

Article

Not peer-reviewed version

SDAF: Strategic Data Governance for Cloud-Enabled Big Data Analytics

[Rahid Zahid Alekberli](#)^{*} and Robert Haussmann

Posted Date: 25 February 2026

doi: 10.20944/preprints202602.1406.v1

Keywords: cloud-enabled big data analytics; data governance; strategic alignment; cloud computing; big data analytics; hybrid cloud; operating model; qualitative study; seaports



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

SDAF: Strategic Data Governance for Cloud-Enabled Big Data Analytics

Rahid Zahid Alekberli ^{1,*} and Robert Haussmann ²

¹ Azerbaijan Technical University

² Walden University

* Correspondence: ralekberli@gmail.com

Simple Summary

Ports increasingly rely on cloud platforms, sensors, and analytics to improve their safety, efficiency, and competitiveness. Many big data programs still fail because teams do not agree on who owns critical data, how data quality is controlled, and how governance supports business goals and compliance. We interviewed 14 senior stakeholders from Caspian Basin seaports and related regulatory and strategic roles to gather data. They described recurring barriers: siloed communication, limited budgets, capability gaps, and difficulty in turning compliance requirements into everyday controls. Based on these insights, we propose a six-step Strategic Data Alignment Framework (SDAF) that helps ports start with a clear diagnosis of strategy–data gaps, set minimum governance foundations, prioritize decision-critical use cases, embed governance into planning and KPI cycles, and continuously review and improve their implementation. The framework provides a practical sequence for implementing data governance in resource-constrained, regulated environments.

Abstract

Cloud-enabled big data programs in hybrid environments often underdeliver, and the conventional assumption that tooling or analytics maturity is the primary constraint is beginning to show cracks. Instead, the limiting factor is the frequent misalignment between governance arrangements and corporate strategy, compounded by ambiguous risk ownership and weak operational integration. This study develops a Strategic Data Alignment Framework (SDAF) for small and medium Caspian Basin seaports to translate this alignment problem into an executable implementation logic. Using an interpretive qualitative design, 14 semi-structured interviews were conducted with port decision-makers, strategy and IT leaders, and relevant regulators, and the transcripts were analyzed to surface recurring mechanisms of failure. The data reveal consistent breakdowns: unclear accountability for data decisions, immature stewardship, siloed collaboration, capability shortfalls, and under-instrumented feedback loops that fail to connect governance controls to the business outcomes. The SDAF addresses these mechanisms through a six-step pathway: (1) diagnose strategy–data misalignment, (2) establish minimum viable governance foundations, (3) specify decision-critical use cases and the role of big data in strategy, (4) embed governance into planning and performance cycles, (5) execute and monitor via measurable KPIs and auditable decision metrics, and (6) iteratively review and improve through a continuous alignment loop, calibrated to regulatory context, platform maturity, and risk appetite.

Keywords: cloud-enabled big data analytics; data governance; strategic alignment; cloud computing; big data analytics; hybrid cloud; operating model; qualitative study; seaports

1. Introduction

Although cloud computing and hybrid architectures offer significant advantages for modern analytics, their adoption in regulated, multi-stakeholder environments often introduces substantial governance complexity. This is because data become more distributed, lineages are harder to maintain, and accountability can be fragmented across organizational and vendor boundaries. In addition, high-velocity provisioning and tool democratization can accelerate “shadow analytics/AI” practices, increasing compliance and risk exposure when decision rights and controls are not operationalized. These issues directly align with the focus of this Special Issue on cloud computing and big data mining, which emphasizes practical approaches to security, privacy, and dependable analytics across distributed infrastructures.

Seaports are increasingly data-intensive systems where excellence, safety, resilience, and competitiveness rely on sensing, integrating, analyzing, and acting on varied data streams in near-real time. Port communities use digital and automation technologies, from terminal operating systems and IoT monitoring to AI decision support and predictive maintenance, to improve throughput, reduce dwell time, enhance security, and optimize energy and emissions. Prior work highlights that “smart port” capabilities emerge from the diffusion of modern technologies, interoperable information systems, and digitally enabled business processes among port stakeholders.[1–4]. These technologies do not generate value automatically; without clear decision rights, data controls, accountability structures, and alignment with objectives, data programs can fragment into isolated initiatives that create compliance, cybersecurity, and performance risks rather than advantages.

This challenge is particularly salient in strategically exposed and rapidly evolving trade geographies, such as the Caspian Basin. The region lies at the Europe-Asia logistics intersection and is influenced by macro-corridor strategies, including the Belt and Road Initiative (BRI) and multimodal routes connecting China, Central Asia, the South Caucasus, and Europe [5–11]. The Trans-Caspian “Middle Corridor” has renewed attention on container flows, multimodal reliability, and the performance of supporting seaports and inland logistics nodes [6]. In this context, small and medium seaports are expected to meet higher expectations for transparency, predictability, and integrated services while operating with tighter budgets, limited specialist capacity, and uneven institutional support.

Simultaneously, “big maritime data” are increasingly operationally consequential. Vessel tracking (e.g., AIS), terminal telemetry, customs and trade data, weather and sea state signals, equipment condition data, security logs, and commercial datasets can be combined to optimize planning and reduce disruption. Survey work on big maritime data management emphasizes both the opportunities for improved decision-making and the practical complexities of data integration, governance, and lifecycle management across diverse actors [3,4]. The adoption of Industry 4.0 in ports similarly underlines that digitization spans technology, process redesign, organizational change, and governance, not merely IT upgrades [2]. Consequently, a port’s ability to convert data into performance outcomes depends on a governance architecture that specifies “who decides what,” under which controls, with which metrics, and how decisions remain connected to strategic priorities.

A persistent problem is the frequent treatment of data governance as merely a compliance or IT issue. However, realizing competitive and operational impacts demands explicit links to corporate strategy and business model selection.” Knowledge management research frames the “big data–business strategy interconnection” as a continuing challenge: organizations must integrate analytics capabilities, governance mechanisms, and strategic intent to avoid misalignment between what data initiatives produce and what the enterprise is trying to achieve [12]. Strategy-oriented frameworks similarly stress the need to translate big data ambitions into coherent strategic choices—prioritizing decision use cases, building enabling capabilities, and ensuring governance and accountability—so that analytics investments support value creation rather than proliferating as disconnected projects [13]. Systematic reviews in supply chain and logistics further warn that big data initiatives can

introduce ethical, privacy, and power-asymmetry concerns, reinforcing the need for governance that goes beyond technical data handling toward responsible use and stakeholder legitimacy [14].

Despite the breadth of research on port digitization, the integration problem remains underexplored for small and medium seaports in the Caspian Basin: what “critical pillars” are required to align Big Data governance (BDG) with corporate strategy under resource constraints, and how stakeholders perceive obstacles, enablers, and implementation pathways in this regional setting [15,16]. This study addresses this gap by examining how stakeholders interpret and operationalize the integration of the BDG and corporate strategy in Caspian Basin seaports, emphasizing how digital transformation and data initiatives can be translated into coherent strategic directions, measurable performance improvements and sustainable governance practices.

1.1. Research Gap

Existing data governance and strategic alignment literature provides valuable definitions, control domains (e.g., stewardship, quality, metadata, privacy), and maturity perspectives for “what to govern” and “what good looks like” [5–7,16,17]. However, organizations still struggle to implement governance that consistently supports corporate strategy in cloud-enabled and hybrid environments, particularly when accountability spans multiple units, platforms, and external stakeholders. Many approaches remain descriptive or target-state oriented; they under-specify implementation sequencing under real constraints, such as limited budgets, siloed decision-making, thin governance capacity, and compliance operationalization demands [16,17]. This gap is especially salient in small and medium Caspian Basin seaport ecosystems, where multi-stakeholder operations and fragmented regulatory requirements interact with constrained investments and mixed legacy/modern technology baselines. Accordingly, this study addresses the lack of empirically grounded, stepwise guidance for aligning BDG with corporate strategy execution in resource-constrained, regulated, and multi-stakeholder contexts [16,17].

1.2. Aim and Contributions

Accordingly, this study proposes the Strategic Data Alignment Framework (SDAF) as an empirically grounded, stepwise guideline that converts the BDG from a policy aspiration into a staged operational pathway for strategy execution in constrained, regulated, multi-stakeholder environments [15,16]. The central research question is as follows: **How can stakeholders improve their understanding of the integration of BDG and corporate strategies in small- and medium-sized seaports in the Caspian Basin?** This study makes three contributions [16,17]:

1. **Empirical evidence** from a regulated, multi-stakeholder domain (Caspian Basin seaports) identifies recurrent governance failure modes and alignment bottlenecks.
2. **A six-step alignment guideline** that is actionable and auditable (strategy → governance operating model → capability building → measurement/feedback).
3. **A traceable mapping** from emergent themes and obstacles to the SDAF pathway to support transferability as implementation logic across similar regulated environments.

Article structure. Section 2 reviews the related literature on data governance, strategic alignment, and cloud-enabled governance complexity. Section 3 details the research design, sampling, data collection and analysis procedures. Section 4 reports the results and presents the SDAF output, including the figures and traceability matrix. Section 5 discusses the interpretation, transferability, implications, and limitations of the study. Finally, Section 6 concludes the paper and outlines future research directions.

2. Background and Related Work

2.1. Data Governance as Decision Rights and Operating Model

Data governance is commonly defined as the specification of decision rights and accountability for data-related decisions to encourage desirable behavior in data use [16,17]. In implementation terms, this definition implies an operating model consisting of (i) governance structures (e.g., data council, data owners, and data stewards), (ii) governance processes (e.g., policy lifecycle, issue escalation, exception handling, and control monitoring), and (iii) governance artifacts (e.g., business glossary, metadata standards, data quality rules, and data classification) that convert policy intent into executable controls [16,17].

2.2. Strategic Alignment and Governance

Strategic alignment concerns the fit between business strategy and enabling information systems and organizational capabilities [16]. Governance operationalizes this fit by assigning ownership, enforcing decision rights, and maintaining consistency between strategic objectives and data practices. When governance is weak or disconnected from strategy, predictable outcomes ensue, including conflicting definitions, duplicated datasets, inconsistent metrics, poor data quality, and reduced trust in analytical outputs, all of which undermine decision-making and performance management [16,17].

2.3. Cloud-Enabled Analytics Increases Governance Complexity

Hybrid and multicloud analytics environments intensify governance demands by expanding control surfaces and accountability boundaries. Key stressors include distributed services and toolchains, shared-responsibility interfaces across internal units and providers, rapid provisioning that can bypass controls, and tighter coupling between security, privacy, compliance, procurement, data engineering, and ML operations. These conditions require governance to be operational and measurable, with explicit integration into strategy and performance cycles, so that accountability, control enforcement, and decision quality remain stable as data and workloads become more distributed [16,17].

3. Materials and Methods

3.1. Research Design and Rationale

This study adopts a **qualitative exploratory design** to understand *how stakeholders in small- and medium-sized Caspian Basin seaports can improve their understanding of the integration of the BDG with corporate strategy* [16,17]. We focused on ‘understanding’ because the problem is primarily interpretive and organizational. It concerns how stakeholders conceptualize alignment, assign responsibilities, negotiate cross-unit collaborations, and interpret compliance and resource constraints. Such mechanisms are best surfaced through qualitative inquiry rather than estimated as population-level effects. A quantitative design would be more suitable for a different question (e.g., “What is the impact of SDAF on governance maturity, compliance incidents, or decision-cycle time?”) This would require standardized measures, a larger sample, and pre/post or comparative data [16,17].

Qualitative inquiry is appropriate here because integration challenges (e.g., governance ownership, decision rights, and cross-functional alignment) are context-dependent and best surfaced through participants’ lived experiences and interpretations. The primary data collection instrument was **semi-structured interviews**, guided by an interview guide designed to cover core topics while allowing emergent themes to surface [16,17].

3.2. Research Context and Sampling Strategy

Participants were recruited using **purposive sampling** to capture information-rich perspectives from stakeholders positioned to observe (or influence) the interface between **data governance** and **corporate strategy** [16,17].

The target informants included senior port executives, relevant experts in national regulatory structures, and strategy-oriented roles related to port operations and digital transformation initiatives [16,17].

Interviews were designed to be **40–60 minutes** (or approximately ~45 minutes depending on participant availability) and were conducted in a setting considered convenient and comfortable for the interviewee [16,17].

3.3. Data Collection Procedures

The principal researcher collected data through one-time, in-depth, semi-structured interviews [16,17]. Interviews were audio-recorded with participant consent and transcribed verbatim to ensure the accuracy and completeness of the textual dataset used for analysis [16,17]. To enhance interpretive richness and support reflexivity, the protocol included contemporaneous field notes documenting the interview dynamics and immediate researcher reflections [16,17]. Participants received clear information regarding (i) the purpose of the study, (ii) the expected scope and time commitment, and (iii) confidentiality and data-handling protections, and provided informed consent prior to their participation [16,17]. The interview protocol and supporting study instruments (e.g., recruitment text, consent materials, and the semi-structured interview guide) are documented in the dissertation appendices associated with this research.

3.4. Data Analysis Approach

The analytical strategy used a **thematic analysis** workflow supported by qualitative data analysis software [16,17].

We imported the interview transcripts into the software and examined them using iterative coding cycles [16,17].

The coding process proceeded in two principal phases.

1. **Open coding (inductive, line-by-line)** was used to identify initial concepts, issues, and meaning units grounded in the interview data [16].

2. **Axial coding (theme consolidation)** was used to connect and condense the initial codes into broader categories and themes that directly addressed the research question and stakeholder narratives [16,17].

This approach supports traceability from raw text → codes → themes → interpretive conclusions, improving the transparency and auditability of qualitative inferences.

3.5. Software Support and Use of AI-Assisted Qualitative Analytics

ATLAS.ti was selected as the primary qualitative analysis platform to organize the transcripts, manage codes, and support thematic pattern detection [16].

In addition, AI-supported features/tools were used to accelerate pattern detection and theme navigation (e.g., surfacing recurring semantic structures); however, interpretive decisions were not delegated to automation. The analytic stance explicitly recognizes that AI tools can enhance speed, while **critical human judgment** remains necessary to preserve validity, particularly when handling large volumes of qualitative texts [16,17].

3.6. Trustworthiness and Quality Assurance

To strengthen qualitative rigor, this study applied the recognized trustworthiness criteria: **credibility, dependability, transferability, and confirmability** [16,17].

Measures included prolonged engagement with the dataset, detailed descriptions of procedures, and maintaining an audit trail from analysis to report writing [16,17].

Discrepant or non-conforming cases were treated as analytically valuable rather than excluded to avoid “forcing” themes and preserve nuance [16,17].

Triangulation was considered a credibility enhancer (e.g., multiple observers/reviewers), consistent with qualitative validity practices [16,17].

3.7. Ethical Considerations

Ethical conduct was addressed through formal safeguards consistent with qualitative interviewing: participation was voluntary, confidentiality protections were emphasized, and informed consent was required prior to engagement [16]. Researcher reflexivity was explicitly recognized as necessary to reduce bias and to manage potential power imbalances during interviews. [16,17].

3.8. Researcher Positionality and Role Separation

Given the potential for prior professional familiarity with the domain, this study considered role separation a methodological and ethical risk. To reduce undue influence and protect candor, participants were informed that responses would be reported in aggregated/anonymized form, interviews followed a consistent protocol with neutral prompts, and the analysis emphasized traceability from raw interview material to codes, themes, and derived framework steps. Where applicable, reflexive memoing was used to document interpretive decisions and potential assumptions during the coding process [16,17].

4. Results

4.1. Research Setting and Analytic Output

This study used a qualitative design based on 14 semi-structured interviews conducted virtually. Participants were recruited via LinkedIn based on professional role, expertise, and involvement in decision-making related to BDG and corporate strategy. The analysis yielded empirically derived themes and obstacle patterns that informed the development of the Strategic Data Alignment Framework for small- and medium-sized ports in the Caspian Sea region [16,17].

4.2. Emergent Themes (T1-T7)

The inductive analysis surfaced a set of themes describing what must be present—organizationally, financially, and institutionally—for BDG to function as an operational capability supporting strategy execution [16,17]. The themes (T1-T7) are synthesized below [16,17]:

1. **Alignment and integration of BDG with corporate strategy**

Integration is not automatic; stakeholders emphasized an explicit integration model linking governance mechanisms to strategic planning and execution cycles (not “governance for governance’s sake”).

2. **The role of BDG in driving corporate strategy**

Governance enables “data-driven strategy,” where strategy formation and refinement are informed by trustworthy information assets and decision analytics.

3. **Organizational culture and operational capacity**

Cultural readiness and operational maturity determine adoption; participants stressed building a data culture, upgrading staff capabilities, and establishing internal ownership of data processes.

4. **Communication and collaboration**

Cross-functional coordination (strategy, operations, IT/data, risk/compliance) is a prerequisite; remedies emphasize reducing silos and formalizing collaboration routines.

5. **Financial resources and investments**

Resource constraints are dominant for small/medium ports; “minimum viable governance” must be staged and justified through measurable operational gains.

6. Regulatory challenges and compliance

Governance must be designed with compliance in mind (not appended later); participants highlighted regulatory readiness as part of the governance model.

7. Long-term strategic benefits

Governance was linked to long-horizon advantages (competitiveness, resilience, sustained optimization), that is, governance as strategic capability rather than only a control layer.

4.3. Obstacle Clusters and Reinforcement Logic

Across participants, obstacles clustered into recurring bottlenecks [16]:

- **Communication barriers** (siloed decision-making and weak cross-unit alignment)
- **Financial constraints** (insufficient resources for sustained capability-building and modernization)
- **Regulatory challenges** (unclear/fragmented requirements or weak compliance structures)
- **Compliance complexity** (difficulty operationalizing governance controls consistently)

These obstacles are mutually reinforcing (e.g., limited budget → underinvestment in skills/controls → weaker compliance posture → slower modernization) [16,17].

4.4. Coding Evidence (Selected Excerpts / Code-To-Cluster Mapping)

These open codes were later consolidated into the obstacle clusters reported in Section 4.3. To improve the transparency of how the obstacle clusters were derived, Table 1 provides examples of inductive codes and representative excerpts illustrating the interview material that underpins the obstacle patterns summarized in Section 4.3 [16,17].

Table 1. Inductive Coding Results.

Obstacle	Initial code	Description	Interview excerpt
Resistance to change	Resistance from experienced business personnel	Experienced business personnel, often older, are resistant to adopting new Big Data strategies due to familiarity with traditional practices and fear of losing control.	“Many very experienced people in business... are still more familiar with older business practices. Sometimes, it becomes challenging to negotiate and explain to them why describing the business goals and strategies is essential.”
	Difficulty in communication between business and technical teams	There is a significant challenge in ensuring clear communication between business management and technical teams, often requiring intervention from higher management levels.	“We sometimes have to go back to the higher up levels, let’s say, the CEO levels or CIO levels, to make sure that all other, let’s say, management people do follow with these protocols.”
Misalignment of goals	Misalignment between business goals and technical solutions	Business goals sometimes necessitate unorthodox technical solutions, leading to friction with engineers who prioritize optimization and standard practices.	“Sometimes, business goals require that some unorthodox or nonstandard solutions be implemented on the technical side due to pricing reasons or costs.”
Unrealistic expectations	Unrealistic expectations	Business stakeholders often have unrealistic	“Sometimes they have a few unbelievable expectations.”

	from Big Data capabilities	expectations about the capabilities of Big Data and AI, influenced by media and hype.	And so, the vital role that I see we must do was to manage the expectations correctly, so they know what to expect.”
Lack of data governance understanding	Inadequate understanding of data governance importance	There is a lack of understanding that effective data governance requires a well-optimized business governance structure.	“For correct data governance, first, business governance itself has to be optimized and almost perfected so that data governance can be applied.”
Skill gaps	Skill gaps in integrating business and technical knowledge	Effective Big Data strategy implementation requires personnel who are skilled in both business and technical domains, which is often a rare combination.	“People who do data governance or design these strategies must be very advanced and skilled in business and engineering.”

4.5. Governance Breakdown Excerpts

To ground the recurring breakdowns summarized in the abstract, interview material provides illustrative evidence consistent with the emergent themes and obstacle clusters [16,17]. The excerpts below show how breakdowns manifested in practice:

- **Accountability gaps (unclear ownership and responsibility boundaries):** unresolved questions about who owns data decisions and what “data ownership” means operationally (e.g., “*What are the responsibilities of the data owners? What are the challenges faced by data owners...?*”) [16,17].
- **Inconsistent data stewardship (immature/partial stewardship operating model):** Stewardship was repeatedly described as incomplete or still being built (e.g., “*We’re still building the framework for data stewardship.*”) [16,17].
- **Insufficient cross-organizational collaboration (silos and escalation-based coordination):** escalation-based coordination used to align business and technical units (e.g., “*We sometimes have to go back to the higher up levels... to make sure that all other... management people do follow with these protocols.*”) [16,17].
- **Limited capability building (skills, literacy, and “data culture” as prerequisites):** training and literacy framed as necessary governance controls rather than optional support (e.g., “*Educate people on why data quality is essential. What is a data catalog?*”) [16,17].
- **Under-instrumented feedback loops (weak cadence for review, learning, and correction):** institutionalized review routines emphasized as a remedy (e.g., “*Conduct quarterly meetings... [with] data owners and stewards... to discuss data quality.*”), indicating that governance feedback mechanisms were not yet consistently embedded as standard operating cadence [16,17].

4.6. Capacity Building as an Integration Mechanism

A key empirical finding is that building capacity is not an option. It serves as the primary method for integrating governance and strategy, allowing them to be implemented jointly rather than sequentially [16,17]. Operationally, the capacity-building direction implied by the results includes: (i) developing skills pipelines (data literacy → analytics → governance execution), (ii) establishing cross-functional coordination routines (shared decision cadence and accountability), and (iii) reinforcing change adoption so that governance becomes “how work is done,” not a parallel initiative [16,17].

4.7. SDAF Framework Derived from Findings

The empirical results—particularly the emergent themes (Section 4.2) and obstacle clusters (Section 4.3)—were synthesized into a staged guideline, the **Structured Strategic Data Alignment**

Framework (SDAF). The overall framework is illustrated in **Figure 1**, and the corresponding six-step execution pathway is summarized in **Figure 2** [16,17]. The SDAF steps are:

1. Gap analysis of strategy–data alignment (T1, T4).
 2. Establish governance foundations (owners, policies, minimum controls) (T3; T6)
 3. Define the role of Big Data in strategic decision processes (T2, T7, T1, and prioritization of decision-critical use cases).
 4. Align data governance processes with corporate planning and performance cycles (T1, T4).
 5. Execute and monitor through measurable KPIs and decision use cases (T2, T5; T7; T6).
 6. Review and improve iteratively (feedback loop/continuous alignment) (T7, T6, T1).
- Note: Themes T1–T7 are summarized in Section 4.2, and obstacles are summarized in Section 4.3.

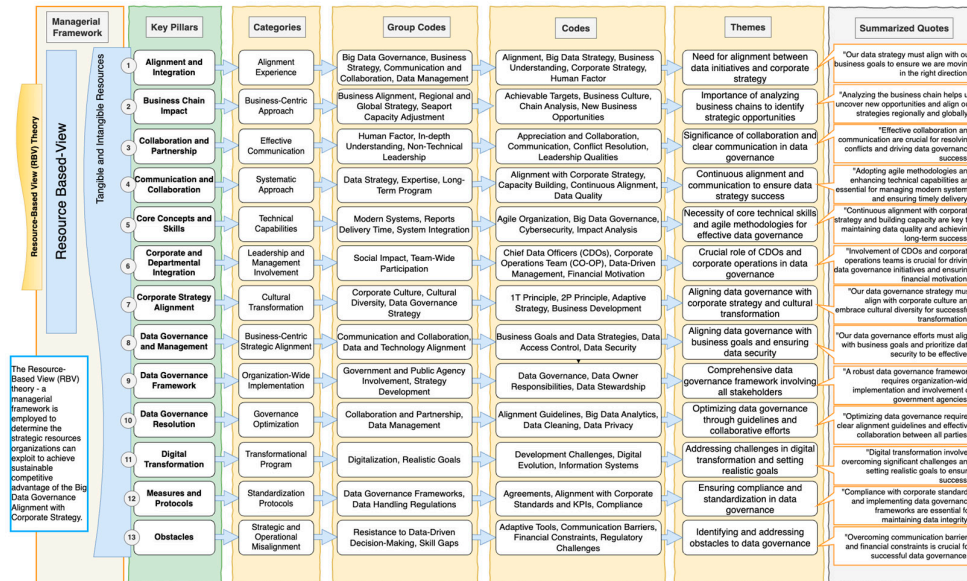


Figure 1. Strategic Data Alignment Framework (SDAF) architecture linking corporate strategy, governance operating model, cross-unit coordination, and compliance-by-design mechanisms in cloud and hybrid environments [16].

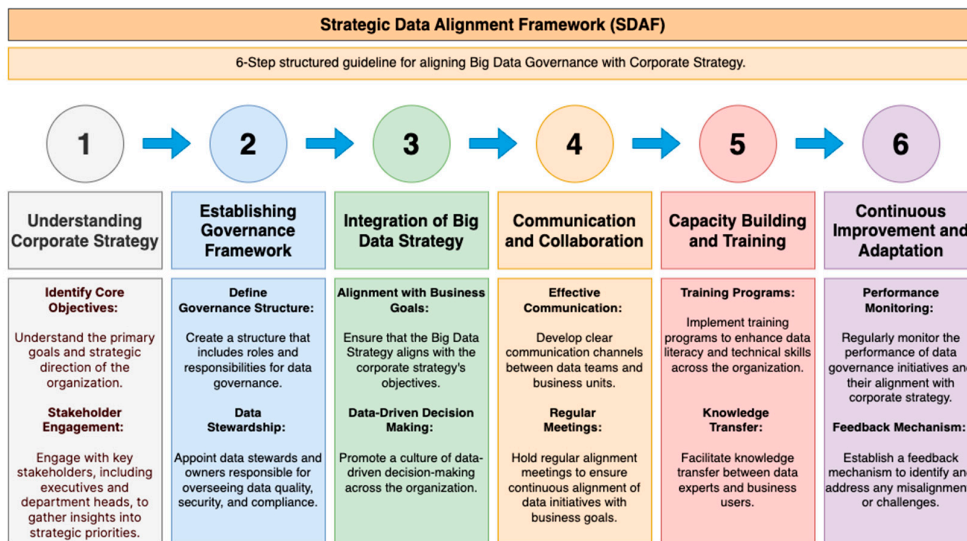


Figure 2. Six-step SDAF execution pathway from strategy–data gap diagnosis to governance foundations, planning-cycle integration, KPI-driven execution, and iterative improvement [16–18].

In practical terms, the SDAF is intended to function as an execution “map” that reduces ambiguity over who decides what, how governance work is prioritized, and how alignment is monitored and adjusted over time [16].

4.8. Traceability Matrix

To strengthen traceability from interview evidence to framework design, **Table 2** maps each SDAF step to its primary emergent themes/obstacles and summarizes the mechanism each step is intended to resolve, alongside illustrative evidence (representative excerpts/examples) supporting the linkage [16,17].

Table 2. Empirical traceability matrix linking SDAF steps to emergent themes/obstacles and illustrative evidence [16].

SDAF step (execution pathway)	Primary empirical drivers (themes/obstacles)	What the step resolves (mechanism)	Illustrative evidence (excerpt/example)
1) Gap analysis of strategy–data alignment	T1 alignment/integration; T2 role of BDG in strategy; Obstacles: communication barriers	Makes misalignment visible (strategy intent vs. data reality), surfaces ownership gaps, and establishes a shared “problem definition”	“Our data strategy must align with our business goals...”; recurring escalation needs to reconcile business and technical interpretations
2) Establish governance foundations (owners, policies, minimum controls)	T3 culture/capacity; T6 regulatory/compliance; compliance complexity; Obstacles: silos	Converts governance from “concept” to minimum viable operating model: ownership, stewardship, definitions, baseline controls	“What are the responsibilities of the data owners...?” and “We’re still building the framework for data stewardship.”
3) Define the role of Big Data in strategic decision processes	T2 data-driven strategy; T7 long-term strategic benefits	Shifts BDG from “IT control” to decision enablement: clarifies decision use-cases, critical data domains, and value logic	Participants framed strategy as process/decision chains that governance must serve (decision-first framing rather than tool-first rollout)
4) Align governance processes to corporate planning & performance cycles	T4 communication/collaboration; T5 financial constraints; Obstacles: silos; budget limits	Institutionalizes alignment as routine (forums, cadence, planning hooks), reducing erosion over time	Evidence of “cadence” remedies: quarterly governance forums to stabilize alignment and coordination
5) Execute and monitor through measurable KPIs and decision use-cases	T7 long-horizon benefits; T6 compliance readiness; Obstacles: operationalization complexity	Turns governance into measurable performance/risk outcomes: KPIs, thresholding, auditability, decision-cycle metrics	Respondents linked governance literacy and data quality discipline to operational outcomes (e.g., customer/service impact logic) and to risk-control posture
6) Review and improve	T3 culture/capacity; T7 sustained advantage;	Creates a learning loop: governance controls and	“Conduct quarterly meetings... [with] data

iteratively (feedback loop / continuous alignment)	Obstacles: reinforcement loops	strategy needs are re-tuned as context changes	owners and stewards..." (institutionalized review)
---	-----------------------------------	---	---

5. Discussion

This study reinforces that in small and medium seaports operating in cloud and hybrid environments, big data governance (BDG) should not be treated as an IT-side compliance layer. Instead, BDG functions as a strategic capability that must be co-designed with corporate strategy to translate analytics investments into measurable performance, resilience, and regulatory assurances. In this framework, strategy defines priority decisions and value logic, whereas BDG provides the decision rights, data quality, integrity, and compliance conditions required to execute those priorities under multi-stakeholder constraints [16,17].

5.1. Alignment as Capability (RBV / Dynamic Capability Interpretation)

The findings suggest that BDG–strategy alignment is sustained (or eroded) through organizational routines rather than one-off projects. Stakeholders have repeatedly emphasized ownership clarity, training, cross-unit communication, and tailoring governance to local constraints as enabling conditions [16,17]. Interpreted through the resource-based view, data assets become strategic resources, and governance routines become capabilities that are difficult to imitate when they are integrated into daily operations. Operationally, this implies institutionalization through (i) explicit ownership (data owners/stewards), (ii) recurring decision forums (e.g., data councils), and (iii) a stable cadence linking data quality, risk, and compliance reviews to strategic KPIs. Under this interpretation, alignment work itself behaves like a dynamic capability: the port’s capacity to repeatedly sense strategic needs, reconfigure data assets, and calibrate governance controls as the context changes [16,17].

5.2. Failure Modes Are Organizational and Regulatory—Not Merely Technical

The dominant breakdowns were not primarily tool-driven; they were structural and contextual, including organizational barriers (siloes decision-making and weak cross-unit alignment), financial constraints, regulatory uncertainty, and the difficulty of operationalizing compliance controls [16,17]. These constraints reinforce each other: communication siloes fragment definitions of “critical data,” financial limitations delay foundational controls (e.g., cataloging, lineage, quality monitoring), and weak compliance integration increases the cost of errors (audit findings and security/privacy exposure), reducing leadership’s willingness to scale data initiatives. Therefore, governance investment should be interpreted as both a performance enabler and a risk-control investment through improved data integrity and compliance [16,17].

5.3. Positioning Relative to Existing Governance Alignment Approaches and Transferability

Existing governance and alignment approaches provide valuable target-state guidance (roles, domains, maturity dimensions). However, they often under-specify implementation sequencing when organizations face simultaneous constraints, such as limited resources, cross-unit siloes, and compliance operationalization burdens [16,17]. The SDAF contributes a stepwise, executable pathway that is empirically grounded in inductive themes and obstacle clusters observed in the field, progressing from alignment diagnosis to minimum viable foundations, planning/performance-cycle integration, KPI-driven execution, and iterative improvement [16,17].

Therefore, transferability should be interpreted as a transfer of mechanisms and sequencing logic rather than a fixed blueprint of artifacts. The six-step pathway is expected to remain stable, while control depth, evidence automation, and investment pacing require calibration to the

regulatory regime, technology baseline (legacy/OT-heavy vs. cloud-enabled), and risk appetite [16,17].

5.4. SDAF as an Executable Roadmap: Linking Findings to Action

SDAF's practical value of the SDAF is that it reduces ambiguity about "where to start" under constrained conditions and helps prevent premature optimization (e.g., advanced analytics without ownership or dashboards without quality thresholds). The traceability from empirical themes/obstacles to SDAF steps is documented in **Table 2 (Section 4.7)** [16,17]. Together, these findings support the interpretation of SDAF as an actionable implementation logic rather than a population-level estimate of effect size [16,17].

5.5. Answering the Research Question and Interpreting Scope of Claims

The results clarify how stakeholders can improve their understanding of cloud-enabled BDG-strategy integration by making integration explicit in operational terms: connecting governance mechanisms to strategy cycles, treating culture/skills/ownership as governance controls, institutionalizing cross-functional routines, and implementing compliance-by-design foundations that remain feasible under limited budgets [16,17]. Because the evidence is qualitative, claims should be treated as analytical generalizations (mechanisms and sequencing logic), with broader applicability requiring replication and implementation studies [16,17].

5.6. Practical Implications

The implications include: (i) institutionalizing cross-functional governance (data council + RACI + escalation paths); (ii) treating governance spending as risk mitigation (auditability, integrity, compliance), not only modernization cost; (iii) prioritizing training and stewardship routines as governance controls; and (iv) rolling out SDAF in stages—foundations first, then scaling analytics once feedback cadence is stable [16,17].

5.7. Social and Policy Impact

Strategy-aligned governance can generate impacts beyond efficiency by strengthening accountability, coordination, and evidence-based oversight in the regulated transport ecosystem. At the individual level, clearer decision-making rights and improved collaboration can reduce bottlenecks and improve safety routines, and capability building can strengthen workforce adaptability. Aligned governance at the organizational and industry levels can improve decision quality, interoperability, and responsible data sharing across ports, supporting corridor-scale coordination. At the national/policy level, coherent governance rules can enable tighter integration with customs, transport, and regulatory networks and support evidence-driven regulation and emergency/environmental coordination [16,17].

5.8. Limitations and Future Research Directions

Limitations include qualitative generalizability, virtual-only interaction, and context sensitivity to local regulatory and resource conditions [16,17]. Future work should test the SDAF through implementation studies (e.g., governance maturity indicators, decision-cycle time, compliance incidents) and compare governance operating models (centralized vs. federated) using mixed evidence sources (document audits, KPI traces, data-quality logs) and longitudinal designs [16,17].

6. Conclusions

This study examined how stakeholders can strengthen their understanding of integrating Big Data governance (BDG) with corporate strategy in small- and medium-sized seaports operating in the Caspian Basin and translated qualitative evidence into actionable managerial guidance. The

findings indicate that alignment is not achieved by adding governance controls after cloud analytics initiatives begin; instead, governance must be embedded early and maintained through strategy and performance cycles so that data initiatives serve corporate objectives and decision-making requirements.

6.1. Main Contributions

Four conclusions were robust across the participants and discrepant cases. First, **stakeholder involvement** is a primary enabling condition: integration improves when executives, operational leaders, and data/technology and risk/compliance stakeholders are engaged from the outset of the integration process. Second, **continuous capability building** is necessary: sustained training and practical literacy are repeatedly described as prerequisites for governance execution, rather than optional support activities. Third, **communication and collaboration operate as governance mechanisms** because structured cross-unit routines are required to stabilize shared definitions, decision rights, and accountability in multi-stakeholder environments. Fourth, **implementation barriers are predictable and must be designed around** communication barriers, financial constraints, and regulatory/compliance challenges that recur and materially shape execution outcomes.

6.2. SDAF as the Operational Outcome

The central operational outcome is the Strategic Data Alignment Framework (SDAF), expressed as a staged six-step execution pathway: (1) gap analysis of strategy–data alignment, (2) establishing governance foundations (owners, policies, minimum controls), (3) defining the role of Big Data in strategic decision processes, (4) aligning governance processes to corporate planning and performance cycles, (5) executing and monitoring through measurable KPIs and decision use cases, and (6) iterative review and improvement through a feedback loop. The framework is best interpreted as an empirically grounded **implementation logic** that reduces ambiguity about where to start and how to sustain alignment under resource- and compliance-related constraints.

6.3. Practical Implications

For seaport leaders and regulators, the results imply four actionable priorities: (i) institutionalize cross-functional governance (clear decision rights, stewardship roles, escalation paths), (ii) treat governance investment as both performance enablement and risk control (integrity, auditability, compliance readiness), (iii) prioritize training as a governance control to build execution capacity, and (iv) stage implementation so that foundational controls precede advanced analytics scaling.

Future Research

Future research should validate the SDAF through implementation-focused designs that move from interpretive mechanisms to quantifiable outcomes. The first priority is comparative replication across different port classes (small/medium vs. hub ports) and governance regimes to test whether the six-step sequencing remains stable under varying compliance demands and technology baselines (legacy OT/ICS-heavy vs. cloud-native). Second, longitudinal or pre/post studies should quantify whether SDAF adoption improves governance maturity, auditability/compliance posture, data quality, and decision-cycle performance (e.g., time-to-decision, exception resolution time, and KPI reliability). Third, mixed-evidence triangulation, which combines interviews with document audits, data quality logs, catalog/lineage coverage, and incident records, can strengthen causal plausibility and reduce reliance on self-reporting. Finally, future work should examine operating-model choices (centralized vs. federated governance, council designs, and RACI variants) and cost-effective capability-building strategies suitable for resource-constrained multi-stakeholder ecosystems, including corridor-scale interoperability requirements and cross-border data-sharing governance.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Walden University Institutional Review Board (approval no. 02-28-24-1155194).

Informed Consent Statement: Informed consent was obtained from all the participants involved in the study. Participants were provided with information regarding the study's purpose, procedures, voluntary nature of participation, and confidentiality protections prior to the interviews, and consent included permission to audio-record and transcribe interviews for analysis.

References

1. Pavlič Skender, H.; Ribarić, E.; Jović, M. An Overview of Modern Technologies in Leading Global Seaports. *Pomorski zbornik* **2020**, *59*, 35–49.
2. De La Peña Zarzuelo, I.; Freire Soeane, M.J.; López Bermúdez, B. Industry 4.0 in the Port and Maritime Industry: A Literature Review. *Journal of Industrial Information Integration* **2020**, *20*, 100173, doi:10.1016/j.jii.2020.100173.
3. Heilig, L.; Voß, S. Information Systems in Seaports: A Categorization and Overview. *Inf Technol Manag* **2017**, *18*, 179–201, doi:10.1007/s10799-016-0269-1.
4. Herodotou, H.; Aslam, S.; Holm, H.; Theodossiou, S. Big Maritime Data Management. In: 2020.
5. Cao, M.; Alon, I. Intellectual Structure of the Belt and Road Initiative Research: A Scientometric Analysis and Suggestions for a Future Research Agenda. *Sustainability* **2020**, *12*, 6901, doi:10.3390/su12176901.
6. Palu, R.; Hilmola, O.-P. Future Potential of Trans-Caspian Corridor. *Logistics* **2023**, *7*, 39, doi:10.3390/logistics7030039.
7. Plekhanov, D.; Franke, H.; Netland, T.H. Digital Transformation: A Review and Research Agenda. *European Management Journal* **2022**, S0263237322001219, doi:10.1016/j.emj.2022.09.007.
8. Lee, P.T.-W.; Hu, Z.-H.; Lee, S.; Feng, X.; Notteboom, T. Strategic Locations for Logistics Distribution Centers along the Belt and Road: Explorative Analysis and Research Agenda. *Transport Policy* **2022**, *116*, 24–47, doi:10.1016/j.tranpol.2021.10.008.
9. Lee, P.T.-W.; Zhang, Q.; Suthiwartnarueput, K.; Zhang, D.; Yang, Z. Research Trends in Belt and Road Initiative Studies on Logistics, Supply Chains, and Transportation Sector. *International Journal of Logistics Research and Applications* **2020**, *23*, 525–543, doi:10.1080/13675567.2020.1828318.
10. Örmeci, O. Sino-Turkish Relations and the Belt and Road Initiative. *Türkiye-Çin ilişkileri ve kuşak yol projesi* **2022**.
11. Rahman, Z.U. A Comprehensive Overview of China's Belt and Road Initiative and Its Implication for the Region and Beyond. *Journal of Public Affairs* **2022**, *22*, doi:10.1002/pa.2298.
12. Ciampi, F.; Marzi, G.; Demi, S.; Faraoni, M. The Big Data-Business Strategy Interconnection: A Grand Challenge for Knowledge Management. A Review and Future Perspectives. *JKM* **2020**, *24*, 1157–1176, doi:10.1108/JKM-02-2020-0156.
13. Mazzei, M.J.; Noble, D. Big Data Dreams: A Framework for Corporate Strategy. *Business Horizons* **2017**, *60*, 405–414, doi:10.1016/j.bushor.2017.01.010.
14. Ogbuke, N.J.; Yusuf, Y.Y.; Dharma, K.; Mercangoz, B.A. Big Data Supply Chain Analytics: Ethical, Privacy and Security Challenges Posed to Business, Industries and Society. *Production Planning & Control* **2022**, *33*, 123–137, doi:10.1080/09537287.2020.1810764.
15. Lumbroso, D.; Tsarouchi, G.; Campbell, A.; Davison, M.; Merrien, A.; Aragones, V.; Kristensen, P.; Li, I.; Sforzi, G.; Barber, D.; et al. Adapting Caspian Sea Ports to Climate-Induced Water Level Declines: The Case of Aktau 2025.
16. Alakbarli, R. Integrating Big Data Governance and Corporate Strategies in Small and Medium Caspian Basin Seaports. PhD Thesis, Walden University: Minneapolis, MN, USA, 2024.
17. Alekberli, R.Z.; Haussmann, R.E. Integrating Big Data Governance and Corporate Strategies in Small and Medium Caspian Basin Seaports. In Proceedings of the 2024 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT); IEEE: Dubai, United Arab Emirates, 2024; Vol. S1 (AI and IoT from Theory to Practice), pp. 1–6.

18. Alekberli, R.; Alguliev, R. Internets Copy: Current State, Problems and Perspectives. In Proceedings of the 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT); IEEE, 2014; pp. 1–7.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.