International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Balancing Autonomy and Consistency: Inter-Agency Data Synchronization Architectures in Government Systems

Abhinav Taduka

Abstract: Inter-agency data synchronization is one of the key architectural and governance challenges of digital government, as many government services, laws, regulations, and complex policy outcomes depend on data shared across agency boundaries in a timely and reliable manner. Most existing government information systems were designed to operate in silos, making such synchronization inherently difficult. This tension between the autonomy of each agency and the consistency of the overall system cannot be resolved through technology alone; it must be addressed through co-design in governance and architecture, in which governance requirements and policy goals are translated into enforceable technical controls within synchronization mechanisms. Drawing on distributed systems theory, governance in public administration, and deployment patterns within and across jurisdictions, a taxonomy is derived that identifies centralized, federated, and hybrid synchronization mechanisms in terms of consistency guarantees, fault tolerance, institutional feasibility, and scalability, while recognizing governance frameworks, semantic interoperability mechanisms, and policy enforcement infrastructures as intrinsic technical preconditions rather than organizational overlays. This work introduces the Policy-Aware Federated Synchronization Architecture (PAFSA), a four-layer framework that enables consistent, auditable, and policy-compliant data exchange among autonomous government agencies. Unlike traditional integration architectures, PAFSA treats governance as an enforceable technical layer, integrates semantic mediation as a first-class concern, and supports event-driven synchronization to balance autonomy and consistency in distributed institutional environments. PAFSA comprises agency systems, integration and synchronization, semantic interoperability, and governance enforcement layers, and serves as a conceptual baseline for policy-aware federated synchronization systems that enable distributed governance alongside coordinated, reliable, and auditable data exchange across the complexity of modern public-sector environments.

Keywords: *Inter-Agency Data Synchronization, Federated Architecture, Governance Interoperability, Semantic Mediation, Policy Enforcement Infrastructure*

Section 1: The Interoperability Imperative of Modern Digital Government

Digital technologies have deeply and rapidly affected how public administrations deliver services and use information over the last 30 years. In the early days of e-government, applications were agency-centric, vertically integrated applications that prioritized mission-based functionality over inter-agency coordination and interoperability. All of this meant that interoperability was not a design consideration but rather an afterthought.

Experiences in responding to the pandemic, coordinating welfare, and meeting cross-agency regulatory obligations in real-time have made plain the limits of siloed data governance in a digital world, where consistent dynamic data exchange across agencies is critical. The scale of the challenge can be indicated by the fact that surveys within scientific and governmental data systems suggest that as many as 28% of scientists have been unable to confirm published results due to data. In a national survey in 2006, a majority of junior scientists reported having experienced data withholding from their collaborators [1]—signifying that the problem of data withholding is more than just a technical issue.

Institutional factors further complicate integration. Public organizations are bound by various forms of legislative mandate and carry different legal liabilities, which makes them difficult to place under the control of a central IT body. Integration thus proceeds through deference to distributed authority while achieving a coherent result.

The stakes are also seen in clinical and research data: of 500,000 clinical trials published on MEDLINE by 2015, only 154 out of 177 approved data-sharing requests have performed a secondary analysis on 3 major data-sharing platforms. This makes it obvious how little the scientific

Independent Researcher, USA

community is reusing shared data despite it being formally available [2]. Of the 37 identified reanalyses, 13 (35%) found materially different results. This has implications for the evidence base that informs subsequent policy and service delivery, as only selective data are shared [2].

Lately, modern trends such as APIs, the cloud, and event-driven architectures have made such synchronization tools more common, but a weak organizational and legal infrastructure still impedes common usage across the agencies. The gap between technical capability and readiness for governance is a central challenge for digital government architects and policymakers [1][2]. Addressing this gap requires not only better tools, but a coherent architectural framework in which governance and synchronization are co-designed from the outset — a need that motivates the PAFSA framework introduced in this work.

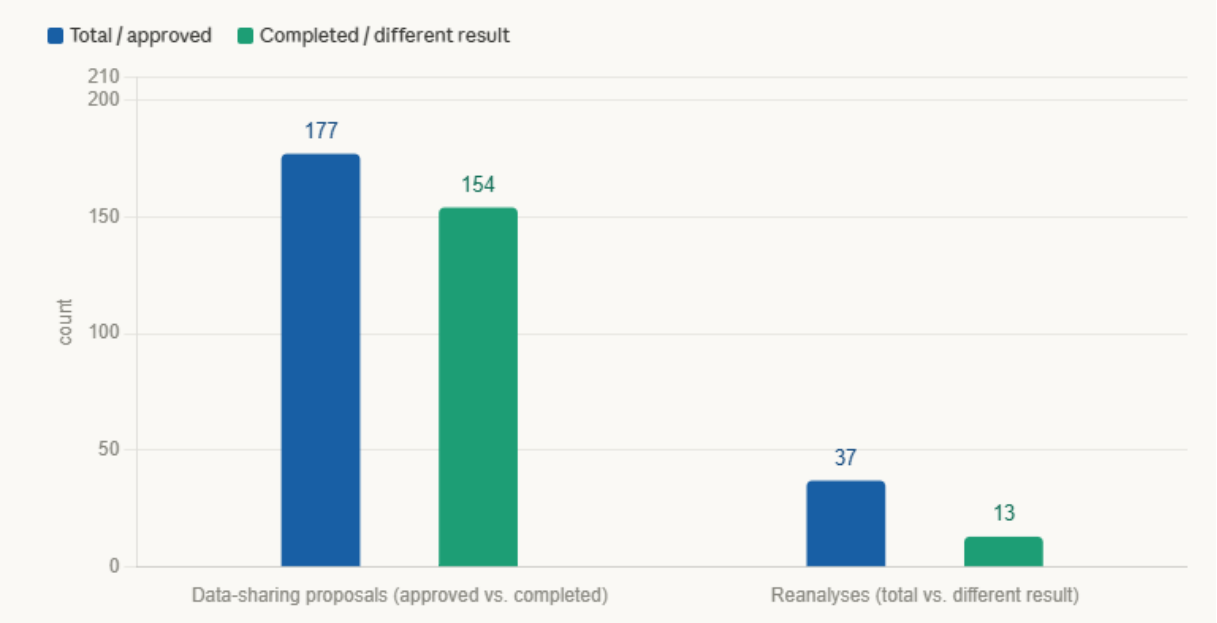

Figure 1: Data Sharing Gaps in Scientific and Government Systems [1, 2]

Section 2: Theoretical Foundations: Distributed Systems and Governance Principles

Inter-agency data synchronization is a distributed systems problem with boundaries defined by institutions, law, and politics. The two academic disciplines that provide the theoretical foundations are computer science and public administration.

In distributed computing, the CAP theorem states that it is impossible for a distributed computer system to simultaneously provide consistency, availability, and partition tolerance. In the context of government, where partitions in the network are likely, and availability is paramount, this becomes a trade-off. A comparison of 20 GIDEPs found that a centralized data-exchange and decentralized storage configuration was used by 12 of the 20 countries studied. This configuration is considered an architectural trade-off to balance the consistency guarantees across the GIDEP and the autonomy of its constituent agencies [3]. Fully centralized configurations, where both the data exchange and storage are centralized, were used by one of the 20 countries studied and are not the norm for GIDEPs [3]. In five of the 20 countries studied, no ecosystem of data exchange and interoperability could be identified, indicating the uneven maturity of synchronization infrastructure worldwide [3]. Applications vary in their acceptability of inconsistency. Benefit eligibility systems and law enforcement grant little room for error, requiring high levels of correctness and auditability, whereas public information systems may tolerate short-term inconsistency. As such, it is a spectrum to be calibrated with respect to the sensitivity of the policy.

Governance theory shows how various countries' data governance frameworks have driven the emergence of federated, negotiated data governance constructs. Globally, digitally enabled services increased to USD 3.8 trillion (four times the level in 2005), and regulatory fragmentation sharply increased. This is evidenced in the public

consultation on cross-border data flows, where support for free-flowing data increased from 49% to 67% when respondents learned about the benefits, such as medical technological innovation. This suggests that institutional legitimacy and the perceived value of what data can deliver influence support for cross-border data flows.

According to socio-technical systems theory, conflicts between design and organizational incentives account for more synchronization failures than faults in technical design. Architectures that conflict with structures intrinsic to legal authority will be resisted, irrespective of their technical merits. Effective inter-agency synchronization, therefore, depends on making decisions that are technically and institutionally sound simultaneously. This dual requirement — technical rigor combined with institutional legitimacy — forms the theoretical grounding for PAFSA, which is designed to resolve both dimensions through co-designed governance and architecture rather than treating either as an afterthought.

Condition	Support Level	Change
Before benefit disclosure	49%	Baseline
After benefit disclosure (e.g., medical innovation)	67%	+18 percentage points

Table 2: Public Support for Cross-Border Data Flows Before and After Benefit Disclosure [4]

Section 3: Architectural Models for Inter-Agency Data Synchronization

There are various architectural models for data synchronization between agencies. These differ based on how data is synchronized and how accountability, authority, and operational risk are distributed. Many can be classified as centralized, federated, or hybrid synchronization architectures. Each has its own technical and institutional characteristics, and each presents trade-offs that inform the design decisions embedded in PAFSA.

Centralized Synchronization Architectures involve a central data warehouse, a master data management system, or a national registry, to which all other systems provide or retrieve data following agreed schemas, validation, and update rules. Synchronization occurs through scheduled batch updates via bulk ingestion, real-time updates via APIs, or extract, transform, load (ETL) pipelines. It maintains data quality, access control, and versioning while the agencies serve as either providers or consumers of data. Primary advantages include strong consistency and a straightforward audit process. Primary disadvantages are single points of failure and institutional resistance. A working paper on information-sharing frameworks between public agencies identifies 4 different types of schema-level conflicts — name conflicts, primary key conflicts, generalization conflicts, and schematic inconsistencies — that must be resolved for centralized integration to function. This is a non-trivial cost of achieving consistency [5].

Federated Synchronization Architectures support agency ownership of data while minimizing the data-sharing workload through a common interface. Each agency is responsible for its own authoritative data stores, pushing out changes through event streams, message brokers, versioned APIs, or change data capture. Although the architecture is designed to allow for eventual consistency and incremental adoption, semantic mismatch continues to be a persistent problem. Testing of the performance of XML-based inter-agency search agents showed that free text methods had lower recall and precision than structured techniques for both database and XML repositories, confirming the negative impact of semantic heterogeneity on quality in federated systems [5].

Hybrid Synchronization Architectures utilize both styles by separating governance and coordination from data ownership, providing selective consistency through layered, application-level control. The security dimensions of hybrid architectures are highlighted by the VPNFilter malware, which infected over 500,000 devices worldwide. The Department of Homeland Security (DHS) identifies 16 critical infrastructure sectors, including emergency services; the attack surface of hybrid government systems is therefore wide and spreads across different sectors. The 2016 Telephony Denial-of-Service attack on 911 services across 3 states illustrates how a single attack can affect multiple related agencies simultaneously [6].

Dimension	Centralized	Federated	Hybrid
Agency Autonomy	Low	High	Medium–High
Consistency Guarantees	Strong	Eventual	Selective
Scalability	Moderate	High	High
Fault Tolerance	Low	High	High
Governance Complexity	Low	High	Very High
Political Feasibility	Low	High	Medium–High

Table 1: Comparative Analysis of Inter-Agency Data Synchronization Architectures [5, 6]

No single model optimizes all dimensions simultaneously. Architectural suitability is contingent on domain sensitivity, legal authority, and the institutional maturity of participating agencies. This comparative analysis motivates a framework — PAFSA — that retains the autonomy-preserving strengths of federated architectures while embedding the consistency and governance enforceability that centralized models provide, without inheriting their institutional and technical liabilities.

Section 4: Governance and Policy Enforcement as Technical Infrastructure

A persistent misconception in technical circles treats governance as an organizational overlay — a set of rules operating alongside technical systems rather than embedded within them. In inter-agency data synchronization, this separation is both analytically misleading and practically harmful. Governance and architecture must be co-designed, with policy requirements translated into enforceable technical controls within synchronization workflows. This principle is foundational to PAFSA: governance is not a constraint applied to the architecture — it is a layer of the architecture itself.

Semantic interoperability is a foundational governance function. Without it, synchronization amplifies inconsistencies rather than resolving them. Agencies frequently employ identical terminology for different concepts or different terminology for the same concept. Metadata standards — including controlled vocabularies, standard data models, and lineage metadata — establish the shared interpretive context that makes exchanged data meaningful across organizational boundaries. This challenge is not hypothetical: cross-border identity ecosystems have demonstrated that even mature frameworks such as eIDAS, implemented in 2014 and spanning 28 EU member states, struggle with incompatible certification authorities and heterogeneous authentication mechanisms that produce silent failures at integration points [7].

Schema registries serve as the institutional memory of governance agreements, maintaining data definitions, ownership information, and compatibility crosswalks. Their absence is a common factor in synchronization failures, as agencies make unilateral schema changes that break downstream consumers. In X-Road-based architectures, the default OCSP certificate validity fetch interval is set to 20 minutes — a parameter that directly bounds the window of exposure during which a revoked credential may still be accepted, illustrating how a single governance-adjacent configuration value carries real security consequences [8].

Policy-driven synchronization embeds governance within technical workflows, reducing dependence on manual compliance. Critical components include access control, data ownership assignments, versioning protocols, and propagation scope limitations. Legal constraints impose additional requirements: privacy laws mandate purpose limitation, auditability demands full provenance logging, and data minimization requires payloads to carry only necessary attributes. Scoping these requirements systematically — as demonstrated by frameworks identifying six discrete identity management processes against which interoperability criteria must be baselined — reduces the risk of gaps that post-hoc auditing cannot remediate [7].

Governance structures require clear institutional roles: data owners, stewards, policymakers, and technical operators. Their effectiveness depends not on formal structure alone, but on the degree to which decisions are technically enforceable — a principle validated by proof-of-concept implementations that satisfy defined design goals only under explicit, machine-verifiable governance assumptions [8]. PAFSA operationalizes this principle by treating the Governance and Policy Enforcement Layer not as a supervisory

mechanism over the architecture, but as a coequal functional layer whose outputs directly condition the behavior of the synchronization infrastructure beneath it.

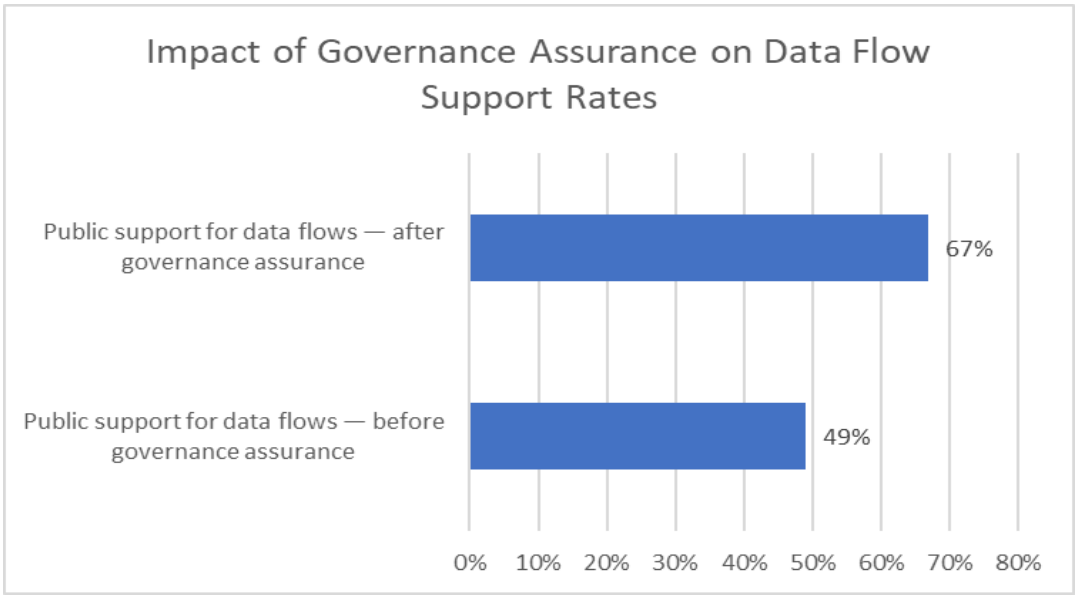

Figure 2: Public Support for Cross-Border Data Flows Before and After Governance Assurance [8,9]

Section 5: Implementation patterns, failures, and lessons from practice

The design premises of PAFSA are best evaluated against actual government practice. An examination of cross-jurisdictional synchronization efforts in identity management, public benefits determination, financial oversight, and regulatory reporting reveals consistent characteristics, even where the specific technologies and administrative arrangements differ.

Key factors for successful implementation include the following. Event-driven global state propagation consistently outperforms periodic bulk updates. Publishing deltas as individual timestamped events with source metadata allows agencies to achieve near real-time updates and loose coupling between systems. It also produces an audit trail mapping each change to the originating source, operator, and policy context — as required by government sectors concerned with data provenance accountability and compliance [9]. This pattern is directly reflected in PAFSA's Integration and Synchronization Layer, which centers event-driven mechanisms rather than treating them as optional enhancements.

Another common success factor is attribute-level data ownership assignment. Having a clear authoritative source for each attribute prevents synchronization conflicts and supports selective consistency models by providing strong consistency only to policy-relevant attributes. With initial deployments of Estonia's X-Road, which plans for more than a thousand information systems and databases, the government estimates that standardized data exchange infrastructure will save 30 percent in administrative and development costs [9], demonstrating that governance investment carries measurable operational returns.

Characteristics of successful approaches also include governance-enabled technical practices such as schema registries, API layer policy enforcement points, automated validation, and detailed audit logging — shifting governance from a controlling constraint to an operational function. This shift achieves better compliance outcomes than documentation-focused governance approaches and is precisely the model PAFSA formalizes through its Governance and Policy Enforcement Layer.

Failure modes include centralized control without adequate legal authority, semantic misalignment, and low enforcement strength. The consequences extend beyond administrative inefficiency: in a study of the USA, privacy concerns caused 586,000 US citizens to postpone cancer treatment, with an annual cost of $1.6 billion [10]. These findings demonstrate that governance failures in data-sharing mechanisms produce measurable harm. Additional evidence confirms that governance guarantees increase public support for cross-border flows from 49% to 67%, confirming that trust is both a necessary precondition and a positive outcome of well-designed synchronization regimes [10].

Four lessons emerge from these implementation experiences: synchronization requires continual attention; governance must be technically enforceable rather than merely documented; phased implementation reduces institutional risk; and trust is simultaneously a precondition for and a product of effective synchronization. Each of these lessons is directly instantiated in PAFSA's architectural commitments.

Section 6: A Reference Architecture for Policy-Aware Federated Synchronization

This work introduces the **Policy-Aware Federated Synchronization Architecture (PAFSA)**, a four-layer architectural framework designed to enable consistent, auditable, and policy-compliant data exchange among autonomous government agencies. Unlike traditional integration architectures that treat governance as an external constraint or organizational overlay, PAFSA treats governance as an enforceable technical layer, integrates semantic mediation as a first-class concern, and supports event-driven synchronization as the primary mechanism for balancing autonomy and consistency in distributed institutional environments. PAFSA is intended as a conceptual and structural reference that can be instantiated across domains, agencies, and levels of government — not as a prescriptive implementation template. Its four layers each address a distinct dimension of synchronization functionality, and together they form a coherent, governance-native architecture for public-sector data exchange.

The **Agency Systems Layer** forms the foundation of PAFSA. Information systems at this layer are developed and maintained independently by individual agencies. Agencies retain full control over their data stores, schemas, and internal business processes, and they publish data via standardized interfaces. Attribute-level authoritative ownership is a defining feature of this layer: each agency specifies which system is responsible for creating or updating which data element. This reduces synchronization conflicts, clarifies data provenance, and avoids imposing a single canonical schema on agencies' internal models — directly preserving institutional autonomy while enabling structured participation in shared data flows.

The **Integration and Synchronization Layer** is the technical backbone for inter-agency data exchange within PAFSA. It includes event brokers, API gateways, change data capture technology for legacy systems, and message queues for guaranteed asynchronous delivery. Events in this layer are self-describing: they carry metadata identifying the originating agency, the version of the information being transferred, the time of event creation, and the policy constraints under which the event is propagated. These properties make events auditable by construction — essential in critical infrastructures spanning up to 16 domains that share dependencies and could suffer cascading failure [11]. The event-driven design eliminates tight coupling between agencies, enabling incremental adoption and resilient operation under partial network or system failure.

The **Semantic Mediation and Data Harmonization Layer** addresses the representation differences that arise inevitably across independently maintained agency systems. A shared metadata and schema registry stores definitions, ownership records, and usage rules, providing the institutional memory needed to detect and resolve conflicts when agencies evolve their internal models. Ontology mapping and metadata translation services enable agencies to preserve their internal semantics while ensuring that exchanged data carries shared meaning across organizational boundaries. Without this layer, synchronization propagates inconsistency rather than resolving it — a failure mode extensively documented in federated deployments [5][7].

The **Governance and Policy Enforcement Layer** is the distinguishing feature of PAFSA and the layer that gives the framework its name. Rather than treating policy compliance as a post-hoc audit function, PAFSA embeds governance directly within synchronization workflows. Attribute-based access control, consent and purpose limitation controls, and pre-propagation validation are executed as part of the synchronization process itself. Governance-as-code automates compliance checking and ensures consistent enforcement at scale. The importance of this design choice is not trivial: infrastructure-targeting malware such as VPNFilter has been known to infect more than 500,000 devices globally, confirming that weak enforcement at integration boundaries creates systemic risk [11]. Additionally, architectures that centralize certificate validation introduce bounded exposure windows — X-Road's default OCSP fetch interval of 20 minutes means revoked credentials may remain accepted during that window [12]. PAFSA addresses this by distributing revocation

status across multiple decentralized validator nodes, removing single points of failure and reducing exposure at governance-critical integration boundaries [12].

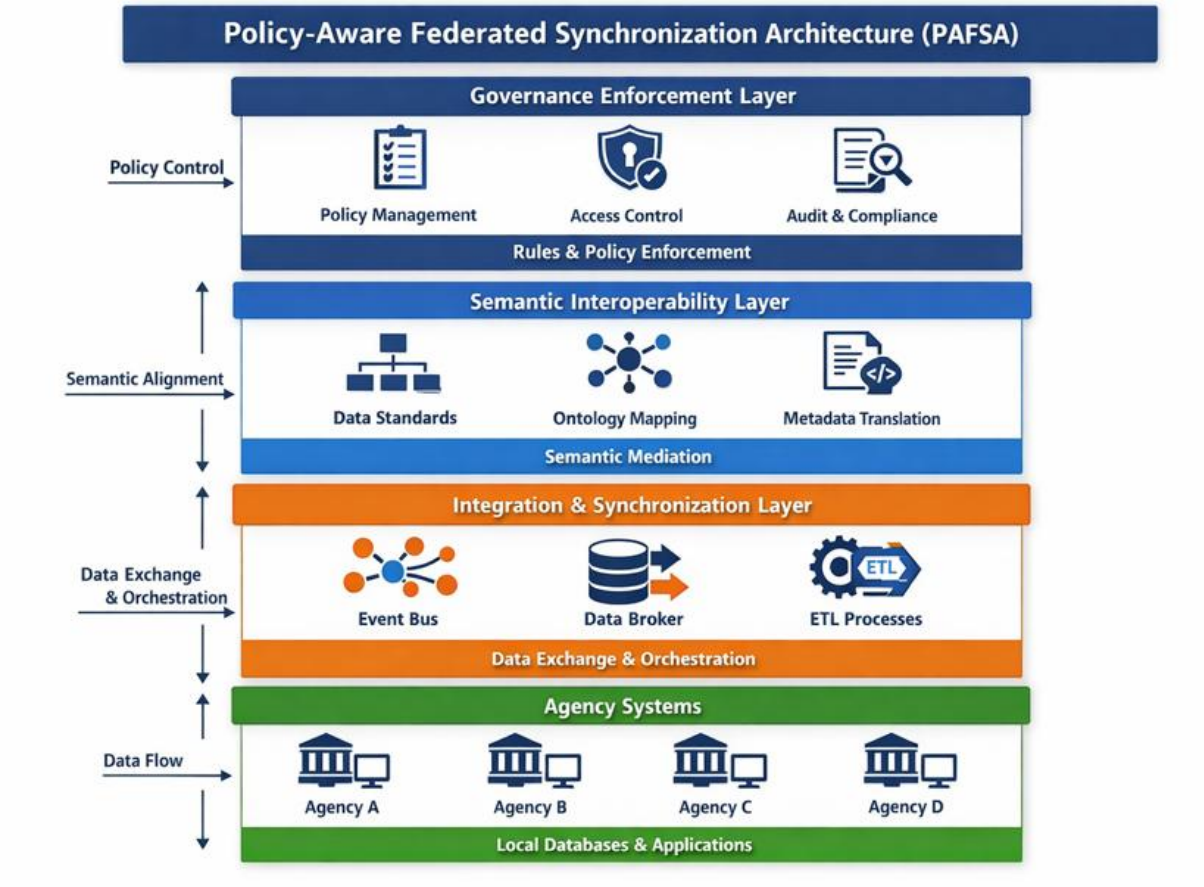

Figure 3: The PAFSA Four-Layer Reference Architecture for Policy-Aware Federated Synchronization [11, 12]

Conclusion

Inter-agency data synchronization is irreducibly both a technical and an institutional problem. Centralized architectures that make consistency practically realizable conflict with the distributed legal authority structures of public administration. Federated architectures maintain agency autonomy but require a broad semantic and governance infrastructure to prevent data degradation. Hybrid architectures are the most feasible and promising in public administration environments, but are also the most governance-complex, presuming coordinated oversight over data accessed through distributed representations.

For all these models, the evidence shows that effective governance consists of the technical enforcement of policy at the level of data exchange itself — not documentation, not audits after the fact, but machine-verifiable controls embedded in the synchronization workflow. Semantic interoperability, attribute-level schema registries, event-driven propagation, and policy enforcement as integral components of APIs and message brokers are preconditions for any synchronization system that must operate reliably at scale.

PAFSA — the Policy-Aware Federated Synchronization Architecture introduced in this work — addresses these requirements through a coherent four-layer design: agency systems, synchronization infrastructure, semantic mediation, and governance enforcement. Unlike traditional integration architectures, PAFSA treats governance as an enforceable technical layer, integrates semantic mediation as a first-class concern, and supports event-driven synchronization to balance autonomy and consistency in distributed institutional environments. It is a conceptually coherent framework applicable across a wide variety of domains, jurisdictions, and levels of institutional maturity. As digital government matures towards real-time delivery of cross-agency services, the ability to synchronize public data across agency boundaries while preserving decentralized decision rights will be key to

effective, equitable, and trusted public service. PAFSA provides the architectural reference for realizing that capability.

References

[1] Stacy Kowalczyk and Kalpana Shankar, "Data Sharing in the Sciences," [Online]. Available: https://courses.washington.edu/geog482/resource/9_Kowalczyk_DataSharingSciences.pdf

[2] Howard Bauchner, Robert M. Golub, "Data Sharing: An Ethical and Scientific Imperative," JAMA, Mar. 2016. [Online]. Available: https://jamanetwork.com/journals/jama/fullarticle/2504790

[3] Keegan McBride et al., "Digital Government Interoperability and Data Exchange Platforms: Insights from a Twenty Country Comparative Study," ACM Digital Library, 2022. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3560107.3560123

[4] James Bacchus et al., "Interoperability of Data Governance Regimes: Challenges for Digital Trade Policy," Centre for Inclusive Trade Policy, Briefing Paper 12, Apr. 2024. [Online]. Available: https://citp.ac.uk/wp-content/uploads/CITP-Briefing-Paper-12-Interopability-of-Data-Governance-Regimes-Challenges-for-Digital-Trade-Policy.pdf

[5] Akhilesh Bajaj, Sudha Ram, "A Comprehensive Framework Towards Information Sharing Between Government Agencies," International Journal of Electronic Government Research, vol. 3, no. 2, pp. 29–44, Apr.–Jun. 2007. [Online]. Available: https://www.researchgate.net/profile/Akhilesh-Bajaj/publication/220526995

[6] Nuzhat Qadir Tunio, "Enhancing the Security of Critical Infrastructure in Emergency Services," Webster University, 2025. [Online]. Available: https://www.researchgate.net/profile/Nuzhat-Tunio/publication/390112018

[7] Ayei E. Ibor et al., "Trustworthy Cross-Border Interoperable Identity System for Developing Countries," arXiv. https://arxiv.org/pdf/2310.16562

[8] Mariia Bakhtina et al., "A Decentralised Public Key Infrastructure for X-Road, " IACM Digital Library, 2023. https://doi.org/10.1145/3600160.3605092

[9] Ahto Kalja, "The X-Road Project: A Project to Modernise Estonia's National Databases," Baltic IT&T Review, 2002. http://www.ebaltics.lv/doc_upl/Kalja(2).pdf

[10] James Bacchus et al., "Interoperability of Data Governance Regimes: Challenges for Digital Trade Policy." Centre for Inclusive Trade Policy, University of Sussex, 2024. https://citp.ac.uk/wp-content/uploads/CITP-Briefing-Paper-12-Interopability-of-Data-Governance-Regimes-Challenges-for-Digital-Trade-Policy.pdf

[11] Nuzhat Qadir Tunio, "Enhancing the Security of Critical Infrastructure in Emergency Services," ResearchGate, 2025. https://www.researchgate.net/profile/Nuzhat-Tunio/publication/390112018

[12] Mariia Bakhtina et al. (2023). "A Decentralised Public Key Infrastructure for X-Road." ACM Digital Library, 2023. https://dl.acm.org/doi/pdf/10.1145/3600160.3605092