

**Gestion de la data dans les grands groupes : Enjeux, pratiques et perspectives d’amélioration**

**APPRENTI**

Nom : Drin-Douguet Prénom : Auxence

Parcours : Business Intelligence and Analytics

**TITRE DU RAPPORT**

Plan détaillé

**ENTREPRISE**

Nom : Société Générale Securities Services

Adresse : 17 Cr Valmy, 92800 Puteaux

Table des matières

[**Abstract** 3](#_Toc201489772)

[**Introduction** 5](#_Toc201489773)

[**I. Position du problème : Les défis de la gestion de la data dans les grands groupes** 6](#_Toc201489774)

[A. Énoncé du problème 6](#_Toc201489775)

[B. Description du contexte professionnel 8](#_Toc201489776)

[C. Enjeux stratégiques et organisationnels 11](#_Toc201489777)

[**II. État de l’art : Pratiques actuelles et limites des solutions existantes** 13](#_Toc201489778)

[A. Panorama des approches existantes 13](#_Toc201489779)

[B. Limites et critiques des approches actuelles 18](#_Toc201489780)

[C. Positionnement de l’entreprise dans cet écosystème 21](#_Toc201489781)

[**III. Étude originale : Analyse, solutions proposées et recommandations** 24](#_Toc201489782)

[A. Approche personnelle du problème 24](#_Toc201489783)

[B. Propositions concrètes de résolution 26](#_Toc201489784)

[C. Conditions de mise en œuvre 30](#_Toc201489785)

[D. Apports, limites et perspectives 31](#_Toc201489786)

[**Conclusion** 34](#_Toc201489787)

# **Abstract**

This report addresses the critical challenges and opportunities associated with data management in large enterprises, based on my practical experience within Société Générale Securities Services (SGSS). The central focus is the gap between the state-of-the-art data management practices such as cloud infrastructures, data lakes, data warehouses, advanced business intelligence tool and the realities observed in complex organizational environments.

The first part of the report reviews the current landscape of data management solutions widely adopted by major corporations. It highlights the growing shift towards cloud platforms like Microsoft Azure, AWS, and Google Cloud, which offer scalability, flexibility, and cost optimization, while also enabling advanced analytics and machine learning capabilities. The use of data lakes and data warehouses is explored in relation to their complementary roles in storing raw and structured data, supporting exploratory and decision-making processes respectively. Additionally, the democratization of data through intuitive BI tools such as Power BI and Tableau is discussed, alongside the importance of automating data pipelines with modern ETL/DataOps technologies to enhance efficiency and reliability.

Subsequently, the report critically examines the limitations of these technologies when implemented within large, heterogeneous organizations. Key challenges include high implementation costs, integration complexity due to legacy and customized systems, and the difficulty of adopting one-size-fits-all solutions given the diverse needs of various business units and geographic regions. The analysis emphasizes that technical sophistication alone does not guarantee success; organizational and human factors play a decisive role.

Drawing from my observations at SGSS, the report identifies significant gaps in data governance, tool standardization, and process automation. Data fragmentation, often with critical information stored in Excel files derived from incompatible software systems, creates bottlenecks that slow down reporting and decision-making. Moreover, the lack of a unified data lake and inconsistent training and time allocation for employees limit the company’s ability to fully leverage its data assets. The ongoing transition towards a Microsoft-centric ecosystem, including Power BI and Azure, is a promising strategic direction, though its full realization will take several years given the group’s scale.

The core of the report proposes actionable recommendations focused primarily on clarifying and formalizing data governance, standardizing tools and processes, and promoting gradual, high-impact improvements (“quick wins”) that demonstrate tangible benefits to teams while aligning with long-term strategic goals. Critical success factors include strong managerial sponsorship, investment in human capital, and an organizational environment conducive to change, where teams are adequately staffed and not overburdened by economic pressures.

Finally, the report reflects on its own scope and limitations, acknowledging the constraints of a single-entity focus and limited time, while highlighting avenues for future extension. These include expanding the approach to other business units, integrating emerging AI technologies for predictive analytics and anomaly detection, and continuously evolving governance and training programs to sustain transformation efforts.

Overall, this work contributes a practical, context-sensitive framework to support SGSS’s data-driven transformation, emphasizing that beyond technical solutions, the key to success lies in aligned leadership, empowered people, and adaptive processes. The insights and recommendations presented here aim to help the organization harness its data as a strategic asset, improving decision-making agility and cross-functional collaboration.

# **Introduction**

Ce mémoire s’intéresse aux enjeux de la gestion de la donnée dans les grandes organisations, à travers l’exemple concret de mon expérience au sein de Société Générale Securities Services (SGSS). Dans un contexte où la donnée est perçue comme un levier stratégique de transformation, de nombreuses entreprises cherchent à moderniser leurs infrastructures, leurs outils et leurs pratiques pour tendre vers une organisation dite data-driven. Pourtant, cette transition reste semée d’obstacles, à la fois techniques, humains et organisationnels.

La première partie du mémoire dresse un état de l’art des solutions les plus couramment mises en œuvre dans les grandes entreprises : infrastructures cloud (Microsoft Azure, AWS, Google Cloud), entrepôts et lacs de données, outils de visualisation (Power BI, Tableau), automatisation des flux de données via des outils ETL ou DataOps. Ces approches offrent des gains importants en matière d’agilité, de performance analytique et de fiabilité, tout en facilitant l’adoption d’outils d’intelligence artificielle ou de traitement avancé.

Dans une seconde partie, j’analyse les limites concrètes rencontrées lors de la mise en œuvre de ces technologies dans des environnements complexes. Ces limites tiennent souvent à la fragmentation des systèmes, à l’hétérogénéité des logiciels existants, ou encore à la difficulté d’unifier les pratiques dans des groupes de grande taille. J’observe également que les facteurs humains, jouent un rôle central dans les difficultés rencontrées : manque de formation, surcharge des équipes, absence de vision partagée, etc.

Mon analyse de terrain, réalisée au sein de SGSS, révèle des freins persistants : données stockées dans des fichiers Excel issus d’outils non connectés entre eux, absence de gouvernance claire, multiplicité d’outils pour un même usage, ou encore faible automatisation des processus. Toutefois, des évolutions sont en cours, avec un recentrage progressif vers l’écosystème Microsoft (Azure, Power BI), et une volonté affichée de construire une feuille de route commune autour des enjeux 'data'.

À partir de ce constat, le cœur du mémoire propose une série de recommandations concrètes et adaptées au contexte de l’entreprise : clarification de la gouvernance, standardisation des outils, automatisation des flux, accompagnement au changement. Ces propositions visent à poser les fondations d’une culture data plus intégrée et plus efficiente.

Enfin, ce travail m’a permis de développer une meilleure compréhension des enjeux liés à la donnée en entreprise, tant d’un point de vue stratégique qu’opérationnel. Il constitue également, je l’espère, une contribution utile à la réflexion collective de SGSS sur les transformations nécessaires pour mieux exploiter son capital informationnel, au service d’une prise de décision plus rapide, plus fiable, et plus partagée.

# **I. Position du problème : Les défis de la gestion de la data dans les grands groupes**

1. **Énoncé du problème**

Dans un contexte où la donnée constitue un levier majeur de performance, de pilotage stratégique et d’innovation, les grandes entreprises investissent massivement dans des infrastructures technologiques et humaines pour collecter, stocker, traiter et exploiter l’information. Pourtant, malgré ces moyens considérables, la gestion de la data reste un défi complexe, multidimensionnel et parfois mal maîtrisé. Ce paradoxe soulève une problématique centrale : comment expliquer que des organisations dotées de ressources avancées rencontrent encore d’importantes difficultés dans la structuration, l’exploitation et la valorisation de leurs données ?

La croissance exponentielle du volume de données produites, issues à la fois des systèmes internes (ERP, CRM, outils métiers) et des sources externes (réseaux sociaux, open data, partenaires), crée une pression permanente sur les capacités de traitement et de gouvernance. Cette expansion s’accompagne d’une hétérogénéité croissante des formats et des origines des données, qui rend leur intégration difficile. On observe fréquemment une fragmentation des sources de données, avec des silos métiers, géographiques ou techniques, qui empêchent une vision globale et cohérente de l’information à l’échelle de l’entreprise.

À cela s’ajoute une difficulté technique souvent sous-estimée : l’incompatibilité entre certains logiciels ou systèmes d’information, qui rend impossible une communication fluide entre les outils utilisés dans différents départements ou filiales. Cette absence d’interopérabilité entraîne des ruptures dans les flux de données, obligeant les équipes à recourir à des manipulations manuelles, des exports multiples, ou des traitements redondants, augmentant ainsi les risques d’erreur et de perte de qualité de l'information.

De plus, il existe des problèmes d’intégrité et de fiabilité de la donnée. Les erreurs de saisie, les doublons, les données obsolètes ou mal qualifiées sont fréquentes et compromettent la qualité des analyses produites. La collecte des données elle-même peut être défaillante, en raison d’outils inadaptés, de processus manuels non fiabilisés, ou d’un manque de définition claire des responsabilités.

Ces défaillances sont souvent aggravées par des lacunes humaines : de nombreux collaborateurs ne disposent pas des compétences nécessaires pour comprendre les enjeux liés à la qualité et à la gestion des données. Le manque de formation, l'absence de référents identifiés ou encore la faible sensibilisation à l’impact stratégique de la donnée conduisent à des pratiques approximatives. Par ailleurs, la responsabilité individuelle en matière de saisie, de validation ou de mise à jour des données est souvent mal définie, ce qui favorise un certain désengagement. Enfin, la donnée est parfois perçue comme une contrainte technique plutôt qu’un actif stratégique, ce qui limite les efforts collectifs pour en garantir la qualité et la cohérence.

Un autre point critique concerne la gouvernance des données [1], souvent insuffisante ou mal définie. L’absence de stratégie 'data' transverse, le manque de coordination entre les métiers et la DSI, ou encore une culture de la donnée peu développée chez les collaborateurs freinent l’alignement entre les besoins opérationnels et les capacités technologiques [2]. Par ailleurs, les enjeux de sécurité et de confidentialité (notamment en lien avec le RGPD ou la cybersécurité) ajoutent une couche de complexité, parfois difficile à concilier avec les impératifs de performance et d’agilité. Cette situation est d’autant plus complexe dans les entreprises anciennes ou historiques, où les systèmes d’information se sont souvent construits de manière progressive, par couches successives, sans réelle homogénéisation.

Ces organisations, parfois dotées de plusieurs entités (Business Units) ou réparties sur différents pays, doivent composer avec des pratiques hétérogènes, des outils disparates, et des référentiels différents d’un service ou d’un site à l’autre. La transformation digitale dans ces environnements s’apparente souvent à un chantier permanent, où l’intégration de nouveaux outils (cloud, plateformes analytiques, data lakes, etc.) coexiste avec des solutions legacy vieillissantes et peu interopérables. Cette cohabitation accroît la complexité de gouvernance, ralentit la circulation fluide des données et rend difficile l’adoption de normes communes. Elle génère également des résistances au changement, tant sur le plan organisationnel que culturel, freinant l’efficacité des projets de transformation.

En résumé, la problématique générale abordée dans ce mémoire porte sur les difficultés persistantes rencontrées par les grandes entreprises dans la gestion, la structuration et l’exploitation de leurs données, malgré la disponibilité croissante d’outils technologiques performants et la montée en compétence des équipes. Cette problématique s’appuie sur un constat empirique formulé à partir de mon expérience en alternance, mais elle reflète également une réalité plus large observée dans de nombreuses organisations confrontées à des défis similaires, quel que soit leur secteur d’activité ou leur maturité numérique.

Ce mémoire cherchera donc à analyser les causes structurelles de cette situation, à mettre en lumière les solutions existantes — qu’elles soient technologiques, organisationnelles ou méthodologiques — et à proposer des pistes concrètes d’amélioration. L’objectif est d’apporter une contribution utile à la réflexion autour de la gouvernance des données dans les grandes entreprises, en croisant les apports théoriques et les observations issues du terrain.

1. Description du contexte professionnel

Durant mon alternance, j’occupe une fonction au sein de l’équipe de pilotage du département de transformation digitale chez SGSS (Société Générale Securities Services). Mon rôle principal consiste à produire des rapports stratégiques sous Power BI à destination de la direction, à partir d’un ensemble de données provenant de nombreuses entités différentes : comptabilité, gestion des risques, ressources humaines, contrôle de gestion, opérations métiers, etc. Cette mission, en apparence analytique, révèle en réalité toute la complexité de la gestion de la donnée dans un grand groupe international.

Chaque entité dispose de ses propres outils, de ses propres normes de saisie, et de ses propres priorités, ce qui engendre une forte hétérogénéité des sources de données. À ce jour, j'intègre régulièrement plus d’une dizaine de flux de données différents, chacun présentant ses propres spécificités : formats variés (Excel, fichiers plats, extractions d’outils métiers au format propriétaire ?), granularité inégale, mises à jour asynchrones, etc. À cela s’ajoute l’absence généralisée de bonnes pratiques data dans certains services, avec des fichiers mal structurés, des colonnes mal nommées ou des saisies manuelles sujettes à erreurs. Cette situation rend le travail de centralisation et de valorisation particulièrement ardu.

J’utilise des scripts Power Query et des processus semi-automatisés pour nettoyer, transformer et agréger les données, mais cela ne suffit pas toujours. En effet, la communication entre les équipes peut parfois être difficile, notamment lorsqu’il s’agit d’accéder à des jeux de données détenus par d’autres services ou considérés comme sensibles. Certaines données restent peu accessibles ou peu documentées, surtout lorsqu’elles sont produites pour un usage local ou opérationnel, sans vision globale de leur exploitation stratégique. L’absence de référents clairs pour certaines sources rend également l’interprétation délicate : il devient difficile de comprendre le sens exact de certains indicateurs ou leur mode de calcul.

Enfin, la qualité et la pertinence des données varient fortement selon les sources, ce qui oblige à des arbitrages ou à des retraitements importants pour produire des rapports fiables. Il m’arrive fréquemment de devoir recroiser les données avec plusieurs services pour confirmer leur validité ou demander des corrections. Cette réalité quotidienne illustre concrètement les tensions structurelles entre systèmes, métiers et usages de la donnée dans un grand groupe. Elle met aussi en lumière les limites techniques, humaines et organisationnelles que je chercherai à analyser dans ce mémoire.

Par ailleurs, mon implication dans des projets de décommissionnement d’outils au sein de SGSS illustre une autre facette de la complexité de la gestion de la donnée dans un contexte de transformation digitale. Ces projets visent à supprimer ou remplacer des systèmes obsolètes, parfois utilisés depuis de nombreuses années par les équipes, et à basculer les processus associés vers des outils plus modernes, plus sécurisés ou mieux intégrés dans l’écosystème informatique global de l’entreprise. Cette démarche, bien que nécessaire pour limiter les coûts de maintenance, réduire les risques techniques et améliorer la performance, se heurte à de nombreux défis.

Tout d’abord, ces outils anciens sont souvent très imbriqués dans les processus métier. Leur remplacement nécessite une compréhension fine de leurs fonctionnalités, de leurs flux de données, et des usages — parfois informels — que les équipes ont développés au fil du temps. Il faut veiller à garantir la continuité des opérations, tout en évitant les ruptures dans les chaînes de traitement ou les pertes de données critiques. Cette phase de transition implique de cartographier précisément les dépendances, de valider les équivalences fonctionnelles dans les nouveaux systèmes, et de construire des plans de migration solides.

Ensuite, ces projets requièrent un effort transverse important, car ils mobilisent simultanément des acteurs issus de directions variées : équipes métiers concernées, équipes IT, sécurité informatique, conformité, contrôle interne, gestion du changement, etc. Il faut aligner les priorités, synchroniser les actions, gérer les ressources et respecter les contraintes réglementaires, tout en maintenant une communication fluide entre des interlocuteurs aux objectifs parfois divergents. Le manque de coordination ou d’arbitrage clair peut rapidement ralentir, voire bloquer les projets.

Enfin, les freins humains et organisationnels sont particulièrement visibles dans ces initiatives. La résistance au changement est forte, surtout lorsque les utilisateurs ne perçoivent pas clairement la valeur ajoutée du nouveau système ou craignent une perte de confort dans leurs habitudes de travail. Le manque d’adhésion peut venir d’une mauvaise communication, d’un défaut de formation, ou d’une non-implication des équipes concernées dès les phases amont du projet. De plus, l’absence de gouvernance claire, le flou sur les responsabilités, ou encore l’impossibilité d’identifier un interlocuteur référent pour certains périmètres rendent la gestion du changement particulièrement laborieuse.

Ces projets révèlent ainsi la dimension humaine de la gestion de la donnée, trop souvent sous-estimée face aux enjeux techniques. Ils soulignent que la réussite d’une transformation digitale ne repose pas seulement sur le choix d’outils performants, mais sur la capacité à accompagner les usages, à structurer la gouvernance et à construire une dynamique collective autour du changement. Cette expérience nourrit directement ma réflexion dans ce mémoire sur les freins rencontrés dans la valorisation de la donnée et les leviers possibles pour y remédier.

Ce contexte professionnel me permet donc d’observer de manière très concrète les problématiques systémiques liées à la gestion de la donnée dans un grand groupe international, à travers des situations opérationnelles réelles. Mon positionnement transversal me donne une vision d’ensemble des enjeux stratégiques et techniques liés à la transformation data-driven de l’organisation. Cette expérience de terrain nourrit et justifie pleinement le choix de ce thème de mémoire, à la fois pertinent pour mon entreprise d’accueil et représentatif des défis que rencontrent de nombreuses grandes structures aujourd’hui.

1. Enjeux stratégiques et organisationnels

La gestion efficace des données ne constitue pas uniquement un enjeu technique : elle revêt aujourd’hui une dimension stratégique majeure pour les grandes entreprises. Dans un environnement économique de plus en plus concurrentiel, volatile et piloté par l'information, la capacité à exploiter les données de manière pertinente conditionne la performance, l’innovation et la résilience des organisations. Les données sont devenues un actif clé, au même titre que les ressources humaines, financières ou matérielles. Leur bonne gestion influence directement la prise de décision, la personnalisation de l’offre, l’efficacité opérationnelle ou encore la conformité réglementaire.

Sur le plan stratégique, la maîtrise de la data représente aujourd’hui un atout concurrentiel majeur, capable d’influencer durablement la performance et la position d’une entreprise sur son marché. Une organisation qui parvient à centraliser, structurer et analyser ses données dispose d’un avantage considérable pour mieux comprendre son environnement et prendre des décisions éclairées.

Par exemple, une entreprise de distribution qui exploite efficacement les données de ses programmes de fidélité peut anticiper les comportements d’achat, segmenter ses clients de manière fine et proposer des offres personnalisées, augmentant ainsi la satisfaction et la fidélisation. C’est le cas des enseignes comme Decathlon ou Carrefour, qui utilisent la donnée client pour adapter leurs campagnes promotionnelles en fonction des habitudes d’achat locales ou saisonnières.

Dans le secteur industriel, la donnée permet d’optimiser les chaînes de production grâce à la maintenance prédictive, en analysant les données issues des capteurs pour anticiper les pannes et ainsi éviter les arrêts non planifiés. Des groupes comme Schneider Electric ou Airbus s’appuient sur ces approches pour améliorer la performance opérationnelle, réduire les coûts de maintenance et accroître la durée de vie des équipements.

Sur le plan financier et commercial, le pilotage en temps réel des performances à l’aide de tableaux de bord dynamiques permet une meilleure réactivité stratégique, notamment pour ajuster les stocks, redéfinir des priorités d’investissement ou réorienter les actions commerciales. Cela est rendu possible grâce à l’intégration de systèmes BI (Business Intelligence) couplés aux données transactionnelles.

Par ailleurs, la donnée constitue le socle de l’innovation digitale : les entreprises qui investissent dans l’intelligence artificielle (IA), le machine learning ou l’analyse prédictive peuvent développer de nouveaux services à haute valeur ajoutée. C’est le cas des banques, comme la Société Générale, qui utilisent des algorithmes d’analyse comportementale pour détecter les fraudes en temps réel, ou encore des plateformes de streaming comme Netflix, qui s’appuient sur les données de visionnage pour recommander des contenus et produire des séries sur mesure.

Enfin, la capacité à exploiter la donnée est un vecteur de transformation organisationnelle. Elle favorise l’émergence d’une culture orientée “data-driven”, où les décisions sont prises sur la base de faits mesurables plutôt que d’intuitions. Cette approche nécessite toutefois une maturité digitale suffisante, une gouvernance claire et un engagement transversal à tous les niveaux hiérarchiques.

Cependant, pour atteindre ce potentiel, les entreprises doivent surmonter plusieurs défis organisationnels majeurs. Le premier est celui de la gouvernance de la donnée, qui constitue le socle de toute stratégie data efficace. Il ne suffit pas de collecter des données : encore faut-il en garantir la qualité, la cohérence, l’accessibilité et la traçabilité.

Cela implique de définir des rôles clairs — tels que le Chief Data Officer (CDO), responsable de la vision stratégique de la donnée, le Data Owner, garant de la qualité et de la conformité des données dans son périmètre, ou encore le Data Steward, chargé de la gestion opérationnelle et de la documentation. Sans une organisation claire, des processus bien définis et des standards de qualité harmonisés, la gestion de la donnée devient chaotique. On observe alors des effets négatifs : redondance d’informations, erreurs dans les 'reportings', décisions prises sur des bases incohérentes, et une perte de confiance globale dans les systèmes d’information.

Le second enjeu majeur concerne l’alignement entre les équipes métiers et les équipes techniques. Cette collaboration est souvent mise à mal par des visions divergentes et un manque de communication. D’un côté, les équipes IT peuvent se concentrer sur des aspects purement techniques (infrastructure, sécurité, outils), sans nécessairement saisir les usages opérationnels de la donnée.

De l’autre, les équipes métiers formulent des besoins parfois flous, contradictoires ou techniquement irréalistes, faute de compréhension des contraintes de mise en œuvre. Cette fracture organisationnelle peut ralentir les projets, générer de la frustration, voire entraîner des échecs dans la transformation digitale. Il est donc essentiel de développer une culture commune autour de la donnée, en instaurant un langage partagé, des outils collaboratifs, et des processus de co-construction des cas d’usage. Les entreprises sous-estiment encore cet aspect, pensant qu’il suffit d’acheter une solution logicielle pour résoudre leurs problèmes de data. En réalité, le facteur humain est souvent le principal frein/levier de la réussite d’une stratégie data.

Ainsi, les enjeux stratégiques et organisationnels de la gestion des données dépassent largement les considérations techniques. Ils engagent une transformation profonde des pratiques, des rôles, et des mentalités, qui conditionne la compétitivité, la capacité d’innovation et l’agilité des entreprises dans un environnement économique de plus en plus piloté par l’information.

# **II. État de l’art : Pratiques actuelles et limites des solutions existantes**

1. **Panorama des approches existantes**

Face aux défis croissants liés à la volumétrie, à la variété et à la complexité des données, les grandes entreprises ont progressivement mis en place un ensemble de solutions technologiques, méthodologiques et organisationnelles pour mieux structurer, exploiter et sécuriser leurs actifs informationnels. Ces approches, souvent complémentaires, ont pour objectif d’améliorer la qualité de la donnée, sa disponibilité, sa traçabilité et sa valeur ajoutée pour la prise de décision. Cette section propose un panorama des pratiques les plus couramment observées dans les grandes organisations.

**Les infrastructures modernes de stockage et de traitement : Cloud, Data Lake, Data Warehouse**

Les grandes entreprises s’appuient de plus en plus sur des infrastructures cloud (comme Microsoft Azure, Amazon Web Services - AWS, ou Google Cloud Platform - GCP) pour répondre à la complexité croissante de la gestion de leurs données. Ces plateformes offrent une grande souplesse d’utilisation, une scalabilité quasi illimitée et une capacité à centraliser rapidement des données issues de sources multiples, souvent très hétérogènes, qu’elles proviennent de systèmes internes (ERP, CRM, outils métiers) ou externes (API, partenaires, open data, etc.).

L’un des principaux atouts du cloud [3] réside dans la réduction de la dépendance aux infrastructures physiques internes. En externalisant le stockage, le traitement et la sécurisation des données, les entreprises se délestent de la gestion de leurs propres serveurs, ce qui représente un soulagement considérable à plusieurs niveaux :

* Économie sur les coûts matériels et de maintenance : plus besoin d’investir dans des serveurs on-premise coûteux, ni dans leur renouvellement ou leur mise à niveau régulière.
* Moins de charge liée à la cybersécurité : les fournisseurs cloud investissent massivement dans la sécurité (chiffrement, authentification, supervision), et proposent des services intégrés de détection des menaces, souvent hors de portée des moyens internes d’une entreprise.
* Fin de l’obsolescence technologique : les mises à jour des infrastructures et logiciels sont assurées directement par les plateformes cloud, garantissant ainsi un accès permanent aux dernières technologies sans effort de migration lourd.
* Réduction de la charge IT interne : les équipes techniques peuvent se concentrer sur des tâches à plus forte valeur ajoutée plutôt que sur la maintenance opérationnelle de l’infrastructure.

Par ailleurs, le cloud facilite l’intégration de services analytiques avancés (machine learning, data visualisation, automatisation des workflows), disponibles en mode « as-a-service », ce qui accélère les cycles d’innovation et rend les projets plus accessibles, même pour les équipes non spécialisées.

Ainsi, le passage au cloud représente bien plus qu’un simple changement d’infrastructure : il constitue un levier de transformation digitale, allégeant les contraintes techniques et financières des entreprises tout en leur permettant de gagner en agilité, en performance et en résilience [4].

Dans ce contexte, on distingue généralement deux types d’architectures complémentaires :

* Le Data Lake est une solution de stockage capable d’accueillir des données brutes, qu’elles soient structurées, semi-structurées ou non structurées, dans leur format d’origine. Contrairement à un entrepôt de données, il n’impose pas de transformation préalable, ce qui permet une ingestion rapide et flexible. Il est particulièrement adapté aux usages exploratoires ou aux traitements lourds comme ceux liés à l’intelligence artificielle, au machine learning ou à l’analyse prédictive, qui nécessitent de grandes quantités de données variées. De plus, les Data Lakes modernes (comme Azure Data Lake, AWS Lake Formation ou Google BigLake) intègrent des fonctions de gouvernance, de sécurité et de catalogage, facilitant leur intégration dans des architectures data plus globales. Ils représentent ainsi un levier stratégique pour les entreprises souhaitant exploiter tout le potentiel de leur patrimoine data, sans contrainte immédiate de format ou de finalité.
* L’entrepôt de données, ou Data Warehouse, est une solution de stockage conçue pour centraliser et structurer les données provenant de multiples systèmes opérationnels, dans un format homogène et optimisé pour l’analyse. Contrairement au Data Lake, qui stocke des données brutes, le Data Warehouse repose sur des modèles de données rigoureux (souvent en étoile ou en flocon), qui permettent d’organiser l’information de manière logique et compréhensible. Cette structuration facilite grandement les analyses décisionnelles, en rendant les données fiables, cohérentes et directement exploitables par les utilisateurs métiers. L’entrepôt de données sert de socle aux outils de Business Intelligence (BI), qui s’appuient sur lui pour produire des tableaux de bord, des rapports ou des indicateurs de performance. Des solutions comme Snowflake, Google BigQuery, Amazon Redshift ou encore Microsoft Synapse permettent aujourd’hui d’avoir des entrepôts de données performants, évolutifs et intégrés dans des écosystèmes cloud.

De nombreuses entreprises mettent en place des architectures hybrides combinant ces deux approches, dans ce que certains chercheurs appellent aujourd’hui le "data lakehouse", un modèle unifiant les capacités analytiques d’un entrepôt avec la flexibilité d’un data lake (cf. Databricks, Snowflake).

**Les outils d’analyse : BI, IA et automatisation**

Sur le plan de l’analyse, les outils de Business Intelligence comme Power BI, Tableau ou Qlik Sense se sont largement démocratisés. Ils permettent aux équipes métiers de visualiser, explorer et partager les données de manière intuitive, grâce à des tableaux de bord dynamiques et interactifs. Ces solutions remplacent efficacement les fichiers Excel, souvent moins intuitifs, moins puissants et plus sujets aux erreurs, en offrant une meilleure gestion des volumes importants de données et une plus grande fiabilité. Ces outils jouent un rôle central dans la démocratisation de la donnée au sein des entreprises, en rapprochant les utilisateurs finaux des indicateurs clés de performance.

L’usage de l’intelligence artificielle (IA), et notamment de techniques d’analyse prédictive, connaît également une forte progression. Des algorithmes de machine learning sont utilisés pour anticiper les comportements clients, détecter des anomalies, optimiser la supply chain ou encore ajuster les prix en temps réel. Cette évolution ouvre la voie à une automatisation partielle des prises de décision, mais repose sur une condition essentielle : disposer de données de qualité et bien gouvernées.

Par ailleurs, les entreprises adoptent des solutions d’automatisation des flux de données, comme le DataOps et les outils modernes d’ETL/ELT (Talend, Apache NiFi, dbt). Ces technologies permettent de simplifier et d’accélérer l’intégration, la transformation et la distribution des données, tout en garantissant leur qualité et leur traçabilité. L’automatisation réduit les erreurs manuelles, améliore la réactivité face aux besoins métiers et optimise la gestion des données dans des environnements complexes.

**La gouvernance des données : un levier stratégique**

Sur le plan organisationnel, la gouvernance des données est devenue une priorité dans les grandes structures. Elle repose sur la mise en place de frameworks structurés (comme DAMA-DMBOK, DCAM ou les principes FAIR), ainsi que sur la définition de rôles clés :

* Le Chief Data Officer (CDO), qui porte la stratégie data de l’entreprise ;
* Le data owner, garant du contenu métier d’un domaine de données donné ;
* Le data steward, responsable opérationnel de la qualité, de la documentation et de la conformité des données.

Ces rôles sont essentiels pour clarifier les responsabilités, réduire les silos et favoriser un usage cohérent des données à travers les différents métiers. Ils s’appuient également sur des outils de catalogage et de data lineage (ex. : Collibra, Informatica, Alation) permettant de tracer les flux, d’identifier les référents et de documenter les données.

Enfin, les exigences réglementaires, telles que le RGPD (Règlement Général sur la Protection des Données) en Europe, ainsi que les normes internationales comme ISO 27001 ou SOC 2, ont considérablement renforcé l’importance d’une gestion rigoureuse des données. Ces réglementations imposent aux entreprises de maîtriser précisément le cycle de vie des données depuis leur collecte jusqu’à leur suppression et en garantissant leur confidentialité, leur intégrité et leur disponibilité. Elles exigent également un contrôle strict des accès, une traçabilité complète des opérations effectuées sur les données, et une conservation conforme aux durées légales.

Le non-respect de ces règles peut entraîner des sanctions financières lourdes, mais aussi des pertes de confiance auprès des clients et partenaires. Ainsi, au-delà des bénéfices opérationnels et stratégiques, la gestion des données doit désormais intégrer une dimension juridique et éthique forte. Cela conduit les grandes entreprises à investir dans des solutions de gouvernance adaptées, des dispositifs de sécurité robustes, et à développer une culture interne de sensibilisation aux enjeux de protection des données.

En résumé, l’état de l’art révèle un écosystème riche et en constante évolution, dans lequel les entreprises cherchent à tirer parti des dernières avancées technologiques tout en structurant leur gouvernance pour gagner en efficacité, conformité et agilité. Ces solutions ne sont pas isolées, mais s’inscrivent dans une logique complémentaire, où la technologie, l’organisation et la culture doivent avancer de concert. Dans la suite de ce mémoire, je m’attacherai à évaluer dans quelle mesure ces approches peuvent être adaptées ou renforcées pour répondre aux difficultés rencontrées dans les contextes observés.

1. Limites et critiques des approches actuelles

Après avoir présenté les principales solutions utilisées pour la gestion des données, il est essentiel d’en analyser les limites et les difficultés rencontrées en pratique. Ces outils, bien que souvent sophistiqués et puissants, ne sont pas sans défauts, en particulier dans le contexte des grandes entreprises.

**Premièrement**, le coût élevé de certaines technologies peut représenter un frein majeur. L’adoption de solutions cloud avancées, de plateformes d’analyse ou d’outils d’automatisation nécessite des investissements importants, tant financiers qu’humains. Ces dépenses ne garantissent pas toujours un retour sur investissement rapide, surtout lorsque les gains en efficacité sont freinés par des obstacles organisationnels.

Un des principaux freins se situe au niveau de la migration depuis des systèmes existants, souvent anciens et profondément ancrés dans les processus métiers. De nombreuses entreprises utilisent encore des solutions locales, parfois développées en interne, qui, bien qu’obsolètes d’un point de vue technologique, restent perçues comme fiables et familières par les équipes. Passer à une architecture cloud ou à des outils plus récents implique non seulement des coûts techniques (refonte des flux, refactoring, mise en conformité…), mais aussi une remise en question des pratiques établies. Cela génère une forme d’inertie difficile à surmonter.

Ce frein est étroitement lié au facteur humain. Les collaborateurs, habitués à leurs outils, peuvent se montrer réticents face au changement, par crainte de perdre en autonomie, en productivité, ou tout simplement par manque de formation. Dans les grandes structures, où les projets de transformation touchent plusieurs couches hiérarchiques et métiers, le changement technologique devient un projet humain à part entière. Sans accompagnement, sans communication claire, et sans montée en compétences progressive, la résistance peut bloquer ou ralentir significativement l’adoption.

En somme, ce n’est pas seulement la technologie qui est coûteuse, mais la transformation qu’elle induit. Et celle-ci ne peut réussir sans une prise en compte sérieuse des enjeux humains, culturels et organisationnels qui l’entourent.

**Deuxièmement**, la complexité d’intégration demeure un défi constant. Les systèmes existants dans les grandes entreprises sont souvent hétérogènes, anciens et fortement personnalisés. Ils ont été conçus pour répondre à des besoins très spécifiques, ce qui rend leur remplacement ou leur interfaçage avec des solutions modernes particulièrement délicat. La mise en place de nouveaux outils nécessite alors des adaptations lourdes, des phases de test prolongées, ainsi que des compétences pointues en architecture, en interopérabilité et en sécurité — autant de facteurs qui rallongent considérablement les délais de déploiement et augmentent les risques d’échec.

À cela s’ajoute une tendance croissante à la standardisation des solutions proposées sur le marché. Les plateformes modernes (notamment cloud ou SaaS) sont souvent très puissantes, mais beaucoup moins personnalisables que les systèmes historiques. Elles imposent des cadres d’utilisation rigides, dans lesquels il est parfois difficile de faire entrer les spécificités métiers d’une entreprise. Cela force parfois les organisations à adapter leurs processus aux outils, plutôt que l’inverse ce qui peut générer des frustrations ou une perte d'efficacité opérationnelle.

Par ailleurs, l’interopérabilité [4] entre ces nouvelles solutions reste un enjeu. Nombre d’éditeurs conçoivent des écosystèmes fermés, dans lesquels les outils ne fonctionnent de manière fluide que s’ils proviennent du même fournisseur (par exemple, l’ensemble Microsoft, ou l’ensemble Salesforce). Cette logique de compatibilité verticale renforce la dépendance à un acteur unique, limitant la liberté de choix technologique, augmentant les coûts à long terme, et rendant plus difficile l’évolution de l’architecture système de manière modulaire. Dans un contexte où l’agilité est essentielle, cette dépendance peut devenir un véritable verrou stratégique.

**Troisièmement**, les besoins spécifiques et variés des différentes entités, métiers ou zones géographiques rendent difficile l’adoption d’une solution unique et universelle. Chaque département, chaque filiale ou chaque région peut avoir ses propres objectifs, contraintes réglementaires, niveaux de maturité numérique ou habitudes de travail. Une solution centralisée, même performante sur le papier, peut ainsi s’avérer inadaptée à certaines réalités locales ou métiers [5]. Par exemple, un outil de reporting conçu pour les besoins du siège peut ne pas convenir aux exigences opérationnelles d’une équipe en charge du suivi de production à l’international, ou aux attentes d’un service RH gérant des données sensibles dans un pays soumis à une législation particulière sur la confidentialité.

Ce manque de flexibilité peut non seulement limiter l’efficacité des outils déployés, mais aussi générer de la frustration, une perte d’engagement et, dans certains cas, un rejet pur et simple de la solution. Certains utilisateurs contournent même les systèmes officiels en créant leurs propres fichiers Excel ou tableaux de bord locaux, ce qui nuit à l’uniformité et à la fiabilité des données.

Cela soulève un dilemme classique dans les grands groupes : faut-il privilégier une approche centralisée, standardisée mais rigide, ou encourager une certaine autonomie locale au risque de perdre en cohérence globale ? Ce compromis entre homogénéité et personnalisation reste l’un des défis majeurs de la gouvernance data à l’échelle d’une organisation complexe.

**Enfin**, un point critique souvent sous-estimé est la rapidité fulgurante des évolutions technologiques dans le domaine de la gestion des données. Le marché voit régulièrement émerger de nouveaux outils, standards ou paradigmes qu’il s’agisse de solutions cloud-native, de frameworks d’orchestration de pipelines, ou de modèles d’intelligence artificielle générative appliqués à la donnée. Cette dynamique impose aux entreprises un effort constant de veille, d’expérimentation et de mise à jour de leurs systèmes pour rester compétitives.

Or, dans les grands groupes, les cycles de décision, de validation budgétaire et de déploiement sont souvent longs et complexes. Entre l’identification d’un besoin et la mise en œuvre effective d’une solution, plusieurs mois voire plusieurs années peuvent s’écouler. Ce décalage structurel entre la vitesse du marché technologique et la lenteur des processus internes crée un fossé. Les outils effectivement utilisés sur le terrain peuvent ainsi être déjà dépassés au moment même où ils sont généralisés. Cela freine l’innovation, alourdit la maintenance (notamment lorsqu’il faut continuer à gérer des solutions hybrides ou obsolètes), et réduit l’agilité de l’organisation.

De plus, cette instabilité technologique peut renforcer la résistance au changement : certains décideurs ou utilisateurs redoutent d’investir dans une solution qui risque de devenir rapidement obsolète, ou de devoir se réadapter en permanence à de nouveaux outils. Cela accentue la nécessité d’une vision stratégique forte en matière de data management, avec des choix technologiques cohérents et durables, mais aussi une culture d’adaptation continue au sein des équipes.

En s’appuyant sur des études de cas sectorielles et des retours d’expérience, il apparaît que malgré leur sophistication, ces solutions ne résolvent pas toujours complètement les problématiques de gouvernance, de qualité et de fiabilité des données. La technologie ne suffit pas : la dimension humaine, les processus, la culture d’entreprise et la stratégie globale jouent un rôle tout aussi déterminant.

1. Positionnement de l’entreprise dans cet écosystème

Au sein du département de la transformation digitale de SGSS (Société Générale Securities Services), plusieurs initiatives ont été lancées pour moderniser la gestion des données et tendre vers les standards du marché. Cela se traduit notamment par le déploiement progressif d’outils de Business Intelligence, la rationalisation des sources de données, et une attention croissante portée à la qualité des indicateurs utilisés par la direction.

Mon rôle au sein d’une équipe transverse, en interaction avec de nombreux métiers (finance, risques, RH, IT…), me place à un poste d’observation privilégié pour identifier les limites structurelles qui freinent cette transformation. Ce positionnement me permet de constater, de manière récurrente, des écarts importants entre les ambitions stratégiques et les pratiques opérationnelles : architecture fragmentée, silos applicatifs, outils redondants, extractions manuelles, absence de référent pour certaines données, ou encore manque de documentation.

Malgré une volonté affirmée de devenir plus “data-driven”, l’organisation reste fortement marquée par son héritage technologique, avec une coexistence de nombreux outils anciens et spécifiques à chaque entité. Cela complique l’harmonisation des pratiques et la mise en place de solutions globales. Par ailleurs, les freins humains et organisationnels — comme la résistance au changement, le cloisonnement des services ou la faible acculturation à la donnée — viennent accentuer ces difficultés, rendant la transformation plus lente qu’espérée.

**Solutions et choix technologiques en place**

L’outil principal utilisé dans mon équipe est Power BI, largement déployé pour concevoir des rapports et tableaux de bord destinés à la direction. Cette solution de Business Intelligence s’appuie sur un langage de requêtage (Power Query, DAX) qui permet de centraliser, transformer et visualiser les données issues de nombreuses sources hétérogènes : comptabilité, contrôle de gestion, risque, ressources humaines, gestion de projet, etc. Les connecteurs Power BI sont utilisés pour interroger, des fichiers Excel, ou des exports automatisés depuis des outils internes.

En matière de stockage et de circulation des données, la situation au sein de l’entreprise reste marquée par une grande hétérogénéité des pratiques. Bien que certaines équipes s’appuient ponctuellement sur des entrepôts de données gérés par la DSI ou les équipes data, la majorité des flux de données transitent encore par des fichiers Excel issus d’extractions manuelles depuis divers logiciels métiers.

Ces fichiers sont souvent le fruit de traitements locaux, sans standardisation ni automatisation, ce qui engendre de nombreux problèmes d’intégration avec des outils comme Power BI. Par exemple, certaines données comptables sont extraites depuis deux outils comptables différents, avec des formats, des périodicités et des référentiels distincts. Ces écarts obligent à un travail de retraitement important avant toute consolidation, avec des risques d’erreur accrus et un temps de traitement allongé.

De plus, cette dépendance à des fichiers Excel rend la chaîne de valeur de la donnée particulièrement fragile et chronophage : il suffit qu’un export change de structure, ou qu’un collaborateur modifie manuellement une cellule, pour perturber l’ensemble du processus en aval.

Enfin, il n’existe pas de data lake transversal à l’échelle de l’organisation, capable d’unifier les données structurées, semi-structurées et non structurées provenant des nombreux logiciels utilisés dans les différents services (comptabilité, risques, RH, pilotage, etc.). Cette absence limite fortement les capacités d’exploration, d’industrialisation et de croisement des données, en particulier pour des usages plus avancés comme le machine learning ou les analyses prédictives.

La gouvernance de la donnée repose encore sur des pratiques très locales. Certains jeux de données critiques ne disposent d’aucun référent métier identifié ni d’une documentation claire, rendant leur exploitation difficile. De même, l’absence de standards communs (qualité, définition des indicateurs, formats) crée des incohérences lors de la consolidation dans les rapports.

**Écarts avec les meilleures pratiques du marché**

Plusieurs écarts apparaissent lorsqu’on compare cette situation aux meilleures pratiques observées dans l’état de l’art :

* Absence de gouvernance formalisée : aucun framework structurant (type DAMA, DCAM) n’est actuellement déployé pour encadrer la qualité et la circulation de la donnée à l’échelle de l’entité.
* Systèmes anciens et cloisonnés : certaines sources de données critiques résident encore dans des outils vieillissants, sans interconnexion native avec les systèmes analytiques modernes. Cela alourdit les processus d’extraction et complexifie leur automatisation.
* Manque d’automatisation des flux : les pipelines d’intégration et de transformation de données sont souvent manuels ou semi-automatisés (extractions Excel, scripts Power Query), ce qui limite leur fiabilité et leur traçabilité.
* Culture data encore en développement : bien que la volonté de “démocratiser” la donnée soit présente, les utilisateurs métiers n’ont pas tous le même niveau de maturité face aux outils analytiques, ce qui freine l’adoption des tableaux de bord.

**Marges de progression identifiées**

Plusieurs leviers d’amélioration sont identifiables :

* Structurer une gouvernance data claire et plus globale, avec des rôles dédiés (data owner, data steward) et des règles de qualité explicites.
* Renforcer l’automatisation des flux grâce à des outils ETL ou DataOps (type Talend, dbt, etc.) pour fiabiliser et industrialiser l’intégration des données.
* Favoriser la convergence des systèmes vers des plateformes communes ou des architectures hybrides plus souples (cloud, data lake, API).
* Investir dans la montée en compétences des collaborateurs, en développant la culture data, la maîtrise des outils BI et la compréhension des enjeux liés à la donnée.
* Mettre en place un dictionnaire des données pour uniformiser les définitions et faciliter leur compréhension et réutilisation dans les différents départements.

# **III. Étude originale : Analyse, solutions proposées et recommandations**

1. **Approche personnelle du problème**

D’après mon expérience au sein de SGSS et du groupe SG en général, la principale cause de blocage dans la gestion des données ne réside pas dans la technologie, mais dans le manque de vision claire et partagée sur la stratégie data à l’échelle de l’organisation. Cette absence de direction structurée se manifeste à plusieurs niveaux, et génère un effet domino sur les processus, les outils et surtout les acteurs impliqués.

Tout d’abord, l’absence de leadership fort sur la donnée crée une zone floue quant aux responsabilités. Faute de gouvernance centralisée ou de roadmap unifiée, chaque équipe ou entité développe ses propres pratiques, souvent en dehors de tout cadre global. Cela se traduit par une fragmentation des outils, une redondance des efforts, et des incohérences dans les indicateurs ou les définitions utilisées.

Ce flou stratégique a un impact direct sur le facteur humain. Les collaborateurs, bien que souvent engagés et de bonne volonté, ne disposent ni du temps ni de l’accompagnement nécessaire pour intégrer pleinement la culture de la donnée dans leur quotidien. Faute de formation adaptée et de reconnaissance institutionnelle, la donnée est encore perçue par beaucoup comme une charge supplémentaire plutôt qu’un véritable outil d’aide à la décision. Cette situation crée une forme de décrochage : les équipes se sentent peu concernées par les enjeux data ou insuffisamment outillées pour les aborder avec sérénité. Cela freine l’émergence d’une dynamique collective autour de la donnée, limite les initiatives locales et réduit la capacité à enclencher un véritable processus d’amélioration continue.

S’ajoute à cela un enjeu réglementaire non négligeable [6]. SGSS opérant dans plusieurs pays, les contraintes de sécurité, de confidentialité et de conformité varient fortement selon les juridictions. Par exemple, les politiques locales en matière de cybersécurité ou de protection des données peuvent empêcher l’adoption de solutions cloud SaaS mutualisées, pourtant plus agiles et pérennes. Ce contexte freine l’harmonisation des outils et complexifie la mise en œuvre d’une architecture data moderne à l’échelle internationale.

Une orientation stratégique claire se dessine au sein du groupe, avec une volonté affirmée de consolider l’ensemble des outils bureautiques et data autour de l’écosystème Microsoft. L’adoption croissante de Power BI pour la visualisation et d’Azure comme infrastructure cloud de référence marque une étape importante vers une meilleure homogénéisation des pratiques et des technologies. Ce choix offre de nombreuses perspectives : il facilite l’intégration des outils entre eux, renforce la sécurité, et simplifie la gestion des environnements informatiques à l’échelle du groupe. Bien que cette transition prenne du temps, notamment en raison de la taille de l’organisation et de la diversité des systèmes existants, elle constitue une avancée structurante. Le déploiement progressif de ces solutions permet d’accompagner le changement en douceur, tout en jetant les bases d’une gouvernance plus cohérente et d’une exploitation plus efficace des données à moyen terme.

Dans ce contexte, il est essentiel que la transformation vers un écosystème unifié soit pilotée de manière ferme et structurée. Un pilotage fort permet non seulement d’aligner les différentes entités autour d’une vision commune, mais aussi d’imposer progressivement des standards de qualité, de sécurité et de gouvernance. Instaurer de bonnes pratiques dès les premières étapes du déploiement est crucial pour garantir la cohérence et la pérennité des solutions mises en place. Par ailleurs, une approche centralisée et rigoureuse de cette transition technologique peut générer des économies significatives : réduction des coûts de maintenance liés à la multiplicité des outils, mutualisation des compétences, et rationalisation des licences ou des infrastructures. Au-delà des gains financiers, cela permet aussi d’augmenter l’agilité de l’organisation et de mieux répondre aux exigences réglementaires et opérationnelles.

En croisant ces constats avec les analyses de l’état de l’art, on peut formuler plusieurs hypothèses explicatives :

* Hypothèse 1 : Sans une stratégie data claire portée au plus haut niveau, les initiatives restent isolées, manquent de cohérence, et peinent à s’inscrire dans la durée.
* Hypothèse 2 : L’acculturation des collaborateurs à la donnée est une condition indispensable à la réussite, au même titre que les outils techniques ou les budgets.
* Hypothèse 3 : L’absence de standardisation et les contraintes réglementaires locales rendent difficile le déploiement de solutions scalables à grande échelle.

Ces hypothèses mettent en lumière les faiblesses structurelles sur lesquelles il est crucial d’agir. Cela passera, comme je le développerai dans la suite du mémoire, par la clarification du pilotage stratégique, la montée en compétences des équipes, et une meilleure articulation entre contraintes locales et objectifs globaux.

1. Propositions concrètes de résolution

Face aux constats dressés dans la partie précédente, plusieurs leviers d’action peuvent être mobilisés de manière réaliste pour améliorer la gestion de la donnée au sein de SGSS. Ces recommandations visent à répondre aux enjeux structurels identifiés tout en tenant compte des spécificités de l’organisation : sa taille, son fonctionnement en silos, ses contraintes réglementaires et son environnement technologique en évolution.

1. **Renforcer la gouvernance des données**

Le levier le plus structurant — et probablement le plus décisif — concerne la clarification de la gouvernance des données. Sans cadre partagé, même les meilleures solutions techniques ou les efforts locaux restent fragmentés et inefficaces à grande échelle. Il devient donc impératif de formaliser une stratégie data transverse, soutenue par la direction, et traduite en une feuille de route claire et opérationnelle, co-construite entre les grandes fonctions de l’entreprise (métier, IT, conformité, sécurité…).

Cette gouvernance constitue le socle sur lequel reposeront la qualité, la cohérence et la valorisation des données dans l’ensemble du groupe. Elle peut se concrétiser à travers :

* La désignation de rôles clairs (Data Owners, Data Stewards) dans chaque direction, avec des responsabilités définies en matière de qualité, documentation et accessibilité des données.
* La création d’un comité de gouvernance des données transverse, réunissant les parties prenantes clés pour assurer le pilotage, l’alignement des priorités et la gestion des arbitrages.
* L’instauration de standards communs : définitions des indicateurs, référentiels de qualité, procédures de validation ou de modification des données critiques.

Une gouvernance bien posée n’est pas une contrainte bureaucratique : c’est une condition de réussite pour toute transformation data-driven. Elle permet de dépasser les silos, de réduire les doublons, et de créer un langage commun autour de la donnée, indispensable pour faire évoluer la culture de l’organisation.

1. **Mieux accompagner les collaborateurs**

Un autre levier fondamental, trop souvent sous-estimé, est l’acculturation et la montée en compétences des équipes. Si l’on veut que la donnée devienne un véritable levier de pilotage et d’amélioration continue, encore faut-il que les collaborateurs s’en sentent légitimes, outillés et responsables. Cela suppose de dépasser la vision technicienne ou centralisée de la donnée pour en faire une compétence partagée à tous les niveaux.

Il s’agit certes d’un investissement en temps et en ressources, mais c’est un effort qui en vaut largement la peine. Une équipe formée et engagée autour des enjeux data est non seulement plus autonome, mais aussi plus à même de produire des analyses fiables, d’interpréter correctement les indicateurs, et de remonter des alertes pertinentes.

Concrètement, cet investissement pourrait se traduire par :

* Des parcours de formation personnalisés autour des données, adaptés aux niveaux et aux besoins des différents profils : compréhension des outils BI (Power BI, Excel), lecture et interprétation des KPI, bonnes pratiques de modélisation, introduction à la gouvernance des données, etc.
* La valorisation des initiatives locales : en donnant de la visibilité à ceux qui développent des outils utiles ou partagent des pratiques efficaces, on stimule l’émulation interne et la diffusion d’une culture data.
* L’inscription d’objectifs liés à la donnée dans les projets d’équipe ou les entretiens annuels, pour ancrer durablement cette culture dans les pratiques quotidiennes.

Former, reconnaître et responsabiliser les utilisateurs est une condition de réussite incontournable pour toute stratégie data ambitieuse. Cela transforme la donnée d’un sujet perçu comme “technique” en un outil concret, valorisant et stratégique pour les collaborateurs comme pour les managers.

1. **Harmoniser les outils et simplifier les flux**

Afin de dépasser la fragmentation des systèmes et de fiabiliser la chaîne de traitement, il conviendrait de poser les bases d’un socle technique plus robuste et cohérent. Cela passe par plusieurs actions concrètes :

* Standardiser les outils de reporting autour de Power BI, en favorisant les connexions directes et sécurisées avec les sources existantes (bases SQL, outils comptables, RH…). Cela permettrait de réduire les manipulations manuelles, d’unifier les indicateurs et de gagner en traçabilité.
* Limiter l’usage d’Excel comme outil pivot dans les chaînes de production de données. Bien qu’utile à des fins ponctuelles, Excel devient vite fragile dès lors qu’il est utilisé pour agréger, transformer ou diffuser des données critiques. Il serait plus pertinent d’automatiser les flux via Power Query, Azure Data Factory ou des solutions ETL internes déjà présentes dans l’environnement SI.
* Identifier les extractions récurrentes (reportings mensuels, KPI réglementaires, tableaux de bord opérationnels…) et les intégrer dans un socle de données partagé, validé, documenté et maintenu. Ce référentiel commun, géré par les équipes IT en lien avec les métiers, permettrait de fiabiliser les analyses tout en réduisant les doublons.

Ces mesures contribueraient à fluidifier les processus, à réduire les erreurs et à rendre l’information plus accessible et exploitable dans un cadre de gouvernance maîtrisé.

1. **Définir des indicateurs de suivi**

Pour mesurer l’impact des changements, plusieurs KPI de maturité data [4] peuvent être suivis certains sont déjà mis en place mais nécessitent plus de visibilité :

* Le taux de couverture des données par des référents identifiés.
* Le nombre de rapports Power BI connectés directement à des sources certifiées.
* Le pourcentage de collaborateurs formés à un socle minimum de compétences data.
* La fréquence des mises à jour de la documentation des données critiques.
* Le taux de réutilisation des jeux de données centralisés.

1. **Adapter les ambitions au rythme de l’organisation**

Enfin, il est crucial d’adopter une approche progressive et pragmatique, en privilégiant les “quick wins” qui apportent des améliorations concrètes et rapides au quotidien des équipes [4]. Par exemple, réduire le temps consacré à la production des rapports, simplifier la lecture et la compréhension des tableaux de bord, ou encore automatiser certaines tâches manuelles peuvent générer un impact positif immédiat. Ces gains rapides permettent de mobiliser l’adhésion des collaborateurs, de renforcer leur confiance dans les nouveaux outils et processus, et de créer une dynamique vertueuse.

Dans le même temps, il faut conserver une vision stratégique à long terme, qui s’inscrit dans les orientations globales du groupe. En particulier, l’évolution vers un écosystème Microsoft plus intégré (Power BI, Azure, Teams, SharePoint, etc.) offre des opportunités fortes de cohérence, d’interopérabilité et de scalabilité. Construire les fondations dès aujourd’hui, avec une feuille de route claire, facilitera l’adoption progressive des nouvelles solutions et assurera une meilleure maîtrise des données dans la durée.

1. Conditions de mise en œuvre

La réussite des propositions formulées repose sur des conditions clés, où le rôle managérial et la prise de décision stratégique jouent un rôle central. En effet, pour qu’une transformation data réussisse, il est impératif que les directions générales et les managers soient pleinement engagés. Ce sont eux qui doivent impulser ces initiatives, porter une vision claire et cohérente, et arbitrer les choix nécessaires. Leur implication active est un levier décisif pour mobiliser les équipes, sécuriser les budgets et assurer la cohérence transverse entre les métiers, l’IT et la conformité.

Dans ce cadre, il est également essentiel que les conditions de travail soient optimales. Les équipes en charge de la donnée ne peuvent donner le meilleur d’elles-mêmes que si elles ne sont pas en sous-effectif et si elles évoluent dans un environnement où la pression liée à la conjoncture économique est maîtrisée. Une surcharge de travail ou un climat anxiogène freinent la créativité, la qualité du travail et la volonté d’adhérer aux changements. Il est donc nécessaire d’investir pour garantir un cadre favorable, avec des ressources humaines suffisantes et un équilibre sain entre les charges opérationnelles et les projets d’amélioration.

Sur le plan des ressources humaines, l’acquisition de profils spécialisés, qu’il s’agisse de 'data analysts', 'data engineers' ou référents gouvernance, est une condition nécessaire. Ces compétences pointues permettent d’accompagner efficacement la montée en puissance des outils, d’assurer la qualité des données et de jouer un rôle de relais entre la technique et le métier. Par ailleurs, un investissement soutenu dans la formation et le développement des compétences internes est indispensable pour pérenniser les acquis et favoriser l’appropriation des nouveaux processus.

Enfin, cette dynamique managériale doit s’accompagner d’une gouvernance claire et structurée, qui délimite les responsabilités et assure un pilotage rigoureux. Seule une coordination forte et une prise de décision rapide permettront de dépasser les inerties organisationnelles et de transformer les initiatives en résultats concrets.

En somme, au-delà des aspects techniques, c’est avant tout un engagement fort des dirigeants, un environnement de travail propice et un investissement humain adapté qui constitueront les fondations nécessaires à la réussite de la transformation data.

1. Apports, limites et perspectives

**Apports :**

Ma démarche a permis de mettre en lumière plusieurs enjeux clés qui, bien que souvent évoqués, restent trop fréquemment sous-estimés ou insuffisamment adressés dans les grandes organisations. Parmi eux, l’importance cruciale d’une gouvernance claire et transverse ressort comme un levier fondamental. Sans une définition précise des rôles, des responsabilités et des processus, les initiatives data peinent à se déployer efficacement, ce qui génère des silos, des redondances et un manque de cohérence.

Par ailleurs, la standardisation des outils et des pratiques apparaît comme une nécessité pour rationaliser les efforts, réduire les coûts et faciliter l’interopérabilité des systèmes. La diversité des outils utilisés aujourd’hui, souvent héritée de choix historiques ou de besoins spécifiques non centralisés, complexifie la maintenance et ralentit les processus décisionnels.

Enfin, une meilleure intégration des flux de données, depuis leur collecte jusqu’à leur mise à disposition dans des formats fiables et exploitables, est indispensable pour garantir la qualité et la pertinence des informations utilisées. Ce travail de consolidation et de validation est la pierre angulaire d’une exploitation efficace des données.

En proposant des recommandations concrètes et pragmatiques, adaptées au contexte particulier de SGSS, ma démarche offre des pistes opérationnelles pour surmonter ces obstacles. Elle vise non seulement à améliorer la qualité, la fiabilité et la cohérence des données disponibles, mais aussi à renforcer la confiance des équipes métiers dans les outils et les rapports qu’elles utilisent quotidiennement.

Ainsi, cette approche contribue à faire de la donnée un véritable levier stratégique pour l’entreprise, en facilitant une prise de décision plus rapide, plus éclairée et mieux coordonnée entre les différentes directions. Elle pose les bases d’une culture data plus mature et partagée, condition sine qua non pour accompagner la transformation digitale durable de SGSS.

**Limites :**

Toutefois, ma contribution présente plusieurs limites qu’il convient de souligner pour bien situer l’apport et les contraintes de cette analyse.

D’une part, le périmètre de mon étude est essentiellement circonscrit à mon expérience au sein d’une seule entité, le département de pilotage de la transformation digitale chez SGSS. Cette focalisation, bien qu’offrant une vision précise et concrète du terrain, restreint la portée des conclusions. En effet, les pratiques, les outils et les problématiques peuvent varier significativement d’une entité à une autre au sein du groupe, notamment en raison des différences métier, géographiques ou organisationnelles. Il est donc possible que certains constats ne soient pas entièrement transposables à l’ensemble du groupe Société Générale.

D’autre part, les contraintes de temps liées à la durée limitée de mon alternance ont restreint la profondeur de certaines investigations. Par exemple, l’analyse technique détaillée des systèmes existants, ainsi que l’évaluation quantitative des impacts des solutions proposées (en termes de gains de productivité, réduction des erreurs ou amélioration de la qualité des données) n’ont pas pu être pleinement réalisées. Ces aspects demanderaient des études complémentaires et plus longues pour valider les hypothèses émises.

Enfin, il faut noter que certains éléments stratégiques, organisationnels et technologiques sont encore en phase d’évolution au sein de l’entreprise. La transformation digitale, notamment dans un grand groupe comme SGSS, est un processus dynamique, marqué par des ajustements constants, des réorientations et des expérimentations. Cette situation rend difficile une évaluation complète et définitive de la maturité data et des solutions optimales à ce stade précis.

Ces limites ne remettent pas en cause la pertinence générale des propositions, mais invitent à les considérer comme des pistes à approfondir et à ajuster dans un cadre évolutif et plus global.

**Perspectives :**

À plus long terme, cette démarche offre un cadre solide pouvant être étendu à d’autres entités ou filiales du groupe, ce qui permettrait d’harmoniser les pratiques autour de la gestion des données et d’instaurer une gouvernance data globale, cohérente et unifiée. Une telle approche favorise la mise en place de standards communs, facilite la collaboration inter-équipes et améliore la qualité ainsi que la fiabilité des données à l’échelle de l’ensemble du groupe.

Par ailleurs, l’intégration progressive de technologies émergentes constitue une véritable opportunité d’innovation. Des solutions d’intelligence artificielle, par exemple, pourraient automatiser la détection d’anomalies dans les données, anticiper les risques grâce à l’analyse prédictive, ou encore optimiser les processus décisionnels via des recommandations intelligentes. Ces avancées technologiques ouvrent la voie à des gains d’efficacité et à une valorisation accrue des données.

Cependant, la réussite de cette transformation repose également sur un développement continu des compétences humaines. Il est essentiel de renforcer les formations et l’accompagnement des collaborateurs afin de faciliter l’adoption des nouveaux outils et méthodes. Le management du changement doit rester au cœur des stratégies, en instaurant un dialogue constructif avec les équipes, en valorisant les bonnes pratiques et en répondant aux préoccupations liées aux évolutions organisationnelles.

Enfin, pour garantir la pérennité et l’efficacité des initiatives engagées, la mise en place d’indicateurs de suivi précis est indispensable. Ces indicateurs permettront de mesurer régulièrement l’impact des actions, d’identifier rapidement les écarts ou dysfonctionnements, et d’ajuster les stratégies en conséquence. Cette démarche d’amélioration continue assurera que la transformation data ne soit pas un simple projet ponctuel, mais un levier durable de performance et d’innovation pour l’entreprise.

# **Conclusion**

En conclusion, ce mémoire a permis de mener une réflexion approfondie sur les défis et opportunités liés à la gestion des données au sein d’une grande organisation comme SGSS. Après avoir dressé un état des lieux des solutions technologiques et méthodologiques actuelles dans le domaine de la data, j’ai pu analyser les forces et les limites de ces approches dans un contexte concret, mettant en lumière les freins humains, organisationnels et techniques qui ralentissent la transformation digitale.

La problématique centrale, centrée sur l’importance d’une gouvernance claire et d’une meilleure intégration des outils et des flux de données, s’est révélée particulièrement pertinente au regard des enjeux stratégiques actuels. L’analyse théorique, enrichie par des observations de terrain, a permis d’identifier des leviers d’amélioration concrets, notamment la formalisation d’une stratégie data transverse, la standardisation des outils, et l’adoption progressive de solutions innovantes. Ces propositions, ancrées dans la réalité de l’entreprise, offrent une feuille de route pragmatique visant à améliorer la qualité, la fiabilité et la pertinence des données exploitées par les équipes décisionnelles.

Au-delà de l’aspect purement technique, ce travail m’a également permis de développer une compréhension approfondie des dimensions humaines et managériales essentielles à la réussite des projets data. La nécessité d’un engagement fort des dirigeants, d’un accompagnement adapté des collaborateurs et d’un environnement de travail favorable s’est imposée comme un facteur clé. Cette prise de conscience nourrit une vision plus stratégique et globale de la gestion de l’information, qui dépasse la simple mise en œuvre d’outils pour embrasser une transformation culturelle et organisationnelle.

Enfin, ce mémoire, bien que modeste dans son ampleur, constitue pour l’entreprise un socle de réflexion utile pour poursuivre sa transition vers une organisation data-driven. En proposant des solutions applicables et mesurables, il cherche à contribuer, dans une certaine mesure, à renforcer la capacité de SGSS à exploiter pleinement son capital data, gage d’une prise de décision plus agile, plus éclairée et mieux coordonnée.

Ce travail a ainsi été pour moi une occasion précieuse d’acquérir des compétences clés en gouvernance, en gestion des données et en conduite du changement, qui me seront utiles tout au long de ma carrière dans un domaine en constante évolution. Je reste convaincu que la maîtrise des enjeux data est aujourd’hui un facteur différenciant majeur pour les entreprises, et je suis enthousiaste à l’idée de continuer à contribuer à cette dynamique dans mes futures expériences professionnelles.

# **Remerciements**

Je souhaite adresser mes sincères remerciements à l’ensemble des collaborateurs du département de pilotage de la transformation digitale de SGSS, au sein duquel j’ai eu la chance d’effectuer mon alternance. Leur accueil, leur disponibilité et la richesse des échanges ont grandement contribué à mon intégration et à la réalisation de ce mémoire.

Travailler au sein d’une équipe impliquée, curieuse et ouverte m’a permis de mieux comprendre les enjeux concrets de la donnée en entreprise, tout en développant des compétences transverses, tant techniques qu’organisationnelles.

Je remercie également toutes les personnes qui ont pris le temps de partager leur expérience, de m’expliquer leurs processus ou de m’aider à mieux saisir les réalités de terrain. Ces contributions ont nourri ma réflexion et donné une réelle valeur appliquée à ce travail.

Enfin, je remercie les équipes pédagogiques pour leur accompagnement méthodologique, ainsi que toutes les personnes qui m’ont soutenu au cours de cette année exigeante mais extrêmement formatrice.

# **Annexe**

Références :

[1] [Farzana Yaqoob (2022*), Data Governance in the Era of Big Data: Challenges and Solutions*](https://www.researchgate.net/publication/374534991_Data_Governance_in_the_Era_of_Big_Data_Challenges_and_Solutions)

[2] [Soares S. (2012), cité dans Chamberlain et al. (2013), *Using Aspects of Data Governance Frameworks to Manage Big Data as an Asset*](https://scholarsbank.uoregon.edu/server/api/core/bitstreams/afecd01f-444d-4173-99fc-55b1e404e263/content)

[3] [N. Gnanasundari et al. (2023), Cloud Computing: Benefits, Risks and Adoption Strategies](https://www.researchgate.net/publication/311972358_Benefits_and_Challenges_of_the_Adoption_of_Cloud_Computing_in_Business)

[4] [Ali Khajeh‑Hosseini et al. (2010), *The Cloud Adoption Toolkit*](https://arxiv.org/abs/1003.3866)

[5] [Lebaea et al. (2024), *The Role of Data Governance in Ensuring System Success*](https://www.researchgate.net/publication/385168037_The_Role_of_Data_Governance_in_Ensuring_System_Success_and_Long-Term_IT_Performance_A_Systematic_Review)

[6] [Nattakarn Phaphoom et al. (2017), *A survey study on major technical barriers affecting the decision to adopt cloud services*](https://arxiv.org/abs/1712.00606)