POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Data governance for Sustainable Performance in Open Finance

TESI DI LAUREA MAGISTRALE IN
MANAGEMENT ENGINEERING
INGEGNERIA GESTIONALE

Author: **Giorgio Civillini**

Student ID: 10789820
Advisor: Cristina Rossi-Lamastra
Co-advisor: Claudio Bonvino
Academic Year: 2024-2025

Abstract

This study examines how data governance influences resource orchestration in evolving digital ecosystems. As open finance and open banking increasingly rely on distributed data infrastructures and inter-organisational collaboration, firms face growing complexity in balancing innovation, regulatory compliance, and risk control. To address this emerging research area, this study adopts a mixed-method approach combining a systematic literature review of academic contributions published between 2010 and 2025 with an exploratory multiple-case study based on 30 in-depth interviews with senior managers and decision-makers operating within open finance ecosystems. The systematic literature review synthesises existing contributions on process engineering, digital transformation, and data governance, identifying conceptual gaps in the literature. Building on these insights, the multiple-case study investigates how organisations operationalise data governance in practice and how governance mechanisms influence managerial action across different phases of resource orchestration. Our findings develop a data governance orchestration framework in which governance operates as a dynamic control system that defines roles, allocates responsibilities, and structures decision rights across organisational boundaries. This system enables firms to structure, bundle, and leverage data resources in innovative and competitive ways while mitigating uncertainty and enhancing accountability. By conceptualising data governance as an embedded organisational capability rather than merely a compliance function, this study contributes to the strategic management and digital transformation literature. It demonstrates how governance mechanisms support sustainable performance and competitive advantage in highly regulated and data-intensive ecosystems.

Key-words: Data governance; Sustainable performance; Open finance; Resource orchestration; Digital ecosystem; Process engineering.

Abstract in italiano

Questo studio esamina come la data governance influenzi l'orchestrazione delle risorse in ecosistemi digitali in evoluzione. Poiché l'open finance e l'open banking si basano sempre più su infrastrutture di dati distribuite e su forme di collaborazione inter-organizzativa, le imprese affrontano una crescente complessità nel bilanciare innovazione, conformità normativa e controllo del rischio. Per affrontare questo ambito di ricerca emergente, lo studio adotta un approccio metodologico misto che combina una systematic literature review di contributi accademici pubblicati tra il 2010 e il 2025 con un multiple-case study esplorativo basato su 30 interviste approfondite a senior manager e decision-maker operanti in ecosistemi di open finance. La systematic literature review sintetizza i contributi esistenti in tema di ingegnerizzazione dei processi, trasformazione digitale e data governance individuando lacune concettuali nella letteratura. Sulla base di tali evidenze, il multiple-case study analizza come le organizzazioni implementino la governance dei dati nella pratica e come i meccanismi di governance influenzino l'azione manageriale nelle diverse fasi di strutturazione, combinazione e valorizzazione delle risorse. I risultati sviluppano un framework di orchestrazione della data governance in cui la governance opera come un sistema dinamico di controllo che definisce ruoli, assegna responsabilità e struttura i diritti decisionali oltre i confini organizzativi. Tale sistema consente alle imprese di strutturare, combinare e sfruttare le risorse di dati in modo innovativo e competitivo, mitigando l'incertezza e rafforzando i meccanismi di responsabilità e controllo. Concettualizzando la data governance come una capacità organizzativa integrata, e non semplicemente come una funzione di compliance, lo studio contribuisce alla letteratura sul management strategico e sulla trasformazione digitale, dimostrando come i meccanismi di governance supportino una performance sostenibile e un vantaggio competitivo in ecosistemi altamente regolamentati e data-intensive.

Parole chiave: Data governance; Performance sostenibile; Open finance; Orchestrazione delle risorse; Ecosistema digitale; Ingegnerizzazione dei processi.

Contents

Introduction

The growing importance attributed to data has increasingly positioned it as a strategic organisational resource across industries. Managerial and consulting reports emphasise that firms rely on data to support strategic and operational decision-making and to pursue digital transformation initiatives aimed at improving efficiency and competitive positioning. At the same time, empirical evidence shows that many organisations struggle to capture sustained value from their data assets. These difficulties are often not related to the availability of data or digital technologies, but rather to the way data-related activities are organised and governed within firms (McKinsey, 2020).

Stakeholder governance is a central strategic issue, as firms must orchestrate heterogeneous interests, accountability measures, and control rights to manage data access, risk, and value creation across increasingly interdependent ecosystems (Aguilera et al., 2008; Jacobides et al., 2018). Within this setting, data governance can be defined as the system of roles, decision rights, policies and standards that guide how data are managed throughout their lifecycle within organisations (Khatri and Brown, 2010). Prior research shows that data governance supports data-driven decision-making and regulatory compliance. However, existing studies largely treat data governance as an organisational or compliance-oriented function (Dumas et al., 2018; Nkomo and Marnewick, 2021), offering limited insight into how governance mechanisms are enacted within organisational processes and shape the design, coordination and execution of resources in data-intensive contexts.

In this evolving setting, financial companies face increasing pressure to redesign internal processes to improve efficiency, ensure regulatory compliance, and support interoperability across systems and stakeholders. Growing reliance on shared data and cross-organisational coordination exposes the limits of traditional operating models built on siloed databases, fragmented infrastructures, and bespoke decisions. Process engineering has emerged as a central organisational response to these pressures. Broadly understood as the systematic analysis, redesign and optimisation of workflows, process engineering encompasses a variety of approaches, including business process reengineering, process redesign, process transformation and workflow optimisation (Davenport, 1993; Hammer and Champy, 1993). Despite differences in terminology, these approaches share a common objective: redefining how work is performed in order to achieve substantial improvements in performance, quality, speed and customer value. Given the absence of a single, universally accepted definition, this thesis uses these terms interchangeably to refer to the same underlying concept, namely the deliberate redesign and improvement of business workflows.

Digital technologies and data analytics are often portrayed as the primary enablers of process transformation. Technologies such as robotic process automation (RPA), artificial intelligence (AI), and application programming interfaces (APIs) expand firms' ability to integrate resources and support data-driven decision-making in complex and regulated environments (Bonvino and Giorgino, 2024; Gotthardt et al., 2020). However, emerging research shows that technological capabilities alone are insufficient for sustainable transformation. Without clear governing rules, digital initiatives may heighten operational and regulatory risk rather than deliver lasting improvements (Davenport and Redman, 2020).

Many studies draw on established strategic frameworks to explain how data assets and related capabilities contribute to organisational change and competitive advantage. A foundational perspective is the resource-based view (RBV), which emphasises the strategic value of resources that are valuable, rare, difficult to imitate and appropriately organised (Barney, 1991). However, in data-intensive environments, the mere availability of data assets does not automatically translate into sustained value, as data must be supported by appropriate organisational structures, routines and coordination mechanisms to be effectively mobilised (Mikalef et al., 2018; Zeng and Khan, 2019).

Building on these limitations, this study adopts resource orchestration theory (ROT) as its analytical lens, shifting attention from resource ownership to the managerial actions through which resources are structured, bundled and leveraged to support value creation (Sirmon et al., 2011). This perspective is particularly relevant in contexts characterised by extensive data sharing and digital interdependence, such as open banking and the broader transition toward open finance.

In these settings, regulatory initiatives including the Revised Payment Services Directive (PSD2) have redefined how financial data are accessed and exchanged by mandating interoperability through APIs, thereby enabling data flows across organisational boundaries (Arner et al., 2022; Stefanelli and Manta, 2023). As a result, financial institutions increasingly operate within ecosystems involving banks, fintech firms, technology providers and third-party service providers, where coordination and control cannot be achieved through hierarchical mechanisms alone. In such environments, understanding how data-related resources are governed and coordinated across organisational boundaries becomes particularly important, making open banking and open finance a suitable empirical setting in which to examine the role of data governance mechanisms. Against this background, this study is guided by the following research question:

How do data governance mechanisms influence the orchestration of resources in the evolving financial sector?

To address the research question, the study adopts a mixed methodological approach combining a systematic literature review with an exploratory multiple-case study design, taking organisational mechanisms as the unit of analysis. The systematic literature review provides a structured overview of existing academic and practitioner knowledge on process engineering in the financial sector, while the case study enables an in-depth, on-the-ground exploration of how organisational mechanisms related to data governance and digital technologies are enacted in practice. The analysis examines how different firms organise, govern and mobilise data when redesigning workflows and developing data driven use cases in regulated and data intensive environments. By focusing on processes rather than firms as a whole, the study opens the black box of operational change and enables a fine-grained understanding of how data governance mechanisms interact with data, technologies and organisational routines in practice.

Drawing on resource orchestration theory, the study investigates how data governance practices influence the coordination of resources across three interconnected layers: the management of data assets, the redesign of workflows and the development of data driven use cases. In particular, it examines how governance mechanisms shape the coordination and deployment of data-related resources, and how they may facilitate or constrain the emergence of new digital initiatives within firm-level processes.

The contributions of this research are threefold. First, it advances the process engineering literature by demonstrating that governance is not a peripheral condition but a central determinant of process redesign effectiveness, particularly in open finance environments. Second, it extends data governance research by repositioning governance as a system of control and orchestration capability that actively shapes operational processes, rather than merely supporting compliance. Third, it contributes to the application of resource orchestration theory by empirically illustrating how governance mechanisms operate across the structuring, bundling and leveraging phases at the process level. Together, these contributions provide a structured framework for understanding how financial organisations can redesign processes and develop new use cases in open banking and open finance while maintaining control, compliance and strategic coherence.

The remainder of the thesis follows a structured progression. After reviewing the relevant literature, the methodological approach adopted in the study is described. The empirical findings from the case analysis are subsequently presented, before the thesis concludes with a discussion of the main implications and directions for future research.

1 Literature Review

1.1. Literature Review Methodology

Literature reviews are particularly appropriate when a research field is fragmented and spans multiple disciplines, as they enable the systematic structuring of existing knowledge and the identification of conceptual gaps (Webster and Watson, 2002). Building on this approach, this systematic literature review examines prior academic research and practitioner-oriented documentation to map the state of the art in process engineering within the financial sector, with particular attention to the role of digital technologies and data as enabling factors. The review follows the methodological framework proposed by Tranfield et al. (2003) and is structured into four phases: planning, conducting, reporting, and dissemination. In the planning phase, we conducted a narrative analysis that confirmed that, although process engineering is widely discussed in chemical and industrial contexts, it is increasingly gaining importance in the financial sector thanks to technological advancements and the rise of concepts such as open banking and open finance.

For data collection, we began by identifying core synonyms for process engineering, such as business process reengineering, process optimization, process transformation and process redesign. Then, after defining the context and enablers, we excluded “design” and “data” because they returned an unmanageable volume of records, and we added “analytics” to capture data driven studies, identifying the following key search terms: (*process* OR analytics) AND (*engineering OR optimization OR transformation)) AND (financ* OR insur* OR bank* OR "asset management*" OR "wealth management*" OR "investment management*" OR "portfolio management*" OR "venture capital*" OR "private equity*" OR "fintech*" OR "hedge fund*") AND (digital* OR AI OR "Artificial Intelligence" OR "Machine learning" OR "Automat*" OR "platform")).

The search was conducted on 15 April 2025 with our identified keywords within the titles, abstracts, and author keywords in Scopus, selected for its extensive coverage of social science literature and its proven reliability, validity, and timeliness (Sauer and Seuring, 2023). The query was limited to journal articles and conference papers published in English in the subject areas business, management and accounting, and economics, econometrics and finance. Lastly, while process engineering has a well-established theoretical and practical foundation dating back several decades (Hammer and Champy, 1993), this review does not aim to provide a historical reconstruction of the concept. Instead, it focuses on its recent transformation within the financial sector, where digital technologies and data have acted as key enabling factors. For this reason, and in order to capture the most relevant and recent developments, the analysis was limited to publications from 2010 to 2025 (1215 documents). The complete search query is reported in Appendix A.

The review was complemented with grey literature documents, that our Scopus search did not capture. We conducted focused searches for strategic reports published by leading industry actors, including Bain, Boston Consulting Group (BCG) and Goldman Sachs, to ensure that our review incorporated the latest practitioner insights (4 reports). Second, we conducted research on Google Scholar using the same keywords (23 documents). Finally, we searched documents with ChatGPT to support and extend the analysis. We developed prompts that searched academic papers in the web, with concepts related to the identified keywords. The prompts where run until saturation of the findings was reached. This method assisted in confirming some of previous documents and identified new documents. These searches uncovered additional landmark articles and reports that enriched our theoretical framework and ensured a more complete historical and conceptual coverage. The sample counted 1255 documents after removing the duplicated documents.

For data analysis, document screening was supported by an AI tool because of the large initial sample (Collins et al., 2021). We used ChatGPT to assess the relevance of papers obtained from Scopus by instructing it to label relevant studies as "1" and non relevant ones as "0". To increase robustness and reduce potential variability in the classification process, the screening procedure was conducted twice using the same criteria. The two rounds produced largely consistent results. In order to adopt a conservative inclusion strategy and minimise the risk of excluding potentially relevant contributions, all studies labelled as "1" in at least one of the two screening rounds were retained for further analysis, even if classified as "0" in the second round. To ensure the validity and reliability of our analysis, a junior assistant reviewed the work.

All papers labelled "0" were manually reviewed by examining titles first and then abstracts and keywords when doubt arose, to decide on inclusion. This process enhanced the accuracy of our findings by reincluding 81 papers. ChatGPT initially excluded these records because the titles or abstracts did not clearly cover all three required areas: process engineering or related terms, the financial sector, and digital or data. Despite this actual lack, after carefully reading the abstracts, we judged them relevant and reincluded them. The papers rated "1" were 120 out of 1215: we randomly sampled 10 of these studies for manual validation and found them all relevant; therefore, we retained the complete set. This process yielded 201 documents from Scopus, resulting in 241 documents in total for full text assessment.

We opened and read each item in detail and applied the same three inclusion criteria used at screening. On this basis 154 records were excluded. None were pure non finance items since every record contained at least a minimal finance anchor. No process content counted 32 records. These studies discussed technologies or business results in finance but did not explain how processes are designed, governed, or improved. 10 papers lacked any digital or data component. In these studies, process change in financial organisations was attributed to organisational, regulatory, or managerial levers, or to traditional improvement programmes, and when technology appeared, it was only background context and not an active driver of process redesign.

The remaining 112 were not clearly aligned with all three criteria once read in depth. Within this set, 66 are conceptual or methodological overviews, such as reviews, surveys, and framework papers that map technologies or capabilities without making explicit how digital or data capabilities reconfigure financial processes. A further 25 falls under elevated level digital and discuss digital transformation or innovation in finance at a strategic level and do not open the black box of operations and decision rights that would show how technology changes work. Then, we classified 14 as technology without process those papers that describe the use of AI, analytics, or automation in financial settings but provide no clear link to process redesign or governance. The last 7 form another misalignment group. These items mix cross context examples or report outcomes such as profitability or customer experience without evidence of the underlying process mechanisms, or they present residual cases that do not achieve simultaneous alignment. In this way, 87 documents remained. We then added 15 documents through backward citation chasing.

Finally, we employed Research Rabbit to identify similar and related works. This procedure resulted in the selection of 3 relevant documents through the "Earlier Work" function (backward approach), 7 documents through the "Similar Work" function, and 5 documents through the "Later Work" function (forward approach). The final sample comprised 117 studies. The steps are reported as per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting protocol for ensuring the reproducibility of the study (Figure 1).

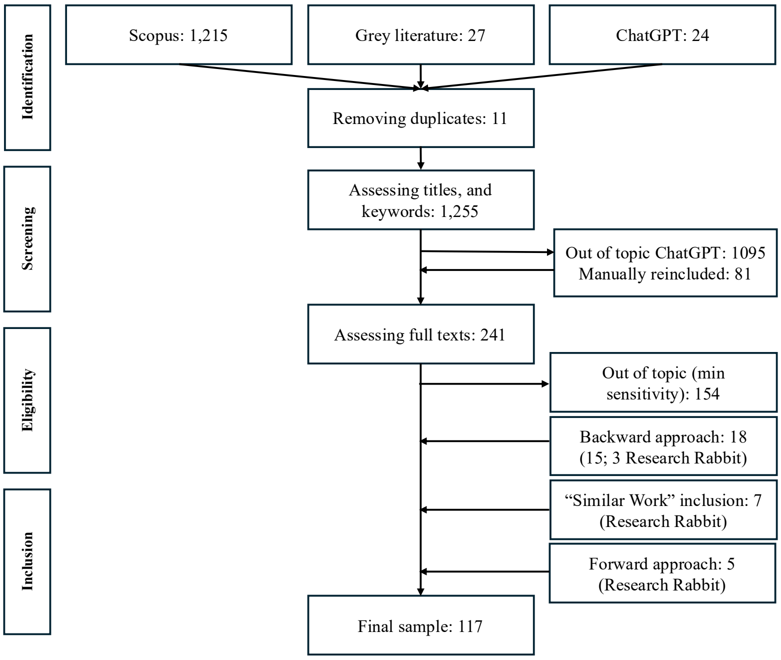


Figure 1. PRISMA framework

For reporting and dissemination, we adopted a systematic, inductive method (Gioia et al., 2021). All key citations were then organised in a concept extraction table with columns for: process engineering and related concepts, finance and related concepts, digital and related concepts, data and related concepts, theories, geographies, sustainability, gender, performance, innovation, and competition (Appendix B). This structure allowed us to identify and cluster significant themes and dimensions across the literature, providing a solid foundation for writing the literature review.

1.2. Process engineering

Process engineering is a systematic, data driven approach aimed at analysing, modelling and continuously improving organisational workflows to enhance efficiency, quality and customer value (Dumas et al., 2018). By contrast, business process reengineering (BPR) is a strategic approach that completely rethinks organisational workflows in order to achieve exceptional improvements in performance. It replaces centralised, hierarchical structures with decentralised models, focused on processes that empower teams, distribute decision authority and ensure that operations support overall strategic objectives (Hammer and Champy, 1993). BPR covers a wide range of areas including customer service activities, operations between organisations such as procurement and logistics, internal routines for planning and control and support functions that, although not directly producing value, are vital to daily operations (Kareem et al., 2024). Despite the difference, we treat them as synonyms because, in practice, both refer to redesigning processes to achieve better performance.

Companies adopt process engineering in order to standardise routines and eliminate waste (Dumas et al., 2018), free specialists from repetitive tasks (Kokina and Blanchette, 2019) and initiate ambitious redesign projects driven by senior leadership (Kaponda et al., 2023). Such projects set explicit goals for cost reduction, quality improvement, service enhancement and faster delivery (Dagher and Fayad, 2024; Onjure et al., 2018) and depend on cross functional teams to dismantle silos while increasing efficiency and adaptability (Kaponda et al., 2023). Beyond the operational benefits, BPR fosters innovation and places the user at the centre by restructuring processes to address actual customer needs (Davenport, 1993).

"Business process Management (BPM) is the art and science of overseeing how work is performed in an organisation to ensure consistent outcomes and to take advantage of improvement opportunities" (Dumas et al., 2018). In BPM lifecycles, process engineering plays a central structural role and is incorporated in two main ways. First, in a continuous and indirect approach, process engineering is embedded directly into the execution cycle. As a process operates, detailed data on execution times, error rates and user feedback are collected. This information is then analysed to diagnose performance gaps and identify opportunities for improvement. Based on these insights, the process is redesigned and redeployed, and the cycle repeats, with execution followed by data collection, analysis and redesign, creating a feedback loop that drives ongoing enhancements. Second, process engineering occurs at set intervals via a dedicated component within the BPM framework. At these predefined moments, a formal reengineering project is launched to examine existing workflows, propose substantive changes and implement a revamped process. Both mechanisms ensure that processes remain aligned with strategic objectives and can evolve in response to changing conditions.

Most BPM frameworks explicitly embed reengineering activities in their lifecycles and architectures, emphasizing the strategic importance of BPR for organisational adaptability and effectiveness (Tsakalidis and Vergidis, 2024). As Muthu et al. (1999) note, organisations undertaking reengineering must first establish a strong rationale for fundamental change and secure stakeholder alignment on the urgency and objectives, which is called the "Discover" phase. During the "Analyse" phase, they perform an As-Is analysis to examine current workflows in depth, identify performance gaps and isolate value adding activities. Using these insights, they develop one or more To-Be models that outline alternative process designs aimed at delivering the desired improvements, known as the "Design" phase. Once the blueprint is in place, implementation begins with pilot tests, targeted training and initiatives to manage resistance and embed the new processes, which corresponds to the "Implement" phase. Finally, continuous monitoring of execution and outcome metrics then supports timely refinements and sustains performance gains, and this is referred to the "Monitor" phase (Andrea and Santoso, 2020; Muthu et al., 1999).

Organisational outcomes from process reconfiguration are influenced by internal factors such as employee education and communication, and by external factors such as regulation and customer expectations (Porfírio et al., 2024). Success also depends on three interdependent dimensions (Marsudi and Pambudi R., 2021): the strategic dimension aligns resources and vision (Mbugua, 2021), the technological dimension exploits modern Information Technology (IT) infrastructures to accelerate data flows and spark innovation (Hadi et al., 2023), and the human dimension empowers employees through continuous learning and skills alignment (Tekka, 2021).

From a market perspective, the global BPR industry was valued at USD 11.78 billion in 2023 and is projected to reach approximately USD 23.72 billion by 2030, growing at a Compound Annual Growth Rate (CAGR) of 10.52%. North America currently holds the largest market share, followed by Europe, while Asia Pacific is expected to experience the fastest growth. Latin America with the Middle East and Africa remain emerging markets with steadily increasing levels of adoption (Maximize Market Research, 2024; Data Insight Market, 2025). Despite its potential, process engineering often stalls in practice due to a range of organisational and executional hurdles. Commonly reported obstacles include resistance to change, lack of change management, unclear definitions of reengineering, absence of a formal framework, insufficient stakeholder appreciation of its benefits, initiatives not linked to strategy, systems not yielding the required results, and incorrect documentation of business processes. Because of these challenges, key stakeholders often fail to provide adequate support, resulting in confusion in BPR initiatives. Workers sometimes embark on projects without proper guidelines, end users do not understand the benefits, and initiatives start without dedicated budgets or resources (Nkomo and Marnewick, 2021).

1.3. Process engineering applications in financial services

Financial institutions seek to redesign their processes to improve customer satisfaction, attain process flexibility, reduce costs (Meadows and Merali, 2003), and accelerate customer onboarding (Bain, 2024). Traditionally, they have relied on structured procedures to ensure compliance, standardisation, and risk mitigation. The move from fragmented systems to integrated infrastructures has positioned data as a decisive element in enabling process transformation in finance (Goldman Sachs 2025; Stefanelli and Manta, 2023)

Open banking and the transition to open finance have emphasised the centrality of data interoperability (Paul and Sadath, 2021; Stefanelli and Manta, 2023). Banks and other financial service providers share customer data with external fintechs and major technology firms via collaborative platforms that use standardized APIs, enabling them to deliver new or enhanced financial services and products that add value for customers (Stefanelli and Manta, 2023). Well designed APIs support reliable real time treasury and banking operations (Goldman Sachs, 2025) and are essential for connecting outdated legacy systems, ensuring uninterrupted data availability (BCG, 2023).

At the same time, the growing complexity of regulatory requirements, intensified customer demands, and mounting competition have highlighted the inadequacy of conventional operational models, pushing firms toward the integration of cohesive, data-driven infrastructures and advanced analytics solutions (Basel Committee, 2003; FDIC, 2006). Market analyses confirm that the Banking, Financial Services and Insurance (BFSI) industry is now one of the largest adopters of BPR solutions worldwide (Maximize Market Research, 2024). In the banking sector, institutions are focusing on reconfiguring both front-office and back-office operations to boost customer experience and operational effectiveness.

Notable interventions include the digitalisation of Know Your Customer (KYC) procedures and reduced loan approval timelines. Moreover, the application of intelligent automation in banking has led to improvements in reporting precision and cost efficiency (Bain, 2024). These improvements contribute directly to financial performance by increasing process productivity, data accuracy and resource optimisation (Gotthardt et al., 2020; McKinsey, 2024).

According to Genty et al. (2025), effective leadership, a strong commitment from top level management and the targeted deployment of digital tools are critical for achieving substantial improvements. As a result, banks are encouraged to reevaluate governance models, invest in technological innovation and nurture a corporate culture oriented toward adaptability and continuous improvement.

In insurance, process engineering has supported the optimisation of core activities such as underwriting, claims management, and policy configuration. These interventions have led to tangible outcomes, including cost reductions and higher fraud detection rates (Agarwal et al., 2022; Eling and Lehmann, 2018). Similarly, in capital markets, big fintech innovations have optimized internal processes, improving investment analysis, forecasting capabilities and operational efficiency (Andronie et al., 2023). In investment management and private equity, processes redefined through operational due diligence, process mapping and workflow optimization (Achleitner, 2010), enhance decision making models, improve portfolio companies' earnings before interest, taxes, depreciation and amortization (EBITDA), and foster more innovative strategies (Liepert, 2024).

Moreover, Big data analytics and the use of AI in the initial stages of the investment process significantly reduce the information asymmetries and even offer more accurate predictions of the probability of success than human analysts. At the same time, emerging technologies help democratize investment decisions (Bajulaiye et al.,2020). FinTech innovations such as green lending and crowdfunding extend sustainability benefits to client services (Andronie et al., 2023). Recent studies in green business process management demonstrate that digital technologies such as advanced analytics, artificial intelligence and Internet of Things (IoT) sensors can embed environmental indicators directly into process dashboards, enabling continuous monitoring, identification of resource waste and implementation of resource efficient process redesign (Couckuyt and Van Looy, 2019; Hoesch-Klohe et al., 2010). Robotic process automation supports sustainability by reducing paper-based processes and enhancing digital workflows.

1.4. Digital technologies as enablers of process engineering

Hammer Champy (1993) and Davenport (1993) are considered foundational contributors to the field of business process reengineering. They highlight the importance of aligning technological capabilities with organisational transformation, a perspective later systematised by Venkatraman and Henderson (1993) through their "strategic alignment model". This model emphasises that organisations must align business strategy, information technology (IT) strategy, organisational and IT infrastructure, and processes to leverage digital technologies as primary enablers of transformation and process optimisation (Venkatraman and Henderson, 1993).

In legacy environments, reconfiguration efforts are often driven by two distinct yet interrelated goals, which are upgrading the underlying technology, and improving the efficiency of embedded business processes (Sneed and Verhoef, 2001). Digital platforms create a unified data backbone that automates and integrates activities, circulates information in real time (Sneed and Verhoef, 2001), enhances control and accuracy (Tinnilä, 1995), and provides the flexibility and scalability needed to reconfigure operations rapidly in response to changing demands (Evans and Fernando, 2025; Volberda et al., 2021). IT solutions also improve communication and coordination and support redesign using heuristics such as task elimination, centralisation and integration. Examples include quality management platforms, adaptive case management systems, web portals and blockchain frameworks, all of which have simplified workflows, removed redundancies and made processes more user centred. In addition, digitalisation removes the physical constraints of paper documents or fixed locations, allowing procedures to be carried out from anywhere (Kubrak et al., 2023).

One of the most effective enablers is robotic process automation (RPA), which uses software bots to automate tasks that are both high in volume and governed by predefined rules. Its adoption has been linked to significant reductions in errors, delays and operational inefficiencies (Pramod et al., 2022; Priyanto et al., 2023). Another information system that plays a strategic role in supporting this evolution is the enterprise resource planning (ERP) system. By centralising and automating a wide range of operational and administrative functions, ERP software platforms improve process transparency, consistency, and organisational integration (Marsudi and Pambudi, 2021). Artificial intelligence (AI) techniques, including machine learning, expert systems and natural language processing, enable the automation of complex tasks (Rahim and Chishti, 2024). These applications improve operational responsiveness by detecting anomalies, interpreting unstructured data and supporting adaptive decision making (Andronie et al., 2024; Gotthardt et al., 2020; Rahim and Chishti, 2024).

Successful AI deployment, however, depends on robust data infrastructure and the organisational capacity to convert analytical insights into tangible process improvements (Mikalef et al., 2018). Finally, other fundamental technologies include application programming interfaces (APIs), defined as "a set of programming tools that enables a program to communicate with another program or an operating system, and that helps software developers create their own applications" (Oxford Learner's Dictionaries, 2025). APIs facilitate system interoperability and modular integration, which are increasingly critical in data-intensive and platform-based environments. Nevertheless, despite their growing relevance, the literature has yet to systematically examine how APIs can be leveraged within structured process engineering and business process reengineering initiatives.

1.5. Data as enablers of process engineering

Big data has revolutionised management by making it possible to measure and therefore manage processes with unprecedented precision, supporting better predictions, smarter decisions and more targeted redesign. By using big data and real time analytics, organisations can switch from fixing problems after they happen to managing processes ahead of time (McAfee and Brynjolfsson, 2012). According to Bonvino and Giorgino (2024), data generates value through four categories of competitive advantages: enhanced risk assessment, optimized operations, customer engagement, and openness. Data are essential for enabling core reengineering methods (Kubrak et al., 2023), and according to Davenport and Redman (2020) digital transformation success depends on bringing together and coordinating technology, data, process, and organisational change capability. The authors liken technology to the engine, data to the fuel, process to the guidance system, and organisational change capability to the landing gear of an airplane. Unless all four components work together, reaching the target destination will not be attainable, and without structured data, even sophisticated technologies cannot deliver improvement. Arsal et al. (2022) argue that making information both relevant and accessible to stakeholders is fundamental to data driven decision making: by transforming raw data into actionable insights, organisations can precisely redesign their processes. Moreover, fostering a culture of information literacy enables them to treat data as a strategic asset for continuous transformation.

1.6. Data governance

Data governance can be defined as the formal system of decision rights, accountability structures and standards that safeguard the quality, security and business value of data throughout its life cycle (Khatri and Brown, 2010). Empirical evidence shows that effective governance configurations vary with contextual contingencies such as regulatory pressure, organisational complexity and strategic focus (Weber et al., 2009). Structurally, firms may adopt centralised, federated or decentralised arrangements, each balancing control and flexibility in different ways. In the centralized model, a single team or department defines and centrally controls all data governance policies and standards, ensuring a high degree of consistency but reducing local flexibility. Regarding the decentralized model, each business unit or geographic area retains full ownership and management of its own data, maximizing local agility and responsiveness but exposing the organization to the risk of inconsistency and duplication. Lastly, the federated or hybrid model strikes a balance by establishing a central governance body that sets overarching policies and quality standards while granting local units' autonomy in implementation, supported by coordination mechanism to maintain alignment across the enterprise (Otto, 2011).

The success of these models depends on social and technical factors such as leadership commitment, data quality processes, a supportive culture and enabling technology, identified as critical success factors for data quality management (Santos and Lucas, 2019). Once digital infrastructures are in place, analytics technologies become central to unlocking the value of data, allowing organisations to automate data collection (van der Aalst et al., 2022), assess the completeness and accuracy of process inputs and outputs (Taymouri et al., 2021) and improve traceability and compliance (Fischer et al., 2024).

Business analytics capabilities translate into superior organisational performance only when they are embedded within and reconfigured through existing routines that fit the firm's context, aligning data quality, analytical insight and managerial actions to prevailing process complexities (Anand et al., 2013). Organisations with mature analytics capabilities are then better able to sense opportunities, reallocate resources and achieve superior performance (Kristoffersen et al., 2021). In open finance ecosystems, PSD2 regulation and API-based business models extend governance beyond firm boundaries, requiring new rules for access, consent and interoperability (Arner et al., 2022). Salerno and Maçada (2025) view data governance as one of the four foundational dimensions of *data orchestration*, which is conceived as a managerial response to the growing complexity of data ecosystems, in which organisations exchange, transform and employ information across organisational and technological boundaries. It is therefore defined as the coordinated management of activities that span those boundaries, moving well beyond the purely technical view of pipelines and workflows. Within this perspective, data governance supplies the formal architecture of decision rights, policies, roles and metrics that coordinate data practices without compromising the autonomy of individual actors. The other dimensions are data strategy, synergy and technological infrastructure. Data strategy works to align every data initiative with the organisation's overarching goals. Synergy encourages partners in the ecosystem to pool their data and expertise, creating value no single actor could achieve alone, while a solid technological infrastructure provides the platforms and networks necessary to move data across boundaries and support analytical workloads.

Everything is grounded in resource orchestration theory (Sirmon et al., 2011), where data orchestration corresponds to the first two managerial phases, structuring and bundling. Within this framework, data strategy and data governance are primarily associated with structuring activities, as they define objectives, decision rights, policies and standards. Synergy and technological infrastructure are instead linked to bundling, as they support the integration of heterogeneous data sources, analytical skills and interoperable platforms. Leveraging is described as a subsequent phase, in which configured resources may be exploited for decision making and innovation (Salerno and Maçada, 2025).

In contrast, recent evidence from manufacturing Small and Medium-sized Enterprises (SMEs) places AI governance within the bundling phase, because it integrates a set of policies and principles that ensure transparent and trustworthy artificial intelligence, mandate reliable data quality assurance, and establish predictive modelling routines that link multiple systems and replace manual planning activities (Peretz-Andersson et al., 2024).

1.7. Theoretical perspectives

Many studies use theoretical frameworks to explain how data assets and related capabilities drive organisational change and competitive edge. A foundational framework is the *resource-based view (RBV)*, which highlights the strategic value of resources that are rare, valuable, hard to imitate and well organised (Barney, 1991). Yet data differ from conventional assets in that they need a supportive organisational environment before they can deliver full value. Research demonstrates that firms, in order to enable meaningful process redesign, must go beyond simply owning data assets and must also build appropriate structures and routines that transform those assets into mechanisms for value creation (Mikalef et al., 2018; Zeng and Khan, 2019)

Building on this foundation, *resource orchestration theory* highlights the importance of managerial actions in organising, combining and deploying resources. It moves the emphasis away from mere ownership toward active coordination: firms achieve success not simply by possessing resources but by orchestrating them into capabilities that support strategic goals (Kristoffersen et al., 2021; Sirmon et al., 2011). This theory views competitive advantage as the outcome of three sequential managerial activities: structuring, bundling and leveraging. Structuring entails configuring the firm's resource portfolio, acquiring what is missing and cutting what no longer fits. Bundling refers to integrating and aligning those resources into coherent capabilities that competitors find hard to replicate. Leveraging involves deploying those capabilities across markets or business models to capture and sustain value (Sirmon et al., 2011). This approach is especially applicable in data intensive settings, where information is dispersed across units, must be integrated, and managed collectively (Mikalef et al., 2018; Saidat, 2023). Bonvino and Giorgino, (2024) demonstrate that orchestrating the six pillars of data valorization, namely sources, types, characteristics, technologies, management practices, and activities, is essential for transforming data assets into competitive advantages and for effectively enabling data driven process improvement.

A complementary perspective comes from the *dynamic capabilities* approach, which highlights the organisation's ability to continuously renew its resource base in order to respond to evolving conditions. Instead of focusing on static advantages, it focuses on sensing opportunities, acting to transform them promptly and reconfiguring processes accordingly. In data contexts this means that analytical and technological capabilities must evolve alongside business goals to support flexible and adaptive process engineering (Mikalef et al., 2018; Zeng and Khan, 2019).

Lastly, *contingency theory* rejects any "one-size-fits-all" approach by asserting that the effectiveness of any intervention hinges on the specific characteristics of the process at hand, such as variability, interdependence, analysability and differentiation, which drive distinct requirements for information processing, uncertainty and equivocality that moderate the impact of digital tools and analytics (Zelt et al., 2019).

While multiple theoretical perspectives offer valuable insights into the role of data and digital capabilities in organisational transformation, this study adopts resource orchestration theory as its main theoretical lens. This choice is motivated by the growing recognition in the literature that the value of data does not derive from ownership alone, but from the organisational and managerial mechanisms through which data-related resources are structured (Salerno and Maçada, 2025). Resource orchestration theory explicitly addresses these mechanisms by focusing on how managerial actions and governance arrangements transform dispersed resources into coordinated capabilities, making it particularly suitable for analysing data governance in complex organisational settings (Sirmon et al., 2011).

Other perspectives, such as the dynamic capabilities approach and contingency theory, provide important complementary insights but are less central for the purposes of this research. Dynamic capabilities emphasise long-term adaptation and renewal, while contingency theory focuses on contextual fit and moderating conditions. However, neither perspective places governance mechanisms at the core of the value-creation process. Accordingly, these frameworks are used to contextualise the analysis, while resource orchestration theory serves as the primary explanatory framework.

1.8. Literature Review framework and gaps identification

Figure 2 provides a synthetic representation of the main themes emerging from the literature review and illustrates how prior studies conceptualise process engineering initiatives through the lens of resource orchestration theory (Sirmon et al., 2011). The framework organises the reviewed contributions according to the three core orchestration activities, structuring, bundling and leveraging, and positions them along a process-oriented progression from initial analysis to implementation and monitoring.

In line with established process engineering life cycles (Andrea and Santoso, 2020; Muthu et al., 1999), the structuring activity aligns with the "Discover" and "Analyse" stages, which focus on defining strategic orientations, establishing organisational foundations and clarifying regulatory constraints. Prior studies typically associate this phase with elements such as strategic alignment, leadership involvement, regulatory framing and the formalisation of governance arrangements that specify roles, responsibilities and decision rights before process redesign takes place.

The subsequent design stage is reflected in the bundling activity, during which organisational components are combined into operational configurations. The literature links this phase to cross-functional collaboration and the integration of technological infrastructures that support coordination across organisational units. Digital platforms, application programming interfaces, automation technologies and advanced analytics are frequently described as enablers that connect previously fragmented activities and facilitate redesigned workflows.

At later stages, leveraging corresponds to the "Implement" and "Monitor" phases, during which developed capabilities are put into use and progressively refined. Existing research associates this activity with pilot deployment, feedback mechanisms and decision-making informed by performance indicators. Leveraging is therefore commonly connected in the literature to outcomes such as operational improvements, product and service innovation, enhanced transparency, customer-related benefits and longer-term competitive positioning.

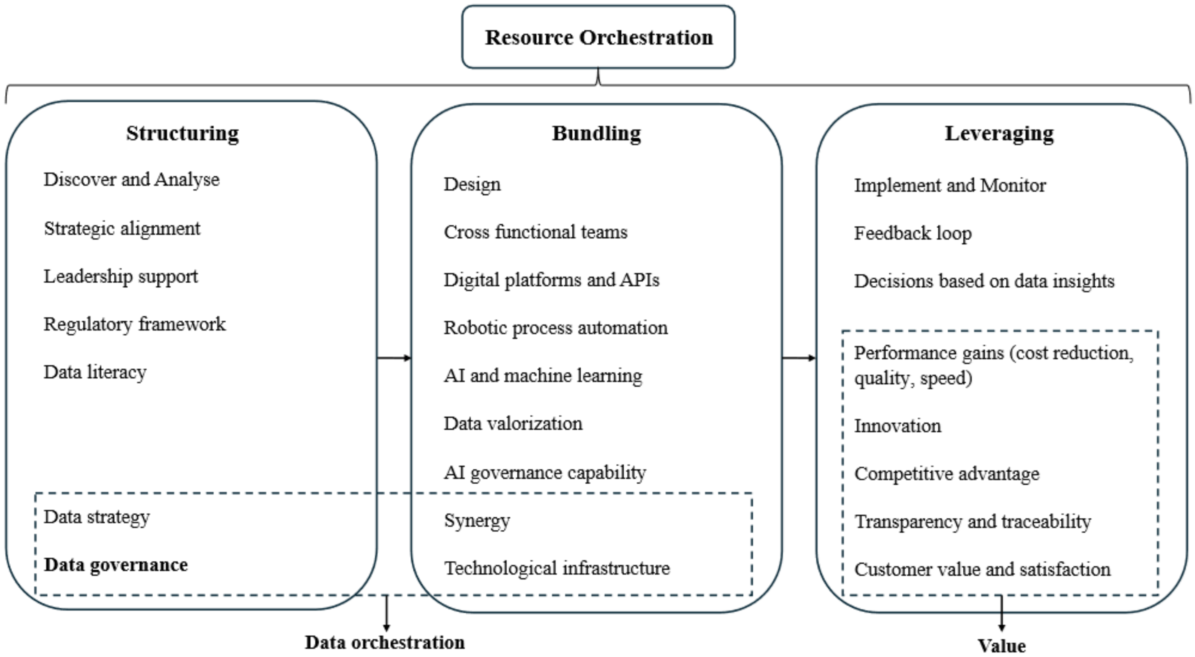


Figure 2. Literature review (LR) framework

The literature still leaves several other critical questions unresolved. Scholars have yet to explain how different data governance arrangements, whether centralised, federated or decentralised, shape the order in which resources are structured, combined and ultimately exploited during process redesign projects. In Salerno and Maçada (2025) data orchestration framework, data governance is firmly rooted in the structuring phase, where it establishes the foundation and configures the data resource portfolio before any integration occurs.

Instead, Peretz-Andersson et al. (2024) show that AI governance in manufacturing SMEs, surfaces in the bundling phase, uniting policies, skills and processes to bind those resources into capabilities. An open issue is whether a similar trajectory, possibly extending into the leveraging phase, appears in an open finance environment, where interoperability is mandatory and data governance occupy a central role in every workflow.

The literature also tends to depict data as the "fuel" for redesign (Davenport and Redman, 2020) without clarifying the concrete governance routines that turn that material into daily execution and measurable performance gains. In addition, empirical studies frequently record obstacles such as resistance to change, misalignment with strategy and poor documentation (Nkomo and Marnewick, 2021), but they rarely examine whether strong governance could ease or intensify these barriers.

A further blind spot concerns the social and environmental dimension, because current research offers little evidence on whether data governance can generate lasting value consistent with environmental, social and governance (ESG) goals or help narrow gender disparities during the implementation of new processes.

Another important gap in the literature lies in the lack of an integrated understanding of how data governance mechanisms influence the coordination and mobilisation of resources across process redesign initiatives. While prior studies recognise the relevance of governance arrangements, digital technologies and organisational capabilities, they tend to examine these elements in isolation or within specific phases, offering limited insight into how governance mechanisms shape resource orchestration as processes unfold over time.

Taken together, these gaps are particularly urgent in data intensive environments such as open banking and open finance, where regulatory pressure and the interoperability delivered by APIs amplify both the potential for process redesign and the complexity of executing it.

1.9. Literature review conclusion and next steps

Overall, the literature highlights the growing relevance of data governance mechanisms, digital technologies and organisational capabilities in shaping how firms operate. Prior studies provide valuable insights into governance arrangements, process redesign practices and resource-based perspectives. However, these streams of research largely remain disconnected, offering limited integration between data governance mechanisms and theories explaining how resources are coordinated and mobilised to create value.

In particular, while resource orchestration theory provides a powerful lens to understand how firms structure, combine and deploy resources, existing research has only marginally explored how data governance mechanisms influence these orchestration dynamics. Governance is often treated as a contextual condition, a compliance requirement or a supporting infrastructure, rather than as an active mechanism shaping how resources are coordinated across organisational and inter-organisational settings.

This limitation becomes especially salient in the evolving financial sector, where increasing data sharing, regulatory intensity and ecosystem-based interactions challenge traditional assumptions about control, coordination and value creation. As financial services move toward more open and interoperable configurations, understanding how data governance mechanisms influence the orchestration of resources emerges as a critical yet underexplored issue. The expected outcome is a framework that answers the central question:

How do data governance mechanisms influence the orchestration of resources in the evolving financial sector?

2 Methodology

2.1. Research approach and case selection

This study adopts an exploratory multiple-case research design, which is well suited to investigating complex organizational phenomena within their real-life contexts and to supporting theory building from empirical evidence (Eisenhardt, 1989; Feagin et al., 1991; Yin, 2003). Accordingly, a qualitative approach is employed to capture contextual dynamics, actors' interpretations, and processual mechanisms that are difficult to fully grasp through quantitative methods alone (Morgan and Smircich, 1980; Neuman, 2006).

The unit of analysis is the data governance process, which is examined through an exploratory approach to understand how data governance practices interact across three interconnected layers: the management of data assets, the redesign of workflows and the development of use cases. Studying governance at each layer allows us to identify its contribution to the structuring, bundling and leveraging phases described by resource orchestration theory and to assess whether data governance facilitates or constrains the emergence of new digital initiatives in open banking and open finance contexts.

A multiple-case research design is particularly appropriate for studying emergent and complex phenomena in which theoretical understanding is still developing (Edmondson and McManus, 2007). This approach enables systematic comparison across cases and supports theory building by revealing patterns and contrasts that would be difficult to observe in single-case designs. It is therefore well suited to examining how organisational arrangements operate across different empirical settings.

The preliminary sample comprises 145 organisations worldwide, identified through the OpenBankingUseCases.com directory, an online repository that catalogs live open banking and open finance use cases across banks, fintech companies and digital platforms on a global scale.

Rather than applying restrictive sampling criteria, the study adopts a census-based approach by contacting the full population of organisations included in the directory. This choice is justified by both the manageable size of the population and its strong relevance for the research objective. In fact, the directory provides broad geographical coverage and spans multiple organisational types and market contexts, making it particularly suitable for investigating data governance practices in heterogeneous financial ecosystems. Importantly, it includes both early-stage fintech and startup initiatives as well as more established and mature financial institutions, allowing the analysis to capture data governance arrangements across different stages of organisational development Moreover, all organisations listed in the directory are directly involved in data-intensive initiatives related to open banking or open finance, ensuring a high level of conceptual alignment with the focus of the study.

Cases were selected based on a set of criteria designed to ensure both relevance and diversity. First, we included only organisations that demonstrated mature data governance practices, reflected in the presence of clearly defined roles such as a Chief Data Officer, formal responsibilities for data ownership and stewardship, and documented data policies. Second, we prioritised cases for which key informants were accessible and willing to provide detailed explanations of their routines and practices, which allowed the systematic collection of rich empirical material. Third, we ensured variation in organisational characteristics, including size, geographical location and sector of activity, in order to capture a broad and representative view of how data governance works across different contexts. This approach enables the study to leverage a globally diverse and theoretically relevant pool of cases while avoiding premature exclusion and preserving the exploratory nature of the multiple-case research design.

2.2. Data collection

Data were collected through semi-structured, in-depth interviews with knowledgeable informants from the selected organizations, following established qualitative case study practices aimed at capturing processual dynamics and organizational complexity (Sjödin et al., 2023). Prior to each interview, participants were informed about the research objectives and explicitly asked for consent to record the interview for research purposes only. All interviews were video recorded with the participants’ consent. To ensure confidentiality, no interviewee consented to being identified by name, nor did any organization authorized the disclosure of its identity. Consequently, all cases are anonymized in the analysis, while the interview recordings are securely stored and documented to ensure accuracy, transparency, and research integrity.

In total, we contacted 280 individuals and, while many did not respond, the data collection process resulted in 30 interviews between August and November 2025. For each organisation, interviewees were selected among individuals with direct involvement in data governance initiatives, such as roles related to data management, analytics, or digital transformation, ensuring informed and experience-based responses. To ensure a comprehensive and balanced perspective, we interviewed individuals with diverse roles and levels of seniority. The sample spans a wide spectrum of organisational positions, ranging from senior executives such as Chief Compliance Officers, Chief Data Officers and Chief Operations Officers to more junior professionals involved in day to day operational and analytical activities. This diversity was intentionally pursued to capture process dynamics across different levels of decision making and execution, rather than relying exclusively on the perspectives of the most senior actors. While senior executives provided strategic and governance level insights, more junior and mid level professionals offered valuable evidence on how data related routines, constraints and controls are implemented in practice.

During the interviews, informants were asked open ended questions guided by a semi-structured interview guide. The questions were developed with reference to the resource orchestration theory and explored how data governance influences the structuring, bundling and leveraging of data. Building on this theoretical foundation, the guide also covered themes such as detailed descriptions of relevant use cases, the role of data governance as an enabler or a constraint in designing, deploying and scaling these use cases, and the capabilities and routines needed for their successful implementation. We also asked participants to reflect on the relationship between data governance and ESG, in order to understand how data related practices contribute to broader environmental, social and governance objectives.

More specifically, respondents were invited to consider a set of open questions that formed the core of the interview guide and were designed to examine whether data governance facilitates or hinders the creation of use cases in open banking and open finance. Appendix C presents the semi-structured interview questions, accompanied by concrete examples illustrating how the questions were phrased and posed during the interviews.

We used extensive follow up questions to clarify points and encourage respondents to provide further detail, which enabled a deep and contextualised understanding of these themes. Our interviews lasted about thirty minutes each and took place via online conference calls. However, several respondents could only offer a very short time slot, and these interviews lasted about ten minutes. In these cases, we focused exclusively on the main questions and did not ask any follow up prompts. All sessions were recorded and transcribed, and the verbatim transcripts served as the primary input for our analysis.

To strengthen the credibility of our findings, we conducted an extensive triangulation process. In addition to the interview material, we reviewed a wide range of documents including solution descriptions, presentation slides, company reports, contractual agreements, published materials and project files.

We also examined content from sector conferences that featured the organisations in our study, including *Il Salone dei Pagamenti (Milan, October 2025), Open Banking Expo UK and Europe (London, October 2025), Open Banking Expo Canada (Toronto, June 2025) and Open Banking Expo USA (New York, June 2025)*. These sources provided further contextual information on the development and operation of open finance initiatives.

Finally, some respondents who were not available for a full interview provided short answers to the main questions through email or LinkedIn messages. These written exchanges offered valuable clarifications and contributed to our understanding of specific data governance practices. Table 1**Table 1.** Summary of interviews summarises the interview sample by reporting interviewees' roles, organisational context, geographic distribution, firm size. Overall, the empirical material consists of 10 hours and 14 minutes of verbatim-transcribed interviews, corresponding to approximately 92,000 words of transcript.

Table 1. Summary of interviews

ID	Role	Category	Geography	Size
A	Software Engineer	Accounting	Global	10.000+

B	Product Data Analyst II	Payments	United Kingdom	10.000+
C	Data Analyst	Payments	Global	10.000+
D	Senior Legal Analyst	API Aggregators, Open Banking Aggregators	Brazil	51–200
E	Chief Product Officer	Account Aggregation	Netherlands	11–50
F	Senior Relationship Manager	KYC, Payments	Sweden	501–1.000
G	Chief Product Officer	API Aggregators	India	201–500
H	Chief Operating Officer	API Aggregators	France	51–200
I	Senior Data Science Manager	Consumer Credit	United Kingdom	501–1.000
J	Engineering Director	API Aggregators	Sweden	501–1.000
K	Chief Commercial Officer	Core Banking	United Kingdom	5001–10.000
L	Chief Operating Officer	Accounting	Canada	201–500
M	Lead Data Science	Payments	United Kingdom	5.001–10.000
N	Product Development Intern	Home Buying	United Kingdom	11–50
O	Senior Product Operations Manager	Business Management	France	1.001–5.000
P	Lead Data Scientist	Card Linked Loyalty	France	51–200
Q	Head of Data Science	Buy Now, Pay Later	Sweden	5.001–10.000
R	Chief Compliance Officer	API Aggregators	Finland	11–50
S	Lead Data Engineer	Wealth Management	United Kingdom	51–200
T	Data Scientist	Wealth Management	Netherlands	51–200
U	Senior Service Delivery Manager	Account Aggregation	Middle East	201–500

V	Internal Auditor	Consumer Credit, Payments, Personal Finance Management	France	201–500
W	Compliance Officer	Accounting, Business Management, Taxes	United Kingdom	51–200
Y	Senior Data Analyst	Consumer Credit, Payments, Personal Finance Management	France	201–500
X	Senior Engineering Manager	Buy Now, Pay Later	Sweden	5.001–10.000
Z	Data Scientist	Payments	Europe	10.000+
AA	Business Developer	Payments	Italy	501–1.000
BB	Lead Data Engineer	Accounting	Global	10.000+
CC	Data Analyst	Buy Now, Pay Later	Ireland	201–500
DD	Data Engineer	Payments	Italy	501–1.000

2.3. Qualitative data analysis

Data analysis followed an iterative and systematic process involving in line with established qualitative analysis practices (Gioia et al., 2010; Miles and Huberman, 1984). In our analysis, we followed Braun and Clarke, 2006 approach to thematic analysis and coded the data using four steps. First, we immersed ourselves in the raw material by reading every interview transcript and all supporting documents multiple times. During this initial phase we highlighted passages that illustrated how data governance is enacted in practice and how it influences the creation, development and scaling of use cases in open banking and open finance.

Second, we generated a set of first level codes by assigning descriptive labels that reflected the participants' own words. These codes captured concrete data governance practices, such as data ownership structures, data access routines, data quality controls, validation procedures and consent management mechanisms. They also captured the constraints and frictions that respondents encountered when mobilising data for new initiatives. In the third phase we examined these initial codes for connections and recurring patterns. We grouped similar items and compared the codes across cases, while consulting the literature on data governance and the resource orchestration theory. This allowed us to refine our emerging categories in ways that linked empirical observations to the structuring, bundling and leveraging of data resources. Throughout this phase we engaged in memo writing to document our interpretations and to trace how each insight related to our central research question on whether data governance facilitates or hinders the development of use cases.

Next, we elevated these categories into a set of core themes. These themes were grounded in theoretical concepts related to roles and responsibilities, policies and standards, data access and control, and the routines through which data are organised, combined and mobilised inside each organisation. By structuring the analysis in this way, we translated individual codes into broader conceptual dimensions that reflect different facets of data governance and their influence on use case development in open banking and open finance.

Finally, we validated our interpretations by presenting a draft version of the data structure to three interviewees. Their feedback helped us consolidate overlapping themes, refine definitions and confirm the relevance of the dimensions we had identified. The outcome of this process is a layered structure that moves from raw excerpts to first level codes, then to second level themes and conceptual dimensions. This structure captures the mechanisms through which data governance enables or constrains the creation of open finance use cases and forms the analytical foundation of our findings.

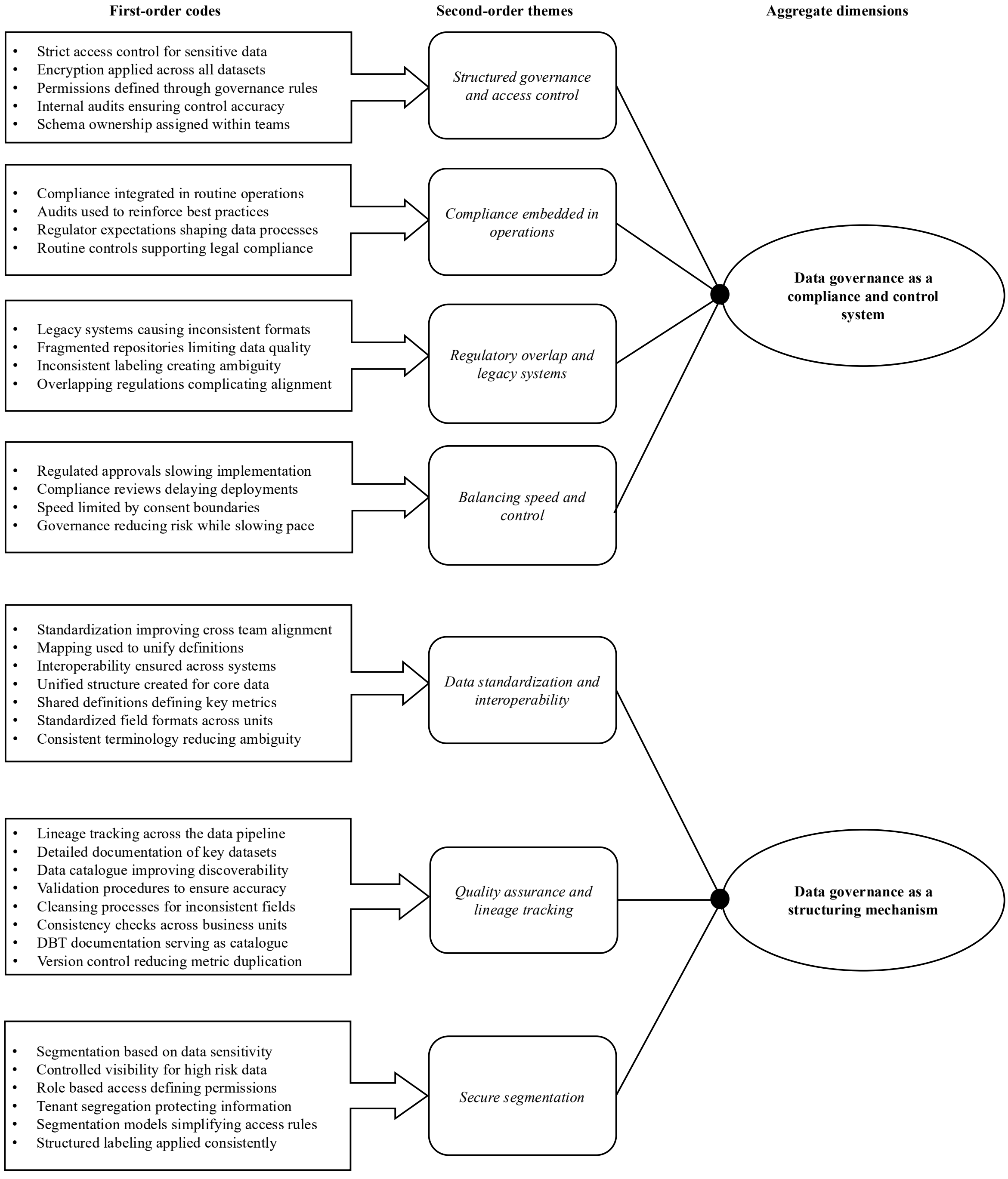
Table 2**Table 2.** Example of the coding structure presents illustrative examples of interview excerpts and demonstrates how raw data were systematically analysed and organised into first-order codes, second-order themes and aggregate dimensions.

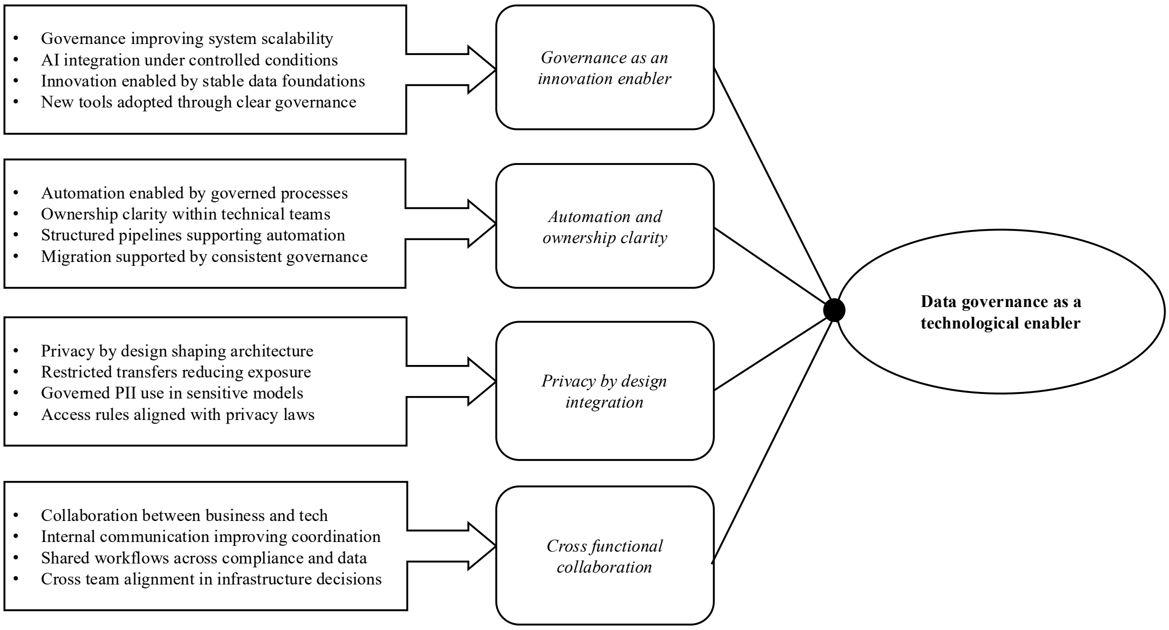
Table 2. Example of the coding structure

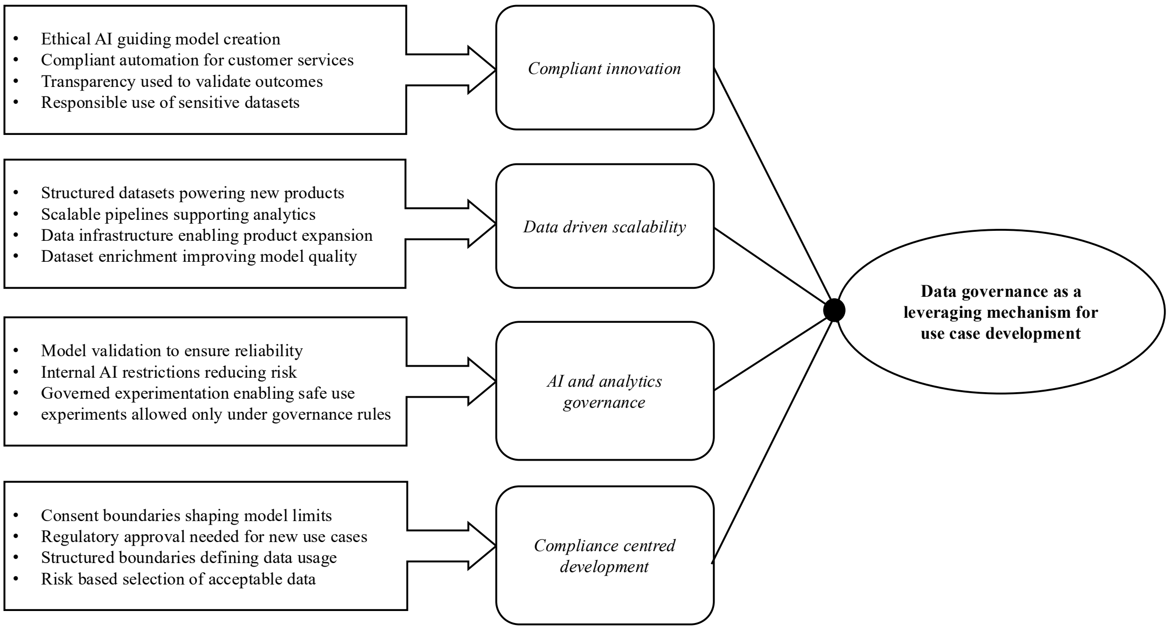
ID	Quote	First-order code	Second-order theme	Aggregate dimension
M	*"Not everyone should have access to the same level of data. Governance comes with a trade-off: as soon as you apply access rules, you inevitably fragment data, but the benefits are much bigger. It protects security and ensures analysts only work with the relevant part of the dataset instead of seeing personal details they do not need. If we allowed everyone to see everything, even basic queries would become messy*	Strict access control for sensitive data	Structured governance and access control	Data governance as a compliance and control system of risk

	and risky. Governance ensures that data is accessed only by the people who truly need it".			
H	*"We fully comply with General Data Protection Regulation (GDPR), Revised Payment Services Directive (PSD2), Commission Nationale de l'Informatique et des Libertés (CNIL) guidelines and all related standards. We conduct a yearly review of 100% of our processes, overseen by our risk and compliance team and formally validated by either myself or our president. Whenever new regulatory recommendations appear, we update our internal procedures. Compliance is not something we do at the end. It shapes how we design every workflow and every feature. It determines how we document processes, how we structure access, and how we monitor activity. Everything starts from compliance, not the other way around".*	Compliance integrated in routine operations	Compliance embedded in operations	Data governance as a compliance and control system of risk
C	*"Legacy datasets often require manual cleaning because basic attributes like dates are stored in incorrect formats. Without fixing those issues, nothing new can be built on top of them. During our migration we saw that even simple errors, like dates saved as text, could break entire workflows. These inconsistencies must be corrected before any new analytical or technological initiative can be introduced".*	Legacy systems causing inconsistent formats	Regulatory overlap and legacy systems	Data governance as a compliance and control system of risk
T	*"Governance often blocks external tools because they conflict with rules around client data. This slows things down, but it also avoids mistakes and misuse, especially on old systems where quality or masking is not guaranteed. If governance forbids something, we stop immediately. It keeps us safe".*	Governance reducing risk while slowing pace	Balancing speed and control	Data governance as a compliance and control system of risk

Figure 3 shows the entire data structure that resulted from the data analysis. Particular attention will be devoted to ensuring transparency, coherence and credibility in the presentation and interpretation of qualitative findings (Golden-Biddle and Locke, 1993; Maxwell, 1997).


First-order codes
Second-order themes
Aggregate dimensions
• Strict access control for sensitive data
• Encryption applied across all datasets
• Permissions defined through governance rules
• Internal audits ensuring control accuracy
• Schema ownership assigned within teams
Structured governance and access control
• Compliance integrated in routine operations
• Audits used to reinforce best practices
• Regulator expectations shaping data processes
• Routine controls supporting legal compliance
Compliance embedded in operations
• Legacy systems causing inconsistent formats
• Fragmented repositories limiting data quality
• Inconsistent labeling creating ambiguity
• Overlapping regulations complicating alignment
Regulatory overlap and legacy systems
• Regulated approvals slowing implementation
• Compliance reviews delaying deployments
• Speed limited by consent boundaries
• Governance reducing risk while slowing pace
Balancing speed and control
Data governance as a compliance and control system
• Standardization improving cross team alignment
• Mapping used to unify definitions
• Interoperability ensured across systems
• Unified structure created for core data
• Shared definitions defining key metrics
• Standardized field formats across units
• Consistent terminology reducing ambiguity
Data standardization and interoperability
• Lineage tracking across the data pipeline
• Detailed documentation of key datasets
• Data catalogue improving discoverability
• Validation procedures to ensure accuracy
• Cleansing processes for inconsistent fields
• Consistency checks across business units
• DBT documentation serving as catalogue
• Version control reducing metric duplication
Quality assurance and lineage tracking
• Segmentation based on data sensitivity
• Controlled visibility for high risk data
• Role based access defining permissions
• Tenant segregation protecting information
• Segmentation models simplifying access rules
• Structured labeling applied consistently
Secure segmentation
Data governance as a structuring mechanism


• Governance improving system scalability
• AI integration under controlled conditions
• Innovation enabled by stable data foundations
• New tools adopted through clear governance
Governance as an innovation enabler
• Automation enabled by governed processes
• Ownership clarity within technical teams
• Structured pipelines supporting automation
• Migration supported by consistent governance
Automation and ownership clarity
• Privacy by design shaping architecture
• Restricted transfers reducing exposure
• Governed PII use in sensitive models
• Access rules aligned with privacy laws
Privacy by design integration
• Collaboration between business and tech
• Internal communication improving coordination
• Shared workflows across compliance and data
• Cross team alignment in infrastructure decisions
Cross functional collaboration
Data governance as a technological enabler


• Ethical AI guiding model creation
• Compliant automation for customer services
• Transparency used to validate outcomes
• Responsible use of sensitive datasets
Compliant innovation
• Structured datasets powering new products
• Scalable pipelines supporting analytics
• Data infrastructure enabling product expansion
• Dataset enrichment improving model quality
Data driven scalability
• Model validation to ensure reliability
• Internal AI restrictions reducing risk
• Governed experimentation enabling safe use
• experiments allowed only under governance rules
AI and analytics governance
• Consent boundaries shaping model limits
• Regulatory approval needed for new use cases
• Structured boundaries defining data usage
• Risk based selection of acceptable data
Compliance centred development
Data governance as a leveraging mechanism for use case development

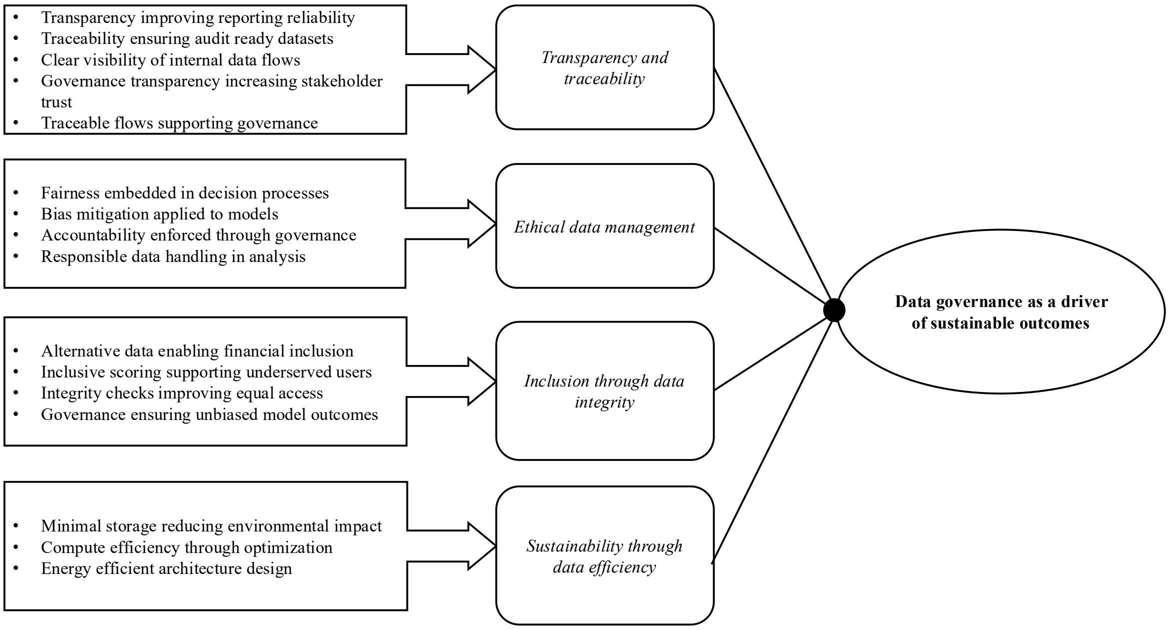


Figure 3. Data structure

3 Findings

Before discussing the substantive findings, it is useful to outline the descriptive characteristics of the interview sample in order to contextualise the empirical evidence. The interviewees include both male and female respondents, specifically 21 men and 9 women, allowing for a diverse representation of perspectives across gender. While the study does not aim to assess gender-based differences, some general tendencies were observed in the way data governance issues were articulated. Female respondents more frequently focused on technical aspects and concrete use situations, grounding their answers in specific examples related to data usage, implementation challenges, and operational trade-offs. Male respondents, by contrast, tended to frame their responses at a more conceptual and systemic level, often structuring their answers around overarching governance principles, organisational rationales, and strategic considerations, and providing a broader interpretative perspective on the questions posed. These observations are not interpreted as causal relationships, but as indicative patterns that help contextualise and enrich the interpretation of the findings.

From a geographical perspective, the sample spans multiple countries and regions, reflecting the global nature of open banking and open finance initiatives. Despite variations in regulatory environments and market maturity, several governance challenges such as data ownership, accountability, and cross unit coordination emerged consistently across countries. At the same time, differences in regulatory intensity and market structure appeared to influence the degree of formalisation of data governance mechanisms.

The sample also includes organisations at different stages of development, ranging from early stage fintech startups to more mature financial institutions with larger employee bases. Respondents from startup environments frequently highlighted challenges associated with rapid organisational growth, including difficulties in keeping data governance frameworks up to date and misalignments between evolving business needs and existing governance rules. In contrast, respondents from more established organisations emphasised challenges related to managing organisational complexity, coordinating governance practices across units, and ensuring consistency and standardisation at scale. Together, these descriptive characteristics provide important contextual lenses for interpreting how data governance supports process engineering across different organisational settings.

We present our findings on data governance in the open banking and open finance context in five parts. First, we examine how data governance functions as a compliance and risk control system, shaping how organizations interpret and operationalize regulatory obligations such as PSD2, GDPR, RBI guidelines, and national privacy laws. In this part, we outline how governance frameworks translate external regulation into internal routines, controls, and access mechanisms. Second, we conceptualize the role of data governance as a structuring mechanism. Here we explore how organizations establish the foundations for data quality, consistency, and interoperability through practices related to standardization, lineage, documentation, and segmentation. This dimension helps explain how firms organize and stabilize their data assets before they can be used effectively. Third, we analyse how data governance acts as a technical enabler, supporting the introduction and operation of new technologies and infrastructures. In this section, we focus on how governance provides the conditions for secure integration of tools such as AI models, cloud platforms, data warehouses, automation pipelines, and API based architectures. Fourth, we investigate how data governance operates as a leveraging mechanism for use case development. This part focuses on how governance shapes the feasibility, scope, and speed of data driven innovation, influencing how organizations design, validate, and scale new products and analytical applications. Finally, we examine the role of data governance as a driver of responsible and sustainable outcomes. This dimension connects governance practices to broader ESG considerations, including transparency, traceability, fairness, inclusion, and environmental efficiency, highlighting how data governance contributes to responsible and accountable value creation within the financial ecosystem.

Following the presentation of the findings, we discuss the resulting framework and elaborate on how data governance, when properly configured, can support the emergence of new open finance use cases and strengthen an organisation's ability to innovate within data driven ecosystems. In doing so, we highlight how governance not only enables the secure mobilisation of data resources but also orchestrates the interaction between regulation, technology, and organisational capabilities. This integrated perspective allows us to explain how firms can move beyond compliance-oriented approaches and leverage data governance as a strategic capability that enhances learning, experimentation, and responsible value creation across the open finance landscape.

3.1. Data governance as a compliance and control system of risk

Our findings indicate that data governance represents the foundational layer upon which organisations operating in open banking and open finance build all data related activities. Across interviews, informants consistently described regulatory obligations such as PSD2, GDPR, national privacy laws or RBI guidelines as forces that do not simply impose external constraints but shape internal routines, access models, approval structures and interpretations of what responsible data use means. In this environment, governance becomes a mechanism through which companies protect sensitive information, reduce operational risk and maintain trust while still enabling innovation within clearly defined boundaries. Respondents repeatedly emphasised that the organisational default is restriction rather than permissiveness. Data is treated as sensitive by design and access is considered an exception that must be justified and documented. Governance is not something that occurs after data has been processed. It is embedded in every step of the data lifecycle and determines what can be done, by whom, and under which controls. To unpack this dimension, we examine four themes that reveal how compliance and risk control shape organisational behaviour.

3.1.1. Structured governance and access control

Interviewees consistently described access control as the clearest and most immediate expression of data governance. Organisations rely on role-based permissions, classification systems, encryption and formal workflows to minimise unnecessary exposure to sensitive information. Access is not granted by default but is treated as an exception that requires justification, approval and clear documentation. This reflects a principle of “least privilege” that determines how employees interact with data in their daily work and ensures that every access request can be traced back to an explicit purpose. Interviewees reinforced the centrality of access control as the most visible expression of data governance.

Participant M highlighted the importance of strict access controls in data governance, emphasising that unrestricted data access may generate significant risks if not properly managed. As the participant stated:

"Not everyone should have access to the same level of data. Governance comes with a trade-off: as soon as you apply access rules, you inevitably fragment data, but the benefits are much bigger. It protects security and ensures analysts only work with the relevant part of the dataset instead of seeing personal details they do not need. If we allowed everyone to see everything, even basic queries would become messy and risky. Governance ensures that data is accessed only by the people who truly need it".

Interviewee B described how access to data is structured along clear organisational boundaries and governed through formalised approval mechanisms. As the interviewee noted:

"Every department has its own schema, and you can only access the tables belonging to your domain. If you need something outside, you must open a ticket, justify the request, and wait for the owner to review it. Sensitive domains often deny access even after several attempts. The default is that access is never granted unless there is a clearly documented purpose, and only then is the request evaluated. This process maintains control and avoids unnecessary exposure of sensitive data".

Respondent H emphasised the strict application of the principle of least privilege, particularly in the context of regulated financial institutions subject to supervisory oversight. As the respondent indicated:

"As a licensed company supervised by the Banque de France, one of our foundational rules is that no one inside the company has access to raw banking data. The data is fully encrypted in transit and at rest, and we strictly enforce a 'least access' principle. Everyone only gets the minimal level of access required to perform their job, and personal data is encrypted multiple times and stored securely. This ensures that every access is controlled and that all interactions with sensitive information are fully traceable".

This mindset influences the entire data lifecycle. Classification frameworks define different levels of sensitivity, encryption protects data in transit and at rest and permissioning structures are periodically reviewed. As a result, access governance is not only a security mechanism but a cultural expectation. Teams work under the assumption that data is always controlled and that compliance must be demonstrated continuously. This strict approach reduces risk, limits the consequences of human error and creates an auditable environment where every interaction with sensitive information is monitored and explainable.

3.1.2. Compliance embedded in operations

A second theme concerns the deep operational integration of compliance in daily practices. Informants explained that regulatory obligations are not handled at the end of a process but shape the process itself. Governance becomes an organising principle that informs the design of workflows, validation procedures, documentation standards, incident response plans and retention rules. Rather than being perceived as a bureaucratic constraint, compliance is internalised as routine organisational behaviour.

Interviewees repeatedly emphasised that compliance plays a foundational role in shaping how operational work is structured within the organisation. In particular, Participant H highlighted how regulatory requirements are embedded directly into process design rather than treated as an ex post constraint. As the participant explained:

"We fully comply with General Data Protection Regulation (GDPR), PSD2, Commission Nationale de l'Informatique et des Libertés (CNIL) guidelines and all related standards. We conduct a yearly review of 100% of our processes, overseen by our risk and compliance team and formally validated by either myself or our president. Whenever new regulatory recommendations appear, we update our internal procedures. Compliance is not something we do at the end. It shapes how we design every workflow and every feature. It determines how we document processes, how we structure access, and how we monitor activity. Everything starts from compliance, not the other way around".

Interviewee G described a similar dynamic, highlighting how regulatory requirements directly influence both day to day operations and longer term product planning. In this context, the interviewee emphasised the central role of financial regulation in shaping organisational processes. As pointed out:

"Reserve Bank of India (RBI) guidelines heavily shape our daily work. Only lenders approved by RBI can legally lend, and every product must comply with guidelines like the Digital

Lending Guidelines. Our roadmap is influenced by upcoming regulation. We have recurring meetings with clients to align on new guidelines for the next quarter and prepare our capabilities. Compliance shapes the entire process, not just the end. It is not a bureaucratic step; it determines how systems are built, how data is stored, and how user flows are designed"

Respondent A highlighted how regulatory requirements are operationalised through concrete technical decisions, underscoring the mediating role played by Governance, Risk and Compliance (GRC) functions in translating legal and policy frameworks into engineering practices. As the respondent remarked:

"The Governance, Risk and Compliance (GRC) team reads 200 - 300 page policy documents and extracts what engineers must do. They translate regulatory text into technical requirements, like how passwords must be hashed or which datasets must be encrypted. Almost every step of the engineering workflow is shaped by this interpretation of compliance. Compliance therefore drives the structure of our internal processes from the very beginning".

This operational embedding of compliance creates a stable environment in which employees know exactly which rules apply, how to interpret them and how to implement them in their daily tasks. It ensures consistency across divisions, reduces ambiguity and allows organisations to demonstrate accountability to regulators and external stakeholders. It also reinforces a shared organisational understanding of data as a regulated asset that requires continuous care.

3.1.3. Regulatory overlap and legacy constraints

Organisations also face the combined challenge of multiple regulatory frameworks and inherited legacy systems that are not always aligned with modern data governance standards. Interviewees described situations in which different regulations apply simultaneously, each with its own definitions, expectations and interpretations. This overlap requires a continuous interpretive effort from teams who must ensure that every procedure respects the strictest applicable rule. Legacy infrastructures make this task even more complex. Datasets created under previous regulatory regimes often lack consistent labels, up to date documentation or adequate masking of sensitive information. This creates structural risks that must be managed before new analytical or technological initiatives can be introduced.

Interviewees provided several examples illustrating how regulatory overlap and legacy inconsistencies shape everyday operational practices. In particular, Participant E highlighted the tensions that arise when multiple regulatory frameworks impose partially conflicting requirements, especially in relation to data retention and deletion. As the participant observed:

"GDPR and PSD2 overlap and sometimes contradict one another, especially regarding data retention. PSD2 obliges us to store data for seven years while GDPR obliges us to delete data when a user offboards. When masking is not applied correctly, we are required to purge the data entirely. This shows how critical proper structuring and documentation are when dealing with legacy systems and multiple regulatory frameworks".

Others highlighted the ongoing need to reconcile legacy infrastructures with contemporary governance and compliance expectations. In particular, Respondent A emphasised how historical design choices and incomplete documentation complicate efforts to align older systems with current regulatory and organisational standards. As the respondent noted:

"Legacy systems were never built with modern governance practices in mind. Documentation may be incomplete or outdated, and sometimes we need to reconstruct the entire context before aligning it with current compliance requirements. It is a constant process of reconciliation because older infrastructures contain fields that have been used differently across time, and mapping or documentation is sometimes missing. This makes harmonising old and new environments one of the most difficult tasks".

These challenges were not limited to internal systems but extended to the broader ecosystem in which organisations operate. Interviewee U described how fragmented practices across banks and fintechs, combined with regulatory overlap and legacy infrastructures, generate additional complexity in process alignment and governance. As the interviewee explained:

"We deal with fragmentation because different banks and fintechs do not always follow the same data standards. Regulatory overlap is another issue: privacy requirements and data sharing mandates sometimes conflict, so we constantly align processes to follow the strictest rule. Legacy infrastructures make this even more complex".

Finally, interviewees highlighted that the consequences of inconsistent or weakly governed historical data also extend to machine learning and data-driven model development. Participant L emphasised how instability in training datasets can undermine system reliability, underscoring the corrective role of governance in ensuring data integrity. As the participant recalled:

"Open source datasets are vulnerable to modifications by unknown contributors. When we trained a skin scanner model with more than 10.000 images, any change in the dataset could destabilise the system. Governance becomes a corrective mechanism to validate sources and ensure the dataset remains consistent".

This example reflects the consequences of legacy misalignment and demonstrates how governance acts as a corrective mechanism. Organisations must continuously reconcile old systems with new requirements through data cleaning, reclassification, migration activities and rigorous control procedures. This ongoing maintenance represents a significant investment of time and resources, yet it is essential for ensuring regulatory compliance and preserving the integrity of data driven initiatives.

3.1.4. Balancing speed and control

The final theme relates to the inherent tension between innovation speed and the need for rigorous control of sensitive information. Interviewees acknowledged that governance requirements often slow down the development and deployment of new technologies, particularly when multiple validation steps are required. However, they consistently framed this slower pace as a necessary trade off for preventing regulatory breaches and protecting the organisation against potential risks.

Participant G described this dynamic particularly clearly, highlighting how certain governance requirements may translate into concrete constraints on innovation speed and project execution. In regulated environments, such constraints are often unavoidable and must be incorporated into process design from the outset. As the participant noted:

"Some governance requirements become roadblocks. For example, a large bank required full on-premise deployment instead of cloud. This delayed the project for months. If a client cannot wait, we lose them. Governance can slow innovation, but we cannot avoid it because we operate under RBI. These constraints are part of the job, and we must work within them even if they reduce speed".

Respondent T emphasised that governance restrictions can, in some cases, halt innovation entirely when they conflict with the protection of client information. The respondent highlighted how such constraints are deliberately enforced to prevent misuse and unintended exposure, particularly in the context of legacy systems. As the respondent remarked:

"Governance often blocks external tools because they conflict with rules around client data. This slows things down, but it also avoids mistakes and misuse, especially on old systems where quality or masking is not guaranteed. If governance forbids something, we stop immediately. It keeps us safe".

This approach ensures that new tools, pipelines and use cases are introduced in a careful and controlled manner. Although governance procedures may delay deployment, they also enhance organisational resilience by obliging teams to verify security, reinforce privacy controls and confirm the legitimacy of data usage. This controlled pace protects the organisation from unintended consequences and fosters a disciplined form of innovation that aligns with regulatory expectations and societal standards. Within this framework, governance becomes a stabilising force that ensures data driven progress does not compromise compliance, user trust or ethical responsibility.

3.2. Data governance as a structuring mechanism

Our findings show that data governance functions as a structuring mechanism that transforms fragmented, heterogeneous and historically inconsistent datasets into coherent organisational assets. While compliance-oriented governance establishes constraints and boundaries, structuring governance focuses on the internal organisation of data: how it is defined, documented, segmented and maintained. Across interviews, participants described an environment in which the reliability of analytical outputs, the scalability of technological infrastructures and the coordination between teams all depend on the presence of clear governance structures. Firms rely on governance mechanisms to create shared definitions, organise data flows, track transformations and ensure that datasets remain usable across different systems and departments. Without this structuring layer, respondents explained that even basic tasks such as reporting, product analysis or onboarding new institutions would become inefficient or technically unfeasible. Governance therefore acts as the organisational capability that turns raw financial information into stable, standardised and interoperable resources. A key insight is that structuring work is not a one-time activity. Interviewees emphasised the continuous nature of this process, particularly in organisations operating with legacy systems or evolving regulatory expectations. The maintenance of documentation, the alignment of definitions and the correction of inconsistencies require ongoing coordination between business, technology and governance teams. We explore this structuring role across four themes: standardisation and interoperability, data quality and lineage, segmentation and architectural organisation and the remediation of legacy data.

3.2.1. Standardisation and interoperability

Standardisation emerged as a critical prerequisite for ensuring that data can circulate meaningfully within the organisation. Interviewees described how governance frameworks translate into shared definitions, common schemas and harmonised metadata structures that allow teams to collaborate around consistent information assets. Without standardisation, data remains confined within departmental boundaries, making it difficult to integrate operational systems, align reporting practices or build scalable analytical models. An essential tool supporting this work is the data catalogue, which consolidates definitions, clarifies ownership and records metadata across the organisation. Catalogues serve as a reference point for teams who need to understand what data exists, how it is structured and which rules govern its use.

Interviewees repeatedly emphasised that mapping activities, standard definitions and shared data catalogues are essential for ensuring consistent data structures across teams and systems. Participant J highlighted how these mechanisms reduce interpretative ambiguity and support coordination across organisational boundaries. As the participant outlined:

"Banks do not provide data in a homogeneous format, so our mapping ensures that clients always receive data in a consistent structure. This is essential, otherwise every partner would have to interpret the banks' data differently. When we fetch transaction data, we map the fields so that clients always receive data in a consistent structure. After the Visa acquisition, we introduced a data catalog that requires every team to declare what data they store, where it sits, its classification and whether it contains personal or biometric elements. This forces all teams to align on shared definitions and maintain consistency across the organisation".

Interviewee H highlighted the breadth and depth of the data governance work required to sustain consistent data practices across the organisation. The interviewee underscored how detailed mapping and access documentation function as shared reference points that align teams around common definitions and usage rules. As the interviewee articulated:

"We maintain a full mapping of all the data we process. There is a highly technical mapping of every element, where it sits and what it is used for. The compliance team also maintains a structured spreadsheet showing exactly who can access which type of data, for what purpose, how long we can keep it and when we must delete it. All teams rely on this structure so that everyone speaks the same language when dealing with the data".

The importance of consistent definitions also emerged in contexts involving multiple internal databases and numerous external data providers. Participant C described how centralised data preparation and mapping practices enable coordination across teams by creating a shared understanding of data flows and responsibilities. As the participant observed:

"The Economics Department collects external data, cleans it, structures it and uploads it into centralised databases. Every team must work on these validated datasets. Mapping is extremely important because it shows how data enters the company, how it moves between teams, where it is stored and what risks are involved at each step. Without mapping, it would be impossible for teams to understand each other's work"

In open finance ecosystems, standardisation also facilitates interoperability with external actors. Incoming datasets often reflect the technical limitations, naming conventions or regulatory interpretations of different financial institutions. Governance ensures that these inputs are mapped onto internal standards, enabling cross party alignment and reducing the operational friction created by semantic inconsistencies. Standardisation therefore underpins both internal coordination and ecosystem level collaboration. As Jane Barratt, Chief Advocacy Officer at MX, explained at the Open Banking EXPO Canada 2025:

"Unless data is flowing in a healthy, permissioned, secure way, we are still just preparing and laying the track"

reinforcing that governed and permissioned data flows constitute the basis upon which all subsequent structuring efforts depend.

3.2.2. Data quality assurance and lineage tracking

Quality assurance and lineage tracking constitute the second structuring pillar of governance. Interviewees emphasised that data quality cannot be assumed by default. Instead, organisations must actively validate, cleanse and document data before it enters analytical pipelines or customer facing systems. Governance routines such as quality checks, anomaly detection and periodic reviews ensure that datasets remain accurate, compliant and fit for purpose. Lineage tracking provides the historical context necessary to understand how data flows across systems. It documents where data originates, which transformations it has undergone and which business processes depend on it. This visibility is critical for replicability, regulatory reporting and internal accountability. Data catalogues again play a central role, serving as the repository for lineage information and ensuring that teams can trace the full lifecycle of each dataset.

Interviewees described how data lineage practices and quality controls are essential for making data reliable and usable across teams. Respondent C illustrated how mapping and documentation provide visibility over data flows and transformations, while also exposing quality issues that can constrain system migration and innovation. As the respondent elaborated:

"Mapping helps us understand how data enters the company, how it moves between teams, where it is stored and what risks are involved at each step. Difficulties emerge when the quality of historical data is not standardised. For example, a date column saved as text created problems when migrating to a new SharePoint because the system could not interpret it. These issues must be fixed before we can build anything new on top of the data. Documentation and mapping are necessary for teams to understand how the data is used and what transformations it undergoes".

Interviewee B emphasised the role of validation routines and anomaly detection mechanisms in maintaining data consistency and interpretability across teams. The interviewee highlighted how automated controls and clearly defined ownership structures help prevent misinterpretations and ensure shared understanding of key metrics. As the interviewee outlined:

"We build checks on the data to ensure consistency. For example, if the number of cards created goes outside the expected range, we get an alert. Every day a Python script runs queries, checks anomalies and posts results to Slack. If a metric is unclear, we contact the owners of the table.

Every dataset must have a clear definition, otherwise teams interpret things differently. This is why checks and documentation are essential to avoid misinterpretations across teams".

Quality and lineage governance are not merely technical safeguards. They constitute social agreements within the organisation about what "trusted data" means and which datasets are authoritative. By stabilising data quality at the source, governance reduces rework, limits operational risk and enhances the reliability of downstream analytics.

3.2.3. Secure segmentation

Segmentation of data across different environments and domains surfaced as another key structuring mechanism. Interviewees described architectures organised around separate spaces for production, testing, analytics and development, each with its own rules, permissions and technical constraints. This layered approach ensures that sensitive information is not exposed beyond what is strictly necessary, while still allowing teams to experiment and iterate within controlled boundaries. Segmentation also addresses performance and scalability concerns. By preventing resource intensive analytical workloads from interfering with operational systems, governance ensures that organisations can maintain service continuity while still pursuing innovation.

Interviewees illustrated how data and environment segmentation protects sensitive information while still enabling experimentation and innovation. Participant G described how access differentiation between testing and production environments allows organisations to balance security, operational stability and analytical flexibility. As the participant detailed:

"We split data into smaller tables and environments. Seventy percent of the organisation has access to User Acceptance Testing (UAT) and only thirty percent has access to production. We have a production environment for real users and a UAT environment for internal testing. Most of the organisation only accesses UAT. This protects sensitive data and lets us experiment safely without interfering with production workflows. It also ensures that resource intensive analytical workloads do not affect operational systems".

Interviewee S described how segmentation practices extend beyond development environments to the architectural level, shaping how multi-tenant systems are designed and governed. The interviewee highlighted how strict data segregation and anonymisation mechanisms enable benchmarking and analytical use cases without compromising confidentiality. As the interviewee clarified:

"We have different tenants whose data is stored separately, and we must ensure access is strictly controlled across these environments. Tenants cannot access each other's data, and governance ensures that segregation is respected at all times. We anonymise data so that tenants can benchmark themselves against competitors without knowing who the competitors are".

A further example illustrated how segmentation is reinforced through strict control over log visibility. Respondent J described how default masking and escalation procedures are embedded into system architecture to ensure that sensitive information is accessed only when strictly necessary. As the respondent pointed out:

"All logs are masked by default and the only way to access unmasked information is through a special authorisation process during an incident. This structure ensures that teams work with data in a controlled environment and only escalate to raw logs when strictly necessary. It creates a multi layered architecture where each environment has its own level of permissible access".

This structure allows organisations to concentrate most activities in user acceptance testing environments, where experimentation and validation can be conducted safely without affecting real operational data. This architectural organisation reflects a broader philosophy within data driven firms: data cannot be managed effectively as a monolithic asset. Instead, it must be partitioned, protected and orchestrated across environments designed for different forms of use. Governance defines the boundaries, permissions and intended purposes of each environment, thereby shaping the workflow and rhythm of organisational data practices.

3.2.4. Challenges emerging from standardization and lineage practices

Although data governance effectively structures organisational data through standardisation, documentation and lineage, participants highlighted persistent challenges that limit the completeness and consistency of these efforts. Several interviewees described how legacy systems make it difficult to maintain coherent datasets across different business units. Fragmentation forces teams to reconcile definitions repeatedly rather than reusing existing structures, and maintaining accurate documentation was repeatedly described as a demanding and resource intensive task.

Participant Q reflected on how mergers and acquisitions introduce significant structural complexity by bringing together heterogeneous data architectures. The participant highlighted how consolidation efforts are necessary to restore consistency and analytical reliability across legacy and newly integrated systems. As the participant observed:

"When you acquire several companies you inherit very different data structures. Serving data across old and new systems becomes extremely complex and requires a lot of work to consolidate definitions and schemas. Some datasets require reclassification; others require full restructuring. Without this effort, teams cannot rely on a single definition of the data and analytical work becomes inconsistent across the company".

Interviewee A reinforced the difficulty of aligning inherited infrastructures with contemporary governance requirements, drawing attention to the structural and historical challenges embedded in legacy systems. The interviewee illustrated how gaps in documentation and inconsistent historical usage complicate efforts to bring older environments into compliance with current standards. As the interviewee remarked:

"Legacy systems were never built with modern governance practices in mind. Documentation may be incomplete or outdated, and sometimes we need to reconstruct the entire context before aligning it with current compliance requirements. It is a constant process of reconciliation because older infrastructures contain fields that have been used differently across time, and mapping or documentation is sometimes missing. This makes harmonising old and new environments one of the most difficult tasks".

Technical inconsistencies further complicate efforts to standardise data structures and enable cross-system integration. Participant C drew attention to how poor data quality and incorrect formatting in legacy datasets can directly constrain migration activities and the introduction of new analytical or technological initiatives. As the participant pointed out:

"Legacy datasets often require manual cleaning because basic attributes like dates are stored in incorrect formats. Without fixing those issues, nothing new can be built on top of them. During our migration we saw that even simple errors, like dates saved as text, could break entire workflows. These inconsistencies must be corrected before any new analytical or technological initiative can be introduced".

Across cases, the most structural challenge was the coexistence of modern governed pipelines with older infrastructures. Participants described this as a dual reality in which new tools follow structured governance, while legacy components remain inconsistent. These tensions make structuring the organisation's data environment an ongoing rather than a completed effort. This misalignment is also evident at industry level. As Saba Shariff, Chief Strategy, Product and Innovation Officer at Symcor, noted during the Open Banking EXPO Canada 2025:

"Fintech and other industry players struggle to access data securely, reliably, and at scale"

illustrating how secure and consistent access remains a systemic challenge beyond individual organisational boundaries.

This fragmentation is mirrored internally as well. Gianluca Sanpietro, Chief Financial Officer at Soldo, observed at the Salone dei Pagamenti 2025 that:

"There is a major mismatch between what financial managers believe is clear and what employees actually perceive," adding that *"this grey area generates inefficiency, lack of clarity, and limited understanding."*

3.3. Data governance as a technical enabler

Data governance appeared not only as a framework for compliance and risk mitigation but as a deeply technical mechanism that enables organisations to build, maintain, and evolve their data infrastructures. Participants consistently highlighted that without structured governance practices, even seemingly simple technological changes become slow, unreliable, or impossible. Governance acts as an invisible infrastructure that coordinates data flows, controls system interactions, and ensures that technological processes can scale safely. Rather than being a barrier to innovation, governance emerged as a prerequisite for technical progress because it creates the stability, clarity, and predictability required for adopting new tools, deploying AI systems, automating operations, and ensuring privacy compliant architectures. The enabling role of governance became particularly clear when interviewees described concrete examples where weak governance had slowed development or caused errors and rework. Conversely, when data lineage, definitions, and access rules were clearly established, teams were able to integrate cloud platforms, scale API frameworks, and automate critical workflows with less friction. Across cases, governance served as a shared technical language through which different teams coordinate changes, avoid inconsistencies, and maintain ownership of key processes. These insights show how data governance operates at the core of technological evolution, shaping not only what organisations can build but also how securely and efficiently they can do so. In the remainder of this section, we unpack the enabling role of data governance in open banking and open finance environments, showing how it provides stable foundations for technological innovation, supports automation through clear ownership structures, embeds privacy by design into technical architectures, and coordinates cross functional collaboration around key infrastructure decisions.

3.3.1. Governance as an innovation enabler

Interviewees frequently underlined that governance creates the structural foundations needed to innovate safely. Without well defined documentation, controlled access and a unified view of core data assets, technical teams face fragmentation that makes any innovation effort unstable.

Interviewees described how innovation becomes feasible only when data definitions, schemas and governance rules are applied consistently across systems and teams. Respondent J reflected on how privacy-driven constraints may slow certain activities, while ultimately creating the conditions for more reliable and trusted innovation. As the respondent articulated:

"Governance sometimes slows us down, especially during incidents where masked logs make debugging harder. But this is intentional: we process people's financial data, and behind every transaction there is a real person whose privacy must be protected. When definitions, schemas and access rules are respected, innovation becomes smoother because the starting point is clear for everyone. Users and clients trust us because we handle the data correctly. Without consistency and strict governance, even simple integrations would break or behave differently across environments".

Participants working in highly technical domains reinforced that data quality and governance constitute essential prerequisites for innovation, particularly in contexts where system reliability is critical. Interviewee L articulated how even minor inconsistencies in training data can undermine model performance, making governance a necessary condition for safe scaling. As the interviewee articulated:

"With medical AI, even a small inconsistency in the dataset breaks the system. Open source datasets are vulnerable because people can modify them without control. Governance becomes essential to stabilise the dataset, validate the sources and make sure the data you use is correct. Without this, you cannot innovate because the model collapses. Governance is what allows us to scale the technology safely".

Others emphasised the need for stable and standardised data inputs when collaborating with external partners. Participant P described how governance practices around mapping and data quality are essential to stabilise integrations with heterogeneous data providers and to ensure that innovation efforts remain reliable and scalable. As the participant clarified:

"We work with very large datasets provided by banks, and to innovate we need them to be clean, structured and mapped. Governance helps us stabilise every integration because each bank sends data differently. Without strong governance around mapping and data quality, our models would produce completely different results for each bank, and we would not be able to innovate on top of that".

Finally, teams operating in fast-paced fintech environments reinforced that governance provides the clarity and stability required to introduce new features safely. Respondent V synthesised this perspective by emphasising how accurate, harmonised and well-governed data constitutes a necessary foundation for reliable innovation. As the respondent articulated:

"For us to deliver new features, the data first needs to be accurate and harmonised. Without that baseline, any new product becomes unreliable. Governance gives us the foundations because it tells us which data can be used, under which permission and which sources are trusted. Without that, innovation would become extremely risky".

Developers and analysts therefore rely on governance to know which datasets represent the source of truth, which systems can be integrated into the architecture and which transformations are allowed. Governance also enforces processes for validating new tools, reviewing access rights and ensuring that the integration of AI modules or cloud based services does not violate internal or regulatory constraints. These practices make innovation predictable rather than ad hoc, enabling organisations to deploy new technologies without compromising security or operational stability. In this sense, governance does not restrict innovation. It channels it, ensuring that technological progress happens on a stable, shared and compliant foundation.

3.3.2. Automation and ownership clarity

Participants described automation as one of the clearest benefits of strong governance. Automated pipelines, monitoring systems and reporting flows work reliably only when ownership of data and processes is explicitly defined. Governance establishes who controls specific datasets, who approves changes to schemas or data models, who documents transformations and who monitors the performance of automated processes. This clarity reduces ambiguity, prevents silent changes and enables infrastructure teams to automate tasks that would otherwise require manual intervention.

Interviewees repeatedly noted that automation functions effectively only when governance ensures consistent definitions and a shared understanding of data across the organisation. Participant D illustrated how systematic mapping activities create the clarity required for both coordination and automation by making data flows, responsibilities and usage rules explicit. As the participant outlined:

"Every process in the company requires a complete data mapping. This mapping must be updated yearly. We interview all teams to understand their daily activities, which data they use, when they share it and with whom. It allows us to understand the full lifecycle of the data. Governance forces us to structure data in a more organised, accessible and transparent way, otherwise it would be impossible for teams to understand each other's work. This clarity is what allows automation, because systems rely on definitions that must be consistent across the entire organisation".

Participants working with highly sensitive operational workflows reinforced that automation depends on strict governance oversight to prevent systemic errors. Interviewee O illustrated how governance mechanisms ensure consistency, traceability and accountability when automated processes have large-scale operational consequences. As the interviewee remarked:

"When you automate payroll, you cannot make mistakes. A single incorrect rule affects thousands of payslips. Governance is what guarantees consistency across the system. Every change to a calculation rule requires validation, documentation and approval. Without this, automation would be impossible because you would not know which version of a rule is correct or who owns it".

In open finance environments, automation further requires alignment across institutions to manage divergent interpretations of regulatory requirements. Respondent U illustrated how governance mechanisms establish shared rules, responsibilities and decision rights that make inter-institutional automation feasible. As the respondent clarified:

"Each institution in the region has its own interpretation of the rules, so without governance we would have hundreds of inconsistent data formats. Governance creates the shared language that allows automation. It tells us who owns the mapping, who owns the transformation, who

approves changes. Without that clarity, you cannot automate anything because every institution would break your pipelines".

Finally, Interviewee W reinforced that automation becomes fragile when data consistency is not maintained across systems and teams. The interviewee highlighted how governance mechanisms play a critical role in aligning definitions and documentation, thereby preventing recurrent automation failures and enabling sustained development efforts. As the interviewee remarked:

"Automation works only when data is consistent. If information comes in different formats, the automations break. Governance forces us to align definitions and to maintain structured documentation, otherwise we would spend our time fixing automation failures instead of building new features".

Automation becomes sustainable when systems consume data that is standardised and governed, when lineage is understood and when documentation is accessible across teams. Migrating to new warehouses, implementing automated workflows or introducing AI-based classifications all depend on this governance-induced clarity. Governance therefore strengthens technical reliability by ensuring that automated steps do not conflict with undocumented definitions, inconsistent naming conventions or team-specific data interpretations. In environments where regulatory audits are frequent, automation built on strong governance also reduces the operational burden associated with demonstrating controls and producing evidence for audits.

3.3.3. Privacy by design integration

Privacy by design arose from the interviews as a pervasive and foundational technical requirement that shapes how data is stored, processed and moved across organisational systems. Informants explained that governance determines what information can travel between environments, which elements must be anonymised or redacted and which operations require full audit trails to ensure compliance with regional privacy laws. These rules influence architectural decisions, including how databases are structured, how pipelines are built and how cloud services are evaluated before deployment. Interviewees described how governance restricts exposure of sensitive data by enforcing strict visibility controls.

Participants emphasised that privacy by design begins with strict segmentation of data across tenants and environments. Interviewee S illustrated how architectural separation, anonymisation practices and access controls are combined to prevent the exposure of sensitive information while still enabling analytical use cases. As the interviewee clarified:

"We have different tenants whose data is stored separately, and we must ensure access is strictly controlled across these environments. Tenants cannot access each other's data. We anonymise data so that tenants can benchmark themselves against competitors without knowing who the competitors are. Governance ensures that no Personally Identifiable Information (PII) is ever exposed outside its allowed boundary. It determines what information can or cannot travel between environments".

Others highlighted that access to personal information is constrained by default and subject to rigorous validation procedures embedded directly into system architecture. Participant A illustrated how role-based access controls, security testing and privacy audits function as core governance mechanisms that limit data exposure and restrict the use of external tools. As the participant specified:

"Accessing personal information is the most heavily regulated procedure. In open finance you cannot use information just because you have it. Role-based access control is fundamental. When we bring something new, we must conduct penetration testing, security audits and privacy audits. External AI tools are restricted because we cannot risk sending sensitive data into systems that do not meet our governance requirements. Every operation must respect strict privacy rules embedded directly into our architecture".

Strict visibility rules also apply to operational processes such as payment verification and Anti-Money Laundering (AML) controls. Participant F illustrated how role-based access and minimal data exposure are embedded directly into operational workflows to enforce privacy-by-design principles. As the participant clarified:

"Access to data is entirely role-based. When we verify payments for Anti-Money Laundering (AML) purposes, only specific people can see the information needed to confirm that the account owner is the same person using the service. Everything else is restricted by design. We never expose more than what is strictly required. Privacy rules determine exactly which data can travel between systems and which cannot".

For firms deploying AI-based products, privacy-by-design is not optional but embedded from the earliest stages of development. Interviewee N illustrated how regulatory approval and governance requirements shape AI integration decisions from the outset, constraining both data usage and system design. As the interviewee specified:

"We work with Financial Conduct Authority (FCA)-approved algorithms, so privacy by design is a requirement from day one. We cannot use any data unless it has gone through the correct governance process. When integrating AI tools, the first question is always whether they comply with the privacy and data-flow constraints. Without this governance layer, no AI-based financial advice would be allowed".

Together, these practices show that privacy by design is not only a legal requirement but a technical architecture in itself. It shapes how organisations design data flows, select tools, build models and structure their environments. By embedding privacy rules and protective mechanisms directly into technical processes, governance ensures that innovation remains compatible with regulatory expectations and protects users from excessive or inappropriate data exposure.

3.3.4. Cross functional collaboration

Cross functional collaboration appeared as a fundamental technical component of data governance across all cases. Interviewees emphasised that governance cannot be implemented by a single function and instead requires continuous coordination among product teams, engineering units, compliance officers, data specialists and operational roles. Governance operates as a shared organisational infrastructure that forces diverse teams to align on definitions, workflows and system behaviours, ensuring that technological decisions are made consistently and transparently. Several participants described governance as a collaborative process of negotiation across departments rather than a set of static rules.

Participant H illustrated how documentation and data mapping function as shared artefacts that support cross-departmental alignment by creating a common reference framework for data usage, access and retention. The participant emphasised how these artefacts facilitate coordination across product, engineering and compliance functions. As the participant clarified:

"We document every internal process and maintain a mapping showing exactly what data is processed, where it comes from, who can access it and for how long. This ensures that every access is controlled and that all interactions with sensitive information are fully traceable. Teams rely on this structure because it aligns product, engineering and compliance around the same definitions and rules"

Cross-functional collaboration also plays a central role in environments where organisations interact with external institutions and data providers. Participant P illustrated how governance mechanisms create the conditions for shared understanding by forcing alignment on data definitions before technical implementation begins. As the participant reflected:

"Every bank we work with has a different understanding of what a field means. Governance forces us to sit together with product, engineering, partners and banks and agree on one definition before building anything. Without that alignment, every team would implement the logic differently and the system would fail".

Internal change processes similarly require close coordination across multiple teams. Interviewee V illustrated how governance functions as a coordinating mechanism that aligns different perspectives on data across product, engineering and legal functions, reducing the risk of downstream failures. As the interviewee reflected:

"Any time we change something related to data, we need product, engineering and legal together. Each team sees the data differently, and if one is left out the change will break later in the process. Governance is the mechanism that brings these teams together so that decisions are consistent"

Collaboration also plays a preventive role in avoiding privacy and security failures when multiple teams influence the same data processes. Participant S illustrated how consistent alignment across engineering, data and compliance functions is essential to ensure that privacy-preserving mechanisms operate correctly. As the participant observed:

"Tenants can benchmark themselves against competitors, but only because anonymisation and governance rules are aligned across engineering, data and compliance. If any team handled the anonymisation differently, the results would leak information".

Interviewee G highlighted similar tensions in regulated environments:

"Sometimes the lending head interprets a guideline one way while a client interprets it differently. Governance requires us to resolve these conflicts internally and align before touching the data. Without this alignment, development would be chaotic".

This demonstrates how governance enforces a structured and collaborative decision making process that reduces future technical debt and improves system resilience. Overall, cross functional collaboration ensures that governance is applied coherently across the organisation. By fostering continuous interaction between technical and non technical teams, governance supports stable architectures, harmonised definitions, secure data flows and coordinated adoption of new technologies. This multidimensional alignment enables organisations to maintain high quality, compliant and scalable infrastructures even in fast evolving open banking and open finance contexts.

3.3.5. Challenges in embedding governance into technology adoption

While data governance enables technological innovation, automation, privacy by design and cross functional alignment, participants also highlighted significant challenges that arise when governance interacts with rapidly evolving infrastructures. A key issue concerns the uneven maturity of technical environments, where legacy structures and fragmented datasets make it difficult to introduce new tools or modify existing workflows efficiently.

Interviewees described how inconsistencies across systems complicate technological change, particularly in regulated environments where governance constraints influence architectural choices. Respondent G illustrated how misaligned data across environments increases the effort required to implement software changes, slowing innovation and shifting attention toward issue resolution. As the respondent observed:

"Some governance requirements become roadblocks. For example, a large bank required full on-premise deployment instead of cloud. This delayed the project for months. Changing the software becomes difficult because the data is not consistent everywhere and we spend time fixing issues instead of improving the system. Governance can slow innovation, but we must work within these constraints because we operate under RBI"

Participants also described challenges in integrating external technologies, particularly cloud-based tools and AI systems, within tightly governed and certification-driven environments. Interviewee A illustrated how stringent security and privacy controls, while necessary for compliance and market access, introduce coordination costs and slow down operational responsiveness. As the interviewee noted:

"Payment systems are rigid, and every new tool must pass penetration testing, security audits and privacy audits. It slows things down, but without meeting the governance requirements we cannot get certifications or sell products. Only a small group can access production, so any issue there slows down the whole company because fixes require coordination and approval".

Strict access controls further contribute to operational bottlenecks, particularly when investigative or corrective actions depend on highly restricted data domains. Participant B illustrated how legal approval processes and limited production access, while essential for protection, can slow down response times and remediation efforts. As the participant observed:

"Access to production data is extremely restricted. Most analysts work with aggregated or UAT-level data. If you need access to a sensitive domain, you must open a legal ticket and justify exactly why you need it. Sensitive domains often refuse access even after several attempts. This can slow down investigations or fixes because only a very small group can touch the data".

Legacy infrastructures were also identified as a significant source of friction in technological change processes. Interviewee W illustrated how inconsistent data across systems constrains software evolution, forcing teams to prioritise remediation over improvement. The interviewee also highlighted how governance makes such inconsistencies unavoidable by requiring them to be addressed before any progress can be made. As the interviewee observed:

"Legacy systems create strong limitations. Changing the software becomes difficult because the data is not consistent across systems, and we spend time fixing issues instead of improving the platform. Governance helps, but it also means you cannot ignore inconsistencies: you must fix them before moving forward".

Cross-institution integration introduces further challenges, particularly in open finance ecosystems where regulatory interpretations vary across partners. Participant U illustrated how fragmentation at the ecosystem level forces governance mechanisms to continuously absorb inconsistencies, slowing down innovation and delaying development activities. As the participant explained:

"The main challenge in our region is fragmentation. Banks implement the rules differently, which means your governance has to handle inconsistencies every time you integrate with a new partner. This slows innovation because your system must adapt to each new interpretation before you can even start building"

Related concerns were raised around data stability and model performance. Respondent L illustrated how governance-driven validation processes are critical to prevent system failure when working with sensitive or complex datasets. As the respondent remarked:

"When a dataset is not stable, we cannot innovate. Even a small inconsistency in the data can break the model. Governance requires us to validate every change and every source. This slows things down, but without it nothing would work reliably".

These coordination costs demonstrate that governance-enabled collaboration is effective but not effortless. Constraints arising from privacy requirements, organisational silos, fragmented infrastructures and regulatory diversity create friction that must be continuously managed. Nonetheless, these mechanisms ensure that innovation proceeds safely, transparently and in line with organisational and regulatory expectations.

3.4. Data governance as a leveraging mechanism for use case development

Across the interviews, data governance arose not only as a set of controls but as a structuring mechanism that shapes how new data driven use cases are conceived, validated, scaled, and monitored. Rather than acting as a mere constraint, governance enables organisations to innovate responsibly by clarifying what types of data can be used, under which circumstances, with which safeguards, and through which processes. This role becomes particularly salient in open banking and open finance settings, where the combination of sensitive financial information, regulatory scrutiny, and inter organisational data sharing requires a disciplined yet enabling governance framework. Beyond ensuring compliance, governance influences the creative and operational phases of innovation by defining the boundaries within which teams ideate, experiment, and deploy new use cases. Interviewees described governance as an "engine" that directs analytical work toward viable and legally sound solutions, determining which datasets are trustworthy, which models can be operationalized, and which forms of data enrichment are permissible. Because open finance ecosystems involve constant data exchanges between banks, fintechs, and third party providers, governance also functions as a coordination mechanism that aligns expectations, standardises data quality requirements, and creates shared rules for interoperability. This ensures that new use cases can move smoothly from experimentation to production without compromising privacy, security, or regulatory commitments. We identify four mechanisms through which data governance leverages organisational capabilities to support the development of robust and compliant use cases: compliant innovation, data driven scalability, AI and analytics governance, and compliance centred development.

3.4.1. Compliant innovation

Compliant innovation refers to the way in which governance structures shape the early phases of use case ideation and determine whether an idea can move from concept to prototype. Multiple interviewees described how governance intervenes at the very beginning, by deciding which datasets are permissible, which require masking or anonymisation, which demand explicit user consent and which are strictly prohibited for certain business purposes. This produces an environment where innovation is encouraged, but only within clearly defined legal and ethical parameters.

Interviewees consistently emphasised that purpose limitation and consent validation represent the first boundary condition for any new idea involving data usage. Participant F illustrated how governance mechanisms act as an immediate filter at the ideation stage, preventing initiatives from moving forward when legal or consent-related requirements are not satisfied. As the participant observed:

"We always check whether the purpose we have allows us to use the data. That is the very first step before doing anything. If the purpose is not aligned with the consent, we simply cannot touch the data. Even if the idea could create value, governance blocks it immediately. It doesn't matter how good the use case looks: if the purpose is not correct, we cannot proceed".

These checks extend to masking requirements, AML constraints and the permissions embedded in the user journey, creating a strict but necessary filter at the ideation stage.

Participants working in regulated environments reinforced the central role of compliance alignment in determining whether new ideas can progress beyond the initial stage. Interviewee G illustrated how regulatory interpretation acts as an early gatekeeper for innovation, requiring alignment before any data-related activity can begin. As the interviewee indicated:

"Whenever we build something new, we need to ensure it complies with RBI guidelines. Some clients want to innovate faster, but if the guidelines do not allow it, we cannot proceed. Before touching the data we must align interpretations; governance forces that alignment. If the

guideline doesn't support the use case, we are not allowed to build it, even if it is commercially interesting".

This example illustrates how regulatory boundaries shape innovation cycles from the outset, filtering ideas before they enter development and constraining experimentation to what is institutionally and legally permissible.

The importance of compliant innovation was particularly evident in sectors handling sensitive or high-risk data, where governance determines not only what can be built but whether systems can operate on real data at all. Respondent L illustrated how governance structures stabilise datasets, validate sources and shape system architecture to ensure both regulatory compliance and safety. As the respondent clarified:

"With medical data you cannot use anything unless governance allows you to. Governance becomes essential to stabilise the dataset, validate the sources and make sure the data you use is correct. It shapes the entire architecture: what we process, how we process it and what we store. Without these rules we would not be able to deliver value, because the models would not be allowed to run on real user data and the output could be clinically unsafe".

A similar dynamic emerged in open banking ecosystems, where the interpretation of consent and permissible data processing varies across institutions. Interviewee P illustrated how governance mechanisms act as an early filter for innovation by determining whether proposed use cases align with partners' interpretations of consent and data usage. As the interviewee observed:

"We work with very large datasets provided by banks, and to innovate we need them to be clean, structured and mapped. When we explore new ideas, the first step is to check whether the bank's interpretation of consent and data use allows the use case. If it doesn't, the idea cannot move forward, even if it creates value. Governance defines what we can build and when we can build it".

Taken together, these insights show that compliant innovation is not about limiting creativity. Instead, governance acts as an early filter that aligns innovation with legal obligations, customer expectations and organisational risk appetite. It ensures that new use cases are both desirable and deployable. This connection was captured clearly by Saba Shariff, Chief Strategy, Product and Innovation Officer at Symcor, who stated at the Open Banking EXPO Canada 2025 that:

"Trust and innovation are foundational. Without trust, you cannot achieve innovation"

illustrating how transparency underpins both responsible governance and competitive advantage.

3.4.2. Data driven scalability

Scalability appeared as one of the most frequently cited benefits of data governance. Interviewees consistently emphasised that the ability to scale use cases across products, teams or markets depends far more on the quality and structure of underlying data than on the creativity of analytics teams. Governance provides this foundation by standardising definitions, ensuring coherent lineage documentation, harmonising taxonomies and enforcing consistent data transformations across shared datasets.

Interviewees noted that scalability becomes achievable only when data is systematically validated, mapped and centralised across the organisation. Participant C illustrated how centralised data preparation and shared mappings enable reuse across multiple use cases, reducing duplication and accelerating innovation. As the participant outlined:

"The Economics Department collects external data, cleans it, structures it and uploads it into centralised databases. Every team must work on these validated datasets. Mapping shows how data enters the company, how it moves between teams, where it is stored and what risks are involved at each step. Once the definitions and mappings are aligned, we can reuse the same dataset for reporting, commercial analytics or benchmarking. Governance lets us build new use cases quickly, because the data is already validated and structured".

In open finance ecosystems, scalability depends on the ability to standardise heterogeneous data inputs originating from external institutions. Participant R illustrated how governance-driven mapping practices create a uniform abstraction layer that enables reuse and reduces integration complexity across partners. As the participant indicated:

"Banks all send data differently, so we remap everything to a standard structure. Without this mapping, clients would not be able to build stable use cases, because every bank would require a different logic. Governance creates the uniform layer that lets developers build features once and reuse them across institutions".

Interviewees also stressed that scalability collapses when data is inconsistent or insufficiently validated, particularly when systems are extended beyond their initial scope. Respondent T illustrated how governance functions as a protective mechanism that preserves scalability over time, even when it introduces short-term friction. As the respondent observed:

"If you build a use case on top of fragmented or inconsistent data, it will not scale because the model or workflow will break once you try to extend it. You need data that is clean, validated and structured. Governance is what protects scalability, even if it sometimes creates friction in the short term".

Fintech teams highlighted that governance can accelerate feature expansion when a solid and consistent data foundation is in place. Interviewee M illustrated how clean baselines reduce rework and enable faster scaling of new functionalities. As the interviewee indicated:

"When the baseline is clean, we can scale features much faster because we don't need to fix definitions or rebuild datasets every time"

Subsequently, reinforced that controlled data scopes support reliability across teams, noting that:

"Analysts only work with the relevant dataset, which makes use cases reliable. Without governance, fragmentation would explode and no use case would scale because the underlying data would be inconsistent across teams".

Overall, interviewees observed that organisations with strong governance practices could scale new models or analytical applications quickly, often by reusing existing pipelines, shared metrics and pre validated datasets. By contrast, firms with weak or fragmented governance faced duplication, inconsistencies and significant manual reconciliation. These issues make scaling extremely challenging because every new use case requires rebuilding or cleaning the underlying data. Data driven scalability is therefore possible because governance creates a stable data foundation upon which use cases can expand efficiently and consistently.

3.4.3. AI and analytics governance

AI related use cases displayed some of the most rigorous governance structures observed in the dataset. Participants repeatedly mentioned that models involving behavioural, transactional or personally identifiable information require predefined approval workflows, strict input validation, detailed documentation of assumptions and controlled experimentation environments. These safeguards ensure that AI systems remain explainable, auditable and consistent with the organisation’s risk appetite.

Interviewees explained that governance plays a foundational role in defining the technical boundaries within which AI systems are allowed to operate, particularly in regulated environments. Participant S illustrated how segregation, anonymisation and controlled testing environments constrain model behaviour and data access, ensuring compliance while still enabling AI deployment. As the participant outlined:

“We have different tenants whose data is stored separately. AI models can only use the anonymised datasets, and they are tested in safe environments. We ensure no identifiable information is ever passed into the modelling layer. Governance defines exactly what the model can see, what it cannot touch and how it should behave. If the model needs a new variable, governance must approve it. Without these restrictions, we could not deploy AI systems in a regulated environment”.

Strict control over sensitive information was repeatedly emphasised as a core governance mechanism in regulated data environments. Interviewee G illustrated how masking, environment separation and restricted production access enable experimentation while preventing exposure of personally identifiable information. As the interviewee clarified:

“PII must always be masked. As a Third-Party Service Provider (TSP), we cannot store raw PII because if we do, we must purge it. Most employees only have access to the UAT environment, where the data is either synthetic or masked. This allows teams to test ideas, run models and experiment without touching real sensitive information. Only a very small group can work with production data. Governance decides what can move between environments, what cannot and which transformations are allowed”.

Ensuring the reliability of analytical outputs requires continuous monitoring and clearly defined accountability structures. Participant B illustrated how automated controls, ownership assignment and shared definitions are essential to prevent model drift and inconsistent analytical results. As the participant indicated:

“We run automatic checks every day. If something goes outside the expected range, we get an alert. Every dataset must have an owner and a clear definition. Without that, models behave inconsistently and analytics outputs become unreliable. Governance ensures that the data feeding the model is accurate, validated and controlled”.

Finally, interviewees stressed that privacy boundaries directly constrain both the behaviour and feasibility of AI models in regulated environments. Respondent S illustrated how governance-defined privacy frameworks establish the technical limits within which models can safely operate, preventing unintended information leakage across tenants. As the respondent observed:

"AI models can only operate within the privacy framework defined by governance. If anonymisation is not consistent, cross tenant outputs could leak information, so governance determines the boundaries of what is technically possible".

Together, these insights show that AI and analytics governance stabilise the development of advanced use cases by introducing safeguards that minimise risks of bias, misuse or unintended consequences. Governance enables teams to innovate while ensuring that every model, workflow and analytical output remains compliant, transparent and aligned with ethical and regulatory expectations.

3.4.4. Compliance centred development

While governance enables many forms of innovation, it also defines strict boundaries that development teams cannot cross. In compliance centred development, governance functions as a decision-making mechanism that determines which use cases are feasible and which pose unacceptable legal or ethical risks. Interviewees explained that compliance is not an optional check but a hard constraint that filters out ideas unable to meet regulatory, ethical or consent requirements.

Participants emphasised that compliance-centred development begins with strict purpose validation, which functions as the first and most decisive checkpoint in the innovation process. Interviewee F illustrated how governance operates as an early gatekeeper by assessing consent alignment and regulatory feasibility before any technical work can begin. As the interviewee clarified:

"We cannot build a use case if the data purpose is not correct. Even if the use case could create value, if the consent does not allow it, governance blocks it immediately. Purpose comes first. If the purpose is aligned, we then check whether the data needs masking, whether AML rules allow us to use it and whether the data flow matches regulatory expectations. If any of these elements fail, the use case stops right away".

This demonstrates how governance acts as an early gatekeeper that prevents non compliant initiatives from progressing beyond ideation.

Interviewees also highlighted that compliance-centred development plays a critical role in protecting fairness and preventing discriminatory outcomes in data-driven innovation. Participant W illustrated how governance mechanisms function as ethical safeguards by blocking features that could undermine equality, transparency or consumer protection, even when they promise commercial gains. As the participant observed:

"Some features would give more flexibility but would not be fair for customers. Governance requires that we treat everyone equally. We cannot deploy models or personalised pricing that would create inequality or reduce transparency. Even if the feature could increase revenue, governance blocks it".

These constraints ensure that innovation aligns with principles of equity and consumer protection.

Strict purpose and consent validation was also shown to shape the architecture of new products from the earliest design stages. Interviewee D illustrated how governance frameworks define clear development boundaries by determining whether personal data can be moved or processed at all, effectively filtering ideas before they reach prototyping. As the interviewee observed:

"We never move personal data unless there is a clear and documented purpose. If the purpose is not necessary or not aligned, the idea cannot even enter the prototype phase. Governance defines the boundaries of development".

This mechanism prevents teams from investing in features that would later be deemed unlawful or unethical, reinforcing governance as an ex-ante constraint that structures product design rather than a control applied after development.

Beyond purpose and fairness considerations, compliance-centred development also requires robust documentation and full auditability of data flows. Participant E illustrated how governance frameworks define the boundaries of what can be built and deployed by ensuring that every use case can be traced, justified and validated during supervisory reviews. As the participant observed:

"We follow strict mapping and documentation practices because during audits we must show exactly where any issue originated and how the data flows across systems. If a use case is built on data that is not properly governed, we would not pass the audit. Governance defines what is possible to build and deploy".

Compliance centred development therefore ensures that new use cases evolve responsibly. It guarantees that innovation remains aligned with core principles of consent, fairness, transparency and legal consistency, even when commercial incentives might pull in other directions. Governance acts as a stabilising force that protects both users and organisations, ensuring that innovation remains trustworthy and sustainable over time.

3.4.5. Challenges in scaling and validating data driven use cases

Even though data governance supports the development of new use cases by ensuring compliance, scalability and analytical reliability, interviews revealed several challenges that influence the speed and flexibility of innovation. One recurring difficulty concerns consent boundaries and purpose limitations. Participants stressed that use cases must fully align with consent conditions, and this sometimes prevents teams from exploring ideas that appear valuable. This strict filtering protects user rights but reduces experimentation space.

Interviewees highlighted how technical inconsistencies complicate the reuse of datasets for new analytical applications, increasing the effort required to extend existing systems. Respondent G illustrated how misaligned definitions across environments force repeated validation cycles, diverting resources away from development and slowing innovation. As the respondent observed:

"Changing the software becomes difficult because the data is not consistent everywhere, and we spend time fixing issues instead of improving the system. Every time we build something new, we need to validate the data again because the definitions may not match. Governance slows innovation but without it we risk breaking the system or violating guidelines".

Participants working with external partners faced similar difficulties when attempting to scale analytical use cases across heterogeneous data sources. Interviewee P illustrated how divergent interpretations of data attributes across institutions require repeated validation efforts, slowing down expansion even when technical logic is already in place. As the interviewee observed:

"We cannot scale use cases quickly because we need to check again what every field really means. Different banks interpret the same attribute differently, so we must validate each definition before building anything new. This slows scaling because even if the logic is ready, the data is not aligned".

AI teams described delays introduced by governance approval cycles required for adopting new tools or models. Participant N illustrated how multi-layered validation involving engineering, compliance and security functions postpones experimentation, even when technical feasibility is already established. As the participant observed:

"Every new AI tool has to be checked by engineering, compliance and security. Even if the tool works, we cannot use it until governance approves it. Sometimes we have ideas but we must wait for the validation cycle to finish before we can prototype anything". These additional checks ensure model safety but reduce development speed.

Governance restrictions also limit the adoption of external tools that could otherwise accelerate experimentation, particularly in regulated infrastructures. Participant T illustrated how strict enforcement of governance rules requires teams to prioritise risk prevention over development speed, even when faster alternatives are available. As the participant remarked:

"If governance forbids a tool or a process, we have to stop immediately even if it would make development much faster. This slows things down, especially when experimenting with new ideas. But it also prevents mistakes and avoids misuse of client data".

Finally, firms operating across multiple infrastructures described the difficulty of maintaining coherence between legacy and newly developed environments. Interviewee Q illustrated how misaligned definitions across systems can undermine the reliability of use cases unless governance mechanisms enforce consolidation before deployment. As the interviewee observed:

"Serving data across old and new systems becomes extremely complex. If we do not consolidate definitions, the same use case will behave differently in different systems. Governance helps, but everything must be aligned before the use case can go live".

These challenges show that while governance strengthens the reliability and safety of use case development, it also introduces delays, validation loops and coordination costs that limit speed and experimentation. As a result, organisations must continuously balance innovation with the structural constraints required to maintain compliance, consistency and user protection.

3.5. Data governance as a driver of responsible and sustainable outcomes

Data governance proved to be a foundational mechanism for ensuring that innovation in open banking and open finance develops in a responsible, transparent, and socially beneficial manner. Beyond compliance, governance structures influence how organisations think about fairness, inclusion, environmental efficiency, and long-term accountability. Participants consistently explained that responsible use of financial data cannot be achieved through technology alone but requires well designed governance that ensures transparency, protects individuals, and minimises risks in data intensive environments. The following sections illustrate four ways in which data governance acts as a driver of responsible and sustainable outcomes: transparency and traceability, ethical data management, inclusion through data integrity, and sustainability through data efficiency.

3.5.1. Transparency and traceability

Informants consistently described transparency and traceability as essential for creating a trustworthy ecosystem in which data movements are visible, verifiable and accountable. Financial information often moves across multiple systems and organisational units, and governance ensures that these flows remain interpretable rather than opaque. Interviewees emphasised that visibility across the entire data lifecycle strengthens both regulatory compliance and internal confidence, enabling organisations to reconstruct data flows, validate transformations and demonstrate control to external authorities.

Participant U described how governance mechanisms guarantee full oversight of data exchanges by embedding validation, traceability and accountability into data flows across the organisation. The interviewee illustrated how governance frameworks support both internal monitoring and regulatory compliance by ensuring that data exchanged through APIs remains controlled and auditable. As the participant observed:

"Our team interacts closely with data governance frameworks at multiple levels. We validate that data exchanged via APIs is accurate, complete and auditable. Governance provides a foundation for consistency, security and accountability across the company's data assets. It enables clear data lineage tracking from source to destination, which is essential for internal monitoring and for demonstrating to regulators that every data flow is controlled and traceable".

Audits further strengthen the need for end-to-end traceability, particularly in highly regulated contexts subject to continuous supervisory scrutiny. Interviewee E illustrated how mapping and documentation practices are mobilised to demonstrate control and accountability during both internal and external audits. As the interviewee observed:

"We follow strict mapping and documentation practices because as a PSD2 provider we are controlled by national and European authorities. We have internal and external audits every year. During audits we must show exactly where any issue originated and how the data flows across systems".

Participant W reinforced this point by highlighting the role of governance in providing a systematic structure for managing data assets across complex and evolving systems. The participant emphasised how maintaining up-to-date lineage documentation is essential for ensuring shared understanding, traceability and control over data flows. As the participant observed:

"Governance provides a systematic structure for managing data assets. One of the challenges is ensuring lineage documentation remains up to date, tracking how data is generated and transmitted and providing a shared understanding of how data flows from various systems".

Transparency also increases internal reliability by ensuring that teams operate only on the information they are authorised to access. Interviewee M illustrated how governance-driven access controls reduce operational risk and prevent unnecessary exposure of sensitive data, even during routine analytical activities. As the interviewee observed:

"Governance ensures analysts only work with the relevant part of the dataset instead of seeing personal details they do not need. If we allowed everyone to see everything, even basic queries would become messy and risky".

Finally, participants stressed that external partners often require full traceability as a prerequisite for approving integrations, particularly in regulated or data-sensitive collaborations. Respondent T illustrated how the ability to demonstrate auditable and transparent data flows becomes a condition for enabling innovation across organisational boundaries. As the respondent pointed out:

"Clients cannot innovate unless their data flows are fully traceable and auditable. If we cannot show the full trail of how data is processed, partners will not approve the integration".

By making data movements observable and reconstructable, governance strengthens organisational reputation, reduces the risk of undocumented processing and supports the creation of a transparent and trustworthy data ecosystem.

3.5.2. Ethical data management

Ethical data handling emerged as a central priority across interviews, especially given the sensitivity of financial, transactional and behavioural information used in open finance ecosystems. Participants explained that governance frameworks do not simply regulate access to data but actively shape how analytical processes, model development and data driven decision making are conducted. Ethical safeguards are embedded throughout the workflow to ensure that outputs remain fair, auditable and aligned with regulatory expectations.

Interviewee E illustrated the organisational emphasis on ethical conduct by showing how regulatory obligations are translated into concrete governance routines that shape both internal behaviour and external relationships. The interviewee highlighted how ethics are operationalised through employee training, access restrictions and rigorous third-party due diligence. As the interviewee remarked:

"As a licensed PSD2 provider, we have to apply regulations to the way we store and process data. We apply sufficient security measures and we educate every employee about access to production data. Every employee gives the banker's oath, swearing they will not use data for any harm to any company or client. We also evaluate every provider through due diligence. We cannot use a random provider; we need one that meets all ethical data requirements and regulations".

This example shows how ethical principles become embedded in governance routines, extending from internal behaviour to third party oversight.

The role of governance in safeguarding ethical data use was particularly evident in open finance contexts, where large volumes of sensitive information require strict alignment between purpose, consent and reporting obligations. Participant U illustrated how governance frameworks define clear boundaries for use case development while ensuring transparency, accountability and trust in data handling practices. As the participant observed:

"Data governance helps define the boundaries and quality requirements for use case development. It supports ethical data usage, ensuring use cases respect consent and purpose limitations. Governance ensures transparency and accountability in how data is collected and reported and provides trusted, auditable data for disclosures".

These structures prevent misuse and ensure that innovation remains consistent with regulatory and societal expectations.

Interviewees also emphasised that ethical data management is grounded in rigorous controls over access to sensitive information, particularly in highly regulated open finance environments. Engineer A illustrated how role-based access, security testing and governance-led validation act as foundational safeguards before any new tool or capability can be introduced. As the engineer specified:

"Accessing personal information is the most heavily regulated procedure. In open finance you cannot use information just because you have it. Role based access control is fundamental. When we bring something new, we must conduct penetration testing, security audits and privacy audits. Engineers cannot use new tools until governance evaluates the privacy risk, the data flow and the compliance gap".

Here, ethics and security are tightly interlinked, with governance acting as a gatekeeper for any new tool or model.

Ethical management also influences the development and validation of AI and analytical use cases by shaping which data assets can be used and under what conditions. Participant W illustrated how governance frameworks support responsible innovation by ensuring that data-driven initiatives comply with legal, ethical and sustainability requirements. As the participant observed:

"Governance supports the creation of data driven use cases by ensuring legal and ethical use of sensitive information. It identifies high value, compliant data assets for analysis and ensures that sustainability related data is accurate, auditable and transparent".

This illustrates how governance aligns responsible AI development with broader organisational and ESG objectives.

Finally, ensuring ethical interpretations of data emerged as a recurrent theme across interviews, highlighting the risks associated with ambiguity in data meaning. Respondent B illustrated how the absence of shared definitions can lead to inconsistent interpretations and potentially harmful decisions. As the respondent observed:

"Checks and documentation are essential because if we do not know exactly what each field means, different teams will interpret things differently, and this leads to inconsistent and risky decisions".

Ethical data management therefore depends not only on privacy controls but also on shared understanding and clarity around data meaning.

Together, these findings highlight that ethical data management is not an aspirational ideal but a concrete governance practice embedded across daily operations. By enforcing rigorous verification steps, restricting high risk processing and ensuring consistent interpretations of data, governance helps organisations protect users, maintain public trust and support responsible innovation.

3.5.3. Inclusion through data integrity

Across interviews, data governance emerged as a key enabler of more inclusive financial services, particularly in contexts where individuals lack traditional credit histories or face structural barriers to financial access. Participants emphasised that inclusion is not achieved merely by using more data, but by ensuring that alternative data is governed, validated and processed in a way that protects users from bias, discrimination and misuse.

A clear illustration of this dynamic emerged from interviewee G, who described the development of alternative credit scoring models designed to expand financial inclusion in contexts where traditional credit histories are unavailable. The interviewee illustrated how governance frameworks structure the collection and use of transactional data to ensure that innovative credit products remain compliant with consent and purpose limitations. As the interviewee outlined:

"G has the right to access transactional data like how much you have spent on food, and based on multiple data points we develop an alternative credit score which can be used by different companies to underwrite loans. This kind of alternative credit product has been pioneered to reach parts of the market where a traditional credit score does not exist. India is still a developing nation and credit penetration is very low, so we came up with this idea to penetrate the market where credit scores do not exist".

This example highlights how governance enables socially oriented innovation by shaping how sensitive transactional data is collected, processed and reused, ensuring that efforts to serve underserved borrowers remain aligned with regulatory, consent and purpose requirements.

The same organisation described rigorous fairness validation procedures to ensure ethical use of alternative data. As the interviewee noted:

"Once we have an alternative credit score, we run a test along with the client. Usually it is done in batches or in parallel with two streams of users. For 500 users we use the alternative credit score and for 500 we use the bureau score. These metrics help us validate the accuracy

and behaviour of the alternative score. This A/B testing is how we make sure the models are performing correctly before using them at scale".

These governance procedures ensure that new credit scoring approaches do not inadvertently introduce biased outcomes.

Inclusion also emerged in the broader context of ESG and responsible innovation, where governance frameworks extend beyond individual products to shape organisational accountability. Participant U illustrated how data governance supports ESG initiatives by ensuring transparency, auditability and ethical data management across open finance ecosystems. As the participant observed:

"Data governance supports ESG initiatives by ensuring transparency and accountability in how data is collected and reported and by providing trusted, auditable data for ESG disclosures. It enables ethical data management, aligning with principles of privacy, inclusivity and fair access to financial data. In the Open Finance context, it ensures responsible innovation, allowing financial inclusion use cases while maintaining strong data ethics".

This reflects how governance contributes to social inclusion not only at the product level but at the level of corporate accountability.

Other participants described inclusion challenges in contexts involving legally legitimate yet high-risk customer groups, highlighting how governance can mitigate exclusion driven by overly rigid risk assessments. Respondent E illustrated how governance frameworks enable more nuanced decision-making by distinguishing between legal risk and illegal activity, thereby supporting ethical inclusion. As the respondent observed:

"We regularly review what kind of customers we accept. If we see a high-risk customer that is high-risk but not unacceptable, we allow them to use our services. A practical example is that in the Netherlands prostitution is legal, and some financial providers tend to reject these individual entrepreneurs. We used to reject them as well, but now, if we see that the customer does not have signs of illegal activity, we accept them. This is an important ethical practice because financial transparency is important for women involved in prostitution. You do not make their life better if you do not allow them to use your tool".

Here, governance structures ensure that financial exclusion is not reproduced through indiscriminate risk aversion, allowing organisations to balance regulatory compliance with ethical responsibility and social inclusion.

Considering these insights collectively, it becomes clear that data integrity, fairness validation and rigorous governance processes enable organisations to extend financial opportunities responsibly to underserved or marginalised groups. Governance ensures that inclusion is achieved not through unchecked experimentation, but through well controlled, transparent and ethically grounded use of data.

3.5.4. Sustainability through data efficiency

Environmental sustainability emerged across interviews not as a standalone initiative but as a natural consequence of disciplined data governance. Participants highlighted how governance driven practices such as data minimisation, compute optimisation, controlled retention and efficient architectural design directly reduce the environmental footprint of digital infrastructures. By limiting unnecessary storage, reducing server workloads and preventing redundant processing, governance enables lighter data pipelines that consume fewer resources while remaining secure and compliant.

Participant H from a digital payments provider illustrated how efficiency, cost optimisation and sustainability objectives are closely linked through data governance and architectural optimisation. The interviewee highlighted how continuous improvement efforts focused on data flows, storage and infrastructure usage can significantly reduce technical costs and energy consumption, even in phases of business growth. As the participant observed:

"We are on a continuous improvement basis, meaning that we have to look after our efficiency and costs and, in our case as a fully digital company, it means optimising all data flow, all storage, all energy consumption, all databases and so on. To give you an example, we reduced by at least 50% our cost in technical processing last year. We made everything we could to make the storage space lower and the way we were using servers lower. Even increasing our business by 20–30%, we managed to decrease our technical cost by 50%".

This example highlights how optimising data architectures not only improves operational efficiency but also directly reduces energy consumption and technical waste, reinforcing the link between governance, performance and sustainability.

The environmental impact of data governance was also evident in discussions surrounding data backup, redundancy and resilience requirements. Interviewee G illustrated how regulatory obligations related to data preservation and continuity can increase infrastructure needs, thereby affecting environmental sustainability. As the interviewee observed:

"If there are any rules related to data saving, like having two or three backup data centres so that data is not lost, these impact the environment more because you need more data centres in the back end. Anything related to the way we are storing or transferring the data might not impact the environment

a lot if it does not involve a lot of computational power, but when you are required to have multiple backup centres, that clearly has an environmental effect".

This highlights the tension between regulatory resilience requirements and environmental footprint.

Other participants stressed that sustainability initiatives depend fundamentally on the integrity and reliability of ESG-related information. Interviewee W illustrated how data governance frameworks ensure that environmental metrics are accurate, auditable and transparent, thereby supporting credible reporting and accountability. As the interviewee observed:

"Data governance is crucial in the pursuit of ESG objectives by ensuring that sustainability related data is accurate, auditable and transparent. Governance frameworks ensure the accuracy of data on energy usage, emissions and waste. Standardised data collection and validation enable credible sustainability reporting and enhance stakeholder trust and accountability".

Here, data governance enables organisations not only to manage their environmental footprint, but also to measure it reliably.

Finally, participant A linked sustainability outcomes to the reliability of the underlying data governance system, emphasising how trustworthy data enables more informed and implementable ESG decision-making. The participant illustrated how governance supports realistic policy formulation by ensuring that sustainability metrics accurately reflect how data is produced, consumed and operationalised within organisations. As the participant observed:

"When policymakers know that the data is reliable, they can confidently define new strategies. Policymakers can define realistic ESG goals because they understand how the metrics and the data are really consumed by engineers and developers. Strong governance aligns with ESG because it promotes realistic, data driven goal setting rather than targets that are disconnected from the underlying systems".

This shows how data governance contributes to sustainability by shaping informed ESG policymaking and avoiding symbolic or unimplementable environmental commitments.

Taken together, these insights show that data governance contributes to environmental sustainability through operational efficiency, disciplined management of data flows, accurate reporting practices and informed decision making. Sustainability emerges not as a separate initiative but as a natural consequence of governance mechanisms that regulate how data is collected, processed, stored and used throughout the organisation.

4 Discussion

The purpose of this study was to examine how data governance mechanisms influence the orchestration of digital resources within organisational processes. We introduce data governance as an essential capability to orchestrate resources in data-intensive and regulated environments such as open banking and open finance. This research contributes to the strategic management literature by explaining how governance mechanisms allow the interdependence and multiple stakeholders, the alignment of interests, and the allocation of responsibilities and accountability (Aguilera et al., 2008). The findings show that governance mechanisms operate at the process level, shaping managerial action and influencing how digital resources are structured, combined and deployed in practice.

Consistent with ecosystem-based views of governance that emphasise the allocation of control rights rather than hierarchical ownership (Jacobides et al., 2018), the findings indicate that governance mechanisms define who can access data, for what purposes and under which conditions across organisational boundaries. In open banking and open finance contexts, where value creation depends on coordinated action among heterogeneous actors, these mechanisms enable collaboration while preserving control by translating ecosystem-level interdependencies into clear processes.

By adopting the lens of resource orchestration theory (Sirmon et al., 2011), the analysis shows how data governance becomes a capability which recurring routines establish mechanisms for structuring, bundling and leveraging resources in a compliant manner. This study advances a perspective in which data governance becomes the central architecture through which organisations stabilise their resource environment, integrate heterogeneous competencies and direct the mobilisation of capabilities toward responsible and scalable innovation. Data governance does not simply delimit what is allowed. It actively shapes how opportunities are recognised, how risks are assessed and how organisational actors coordinate their actions across technical, legal and strategic domains. In this sense, governance operates as a mediating mechanism between uncertainty and action, allowing organisations to engage in innovation without losing control over exposure, accountability and legitimacy.

These findings align with managerial perspectives that view governance as a value-enabling architecture rather than a purely compliance-oriented function (McKinsey, 2020). When embedded in everyday processes, data governance supports managerial decision-making by creating auditable conditions for mobilising and scaling digital resources within acceptable boundaries. It emerges as a dynamic and evolving system of risk control that governs how organisations interpret, structure, combine and deploy their data resources.

Financial organisations operating in open banking and open finance are embedded in ecosystems characterised by multiple and often overlapping regulatory regimes, interdependent digital infrastructures and persistent uncertainty regarding the legitimacy, ethical implications and systemic risk profiles of data driven initiatives. Data flows frequently cross organisational and national boundaries, interact with evolving interpretations of privacy and consumer protection, and support increasingly automated and algorithmic forms of decision making. Under these conditions, uncertainty extends beyond technical feasibility and market demand. It concerns how decisions should be coordinated, how responsibilities should be allocated and how risk should be managed when data is shared, recombined and reused across institutional boundaries.

These conditions generate ambiguity not only about what can be done with data, but also about how organisations can act coherently and responsibly in environments where failure may have financial, reputational and systemic consequences. Data governance provides the interpretive and operational lens through which organisations can transform such ambiguity into structured knowledge and actionable procedures. By clarifying meanings, defining responsibilities and embedding controls into everyday practices, data governance enables organisations to translate regulatory principles and ethical expectations into concrete organisational routines. As such, it is not merely an antecedent to innovation, nor a reactive safeguard applied after technological decisions have been made. It is a constitutive element of the innovation process itself, shaping both the direction and the feasibility of organisational change.

The interviews indicate that, within the current data-driven evolution, data assets cannot be effectively leveraged through technology alone. Instead, data and digital technologies jointly require a governance framework that provides control, coordination, and accountability. The proposed framework illustrated in Figure 4 shows how managers can govern data resources to transform governance challenges into sustainable outcomes, while also highlighting key constraints such as legacy system fragmentation, trade offs between control and flexibility, and increasing coordination and compliance costs. Importantly, this enabling role operates along two complementary dimensions of value creation.

On the one hand, governance supports cost control by improving data quality, reducing redundancy, limiting rework, and increasing operational efficiency. On the other hand, the same governance mechanisms facilitate revenue generation by making data trustworthy, scalable, and reusable across products and services, thereby supporting the design, validation, and scaling of new use cases. In this sense, data governance emerges as a foundational mechanism that links control, innovation, and value creation in open banking and open finance ecosystems.

They also support ESG-aligned performance by enabling traceable data processes, long-term regulatory alignment, and greater accountability and transparency. Data governance orchestration thus creates dynamic conditions that allow managers to enable data-driven initiatives under legitimation and regulatory constraints while continuously refining use cases in response to evolving risks, technologies, and organisational practices.

The discussion then elaborates on two dynamic effects that characterise the interaction between governance and organisational practice, namely the enabling effect and the evolving effect. Together, these effects explain how data governance both facilitates innovation and evolves through the experience generated by innovation itself. Finally, the Discussion introduces the notion of a radical status recategorization. Building on the empirical findings and theoretical analysis, it argues that data governance should be repositioned at the centre of organisational theory, not as a compliance function or a control mechanism detached from value creation, but as a dynamic system of risk control that mediates the relationship between risk and innovation in data intensive ecosystems. By doing so, the study contributes to a deeper understanding of how organisations can pursue process engineering innovation in open finance while maintaining coherence, trust and long-term sustainability.

Accordingly, managers should design data governance as a strategic architecture rather than treat it as an ex-post compliance obligation. Investments in data governance reduce coordination costs, support cross-functional collaboration, and enable the scalable reuse of validated datasets. By providing a shared language across legal, technical, and business functions, governance reduces rework and regulatory risk.

4.1. Data governance as a risk control system

The findings indicate that, in open banking and open finance contexts, data governance operates as an organisational risk-control system that is embedded within everyday processes rather than applied through isolated or ex post compliance checks. Risk emerges from the continuous circulation of data across organisational boundaries, regulatory regimes and technological systems, making it a systemic condition rather than a discrete event. Data governance addresses this condition by shaping how data can be accessed, interpreted and used throughout the organisation.

Empirically, governance performs risk control by establishing a restrictive default logic that frames data access as an exception requiring justification, approval and documentation. Through role-based permissions, approval workflows and traceability mechanisms, governance structures daily operational behaviour and limits unnecessary exposure to sensitive information. In doing so, governance transforms uncertainty into a form of exposure that can be anticipated and managed rather than merely reacted to. This interpretation is consistent with evidence from highly regulated and AI-intensive environments, where organisations increasingly rely on strengthened control environments and data governance practices to sustain decision-making under conditions of elevated regulatory and technological uncertainty (KPMG, 2025).

A second mechanism concerns the integration of compliance into the design of operational processes. Regulatory requirements are not addressed after data has been processed or products have been deployed, but are translated into workflow rules, technical specifications and documentation practices that shape organisational routines from the outset. Governance thus functions as an interpretive layer that converts regulatory expectations into concrete operational requirements, ensuring that risk control is enacted consistently across units and over time. Rather than constraining organisational action, governance performs the procedural work needed to render action possible in contexts characterised by uncertainty and high accountability demands.

Data governance also operates as a continuous risk-control mechanism in environments characterised by regulatory overlap and legacy infrastructures. Organisations must constantly reconcile historical datasets, inherited system architectures and evolving regulatory interpretations. Activities such as data reclassification, documentation, validation and migration serve to mitigate structural risks embedded in legacy systems and prevent their propagation into new analytical models or digital services. These findings align with broader evidence on the expansion from open banking to open finance, which multiplies data types, regulatory authorities and coordination requirements, thereby increasing governance complexity across interconnected financial ecosystems (Cambridge Centre for Alternative Finance, 2024).

By clarifying the semantic meaning of datasets, defining the purposes for which data can be mobilised and specifying responsibilities associated with data transformation and reuse, governance stabilises expectations among organisational actors. This stabilisation is particularly critical in environments where regulatory requirements evolve over time and where automated systems may behave in ways that are difficult to fully foresee. When actors operate on consistent assumptions about data meaning, ownership and legitimacy, they are better able to coordinate decisions and manage exposure without paralysing innovation.

Finally, our findings show that the tension between innovation speed and control is managed through governance procedures that deliberately introduce validation steps and approval stages into development processes. While these controls may slow down deployment, they reduce the likelihood of regulatory violations and unintended data exposure. Governance therefore reshapes the procedural conditions under which innovation occurs, allowing organisations to pursue data-driven initiatives within boundaries that preserve compliance, trust and institutional legitimacy.

4.2. Structuring as the foundation of orchestrated risk

Data governance performs a structuring function by shaping how organisational data are stabilised, organised and rendered reusable across processes, systems and teams. In open banking and open finance contexts, data governance does not merely impose constraints but actively configures the internal organisation of data by defining shared meanings, organising data flows and maintaining coherence across heterogeneous and historically layered infrastructures. Through this structuring work, fragmented data are transformed into resources that can support coordination, analytical reliability and scalable technological development.

Consistent with Salerno and Maçada (2025), who conceptualise data governance as a foundational capability within the structuring phase of resource orchestration, our findings confirm the centrality of governance in shaping how organisational data resources are initially stabilised and organised. However, this study extends their conceptualisation by explicating how structuring is enacted in practice and by showing that governance-related mechanisms are not confined to a single configuration phase but continue to influence organisational processes as data practices evolve.

Empirically, structuring is accomplished through standardisation and interoperability practices that reduce interpretative ambiguity and align organisational actors around shared definitions and schemas. Governance artefacts such as data catalogues, mapping frameworks and metadata repositories function as coordination mechanisms that make data understandable and comparable across units and systems. By creating a common organisational language, these practices enable teams to work on consistent datasets and prevent fragmentation that would otherwise undermine reporting, onboarding and cross-system integration.

A second structuring mechanism concerns data quality assurance and lineage tracking. Our findings show that data quality must be actively produced through validation routines, anomaly detection and systematic documentation rather than assumed by default. Lineage practices provide visibility into how data enters the organisation, how it is transformed and which processes depend on it. This visibility allows organisations to identify dependencies, correct inconsistencies and prevent historical errors from propagating into analytical models or operational systems. Structuring thus creates the conditions under which data can be reused across organisational units without constant revalidation.

Structuring also takes a material form through the segmentation of data environments and domains. Governance architectures based on the separation of production, testing, analytical and development environments allow organisations to protect sensitive information while still enabling experimentation within controlled boundaries. By partitioning data according to purpose and risk profile, governance shapes the sequencing and rhythm of organisational data practices and prevents analytical workloads from disrupting operational systems.

These process-level findings are consistent with an operational view documented by Deloitte (2024) which shows how financial institutions implement data governance through shared definitions, ownership models, data lineage and data health monitoring in order to make data risks visible and manageable before data is reused across organisational units. Importantly, our findings indicate that these structuring mechanisms are mutually reinforcing. Shared definitions ground responsibilities in common meanings, clarity over ownership and stewardship makes accountability traceable, and lineage and quality monitoring enhance risk visibility by revealing dependencies and points of exposure.

Finally, our findings highlight that structuring governance must continuously address the challenges posed by legacy infrastructures and organisational change. Mergers, historical design choices and incomplete documentation generate inconsistencies that require ongoing remediation through reclassification, restructuring and manual data cleaning. The coexistence of modern governed pipelines with older systems makes structuring an enduring organisational effort rather than a completed task.

Taken together, these findings show that data governance structures organisational data through a set of interdependent processes that stabilise meaning, clarify responsibility and render data usable over time. By unpacking how structuring is practically accomplished, this study moves beyond conceptual placement and demonstrates how data governance enables the coordinated mobilisation of data resources in complex and regulated financial ecosystems.

4.3. Bundling as an integrative and enabling capability

At this stage, data governance enables bundling by providing the technical and organisational conditions through which structured data resources are integrated into coherent and deployable capabilities. In open banking and open finance contexts, bundling does not occur through the simple aggregation of data assets, but through the deliberate integration of infrastructures, automation routines, privacy constraints and decision rights. Data governance functions as the enabling infrastructure that allows these heterogeneous elements to be combined in a controlled manner, preventing inconsistency, security breaches and loss of control.

Empirically, governance supports bundling by stabilising the technical foundations upon which innovation depends. Interviewees consistently emphasised that cloud integration, automation and the deployment of AI-based systems become feasible only when data definitions, lineage and access rules are clearly established. Governance provides a shared technical language that allows development teams to integrate new tools into existing architectures while preserving consistency across environments. Without this foundation, even minor technological changes require extensive rework and produce unpredictable outcomes. These findings extend prior work by Peretz-Andersson et al. (2024), who positioned AI governance primarily within the bundling phase of resource orchestration, in the context of manufacturing small and medium-sized enterprises. While their study highlights how governance mechanisms support the coordination and combination of AI-related resources, the present research shows that, in data-intensive and highly regulated open finance ecosystems, bundling activities depend more fundamentally on data governance mechanisms that precede and enable AI governance itself. By structuring data assets, stabilising data flows, and defining permissible uses, data governance creates the conditions under which AI governance and other advanced technological capabilities can be effectively bundled into organisational processes.

This enabling role of governance in technical integration is consistent with operational evidence reported by Deloitte (2024). The report documents how financial institutions integrate analytics, automation, risk management and regulatory requirements through formal data governance practices, highlighting governance as a key mechanism for aligning technical systems and organisational functions within complex financial infrastructures. Our findings echo this perspective by showing that contemporary financial architectures consist of distributed databases, cloud services, APIs, orchestration layers and analytical models that interact continuously and often asynchronously. These components are frequently developed by different teams, updated at different times and subject to heterogeneous constraints. By embedding documentation practices, validation rules and monitoring routines directly into these technical processes, data governance allows organisations to maintain coherence as systems evolve and ensures that bundled capabilities remain stable, auditable and trustworthy even as technological complexity increases.

A central bundling mechanism concerns automation and ownership clarity. Our findings show that automated pipelines and monitoring systems function reliably only when governance clearly defines who owns datasets, who approves changes to schemas and transformation rules and who is accountable for system behaviour.

Governance embeds these decision rights directly into technical workflows, enabling automation to scale without introducing silent errors or undocumented changes. In this sense, governance transforms automation from a fragile optimisation into a stable organisational capability.

Bundling is further enabled through the integration of privacy-by-design principles into technical architectures. Rather than operating as an external compliance requirement, privacy constraints are embedded into system design through segmentation of environments, anonymisation routines and role-based access controls. Governance determines which data elements can move across systems, under which conditions and for which purposes. By encoding privacy rules directly into pipelines and infrastructures, governance allows organisations to combine analytical, operational and customer-facing components without exposing sensitive information or violating regulatory constraints.

Cross-functional coordination represents another critical bundling mechanism. Our findings show that governance facilitates integration not only by aligning technical systems, but also by synchronising the contributions of engineering, product, compliance, risk and legal functions. Documentation, data mapping and approval workflows act as shared artefacts that align these actors around common definitions and system behaviours. This embedded coordination reduces the need for continuous renegotiation and enables complex capabilities to be developed and modified in a controlled and repeatable manner.

In line with the conceptual framework developed in this study (Figure 4), bundling can therefore be understood as the process through which data, technology, regulatory requirements and business objectives are actively integrated into coherent organisational capabilities. Standardised documentation, access control models and data lineage requirements operate as integrative mechanisms that allow heterogeneous digital resources to be combined across processes and organisational units. Through these mechanisms, data governance enables coordination across legal, technological and business functions by establishing shared reference points and codified decision rights. Standardised documentation supports a common understanding of data meaning and usage, access controls translate regulatory and risk constraints into technical rules, and lineage requirements make dependencies and responsibilities visible across systems. Together, these mechanisms align organisational routines and decision rights and allow data-driven capabilities to be deliberately configured rather than emergent.

At the same time, bundling through governance introduces coordination costs and temporal frictions. Validation procedures, restricted access to production environments and certification requirements can slow development and delay deployment. However, our findings indicate that these constraints are not merely obstructive. By forcing inconsistencies to be resolved before integration, governance prevents technical debt and reduces the likelihood of systemic failure. Bundling therefore reorganises risk rather than eliminating it, embedding control and accountability directly into organisational capabilities.

Taken together, these findings show that data governance enables bundling by integrating technical infrastructures, automation routines, privacy constraints and organisational roles into coherent configurations that can be deployed in practice. Bundling moves organisations beyond isolated experiments toward capabilities that are stable, scalable and institutionally robust. In this sense, data governance actively shapes how capabilities are assembled and prepares the ground for the leveraging phase, in which these bundled capabilities can be mobilised for innovation while preserving control, trust and legitimacy.

4.4. Leveraging as a driver of innovation, scalability and responsible use

Leveraging refers to the phase in which data-driven capabilities are deployed operationally, extended across contexts and translated into value-creating outcomes. It is at this stage that the outcomes of structuring and bundling become visible in day-to-day processes, customer-facing services and automated decision systems. In open banking and open finance, leveraging is particularly critical because organisational action becomes consequential beyond internal boundaries. Decisions made during this phase affect customers, partners and regulators, making leveraging the point at which innovation, risk and accountability intersect most directly.

Data governance shapes how organisations leverage capabilities in practice by defining the conditions under which data-driven solutions can be operationalised. It does not operate as a background control activated only when issues arise, nor as a compliance layer added after deployment. Instead, it frames permissibility, responsibility and legitimacy before and during execution. Through governance mechanisms, organisations determine which data can be used, for which purposes, under which safeguards and through which operational procedures. In regulated, data-intensive environments, leveraging cannot be guided solely by technical feasibility or strategic ambition; it must remain aligned with regulatory requirements, ethical standards and institutional expectations that extend beyond the firm.

A central mechanism through which data governance enables leveraging is translation. Governance converts abstract regulatory and ethical principles into operational constraints and design choices that shape workflows, system architectures and decision processes. Requirements related to privacy, consent, security and consumer protection are embedded directly into how capabilities are deployed, monitored and audited, preventing innovation from drifting away from the normative environment in which it must remain legitimate. As organisations operationalise analytics, automation and AI, governance becomes even more consequential because these technologies introduce opacity, complexity and new forms of systemic risk. Governance enables responsible deployment by defining how automated decisions are generated, monitored and justified, and by making transparency, traceability and oversight executable in practice rather than aspirational values.

In open finance ecosystems, data governance also enables leveraging by operationally sustaining trust as data flows expand across organisational boundaries. PwC (2024) describes how trust is supported through mechanisms such as consent management, accountability frameworks and controls over data reuse. These governance arrangements determine how data can circulate among actors while remaining permissioned, auditable and socially acceptable. By stabilising consent interpretation, constraining reuse and making accountability enforceable, governance allows organisations to deploy data-driven capabilities without eroding the trust required for continued ecosystem participation.

Scalability constitutes another defining challenge of the leveraging phase. Leveraging requires not only deploying a capability once, but extending it across products, customer segments and jurisdictions characterised by heterogeneous data sources and regulatory interpretations. Accenture (2020) shows that many open banking initiatives fail to progress beyond pilots because data management, consent handling and governance practices are not sufficiently embedded into operational processes and ecosystem coordination models. Data governance enables scalability by making deployments repeatable across contexts: stable definitions, documented procedures, auditable controls and traceable data flows allow organisations to extend capabilities without introducing hidden vulnerabilities, inconsistent outcomes or trust breakdowns. In this sense, scalability is not primarily a technological achievement, but a governance outcome produced through standardisation, accountability and controlled reuse.

Leveraging also reveals the adaptive nature of governance. Each deployment generates feedback about system behaviour, regulatory interpretation and organisational readiness. Edge cases, unexpected outcomes and operational frictions expose limitations in existing arrangements and trigger refinement of rules, controls and responsibilities. This learning feeds back into governance practices, reshaping how data resources are structured and how capabilities are reconfigured for subsequent deployments. Leveraging therefore both concludes earlier orchestration efforts by translating capabilities into operational action and renews them by generating insights that inform future structuring and bundling. Data governance remains central throughout this cycle by ensuring that value creation is realised under conditions of accountability, legitimacy and trust in open finance ecosystems. Rather than representing a final step, leveraging opens a new cycle in which governance and innovation co evolve. Leveraging thus demonstrates how data governance functions as a driver of innovation, scalability and responsible use by closing the loop between risk control and the conditions under which value can be legitimately realised in open finance ecosystems.

4.5. The enabling effect and the evolving effect

The dynamic relationship between data governance and organisational action can be understood through two interrelated effects: the enabling effect and the evolving effect. These effects capture how governance and innovation are not sequential or independent phenomena, but mutually constitutive processes that continuously shape one another. Rather than framing governance as a static framework applied to innovation, these effects illustrate how governance actively conditions innovation while simultaneously being transformed by it.

The enabling effect refers to the capacity of data governance to create the stability, clarity and procedural coherence required for innovation to emerge in complex and regulated environments. By reducing interpretive ambiguity and establishing shared expectations, governance provides organisations with a controlled space in which experimentation becomes possible. It clarifies what constitutes legitimate data use, how responsibilities are allocated and which safeguards must accompany new initiatives. This clarity supports confidence in decision making, allowing organisational actors to engage in exploration without fear of uncontrolled exposure or retrospective sanction. Importantly, the enabling effect does not eliminate uncertainty. Instead, it renders uncertainty intelligible and manageable. Through shared definitions, documented processes and visible controls, data governance transforms uncertainty into a bounded condition within which organisations can act deliberately. Innovation is therefore enabled not because risk disappears, but because it is recognised, structured and embedded into organisational routines. Without this effect, innovation would either be reckless, proceeding without sufficient awareness of consequences, or impossible, paralysed by fear of regulatory, ethical or reputational failure.

The evolving effect captures the complementary dynamic through which innovation reshapes governance itself. As organisations deploy new capabilities, introduce automated decision processes and scale data driven services, they encounter situations that could not be fully anticipated during design. Edge cases, unexpected interactions and emergent risks expose limitations in existing governance arrangements. These experiences generate learning that feeds back into governance practices, prompting refinement of definitions, adjustment of controls and reconfiguration of responsibilities. Over time, this feedback loop allows governance to become more precise, more anticipatory and more closely aligned with organisational practice. Governance evolves not through abstract redesign but through sustained engagement with real use cases and operational realities.

Taken together, the enabling effect and the evolving effect demonstrate that data governance functions as a dynamic learning system rather than as a fixed set of rules. Governance enables innovation by creating the conditions for controlled exploration, while innovation refines governance by revealing where existing arrangements are insufficient or misaligned. Importantly, this dynamic applies across all three phases of resource orchestration. Structuring, bundling and leveraging do not unfold in a linear sequence, but continuously inform and reshape one another through ongoing governance and learning processes. This iterative cycle explains how organisations can progressively expand their innovative capacity while maintaining coherence, accountability and legitimacy.

By articulating these two effects, the study highlights a fundamental shift in how data governance should be conceptualised. Governance is not positioned outside the innovation process, nor does it intervene only at moments of failure. Instead, it is embedded within the continuous interaction between risk control and the conditions under which value can be legitimately realised. As organisations pursue new technological and strategic ambitions, governance evolves alongside them, shaping and being shaped by the trajectory of innovation. This dynamic interplay lays the conceptual groundwork for the radical status recategorization developed in the following section, in which data governance is repositioned as a central organisational capability that mediates the relationship between risk and innovation in data intensive ecosystems.

4.6. Performance implications of data governance orchestration

When data governance is effectively integrated across the structuring, bundling and leveraging phases, and sustained through the enabling and evolving effects, it can lead to a set of cumulative performance outcomes that extend beyond short-term compliance or efficiency gains. Rather than generating value through isolated controls, data governance can contribute to sustainable performance by shaping how firms operate, decide and scale data-driven activities over time.

Data governance contributes to organisational value creation through two closely connected mechanisms that operate in parallel. From an efficiency perspective, governance enhances cost control by improving data quality, limiting duplication across systems, and reducing the need for corrective rework caused by inconsistent definitions or undocumented changes. By stabilising data flows and clarifying responsibilities, governance streamlines operational activities and increases process reliability, allowing organisations to manage growing data volumes and regulatory complexity with lower marginal costs.

At the same time, governance enables revenue-oriented outcomes by establishing the conditions under which data can be safely reused and recombined across products, services, and organisational units. When data assets are trusted, well documented, and scalable, teams can more easily design, validate, and scale new data-driven use cases, including advanced analytical applications and AI-based solutions. In this way, governance does not merely constrain innovation but actively supports it by transforming data into a dependable input for experimentation and commercial development.

These value creation mechanisms translate into performance improvements across multiple organisational dimensions. Clearly governed data flows, documented responsibilities, and traceable transformations reduce operational friction, limit rework, and help prevent errors that typically arise from inconsistent interpretations or undocumented changes. At the same time, validated datasets, visible lineage, and clearly defined usage conditions improve the quality and consistency of data-driven decision making, supporting more informed managerial choices and reducing the risk of conflicting insights derived from the same underlying data.

In addition, data governance embeds traceability and auditability directly into organisational operations. Rather than treating compliance as an ex post reporting exercise, governance integrates accountability and transparency into workflows and technical systems. This strengthens regulatory alignment and supports long-term compliance without requiring continuous reactive controls, thereby increasing organisational robustness as data-driven activities scale.

Finally, data governance contributes to accountability, transparency, and trust in areas such as ESG reporting and responsible analytics. By ensuring that sustainability-related data is accurate, auditable, and consistently interpreted, governance supports credible disclosures and reduces the risk of symbolic or unreliable performance claims. In this sense, sustainable performance emerges not only in economic terms, but also in firms' ability to demonstrate legitimacy, responsibility, and long-term alignment with institutional expectations.

Taken together, these outcomes indicate that data governance orchestration can transform risk control into durable firm performance. Sustainable performance is more likely to emerge not from minimising risk or maximising speed in isolation, but from maintaining reliable processes, consistent decisions, and trustworthy data use as organisations evolve within data-intensive and highly regulated ecosystems.

4.7. Theoretical contribution and the radical status recategorization

The findings of this study contribute to academic theory by proposing a fundamental reconceptualization of data governance. Rather than treating data governance as a compliance-oriented mechanism or as a set of administrative controls designed primarily to prevent misuse, this research recategorizes data governance as a central system of risk control that actively shapes organisational action. This shift is theoretically significant because it moves governance from the periphery of organisational analysis to its centre, redefining how scholars understand the relationship between data, risk and innovation in data intensive environments.

A first contribution lies in demonstrating that data governance is integral to the core phases through which organisational resources are mobilised. The analysis shows that governance is not external to structuring, bundling and leveraging, nor does it merely impose constraints on these activities. Instead, governance operates through all three phases by shaping how resources are recognised, stabilised and rendered meaningful during structuring, how they are integrated and aligned during bundling, and how they are deployed, scaled and legitimised during leveraging. By embedding governance within the orchestration process itself, the study challenges perspectives that position governance as an ex-post control applied after strategic or technological decisions have already been made.

A second contribution concerns the dynamic and learning oriented nature of data governance. Through the articulation of the enabling effect and the evolving effect, the study shows that governance is not static but continuously produced and refined through organisational practice. Governance enables innovation by creating clarity, coherence and bounded conditions for exploration, while innovation simultaneously reshapes governance by revealing new risks, ambiguities and coordination challenges. This reciprocal relationship positions data governance as an adaptive system that evolves alongside technological development, regulatory change and organisational learning, extending existing theories that emphasise static governance frameworks.

A third contribution reframes data governance as a strategic mechanism that mediates the relationship between risk and innovation. Traditional perspectives often depict governance as either protective, focused on limiting exposure and ensuring compliance, or permissive, designed to facilitate innovation by reducing restrictions. The findings of this study suggest a different interpretation. Data governance functions as a risk controlling and directional system that allows organisations to pursue innovation while maintaining ethical, operational and regulatory integrity. It neither suppresses experimentation nor allows it to proceed unchecked. Instead, it shapes the trajectory of innovation by defining which opportunities are legitimate, how they can be pursued and under which conditions they can be scaled responsibly.

Together, these insights constitute the radical status recategorization proposed by this study. Data governance should no longer be understood as an administrative burden or a supporting function subordinate to strategy and technology. It should instead be recognised as a central theoretical category that explains variation in organisational outcomes within data intensive ecosystems. Organisations that succeed in responsible and scalable innovation are not simply those with superior technology or greater risk appetite, but those that are able to orchestrate data governance as a dynamic system of risk control that aligns resources, governs exposure and supports continuous learning.

The final conceptual framework (Figure 4**Figure 4Figure 4Figure 4**) synthesises these insights by positioning data governance as a central organizational system of risk control that continuously orchestrates structuring, bundling and leveraging processes in open finance ecosystems.

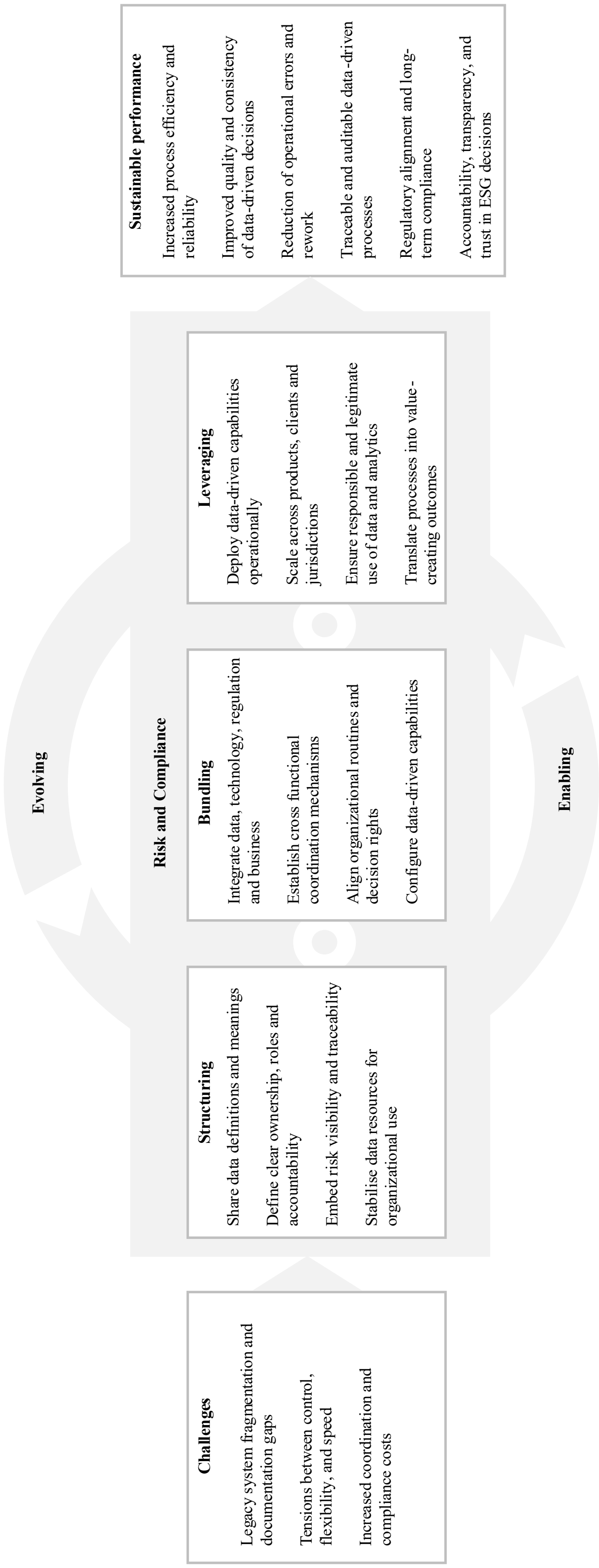


Figure 4. Data governance orchestration

4.8. Managerial implications

For practitioners operating in open banking and open finance, the findings of this study highlight the need to fundamentally reconsider the role of data governance within the organisation. Rather than approaching governance as a constraint imposed by regulation or as a technical requirement to be satisfied ex post, managers are encouraged to treat data governance as a strategic architecture that actively enables coordinated action, responsible innovation and the conditions for long-term value creation. Viewing governance through this lens requires a shift in managerial attention, investment priorities and organisational design.

One key implication concerns the importance of clarity. Establishing shared definitions of data, consistent documentation practices and clearly articulated accountability structures reduces ambiguity across the organisation and creates the conditions for effective collaboration. When teams operate with a common understanding of what data represents, how it can be used and who is responsible for it, they are better equipped to innovate without exposing the organisation to unnecessary risk. In complex and regulated environments, such clarity does not slow down decision making but instead accelerates it by reducing friction, misalignment and the need for repeated negotiation.

The findings also underscore the central role of data governance in supporting cross functional coordination. Data driven innovation requires sustained interaction between legal, technical, risk and business functions, each of which brings distinct expertise and priorities. Without a unifying governance framework, these interactions often become fragmented or conflictual, leading to delays, inconsistent outcomes or overly conservative decisions. Strong data governance provides the shared reference point that allows these actors to collaborate productively, aligning technological development with regulatory interpretation and strategic objectives. Managers who invest in governance as a coordination mechanism can therefore improve both the quality and the speed of innovation.

Scalability emerges as another critical managerial concern shaped by data governance. Many organisations are able to develop promising pilot initiatives, yet struggle to extend them across products, customer segments or jurisdictions. The findings suggest that this challenge is rarely caused by technological limitations alone. Instead, it often reflects weaknesses in governance structures that undermine consistency, transparency and trust as initiatives grow. Firms with robust data governance are better positioned to replicate innovations reliably, because governance ensures stable definitions, traceable processes and auditable controls across contexts. In this sense, scalability should be understood not as a purely technical achievement, but as the outcome of sustained governance capability.

The study further highlights the role of data governance in embedding ethical considerations into everyday organisational practice. As automation and artificial intelligence become more prevalent, managers face increasing scrutiny regarding fairness, accountability and transparency. Data governance offers mechanisms through which these concerns can be integrated into system design and deployment rather than addressed reactively. By defining expectations around oversight, explainability and monitoring, governance helps organisations anticipate ethical risks and maintain stakeholder trust in environments where reputational damage can occur rapidly and be difficult to reverse.

Finally, the findings encourage managers to treat data governance as a dynamic and evolving system of risk control rather than a fixed framework. Each new use case generates insights into organisational readiness, regulatory interpretation and system behaviour. When governance structures are designed to absorb and act on this feedback, they support continuous learning and adaptation. Over time, this creates a virtuous cycle in which governance becomes more precise and more supportive of innovation, while innovation benefits from increasingly mature and anticipatory governance practices. Managers who embrace this dynamic view are better equipped to navigate technological change and regulatory uncertainty without sacrificing coherence or control.

4.9. Cross-jurisdictional differences in data governance models

Data governance practices vary significantly across jurisdictions as a result of different institutional arrangements, sources of authority and enforcement mechanisms. In regulatory based models, data governance is primarily shaped by binding legal frameworks that mandate data sharing, define compliance requirements and establish supervisory oversight. This approach is exemplified by jurisdictions such as the

European Union and the United Kingdom, where data sharing obligations are grounded in formal regulation, as well as by countries including Australia, Brazil and India, which have introduced mandatory open banking or open finance frameworks. In these contexts, governance structures tend to be formalised, compliance oriented and closely aligned with regulatory expectations, with organisations investing heavily in controls, reporting mechanisms and risk management processes to ensure adherence to legal obligations.

By contrast, market-based models rely predominantly on competitive dynamics, contractual arrangements and reputational incentives. Jurisdictions such as the United States and Switzerland illustrate this approach, where data sharing practices have largely emerged through bilateral or multilateral agreements between market actors rather than through comprehensive mandatory regulation. In these settings, governance practices are more strongly oriented toward commercial value creation, operational efficiency and trust building, with formal compliance playing a less central role.

Voluntary-based models occupy an intermediate position, where governance frameworks are promoted through industry standards, guidelines and collaborative initiatives rather than imposed by law. Countries such as Singapore, Japan and South Korea provide examples of this approach, in which public authorities often support coordination efforts while leaving participation largely discretionary. In these contexts, governance effectiveness depends on shared norms, coordination mechanisms and collective commitment among ecosystem participants.

Importantly, these models are not mutually exclusive and often coexist in practice, giving rise to hybrid configurations, particularly in cross jurisdictional and multi market environments. Organisations operating across multiple regulatory contexts are frequently required to combine mandatory compliance driven governance with market oriented and voluntary mechanisms, resulting in governance arrangements that are adaptive, layered and context specific. Such hybrid models reflect the need to reconcile regulatory obligations with commercial objectives and innovation imperatives.

Appendix D summarises the regulatory jurisdictions considered and distinguishes between regulatory-based, market-based, voluntary-based and hybrid configurations.

4.10. Limitations and future research directions

This study presents several limitations that should be acknowledged and that simultaneously point toward relevant directions for future research. First, the empirical analysis was based on an extensive outreach effort, during which approximately 280 individuals were contacted across different organisations and roles, resulting in 30 in depth interviews with knowledgeable informants. While this sample size allowed for the generation of rich qualitative insights and the identification of recurring patterns, the limited number of responses inevitably constrains the breadth of perspectives represented. Future studies could therefore extend the empirical sample in order to strengthen the interpretive power of the findings and capture a wider range of organisational experiences related to data governance and process engineering.

Second, the study intentionally sought to capture heterogeneous contexts across different jurisdictions, characterised by varying regulatory frameworks and different levels of data governance maturity. This diversity was central to the research design, as it enabled a broader understanding of how governance practices operate under distinct institutional conditions. At the same time, such heterogeneity suggests that future research could further disentangle jurisdiction specific dynamics by focusing on more homogeneous regulatory settings, thereby allowing a more detailed examination of how governance configurations differ across contexts.

In this regard, future studies could explicitly analyse alternative governance models, such as market based and voluntary based approaches, examining each configuration separately to gain a clearer understanding of the internal dynamics, incentives and constraints that characterise different governance regimes.

Finally, this research treated digital technologies as a collective enabling layer supporting process engineering initiatives, approaching technological projects as a unified set rather than differentiating among specific technologies. Future research could adopt a more fine-grained technological focus by investigating how individual technologies, such as artificial intelligence or blockchain, influence processes in distinct ways. This line of inquiry appears particularly relevant given that an increasing number of process engineering projects are being developed around blockchain based infrastructures, which may introduce specific governance mechanisms and organisational implications.

Taken together, these limitations do not detract from the contribution of the study but rather delineate the scope of the analysis and clarify the conditions under which the findings should be interpreted. They also highlight relevant directions for future research aimed at further deepening the understanding of data governance and process engineering in open banking and open finance contexts.

5 Conclusion

This study examined how data governance shapes the creation, development and scaling of process engineering use cases in open banking and open finance contexts. By moving beyond perspectives that frame data governance as a peripheral compliance function or a purely technical requirement, the analysis highlights its central organisational role in environments characterised by pervasive data flows, high regulatory scrutiny and increasing reliance on data-driven innovation.

The study shows that data governance functions as an integrated system of risk control that structures, coordinates and guides firm behaviour across the phases of resource orchestration. Rather than operating outside innovation processes or intervening only ex post, data governance actively shapes how data resources are stabilised, combined and mobilised in practice. Through its involvement in structuring, bundling and leveraging activities, governance enables firms to transform fragmented and historically layered data environments into coherent foundations for scalable and legitimate innovation.

By articulating the enabling effect and the evolving effect, the study demonstrates that the relationship between data governance and innovation is inherently dynamic. Governance creates the conditions for controlled experimentation by clarifying boundaries, responsibilities and acceptable risk exposure, while the deployment and scaling of data-driven use cases continuously reshape governance through learning, feedback and exposure to real-world operational conditions. This interaction explains how firms can expand their innovative capacity over time without resorting to either excessive restriction or unmanaged risk-taking.

Importantly, the findings show that data governance orchestration translates regulatory constraints and legacy challenges into sustainable organisational outcomes. When embedded across the orchestration process, governance supports not only innovation but also scalability, trust, accountability and long-term performance. In this sense, data governance orchestration is not optional. It represents a critical capability for firms seeking to compete in open finance ecosystems, where value creation depends on the responsible reuse of data across organisational and institutional boundaries.

The final conceptual framework synthesises these insights by positioning data governance as a system of risk control that continuously orchestrates structuring, bundling and leveraging processes. By doing so, the study contributes a new lens for understanding the interrelationship between governance, innovation and competition in complex and dynamic open data environments. More broadly, it invites scholars and practitioners to reconsider data governance not as an administrative burden, but as a foundational organisational capability that shapes strategic action and enables sustainable innovation and competitive performance in data-intensive industries.

References

Accenture. (2020). *Catching the Open Banking Wave*. https://www.accenture.com/content/dam/accenture/final/industry/banking/document/Accenture-Catching-Open-Banking-Wave.pdf

Achleitner, A.-K., Braun, R. ,. &. Puche, B. (2010). *Value Creation Drivers in Private Equity Buyouts: Empirical Evidence from Europe*. *13.0*(2.0). https://doi.org/10.3905/JPE.2010.13.2.017

Agarwal S., Bhardwaj G., Saraswat E., Singh N., Aggarwal R., & Bansal A. (2022). *Insurtech Fostering Automated Insurance Process using Deep Learning Approach. Scopus*. https://doi.org/10.1109/ICIPTM54933.2022.9753891

Aguilera, R. V., Filatotchev, I., Gospel, H., & Jackson, G. (2008). *An organizational approach to comparative corporate governance. Organization Science*, *19*(3), 475–492. https://doi.org/10.1287/orsc.1070.0322

Anand, A., Sharma, R., & & Kohli, R. (2013). *Routines, reconfiguration and the contribution of business analytics to organisational performance*. https://ro.uow.edu.au/eispapers/2545

Andrea, G., &. Santoso, S. (2020). *Improving Economy of the Community Based on Sustainable Tourism and Creative Economy through Business Process Re-Engineering (BPR) With Geopark Development in Lebak Regency Banten Province*. *5.0*(1.0). https://ijisrt.com/assets/upload/submitted_files/1578462846.pdf

Andronie M., Blazek R., Iatagan M., Skypalova R., Uță C., Dijmărescu A., Kovacova M., Grecu G., Pârvu I., Strakova J., Guni C., Zabojnik S., Chiru C., Novák Sedláčková A., Novák A., & Dijmărescu I. (2024). *Generative artificial intelligence algorithms in Internet of Things blockchain-based fintech management. Scopus*, *15.0*(4). https://doi.org/10.24136/oc.3283

Andronie M., Iatagan M., Uță C., Hurloiu I., Dijmărescu A., & Dijmărescu I. (2023). *Big data management algorithms in artificial Internet of Things-based fintech. Scopus, 14.0*(3). https://doi.org/10.24136/oc.2023.023

António Porfírio J., Augusto Felício J., & Carrilho T. (2024). F*actors affecting digital transformation in banking. Scopus, 171.0*. https://doi.org/10.1016/j.jbusres.2023.114393

Arsal R.E., Durdu H.T., & Tongarlak M.H. (2022). *Organizational Enablers of Data-Driven Digital Transformation: A Case Study from Banking Industry. Scopus*. https://doi.org/10.1109/TEMSCONEUROPE54743.2022.9802044

Bain & Company (2024). *How Banks Can Parlay Technology into a Competitive Edge*. https://www.bain.com/insights/how-banks-can-parlay-technology-into-a-competitive-edge/

Barney, J. (1991). *Firm resources and sustained competitive advantage*. *17.0*(1.0). https://doi.org/10.1177/014920639101700108

Basel Commitee. (2003). *Sound Practices for the Management and Supervision of Operational Risk*. https://www.bis.org/publ/bcbs91.pdf?utm

BCG Platinion. (2023). *Modern Banking APIs: From "Nice to Have" to a Critical Success Factor*. https://www.bcgplatinion.com/insights/modern-banking-apis---from-nice-to-have-to-a-critical-success-factor

Boris Otto. (2011). *A morphology of the organisation of data governance*. (PDF) A morphology of the organisation of data governance

Braun, V., & Clarke, V. (2006). *Using thematic analysis in psychology. Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cambridge Centre for Alternative Finance. (2024). *The Global State of Open Banking and Open Finance*. https://www.jbs.cam.ac.uk/wp-content/uploads/2024/11/2024-ccaf-the-global-state-of-open-banking-and-open-finance.pdf

Claudio Bonvino & Marco Giorgino. (2024). *A valorization framework to strategically manage data for creating competitive value*. *269.0*. https://doi.org/10.1016/j.ijpe.2024.109152

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). *Artificial intelligence in information systems research: A systematic literature review and research agenda. International Journal of Information Management*, *60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Couckuyt, D., & Van Looy, A. (2019). *A systematic review of Green Business Process Management*. *26.0*(2.0). https://doi.org/10.1108/BPMJ-03-2019-0106

Dagher J. & Fayad L. (2024). *Business Process Reengineering: A Crucial Approach for Enhanced Organizational Sustainability*. https://doi.org/10.1007/978-981-99-8572-2_2

Davenport, T. H. (1993). *Process Innovation: Reengineering Work Through Information Technology*. https://books.google.com/books/about/Process_Innovation.html?id=kLlIOMGaKnsC

Deloitte. (2024). *Data Management Trends in the Financial Services Sector*. https://www2.deloitte.com

Dominik A. Fischer, Laura Marcus, Maximilian Ro¨glinger. (2024). *A Portfolio Management Method for Process Mining-Enabled Business Process Improvement Projects*. *66.0*(5.0). https://doi.org/10.1007/s12599-024-00906-2

Douglas W. Arner, Giuliano G. Castellano, & Ēriks K. Selga. (2022). *Financial Data Governance: The Datafication of Finance, the Rise of Open Banking and the End of the Data Centralization Paradigm*. https://doi.org/10.2139/ssrn.4040604

Dumas M., La Rosa M., Mendling J., & Reijers H.A. (2018). *Fundamentals of Business Process Management*. https://doi.org/10.1007/978-3-662-56509-4

Edmondson, A. C., & McManus, S. E. (2007). *Methodological fit in management field research. Academy of Management Review*, *32*(4), 1246–1264. https://doi.org/10.5465/amr.2007.26586086

Eisenhardt, K. M. (1989). *Building theories from case study research. Academy of Management Review*, *14*(4), 532–550. https://doi.org/10.5465/amr.1989.4308385

Eling M. & Lehmann M. (2018). *The Impact of Digitalization on the Insurance Value Chain and the Insurability of Risks. Scopus*, *43.0*(3). https://doi.org/10.1057/s41288-017-0073-0

Evans N. & Fernando A. (2025). *The Enterprise 4.0 Business Analyst: Digital Business Transformation Requires a New Mindset. Scopus*. https://doi.org/10.1002/kpm.1799

Data Insight Market. (2025). *Exploring Barriers in Business Process Reengineering Market: Trends and Analysis 2025-2033*. https://www.datainsightsmarket.com/reports/business-process-reengineering-1947088#

FDIC. (2006). *Operational Risk Management: An Evolving Discipline*. https://www.fdic.gov/regulations/examinations/supervisory/insights/sisum06/sisummer06-article1.pdf

Feagin, J. R., Orum, A. M., & Sjoberg, G. (1991). *A case for the case study*. University of North Carolina Press.

Genty K. I., A. S. O., Villet H. J. (2025). *Business Process Re-Engineering and Commercial Banks' Performance: Empirical Evidence from Nigeria*. *6.0*(1.0). https://doi.org/10.47616/jamrems.v6i1.568

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2010). *Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2021). *A systematic methodology for doing qualitative research*. Sage Publications.

Golden-Biddle, K., & Locke, K. (1993). *Appealing works: An investigation of how ethnographic texts convince. Organization Science*, *4*(4), 595–616. https://doi.org/10.1287/orsc.4.4.595

Goldman Sachs. (2025). *Transaction Banking—Foundations for a New Horizon*. https://www.goldmansachs.com/what-we-do/transaction-banking/treasury-transformation-new-horizons/

Gotthardt M., Koivulaakso D., Paksoy O., Saramo C., Martikainen M., & Lehner O. (2020). *Current state and challenges in the implementation of smart robotic process automation in accounting and auditing. Scopus*, *9.0*(1). https://doi.org/10.35944/JOFRP.2020.9.1.007

Hadi, A. H., Abdulhameed, G. R. ,. Malik, Y. S. ,. &. Flayyih, H. H. (2023). *The Influence of Information Technology (IT) on Firm Profitability and Stock Returns*. *124.0*(13.0). https://www.researchgate.net/publication/373485742_THE_INFLUENCE_OF_INFORMATION_TECHNOLOGY_IT_ON_FIRM_PROFITABILITY_AND_STOCK_RETURNS

Hammer M. & Champy J. (1993). *Reengineering the corporation: A manifesto for business revolution*. *36.0*(5.0). https://doi.org/10.1016/S0007-6813(05)80064-3

Harry Sneed, C. V. (2001). *Reengineering the Corporation—A Manifesto for IT Evolution*. https://www.researchgate.net/publication/2366189_Reengineering_the_Corporation_-_A_Manifesto_for_IT_Evolution

Hoesch-Klohe, K., Ghose, A. &. Le, L. (2010). *Towards green business process management*. https://core.ac.uk/download/pdf/37006552.pdf

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). *Towards a theory of ecosystems. Strategic Management Journal*, *39*(8), 2255–2276. https://doi.org/10.1002/smj.2904

Kareem D.A., Obaid H.A.H., & Taha H.Y. (2024). *The Role of Financial Process Re-Engineering in Alleviating the Pressures in Commercial Banks of Al-Muthanna*. *5.0*(3.0). https://doi.org/10.48165/sajssh.2024.5304

Kokina J. & Blanchette S. (2019). *Early evidence of digital labor in accounting: Innovation with Robotic Process Automation. Scopus, 35.0*. https://doi.org/10.1016/j.accinf.2019.100431

KPMG. (2025). *KPMG Global Tech Report: Financial Services Insights*. https://assets.kpmg.com/content/dam/kpmg/ae/pdf-2025/03/global-tech-report-financial-services-insights.pdf

Kristin Weber, Boris Otto, & Hubert Österle. (2009). *One Size Does Not Fit All—A Contingency Approach to Data Governance*. *1.0*(1.0). https://doi.org/10.1145/1515693.1515696

Kristoffersen E., Mikalef P., Blomsma F., & Li J. (2021). *The effects of business analytics capability on circular economy implementation, resource orchestration capability, and firm performance*. *239.0*. https://doi.org/10.1016/j.ijpe.2021.108205

Kubrak K., Milani F., & Nava J. (2023). *Digital Technology-Driven Business Process Redesign: A Classification Framework*. *476.0*. https://doi.org/10.1007/978-3-031-33080-3_13

Liepert M. (2024). *A meta-analysis on private equity technology-driven value creation. Scopus*, *11.0*(1). https://doi.org/10.52566/msu-econ1.2024.29

M Meadows & Y Merali. (2003). *Process Improvement with Vision: A Financial Services Case Study*. *16.0*(3.0). https://doi.org/10.1023/A:1023807821671

Maria Paula Santos & Ana Lucas. (2019). *Identifying Critical Success Factors for Data Quality Management through a Delphi Study*. *13.0*(8.0). https://doi.org/10.5281/zenodo.3455599

Marsudi A.S. & Pambudi R. (2021). *The effect of enterprise resource planning (ERP) on performance with information technology capability as moderating variable*. *24.0*(1.0). https://doi.org/10.14414/jebav.v24i1.2066

Maximize Market Research. (2024). *Business Process Re-engineering Market: Global Industry Analysis and Forecast (2024-2030)*. https://www.maximizemarketresearch.com/market-report/global-business-process-re-engineering-market/95357/

Maxwell, J. A. (1997). *Designing a qualitative study*. Sage Publications.

Mbugua, A. N. (2021). *Effects of Business Process Re-engineering on project implementation of Core Banking Software: A case of Kenyan commercial banks in Nairobi County*. https://su-plus.strathmore.edu/items/d0919aed-9e74-4c88-a072-b157f93869e7

McAfee A. & Brynjolfsson E. (2012). *Big Data: The Management Revolution*. *90.0*(10.0). https://hbr.org/2012/10/big-data-the-management-revolution

McKinsey & Company. (2020). *Designing data governance that delivers value*. https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/designing-data-governance-that-delivers-value

McKinsey & Company. (2024). *How Banks Can Supercharge Technology Speed and Productivity*. https://www.mckinsey.com/industries/financial-services/our-insights/how-banks-can-supercharge-technology-speed-and-productivity

Mikalef P., Pappas I.O., Krogstie J., & Giannakos M. (2018). *Big data analytics capabilities: A systematic literature review and research agenda*. *16.0*. https://doi.org/10.1007/s10257-017-0362-y

Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Sage Publications.

Morgan, G., & Smircich, L. (1980). *The case for qualitative research*. *Academy of Management Review*, *5*(4), 491–500.

Neuman, W. L. (2006). *Social research methods: Qualitative and quantitative approaches* (4th edn). Allyn and Bacon.

Nkomo A. & Marnewick C. (2021). *Improving the success rate of business process re-engineering projects: A business process re-engineering framework*. *23.0*(1.0). https://doi.org/10.4102/sajim.v23i1.1259

Omololu Bajulaiye Mark Fenwick Ivona Skultetyova Erik P.M. Vermeulen. (2020). *Digital Transformation in the Hedge Fund and Private Equity Industry*. https://www.ecgi.global/sites/default/files/working_papers/documents/bajulaiyefenwickskultetyovavermeulenfinal.pdf?utm

Onjure C.O., Wanyoike D.M., & Mung'atu J. (2018). *Influence of Business Process Reengineering on Electronic Commerce Strategy as Adopted by Commercial Banks in Kenya*. *7.0*(1.0). https://doi.org/10.24940/ijird/2018/v7/i1/JAN18003

Oxford Learner's Dictionaries. (2025). *Application programming interface*. https://www.oxfordlearnersdictionaries.com/definition/english/application-programming-interface

Paul L.R. & Sadath L. (2021). *A systematic analysis on fintech and its applications. Scopus*. https://doi.org/10.1109/ICIPTM52218.2021.9388371

Peggy Kaponda, S. K., Charity Meki-Kombe, Austin Mwange. (2023). *Models and methodologies for improving and redesigning business processes: A theoretical review*. *4.0*(5.0). https://doi.org/10.54660/.IJMRGE.2023.4.5.651-660

Peretz-Andersson, E., Tabares, S., Mikalef, P., & Parida, V. (2024). *Artificial intelligence implementation in manufacturing SMEs: A resource orchestration approach. International Journal of Information Management*, *77*, 1–23. https://doi.org/10.1016/j.ijinfomgt.2024.102781

Pramod D. (2022). *Robotic process automation for industry: Adoption status, benefits, challenges and research agenda. Scopus*, *29.0*(5). https://doi.org/10.1108/BIJ-01-2021-0033

Priyanto P., Murwaningsari E., & Augustine Y. (2023). *Exploring the Relationship between Robotic Process Automation, Digital Business Strategy and Competitive Advantage in Banking Industry. Scopus*, *13.0*(3). https://doi.org/10.33168/JSMS.2023.0320

PwC. (2024). *Open Finance: Genesis of a Revolution*. https://www.pwc.com

Rahim R. & Chishti M.A. (2024). *Artificial Intelligence Applications in Accounting and Finance. Scopus*. https://doi.org/10.1109/ICETSIS61505.2024.10459526

Saidat, Z., Abdelrahim, H. J. ,. Ahmad Alkhodary, D. ,. Qasaimeh, G. ,. (2023). *Impact of open big data and insurtech on business digitalization.* https://doi.org/10.1109/ICBATS57792.2023.10111209

Salerno, F. F., & Maçada, A. C. G. (2025). *Developing a Data Orchestration Scale: A Validity and Reliability Study. AMCIS 2025 Proceedings*, *3*. https://aisel.aisnet.org/amcis2025/data_eco/data_eco/3/

Sauer, P. C., & Seuring, S. (2023). *How to conduct systematic literature reviews in management research: A guide in 6 steps and 14 decisions. Review of Managerial Science*, *17*(5), 1899–1933. https://doi.org/10.1007/s11846-023-00668-3

Sirmon D.G., Hitt M.A., Ireland R.D., & Gilbert B.A. (2011). *Resource Orchestration to Create Competitive Advantage: Breadth, Depth, and Life Cycle Effects*. *37.0*(5.0). https://doi.org/10.1177/0149206310385695

Sjödin, D., Parida, V., & Kohtamäki, M. (2023). *Artificial intelligence enabling circular business model innovation in digital servitization: Conceptualizing dynamic capabilities, AI capacities, business models and effects. Technological Forecasting and Social Change*, *197*, 122903. https://doi.org/10.1016/j.techfore.2023.122903

Stefanelli V. & Manta F. (2023). *Digital Financial Services and Open Banking Innovation: Are Banks Becoming 'invisible'? Scopus*. https://doi.org/10.1177/09721509231151491

Subramanian Muthu, L. W. & S. Hossein Cheraghi. (1999). *Business Process Reengineering: A Consolidated Methodology*. https://nzbef.org.nz/wp-content/uploads/2019/05/Paper-Business-Process-Re-engineering-A-Consolidated-Methodology.pdf

Taymouri F., Rosa M.L., Dumas M., & Maggi F.M. (2021). *Business process variant analysis: Survey and classification. Scopus, 211.0*. https://doi.org/10.1016/j.knosys.2020.106557

Tekka, R. S. (2021). *Business Re-engineering Strategic Process Model for Small-Medium Local Contractor's Sustainable Competitiveness: A Case of Tanzania*. *6.0*(4.0). https://doi.org/10.11648/j.ajmse.20210604.15

Thomas H. Davenport & Thomas C. Redman. (2020). *Digital Transformation Comes Down to Talent in 4 Key Areas*. https://hbr.org/2020/05/digital-transformation-comes-down-to-talent-in-4-key-areas

Tinnilä, M. (1995). *Strategic perspective to business process redesign*. *1.0*(1.0). https://doi.org/10.1108/14637159510798202

Tranfield, D., Denyer, D., & Smart, P. (2003). *Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. British Journal of Management*, *14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Tsakalidis G. & Vergidis K. (2024). *Business Process Redesign: A Systematic Review of Evaluation Approaches Prior to Implementation*. *7.0*(1.0). https://doi.org/10.31181/dmame712024889

van der Aalst M. P. & Carmona J. (2022). *Process Mining Handbook*. https://doi.org/10.1007/978-3-031-08848-3

Venkatraman & Henderson. (1993). *Strategic Alignment: A Model for Organizational Transformation via Information Technology*. https://doi.org/10.1093/oso/9780195068061.003.0009

Vijay Khatri & Carol V. Brown. (2010). *Designing Data Governance*. *53.0*(1.0). https://doi.org/10.1145/1629175.1629210

Volberda H.W., Khanagha S., Baden-Fuller C., Mihalache O.R., & Birkinshaw J. (2021). *Strategizing in a digital world: Overcoming cognitive barriers, reconfiguring routines and introducing new organizational forms*. *54.0*. https://doi.org/10.1016/j.lrp.2021.102110

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), XIII–XXIII.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd edn). Sage Publications.

Zelt, S., Recker, J., Schmiedel, T., & vom Brocke, J. (2019). *A Theory of Contingent Business Process Management*. *25.0*(6.0). https://doi.org/10.1108/BPMJ-05-2018-0129

Zeng J. & Khan Z. (2019). *Value creation through big data in emerging economies: The role of resource orchestration and entrepreneurial orientation*. *57.0*(8.0). https://doi.org/10.1108/MD-05-2018-0572

Appendices

Appendix A. Scopus search query

((TITLE-ABS-KEY (*process* OR analytics) AND TITLE-ABS-KEY (*engineering OR optimization OR transformation))

AND TITLE-ABS-KEY (financ* OR insur* OR bank* OR "asset management*" OR "wealth management*" OR "investment management*" OR "portfolio management*" OR "venture capital*" OR "private equity*" OR "fintech*" OR "hedge fund*")

AND TITLE-ABS-KEY (digital* OR AI OR "Artificial Intelligence" OR "Machine learning" OR "Automat*" OR "robot*" OR platform))

AND PUBYEAR > 2009 AND PUBYEAR < 2026 AND (LIMIT-TO (SRCTYPE,"j") OR LIMIT-TO (SRCTYPE,"p")) AND (LIMIT-TO (SUBJAREA,"BUSI") OR LIMIT-TO (SUBJAREA,"ECON")) AND (LIMIT-TO (LANGUAGE,"English"))

Appendix B. Literature review extraction table

Authors	Year	Database	Process Engineering	Finance	Digital	Data	Theories	Geographies	Sustainability	Gender	Performance	Innovation	Competition
Achleitner et al.	2010	Grey Literature		x									
Agarwal et al.	2022	Scopus		x							x		
Al-Ababneh et al.	2023	Scopus			x	x							
Alhassan et al.	2016	Research Rabbit				x							
Amit et al.	2013	Research Rabbit	x										
Anand et al.	2013	Grey Literature				x							
Andrea and Santoso	2020	Chat GPT	x										
Andronie et al.	2023	Scopus		x	x				x		x		
Andronie et al.	2024	Scopus			x								
Arner et al.	2022	Scopus		x									
Arsal et al.	2022	Scopus		x	x	x					x		
Bain	2024	Grey Literature		x		x					x		
Bajulaiye t al.	2020	Chat GPT	x										
Barney	1991	Grey Literature					x						
Basel Committee	2003	Grey Literature		x									
BCG	2023	Grey Literature			x								
Bharadwaj et al.	2013	Research Rabbit		x		x							
Bhaskar et al.	2025	Scopus		x		x							
Bonvino and Giorgino	2024	Grey Literature				x							
Braz dos Santos Moderno et al.	2023	Research Rabbit			x								
Brous et al.	2016	Research Rabbit				x							
Chanias et al.	2019	Scopus			x								x
Cheong and Chang	2007	Research Rabbit				x							
Couckuyt and Van Looy	2019	Grey Literature							x				
Dagher and Fayad	2024	Chat GPT	x										
Data Insights Market	2025	Grey Literature	x					x			x		x
Davenport	1993	Grey Literature			x								
Davenport and Redman	2020	Grey Literature			x								
Dong and Quan	2025	Scopus						x					
Dumas et al.	2018	Scopus	x										
Eden et al.	2023	Scopus			x			x					x
Eling and Lehmann	2018	Scopus			x				x		x	x	x
Evans and Fernando	2025	Scopus	x		x								x
Fares et al.	2023	Scopus		x								x	
FDIC	2006	Grey Literature		x									
Firmansyah and Arman	2022	Scopus		x	x	x							
Fischer et al.	2024	Chat GPT				x							
Galazova and Magomaeva	2019	Scopus		x	x								
Genty et al.	2025	Chat GPT		x									
Goldman Sachs	2025	Grey Literature		x									
Gotthardt et al.	2020	Scopus			x	x					x		x
Gul and Al-Faryan	2023	Scopus			x								
Hadi et al.	2023	Grey Literature		x								x	
Hammer and Champy	1993	Grey Literature	x										
Henderson and Venkatraman	1993	Grey Literature	x		x								
Hoesch-Klohe et al.	2010	Grey Literature							x				
Holland et al.	2025	Chat GPT		x									
Hughes et al.	2024	Chat GPT		x									
Kaponda et al.	2023	Chat GPT	x									x	
Kareem et al.	2024	Chat GPT	x										x
Khalatur et al.	2024	Scopus			x			x					

Authors	Year	Database	Process Engineering	Finance	Digital	Data	Theories	Geographies	Sustainability	Gender	Performance	Innovation	Competition
Khatri and Brown	2010	Scopus				x							
Kitsios et al.	2021	Scopus						x					
Kokina and Blanchette	2019	Scopus	x								x		x
Kolisnyk et al.	2023	Scopus						x					
Kooper et al.	2011	Research Rabbit			x	x							
Koshelev	2022	Chat GPT						x					
Kristoffersen et al.	2021	Grey Literature				x	x	x					x
Kubrak et al.	2023	Chat GPT			x								
Lähteenmäki et al.	2022	Scopus	x	x	x								
Lameijer et al.	2024	Scopus	x		x								x
Li et al.	2024	Scopus				x							
Liepert	2024	Scopus									x		
Maletic et al.	2023	Scopus		x									
Marsudi and Pambudi	2021	Chat GPT	x		x	x					x		
Maximize Market Research	2024	Grey Literature	x	x				x					
Mbugua	2021	Chat GPT	x										x
McAfee and Brynjolfsson	2012	Grey Literature									x		x
McKinsey	2024	Grey Literature									x		
Meadows and Merali	2003	Grey Literature		x									
Mikalef et al.	2018	Grey Literature				x	x						
Monk et al.	2017	Research Rabbit				x							
Muthu et al.	1999	Grey Literature	x										
Naimi-Sadigh et al.	2022	Scopus			x								x
Nesindande et al.	2024	Scopus		x									
Nkomo and Marnewick	2021	Chat GPT	x										
Onjure et al.	2018	Chat GPT	x					x					x
Otto	2011	Grey Literature			x	x							
Otto et al.	2007	Research Rabbit				x							
Paul and Sadath	2021	Scopus			x	x		x					
Peretz-Andersson et al.	2024	Chat GPT			x								
Pisoni et al.	2023	Scopus		x		x							x
Porfírio et al.	2024	Chat GPT	x		x				x	x	x		
Pramod et al.	2022	Scopus				x							
Priyanto et al.	2023	Scopus			x	x							
Pronoza et al.	2023	Scopus		x									
Rahim and Chishti	2024	Scopus		x									
Reis et al.	2023	Scopus	x										
Rodrigues et al.	2022	Scopus		x	x						x	x	x
Rosenbaum	2010	Research Rabbit				x							
Saidat et al.	2023	Research Rabbit		x	x								
Salerno and Maçada	2025	Chat GPT				x							
Sambamurthy and Zmud	1999	Research Rabbit			x								
Santos and Lucas	2019	Grey Literature			x								
Sirmon et al.	2011	Grey Literature					x						x
Smith and McKeen	2008	Research Rabbit				x							
Sneed and Verhoef	2001	Grey Literature	x										
Spilbergs	2023	Scopus						x					
Stefanelli and Manta	2023	Scopus		x				x	x		x		x
Taneja et al.	2024	Scopus	x										
Taymouri et al.	2021	Scopus			x	x							
Teece et al.	1998	Research Rabbit					x						

Authors	Year	Database	Process Engineering	Finance	Digital	Data	Theories	Geographies	Sustainability	Gender	Performance	Innovation	Competition
Tekka, R. S.	2021	Chat GPT	x										x
Tinnilä	1995	Grey Literature			x				x				
Tran et al.	2022	Scopus			x			x					
Tsakalidis and Vergidis	2024	Chat GPT	x										
Tsakalidis and Vergidis	2025	Chat GPT		x	x								
van der Aalst et al.	2022	Scopus				x					x		
Vial	2023	Research Rabbit		x									
Volberda et al.	2021	Chat GPT			x								
Weber et al.	2009	Grey Literature			x								
Weill and Ross	2005	Research Rabbit			x	x							
Xiang et al.	2019	Scopus		x	x	x							
Zamula et al.	2020	Scopus		x		x							
Zelt et al.	2019	Scopus	x										
Zeng and Khan	2019	Chat GPT					x						
Zou	2023	Scopus		x	x								

Appendix C. Semi-structured interview questions

1. Can I record this interview for academic research purposes only?

e.g. *"Before we start, I just want to clarify a couple of practical points. This interview is part of my academic research and will be used only for research purposes. The recording will be transcribed and anonymised, and no personal or company-specific information will be disclosed. Is it ok for you if I record the interview for academic research purposes only?"*

2. Which company do you work for, and which is your role?

e.g *"To better contextualise your answers, it would be useful to understand your background. Could you briefly tell me which company you work for and what your role is, especially in relation to data, governance or operational processes? This will help me interpret your perspective throughout the interview."*

3. How does your team interact with data governance practices?

e.g. *"Before going into more technical aspects, I would like to understand how data governance fits into your daily work. I am not referring only to formal policies, but also to how governance is experienced in practice by your team. I am interested in understanding how often you interact with governance rules, controls or procedures, and whether they influence everyday decisions or workflows. Could you describe how your team interacts with data governance practices in practice?"*

4. How data governance helps structure the company data? Challenges?

e.g. *"Many organisations struggle with fragmented data, legacy systems or inconsistent definitions. I would like to understand how governance helps bring order to this complexity in your organisation. From your experience, how does data governance contribute to structuring data across the company, for example in terms of definitions, documentation or data flows? And what are the main challenges you encounter when trying to maintain this structure over time?"*

5. How data governance helps the technology and infrastructure adoption? Challenges?

e.g. *"Now I would like to focus more on the technical side. In many companies, adopting new technologies such as cloud platforms, APIs or AI tools depends heavily on how data is governed. Thinking about your experience, how does data governance enable or support the adoption of new technologies and infrastructures? And are there situations where governance requirements make this process more complex or slower?"*

6. How data governance helps developing use cases? Challenges?

e.g. *"I am also interested in understanding the role of governance in the development of concrete data-driven use cases. When teams work on new products, analytical models or automation initiatives, governance can play different roles. From your perspective, how does data governance support the development of new use cases? And what are the main difficulties or limitations that teams typically face in this phase?"*

7. Can you describe one successful or failure case where data governance had a relevant role?

e.g. *"To make this more concrete, I would like to ask you about a specific example. It can be either a success or a failure. Can you think of a case where data governance had a significant impact on the outcome of a project, for instance enabling a use case, slowing it down, or even blocking it? I am interested in understanding what happened and why governance was relevant in that situation."*

8. How can data governance influence the achievements of ESG goals?

e.g. *"Finally, I would like to broaden the perspective to sustainability and responsibility. There is increasing attention on ESG aspects, and data plays a central role in measuring and reporting them. From your point of view, how can data governance influence the achievement of ESG goals, for example in terms of transparency, fairness, inclusion or environmental efficiency?"*

Appendix D. Overview of regulatory jurisdictions

Regulatory configuration	Key characteristics	Representative jurisdictions
Regulatory-based	Binding legal frameworks, mandatory participation, supervisory oversight	EU, UK, Australia, Brazil, India
Market-based	Contractual agreements, competitive dynamics, limited mandatory regulation	United States, Switzerland
Voluntary-based	Guidelines, industry standards, discretionary participation	Singapore, Japan, South Korea
Hybrid	Combination of mandatory and voluntary mechanisms across markets	Cross-border and multi-market contexts

List of Figures

List of Tables