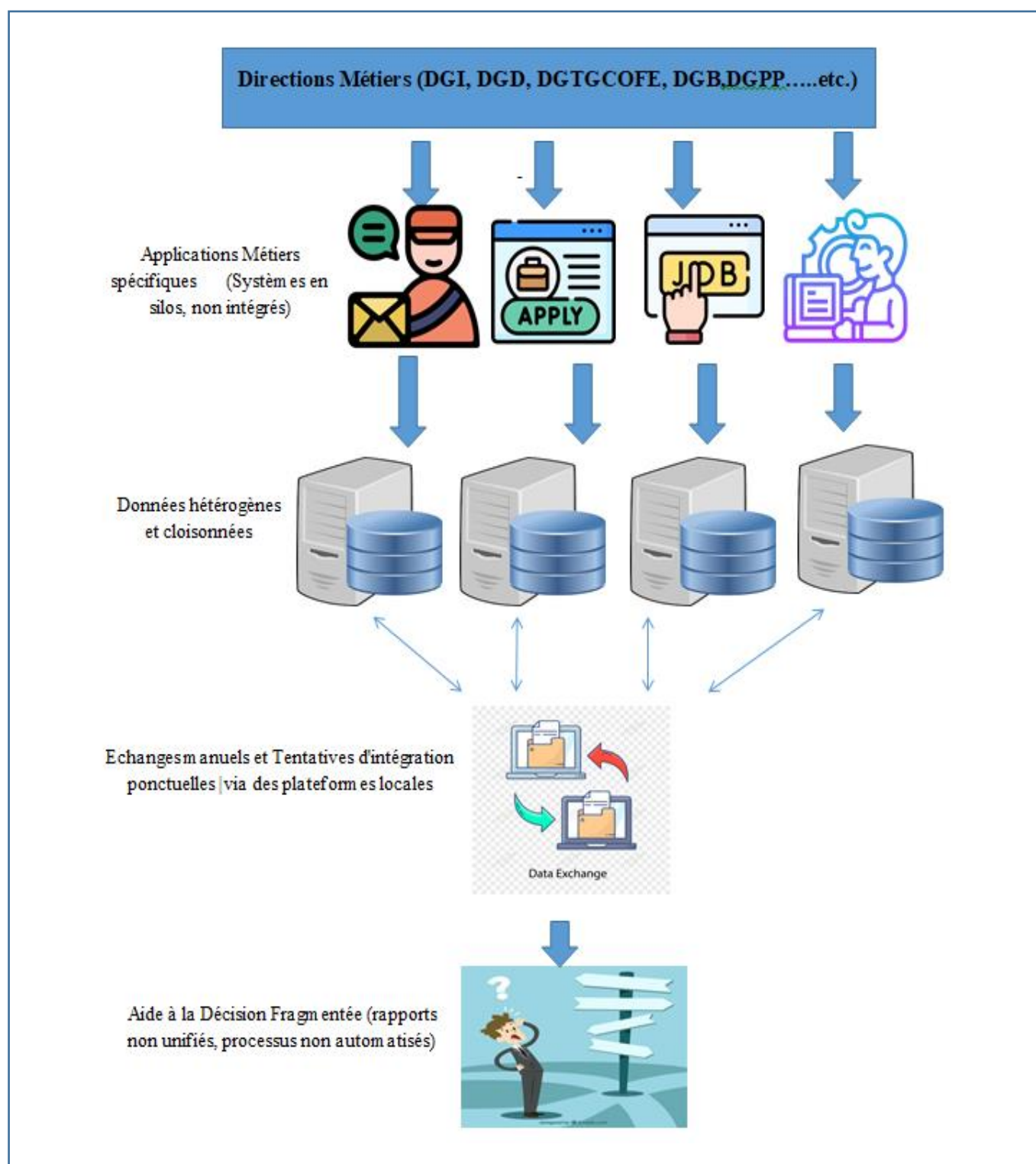
- **Compréhension des données:** La métadonnée fournit des informations sur les données (origine, signification, format, qualité, etc.), ce qui facilite leur compréhension et leur utilisation.
- **Gouvernance des données:** Elle permet de gérer et de contrôler les données, en assurant leur qualité, leur cohérence et leur conformité.
- **Traçabilité des données:** Elle permet de suivre l'origine et les transformations des données, ce qui est crucial pour la fiabilité des analyses et la conformité réglementaire.

Au niveau du Ministère des Finances, ces avantages peuvent être particulièrement importants pour :

- **Consolider les données financières:** Obtenir une vue globale et cohérente des finances publiques.
- **Améliorer la prise de décision:** Fournir aux décideurs des informations précises et pertinentes pour la gestion des finances.
- **Renforcer la transparence et la responsabilité:** Faciliter l'accès aux données financières pour les citoyens et les institutions.
- **Optimiser les processus:** Automatiser et rationaliser les processus financiers.

À partir des éléments du *Schéma Directeur Stratégique des Systèmes d'Information 2024-2028* du Ministère des Finances, voici une **schématisation synthétique du système décisionnel existant (actuel)**

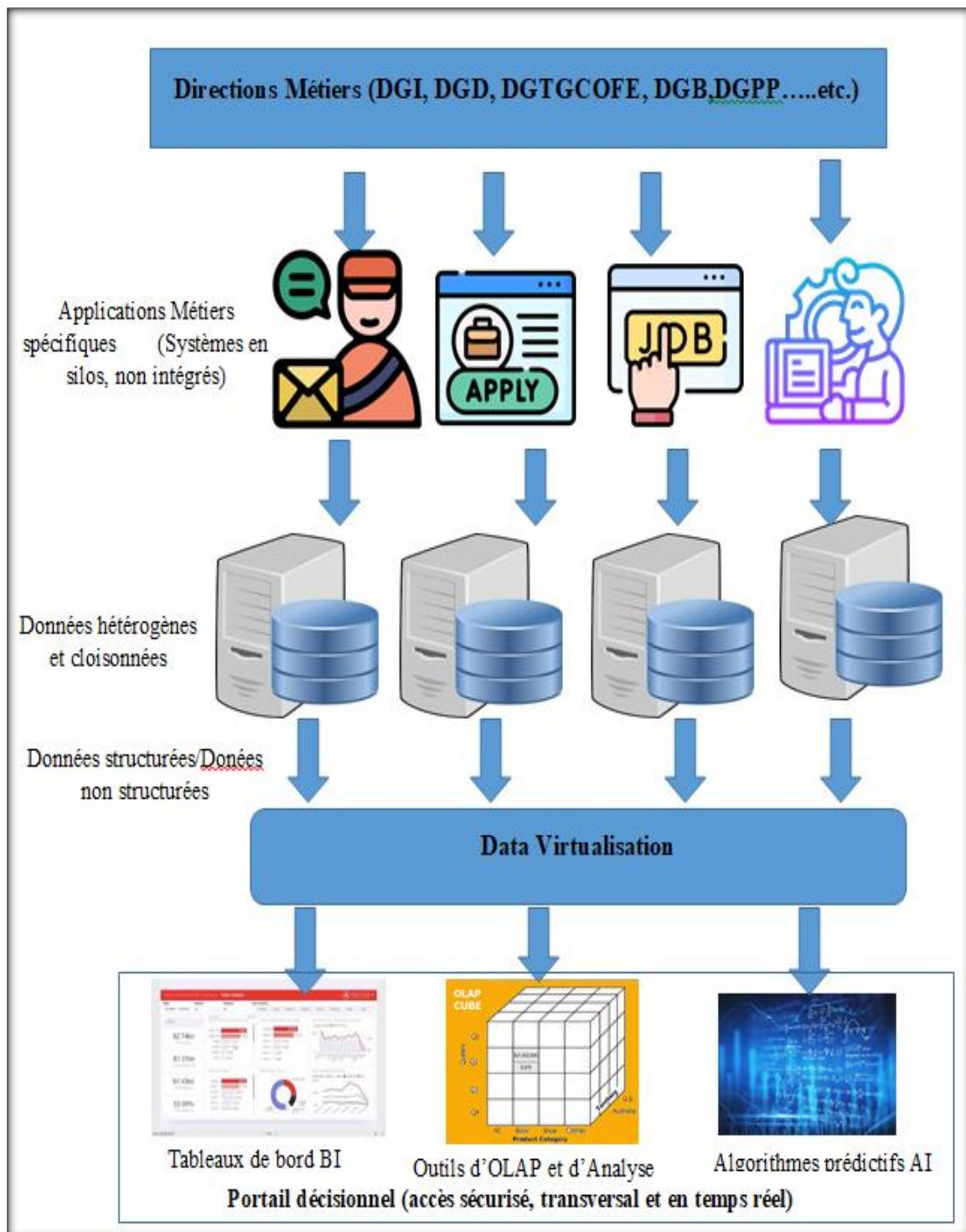
Figure 6: Schéma du Système Décisionnel Actuel du Mdf



source : [élaboré par nos soins]

Voici une vue du système décisionnel cible du Ministère des Finances, intégrant le principe de **la virtualisation de données**:

Figure 7: Schéma du SID cible avec Data virtualisation



source : [élaboré par nos soins]

2.2.2. La position épistémologique et méthode de recherche

2.2.2.1. La position épistémologique

La position épistémologique concernant l'intégration d'une solution de virtualisation des données dans un système d'aide à la décision repose sur une approche **constructiviste et pragmatique** de la connaissance. En effet, ce sujet implique la compréhension des phénomènes techniques, organisationnels et humains liés à cette innovation numérique, ce qui nécessite une posture ouverte à la complexité et à la multi dimensionnalité. (Kitchin, 2014) (Chen, 2012)

La virtualisation des données peut être vue comme une couche d'abstraction qui permet d'intégrer des données hétérogènes issues de multiples sources sans duplication physique, rendant les données accessibles « à la demande » tout en sécurisant et en optimisant leur gestion. Cette réalité technique est construite dans un contexte socio-organisationnel, où la valeur des données et leur gouvernance prennent un sens différent selon les usages et les acteurs impliqués. Ainsi, l'épistémologie dominante s'appuie sur une démarche qualitative, exploratoire et interprétative qui prend en compte les interactions complexes entre la technologie, les pratiques managériales et les objectifs stratégiques (Benbasat, 1987) (Orlikowski, 1991)

Par ailleurs, la position pragmatique soutient que la connaissance produite doit être directement utile à la pratique, notamment pour fournir des recommandations opérationnelles au ministère des Finances. Cette posture favorise l'intégration entre savoirs théoriques et savoirs d'action, permettant d'éclairer les conditions d'une adoption réussie de la virtualisation dans un système d'aide à la décision. L'épistémologie s'inscrit donc aussi dans un cadre critique, évaluant les impacts, bénéfices et limites, tout en cherchant à optimiser la gouvernance et l'efficacité du système. (Laudon, 2015) (Armstrong, 2019)

En résumé, la réflexion épistémologique autour de ce sujet navigue entre :

- Un constructivisme méthodologique qui valorise la compréhension contextuelle, l'analyse des pratiques et la signification des données et technologies dans leur environnement.
- Un pragmatisme orienté vers les résultats opérationnels et l'amélioration continue des systèmes d'aide à la décision par la virtualisation.
- Une approche interdisciplinaire mobilisant informatique, sciences de gestion, et études organisationnelles, pour saisir la complexité du déploiement et de l'usage de cette technologie innovante. (Simmonds, 2020) (Kitchin, 2014)

Cette posture favorise une recherche qualitative, agile, capable de dialoguer avec les différents acteurs et d'intégrer les dimensions techniques, humaines et stratégiques dans l'analyse et la production de connaissances.

2.2.2.2. La méthode de recherche

Cette méthode de recherche assure une investigation rigoureuse et complète des enjeux liés à l'intégration d'une solution de virtualisation des données dans un système d'aide à la décision, en conciliant profondeur analytique et utilité opérationnelle.

2.2.2.2.1. Les méthodes utilisées dans le terrain d'enquête

La méthodologie de cette recherche s'appuie principalement sur une approche qualitative exploratoire visant à comprendre en profondeur les dynamiques d'intégration de la virtualisation des données dans un système d'aide à la décision. Sur le terrain, la collecte des données se fera par le biais d'entretiens semi-directifs auprès des utilisateurs finaux, des décideurs et des experts techniques impliqués dans le projet. Ces entretiens permettront de recueillir des témoignages riches, détaillés et contextualisés sur les perceptions, les pratiques et les difficultés rencontrées au cours du processus d'intégration. Par ailleurs, des observations participantes et l'analyse de documents internes (rapports, cahiers des charges, protocoles de gestion) viendront compléter le recueil d'informations numériques et qualitatives.

Cette triangulation des sources garantit une connaissance fine et plurielle du phénomène étudié, indispensable dans un contexte sociotechnique aussi complexe que l'intégration de la virtualisation. L'analyse croisée de ces données permettra de dégager des tendances et des leviers spécifiques à ce contexte organisationnel, tout en identifiant les freins susceptibles de freiner l'adoption.

- **L'étude qualitative**

L'étude adoptera une démarche qualitative afin de saisir non seulement les aspects techniques, mais aussi les dimensions humaines, culturelles et organisationnelles qui façonnent le succès ou l'échec de l'intégration. Contrairement à une approche quantitative qui viserait à mesurer des indicateurs précis, la recherche qualitative se focalise sur la signification donnée par les acteurs aux usages et transformations induits par la virtualisation des données. Elle offre ainsi une compréhension riche, nuancée et évolutive des enjeux liés à l'innovation, notamment en

ce qui concerne l'alignement stratégique, la gouvernance des données et la satisfaction des utilisateurs.

L'analyse thématique sera la principale méthode d'analyse des données recueillies, permettant d'identifier et de catégoriser les grands thèmes émergents autour des bénéfices perçus, des défis rencontrés et des conditions favorables. Cette méthode donne la latitude d'intégrer les variations contextuelles et les divergences d'opinions, facteurs fondamentaux dans un cadre décisionnel centralisé mais multi-acteurs.

- **Le choix d'échantillon**

Le choix de l'échantillon sera raisonné et intentionnel, basé sur un échantillonnage de type « par critères ». Il s'agira de sélectionner des participants qui représentent de façon pertinente les différentes catégories d'acteurs concernés par l'intégration de la virtualisation :

Des utilisateurs opérationnels du système d'aide à la décision, qui utilisent au quotidien les données virtualisées pour leurs analyses ;

Des décideurs stratégiques responsables de la mise en œuvre et de l'adoption de la technologie ;

Des experts techniques en charge de la conception, du déploiement et de la maintenance de la solution.

- **La population interviewée**

La population interviewée sera constituée des acteurs clés concernés par l'intégration de la virtualisation dans le système d'aide à la décision. Les participants incluront principalement :

Les utilisateurs opérationnels qui exploitent le système au quotidien pour la prise de décisions, apportant un éclairage sur l'expérience d'usage, les bénéfices perçus et les difficultés pratiques.

Les décideurs stratégiques du ministère, responsables des orientations et des investissements technologiques, chargés de piloter l'adoption et d'aligner la solution avec les objectifs métiers.

Les experts techniques et informaticiens, qui conçoivent, implémentent et maintiennent la plateforme de virtualisation, intervenant sur les aspects technologiques, d'interopérabilité et de sécurisation des données

Tableau 2: Liste des interviewés

Catégorie	Poste occupé	Durée de l'entretien
Utilisateurs opérationnels	1. Cadre 1, Direction générale des Finances publiques (DGFIP)	45 minutes
	2. Cadre 2, Direction générale des Finances publiques (DGFIP)	45 minutes
Décideurs stratégiques	1. Sous-Directeur de la coordination des projets de numérisation	60 minutes
	2. Directeur Générale de la Numérisation, de la Digitalisation et des Systèmes d'Information Économiques	60 minutes
Experts techniques	1. DB administrator, Service informatique	45 minutes
	2. Ingénieur systèmes et réseaux Service informatique	45 minutes
	3. Responsable sécurité, Service informatique	45 minutes

Source: élaboré par nos soins

La diversité de cette population permet de recueillir des données qualitatives riches et complémentaires, couvrant les dimensions humaines, techniques et organisationnelles du projet. La sélection se fera en veillant à représenter différents niveaux hiérarchiques, plusieurs services et si possible divers sites géographiques, afin de capter des perspectives variées.

Cette population cible est essentielle pour identifier les conditions favorables, les freins et les leviers d'une intégration réussie, tout en garantissant la pertinence et la validité des recommandations opérationnelles proposées.

• Méthode de collecte des données

La collecte des données est un élément essentiel qui conditionne la qualité et la pertinence des résultats de la recherche. Pour notre étude qualitative portant sur l'intégration d'une solution de virtualisation de données dans un système d'aide à la décision, plusieurs méthodes seront mobilisées afin d'obtenir une vision riche, contextualisée et multidimensionnelle du phénomène étudié.

- Entretien individuel

L'entretien individuel semi-directif constitue la méthode principale de collecte. Il s'agit d'une conversation structurée autour d'un guide d'entretien, offrant la possibilité d'explorer en

profondeur les expériences, perceptions, enjeux et attentes des participants. Cette flexibilité dans la conduite de l'entretien permet d'adapter les questions aux spécificités de chaque interlocuteur, tout en conservant un cadre cohérent pour assurer la comparabilité des données.

Pour les utilisateurs opérationnels, le guide d'entretien portera sur leur expérience quotidienne d'utilisation du système décisionnel existant, les bénéfices à percevoir, les difficultés rencontrées, et leurs attentes en termes de fiabilité et d'ergonomie.

Pour les décideurs stratégiques, il s'agira d'explorer les motivations et objectifs derrière l'adoption de la virtualisation, les critères d'évaluation du succès du projet, ainsi que les enjeux organisationnels et stratégiques.

Pour les experts techniques, le focus sera mis sur les choix techniques, les défis liés à l'intégration, la gestion des données et de la sécurité, ainsi que sur les contraintes d'interopérabilité.

- Analyse documentaire

En complément des entretiens, une analyse documentaire sera effectuée. Elle portera sur les documents internes, des rapports de spécifications techniques et des procédures de gouvernance des données. Cette méthode permettra de trianguler les données recueillies lors des entretiens, d'objectiver certains éléments et d'enrichir la compréhension du contexte organisationnel et technique.

- Observation participante ou non participante

Dans la mesure du possible, une ou plusieurs séances d'observation seront menées pour voir comment le système décisionnel existant est utilisé en situation réelle, quelles interactions techniques et humaines émergent, et comment les collaborateurs vont s'adapter à cette nouvelle solution. Cette méthode permettra d'identifier des pratiques informelles, des résistances ou des ajustements non exprimés explicitement lors des entretiens.

2.2.2.2.2. Guide d'entretien pour les utilisateurs opérationnels

L'objectif est de recueillir leur expérience quotidienne et leurs ressentis vis-à-vis du système décisionnel existant, ainsi que la facilité d'utilisation, les avantages perçus, les problèmes rencontrés et leurs attentes pratiques pour améliorer l'ergonomie et la fiabilité.

Axes :

- Description du rôle et des tâches ;
- Facilité d'accès et rapidité d'utilisation ;
- Fiabilité et disponibilité des données ;
- Difficultés rencontrées (techniques, organisationnelles) ;
- Suggestions pour améliorer l'outil (ergonomie, fonctionnalités).

2.2.2.2.3. Guide d'entretien pour les décideurs stratégiques

Le but est ici d'explorer les motivations de la virtualisation, les objectifs institutionnels, les critères d'évaluation du projet, les implications organisationnelles, enjeux de gouvernance des données et les recommandations stratégiques.

Axes :

- Raisons du choix de la virtualisation de données ;
- Objectifs attendus sur la prise de décision ;
- Indicateurs de succès et suivi des résultats ;
- Adaptation des processus organisationnels ;
- Risques, obstacles et mécanismes de gestion ;
- Leçons tirées et conseils pour l'avenir

2.2.2.2.4. Guide d'entretien pour les experts techniques

L'entretien vise à comprendre les aspects techniques, Architecture et choix technologiques à déployés, des aspects sécuritaires, des processus de maintenance, les contraintes associées à la virtualisation de données et les difficultés d'intégration et d'interopérabilité ainsi que des retours utilisateurs sur la technique.

Axes :

- Description technique détaillée ;
- Problèmes rencontrés et solutions adoptées ;
- Protocoles de sécurité et intégrité des données ;
- Organisation de la maintenance et mises à jour ;
- Retour utilisateur sur la technique ;
- Perspectives techniques futures ;

Ces guides d'entretien semi-directifs fournissent un cadre structuré mais souple, permettant de recueillir des données qualitatives denses et exploitables pour analyser de manière approfondie l'intégration de la virtualisation des données dans le système d'aide à la décision. Ils favorisent ainsi une compréhension fine des dimensions techniques, organisationnelles et humaines du projet.

Chapitre 3: Modélisation de la solution de virtualisation de données

Chapitre 3 : Modélisation de la Solution de virtualisation de données

La modélisation d'une solution de virtualisation des données représente une étape fondamentale qui consiste à concevoir une représentation logique et fonctionnelle intégrée des diverses sources de données hétérogènes auxquelles l'organisation doit accéder.

Cette modélisation vise à créer une couche d'abstraction centralisée qui simplifie la gestion, l'accès et l'intégration des données sans nécessité de duplication physique, tout en assurant gouvernance, sécurité et performance. Concrètement, la solution virtualisée permet de présenter aux utilisateurs finaux une vue unifiée et cohérente des données, quel que soit leur emplacement, leur format ou leur système d'origine.

Objectifs :

- Établir une architecture cible du système décisionnel virtualisé ;
- Concevoir des vues de données virtuelles pour les décideurs ;
- Mettre en place un prototype fonctionnel ;
- Définir les règles de transformation et d'intégration des données ;
- Élaborer des maquettes de tableaux de bord et rapports décisionnels.

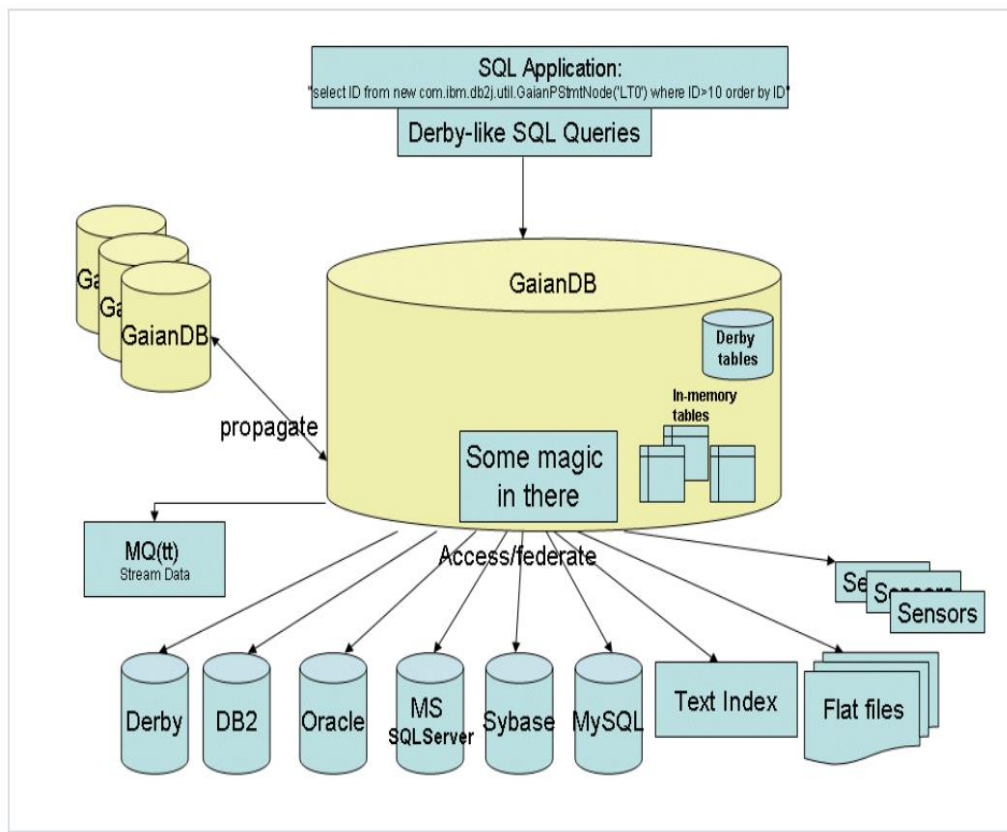
3.1. L'Architecture cible du système décisionnel virtualisé

Le tableau ci-dessous illustre l'architecture cible à quatre niveaux d'un système décisionnel virtualisé, dérivée de l'architecture décisionnelle classique à trois niveaux par l'ajout d'une **couche de virtualisation de données** :

Couche	Description
Sources de données	Bases Opérationnelles, Systèmes financiers, Douanes , Achats et des référentiels externes
Couche de virtualization de données	Moteur de virtualisation de données (Ex : Denodo plateforme)
Couche décisionnelle	Outils de BI et tableaux de bord connectés à la couche virtualisée
Sécurité et gouvernance	Contrôle d'accès, chiffrement, audit, gestion des métadonnées4.

Tableau 3 Architecture cible du système décisionnel virtualisé

Figure 8: Architecture de haut niveau de GaianDB



Source : (David Vyvyan, 2023)

3.2. Vues de données virtuelles pour les décideurs

3.2.1. Définition d'une vue de données

Selon (David Vyvyan, 2023), une vue de données (ou vue en abrégé) est une représentation virtuelle d'une table ou d'un ensemble de tables dans une base de données. Elle est généralement définie par une requête SQL et ne stocke pas physiquement les données, mais présente les résultats d'une requête comme s'ils étaient dans une table.

En d'autres termes, une vue permet d'abstraire et simplifier l'accès aux données en exposant uniquement certaines colonnes ou lignes, ou en combinant plusieurs tables de manière logique. Elle est souvent utilisée pour des raisons de sécurité, de facilité d'utilisation, ou pour masquer la complexité des structures sous-jacentes.

Deux étapes essentiels:

- **Recueil des besoins des décideurs** (indicateurs clés, axes d'analyse);

- **Modélisation des vues de données** : les vues de base (correspondant aux tables/fichiers sources), les vues d'intégration (jointures, enrichissements, élimination de la redondance) et les vues métiers (pré-calculs, agrégations selon les besoins des décideurs).
- **Sources** : Bases SQL Server, SIGAPE, MEGA, GAMS, etc.

3.2.2. Modélisation logique avec un moteur de virtualisation

- Vues de base : importées depuis les SGBD existants (SQL Server, MySQL...).
- Vues d'intégration : jointures, nettoyages (normalisation, suppression des doublons).
- Vues de présentation (métiers) : prêtes pour être consommées dans les tableaux de bord (ex. indicateurs macroéconomiques, fiscalité, dépenses publiques).

3.2.3. Exemples des vues métiers :

-Vue_Macro_Eco_Indicateurs

- Agrégation les données de prévision économique (SIGAPE), dépenses (MEGA), recettes fiscales ;
- Transformation et standardisation des formats temporels, conversion de devises, anonymisation si nécessaire.

-Vue_budget_global

- Agrégation des dépenses et recettes par entité ;
- Suivi budgétaire global.

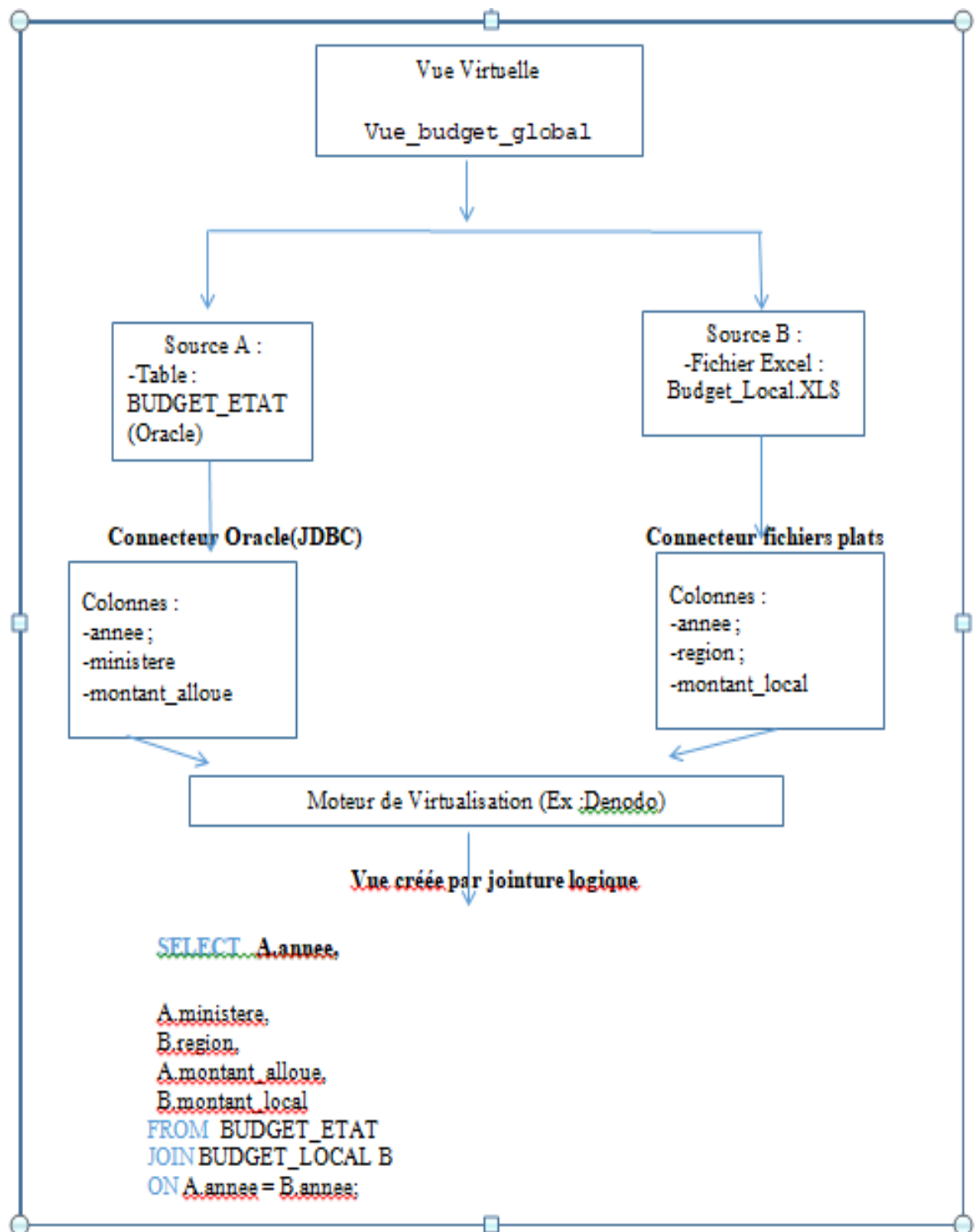
-Vue_recettes_fiscales

- Recettes par type d'impôt et par région ;
- Evaluation de la performance fiscale.

*Exemple d'un schéma conceptuel d'une vue virtuelle avec la requête Sql :

La figure ci-dessous représente le schéma conceptuel de la vue budget global comme exemple

Figure 9 : Schéma conceptuelle d'une vue virtuelle



Source : élaboré par nos soins.

3.2.4. Illustration des vues virtuelles Métier en Sql:

-vue_synthese_budgetaire :

```
CREATE VIEW vue_synthese_budgetaire AS
SELECT
b.annee,
b.code_ministere,
m.nom_ministere,
SUM(b.montant_budget) AS montant_total,
SUM(r.montant_realise) AS montant_realise
FROM
    budgets b
JOIN ministeres m ON b.code_ministere = m.code_ministere
LEFT JOIN realisations r ON b.id_budget = r.id_budget
GROUP BY
b.annee, b.code_ministere, m.nom_ministere;
```

Objectif de la vue : le reporting budgétaire, accessible aux décideurs sans requêtes complexes à partir de systèmes sources hétérogènes.

-vue_recette_fiscale,consolidee :

```
CREATE VIEW vue_recette_fiscale_consolidee AS
SELECT
f.annee,
f.type_recette,
SUM(f.montant) AS montant_total
FROM
    fiscalite f
GROUP BY f.annee, f.type_recette;
```

Objectif de la vue : Vue consolidée des recettes fiscales de différentes sources et applications de fiscalité.

3.3. Mise en place d'un prototype fonctionnel

L'objectif du prototype : valider la faisabilité technique avec un jeu réduit de données pour fournir un accès unifié, sécurisé et en temps réel aux données provenant de différentes sources hétérogènes (bases relationnelles, fichiers plats, entrepôts, services web...) sans déplacer physiquement les données.

3.3.1. Les quatre (04) étapes principales du modèle fonctionnel

- **Etape 1:** Définition du Périmètre et des Objectifs

- Identification des besoins décisionnels prioritaires: quels sont les domaines décisionnels clés du ministère qui pourraient bénéficier le plus de l'intégration de données virtualisées? (ex: suivi budgétaire, GRH, évaluation des politiques publiques, etc.).

- Identifier au moins deux ou trois sources de données pertinentes pour les besoins identifiés Antérieurement (ex: bases de données transactionnelles existantes, fichiers Excel, applications spécifiques, etc.). à condition qu'elles doivent être de formats différents si possible pour illustrer la capacité de la virtualisation à gérer l'hétérogénéité.

- Choisir quelques KPIs simples mais significatifs qui pourront être calculés et visualisés à partir des données virtualisées.

- Sélectionner un outil de virtualisation de données approprié pour le prototype. Plusieurs options existent, y compris des solutions open source (voir Etape 2). Les critères de choix peuvent inclure la facilité d'utilisation, la connectivité aux sources de données sélectionnées et les capacités de manipulation et de transformation des données.

- **Etape 2 :** Sélection des Outils et de l'Architecture

- Choisir l'Outil de Virtualisation de Données, nous mentionnons quelques outils et plateformes de virtualisation comme **Apache NiFi, Teiid et Denodo (Édition Communautaire)**.

Mais, il est recommandé de choisir un outil open source.

- Accéder aux sources de données **PostgreSQL ou SQL Server** (SGBD existants au MdF) et les Fichiers CSV/Excel

- Choisir l'outil BI pour la visualisation, **Apache Superset, Metabase et Tableau Public**.

- **Etape 3 :** Mise en œuvre du Prototype

- Installer et configurer l'outil de virtualisation de données choisi sur un serveur ou une machine virtuelle accessible.
- Établir les connexions entre l'outil de virtualisation et les sources de données identifiées en utilisant les connecteurs appropriés (JDBC, Fichiers plats...etc)
- Définir le modèle de données virtuel en créant des vues abstraites qui combinent et transforment les données provenant des différentes sources. Cela implique l'identification des champs clés et les relations entre les données, la transformation de données si nécessaire (nettoyage, formatage, agrégation) et la création des vues logiques pour les KPIs définis.
- Connecter l'outil de BI à la vue de données virtuelle créée dans le modèle de données virtuel. L'outil de BI doit pouvoir interroger la vue virtuelle comme s'il s'agissait d'une base de données unique.
- Utiliser l'outil de BI pour créer des tableaux de bord et des rapports.

- **Etape 4 : Tests et Evaluation**

- Vérifier que l'outil de virtualisation se connecte correctement à toutes les sources de données et que l'outil de BI peut accéder à la vue de données virtuelle.
- S'assurer que les données présentées dans les tableaux de bord et les rapports sont exactes et cohérentes avec les données sources.
- Vérifier que le prototype répond aux besoins décisionnels identifiés et que les KPIs peuvent être suivis efficacement.
- Recueillir les commentaires des utilisateurs potentiels du système décisionnel (feedback) sur la facilité d'utilisation et la pertinence des informations présentées.

3.3.2. Les Résultats attendus du prototype:

- Accès unifié aux données: Les utilisateurs peuvent accéder aux informations nécessaires sans se soucier de l'emplacement ou du format des données sources.
- Réduction de la duplication des données: La virtualisation évite la nécessité de créer des copies physiques des données, réduisant ainsi le stockage et les efforts de maintenance.
- Agilité et flexibilité: L'ajout de nouvelles sources de données ou la modification du modèle de données virtuel est plus rapide et moins coûteux que dans les approches traditionnelles d'intégration de données.
- Amélioration de la qualité des données: La virtualisation permet d'appliquer des règles de qualité et des transformations de données de manière centralisée.

-Accélération du processus décisionnel: Un accès plus rapide et plus facile à des informations cohérentes permet aux décideurs de prendre des décisions plus éclairées et plus rapidement.

3.4. Les règles de transformation et d'intégration des données

- La transformation de données s'agit de :
 - **Nettoyage** : suppression des doublons, correction des erreurs, traitement des valeurs manquantes.
 - **Normalisation** : standardisation des formats (dates, unités de mesure, codifications).
 - **Enrichissement** : ajout de données issues de référentiels ou calculs dérivés.
 - **Historisation** : gestion des versions et de l'évolution temporelle des données (dimension temporelle pour les faits).
- Les règles d'intégration s'appliquent lors de la collecte et du chargement des données dans le système décisionnel:
 - **Interopérabilité** : Respect strict du **Référentiel National d'Interopérabilité (RNI)** qui impose l'usage de normes ouvertes, protocoles standards, formats d'échange (XM...etc.).
 - **Gestion des métadonnées** : Une gestion rigoureuse des métadonnées doit être mise en place pour documenter l'origine et la qualité de données intégrées.
 - **Évolutivité** : L'intégration doit permettre l'ajout progressif de nouvelles sources de données sans refonte.
 - **Suivi et contrôle qualité** : mécanismes de contrôle de la qualité de données intégrées doivent être mis en considération, avec des indicateurs de performance.
- Pour garantir la cohérence globale du système décisionnel, un **cadre de gouvernance** est crucial pour :
 - Définition des **rôles et responsabilités**.
 - **Politiques de sécurité, confidentialité et conformité réglementaire**.
 - Suivi de **la qualité et de la traçabilité** des données utilisées dans les analyses.

3.5. Élaboration des maquettes des tableaux de bord et des rapports décisionnels

Objectifs des tableaux de bord et des rapports décisionnels

- Valoriser les données consolidées et virtualisées.

- Appuyer le pilotage opérationnel et stratégique.
- Faciliter la prise de décision pour les directions générales.
- Assurer la transparence, la performance et l'efficacité.

Les résultats des tableaux de bord décisionnels doivent être **alignés avec les besoins spécifiques des décideurs**. Cela signifie qu'ils doivent fournir des informations claires, pertinentes et exploitables pour appuyer la prise de décision stratégique.

- **Exemples des tableaux de bord et rapports décisionnels (de quelques directions)**
 - Tableau de bord de performance fiscale – DGI

Figure 10 : Rapport décisionnel : Situation macroéconomique trimestrielle – DGPP/DGP

Indicateur	Valeur actuelle	Objectif mensuel	Tendance	État
Taux de recouvrement	83%	90%	 +2 pts	 Moyenne
Assiette fiscale consolidée	9 400 MDA	10 000 MDA	 -1.5%	 À surveiller
Délai moyen de traitement	7 jours	≤ 5 jours	 Stable	 Acceptable
Contentieux en cours	1 340 dossiers	≤ 1 000	 +6%	 Risque

- **Tableau** :
 - Taux d'inflation (ONS)
 - Taux de change moyen (Banque d'Algérie)
 - Solde budgétaire
 - Balance commerciale (DGD)
- **Visualisations** :
 - Graphiques en barres empilées (recettes/dépenses)
 - Courbes temporelles (évolution du PIB, taux de chômage)
 - Cartes économiques (exportations par wilaya)
- **Annotations** : commentaires d'experts de la DGPP
- **Export** : PDF interactif + Excel détaillé
- Tableau de bord de ressources humaines DRH/DFMI

Figure 11: Rapport trimestriel de Gouvernance des finances publiques -- DGND SIE

Indicateur RH	Valeur	Objectif	Variation	Observations
Effectif total MdF	14 230	-	-1.2% vs N-1	 Retraites
% d'agents formés en 2025	36%	50%	+4 pts	 Formations SAP prévues
Taux de rotation	3.8%	≤ 3%		Départs DGB

- **Contenu :**

- Synthèse de l'exécution budgétaire global.
- Analyse des écarts principaux.
- Etat des projets de Numérisation du SI.
- Indicateurs de performance des directions métiers
- Evaluation de la qualité de la sécurité de données.

-Export : Tableaux, graphiques et commentaires explicatifs.

Dans cette section, nous avons jeté les bases d'un système décisionnel modernisé pour le Ministère des Finances, en s'appuyant sur l'approche de la virtualisation des données. Nous avons débuté par une présentation détaillée du contexte organisationnel du Ministère des Finances, soulignant ses missions essentielles dans le développement économique et la gestion des finances publiques. Une analyse approfondie du système décisionnel existant a révélé ses caractéristiques, ses forces, notamment un socle de données riche et des capacités techniques internes, mais surtout ses faiblesses, telles que la fragmentation de l'information et l'absence d'un système décisionnel intégré et performant.

Face à ces défis, nous avons démontré pourquoi la virtualisation des données représente une solution révolutionnaire pour un système décisionnel sécurisé et performant, en mettant en évidence ses avantages clés comme l'adaptation à la confidentialité des données, la réponse à l'hétérogénéité des sources, l'agilité et la réduction des coûts par rapport aux data warehouses traditionnels. Cette approche permet de consolider les données financières, d'améliorer la prise de décision et de renforcer la transparence

Le cœur de ce chapitre a été consacré à la conception et à la modélisation de la solution de data virtualisation. Nous avons défini une architecture cible à quatre niveaux intégrant une couche de

virtualisation et élaboré des vues de données virtuelles pour les décideurs, à travers des vues de base, d'intégration et métiers. Un prototype fonctionnel a été détaillé, définissant son périmètre, les outils à sélectionner, et les étapes de mise en œuvre et d'évaluation, avec des résultats attendus prometteurs. Enfin, nous avons formulé les règles de transformation et d'intégration des données, et présenté des maquettes de tableaux de bord et de rapports décisionnels alignés avec les besoins stratégiques du ministère.

Dans la prochaine section, nous allons passer à la phase d'implémentation et de validation de cette solution proposée, en nous concentrant sur l'installation et la configuration de la plateforme de virtualisation, l'intégration concrète des sources de données, et le développement effectif des tableaux de bord avec des outils de BI.

3.6. Implémentation et validation de La solution proposée

Après avoir méticuleusement conçu l'architecture et les composants de notre solution de virtualisation de données dans le chapitre précédent, nous entrons à présent dans une phase cruciale, celle de l'implémentation et de la validation. Cette section marque le passage de la théorie à la pratique, où les concepts modélisés seront mis à l'épreuve dans un environnement réel. L'objectif principal est de démontrer la faisabilité technique et l'efficacité de la solution proposée pour le Ministère des Finances. Nous allons détailler les étapes concrètes d'installation et de configuration de la plateforme de virtualisation choisie, l'intégration des sources de données identifiées pour un cas d'usage initial, et la création des vues virtuelles nécessaires pour répondre aux besoins métiers spécifiques. Enfin, nous aborderons le développement des tableaux de bord et des rapports décisionnels, en assurant leur connexion avec la couche de virtualisation via des outils de Business Intelligence pertinents. Cette phase d'implémentation et de validation est essentielle pour confirmer les avantages de la virtualisation des données et préparer le terrain à une éventuelle généralisation de la solution au sein du Ministère des Finances.

Objectifs :

- Installation et configuration de la plateforme de virtualisation choisie dans un environnement de test
- Intégrer les sources de données identifiées (2 à 3 sources pour un premier cas d'usage).
- Création de vues virtuelles pour répondre à un besoin métier concret.

- Développer les tableaux de bord et les rapports décisionnels, par la Connexion de la couche de virtualisation à Power BI.

Remarque:

Compte tenu de la nature confidentielle des données du ministère, nous travaillerons avec des données génériques pour cette étape.

3.6.1. Installation et Configuration de la plateforme de virtualisation choisie dans un environnement de test

Notre approche méthodologique pour ce travail pratique repose sur l'intégration des solutions technologiques suivantes : **Denodo Express 9** (Plateforme de Virtualisation de Données Open Source), **PostgreSQL 16** (Système de Gestion de Bases de Données Relationnel Open Source), et **Power BI Desktop** (outil de Business Intelligence).

L'installation de la plateforme de virtualisation de données **Denodo Express 9** nécessite une installation d'un système de gestion de base de données (SGBD) et un outil d'analyse BI, et comme il existe une panoplie de SGBD sur le marché et plusieurs outils d'analyse, nous avons choisis d'installer le SGBD **PostgreSQL 16** et l'outil d'analyse **POWER BI**, considérant la simplicité de leur déploiement, l'installation des deux logiciels **PostgreSQL 16** et **Power BI** ne sera pas détaillée. Nous nous concentrerons plutôt sur les étapes d'installation et de configuration de la plateforme de virtualisation de données **Denodo Express 9**.

L'installation de la Plateforme Denodo Express 9 sera effectuée sur un environnement de test sous **Windows 10 (64 bits)**. Néanmoins, pour tirer pleinement parti de toutes ses capacités, une installation sur **Windows Server** est préconisée.

3.6.1.1. La plateforme Denodo

La Plateforme Denodo est une solution de gestion logique des données et de virtualisation des données. Elle permet aux organisations d'intégrer, de gérer et de délivrer des données en temps réel, quelle que soit leur source, leur format ou leur emplacement physique. Agissant comme une couche d'abstraction unique au-dessus de sources de données disparates (bases de données, applications, services cloud, Big Data, etc.), elle crée une vue unifiée et virtuelle des données sans nécessiter de réplique physique. Cela permet aux utilisateurs métier et aux applications d'accéder à des informations fiables et cohérentes pour l'analyse, le reporting,

l'intelligence artificielle (IA) et d'autres initiatives, tout en améliorant la gouvernance des données et en réduisant les coûts d'intégration traditionnels (denodo, n.d.).

3.6.1.2. Types de connexion de la plateforme Denodo

La plateforme Denodo offre plusieurs méthodes pour se connecter et interagir avec ses services, permettant une intégration flexible avec diverses et outils tiers (Denodo platform, 2025):

- **JDBC (Java Database Connectivity)** : C'est l'une des options privilégiées pour connecter des outils tiers à Denodo. Le Virtual DataPort distribue un pilote JDBC qui permet aux clients Denodo de se connecter aux serveurs Denodo via cette API.
- **ODBC (Open DataBase Connectivity)** : API par défaut pour les produits Microsoft lorsqu'il s'agit de travailler avec des magasins de données. Le Virtual DataPort fournit un pilote ODBC pour établir des connexions (comme la connexion à Power BI).
- **Arrow Flight SQL** : À partir de Denodo 9.1, les connexions via Arrow Flight SQL sont prises en charge, offrant une option supplémentaire pour l'accès aux données.
- **Services Web (REST et SOAP)** : Denodo facilite la consommation de données en permettant la publication de services Web REST et SOAP. Dans ce cas, les applications clientes se connectent au conteneur Web qui héberge ces services en utilisant les protocoles HTTP ou HTTPS, et les services Web se connectent ensuite au serveur Virtual DataPort.
- **RMI (Remote Method Invocation)** : Principalement utilisé pour la communication interne entre les composants du serveur Denodo et pour la connexion des clients Denodo (comme les outils d'administration) au serveur Virtual DataPort.

3.6.1.3. Installation et Configuration de la plateforme Denodo 9 Express sur Windows

-Avant d'installer Denodo, il est important d'avoir un JDK (Java Development Kit) avec une version récente déjà installé (par exemple : Java SE Development Kit 24.0.1)

-Pour télécharger Denodo 9 Express, qui est la version gratuite de la plateforme Denodo, nous devons nous rendre sur le site officiel de Denodo.

-Voici le lien direct pour accéder à la page de Denodo Express et démarrer le téléchargement : <https://www.denodo.com/fr/denodo-platform/denodo-express>

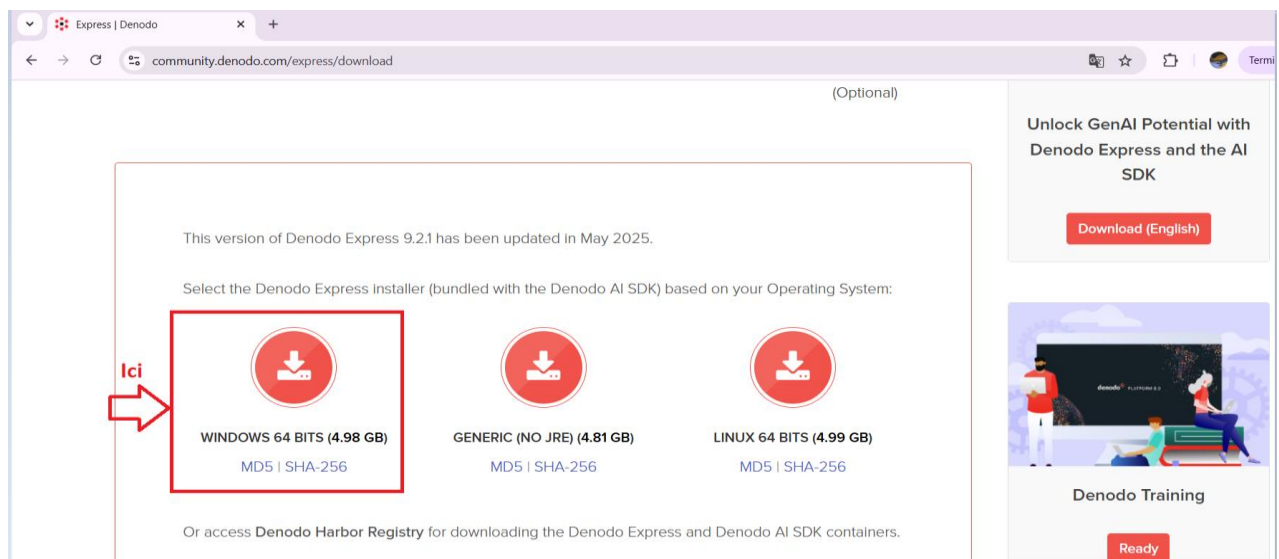
-Il est necessaire de créer un compte utilisateur avec un mot de passe sur cette page pour pouvoir télécharger cette version de denodo, comme illustré dans la figure ci-dessous :

Figure 12: Authentification pour accéder au site officiel Denodo



-L'accès au site nous permettra de procéder au **téléchargement gratuit** de cette version, comme le montre la figure ci-dessous

Figure 13: Lancement du téléchargement sur site.

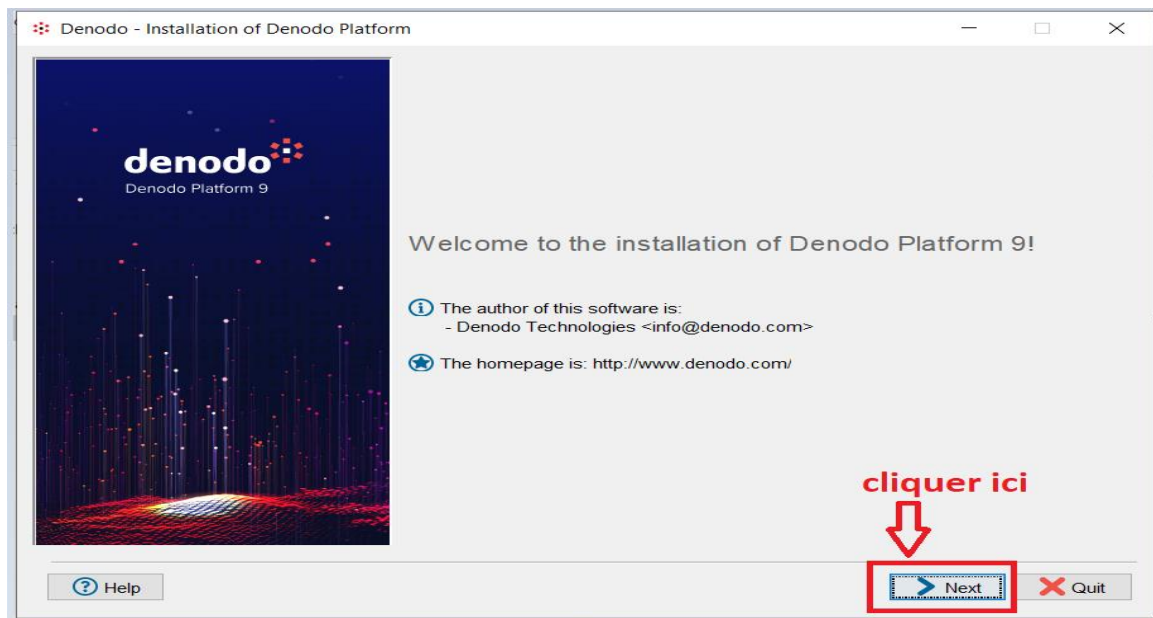


-Une fois le téléchargement lancé, nous obtiendrons un **dossier compressé nommé denodo-express-install-9-win64.zip** ;

-Pour commencer l'installation, nous devons d'abord **décompresser ce dossier**,

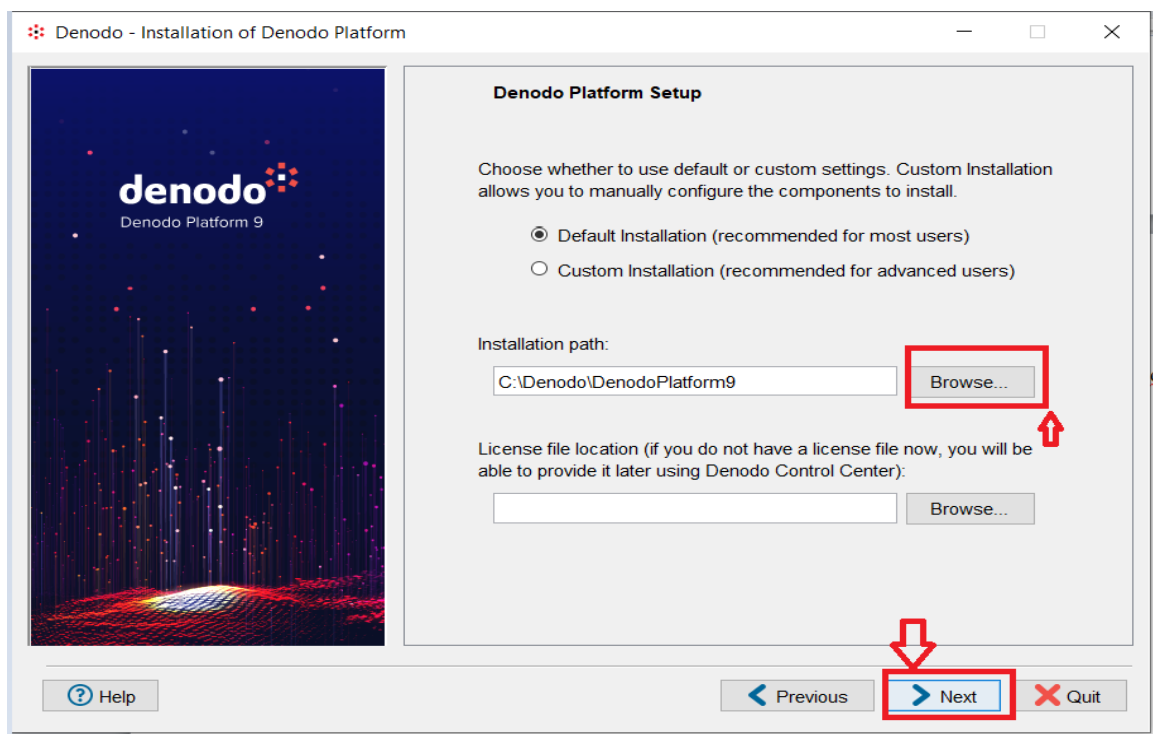
-Ensuite, **accédez au dossier décompressé et lancez l'application install.exe**. La fenêtre d'installation s'affichera alors :

Figure 14: Fenêtre d'installation de la plateforme Denodo.



-Et puis, choisissez l'emplacement d'installation (Répertoire) et cliquez sur Next:

Figure 15: Choix du répertoire d'installation.



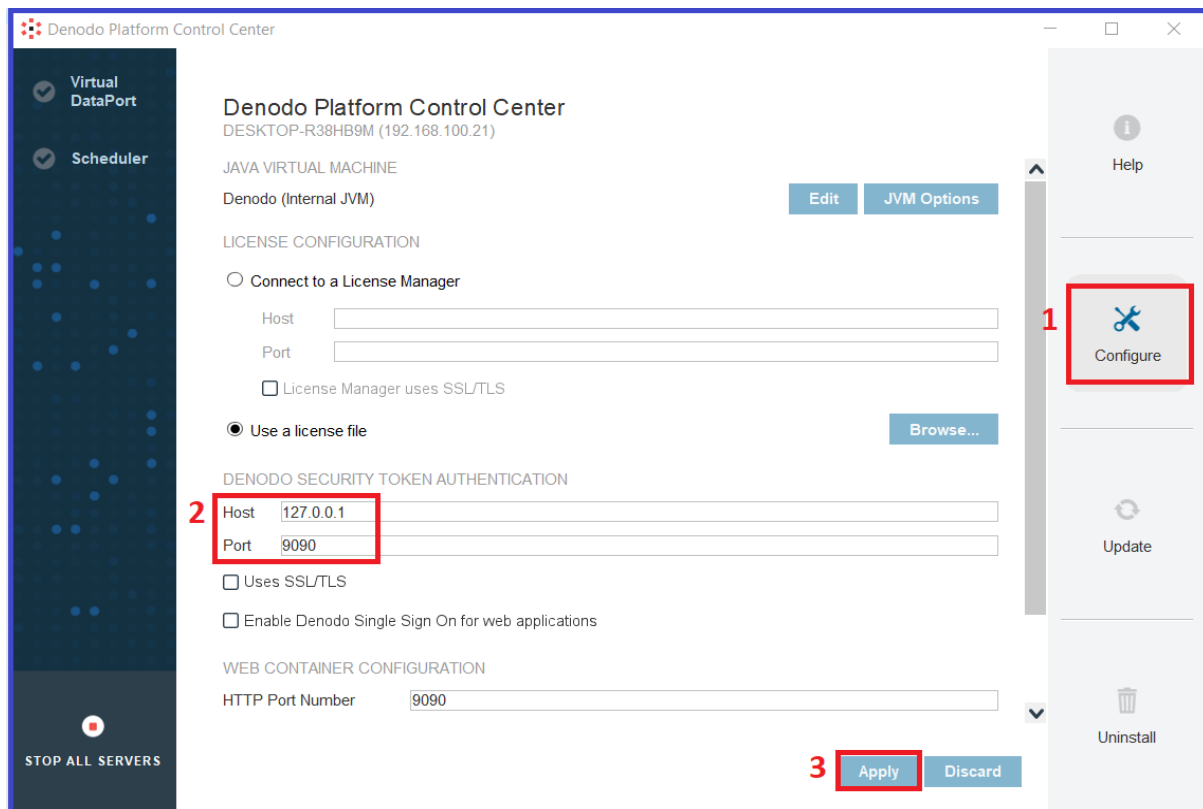
- L'installation de la plateforme peut prendre **entre 10 et 15 minutes**.
- Une fois l'installation terminée, une icône "**Denodo Platform 9**" apparaîtra sur votre bureau.

Figure 16: L' icône Denodo Platform 9.



- Une fois l'installation de la plateforme denodo express 9 terminé avec succès, nous devons passer à la **configuration de la solution**.
- En cliquant sur l'icône "**Denodo Platform 9**" sur le bureau, nous accéderons au **Denodo Platform Control Center**. Ce centre de contrôle est l'interface principale qui nous permet de démarrer et d'arrêter tous les serveurs et outils de la plateforme Denodo, et il offre également de nombreuses fonctions supplémentaires.
- Cliquez ensuite sur **Configuration**, puis saisissez l'adresse de l'hôte : **127.0.0.1** et le port **9090**. Pour valider les modifications, cliquez sur **Apply**.

Figure 17: Configuration de Denodo Platform Control Center.



-Ensuite, nous devons **démarrer le serveur Virtual DataPort (VDP)**. C'est le **cœur de la plateforme Denodo**, agissant comme le **moteur principal** qui la fait fonctionner. Concrètement, il est responsable de :

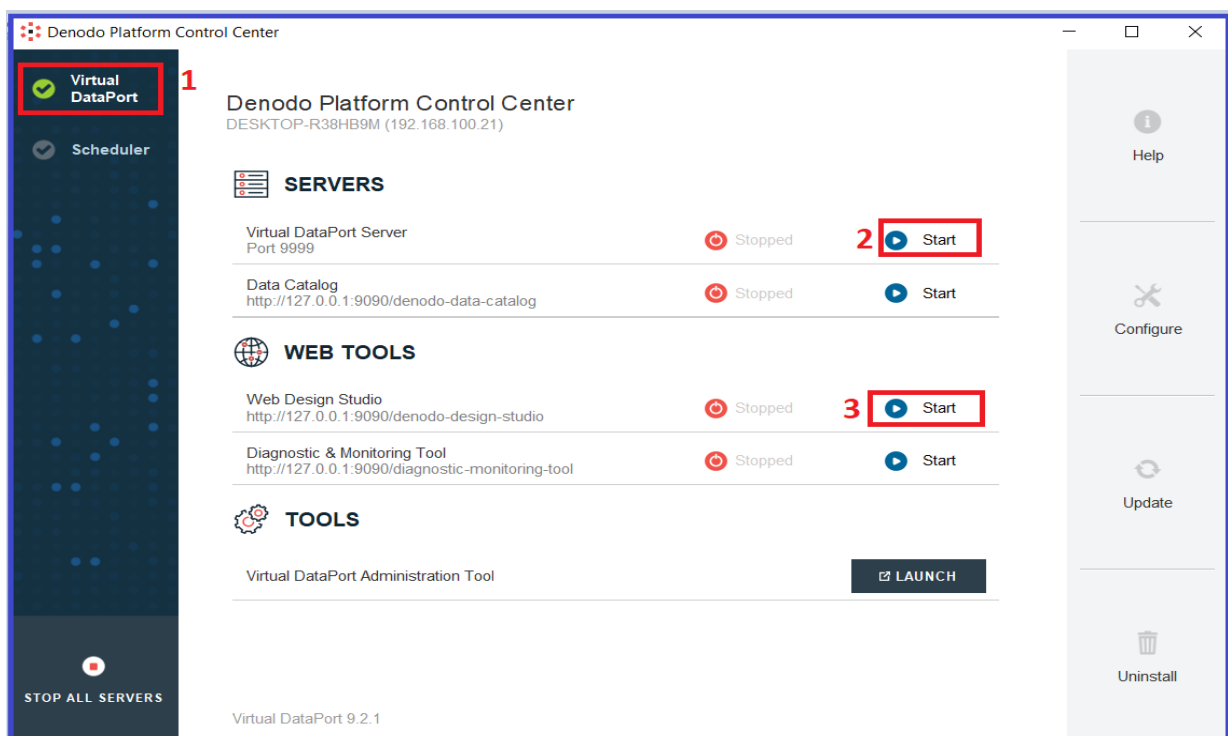
- **Intégration de données en temps réel** : Le VDP Server est capable de se connecter à une multitude de sources de données différentes, qu'il s'agisse de bases de données relationnelles (Oracle, SQL Server, etc.), de services web, de fichiers plats (CSV, Excel), de documents XML, de sources NoSQL (MongoDB), etc. Il ne déplace pas les données, mais il les "virtualise", c'est-à-dire qu'il crée une **couche d'abstraction logique**.
- **Création de vues virtuelles** : Il permet de créer des "vues virtuelles" qui sont en fait des agrégations ou des combinaisons de données provenant de différentes sources. Ces vues ne stockent pas les données, elles définissent simplement comment les données doivent être accédées et transformées à partir des sources originales.
- **Optimisation des requêtes** : Lorsqu'une application ou un utilisateur interroge une vue virtuelle, le VDP Server optimise la requête pour la traduire en requêtes spécifiques à chaque source de données sous-jacente. Il gère l'exécution de ces

requêtes, combine les résultats et les renvoie à l'utilisateur comme si les données provenaient d'une seule source.

- **Publication de services de données** : Le VDP Server peut exposer ces vues virtuelles sous forme de services de données standard (comme JDBC, ODBC, REST ou SOAP API), ce qui permet à n'importe quelle application (outils BI, applications métier, etc.) d'accéder facilement aux données sans avoir à connaître la complexité des sources originales.
- **Sécurité et gouvernance des données** : Il centralise la sécurité et la gouvernance des données. Nous définissons les droits d'accès aux vues virtuelles dans le VDP Server, ce qui garantit que les utilisateurs n'accèdent qu'aux données pour lesquelles ils sont autorisés, et ce, quelle que soit la source physique des données.

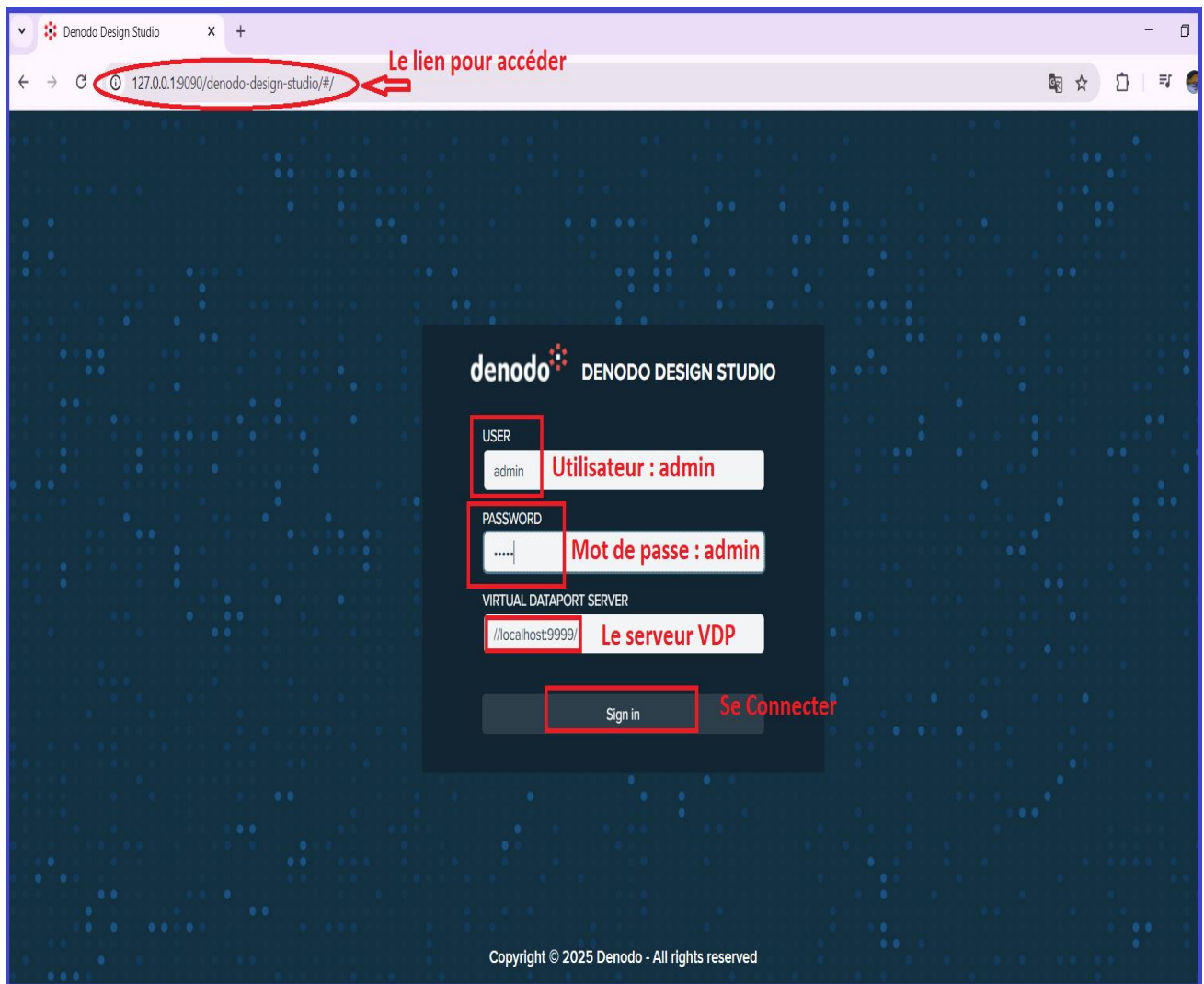
-Démarrerez maintenant le service web design studio, **l'interface web** de la plateforme **Denodo** qui permet de concevoir, configurer et gérer des **vues virtuelles** (Virtual Views) à partir de différentes sources de données. C'est une **interface graphique conviviale**,

Figure 18: Démarrage du serveur VDP et le Web Design Studio.



Afin d'accéder à la page web de la plateforme, veuillez saisir le lien suivant dans votre navigateur (tel que Chrome ou Mozilla Firefox) : <http://127.0.0.1:9090/denodo-design-studio/#/> . Ceci est illustré dans la figure ci-dessous:

Figure 19: Page de connexion Denodo Design Studio.



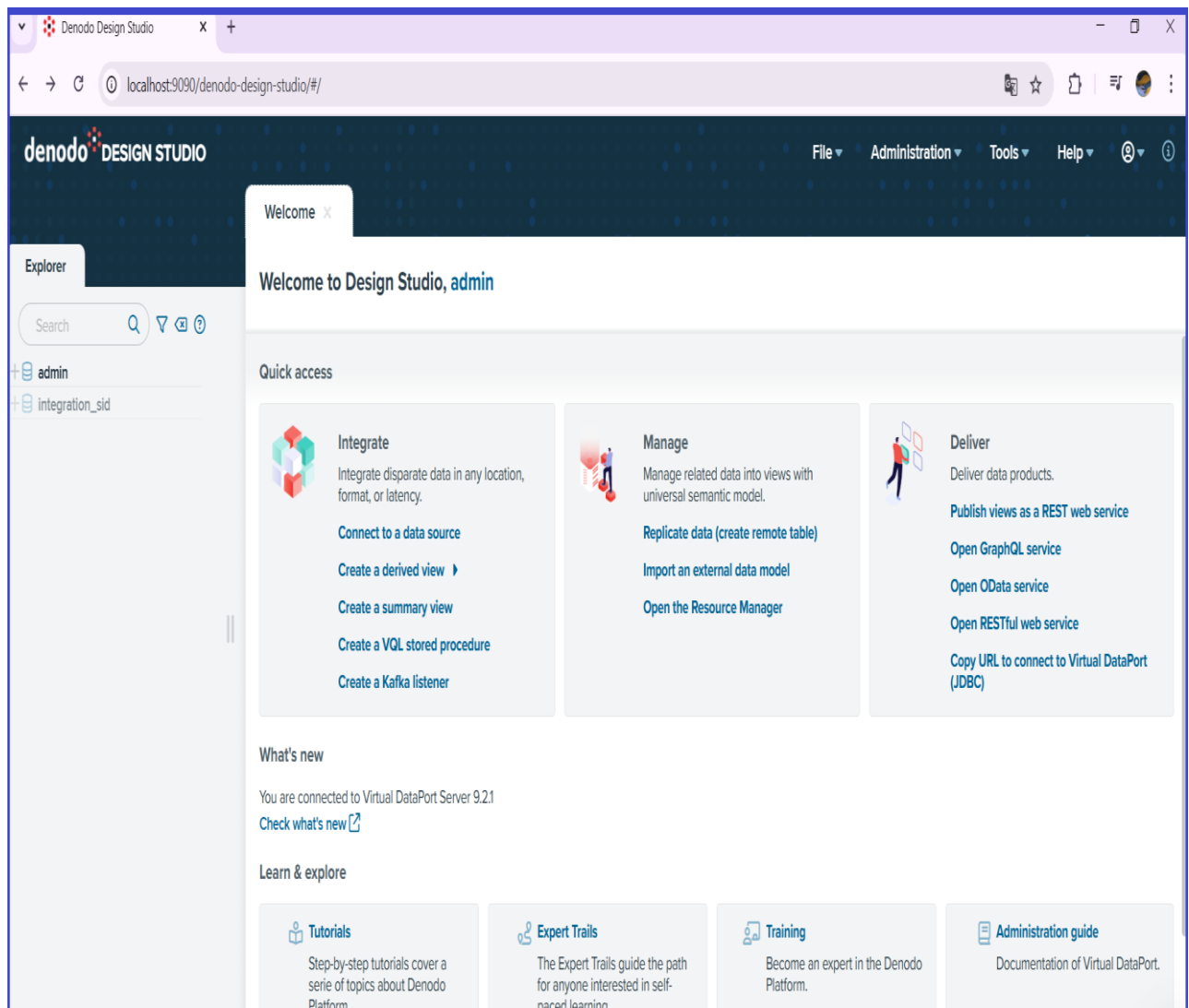
-Denodo Design Studio est l'interface utilisateur graphique principale de la plateforme Denodo. C'est un outil essentiel pour les développeurs et les administrateurs.

En résumé, il permet de :

- **Se connecter à diverses sources de données** (bases de données, applications, fichiers, etc.).
- **Virtualiser ces données** en créant des vues logiques (abstraites des sources physiques).
- **Combiner et transformer** ces données virtualisées pour créer de nouvelles vues plus complexes, adaptées aux besoins métier.

- **Publier ces données virtualisées** sous forme de services (REST, GraphQL, OData, etc.) pour qu'elles puissent être consommées facilement par d'autres applications ou outils d'analyse.
- **Gérer l'ensemble de l'environnement** de virtualisation de données, y compris les utilisateurs, les permissions, la sécurité et la surveillance.

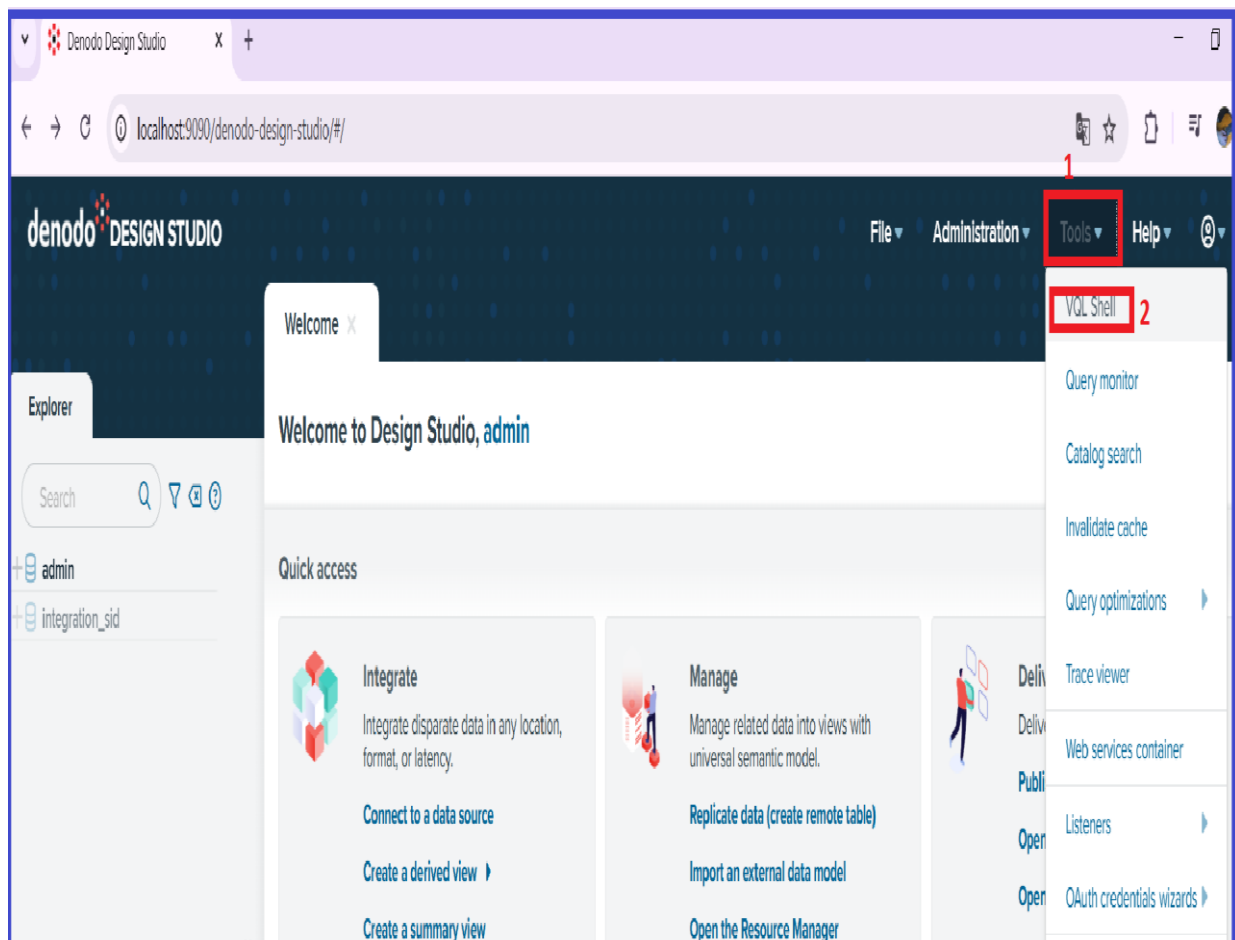
Figure 20: Page d'accueil Denodo Design Studio.



-Lors du premier accès à Denodo Design Studio, nous remarquerons qu'il contient par défaut une seule base de données virtuelle nommée 'admin'. Pour vos propres travaux et pour exploiter pleinement la plateforme, nous devrons créer une nouvelle base de données virtuelle. C'est dans cette nouvelle BDD que nous intégrerons vos sources de données et créerez les vues virtuelles nécessaires à la prise de décision.

-Pour créer cette nouvelle base de données, cliquez simplement sur 'Tools', puis sélectionnez 'VQL SHELL', comme illustré dans la figure suivante :

Figure 21: Lancement de VQL Shell.



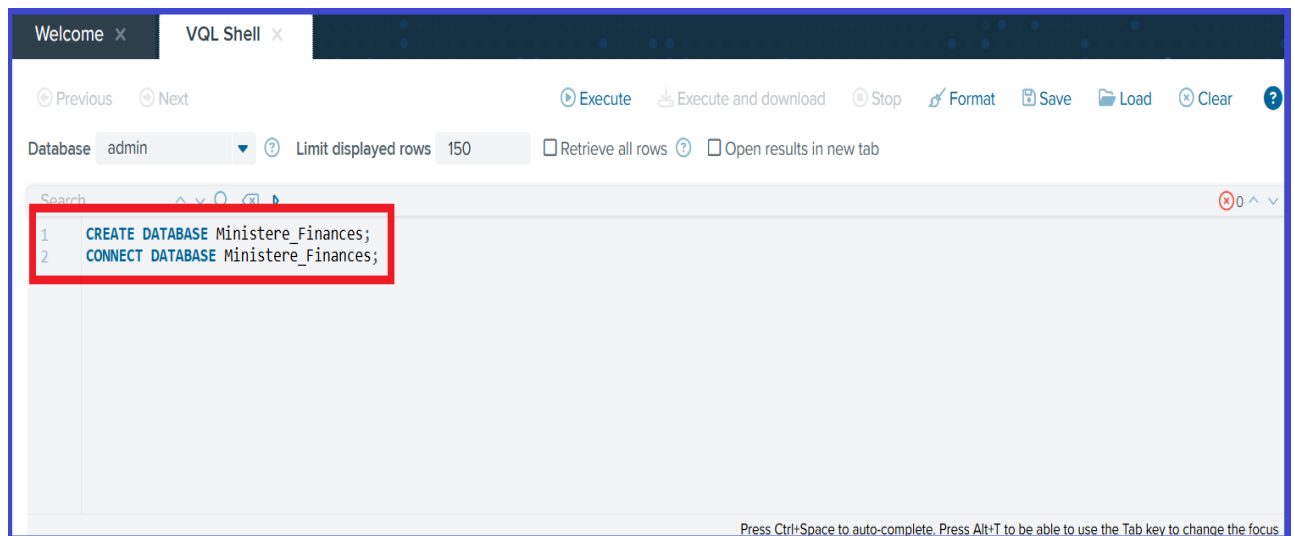
Ensuite, dans la zone SQL, saisissez la commande : `CREATE DATABASE [nom_de_la_base];`

Par exemple :

```
CREATE DATABASE Ministere_Finances;
```

```
CONNECT DATABASE Ministere_Finances;
```

Figure 22: Requette SQL pour la creation d'une BDD Virtuelle



3.6.2. Intégration des sources de données dans la Plateforme

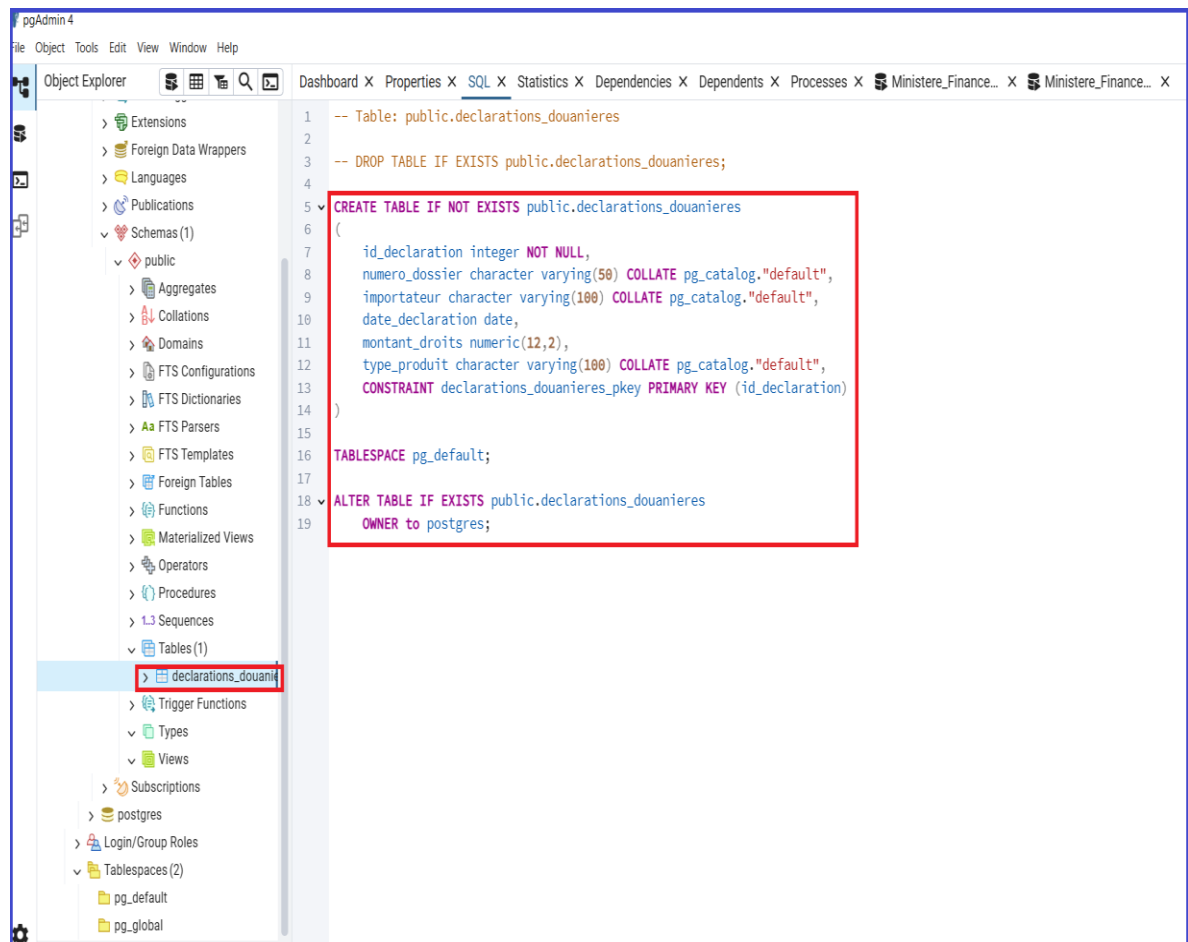
La méthodologie adoptée pour la simulation de l'intégration de sources et la création de vues virtuelles s'articulera autour de deux scénarios. Le premier scénario traitera de l'intégration de sources de données hétérogènes, incluant des bases de données SQL, des fichiers plats CSV et des fichiers au format JSON. Le second scénario se concentrera sur l'utilisation d'une base de données SQL générique, mise à disposition par le Ministère des Finances, pour des besoins de reporting et de visualisation via l'outil Power BI.

➤ 1^{er} Scénario :

Nous allons commencer par la création de la BDD SQL et les fichiers

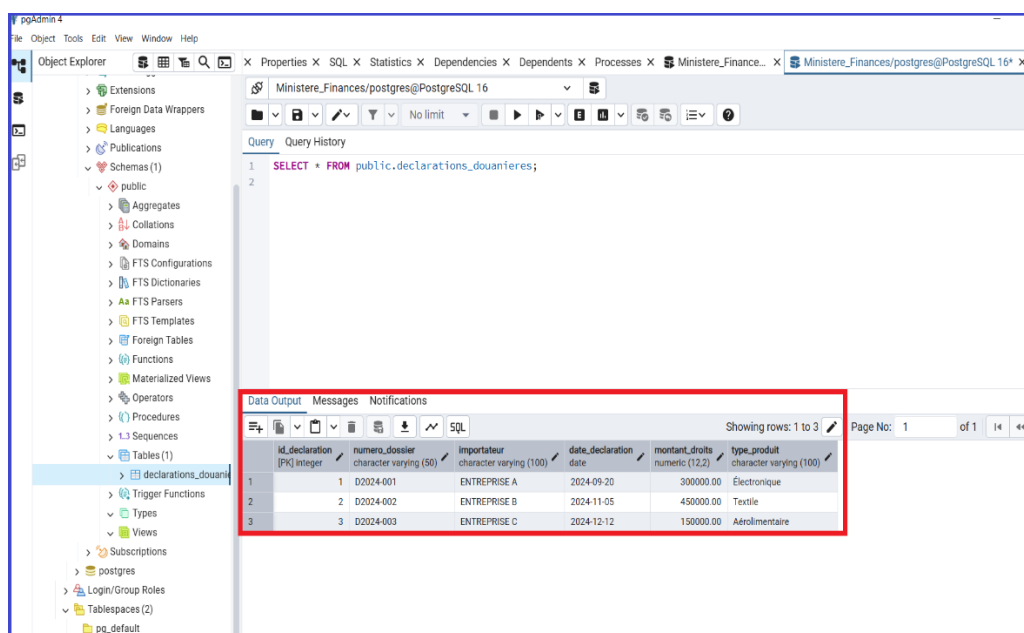
- Sous SQL ; create Database Ministere_Finances ; pour créer une Nouvelle BDD.
- Et puis, la création de la table déclarations_douanier comme illustré dans la figure suivante :

Figure 23: Création d'une table sous PostgreSQL 16.



- Nous Remplissons par la suite cette table avec les enregistrements suivantes :

Figure 24: Données de la Table declarations_douaniers.



- Nous créons des fichiers Plats Imprtateur.csv, engagements_budgétaires.csv, loi_finances.csv, paiements_effectues.csv et Structures_Administratives.csv avec l'outil bloc-notes comme suite (résumés dans une seule figure) :

Figure 25: Création des fichiers plats de type .csv

Importateurs - Bloc-notes Fichier Edition Format Affichage Aide hif,nom,secteur,pays 4001234567,ENTREPRISE A,Électronique,Chine 4009876543,ENTREPRISE B,Textile,Turquie 4005432198,ENTREPRISE C,Aéroalimentaire,France	engagements_budgétaires - Bloc-notes Fichier Edition Format Affichage Aide id_engagement,id_structure,code_programme,mois,montant_engage ENG001,STR001,PRG01,2025-01,3000000 ENG002,STR001,PRG01,2025-02,5000000	Structures_Administratives - Bloc-notes Fichier Edition Format Affichage Aide id_structure,nom_structure,type,wilaya STR001,Direction Générale de Budget,centrale,Alger STR002,Trésorerie de Wilaya d'Oran,locale,Oran
loi_finances - Bloc-notes Fichier Edition Format Affichage Aide exercice,code_programme,intitule_programme,nature_depense,montant_accord 2025,PRG01,Équipement des écoles,capital,30000000 2025,PRG02,Informatique Trésorerie,fonctionnement,10000000	paiements_effectues - Bloc-notes Fichier Edition Format Affichage Aide id_paiement,id_engagement,date_paiement,montant_paye PAY001,ENG001,2025-02-15,2500000 PAY002,ENG002,2025-03-10,4800000	

- De la meme façon, nous créons des sources API.Json avec l'outil bloc-notes comme suite (résumés dans cette figure) :

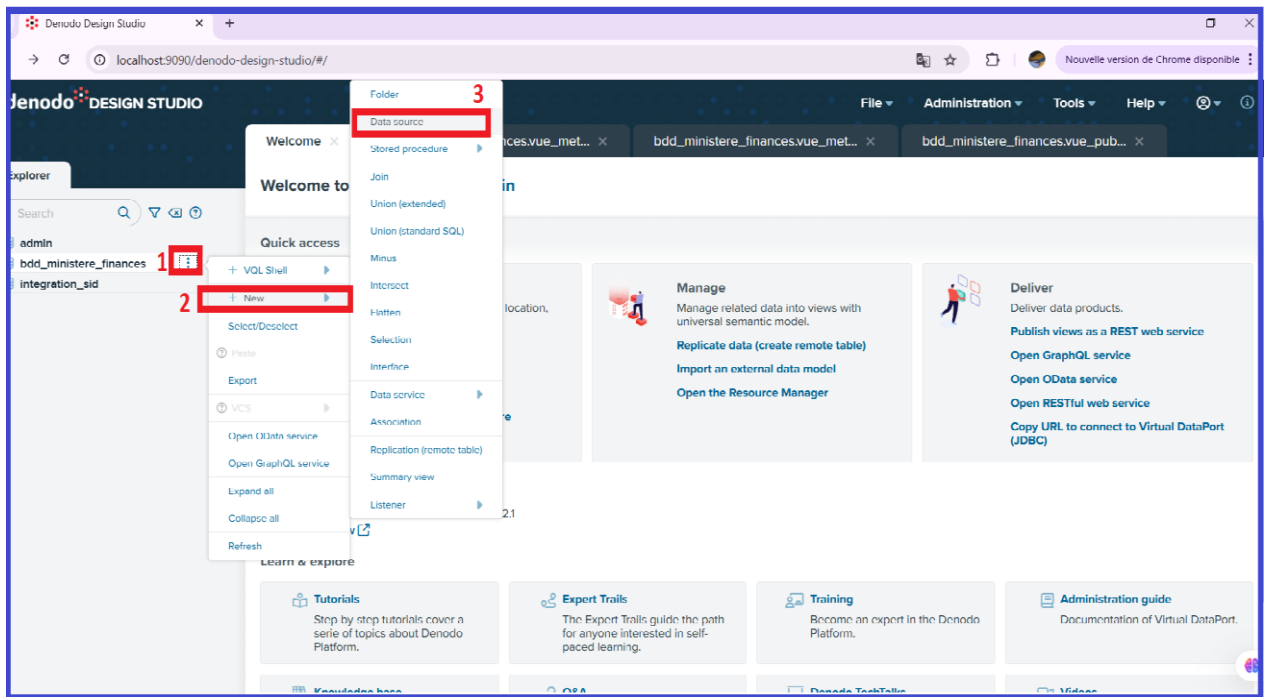
Figure 26: Création des sources de type .Json

taux_de_change - Bloc-notes Fichier Edition Format Affichage Aide <pre>{ "date": "2025-06-10", "usd_dzd": 138.45, "eur_dzd": 149.22, "baril_brutle_dollars": 78.3 }</pre>	indicateurs_performance - Bloc-notes Fichier Edition Format Affichage Aide <pre>[{ "id_indicateur": "IND001", "code_programme": "PRG01", "indicateur": "Taux de réalisation physique", "valeur": 75, "unite": "%" }, { "id_indicateur": "IND002", "code_programme": "PRG02", "indicateur": "Délai moyen de paiement", "valeur": 21, "unite": "jours" }]</pre>	projets_investissement - Bloc-notes Fichier Edition Format Affichage Aide <pre>[{ "id_projet": "PRJ001", "intitule_projet": "Construction Lycée Bouira", "wilaya": "Bouira", "code_programme": "PRG01", "budget_prevu": 5000000, "taux_avancem": 60 }, { "id_projet": "PRJ002", "intitule_projet": "Système GFMIS Trésor", "wilaya": "Alger", "code_programme": "PRG02", "budget_prevu": 10000000, "taux_avancem": 40 }]</pre>
ajustements_budgetaires - Bloc-notes Fichier Edition Format Affichage Aide <pre>[{ "id_ajustement": "AJ001", "code_programme": "PRG01", "type_ajustement": "+", "montant": 2000000, "date_ajustement": "2025-04-01" }, { "id_ajustement": "AJ002", "code_programme": "PRG02", "type_ajustement": "-", "montant": 1500000, "date_ajustement": "2025-05-10" }]</pre>		

Après la création de différents types de sources, nous passons maintenant à l'intégration de ces sources sur la plateforme denodo, pour ce faire nous devons suivre les étapes suivantes :

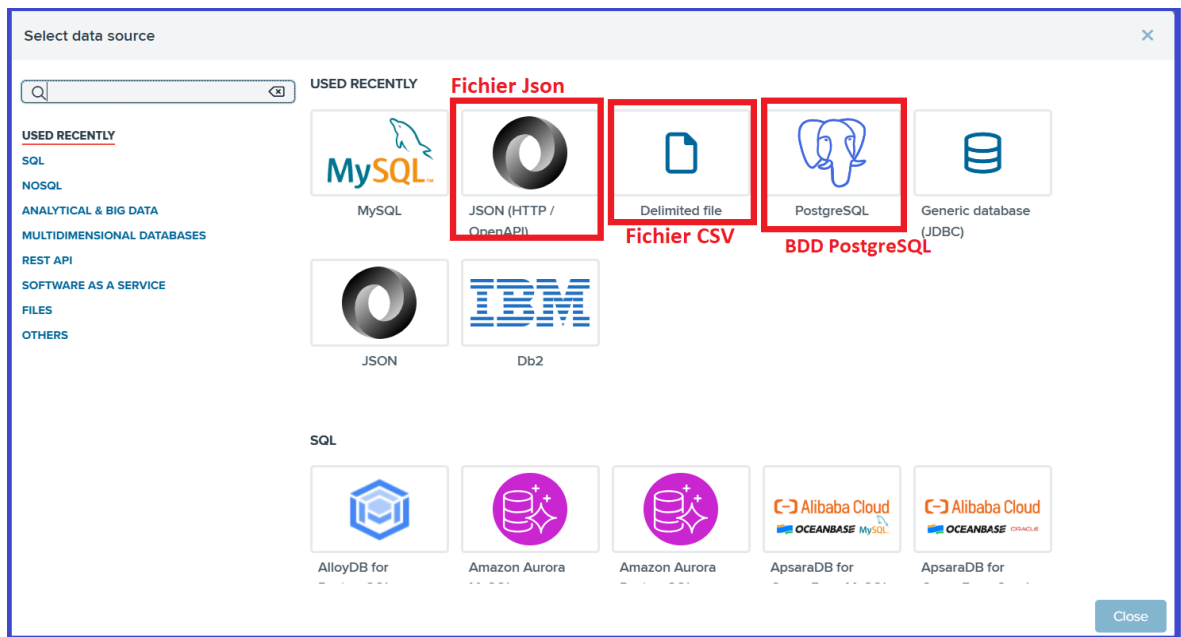
- Pour intégrer une nouvelle source, avec un simple clic droit sur le nom de la BDD virtuelle, nous cliquons sur « New », et puis « Data source », comme illustré dans cette figure :

Figure 27: Intégration d'une nouvelle source de données.



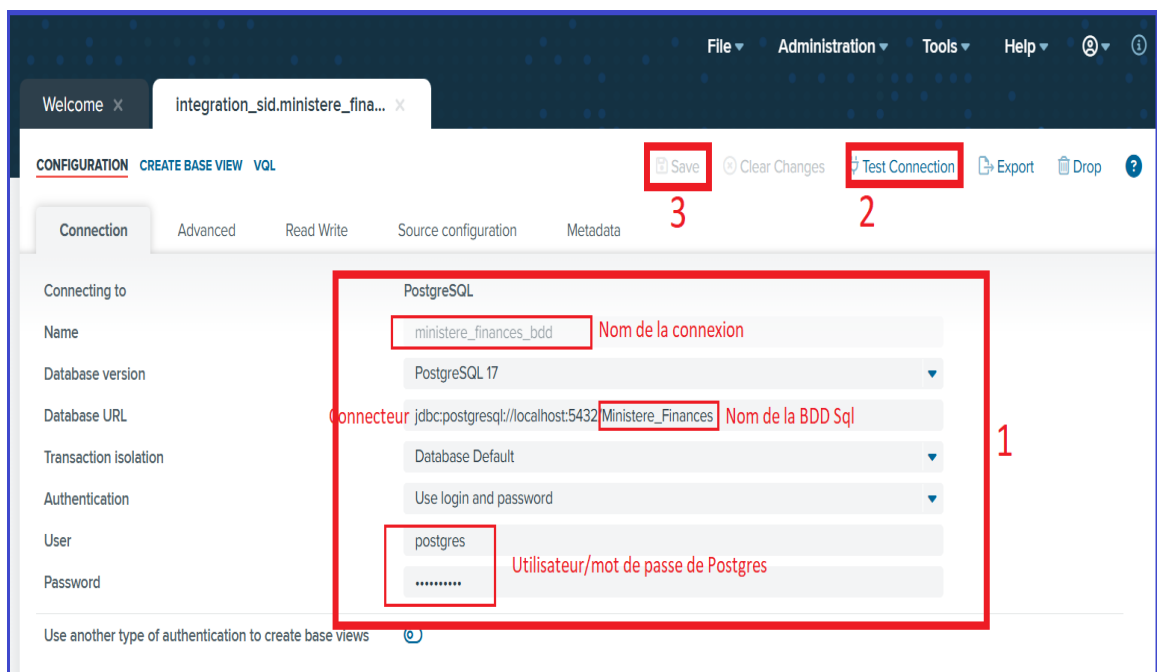
- La fenêtre suivante va être affichée, la méthode est la même pour toutes les sources, il suffit juste de choisir le type de la source de données, nous pouvons intégrer des centaines de types mais dans notre exemple nous allons choisir une BDD Sql, des fichiers CSV et des sources Json :

Figure 28: Interface pour le choix de type de la source.



- Maintenant, nous passons à l'intégration de la BDD PostgreSQL (la table déclarations_douaniers), pour cela nous choisissons une source PostgreSQL (comme illustré dans la figure 26 précédente)

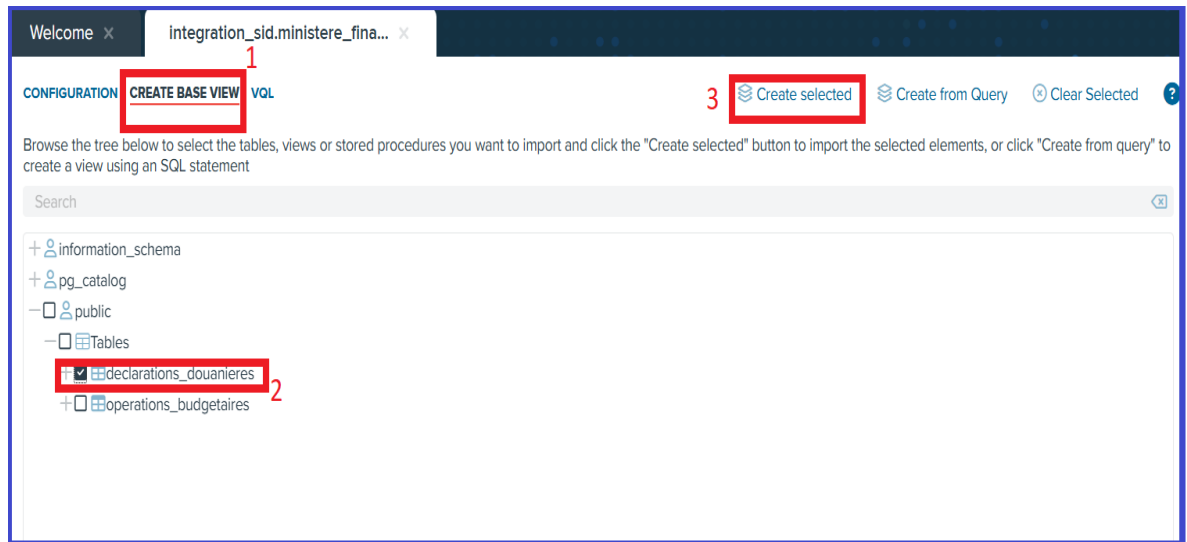
Figure 29: Connexion à une BDD Sql.



- Par la suite, il est important de tester la connexion avec cette BDD, une fois le test Ok nous pouvons sauvegarder cette connexion.

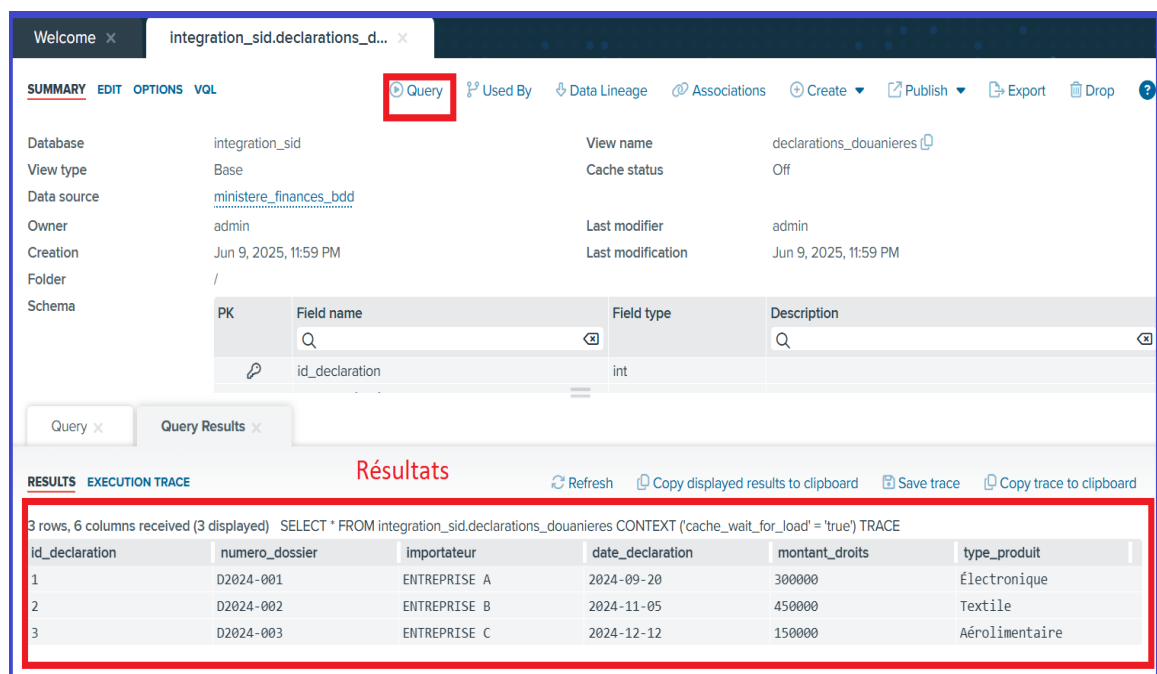
- Après la sauvegarde, nous devons créer une vue virtuelle de cette connexion (vue de base) comme suite :

Figure 30: Création de la vue virtuelle declaration_douaniers.



- Pour chaque intégration d'une source de données, nous devons créer une vue virtuelle de base correspondante à cette source, pour pouvoir créer des autres types de vue virtuelles (agrégation, jointure, métiers, précalculs...etc)

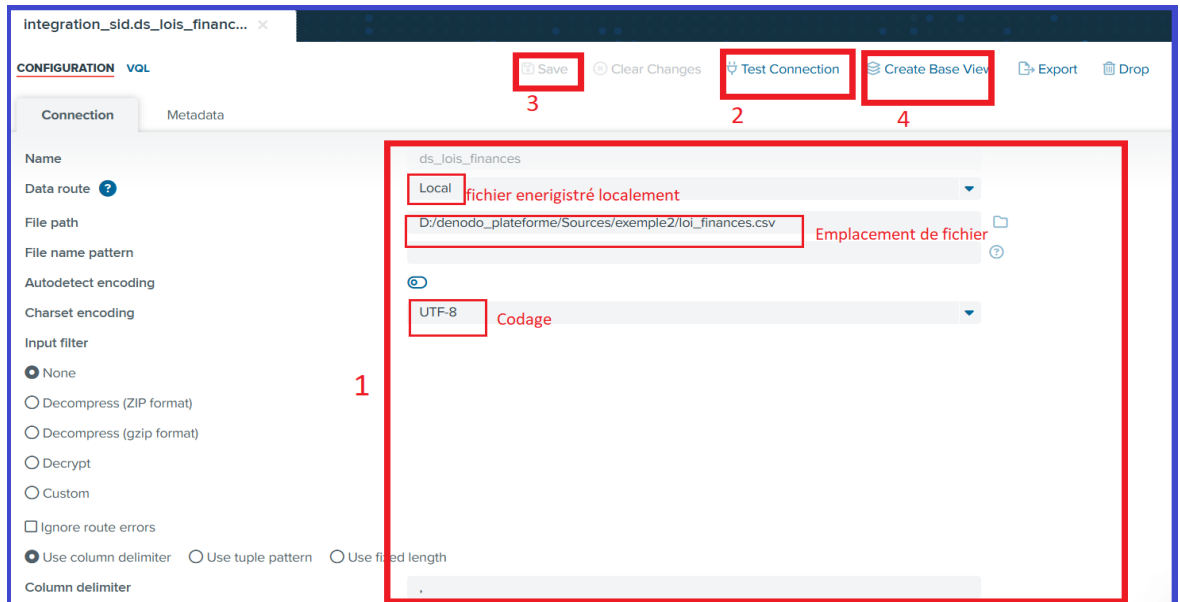
Figure 31: Résultats de l'exécution de la vue virtuelle declarations_douaniers.



Nous remarquons que les résultats sont les mêmes (avec la figure 22)

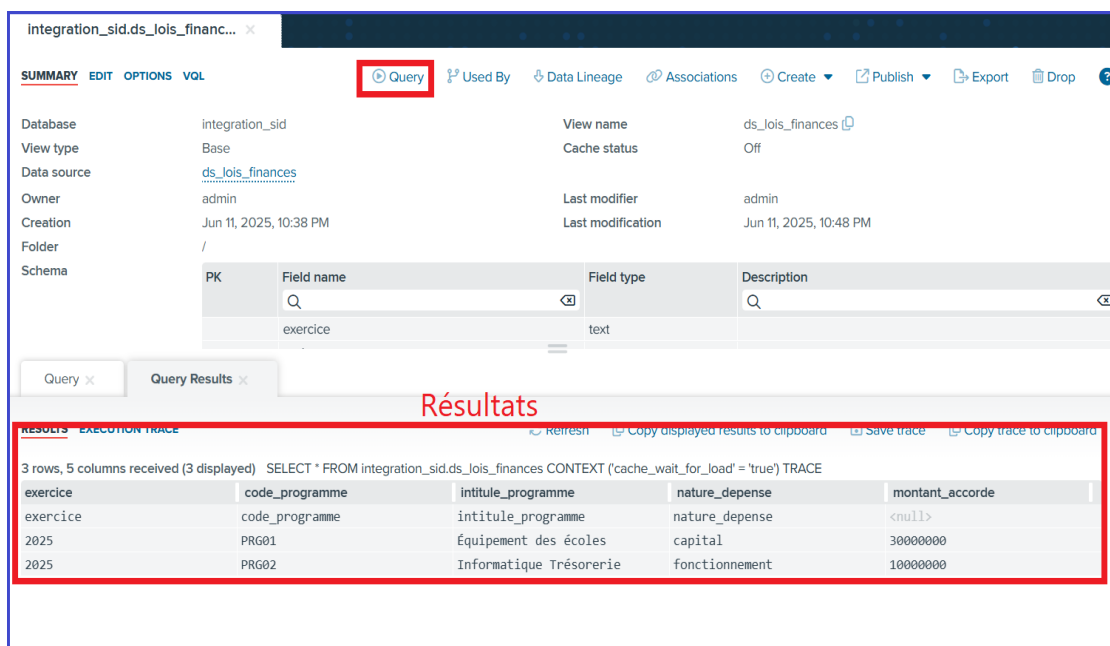
- Nous passons maintenant à l'intégration des fichiers plats (CSV), pour cela nous choisissons une source « Delimited File » (comme illustré dans la figure 26), par exemple le fichiers loi_finances :

Figure 32: Intégration du fichier plat loi_finances.csv.



-Nous créons la vue virtuelle correspondante à cette connexion, après l'exécution de cette vue nous obtenons les résultats suivants :

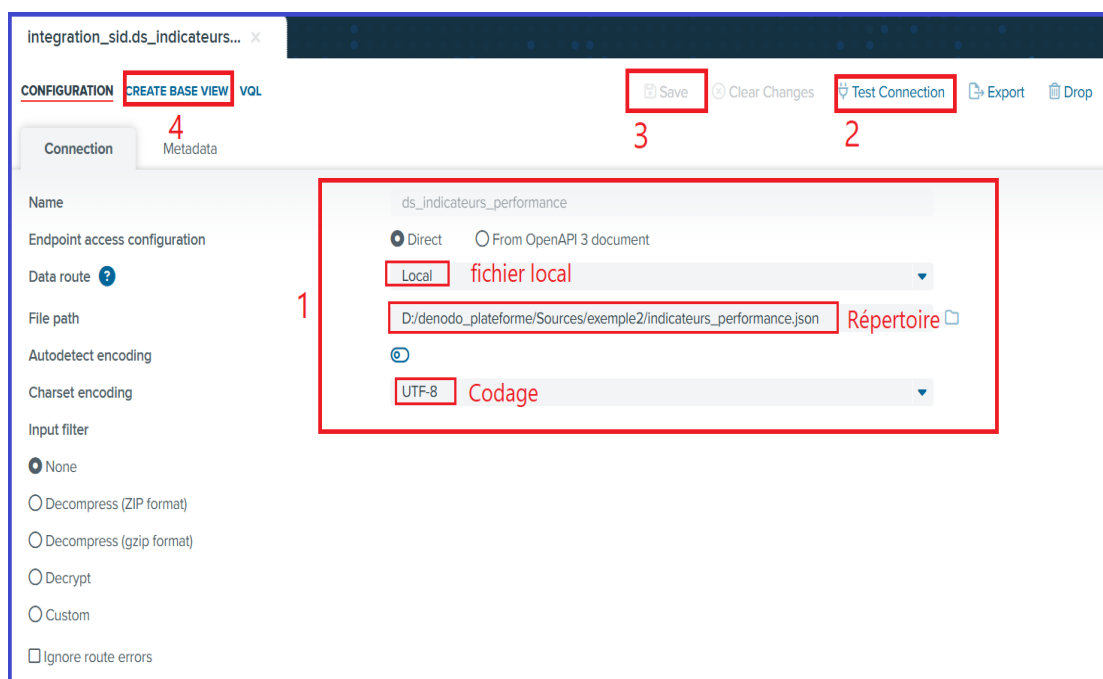
Figure 33: Exécution de la vue virtuelle loi_finances.



-De même façon, nous intégrons les autres fichiers plats sur denodo

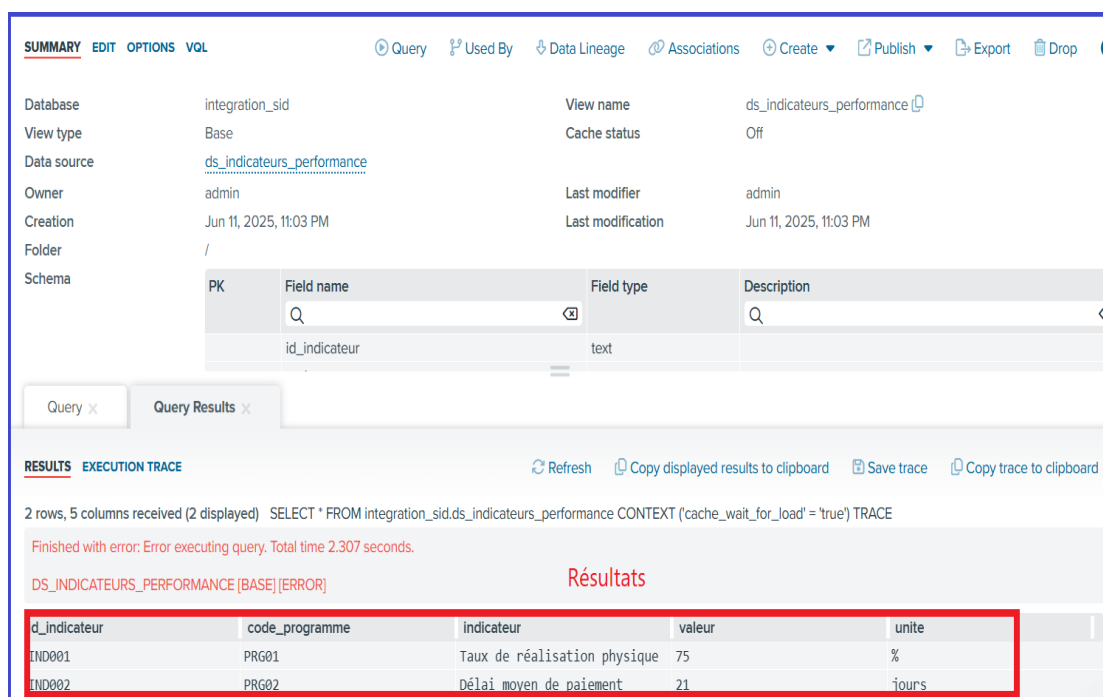
-Nous passons maintenant à l'intégration des sources API Json (Indicateurs_performance.json) de la même manière, en choisissant JSON une source (HTTP/Open API) dans la fenêtre illustrée dans la figure.26

Figure 34 : Intégration d'une source API Json



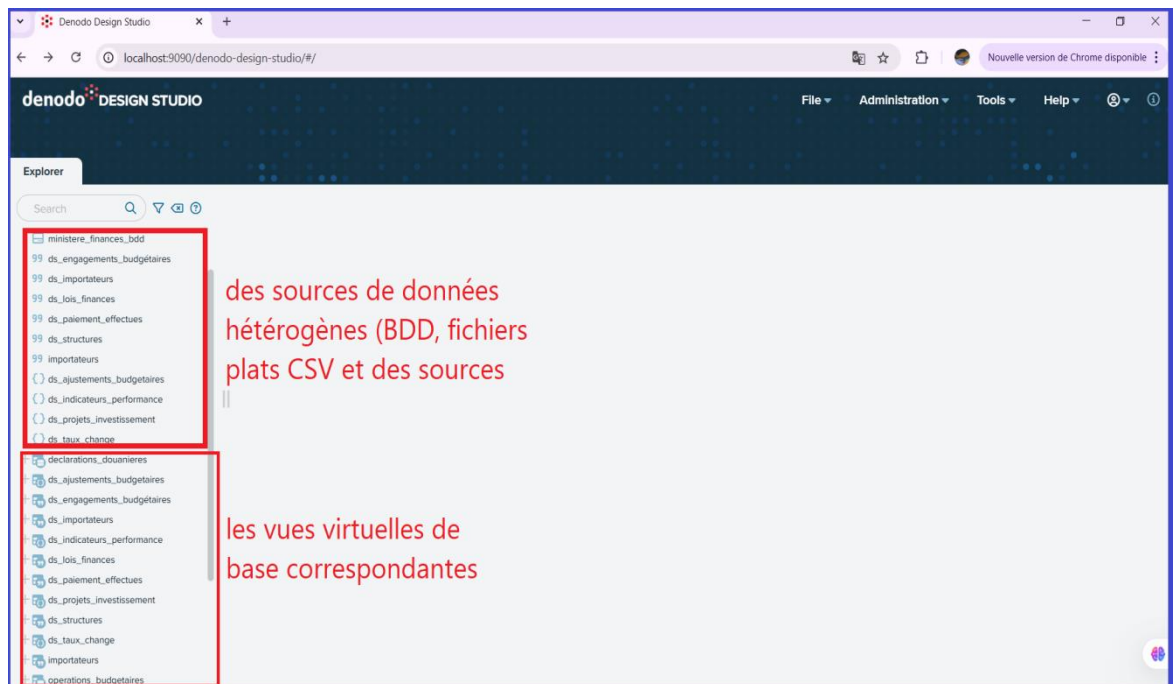
-Nous créons la vue virtuelle correspondante à cette connexion, après l'exécution de cette vue nous obtenons les résultats suivants :

Figure 35: Résultats de l'exécution de la vue virtuelle indicateurs_performance



-De même façon, nous intégrons les autres sources Json sur denodo, nous obtenons alors les différentes sources et les vues virtuelles correspondantes sur une seule base de données virtuelle :

Figure 36: Sources de données hétérogènes sur une seule Base de données virtuelle.



Après l'agrégation des données issues des différentes sources, nous procédons à la conception de vues virtuelles, offrant ainsi une base analytique optimisée pour appuyer le processus décisionnel.

3.6.3. Création de quelques vues d'agrégation

-Vue Virtuelle Déclarations Douaniers Détaillées

Nom : vue_declarations_enrichies

Jointure : entre la source declarations_douaniers (Sql) et la source Importateurs (fichier CSV)

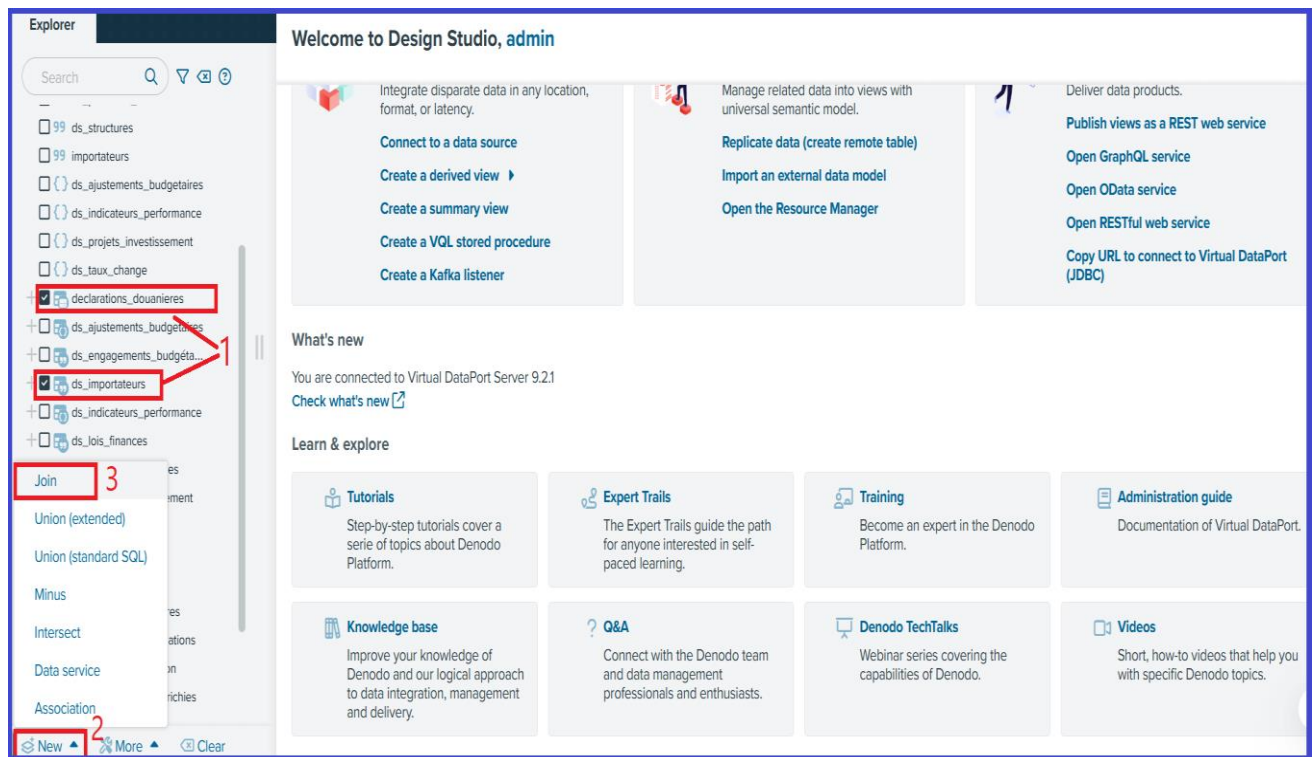
Condition :

ds_importateurs.nom = declarations_douanieres.importateur

Objectif :

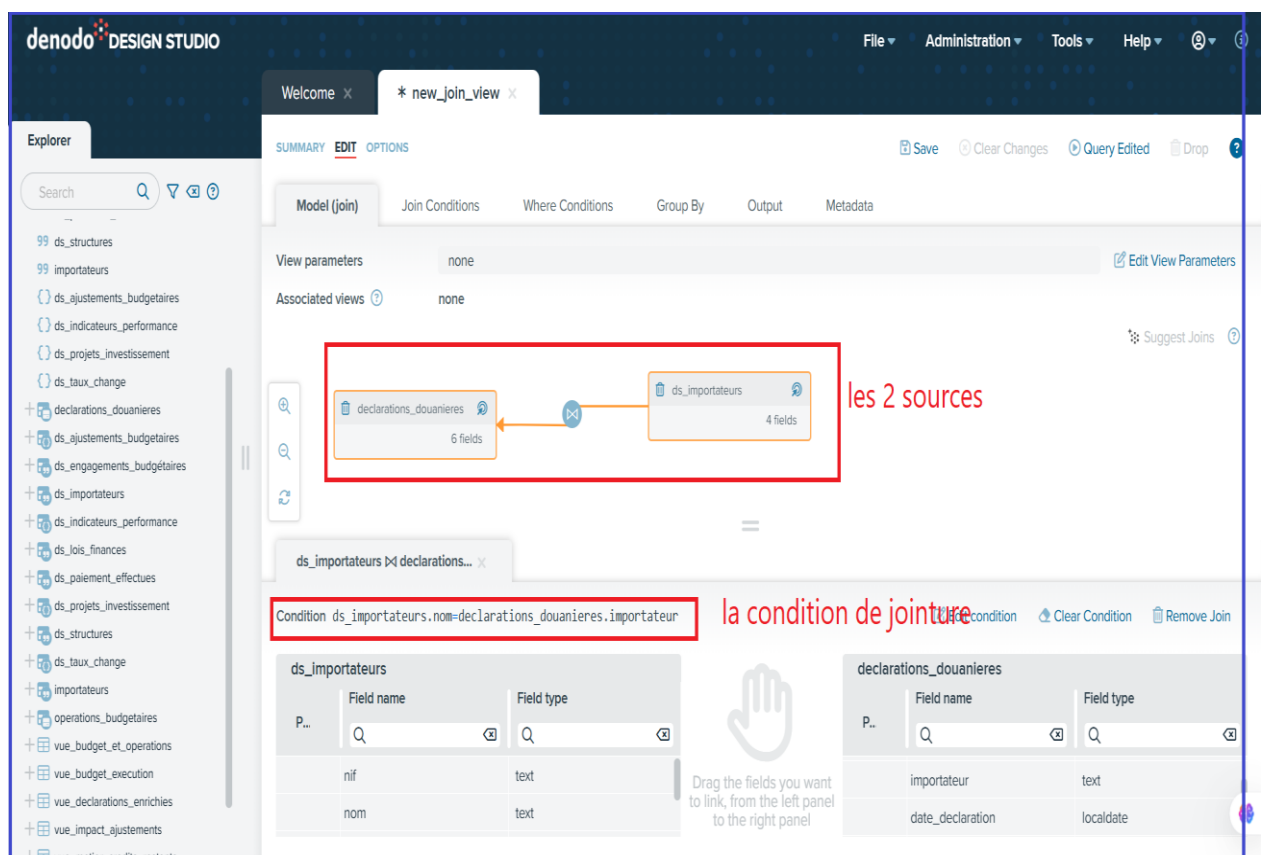
Avoir le détail de l'importateurs pour chaque déclaration douanière. Pour créer cette jointure, nous sélectionnons les deux vues de base, et puis nous cliquons sur "New" et par la suite sur "Join" comme suite (Figure 34):

Figure 37: Création de la jointure pour la vue declarations_enrichies.



Maintenant nous sélectionnons la condition de la jointure comme suite :

Figure 38: Condition de jointure pour la vue virtuelle declarations_enrichies.



Après la sauvegarde de cette nouvelle vue virtuelle, nous exécutons cette dernière et nous obtenons les résultats suivants:

Figure 39: Résultats de l'exécution de la vue virtuelle declarations_enrichies.

Summary Metadata:

Database	integration_sid	View name	vue_declarations_enrichies
View type	Derived	Cache status	Off
Owner	admin	Last modifier	admin
Creation	Jun 10, 2025, 9:09 PM	Last modification	Jun 10, 2025, 9:12 PM
Folder	/		
Schema			

Fields:

PK	Field name	Field type	Description
	id_declaration	int	
	numero_dossier	text	
	importateur	text	

Results:

3 rows, 10 columns received (3 displayed) SELECT * FROM integration_sid.vue_declarations_enrichies CONTEXT ('cache_wait_for_load' = 'true') TRACE

id_declaration	numero_dossier	importateur	date_declaration	montant_droits	type_produit	nif	nom	secteur	pays
1	D2024-001	ENTREPRISE A	2024-09-20	300000	Électronique	4001234567	ENTREPRISE A	Électronique	Chine
2	D2024-002	ENTREPRISE B	2024-11-05	450000	Textile	4009876543	ENTREPRISE B	Textile	Turquie
3	D2024-003	ENTREPRISE C	2024-12-12	150000	Aérolimentaire	4005432198	ENTREPRISE C	Aérolimentaire	France

Donc, nous remarquons que nous avons réussi à intégrer deux sources hétérogènes (une BDD et un fichier CSV) sur Denodo dans une seule vue virtuelle de données.

-Vue Virtuelle Suivi Budgétaire global

Nom : vue_budget_execution

Jointure : entre les trois sources loi_finances ,engagements_budgétaires et paiements_effectues (trois fichiers CSV)

Conditions :

ds_paiement_effectues.id_engagement = ds_engagements_budgétaires.id_engagement

ds_engagements_budgétaires.code_programme = ds_lois_finances.code_programm

Objectif :

Suivre l'évolution des crédits, engagements et paiements par programme.

« Le DGB souhaite disposer d'une vue centralisée des engagements et paiements pour évaluer l'absorption budgétaire mensuelle. »

-Les étapes pour créer cette vue sont les mêmes, il suffit de choisir les trois vues de base de ces sources, et puis nous créons les deux jointures, la 1ère jointure entre paiement_effectues et engagements_budgetaires, et la 2ème entre engagement_budgetaires et loi_finances.

Après l'exécution de cette vue virtuelle, nous obtenons les résultats suivants :

Figure 40: Résultats de l'exécution de la vue virtuelle budget_execution.

The screenshot shows a database management tool interface. The top part displays the 'SUMMARY' tab for the view 'vue_budget_execution'. Below this, the 'RESULTS' tab is active, showing the execution results of the query. The results are displayed in a table with 10 columns. The first two rows of data are highlighted with a red border.

rcice	intitule_progra...	nature_depense	montant_accorde	id_engagement...	id_structure	code_program...	mois	montant_engage	id_paieme
rcice	intitule_progr...	nature_depense	<null>	id_engagement	id_structure	code_programme	mois	<null>	id_paieme
5	Équipement des...	capital	30000000	ENG001	STR001	PRG01	2025-01	3000000	PAY001
5	Équipement des...	capital	30000000	ENG002	STR001	PRG01	2025-02	5000000	PAY002

-Vue Virtuelle Impact des ajustements

Nom : vue_impact_ajustements

Jointure : entre la sources ajustements_budgetaires (API Json) et la source engagements_budgétaires (fichier CSV)

Conditions :

ds_engagements_budgétaires.code_programme = ds_ajustements_budgetaires.code_programme

Objectif :

« La cellule d'analyse de la loi de finances cherche à mesurer l'impact des ajustements sur la capacité de dépense réelle. »

Les étapes pour créer cette vue sont les mêmes, il suffit de choisir les deux vues de base correspondantes à ces sources, et puis nous créons une jointure entre engagements_budgetaires et ajustements_budgetaires

Voilà les résultats de l'exécution de cette vue virtuelle par agrégation :

Figure 41: Résultats de l'exécution de la vue virtuelle Impact_ajustements.

The screenshot displays a database management interface. At the top, there's a tab for 'integration_sid.vue_impact_aju...'. Below the tabs, there's a menu with 'SUMMARY', 'EDIT', 'OPTIONS', and 'VQL'. The 'SUMMARY' tab is active, showing metadata for the view 'vue_impact_ajustements' in the 'integration_sid' database. The view is derived, owned by 'admin', and was created and last modified on 'Jun 12, 2025, 12:11 AM'. Below the metadata, there's a table with columns 'PK', 'Field name', 'Field type', and 'Description'. The table lists 'id_ajustement' and 'type_ajustement' as text fields. Below this, there's a 'Query Results' tab showing the execution results. The results are displayed in a table with 9 columns: 'id_ajustement', 'type_ajustement', 'montant', 'date_ajustement', 'id_engagement', 'id_structure', 'code_programme', 'mois', and 'montant_enga...'. The results show two rows of data for 'AJ001' with different engagement IDs and amounts.

id_ajustement	type_ajustement	montant	date_ajustement	id_engagement	id_structure	code_programme	mois	montant_enga...
AJ001	+	2000000	2025-04-01	ENG001	STR001	PRG01	2025-01	3000000
AJ001	+	2000000	2025-04-01	ENG002	STR001	PRG01	2025-02	5000000

3.6.4. Création de quelques vues métier (avec précalculs)

-Vue métier Exécution par structure

Nom : vue_metier_engagements_payement

Jointure : entre la source engagementnts_budgetaires (fichier CSV) et la source paiement_effectue (fichier CSV)

Condition :

ds_paiement_effectues.id_engagement = ds_engagements_budgétaires.id_engagement

Group by: Id_Structure, Code_Programme

Les Calculs :

Somme_engagement=sum(ds_engagements_budgétaires.montant_engage) /la somme des montants d'engagement par structure.

Somme_paiement=sum(ds_paiement_effectues.montant_paye) /la somme des montants de paiement par structure.

Objectif : Avoir le détail sur la somme des engagements et des paiements par structure/programme. Pour créer cette jointure métier, nous créons une jointure entre les deux vues de base comme nous avons déjà vu dans les exemples précédents, mais cette fois nous utilisons Group By (structure, programme) et les formules de calculs, comme illustré dans la figure.39 et la figure.40

Figure 42: Utilisation Group By pour le filtrage.

The screenshot shows a database management interface with a tab titled 'integration_sid.vue_impact_aju...'. The 'SUMMARY' tab is active, displaying metadata for the view 'vue_impact_ajustements' in the 'integration_sid' database. Below the metadata, a table lists the fields of the view:

PK	Field name	Field type	Description
	id_ajustement	text	
	type_ajustement	text	
	montant	text	

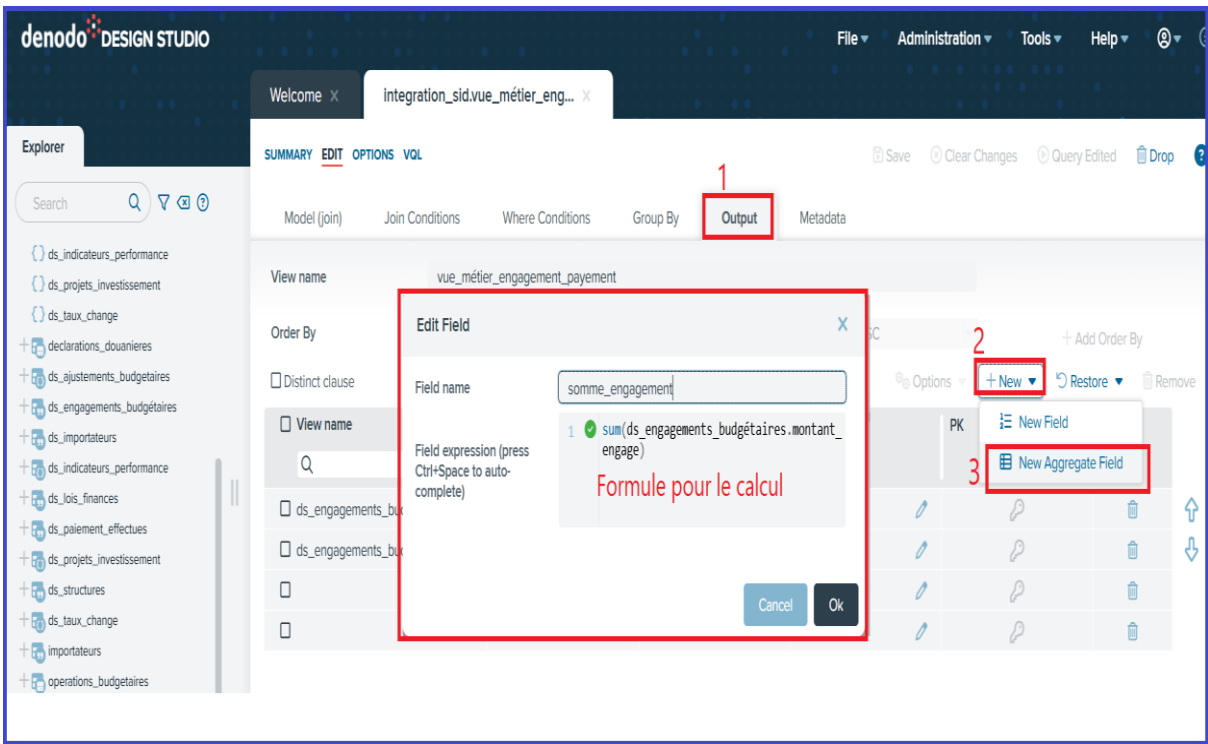
Below the metadata, the 'Query Results' tab is active, showing the results of a query. The results are displayed in a table with 9 columns: id_ajustement, type_ajustement, montant, date_ajustement, id_engagement, id_structure, code_programme, mois, and montant_enga... The results are grouped by id_structure and code_programme.

Résultats

2 rows, 9 columns received (2 displayed) SELECT * FROM integration_sid.vue_impact_ajustements CONTEXT ('cache_wait_for_load' = 'true') TRACE

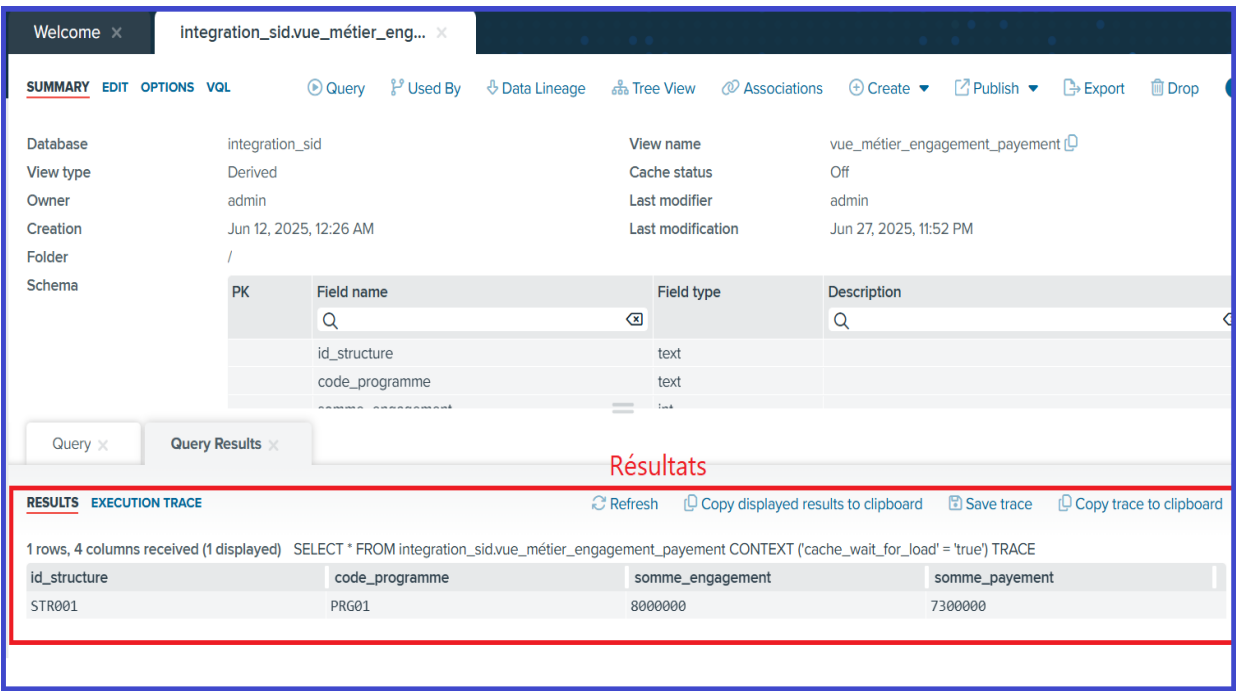
id_ajustement	type_ajustement	montant	date_ajustement	id_engagement	id_structure	code_programme	mois	montant_enga...
AJ001	+	2000000	2025-04-01	ENG001	STR001	PRG01	2025-01	3000000
AJ001	+	2000000	2025-04-01	ENG002	STR001	PRG01	2025-02	5000000

Figure 43: Calcul de la somme des montants d'engagement.



L'exécution de cette vue métier permet d'obtenir les résultats suivants :

Figure 44: Résultats de l'exécution de la vue métier engagement_paiement.



Par exemple ici la somme des montants d'engagement est à 8000000DA et la somme des montants de paiement est à 7300000DA pour la 1ère structure.

-Vue métier Taux d'exécution par programme

Nom : vue_metier_taux_execution_par_programme

Jointure : entre la vue métier de l'exemple précédent et la source loi_finances (fichier CSV)

Condition de jointure:

vue_métier_engagement_paiement.code_programme = ds_lois_finances.code_programme

Group by : Code_Programme, Montant_accorde, Somme_paiement

Les Calculs:

Taux_execution=((vue_métier_engagement_paiement.somme_paiement*100)/ds_lois_finances.montant_accorde)

Objectif : Avoir le taux d'exécution par programme.

Figure 45: Résultats de l'exécution de la vue métier engagement_paiement.

Database	integration_sid	View name	vue_metier_taux_execution_par_programme
View type	Derived	Cache status	Off
Owner	admin	Last modifier	admin
Creation	Jun 12, 2025, 10:34 PM	Last modification	Jun 27, 2025, 11:39 PM
Folder	/		
Schema			
PK	Field name	Field type	Description
	code_programme	text	
	intitule_programme	text	
	montant_accorde	text	
	somme_paiement	text	
	taux_execution	text	

code_programme	intitule_programme	montant_accorde	somme_paiement	taux_execution
PRG01	Équipement des écoles	30000000	7300000	24

Alors le taux d'exécution du premier programme est 24%.

➤ 2^{ème} Scénario :

Nous allons utiliser une BDD (Sql) fournie par le Ministère des Finance (avec des données génériques).

Nous commençons par la connexion de la plateforme Denodo avec cette BDD, puis nous créons des vues de base pour les quatre tables de cette dernière :

- Le nom de la BDD : 'bdd_ministere_finances' ;

- La 1^{ère} table 'Axe publicité' avec les champs (**code_structure, id_type_acte, date_depot, nombre**) ;
- La 2^{ème} table 'Dimension date dépôt' avec les champs (**date_depot, jour, mois, trimestre, semestre, annee**) ;
- La 3^{ème} table 'Dimension structure' avec les champs suivants (**code_structure, nom, wilaya, code_wilaya, region**) ;
- La 4^{ème} table 'Dimension type acte' avec les champs suivants (**id_type_acte, nom, type**) .

Nous commençons par la création d'une **vue dérivée (intégration)** en joignant les 4 vues de base.

-Vue Virtuelle Publicité détaillée

Nom : `vue_publicite_detaille`

Jointure : entre les quatres tables de cette BDD.

Conditions de jointure :

`axe_publicite.code_structure = dim_structure.code_structure`

`axe_publicite.id_type_acte = dim_type_acte.id_type_acte`

`axe_publicite.date_depot = dim_date_depot.date_depot`

-Objectif :

Avoir le détail des publicités sur les trois dimensions (la structure, la date de dépôt, le type d'acte et le nombre de fois).

Figure 46 : Détails sur l'axe publicité (BDD Ministère des Finances).

The screenshot shows a database management interface with a top navigation bar and a main content area. The main content area is divided into two sections: 'Summary' and 'Query Results'.

Summary Section:

- Database: bdd_ministere_finances
- View name: vue_publicite_detailee
- View type: Derived
- Cache status: Off
- Owner: admin
- Last modifier: admin
- Creation: Jun 20, 2025, 11:49 PM
- Last modification: Jun 29, 2025, 12:25 AM
- Folder: /
- Schema: PK, Field name, Field type, Description

Query Results Section:

Query Results x

RESULTS EXECUTION TRACE

150 rows, 9 columns received (150 displayed) SELECT * FROM bdd_ministere_finances.vue_publicite_detailee CONTEXT ('cache_wait_for_load' = 'true') TRACE

nom_structure	wilaya	region	nom_acte	type	annee	mois	trimestre	nombre
Conservation Ti...	Tiaret	Ouest	Mise sous séque...	INS	2008	10	4	4
Conservation Be...	Bejaia	Est	Droit d'affecta...	INS	2008	10	4	0
Conservation Be...	Bejaia	Est	Immatriculation...	IMM	2008	10	4	1
Conservation22	Blida	Centre	Donation usufru...	PUB	2008	10	4	0
Conservation Ti...	Tiaret	Ouest	Mise sous séque...	INS	2008	10	4	2
Conservation Ti...	Tiaret	Ouest	Acte de Wakf (H...	PUB	2008	10	4	0
Conservation22	Blida	Centre	Immatriculation...	IMM	2008	10	4	3
Conservation Be...	Bejaia	Est	Cabier des cher...	PUB	2008	10	4	0

Les résultats de l'exécution de cette vue virtuelle contiennent 150 lignes (enregistrements).

Maintenant nous nous intéressons au nombre d'actes par type et par année comme suite :

-Vue Virtuelle Nombre d'actes par année et par type

Nom : vue_agregation_nombre_actes_par_annee_et_type

Nous appliquons des filtrages sur la vue virtuelle précédente (vue_publicite_detailee)

Group by: annee, nom_type_acte

Avec le calcul suivant : **nombre_total_acte= sum(axe_publicite.nombre)**

-Objectif :

Calcul du nombre d'actes par type et par année.

Les résultats de l'exécution de cette vue virtuelle :

Figure 47: Nombre d'actes par type et par année.

The screenshot displays a database management interface. At the top, there's a tab for 'bdd_ministere_finances.vue_agr...'. Below it, a 'SUMMARY' tab is active, showing metadata for the view 'vue_agregation_nombre_actes_par_annee_et_type'. The metadata includes the database name, view type (Derived), owner (admin), creation date, and last modification date. Below this, a table lists the fields of the view: 'nom_type_acte' (text), 'annee' (int), and 'nombre_total_acte' (int). The 'nombre_total_acte' field is highlighted with a red box. Below the metadata, the 'Query Results' tab is active, showing the execution results of the view. The results are displayed in a table with 18 rows and 3 columns: 'nom_type_acte', 'annee', and 'nombre_total_acte'. The 'nombre_total_acte' column is highlighted with a red box. The results show the number of acts for various types of acts across different years.

nom_type_acte	annee	nombre_total_acte
Acte de Wakf (Habous)	2008	19105
Donation usufruit	2008	19485
Immatriculation Provisoire 4 mois	2008	19517
Immatriculation Définitive	2008	19314
Mise sous séquestre	2008	19025
Cahier des charges	2008	19034
Saisie	2008	18953
Droit d'affectation	2008	19208

Maintenant, si nous nous intéressons au nombre des actes par Région :

-Vue Virtuelle Nombre d'actes par Région

Nom : vue_agregation_repartition_par_region

Group by : region (pour faire de filtrage par region)

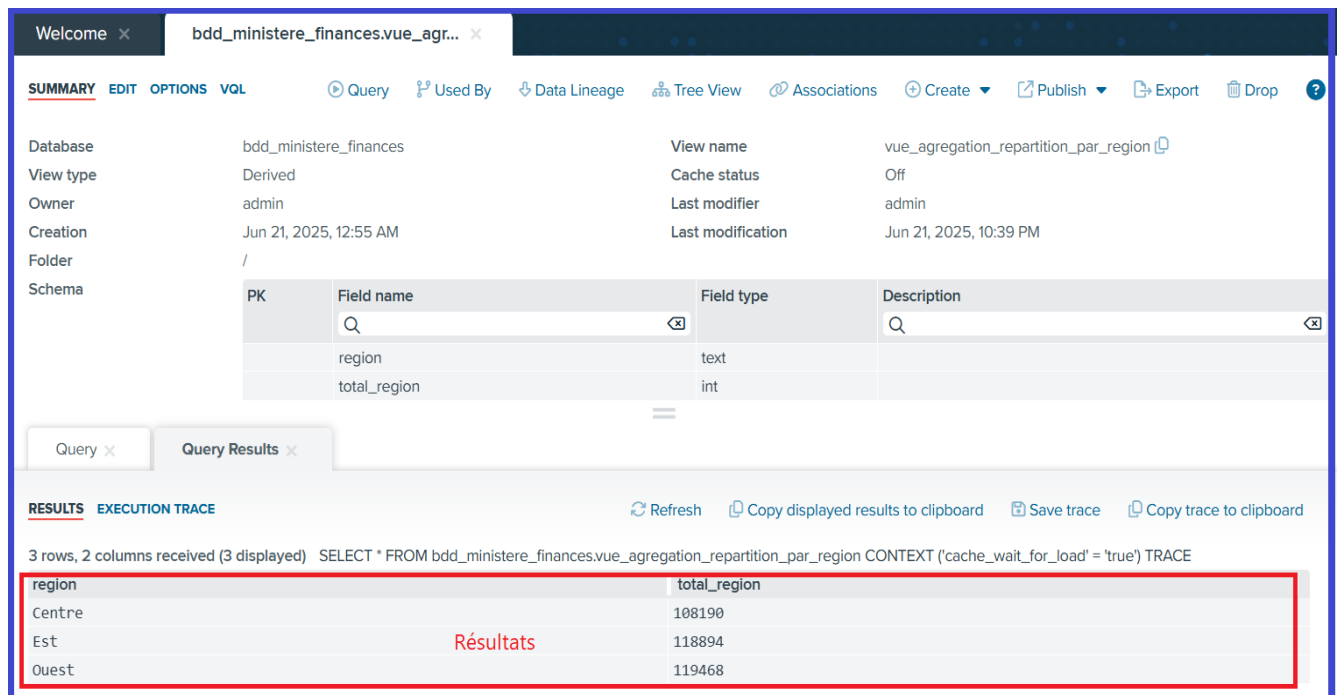
Avec le calcul suivant : **total_region= sum(axe_publicite.nombre)**

-Objectif :

Calcul du nombre d'actes par région (Est, Centre, Ouest)

Les résultats de l'exécution de cette vue virtuelle :

Figure 48: Nombre d'actes par région.



The screenshot displays a database management interface. The top navigation bar includes tabs for 'Welcome', 'bdd_ministere_finances.vue_agr...', and a search bar. Below the navigation bar, there are tabs for 'SUMMARY', 'EDIT', 'OPTIONS', and 'VQL'. The 'SUMMARY' tab is active, showing details about the view 'vue_agregation_repartition_par_region'. The view is derived, owned by 'admin', and was created on 'Jun 21, 2025, 12:55 AM'. It has two fields: 'region' (text) and 'total_region' (int). Below the summary, there are tabs for 'Query' and 'Query Results'. The 'Query Results' tab is active, showing the results of a query. The query is 'SELECT * FROM bdd_ministere_finances.vue_agregation_repartition_par_region CONTEXT ('cache_wait_for_load' = 'true') TRACE'. The results show 3 rows, 2 columns. The columns are 'region' and 'total_region'. The rows are 'Centre' (108190), 'Est' (118894), and 'Ouest' (119468). The word 'Résultats' is displayed in the center of the results table.

region	total_region
Centre	108190
Est	118894
Ouest	119468

Maintenant nous nous intéressons aux structures les plus actives, dont le nombre d'actes est supérieur à 1000 durant les années 2007 et 2008.

-Vue Virtuelle Structures plus actives :

Nom : vue_metier_structures_actives

Group by : nom_structure, annee

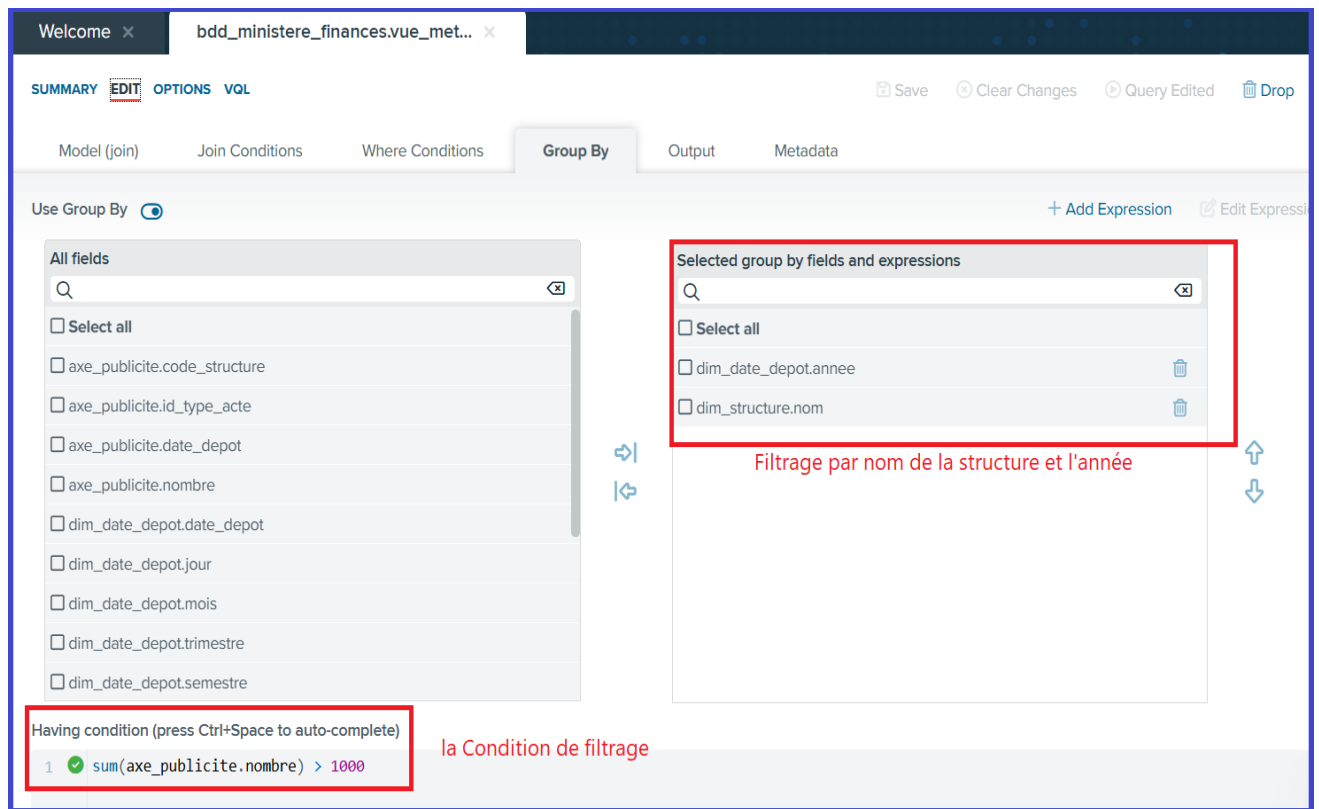
Avec le calcul suivant : **total acte= sum(axe_publicite.nombre)** avec la condition (**total acte > 1000**)

-Objectif :

Connaitre les structures les plus actives dans les trois régions.

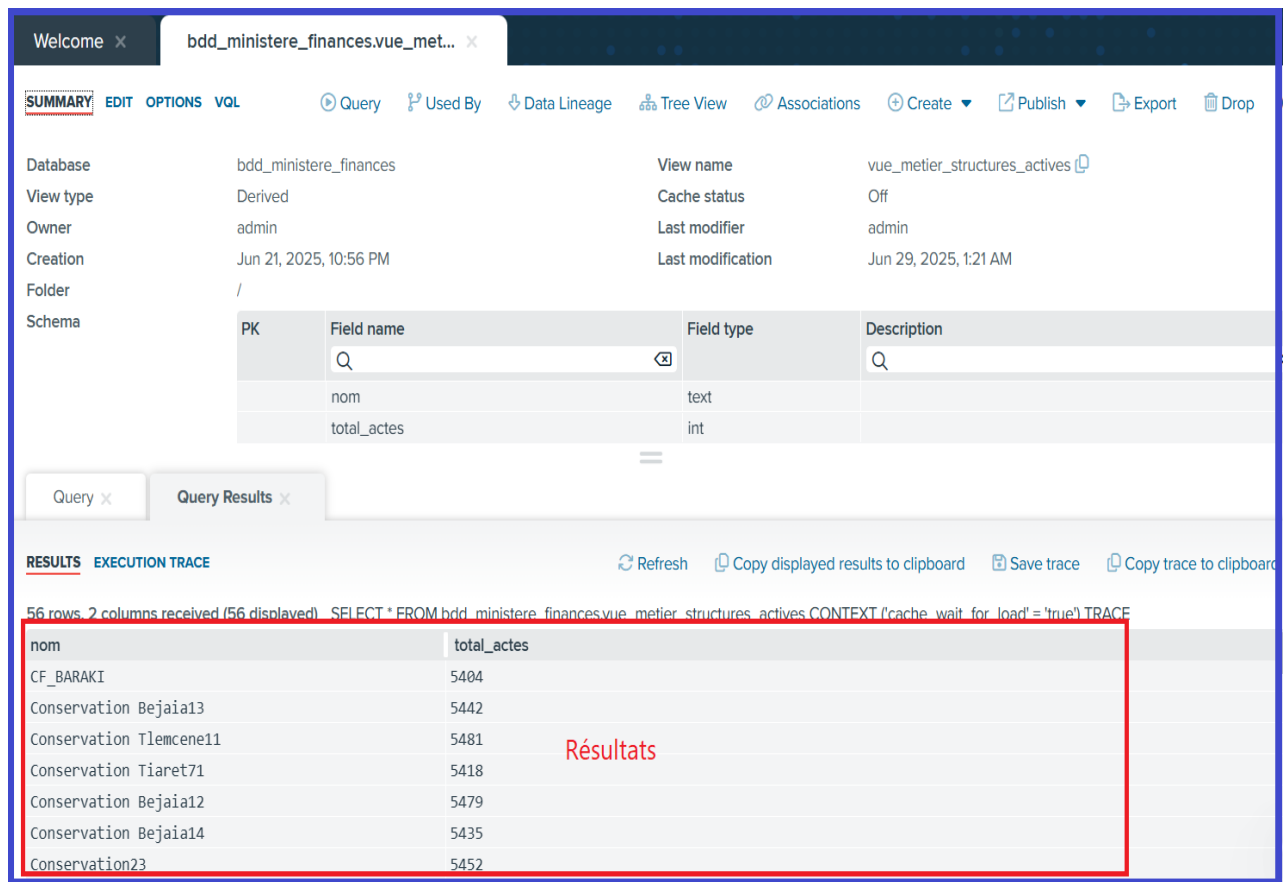
Nous rajoutons la condition de filtrage comme suite :

Figure 49: Filtrage de structures selon le nombre d'actes.



Après l'exécution de cette vue, nous aurons les résultats suivants :

Figure 50: Les Structures les plus actives.



PK	Field name	Field type	Description
	nom	text	
	total_actes	int	

nom	total_actes
CF_BARAKI	5404
Conservation Bejaia13	5442
Conservation Tlemcene11	5481
Conservation Tiaaret71	5418
Conservation Bejaia12	5479
Conservation Bejaia14	5435
Conservation23	5452

-Maintenant, nous nous intéressons au meilleur type d'acte pour chaque région :

Vue Virtuelle Top type d'acte par région

Nom : vue_metier_meilleur_type_acte_par_region

Group by : region, nom_acte, total avec la condition (**total = Max(nombre des actes)**)

Objectif :

Afficher, pour chaque région, le type d'acte le plus fréquent (c'est-à-dire celui ayant la valeur maximale du nombre d'actes).

Les résultats de l'exécution de cette vue virtuelle :

Figure 51: Top Type Acte par Région

The screenshot displays a database management interface for the 'bdd_ministere_finances' database. The selected view is 'vue_metier_meilleur_type_par_region'. The view's schema is as follows:

PK	Field name	Field type	Description
	region	text	
	nom_type_acte	text	
	total	text	

The 'RESULTS' tab shows the following data:

region	nom_type_acte	total
Est	Immatriculation Provisoire 4 mois	13338
Ouest	Immatriculation Provisoire 4 mois	13442
Centre	Immatriculation Définitive	12145

3.6.5. Connexion de la plateforme Denodo avec un outil de visualisation (PowerBI)

La connexion entre Denodo Platform 9 et Power BI est une fonctionnalité essentielle pour les entreprises qui souhaitent tirer parti de la virtualisation des données de Denodo avec les capacités d'analyse et de visualisation de Power BI. Denodo agit comme une couche d'intégration et d'abstraction des données, permettant à Power BI d'accéder à des données hétérogènes provenant de diverses sources sans avoir à comprendre leur complexité sous-jacente.

La connexion entre Denodo Platform 9 et Power BI est intrinsèquement liée à l'amélioration de la prise de décision au sein d'une organisation. Cette synergie permet de transformer des données brutes et fragmentées en informations exploitables et pertinentes pour les décideurs.

Voici comment cette intégration facilite et optimise la prise de décision :

- Accès unifié et complet aux données pour une vision à 360° :

-Le rôle de Denodo : Denodo agit comme une couche de virtualisation de données, intégrant des données provenant de sources hétérogènes (bases de données relationnelles, NoSQL, applications cloud, services web, fichiers plats, etc.) sans les déplacer. Il brise les silos de données.

-Lien avec la décision : Pour prendre des décisions éclairées, les dirigeants ont besoin d'une vue d'ensemble et consolidée de leurs opérations (ventes, marketing, finance, logistique, production, etc.). Denodo fournit cette vision à 360 degrés, permettant à Power BI d'accéder à toutes les données pertinentes via un point d'entrée unique. Cela évite les décisions basées sur des informations incomplètes ou incohérentes.

- Informations en temps quasi réel pour des décisions agiles :

-Le rôle de Denodo et Power BI (Direct Query) : Grâce au mode Direct Query dans Power BI, les requêtes sont exécutées en temps réel sur Denodo. Denodo, à son tour, optimise ces requêtes vers les sources de données sous-jacentes.

-Lien avec la décision : Dans un environnement commercial en constante évolution, la capacité à prendre des décisions rapidement est cruciale. Accéder à des données à jour et non à des instantanés d'il y a 24 heures permet aux décideurs de réagir plus vite aux tendances du marché, aux changements de comportement client ou aux problèmes opérationnels, transformant la réactivité en avantage concurrentiel.

- Amélioration de la qualité et de la confiance dans les données :

-Le rôle de Denodo : Denodo permet d'appliquer des règles de nettoyage, de transformation et de gouvernance des données au niveau de la couche virtuelle. Il assure la cohérence et la fiabilité des données.

-Lien avec la décision : Des décisions sont aussi bonnes que les données sur lesquelles elles sont basées. Si les données sont erronées ou incohérentes, les décisions seront forcément imparfaites. En garantissant la qualité des données avant qu'elles n'atteignent Power BI, Denodo renforce la confiance des décideurs dans les rapports et les analyses qu'ils consultent, les incitant à s'appuyer davantage sur les données.

- Analyse approfondie et découverte d'insights :

-Le rôle de Power BI : Power BI excelle dans l'analyse de données et la visualisation. Il permet de créer des tableaux de bord interactifs, des rapports dynamiques et d'appliquer des techniques d'analyse avancées.

-Lien avec la décision : Une fois que Denodo a fourni un jeu de données propre et unifié, Power BI permet aux analystes et aux décideurs d'explorer ces données en profondeur. Ils peuvent identifier des tendances, des corrélations, des anomalies et des opportunités qui n'auraient pas été visibles autrement. Ces "insights" sont le moteur de décisions stratégiques, qu'il s'agisse d'optimiser les campagnes marketing, de rationaliser la chaîne d'approvisionnement ou de développer de nouveaux produits.

- Réduction de la complexité technique et focalisation sur la valeur métier :
- Le rôle de Denodo : Denodo abstrait la complexité technique des sources de données sous-jacentes. Les utilisateurs de Power BI interagissent avec des vues métiers simplifiés.
- Lien avec la décision : Les décideurs peuvent se concentrer sur l'analyse des résultats et l'élaboration de stratégies, plutôt que de se soucier de la provenance ou du format des données. Cette simplification technique libère du temps et des ressources qui peuvent être réinvestis dans l'innovation et la stratégie.

3.6.5.1. Étapes Générales de Connexion

- **Configuration de la connexion dans Power BI Desktop :**

- Ouvrir Power BI Desktop.
- Cliquer sur "Obtenir des données" (Get Data).
- Rechercher "Denodo" ou "ODBC".
- Sélectionner l'option de connexion (par exemple, "Denodo").
- Configurer la connexion en fournissant le DSN ou la chaîne de connexion directe, ainsi que vos identifiants Denodo.
- Dans le navigateur Power BI, sélectionner les vues Denodo que vous souhaitez importer ou interroger.

- **Mode de Connexion (Import vs. Direct Query) :**

-Import (Importation) : Les données sont chargées directement dans le modèle de données de Power BI. Cela offre de meilleures performances pour les rapports, mais les données ne sont pas en temps réel et nécessitent une actualisation.

-Direct Query : Les requêtes sont directement envoyées à Denodo chaque fois qu'un utilisateur interagit avec le rapport. Cela garantit des données en temps réel mais peut être plus lent pour les rapports complexes. C'est l'un des points forts de l'intégration avec Denodo, car Denodo peut optimiser ces requêtes vers les sources sous-jacentes.

3.6.5.2. Connexion de PowerBI avec la BDD virtuelle (BDD_ministere_finances)

Après avoir ouvert Power BI Desktop, nous allons suivre les étapes générales de connexion que nous avons vues précédemment.

Figure 52: Connexion avec la plateforme Denodo.

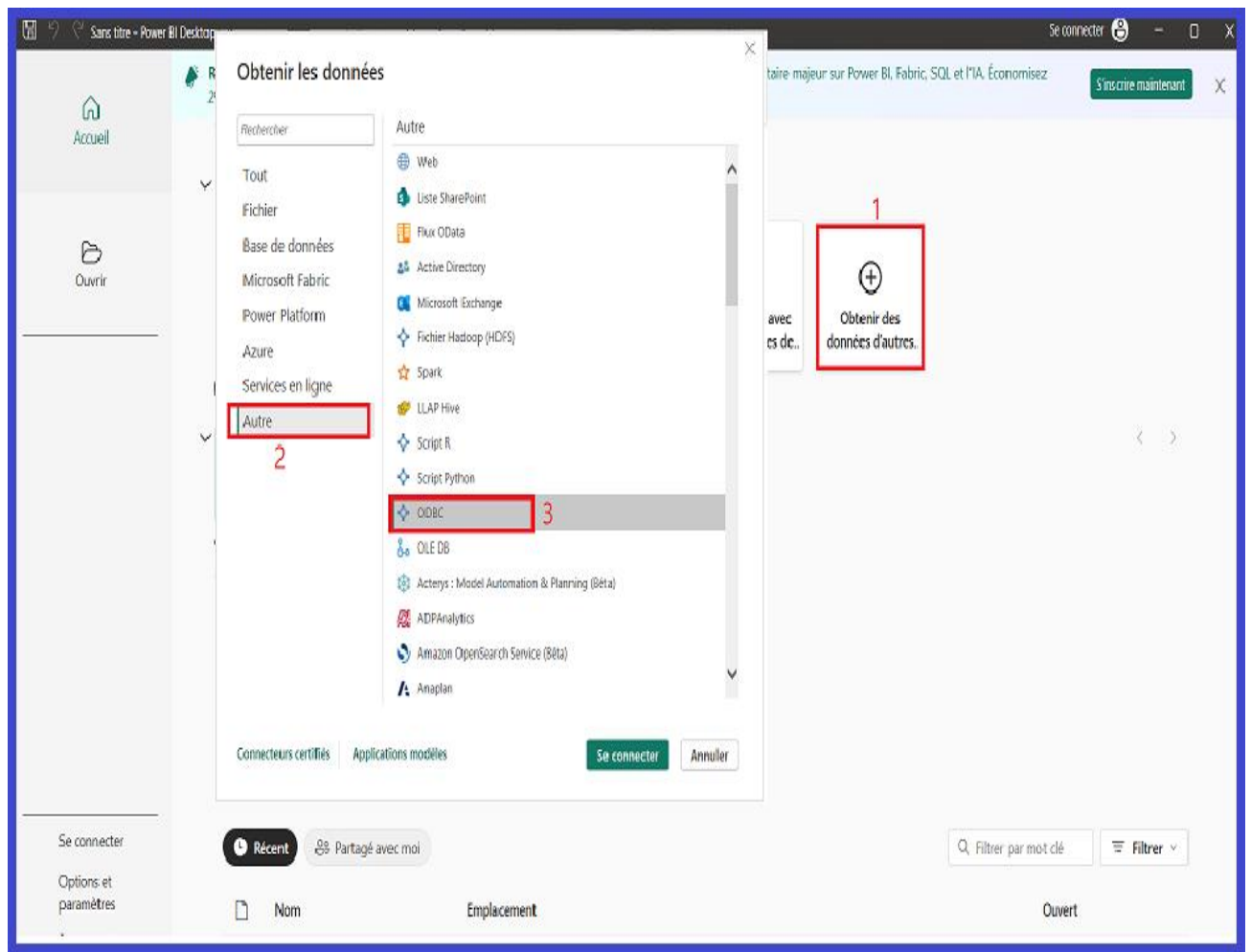
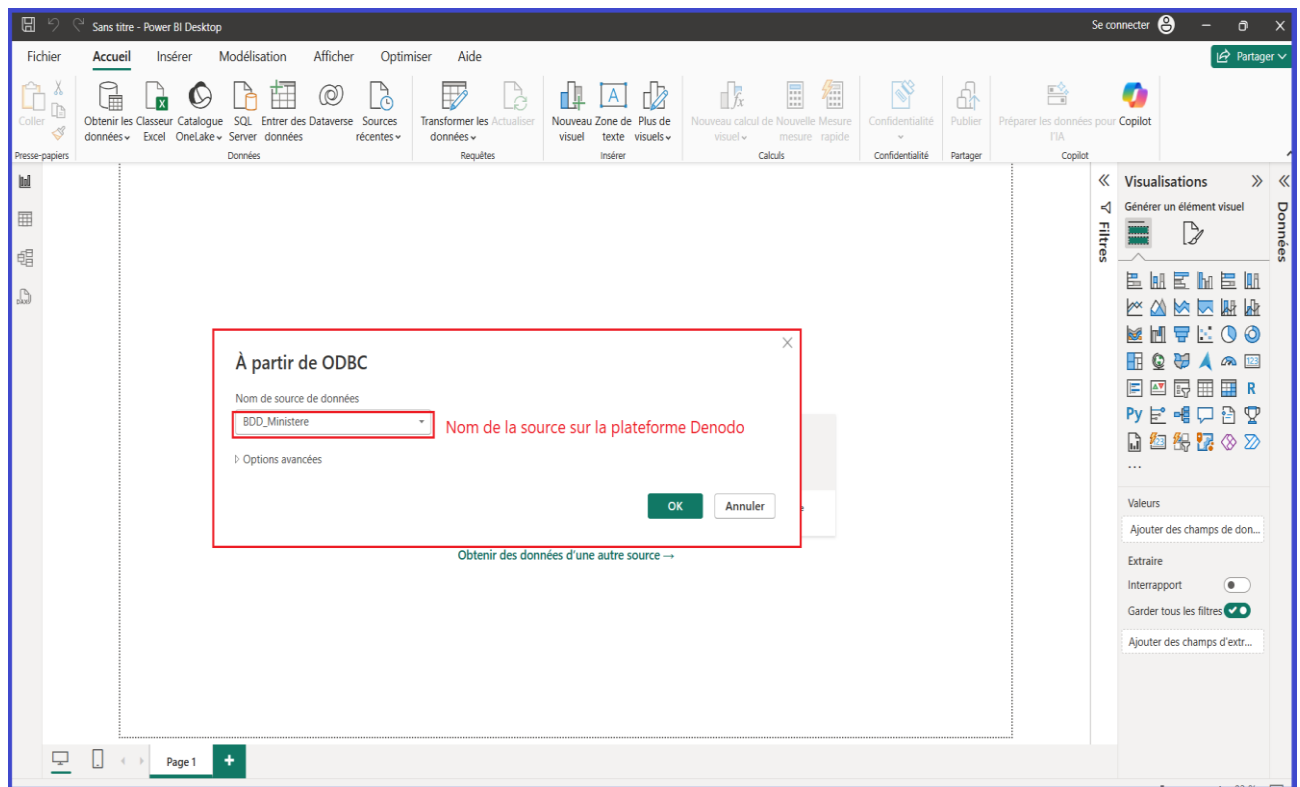
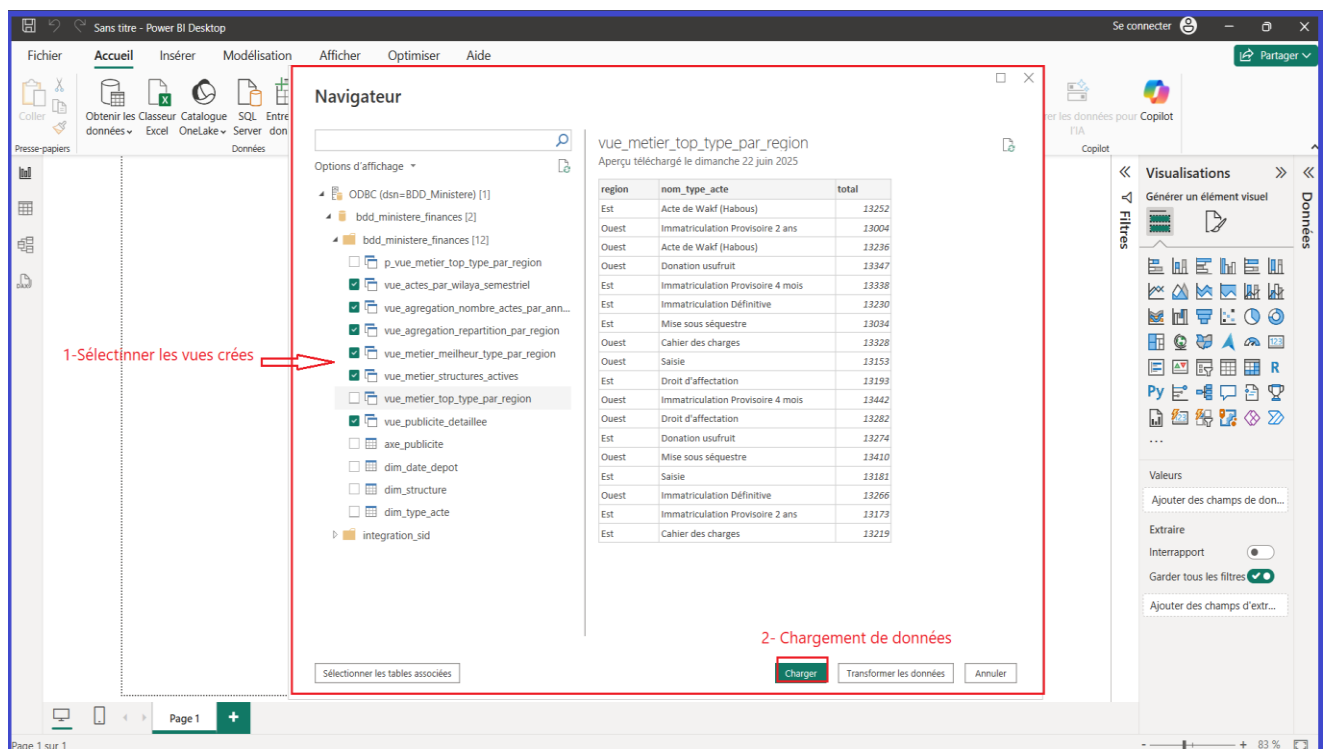


Figure 53: Configuration de la connexion avec la plateforme.



Après avoir choisi le nom de la connexion, nous verrons l'ensemble des **vues virtuelles** déjà créées sur la plateforme Denodo.

Figure 54: Chargement de données sur PowerBI.

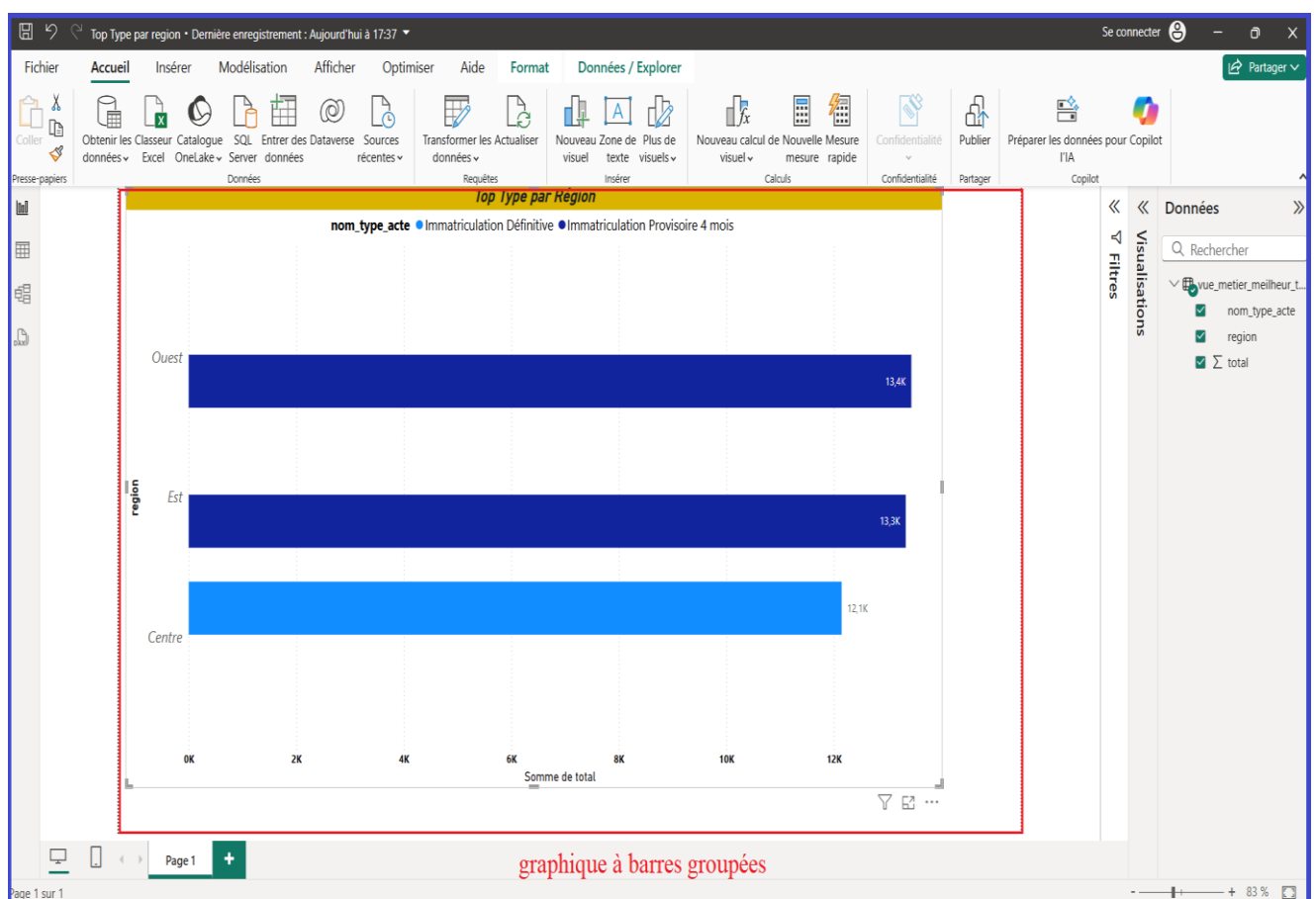


Une fois les données importées et préparées dans Power BI, nous pouvons les visualiser efficacement pour élaborer des rapports pertinents et soutenir la prise de décision stratégique.

-Analyse du Meilleur Type d'Acte par Région

Pour obtenir une vue d'ensemble du meilleur type d'acte par région, l'utilisation d'un graphique à barres groupées est en effet un excellent choix. Ce type de visualisation permet de comparer facilement les différentes catégories (types d'actes) au sein de chaque groupe (régions), et d'identifier rapidement les tendances ou les points saillants.

Figure 55: Visualisation du meilleur type d'acte par région avec PowerBI.



-L'interprétation des résultats

➤ Région Est :

Le type d'acte "**Immatriculation Provisoire 4 mois**" est le plus utilisé ou le plus enregistré dans cette région, avec **13 338 actes**.

➤ Région Ouest :

Le même acte "**Immatriculation Provisoire 4 mois**" est aussi le plus fréquent ici, avec **13 442 actes**. Cela peut refléter une politique ou une pratique administrative commune dans l'Ouest.

➤ Région Centre :

C'est "**Immatriculation Définitive**" qui est le plus utilisé avec **12 145 actes**, ce qui suggère que dans cette région, les véhicules atteignent plus rapidement la phase d'immatriculation finale, ou bien que les demandes provisoires sont moins nombreuses.

-Analyse de l'impact des ajustements

Concentrons-nous maintenant sur la vue virtuelle intitulée « **Impact des Ajustements du 1er scénario** ». Cette vue, obtenue en croisant les engagements budgétaires et les ajustements, est essentielle pour le suivi financier et la prise de décision budgétaire éclairée.

Nous commencerons par charger les données de cette vue dans Power BI :

Figure 56: Chargement de données de la vue Impact des ajustements.

The screenshot shows the Denodo Navigator interface. On the left, a tree view under 'bdd_ministere_finances [12]' and 'integration_sid [23]' lists various views. The view 'impacts_des_ajustements' is selected and highlighted with a red box. A red annotation '1-Selectionnement de la vue Impact des ajustements' points to this selection. On the right, a table titled 'impacts_des_ajustements' displays data. Below the table, a red annotation '2-Chargement de Données' points to a 'Charger' button in the bottom right corner. The table data is as follows:

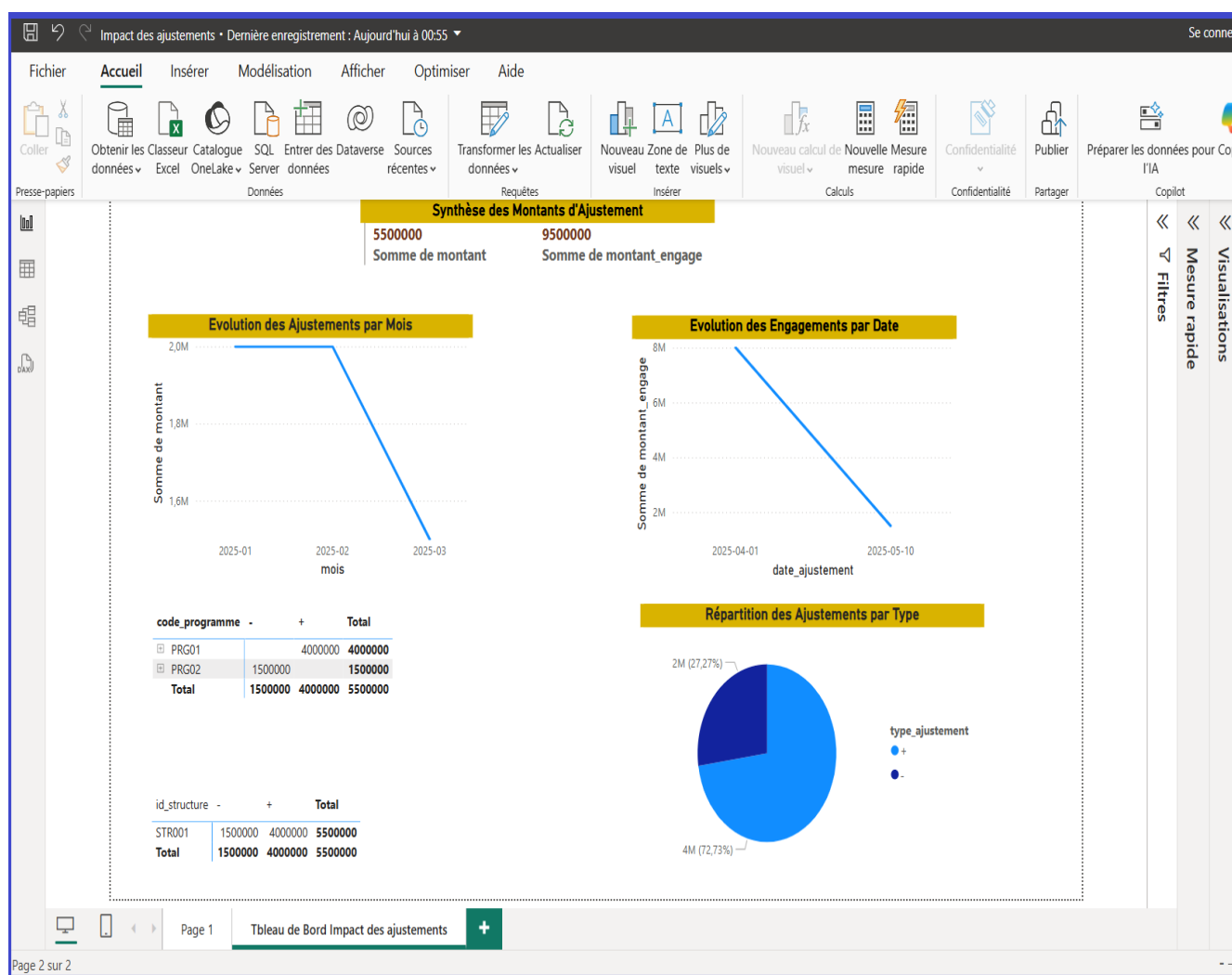
id_ajustement	type_ajustement	montant	date_ajustement	id_engagement	id_structure	code_programme	mois	montant_engage
AI001	+	2000000	2025-04-01	ENG001	STR001	PRG01	2025-01	3000000
AI001	+	2000000	2025-04-01	ENG002	STR001	PRG01	2025-02	5000000
AI002	-	1500000	2025-05-10	ENG003	STR001	PRG02	2025-03	1500000

Pour cet exemple, nous allons **développer un tableau de bord avec Power BI**. Il illustre l'un des bénéfices majeurs de l'intégration de la virtualisation de données Denodo: la capacité à fournir une vue consolidée et en temps quasi-réel de l'impact des ajustements budgétaires. Ce tableau de bord permet une prise de décision agile et une meilleure gestion financière, sans les contraintes des silos de données traditionnels.

Pour ce faire, nous utiliserons les visuels suivants :

- Une Carte à plusieurs lignes : pour une Synthèse des Montants d'Ajustement ;
- Deux Graphiques en courbes : pour l'Évolution des Ajustements et Engagements par Mois ;
- Une Graphique en secteurs : pour la Répartition des Ajustements par Type ;
- Deux Matrices : pour l'Ajustements par Programme et Structure ;

Figure 57: Tableau de Bord Impact des ajustements.



Analyse Détaillée du Tableau de Bord :

- "Synthèse des Montants d'Ajustement" (Cartes KPI en haut)

- **Observation :** Nous avons deux indicateurs clés affichés : "Somme de montant" (5 500 000) et "Somme de montant_engage" (9 500 000).
- **Interprétation :**
 - Le "Somme de montant" représente le total des ajustements appliqués au budget. Dans cet exemple, il s'agit d'un total cumulé des ajustements (positifs et négatifs), avec des signes.
 - Le "Somme de montant_engage" représente le montant total des engagements budgétaires initiaux.
 - Ces chiffres offrent une vue d'ensemble immédiate de l'ampleur des mouvements financiers. Le fait que la somme des ajustements soit significative par rapport au montant engagé (5.5M sur 9.5M) suggère une activité budgétaire

intense, ce qui peut indiquer une flexibilité, mais aussi un besoin de surveillance.

- **Aide à la Décision** : Ces KPI en temps réel, rendus possibles par Denodo, permettent aux dirigeants d'évaluer la situation financière globale en un coup d'œil. Si la "Somme de montant" est excessivement élevée par rapport au "Somme de montant_engage", cela pourrait indiquer une planification budgétaire initiale inefficace ou des événements imprévus majeurs, nécessitant une révision des processus.

- "Évolution des Ajustements par Mois" (Graphique en courbes, à gauche)

- **Observation** : Ce graphique montre une ligne bleue décroissante, avec des points de données pour "2025-01", "2025-02" et "2025-03". Le montant passe de plus de 2M à environ 1.5M.
- **Interprétation** : Ce visuel révèle la tendance des ajustements au fil des mois. La diminution du montant total des ajustements de janvier à mars 2025 peut indiquer soit une période de moindre activité d'ajustement, soit que des ajustements négatifs plus importants ont été appliqués, ou encore que les ajustements positifs ont diminué.
- **Aide à la Décision** : Ce graphique permet d'identifier les périodes où les ajustements sont les plus fréquents ou les plus importants. Une tendance à la baisse pourrait être positive si elle signifie une meilleure planification, mais négative si elle indique des coupes budgétaires importantes. La capacité de suivre cette tendance mensuelle, directement depuis les systèmes sources via Denodo, est cruciale pour une réactivité financière.

- "Évolution des Engagements par Date" (Graphique en courbes, à droite)

- **Observation** : Similaire au précédent, ce graphique montre l'évolution du "montant_engage" par date_ajustement. On observe également une décroissance, passant de plus de 8M à environ 2M.
- **Interprétation** : Ce graphique suit l'évolution des montants engagés à chaque date d'ajustement. La forte décroissance pourrait signifier que les montants engagés initiaux ont été significativement réduits au fil du temps par les ajustements négatifs, ou que les ajustements s'appliquent à des engagements de plus en plus petits.
- **Aide à la Décision** : En comparant l'évolution des ajustements avec celle des engagements, les décideurs peuvent évaluer l'impact net des modifications budgétaires. Une forte réduction des engagements globaux pourrait signaler une contraction des dépenses ou une réorientation stratégique. La virtualisation de données assure que cette vue est toujours à jour avec les dernières modifications.

- "Répartition des Ajustements par Type" (Graphique en anneau)

- **Observation** : Le graphique montre deux segments : le segment principal, plus grand, représente **72.73%** du total des ajustements et correspond au type + (4M). Le segment plus petit représente **27.27%** du total et correspond au type - (1.5M).
- **Interprétation** : Ce visuel est très clair : une **écrasante majorité (près des trois quarts, soit 72.73%) des ajustements sont de type positif (+)**, ce qui indique des augmentations budgétaires ou des ajouts de fonds. Inversement, une minorité significative (27.27%) des ajustements sont de type négatif (-), signifiant des réductions budgétaires ou des annulations.
- **Aide à la Décision** : C'est un insight critique pour la gestion financière. La prédominance des ajustements positifs (plus de deux fois les ajustements négatifs) pourrait indiquer :
 - **Croissance ou Expansion** : L'entreprise est peut-être en phase de croissance, nécessitant des budgets supplémentaires pour de nouveaux projets, initiatives ou des extensions.
 - **Sous-estimation des Coûts Initiaux** : Les budgets initiaux pourraient être systématiquement sous-estimés, nécessitant des ajustements à la hausse au cours de l'exercice.
 - **Flexibilité Budgétaire** : Une capacité à allouer des fonds supplémentaires rapidement pour saisir des opportunités ou répondre à des besoins non prévus.
 - **Action** : Les gestionnaires financiers devraient enquêter sur les raisons de cette prédominance des ajustements positifs. Est-ce un signe sain de croissance et d'opportunités, ou cela révèle-t-il des lacunes dans le processus de prévision budgétaire initial qui nécessiteraient d'être affinées ? La capacité de voir cette répartition et son évolution en temps réel, grâce à la virtualisation des données, est un atout majeur pour la gouvernance financière et la planification stratégique.

- Matrices : "code_programme" et "id_structure"

- **Observation** : Deux matrices affichent le montant des ajustements ventilé par code_programme et id_structure, avec des colonnes pour + et - et un Total général.
 - **Matrice code_programme** : PRG01 a 4 000 000 d'ajustements positifs et 0 négatifs (Total 4 000 000). PRG02 a 1 500 000 d'ajustements négatifs et 0 positifs (Total 1 500 000). Le total général est 5 500 000.

- **Matrice id_structure** : STR001 a 4 000 000 d'ajustements positifs et 1 500 000 négatifs (Total 5 500 000).
- **Interprétation :**
 - Ces matrices fournissent une granularité essentielle sur l'origine des ajustements.
 - Le PRG01 semble avoir reçu des ajustements exclusivement positifs, tandis que le PRG02 a subi des ajustements exclusivement négatifs. C'est une divergence frappante.
 - Toutes les activités d'ajustement semblent être concentrées sur la STR001, qui gère à la fois les ajustements positifs (de PRG01) et négatifs (de PRG02).
- **Aide à la Décision** : Ces matrices sont des outils puissants pour cibler les interventions :
 - **Direction des Programmes** : Pourquoi PRG01 reçoit-il tant d'ajustements positifs ? Est-ce une croissance inattendue ou une sous-estimation initiale ? Pourquoi PRG02 subit-il des réductions ? Cela peut indiquer un problème de performance ou un changement de priorité.
 - **Responsables de Structure** : La STR001 est le point central de tous les ajustements. Une surveillance particulière est nécessaire pour cette structure pour comprendre pourquoi elle est si active en termes d'ajustements et si elle est bien équipée pour gérer ces fluctuations.
 - **Audit et Conformité** : Permet de tracer précisément l'impact budgétaire de chaque ajustement et de chaque programme/structure, facilitant les audits et la redevabilité.

Chapitre 4 : Discussion des résultats

Chapitre 4 : Discussion des résultats

4.1. Synthèse analytique des réponses des utilisateurs opérationnels

Les interviews menées auprès des cadres de la Direction générale des Finances publiques exposent clairement les forces et faiblesses du système décisionnel actuel, ainsi que les attentes précises vis-à-vis de l'intégration de la virtualisation des données.

4.1.1. Difficultés majeures avec le système actuel

Les utilisateurs soulignent principalement des problèmes de performance et de réactivité. La lenteur dans l'exploitation de données provenant de sources multiples ralentit considérablement la production des rapports essentiels à la prise de décision, limitant ainsi la capacité du ministère à réagir rapidement aux évolutions économiques. Par ailleurs, des interruptions fréquentes et des accès instables perturbent le travail quotidien, engendrant frustration et perte de confiance. L'ergonomie jugée insuffisante, notamment pour les nouveaux utilisateurs ou les profils non techniques, accroît la complexité d'utilisation et nécessite un accompagnement renforcé.

4.1.2. Attentes pour la solution future

L'intégration d'une solution de virtualisation suscite un fort espoir de gains en efficacité. Les utilisateurs attendent un accès plus rapide aux données, une meilleure fiabilité et un meilleur contrôle centralisé de l'information, permettant d'éliminer les redondances et dysfonctionnements liés aux copies multiples et aux silos de données. L'amélioration de l'interface, avec des fonctionnalités personnalisables et plus intuitives, est également une requête récurrente. Enfin, le besoin d'un accompagnement renforcé en termes de formation et de support est clairement exprimé pour assurer l'appropriation de la technologie.

4.1.3. Perception des bénéfices de la virtualisation

Les cadres estiment que la virtualisation permettra une unification des sources de données et une simplification de l'accès, conduisant à une prise de décision plus rapide et mieux informée. La perspective d'une réduction des interruptions, de la centralisation de la gouvernance des données et d'une meilleure qualité de l'information est vue comme un facteur clé de succès. Ce résultat attendu s'inscrit dans une stratégie plus large de modernisation des systèmes numériques pour répondre aux exigences croissantes du pilotage public.

Cette synthèse met en lumière la nécessité d'adresser simultanément les aspects technologiques, ergonomiques et humains lors de la mise en œuvre de la virtualisation dans le système d'aide à la décision. Les difficultés constatées justifient les choix stratégiques engagés, tandis que les attentes formulées définissent un cadre clair pour orienter le développement et l'accompagnement du changement.

4.2. Analyse critique des réponses des décideurs stratégiques

Les réponses des deux décideurs stratégiques soulignent des motivations et des objectifs clairement alignés sur la nécessité d'une modernisation profonde du système d'information du ministère des Finances, avec un focus marqué sur la rapidité d'accès aux données, la qualité et la fiabilité de l'information. La volonté d'« assurer une meilleure gouvernance des données » et de disposer d'une infrastructure « flexible et agile » correspond bien aux tendances actuelles dans la transformation numérique des administrations publiques, où les défis liés à la volumétrie et à l'hétérogénéité des données sont majeurs.

4.2.1. Forces observées

La vision stratégique partagée par les deux décideurs, orientée vers le renforcement de l'architecture des données et vers l'optimisation des processus décisionnels, est un point fort.

Les critères de succès définis, basés sur des indicateurs mesurables comme le temps d'accès aux données, le taux d'adoption des outils et la qualité des tableaux de bord, témoignent d'une approche pragmatique tournée vers des résultats concrets.

L'attention portée sur la gouvernance des données, avec une centralisation accrue et une responsabilisation des directions, traduit une maturité organisationnelle importante et une volonté d'assurer la sécurité et la conformité réglementaire.

Le constat d'un impact stratégique positif sur la collaboration inter-services et la prise de décision fondée sur des données consolidées est en ligne avec les bénéfices attendus des plateformes de virtualisation modernes.

4.2.2. Limites et défis majeurs

Malgré une anticipation claire des risques liés au changement humain (formation, résistance), les réponses manquent encore de détails sur les mécanismes concrets déployés pour y faire face ou les niveaux d'acceptation réelle observés jusqu'ici. La résistance au changement

constitue en général un frein critique qui peut compromettre durablement le succès du projet si elle n'est pas gérée de manière appropriée.

Sur le plan technique, les références générales à la complexité d'intégration des systèmes anciens et nouveaux ne sont pas accompagnées d'éléments sur la manière dont ces défis techniques sont surmontés. Une analyse plus approfondie des outils et méthodologies spécifiques utilisés serait utile pour renforcer cette partie.

La recommandation de démarrer par des projets pilotes et de garantir une formation continue est pertinente, mais on perçoit un besoin d'une stratégie plus holistique qui intègre également la gestion des connaissances, la documentation dynamique, et une communication proactive pour mobiliser toutes les parties prenantes.

4.2.3. Perspective critique

Ces réponses reflètent une posture réaliste et pragmatique, mettant en avant les aspirations stratégiques d'une administration moderne et consciente des enjeux liés à la data. Toutefois, la réussite effective du projet dépendra en grande partie de la capacité à maîtriser l'accompagnement du changement aussi bien que les défis techniques, deux dimensions qui doivent être traitées en synergie.

Il est essentiel que la gouvernance ne se limite pas à la simple centralisation des données, mais qu'elle intègre aussi une gouvernance des usages et une politique d'amélioration continue basée sur l'écoute des retours des utilisateurs et la mesure d'impact régulière.

En résumé, la discussion devrait s'appuyer sur cet équilibre entre les promesses techniques et stratégiques de la virtualisation des données d'une part, et les besoins critiques en termes d'accompagnement humain, de gestion des risques et de pilotage opérationnel d'autre part. Cet équilibre est le socle sur lequel fonder des recommandations opérationnelles adaptées et un plan d'action réaliste.

4.3. Analyse critique des réponses des experts techniques

Les réponses des experts techniques démontrent une compréhension approfondie et réaliste des enjeux complexes associés à la mise en œuvre d'une solution de virtualisation des données dans un environnement public sensible.

4.3.1. Points forts relevés

La prise en compte systématique des contraintes d'interopérabilité et d'hétérogénéité des données, avec des choix technologiques guidés par la modularité, la compatibilité et l'évolutivité, rassure sur la viabilité technique du projet.

L'attention portée à la sécurité, à la confidentialité et à la qualité des données est essentielle, notamment dans un contexte où la donnée publique nécessite une protection rigoureuse pour garantir confiance et conformité.

La mise en place de processus de maintenance structurés et une perspective claire d'intégration de technologies avancées (comme automatisation DevOps) montre une vision tournée vers la pérennité et l'adaptabilité du système.

4.3.2. Défis identifiés

La gestion du changement et de l'appropriation utilisateur reste un défi, malgré la présence d'un support actif ; cela souligne l'importance d'un accompagnement constant et d'une sensibilisation aux nouvelles technologies, surtout dans un secteur public où les compétences techniques peuvent varier.

Les difficultés liées à la continuité du service et à la gestion des interruptions mettent en relief la nécessité d'un plan de résilience robuste, ce qui est critique pour un système décisionnel à usage stratégique.

En termes de complexité d'intégration technique, il faudra veiller à une coordination efficace entre équipes réseau, bases de données et sécurité, pour éviter les silos ou les risques de faille.

Cette analyse technique complète la perspective stratégique et fonctionnelle des décideurs et utilisateurs, en montrant que la réussite de la virtualisation n'est pas seulement une question d'innovation technologique, mais résulte d'un équilibre délicat entre robustesse technique, sécurité renforcée et gestion du facteur humain.

L'expérience ici simulée met en exergue la nécessité d'une gouvernance technique intégrée, accompagnée d'un pilotage proactif et d'un support organisationnel solide, pour transformer les promesses de la virtualisation en bénéfices opérationnels tangibles.

4.4. Synthèse analytique du déploiement du prototype de la plateforme de virtualisation des données

Nous avons bien marqué le passage fondamental de la beauté de la théorie à la technique concrète, en se consacrant à l'implémentation et à la validation de la solution de virtualisation de données que nous avons conçue pour le Ministère des Finances. Après avoir établi une architecture cible et modélisé les vues de données virtuelles dans le chapitre précédent, nous avons ici concrétisé notre approche par des étapes d'exécution précises et vérifiables.

Nous avons détaillé l'installation et la configuration de la plateforme de virtualisation Denodo Express, confirmant sa capacité à servir de couche d'abstraction unifiée pour des sources de données hétérogènes. L'intégration de bases de données (telles que PostgreSQL) et le développement de vues virtuelles ont démontré la faisabilité technique de notre solution, permettant de consolider l'information sans déplacement physique des données. La connexion de cette couche de virtualisation à des outils de Business Intelligence comme Power BI a permis d'élaborer des maquettes de tableaux de bord décisionnels, prouvant ainsi la capacité à transformer des données complexes en insights actionnables pour les décideurs.

Les objectifs étaient d'installer la plateforme de virtualisation des données **Denodo Express 9**, d'intégrer les différents types de sources (SGBD, fichiers plats : CSV, Json, ...etc), de créer des vues virtuelles et développer des tableaux de bord connectés – ont été atteints avec succès, au mépris de l'utilisation de données génériques pour des raisons de confidentialité. Cette implémentation prototypée valide les avantages théoriques de la virtualisation des données, notamment l'accès unifié à l'information, la réduction de la duplication, une agilité accrue, et une amélioration de la qualité des données, le tout contribuant à l'accélération du processus décisionnel.

En somme, nous avons démontré que la virtualisation des données n'est pas seulement une innovation conceptuelle, mais une solution techniquement réalisable et hautement pertinente pour moderniser le système d'information décisionnel du Ministère des Finances. Les résultats obtenus ouvrent la voie à une amélioration significative de la gouvernance financière et à une prise de décision plus éclairée et réactive au sein de l'administration publique.

Conclusion

Conclusion

Ce travail de recherche a porté sur l'intégration d'une solution de virtualisation de données dans un Système d'Information Décisionnel (SID) hétérogène au sein du Ministère des Finances algérien. Face à la complexité croissante des sources de données, cette étude a proposé de concevoir, modéliser et implémenter une architecture cible basée sur la plateforme Denodo, en vue d'améliorer l'exploitation et la gouvernance des données pour améliorer la prise de décision stratégique.

L'analyse approfondie de l'état actuel du SID a mis en lumière plusieurs défis majeurs, notamment la fragmentation des données, les longs délais d'accès, les difficultés liées à l'intégration physique des données, ainsi que les limitations des approches traditionnelles basées sur les data warehouses et les processus ETL. Ces lacunes nuisent à la réactivité et à la pertinence des décisions ministérielles dans un contexte économique et réglementaire en constante évolution.

La virtualisation des données a été identifiée comme une solution prometteuse permettant d'adresser ces limites. Cette technologie offre une couche d'abstraction unifiée, garantissant un accès en temps réel à des données diverses sans nécessiter de duplication physique, tout en améliorant la flexibilité, la sécurité et la qualité des données.

Les résultats empiriques issus des entretiens et de la mise en œuvre d'un prototype fonctionnel avec Denodo et Power BI ont confirmé cette hypothèse. La virtualisation de données facilite la consolidation, la transformation et l'exploitation des données hétérogènes, permettant aux utilisateurs métiers et aux décideurs d'accéder à des informations précises, fiables et actualisées pour une prise de décision plus agile et informée.

L'étude a également formulé un modèle détaillé d'architecture cible intégrant Denodo comme couche de virtualisation de données, accompagné d'une méthodologie d'implémentation progressive s'appuyant sur des étapes claires, depuis l'installation et la configuration jusqu'à la création des vues virtuelles et des tableaux de bord décisionnels.

Le SID du ministère est fragmenté, marqué par des systèmes cloisonnés, des outils rigides, une faible intégration des sources de données, et une gouvernance des données à consolider.

La virtualisation de données, par sa capacité à fédérer dynamiquement des données diverses, à offrir une vue consolidée en temps réel, à réduire les délais et coûts liés à l'intégration, et à sécuriser l'accès aux données sensibles améliore l'efficacité du système du ministère existant.

Plusieurs défis ont été identifiés lors de l'intégration de la virtualisation des données, notamment sur le plan technique avec l'interopérabilité, la sécurité et la performance, mais aussi au niveau organisationnel avec la gestion du changement et la formation des utilisateurs. Stratégiquement, le pilotage et la gouvernance restent aussi des enjeux importants. Pour la modélisation de l'architecture avec Denodo, on ajoute une couche de virtualisation à l'existant, on crée des vues virtuelles métiers, et on centralise la gestion de la sécurité et des métadonnées. L'implémentation se fait progressivement, avec des phases de prototypage, l'intégration des sources, des tests utilisateurs, et un déploiement accompagné par la formation et un suivi gouverné activement.

Notre étude reste limitée par la portée du prototype, qui s'appuie sur des données génériques et partiellement intégrées. De plus, le degré d'acceptation des utilisateurs à long terme et les impacts organisationnels profonds nécessitent un suivi sur une plus longue durée. Le contexte technique spécifique au ministère, incluant certaines contraintes réglementaires et infrastructures existantes, peut aussi limiter la généralisation immédiate des résultats.

Enfin, l'évolution rapide des technologies de gestion des données invite à considérer cette étude comme une étape dans un processus continu d'amélioration et d'adaptation.

Au-delà de la validation technique, il est essentiel d'approfondir l'accompagnement humain, notamment par des programmes de formation adaptés et la mise en place d'une gouvernance des usages des données. Un suivi rigoureux des indicateurs d'adoption et de performance est aussi crucial pour ajuster la solution en fonction des retours du terrain.

La recherche future pourrait étendre le prototypage à des environnements plus complexes et intégrés, inclure l'intelligence artificielle pour améliorer la qualité des données et l'analyse prédictive, et explorer des modèles hybrides combinant virtualisation, data lakes et entrepôts pour capitaliser sur leurs complémentarités.

Bibliographies

- Aaliyah Jordon, D. A. (2023, January). Data Virtualization: A Game-Changer for Big Data Analytics. (DOI:10.5281/zenodo.8357544).
- APS. (2023, novembre 16). Transformation numérique: l'Algérie réalise de grands progrès. Récupéré sur Algerie presse service: <https://www.aps.dz/sante-science-technologie/162753-transformation-numerique-de-grands-progres-vers-la-transparence-dans-la-gestion-de-la-chose-publique>
- Armstrong, A. (2019). Virtualization in Decision Support Systems: Bridging Data and Analytics. *International Journal of Information Systems*.
- AtScale. (2025, Janvier 28). ETL : Comprendre le Processus Extract Transform Load et l'Intégrer à Votre Système d'Information. Retrieved from [tenorsolutions: https://tenorsolutions.com/qu-est-ce-que-etl/](https://tenorsolutions.com/qu-est-ce-que-etl/)
- Benbasat, I. G. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*.
- Bologa, A.-R. B. (2011, January). A Perspective on the Benefits of Data Virtualization Technology. *Informatica Economica*.
- Boumediene, N. (2024). L'Innovation Digitale dans les Services en Algérie: Pratiques. *Industrial Economics Review*, pp. 270-290.
- Caserta, R. K. (2004). *The Data Warehouse ETL TOOLKit*. Canada & USA: Wiley Publishing.
- Chen, H. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*.
- Christophe Auffray. (avr 2025). Gaia-X et Data Spaces : 2025, année du passage à l'échelle ? TechTarget.
- David Vyvyan, P. D. (2023, janvier 12). [gaiandb/gaiandb](https://github.com/gaiandb/gaiandb). Retrieved from [github: https://github.com/gaiandb/gaiandb](https://github.com/gaiandb/gaiandb)
- denodo. (n.d.). [denodo-platform](https://www.denodo.com/en/denodo-platform/denodo-platform). Retrieved from [denodo: https://www.denodo.com/en/denodo-platform/denodo-platform](https://www.denodo.com/en/denodo-platform/denodo-platform)

- Denodo platform. (2025, janvier 07). Denodo Community. Retrieved from denodo: <https://community.denodo.com/kb/en/view/document/Virtual%20DataPort%20Connectivity>
- Domo resource. (2022, janvier 22). A Guide to Data Federation: Everything you need to know. Retrieved from domo: <https://www.domo.com/learn/article/a-guide-to-data-federation-everything-you-need-to-know>
- Elmasri, R. &. (2015). Fundamentals of Database Systems (7th Edition ed.). Pearson Education.
- Fernandez, A. (2020, février 11). Business Intelligence, Architecture du SID. Retrieved from Piloter.org: <https://www.piloter.org/business-intelligence/business-intelligence.htm>
- GDH. (2024, juin 20). 7 Common Virtualization Challenges – And How to Overcome Them. Retrieved from GDH: <https://gdhinc.com/7-common-virtualization-challenges-and-how-to-overcome-them/>
- GitHub Contributors. (2022). A light-weight data-federation technology built on Apache Derby. Gaian accesses and transforms data securely "at the edge" and can discover other instances of itself to form a scalable data network. Retrieved from GitHub Platform: <https://github.com/gaiandb/gaiandb>
- IBM Software Hub. (2024). Known issues and limitations for Data Virtualization. Retrieved from IBM: <https://www.ibm.com/docs/en/software-hub/5.1.x?topic=issues-data-virtualization>
- Inmon, W. H. (2005). Building the Data Warehouse. Indiana: Wiley Publishing.
- ITU. (2012). Cloud Computing en Afrique Situation et perspectives.
- Janaushek, B. S.-K.-E.-L. (2024). Digitalisation des données de santé en Afrique Libérer le potentiel. AHOP Policy Briefs.
- Kitchin, R. (2014). The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. Sage Publications.
- Lans, R. F. (2012). Data Virtualization for Business Intelligence Systems. United States of America: Morgan Kaufmann Publishers.

- Laudon, K. C. (2015). *Management Information Systems: Managing the Digital Firm*. Pearson. New York: New York University - Stern School of Business.
- M&A. (Dec 27, 2024). *Marché des data rooms virtuelles : Statistiques, tendances et top fournisseurs*. M&A.
- Ministere des Finances / DGNDISIE. (2024-2028.). *Schema directeur Strategique des Systemes d'information* .
- Orlikowski, W. J. (1991). *Studying information technology in organizations: Research approaches and assumptions*. *Information Systems Research*.
- Radio Algérienne. (2023). *Le «Data center» du secteur inauguré le 15 décembre prochain*. Radio Algérienne.
- Serano, W. (2018). *Digital Systems in Smart City and Infrastructure: Digital as a Service*. London: Intelligent Systems and Networks Group, Imperial College London.
- Simmonds, J. (2020). *Data Virtualization: Techniques and Best Practices*. *Journal of Data Management*.
- tenorsolutions. (2025, Mai 14). *ETL : Comprendre le Processus Extract Transform Load et l'Intégrer à Votre Système d'Information*. Retrieved from tenorsolutions: <https://tenorsolutions.com/qu-est-ce-que-etl/>
- Thangaraj, J. M. (2024). *La virtualisation des données unifie les données pour une IA et une analyse transparentes*. IBM, 1.
- webservices by ebs. (s.d.). *Souveraineté numérique en Algérie : Construire un avenir sûr et prospère*. Récupéré sur webservices: <https://webservices.dz/actualites-dz/121-souverainete-numerique-en-algerie-construire-un-avenir-sur-et-prospere>
- What is a data federation? (n.d.). *What is a data federation?* Retrieved from tibco: <https://www.tibco.com/glossary/what-is-a-data-federation/>
- Work, F. (2023, Mai 24). *What is Data Virtualization?* Retrieved from atscale: <https://www.atscale.com/glossary/data-virtualization/>

ANNEXES

Annexes

Guide d'entretien

1. Pour les utilisateurs opérationnels

Expérience et rôle

- Pouvez-vous décrire votre rôle au sein du système décisionnel actuel ?

Difficultés rencontrées

- Quelles sont les principales difficultés que vous avez rencontrées avec système décisionnel actuel ?
- Avez-vous des exemples où le système décisionnel actuel n'a pas répondu à vos attentes ?

Ergonomie et fiabilité

- Trouvez-vous l'interface du système décisionnel actuel est ergonomique et facile à utiliser ?
- Avez-vous déjà subi des interruptions ou des problèmes d'accès aux données ?

Attentes et suggestions

- Quelles améliorations souhaiteriez-vous voir mises en place ?
- Comment la du futur système décisionnel pourrait mieux répondre à vos besoins ?

Bénéfices à percevoir

- Quels avantages concrets l'intégration de la virtualisation de données vous apporte-t-elle dans vos analyses ?
- Pensez-vous que le système sera plus rapide ou plus fiable grâce à cette technologie ?

2 . Pour les décideurs stratégiques

Motivations et objectifs

- Qu'est-ce qui a motivé la décision d'intégrer la virtualisation des données dans le système décisionnel actuel ?
- Quels objectifs stratégiques cherchez-vous à atteindre avec ce projet ?

Critères de succès

- Comment mesurez-vous le succès ou l'efficacité de la virtualisation des données ?
- Quels indicateurs utilisez-vous pour évaluer l'impact sur la prise de décision ?

Enjeux organisationnels

- Quels sont les principaux défis humains et organisationnels liés à ce projet ?
- Comment la gouvernance des données a-t-elle évolué ?

Impacts stratégiques

- En quoi cette technologie modifie-t-elle la manière dont les décisions sont prises ?
- Quels risques ou freins avez-vous dû anticiper ?

Recommandations

- Quelles leçons tirez-vous de cette expérience pour d'autres projets similaires ?
- Quels conseils donneriez-vous pour réussir l'implantation d'une telle solution ?

3 . Pour les experts techniques

Défis d'intégration

- Quelles difficultés techniques majeures avez-vous rencontrées ?
- Comment avez-vous géré les questions d'interopérabilité entre systèmes ?

Gestion et sécurité des données

- Quelles mesures de sécurité et de confidentialité seront mises en place ?
- Comment assurez-vous la qualité et l'intégrité des données virtualisées ?

Maintenance et évolutivité

- Comment est organisée la maintenance et la mise à jour de la plateforme ?
- Quelles sont les perspectives d'évolution technique de la solution ?

Architecture et choix techniques

- Quels critères ont orienté le choix des technologies utilisées ?
- Pouvez-vous décrire l'architecture à mettre en place pour la virtualisation des données ?

Retour utilisateur

- Avez-vous constaté des besoins ou résistances spécifiques chez les utilisateurs ?
- Comment le support technique répond-il à ces besoins ?

Réponses des interviewés

1. Les utilisateurs opérationnels

Expérience et rôle

Cadre 1 DGFIP : "Je supervise les analyses fiscales au sein de la DGFIP. J'utilise quotidiennement le système décisionnel pour collecter et croiser des données provenant de multiples bases afin de produire des rapports pour la prise de décision au niveau ministériel."

Cadre 2 DGFIP : "Mon rôle est d'assurer le suivi des indicateurs de performance financière. Le système actuel permet d'extraire des données mais nécessite souvent des manipulations complexes."

Difficultés rencontrées

Cadre 1 DGFIP : "Le principal problème est la lenteur dans l'exploitation des données, notamment lors de la compilation de rapports impliquant plusieurs sources. Cela impacte la réactivité des analyses."

Cadre 2 DGFIP : "Le système manque parfois de fiabilité, avec des interruptions fréquentes et des difficultés d'accès aux données en temps réel, ce qui nuit à la qualité des décisions prises."

Exemples : "Lors d'une campagne fiscale récente, nous n'avons pas pu obtenir à temps certaines données clés, ce qui a retardé la synthèse et diminué l'efficacité des opérations."

Ergonomie et fiabilité

Cadre 1 DGFIP : "L'interface est fonctionnelle mais peu intuitive, ce qui complique la prise en main pour les nouveaux utilisateurs."

Cadre 2 DGFIP : "J'ai constaté plusieurs cas d'interruptions serveur, particulièrement durant les périodes de forte sollicitation. Cela a coupé l'accès aux données critiques pendant plusieurs heures."

Attentes et suggestions

Cadre 1 DGFIP : "Il serait utile de moderniser l'interface avec des outils plus conviviaux et de prévoir une formation utilisateur régulière."

Cadre 2 DGFIP : "La solution devrait offrir une meilleure continuité de service et permettre un accès plus rapide aux données, sans perte de qualité ou disponibilité."

Suggestions : "Intégrer des fonctions de personnalisation des tableaux de bord, faciliter les croisements de données sans requêtes complexes."

Bénéfices à percevoir avec la virtualisation

Cadre 1 DGFIP : "L'intégration de la virtualisation pourra grandement accélérer le traitement, en offrant un accès virtuel simultané aux bases multiples sans duplication."

Cadre 2 DGFIP : "Nous espérons que le système sera plus fiable et évitera les interruptions précitées. La virtualisation doit également améliorer la qualité des données utilisées, grâce à une gouvernance plus robuste."

2. Les décideurs stratégiques

Motivations et objectifs

1. Sous-Directeur de la coordination des projets de numérisation :

"La décision d'intégrer la virtualisation des données s'est imposée dans le cadre plus large de la modernisation et de la digitalisation du ministère pour répondre aux enjeux de rapidité et de qualité d'accès à l'information dans un environnement de données hétérogènes et volumineuses. Notre objectif stratégique est d'assurer une meilleure gouvernance des données, une infrastructure flexible et agile, tout en soutenant la prise de décision efficace et transparente."

2. Directeur Général de la Numérisation, de la Digitalisation et des Systèmes d'Information Économiques :

"L'objectif est d'optimiser les processus décisionnels en éliminant les silos d'information et en proposant une couche unifiée d'accès aux données sans duplication, facilitant ainsi l'analyse intégrée des données financières. Nous visons également à réduire les coûts opérationnels liés aux systèmes d'information tout en augmentant leur fiabilité et leur sécurité."

Critères de succès

1. Sous-Directeur de la coordination des projets de numérisation :

"Nous mesurons le succès par la réduction du temps d'accès aux données critiques, l'augmentation du taux d'utilisation des outils décisionnels par nos équipes, ainsi que la qualité et la fiabilité des tableaux de bord produits. L'impact sur la prise de décision est suivi via des indicateurs de performance financière et budgétaire démontrant une amélioration des prévisions et une meilleure réactivité."

2. Directeur Général de la Numérisation, de la Digitalisation et des Systèmes d'Information Économiques :

« Le succès sera mesuré à travers la réduction significative du temps d'accès aux données pertinentes, le taux d'adoption par les utilisateurs métiers, ainsi que par l'amélioration des indicateurs de performance budgétaire et fiscale.

Nous suivrons également l'impact sur la qualité des décisions prises, notamment leur rapidité et précision, ainsi que la diminution des erreurs dues à des données obsolètes ou erronées. »

Enjeux organisationnels

1. Sous-Directeur de la coordination des projets de numérisation :

"Le principal défi humain a été l'accompagnement du changement, notamment la formation des utilisateurs aux nouveaux outils et la diffusion d'une culture data-driven. Organisationnellement, il a fallu adapter la gouvernance des données pour intégrer la virtualisation, avec des politiques claires sur la sécurité et la gestion des accès

La gouvernance a évolué vers un pilotage centralisé des données, mieux coordonné entre les directions, avec une responsabilisation accrue et une meilleure traçabilité."

« Les principaux défis humains concernent l'accompagnement du changement, avec la nécessité de former les personnels à ces nouveaux outils et de combattre les résistances liées à la transformation numérique.

L'organisation a dû mettre en place de nouvelles politiques de gouvernance des données, avec une responsabilisation accrue des directions sur la qualité des données et le respect des procédures de sécurité. »

Impacts stratégiques

1. Sous-Directeur de la coordination des projets de numérisation :

"La technologie transforme la manière dont nous prenons des décisions, en favorisant une collaboration inter-services et en fournissant des informations consolidées et fiables. Cela a permis d'adopter une approche plus proactive et fondée sur des preuves concrètes

La virtualisation transforme la prise de décision en rendant possible un travail collaboratif multi-directionnel basé sur une vue unique et partagée des données. Cela permet des ajustements stratégiques plus rapides en réaction aux réalités économiques. »

2. Directeur Général de la Numérisation, de la Digitalisation et des Systèmes d'Information Économiques :

"Les freins anticipés comprenaient des résistances au changement, des craintes sur la sécurité des données, ainsi que la complexité technique à intégrer des systèmes anciens et nouveaux. Ces risques ont été les leviers d'une stratégie d'accompagnement intense

Parmi les risques anticipés figurent la résistance au changement, les contraintes techniques d'interopérabilité et de sécurité, ainsi que la complexité d'intégration avec les systèmes existants. »

Recommandations

1. Sous-Directeur de la coordination des projets de numérisation :

"La leçon principale est qu'une telle intégration ne peut réussir sans un pilotage fort, associant les équipes métiers et techniques dès le départ. La communication régulière et la formation continue sont indispensables

Nous recommandons également de démarrer par des projets pilotes pour ajuster les solutions avant un déploiement massif, et de prévoir des ressources dédiées à la maintenance et à l'évolution de la plateforme."

2. Directeur Général de la Numérisation, de la Digitalisation et des Systèmes d'Information Économiques :

« Impliquer les utilisateurs dès les premières phases du projet est crucial pour garantir l'appropriation.

Il est important de constituer une équipe dédiée à la conduite du changement avec une communication régulière et transparente.

La mise en œuvre doit être progressive, en commençant par des pilotes pour expérimenter avant un déploiement à grande échelle.

Enfin, il faut assurer une veille technologique constante pour anticiper les évolutions et adapter la plateforme en conséquence. »

3. Les experts techniques

1. DB Administrator, Service informatique

-Défis d'intégration

"L'un des défis majeurs a été l'intégration des multiples bases de données aux structures différentes, avec des formats et schémas hétérogènes. La synchronisation en temps réel sans perte de qualité était complexe."

"Nous avons utilisé des middlewares spécialisés pour assurer l'interopérabilité, ainsi que des protocoles d'échange standardisés comme REST et ODBC."

-Gestion et sécurité des données

"Nous avons mis en place un système de contrôle d'accès basé sur des rôles stricts, avec authentification multi-facteurs et chiffrement des données en transit et au repos."

"La qualité des données est assurée par des validations automatiques à l'entrée, des politiques de nettoyage régulier et des audits périodiques de cohérence."

-Maintenance et évolutivité

"La maintenance est organisée en cycles mensuels de mises à jour fonctionnelles et correctives, avec un support en continu."

"Nous envisageons d'intégrer des fonctions de machine learning pour l'optimisation des requêtes et la détection proactive d'anomalies."

-Architecture et choix techniques

"Nous avons choisi des technologies compatibles avec l'ensemble des systèmes existants, basées sur une architecture orientée services (SOA) pour garantir modularité et évolutivité."

"L'architecture comprend une couche de virtualisation qui interagit avec les différents entrepôts et bases de données via des APIs normalisées."

-Retour utilisateur

"Certains utilisateurs éprouvent des difficultés à comprendre les nouvelles fonctionnalités. Le support technique organise régulièrement des sessions de formation et des FAQ interactives pour les accompagner."

2. Ingénieur systèmes et réseaux, Service informatique

-Défis d'intégration

"Garantir la continuité du service lors des phases de basculement et éviter les interruptions a été un challenge clé."

"Nous avons dû déployer des passerelles réseau et des solutions VPN sécurisées pour assurer une interconnexion fiable entre les différentes infrastructures."

-Gestion et sécurité des données

"Nous utilisons des pare-feux avancés, des systèmes de détection d'intrusion (IDS) et des politiques strictes de gestion des correctifs pour protéger l'infrastructure."

"Un monitoring en temps réel est déployé pour surveiller les accès et détecter rapidement toute anomalie."

-Maintenance et évolutivité

"Les mises à jour sont planifiées en heures creuses pour minimiser l'impact, avec une équipe dédiée pour intervenir rapidement en cas d'incident."

"Nous prévoyons d'intégrer progressivement l'automatisation via des outils DevOps pour accélérer le déploiement et la maintenance."

-Architecture et choix techniques

"La virtualisation repose sur des hyperviseurs de dernière génération et des solutions cloud hybrides pour garantir flexibilité et scalabilité."

"L'architecture réseau est conçue pour supporter les flux sécurisés et admettre un fort trafic de données en simultané."

-Retour utilisateur

"Les retours mettent en avant des attentes fortes sur la stabilité et la vitesse d'accès. Le support intervient via un helpdesk 24/7 et des formations spécifiques sont organisées."

3. Responsable sécurité, Service informatique

-Défis d'intégration

"Un grand défi a été d'assurer la sécurité dans un contexte multiplateformes, avec des utilisateurs distants et des applications diverses."

"Nous avons adopté des normes ISO 27001 et déployé une gouvernance forte autour des accès et de la gestion des clés de chiffrement."

-Gestion et sécurité des données

"La confidentialité est garantie par des contrôles d'accès granulaires, l'audit des accès et des journaux de traçabilité complets."

"Des simulations d'attaques régulières sont menées pour tester la résistance du système."

-Maintenance et évolutivité

"Les mises à jour de sécurité sont appliquées de façon prioritaire, avec des contrôles avant et après déploiement."

"Nous travaillons à l'intégration de solutions de sécurité basées sur l'intelligence artificielle pour anticiper les menaces."

-Architecture et choix techniques

"Les critères principaux ont été la robustesse, la conformité réglementaire, et la capacité à gérer un accès sécurisé multi-niveaux."

"L'architecture est segmentée en zones de sécurité distinctes avec passerelles contrôlées."

-Retour utilisateur

"Nous constatons une certaine appréhension à utiliser des applications cloud ; le support organisationnel travaille sur la sensibilisation et la formation à la cybersécurité."