International Journal of Computer (IJC)
ISSN 2307-4523 (Print & Online)
https://ijcjournal.org/index.php/InternationalJournalOfComputer/index

# Architecture for Integrating Federal Databases into a Unified State Monitoring Perimeter

Pratikkumar Chaudhari*

*Project Management Analyst, Datics Inc., 13717 S. Route 30, Unit 105B, Plainfield, IL - 60544*
*Email: pratikkumar.chaudhari95@gmail.com*

**Abstract**

The article examines the strategic relevance of database integration amid fragmented regulatory data, cross-agency dependence, and a growing demand for evidence-based public oversight. Its purpose is to develop a theoretical model and target architecture that connect governance arrangements, metadata management, data quality rules, lineage requirements, and platform design within a single supervisory system. The study proceeds from the premise that effective state monitoring depends on semantically aligned, traceable, and standardized data flows across heterogeneous federal and departmental sources. The novelty of the article lies in the proposed model and architecture of a unified state monitoring perimeter grounded in DAMA-DMBOK principles and adapted to regulatory banking practice. The main conclusions show that integration should be organized through a layered architecture including source systems, an integration layer, a harmonization layer, a quality and lineage control layer, and a unified analytical platform functioning as a Single Source of Truth. The article also demonstrates that federated data governance with centralized standards provides the institutional basis for accountability, coherence, and supervisory use of integrated data. The article will be useful for researchers, public administrators, financial regulators, and designers of state information systems.

***Keywords:*** Federal databases; unified state monitoring perimeter; data governance; integration architecture; DAMA-DMBOK; metadata; data lineage.

---

*Received: 3/24/2026*
*Accepted: 5/24/2026*
*Published: 6/2/2026*

---

* *Corresponding author.*

## 1. Introduction

The integration of federal databases into a unified state monitoring perimeter has become a strategic issue for public supervision because digital regulation now depends on high-volume, cross-institutional data flows that must support licensing, risk assessment, compliance validation, and enforcement. Research on data governance shows that public institutions require governance arrangements that connect data quality, accountability, metadata, and decision-making use in a single operating model, which makes integration architecture a governance problem as much as an infrastructure problem [1]. In the financial sector, the expansion of RegTech and SupTech has increased the dependence of supervisory capacity on interoperable data environments, shared standards, and machine-readable regulatory information [2]. Under these conditions, the question of how federal databases can be aligned within a single monitoring perimeter is directly relevant to state authorities that oversee banking and financial institutions through data-intensive supervisory processes. The core obstacle lies in fragmentation. Regulatory information is dispersed across reporting platforms, licensing systems, document repositories, payment channels, and external federal sources, each designed for distinct legal and operational purposes. Studies of data ecosystems indicate that fragmented institutional arrangements weaken coordination, reduce semantic consistency, and obstruct shared accountability across organizational boundaries [3]. Data governance research also shows that weak integration leads to recurring problems with lineage, access control, stewardship, and the reliable reuse of data in high-stakes settings [4]. In a state monitoring context, this fragmentation affects the comparability of supervisory records, the traceability of regulatory evidence, and the integrity of institution-level risk profiles. The result is a monitoring environment in which data may exist in large quantities while remaining difficult to reconcile into a trusted supervisory picture.

This article aims to develop a theoretical architecture for integrating federal databases into a unified state monitoring perimeter through a data governance lens grounded in DAMA-DMBOK principles and adapted to regulatory banking practice. The study treats integration as a structured combination of governance roles, metadata controls, quality rules, lineage requirements, and platform design that can support a Single Source of Truth for supervisory use. This framing is consistent with recent scholarship that links strong data governance to trustworthy data operations, cross-domain coordination, and durable institutional control in complex data environments [3]. The article's scientific value lies in connecting public-sector data integration with regulatory oversight theory. The practical value lies in offering an architectural foundation for state bodies that need audit-ready, standardized, and analytically usable data to detect systemic risk and sustain data-driven supervision.

## 2. Materials and Methodology

The study is grounded in a targeted review of twelve sources drawn from peer-reviewed scholarship, supervisory guidance, and conceptual work on public-sector data governance, interoperability, provenance, and regulatory oversight. Its material base includes studies on trustworthy data governance in state institutions, integrated governance in data ecosystems, inter-organizational process governance, heterogeneous database environments, interoperability, infrastructure-level bureaucracy, semantic integration, and provenance, together with the 2024 ECB supervisory guide on risk data aggregation and risk reporting (Janssen and his colleagues, 2020; Bagherifam and his colleagues, 2025; Bartolomucci and his colleagues, 2026; Bernardo and his colleagues, 2024; Ribeiro and

his colleagues, 2024; Abbasi and his colleagues, 2026; ECB Banking Supervision, 2024; Berg, 2024; Widlak and Peeters, 2025; Kalampokis and his colleagues, 2023; Mitzutani and his colleagues, 2021; Gierend and his colleagues, 2024). These materials were selected because they capture the institutional, semantic, and evidentiary dimensions of database integration in supervisory environments where licensing, compliance control, risk aggregation, and auditability depend on shared standards, lineage visibility, and durable governance arrangements.

Methodologically, the article applies a conceptual-analytical design oriented toward architecture formation. The first procedure consists of a comparative synthesis of the selected literature in order to identify recurrent integration problems across fragmented regulatory data environments, including weak cross-agency coordination, schema heterogeneity, metadata inconsistency, data quality defects, and limited traceability of transformations (Bartolomucci and his colleagues, 2026; Bernardo and his colleagues, 2024; Abbasi and his colleagues, 2026; Berg, 2024; Gierend and his colleagues, 2024). The second procedure consists of normative-structural modeling through the DAMA-DMBOK lens adapted to banking supervision, where governance roles, metadata controls, quality rules, lineage requirements, and access principles are assembled into a theoretical architecture for a unified state monitoring perimeter and a Single Source of Truth suitable for supervisory use (Janssen and his colleagues, 2020; Ribeiro and his colleagues, 2024; ECB Banking Supervision, 2024). This methodology enables the derivation of an architecture that treats integration as an institutional system for controlled data circulation, semantic harmonization, and evidentiary consistency across federal and departmental sources.

## 3. Results and Discussion

The body of prior research outlines several complementary trajectories that frame the present study. Janssen and his colleagues develop a foundational account of data governance as an organizing system for trustworthy data operations in public institutions and link governance arrangements to the reliability of artificial intelligence applications in state settings [1]. Bagherifam and his colleagues examine the trajectory of digital regulatory governance and document the role of RegTech and SupTech in reshaping financial oversight and administrative capacity [2]. Bartolomucci and his colleagues consolidate collaborative and data governance traditions into an integrated framework for data ecosystems and address the institutional preconditions of cross-organizational coordination [3]. Bernardo and his colleagues survey data governance and quality management across multiple domains and trace the connection between governance arrangements and operational data reliability [4]. Ribeiro and his colleagues model inter-organizational business process governance in collaborative networks and clarify the allocation of ownership and accountability for shared data assets [5].

Abbasi and his colleagues develop an API-driven architecture for unified data governance across heterogeneous database environments and address policy enforcement at the platform level [6]. The ECB Banking Supervision guide on risk data aggregation and risk reporting documents persistent structural weaknesses in supervisory data pipelines and specifies requirements for ownership, lineage, validation, and audit-trailled control [7]. Berg analyzes interoperability as a socio-technical condition that integrates technical standards, information architecture, and organizational governance [8]. Widlak and Peeters formulate a theory of infrastructure-level bureaucracy and trace how inter-organizational data exchange redistributes power and responsibility across connected agencies [9].

Kalampokis and his colleagues review the use of emerging technologies in the public sector and document the conditions under which fragmented operational traces can be reassembled into integrated supervisory views [10]. Mitzutani and his colleagues address semantic data integration with DevOps practices and connect schema heterogeneity with engineering processes [11]. Gierend and his colleagues provide a scoping review of provenance information for biomedical data and workflows and consolidate the formal apparatus for reconstructing data origins across distributed systems [12]. The convergence of these works around governance, semantic alignment, quality control, and provenance motivates the unified architectural perspective adopted in this article. In the regulatory sphere, Data Governance defines the allocation of decision rights, stewardship duties, control procedures, and metadata rules that shape how supervisory data are created, validated, shared, and used across institutional boundaries [4]. In public and inter-organizational settings, governance must also specify ownership of shared data assets, accountability for operations, and coordination rules for data exchange among multiple actors participating in a single control process [5]. Within this logic, the DAMA-DMBOK framework remains a useful conceptual basis because it links data architecture, data quality, metadata, lineage, stewardship, and policy management into a single management system that supports state supervision of banking and financial institutions [6]. For a unified state monitoring perimeter, this framework provides a language for connecting federal sources, internal repositories, reporting workflows, and decision layers through common definitions, controlled transformations, and assigned responsibilities.

Supervisory data pipelines in banking regulation carry a distinct operational burden because they support risk aggregation, prudential reporting, case analysis, remediation tracking, and management oversight within a single chain of evidence. Recent supervisory guidance from ECB Banking Supervision shows that institutions still face structural weaknesses in risk data aggregation and reporting, including reconciliation errors, manual adjustments, incomplete source data, weak quality controls, and long production times for monthly risk reports [7]. In this environment, data quality, lineage, and auditability become core supervisory conditions. The same guidance requires clear data ownership, uniform glossaries, validation rules, complete and up-to-date lineage at the data-attribute level, issue registers, remediation processes, and audit-trailled control of manual workarounds. These requirements show why supervisory pipelines cannot be treated as simple technical transfer channels. They operate as regulated evidence systems in which every critical data element must remain traceable from capture to reporting use, with documented quality status and a form fit for external validation and internal accountability.

The current model of working with federal and departmental data is constrained by fragmentation across institutions, platforms, and legal mandates. Research on interoperability describes cross-system exchange as a socio-technical condition that depends on technical standards, information architecture, and organizational governance, which means fragmented supervision cannot be resolved through interfaces alone [8]. Studies of public-sector data exchange show that growing interdependence among government bodies creates infrastructure-level dependencies that redistribute power, responsibility, and control across connected organizations [9]. In a regulatory environment, this fragmentation separates reporting platforms, licensing systems, document repositories, payment channels, and federal reference datasets into parallel data domains that produce partial views of the same supervised entity. This condition weakens the continuity of monitoring and complicates the construction of a single institution-level risk profile from multiple operational traces [10].

Diverse databases also provide a second fault line at the semantic layer, in which entities, attributes, statuses and reference values can be defined using different vocabularies and databases. Governance work in heterogeneous database environments has continued to highlight the challenges of policy fragmentation, schema heterogeneity and inconsistent metadata [6]. In the area of semantic integration, similar mismatches between formal schemas and business meaning linked to heterogeneity of data environments degrade shared record interpretability and integrated query accuracy [11]. These semantic fractures further interact with problems of data quality. However, the 2024 ECB Guide on risk data aggregation and risk reporting states that many important risk indicator ratios and limits were miscalculated because of weaknesses in reconciliations, large or critical manual corrections, gaps or inconsistencies in the underlying data, and insufficient quality control. It also reports production times of 40 or more working days for monthly risk reports in many cases. The persistence of such defects matters for supervisory work because data quality determines whether records can support risk aggregation, trend analysis, remediation tracking, and prudential judgment under one evidentiary standard.

A third structural weakness concerns the limited traceability of data origin and transformation. Provenance research treats lineage as a formal record of where data came from, how they changed, and which processes and actors shaped each output, which makes lineage a precondition for assessing reliability and reuse [12]. The same logic appears in recent governance models for heterogeneous databases, where standardized metadata schemas and operation-level records are used to reconstruct data provenance across distributed systems [6]. When lineage is incomplete, a regulator cannot test whether a supervisory indicator reflects original source values, a manual intervention, or a derived transformation performed outside controlled workflows. This gap increases the likelihood of weak regulatory decisions and slows the detection of systemic risk concentrations that depend on integrating many signals across institutions and domains. The problem reaches beyond documentation. It affects the evidentiary basis of supervision because any assessment of capital, conduct, licensing status, or cross-entity exposure requires data that remain attributable, reconcilable, and open to audit across the full monitoring perimeter. Key structural weaknesses in the current model of working with federal and departmental data are summarized in Table 1.

**Table 1:** Key structural weaknesses in the current model of working with federal and departmental data

| Structural weakness | Description | Supervisory effect |
|---|---|---|
| Institutional fragmentation | Data are split across agencies, platforms, and legal domains | Partial entity views, weak monitoring continuity, difficult risk profiling |
| Semantic inconsistency | Different vocabularies, schemas, and metadata define similar records differently | Lower interpretability, weaker query precision, reduced cross-system coherence |
| Data quality defects | Records are inconsistent, incomplete, or heavily adjusted manually | Unreliable aggregation, slower reporting, weaker analytical confidence |
| Limited data lineage | Data origin, transformation, and intervention history are not fully traceable | Harder auditability, weaker evidentiary basis, higher supervisory risk |

A unified state monitoring perimeter can be understood as an institutional and technological environment in which heterogeneous federal, departmental, and supervisory data are brought into one governed observation space. Within this space, data from licensing systems, reporting channels, document repositories, payment platforms, and external regulatory sources are integrated into a single monitoring logic. The perimeter does not imply the physical merger of all databases into one storage layer. It defines a common control framework in which data remain linked through shared rules, common identifiers, coordinated access, and a stable interpretation of regulated entities, events, statuses, and obligations. Such a perimeter converts dispersed records into an integrated supervisory field that supports state oversight across the full lifecycle of regulatory interaction.

The purpose of this perimeter is to create a trusted foundation for public monitoring, supervisory analytics, compliance assessment, and risk detection. Its functions include consolidating fragmented signals about supervised institutions, reducing semantic conflicts between source systems, preserving traceability across data flows, and supporting audit-ready decision-making. Standardization forms the first construction principle because a monitoring perimeter requires harmonized definitions, common validation rules, and stable data structures. Centralization is the second principle because strategic standards, control policies, and core metadata must be managed at a single governing level. Federated governance forms the third principle because domain units retain responsibility for the quality and meaning of their own data assets within a shared regulatory architecture. Security forms the fourth principle because supervisory data include sensitive institutional information, access restrictions, and obligations linked to privacy, cybersecurity, and lawful use.

A unified data model gives this perimeter its internal coherence. It establishes the formal representation of supervised entities, legal relationships, reporting objects, control actions, and risk indicators across all connected sources. Without such a model, integration remains a sequence of local mappings that break under institutional variation and policy change. Normatively aligned metadata extend this model into the realm of governance by defining the business meaning, ownership, classification, lineage, permissible use, and validation status for each critical data element. Metadata operate as a regulatory instrument that stabilizes interpretation across systems and actors. The monitoring perimeter gains durability when the data model and metadata structure are treated as common state infrastructure that binds technical integration to legal accountability and supervisory purpose.

The target architecture for integrating federal databases into a unified state monitoring perimeter begins at the source level, where data originates in heterogeneous operational environments with different legal functions, ownership models, update cycles, and structural formats. These sources include federal registries and reporting systems, licensing and payment systems, document systems, case management systems, and departmental or programmatic data sources. Each of these sources has its own administrative logic and data quality constraints. The architecture required must both represent these constraints and support common monitoring tasks. Thus, the source system is responsible for the primary capture and legal meaning of a core record, while the integration architecture is responsible for controlled extraction, alignment, and reuse.

The integration layer that interconnects these distributed ecosystems can be based on ETL and ELT, API and asynchronous data transfer enabling batch and near real-time and event-based pipelines. This layer performs technical acquisition, schema mapping, format conversion, scheduling, and transport control. It also establishes

the transition point at which isolated operational data enters a governed-state monitoring flow. The design of this layer must support repeatability, version control, recoverability, and observability of each data movement process. Integration pipelines should therefore record source references, timestamps, transformation logic, and processing outcomes for each critical dataset. Such a design reduces opacity in multi-source exchange and creates a stable backbone for supervisory data circulation.

The standardization and harmonization layer transforms incoming data into a common semantic structure that can support cross-system interpretation. At this stage, source-specific labels, codes, entity representations, and event descriptions are aligned with a unified conceptual model. Harmonization includes reference data alignment, identifier matching, schema reconciliation, normalization of statuses and classifications, and resolution of semantic conflicts between institutional vocabularies. This layer has strategic value because a state-monitoring perimeter cannot function coherently when equivalent entities carry divergent meanings across systems. With a common, shared data language across the perimeter, records referencing the same institution, obligation, document, transaction, or supervisory event can be uniquely identified and easily linked.

The data quality control layer ensures the data integrated are a good basis for supervisory decision-making, regulatory reporting and risk analysis. It involves the dimensions of completeness, validity, consistency, timeliness, uniqueness and regulatory compliance. The architecture's interpretive core should have automatable rules for validation, exception processing, issue escalation, remediation, and the provision of feedback to the owners of the source and the data stewards. It should provide the metadata management, business glossary, and data lineage layers. Metadata defines the meaning, ownership, classification, and permitted use of each critical data element. The business glossary stabilizes domain language across organizational units. Data lineage links each analytical output to its source records, transformation stages, and control actions. TTogether these mechanisms create traceability which supports audit, review, and institutional accountability.

The top of the stack is a Single Source of Truth for state monitoring, which lets regulators use the governed data for reporting and dashboarding, longitudinal analyzes, anomaly detection, and assessments of institution-level risk, all from a unified analytical stack. It exposes trusted datasets and controlled analytical views. Authentication, authorization, session monitoring, encryption, and gateway-based policy enforcement are integrated as part of a single architecture. In such a model, the analytical platform becomes the operational center of the monitoring perimeter, while access and security mechanisms protect the integrity, confidentiality, and legal defensibility of the supervisory data space. The State Monitoring Architecture is visualized in Figure 1.

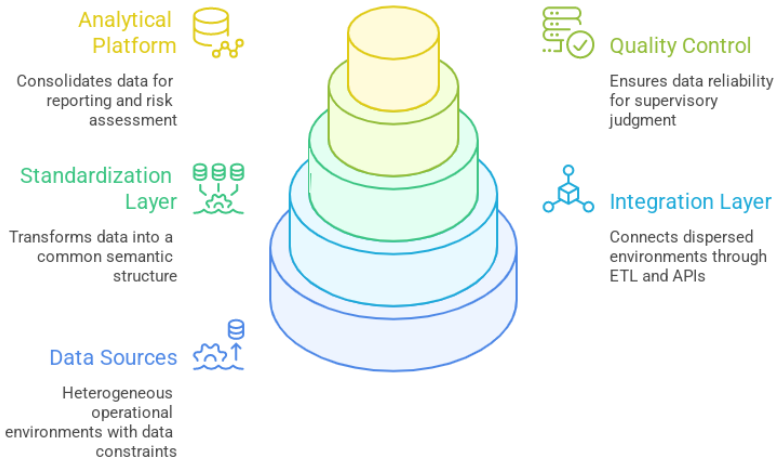


**Figure 1:** State Monitoring Architecture

The organizational model for Data Governance for a unified state monitoring perimeter must be based on a federated structure with centralized standards. Such a model reflects the institutional reality of public supervision, where data are produced, interpreted, and maintained across multiple regulatory domains with distinct mandates and operational routines. A centralized governance core is needed to define common policies, shared terminology, metadata requirements, quality rules, access principles, and control procedures. At the same time, domain units must retain responsibility for the substance of the data they create and use in supervisory practice. This combination allows the governance model to preserve domain knowledge while maintaining a single regulatory logic across the integrated data environment.

Within this structure, the Data Owner is accountable for the meaning, use, and control of a data domain. The Data Owner determines which data elements carry regulatory significance, which rules define their acceptable state, and which business processes depend on them. The Data Steward operates at the level of daily governance practice. This function is responsible for metadata maintenance, monitoring and managing data quality exceptions, glossary maintenance, as well as the collaboration between the operational data governance teams and the Governance function at the center of the model, known as the Governance Office. The Governance Office sets the standards, governs the forums, arbitrates cross-domain issues, monitors conformity to policy, and ensures the sustainability of governance decisions in the procedures. These roles form a stable chain of authority that links policy, operations, and supervisory use.

The distribution of responsibility between domain units and the central function determines whether the monitoring perimeter can produce regulatory consistency across heterogeneous sources and workflows. Domain teams possess knowledge of the institutional context, legal nuances, and operational details. The central function holds responsibility for coherence, comparability, and control across the full perimeter. Governance serves as the mechanism that binds these layers into a single supervisory system. It aligns local data practices with common state requirements and turns fragmented administrative records into evidence that supports integrated monitoring, audit, and risk-based decision-making. Data Governance functions as an institutional architecture of consistency that allows the state to observe complex supervised environments through a single, managed, and traceable data

space. Core roles in the federated Data Governance model for a unified state monitoring perimeter are summarized in Table 2.

**Table 2:** Core roles in the federated Data Governance model for a unified state monitoring perimeter

| Governance element | Main responsibility | Governance contribution |
|---|---|---|
| Central governance core | Defines common policies, standards, and controls | Ensures coherence across the monitoring perimeter |
| Domain units | Manage the substance and use of domain data | Preserve domain expertise and regulatory context |
| Data Owner | Accountable for meaning, quality rules, and business use of data | Establishes formal domain responsibility |
| Data Steward | Maintains metadata, monitors quality, and resolves issues | Supports daily governance operations |
| Governance Office | Coordinates standards, compliance, and conflict resolution | Connects policy with operational practice |

In a banking regulatory environment, the practical value of this architecture emerges when integrated data support a single supervisory view across licensing, compliance reviews, institutional monitoring, and risk analysis. Data from filings, examination records, payment activity, internal case workflows, and external federal sources can be assembled into a unified profile of each regulated entity. This profile allows supervisory teams to trace links between legal status, reporting behavior, operational events, and early indicators of instability. Integrated datasets also improve the quality of supervisory analytics by making patterns that remain hidden within separate systems visible once records are aligned through common identifiers and shared business definitions. The same architecture strengthens licensing processes by connecting application data, documentary evidence, historical actions, and control results into a single, governed flow. Compliance validation gains a stronger evidentiary base because obligations, submissions, exceptions, and remediation actions can be assessed inside a common data space. Risk monitoring also benefits from this design because supervisors can compare institutions within a single normalized analytical perimeter to detect exposure clusters, report anomalies, and identify recurring control failures.

The role of the Single Source of Truth becomes decisive in supervisory decision-making. A regulator requires more than data availability. A regulator requires a trusted analytical environment in which critical facts have stable meaning, known origins, controlled transformation paths, and documented quality states. The SSOT provides this environment by transforming fragmented operational records into curated supervisory evidence. It supports decisions on licensing, escalation, remediation, and ongoing oversight because the same governed data foundation is used across functions and domains. Modernization initiatives fit into this logic when they reduce fragmentation at entry points, in internal processing, and in analytical consumption. Unified digital portals can standardize interactions with regulated entities and reduce variation in submitted information. Internal system renewal can

improve record consistency and workflow visibility. Cloud platforms can create a scalable space for storage, harmonization, and analytical reuse. These initiatives add value when embedded in governance rules that preserve traceability, access controls, and data accountability across the monitoring perimeter.

Implementing this architecture requires an institutional pathway that accounts for technological diversity and organizational variation across the supervisory environment. Legacy systems remain part of this landscape because they preserve critical records, domain history, and procedural continuity, making their integration a high-priority design task. Organizational differentiation also shapes implementation, since domain units operate through distinct vocabularies, control practices, and operational priorities. For this reason, the transition toward a unified monitoring perimeter is most sustainable when it follows a phased and maturity-based model. Such an approach allows the regulator to begin with high-value supervisory domains, establish common standards through controlled implementation, and extend the perimeter as stewardship capacity, metadata discipline, and quality controls become embedded in routine regulatory practice.

The synthesis of the reviewed material reveals a coherent gradient running through the four structural weaknesses and the layered architectural response. Institutional fragmentation creates the outer boundary of the problem space and produces partial entity views, semantic inconsistency operates at the interpretive layer and fractures shared meaning across vocabularies, data quality defects propagate through reconciliation chains and degrade the evidentiary value of supervisory outputs, and limited lineage removes the analytical thread that ties reported indicators back to original source records. The proposed layered architecture maps onto these weaknesses through a chain of compensating mechanisms.

The integration layer absorbs structural fragmentation by establishing controlled extraction pathways, the harmonization layer dissolves semantic conflicts through unified identifiers and reference data alignment, the quality control layer addresses defect propagation through validated rule sets and exception handling, and the metadata and lineage layer reconstructs traceability across the full transformation chain. The Single Source of Truth functions as the convergence point of these mechanisms and translates governed data into supervisory evidence usable for licensing, prudential reporting, and risk aggregation. The discussion further indicates that the institutional weight of the architecture rests on the federated governance model with centralized standards. This arrangement reconciles domain expertise with cross-perimeter coherence and converts the technical stack into an accountability structure where each role anchors a specific regulatory function within the supervisory chain.

## 4. Conclusion

The study demonstrates that integrating federal databases into a unified state monitoring perimeter must be approached as an institutional architecture that binds governance, metadata, data quality, lineage, and platform design into a single supervisory system. Fragmentation across agencies, platforms, legal domains, and technical environments undermines continuity of monitoring, weakens semantic coherence, and limits the construction of institution-level risk profiles from dispersed operational traces. In this context, the article shows that the DAMA-DMBOK provides a suitable conceptual foundation for regulatory integration by linking stewardship, architecture, quality management, metadata control, and policy coordination within a single management logic. A unified

monitoring perimeter emerges here as a governed observation space that connects heterogeneous federal, departmental, and supervisory data through shared rules, common identifiers, coordinated access, and stable interpretation of regulated entities, events, statuses, and obligations.

The results indicate that the target architecture of such a perimeter requires a layered design in which source systems preserve responsibility for primary data capture and legal meaning, while the integration environment ensures controlled extraction, harmonization, traceability, and analytical reuse. The integration layer, the standardization and harmonization layer, the quality control layer, and the unified analytical platform together form the structural basis of a Single Source of Truth for state monitoring. This architecture gains scientific weight by treating supervisory data pipelines as evidence systems whose outputs must remain attributable, reconcilable, and open to audit across the full monitoring perimeter. In that setting, metadata cease to be an auxiliary technical resource and assume the role of a regulatory instrument that stabilizes meaning, ownership, classification, lineage, permissible use, and validation status for each critical data element. The same logic applies to data quality and lineage, since prudential judgment, compliance assessment, remediation tracking, and risk aggregation depend on records whose origins, transformation paths, and control histories can be reconstructed without ambiguity.

The article also establishes that the organizational form of this architecture must rely on federated Data Governance with centralized standards. Such a model reflects the distribution of responsibility across public supervision, where domain units retain substantive knowledge of their own data while a central governance core secures coherence, comparability, and control across the perimeter. The defined roles of Data Owner, Data Steward, domain units, and Governance Office create a chain of accountability that links policy, operational practice, and supervisory use. On this basis, the proposed architecture delivers practical value for banking regulation by enabling a unified supervisory view across licensing, compliance validation, institutional monitoring, and risk analysis. Its implementation requires phased development shaped by stewardship capacity, metadata discipline, and quality controls embedded in routine practice. The article connects public-sector data integration to regulatory oversight theory. In practice, it provides an architectural foundation for state bodies seeking audit-ready, standardized, and analytically usable data within a single, managed monitoring space.

The scope of the present study is shaped by several boundary conditions that situate the obtained conclusions within a defined analytical perimeter. The argument rests on twelve sources spanning peer-reviewed scholarship, supervisory guidance, and conceptual work on public-sector data governance, which establishes a coherent thematic field and opens a natural extension toward jurisdiction-specific regulatory case archives, institutional implementation reports, and longitudinal observations of supervisory data programs. The conceptual-analytical design treats the proposed architecture as a normative-structural model anchored in DAMA-DMBOK principles, and the calibration of governance roles, lineage requirements, and quality rules against measured supervisory pipelines marks a productive direction for subsequent inquiry.

The interpretation of structural weaknesses operates through qualitative criteria distilled from the literature, and the construction of quantitative metrics tied to reconciliation latency, manual adjustment frequency, and lineage completeness represents a further analytical horizon. The architectural framework treats institutional fragmentation, semantic inconsistency, quality defects, and limited lineage as a unified set of structural conditions,

and the disaggregation of these dimensions through dedicated empirical instruments offers a path toward more granular decision support for the design of unified state monitoring perimeters across diverse regulatory environments.

**References**

[1] M. Janssen, P. Brous, E. Estevez, L. S. Barbosa, and T. Janowski, "Data governance: Organizing data for trustworthy Artificial Intelligence," Government Information Quarterly, vol. 37, no. 3, p. 101493, Jul. 2020, doi: https://doi.org/10.1016/j.giq.2020.101493.

[2] N. Bagherifam, S. Naghdi, V. Ahmadian, A. Fazlzadeh, and M. Baghalzadeh Shishehgarkhaneh, "Digital Regulatory Governance: The Role of RegTech and SupTech in Transforming Financial Oversight and Administrative Capacity," International Journal of Financial Studies, vol. 13, no. 4, p. 217, Nov. 2025, doi: https://doi.org/10.3390/ijfs13040217.

[3] F. Bartolomucci, E. Ramalli, and V. M. Urbano, "Integrated governance in data ecosystems: A conceptual framework consolidating collaborative and data governance," Government Information Quarterly, vol. 43, no. 1, p. 102107, Mar. 2026, doi: https://doi.org/10.1016/j.giq.2026.102107.

[4] B. M. V. Bernardo, H. S. Mamede, J. M. P. Barroso, and V. M. P. D. dos Santos, "Data governance & quality management—Innovation and breakthroughs across different fields," Journal of Innovation & Knowledge, vol. 9, no. 4, p. 100598, Oct. 2024, doi: https://doi.org/10.1016/j.jik.2024.100598.

[5] V. Ribeiro, J. Barata, and P. R. da Cunha, "Modeling inter-organizational business process governance in the age of collaborative networks," Electronic Markets, vol. 34, p. 51, Oct. 2024, doi: https://doi.org/10.1007/s12525-024-00730-2.

[6] M. Abbasi, P. Váz, J. Silva, F. Cardoso, F. Sá, and P. Martins, "Unified Data Governance in Heterogeneous Database Environments: An API-Driven Architecture for Multi-Platform Policy Enforcement," Data, vol. 11, no. 3, p. 54, Mar. 2026, doi: https://doi.org/10.3390/data11030054.

[7] "Guide on effective risk data aggregation and risk reporting," ECB Banking Supervision, 2024. Available: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guides240503_riskreporting.en.pdf

[8] C. Berg, "Interoperability," Internet Policy Review, vol. 13, no. 2, Apr. 2024, doi: https://doi.org/10.14763/2024.2.1749.

[9] A. C. Widlak and R. Peeters, "A theory of the infrastructure-level bureaucracy: Understanding the consequences of data-exchange for procedural justice, organizational decision-making, and data itself," Government Information Quarterly, vol. 42, no. 2, p. 102021, Mar. 2025, doi:

https://doi.org/10.1016/j.giq.2025.102021.

[10] E. Kalampokis, N. Karacapilidis, D. Tsakalidis, and K. Tarabanis, “Understanding the Use of Emerging Technologies in the Public Sector: A Review of Horizon 2020 Projects,” Digital Government: Research and Practice, vol. 4, no. 1, pp. 1–28, Jan. 2023, doi: https://doi.org/10.1145/3580603.

[11] I. Mitzutani, G. Ramanathan, and S. Mayer, “Semantic data integration with DevOps to support engineering process of intelligent building automation systems,” BuildSys ’21: Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, pp. 294–297, Nov. 2021, doi: https://doi.org/10.1145/3486611.3492413.

[12] K. Gierend, “Provenance Information for Biomedical Data and Workflows: Scoping Review,” Journal of Medical Internet Research, vol. 26, p. e51297, Aug. 2024, doi: https://doi.org/10.2196/51297.