

# DATA-PROOF: Multi-Dimensional Governance-State Assurance for Enterprise Data Transformation and AI Consumption

Naresh Somara

Independent Researcher, Erie, Pennsylvania, USA

Email: [naresh.somara.us\[at\]gmail.com](mailto:naresh.somara.us[at]gmail.com)

ORCID: 0009-0008-4997-3133

**Abstract:** Enterprise data governance distributes controls across identity systems, catalogs, policy engines, lineage services, quality platforms, privacy classifications, and audit stores. Yet a consequential data operation can change governance characteristics after source access has been approved: joins may increase inference risk, reuse may violate purpose, cross-region processing may alter jurisdictional constraints, and AI retrieval may over-collect fields. This paper presents DATA-PROOF, a multi-dimensional governance-state assurance framework that evaluates operations as transitions from a pre-operation state to a predicted post-operation state. The framework integrates purpose compatibility, lineage, quality and freshness, residency, composition risk, minimization, downstream exposure, and policy-version freshness, returning ALLOW, CONSTRAIN, REVIEW, or DENY and recording the decision in a Governance State Transition Certificate (GSTC). A controlled benchmark evaluates 16 governance scenario families with baseline, ablation, certificate-integrity, and latency tests. DATA-PROOF conformed to the configured benchmark policy in 16/16 families, detected 5,000/5,000 certificate mutations, and averaged 54.55 microseconds in-process across 32,000 local executions. The API decision, persistence, and GSTC-verification path was also reproduced successfully under Docker Compose. Results establish prototype feasibility and policy coverage under synthetic semantics, not production effectiveness or legal compliance.

**Keywords:** Data governance; Derived data; Policy enforcement; AI governance; Runtime assurance

## 1. Introduction

Data governance is commonly understood as the exercise of decision rights, accountability, and control over data assets. Foundational research emphasizes governance structures and decision domains, while later reviews broaden the field to organizational scope, data scope, mechanisms, and outcomes [1], [2]. Enterprises increasingly operationalize these ideas through catalogs, stewardship workflows, policy engines, quality systems, privacy classification, data lineage, and audit tooling. Research has also moved toward automated governance in data lakes and service-based pipelines [3], [4].

Cloud data warehouses, data products, generative AI, retrieval-augmented generation, and autonomous agents can compose, transform, move, and reuse data with little human intervention. In these settings, a permitted operation can create a result whose classification, inference risk, residency, purpose compatibility, quality, or disclosure exposure differs materially from its inputs. Temporal and derived-data authorization, data-fusion policies, proof-carrying authorization, purpose-based access control, provenance, usage control, and data-flow enforcement already address important parts of this problem [5]-[12]. DATA-PROOF therefore does not claim that derived-data control or policy evidence is new.

The research question is narrower: can one operation-level assurance layer assemble multiple governance dimensions, predict the candidate post-operation state, determine whether a safe constraint can preserve task utility, and generate a compact evidence record that explains the decision? DATA-PROOF is designed to coordinate existing catalogs, policy engines, lineage services, privacy classifiers, data-quality systems, and AI gateways rather than replace them.

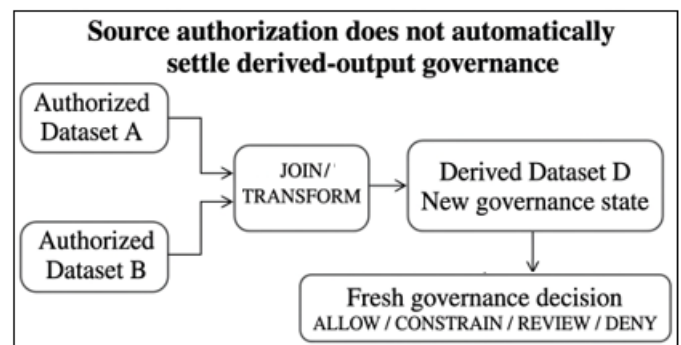


**Figure 1:** DATA-PROOF problem framing: source authorization does not settle the governance state of a derived result.

### 1.1 Research Questions and Contributions

RQ1- Coverage: Do multi-dimensional operation-level checks identify governance conditions that simplified source-only and purpose/context authorization baselines omit?

RQ2- Component contribution: Which governance dimensions materially affect policy coverage under controlled ablation?

RQ3- Evidence integrity: Can the GSTC consistently bind the evaluated state and detect post-generation payload modification?

RQ4- Overhead: What local decision-function and certificate-generation overhead does DATA-PROOF introduce in the reference prototype?

The paper makes five bounded contributions: a multi-dimensional governance-state representation; an operation-level transition procedure; a constrained-repair abstraction for masking, tokenization, field reduction, or equivalent

Volume 15 Issue 10, October 2026

Fully Refereed | Open Access | Double Blind Peer Reviewed Journal

[www.ijsr.net](http://www.ijsr.net)

policy-preserving changes; the GSTC transition witness; and a reproducible benchmark with baseline, ablation, integrity, and latency measurements. The contribution is architectural and empirical, not a claim to invent data minimization, purpose limitation, proof-carrying authorization, data-fusion control, provenance, policy repair, or data-flow enforcement.

## 2. Literature Survey

### 2.1 Data Governance and Automated Governance

Khatri and Brown frame data governance around decision rights and accountability across data principles, quality, metadata, access, and lifecycle [1]. Abraham, Schneider, and vom Brocke synthesize the field into a conceptual framework and research agenda [2]. Nadal et al. operationalize and automate governance using standardized flows and metadata artifacts in a data-lake architecture [3]. Polimeno et al. study service-based pipelines in which data quality is maximized while data-protection requirements are satisfied [4]. These works motivate governance as an organizational and technical capability but do not by themselves define a single operation-level assurance model across heterogeneous transformations and AI consumption.

### 2.2 Derived Data, Fusion Control, and Proof-Oriented Authorization

Atluri and Gal explicitly address authorization for temporal and derived data [5]. Den Hartog and Zannone define policy mechanisms for data fusion and derived-data control [6], while Bertino, Verma, and Calo further study policy control over data-fusion processes and resulting data [7]. Proof-carrying authorization established that authorization decisions can be supported by machine-checkable evidence [8], [9]. These studies are direct prior foundations and

prevent any credible claim that derived-data authorization or proof-oriented policy evidence is itself novel.

### 2.3 Purpose, Privacy, Provenance, and Data Flow

Data Guard implements fine-grained purpose-based access control by routing warehouse access through purpose-specific masking views [10]. ODRL provides machine-readable permissions, prohibitions, duties, and constraints [11], while PROV-O standardizes provenance relationships [12]. GDPR purpose limitation and data-minimization principles motivate downstream processing controls [13]. Navaie argues that these privacy principles should become runtime controls for agentic AI, including purpose-aware egress, field-level minimization, regional checks, and traceability [14]. C-Trace verifies GDPR-derived predicates over agent traces [15]. Data Flow Control (DFC) enforces provenance-based predicates over tuple-level database flows during query execution [16], and Environment Steering extends data-flow control toward runtime agent safety [17].

### 2.4 Policy Repair and Current Governance Research

Policy-repair research shows that over-permissive access-control policies can be automatically repaired using symbolic analysis and formal constraints [18], while CloudFix combines formal localization with large language models for cloud-policy repair [19]. Recent governance research also examines how AI changes governance practices and how governance capability can be measured [20], [21]. DATA-PROOF therefore treats repair, AI governance, and governance measurement as prior or adjacent work.

### 2.5 Novelty Boundary

| Established capability           | Representative prior work | DATA-PROOF position                                                          |
|----------------------------------|---------------------------|------------------------------------------------------------------------------|
| Organizational data governance   | [1], [2]                  | Runtime evidence inputs                                                      |
| Automated governance / pipelines | [3], [4]                  | Operation-level assurance over existing automation                           |
| Derived-data / fusion control    | [5]-[7]                   | Historical policy foundation                                                 |
| Proof-oriented authorization     | [8], [9]                  | Evidence foundation; not claimed as new                                      |
| Purpose-based access             | [10]                      | One transition predicate                                                     |
| Usage / provenance vocabularies  | [11], [12]                | Reusable policy/evidence vocabulary                                          |
| Runtime privacy / compliance     | [13]-[15]                 | Privacy and AI-governance context                                            |
| Data-flow enforcement            | [16], [17]                | Closest modern enforcement work                                              |
| Policy repair                    | [18], [19]                | Repair is established; DATA-PROOF constrains data operations                 |
| DATA-PROOF                       | This work                 | Integrated multi-dimensional state transition + GSTC + controlled evaluation |

## 3. Problem Definition

A source-access decision can be correct while a subsequent transformation still changes the governance state of the data. DATA-PROOF models this condition explicitly. Let the governance state at time  $t$  be:

$$S_t = (C_t, P_t, L_t, Q_t, R_t, V_t, M_t, E_t, W_t)$$

where  $C$  is classification/sensitivity,  $P$  purpose compatibility,  $L$  lineage/provenance completeness,  $Q$  quality and freshness,  $R$  residency/jurisdictional state,  $V$  privacy/inference/disclosure risk,  $M$  minimization state,  $E$  downstream exposure, and  $W$  policy-version freshness.

For a proposed operation  $o_t$  with additional inputs  $D_t$  and execution/policy context  $K_t$ , the candidate post-operation state is:

$$\hat{S}_{t+1} = F(S_t, o_t, D_t, K_t)$$

The governance delta is a diagnostic indicator:

$$\Delta G_t = \text{Risk}(\hat{S}_{t+1}) - \text{Risk}(S_t)$$

A scalar delta cannot override hard policy predicates. The admissibility condition is therefore:

$$\text{Admissible}(o_t) = \text{Source} \wedge \text{Purpose} \wedge \text{Lineage} \wedge \text{Quality} \wedge \text{Residency} \wedge \text{Policy} \wedge \text{Derived} \wedge \text{Min}$$

The operational invariant is:

$$\text{Authorized}(D_1) \wedge \text{Authorized}(D_2) \not\Rightarrow \text{Authorized}(T(D_1, D_2))$$

This is not presented as a novel theorem. It is an implementation rule grounded in prior derived-data and fusion-control research [5]-[7]. The problem addressed here is how to evaluate this rule together with purpose, quality, residency, minimization, policy freshness, and downstream AI exposure in one assurance decision.

## 4. Methodology/ Approach

### 4.1 Composition Risk, Purpose Continuity, and Minimization

$$\text{CompositionFlag}(o_i) = 1, \text{ if Risk}(\hat{S}_{t+1}) \geq \tau_{\text{comp}}$$

The composition threshold may represent classification escalation, linkage or re-identification potential, cross-domain inference, external exposure, or jurisdictional policy. The current benchmark uses deterministic labels; production use requires calibrated organizational definitions.

$$\text{PurposeCompatible}(P_0, P_t) \in \{0,1\}$$

$P_0$  is the originating approved purpose and  $P_t$  is the current declared purpose. Purpose limitation is prior art and law [10], [13], [14]; DATA-PROOF keeps purpose compatibility active during later transformation and AI-consumption decisions.

$$D^* = \arg \min \text{GovernanceExposure}(D'), D' \subseteq D$$

$$\text{subject to } \text{Utility}(q, D') \geq \tau U \text{ and } \text{PolicyCompatible}(D', q) = 1$$

This formulation operationalizes task-scoped minimization; data minimization itself is not claimed as novel.

### 4.2 Constraint / Repair Selection

$$r^* = \arg \min \text{Cost}(r), r \in R(o_i)$$

$$\text{subject to } \text{Admissible}(r(o_i)) = 1 \text{ and } \text{Utility}(r(o_i)) \geq \tau U$$

$R(o_t)$  can include field removal, masking, tokenization, aggregation, region redirection, or another organization-approved transformation. This is constrained operation repair, not a new theory of policy repair [18], [19].

### 4.3 DATA-PROOF Architecture

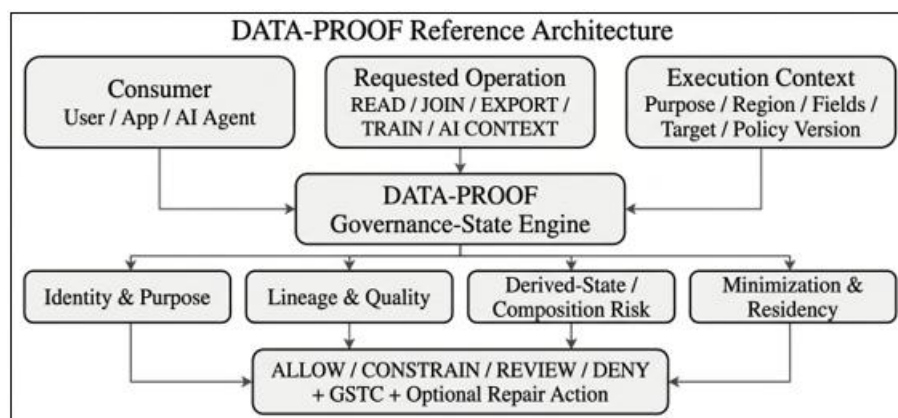


Figure 2: DATA-PROOF reference architecture

The request layer captures actor, operation, purpose, sources, target, fields, region, and policy version. The evidence layer reads identity, policy, lineage, classification, quality, freshness, and residency metadata. The transition engine predicts the candidate state, evaluates composition risk and minimization, and optionally synthesizes a permitted constraint. The outcome is recorded in a GSTC.

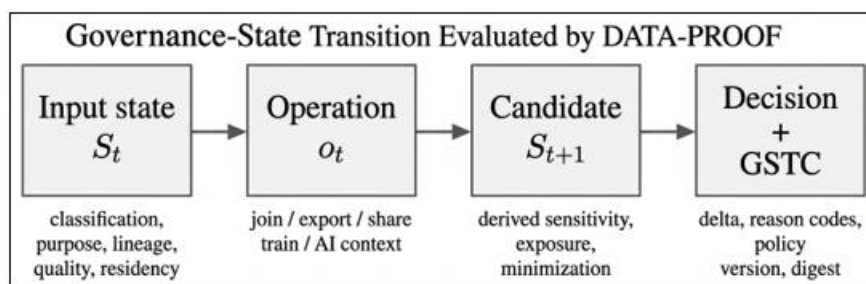


Figure 3: Governance-state transition evaluated by DATA-PROOF.

### 4.4 Governance State Transition Certificate

The GSTC is an operation-level transition witness, not a cryptographic proof of legal compliance. It records the evidence evaluated, policy version, candidate state, decision rule, applied constraint, and integrity digest.

| GSTC field             | Meaning                                                      |
|------------------------|--------------------------------------------------------------|
| request_id             | Unique operation request identifier                          |
| actor / operation      | Requester and requested processing action                    |
| source_assets          | Governed inputs contributing to the operation                |
| purpose                | Declared and evaluated purpose                               |
| pre_state / post_state | Governance state before and predicted state after processing |
| governance_delta       | Configured risk-change indicator                             |
| lineage / quality      | Provenance, freshness, quality, and evidence sufficiency     |
| constraint / repair    | Masking, minimization, tokenization, or approved change      |
| decision               | ALLOW / CONSTRAIN / REVIEW / DENY                            |
| reason_codes           | Machine-readable rationale                                   |
| policy_version         | Policy basis used for the evaluation                         |
| digest                 | SHA-256 integrity digest over canonical certificate payload  |

| ID  | Scenario family                            | Expected  |
|-----|--------------------------------------------|-----------|
| S1  | Normal governed read                       | ALLOW     |
| S2  | Wrong-purpose export                       | DENY      |
| S3  | Disallowed region transfer                 | DENY      |
| S4  | Stale/low-quality evidence                 | REVIEW    |
| S5  | Missing lineage                            | REVIEW    |
| S6  | Sensitive derived join                     | CONSTRAIN |
| S7  | AI overcollection                          | CONSTRAIN |
| S8  | Purpose drift at model training            | DENY      |
| S9  | Allowed masked join                        | ALLOW     |
| S10 | Quality defect requiring review            | REVIEW    |
| S11 | Region violation despite possible redirect | DENY      |
| S12 | Stale policy version                       | REVIEW    |
| S13 | High-risk external release                 | CONSTRAIN |
| S14 | Benign aggregation                         | ALLOW     |
| S15 | Unauthorized source role                   | DENY      |
| S16 | Unapproved AI fine-tuning reuse            | DENY      |

#### 4.5 Decision Procedure

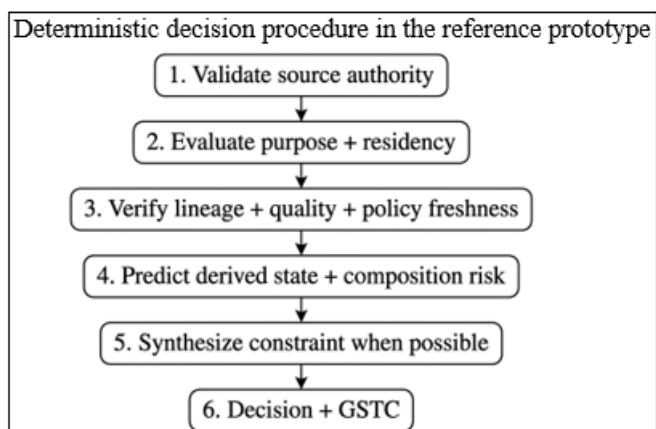


**Figure 4:** Deterministic decision procedure in the reference prototype

The implementation validates source authority; evaluates purpose and residency; verifies lineage, quality, freshness, and policy version; computes the candidate derived state; evaluates composition risk and whether a permissible constraint exists; then emits ALLOW, CONSTRAIN, REVIEW, or DENY together with the GSTC. Hard source, purpose, or residency violations are denied. Insufficient evidence or stale policy is escalated to review. Repairable composition or minimization violations are constrained.

#### 4.6 Prototype and Experimental Design

The prototype is implemented in Python and isolates control semantics from cloud infrastructure. Each request contains source-authorization state, declared purpose, region compatibility, quality/freshness, lineage status, composition-risk flag, minimization requirement, policy freshness, and expected decision. GSTCs are serialized canonically and hashed with SHA-256.

RBAC models source-role authorization only. A simplified ABAC baseline adds purpose and residency checks. These baselines provide transparent lower bounds on control coverage and are not proxies for mature PBAC, information-flow-control, DFC, or commercial governance systems.

Each scenario family was executed 2,000 times per primary method for local latency measurement, producing 32,000 executions per method. Seven ablations independently disable purpose, residency, quality, lineage, composition, minimization, or policy-version checks. The integrity test generates 5,000 valid GSTCs and mutates a decision field after certificate generation. The HTTP path was additionally exercised through FastAPI and then repeated under Docker Compose, including decision generation, SQLite certificate persistence, retrieval, and digest verification. Repeated executions are performance measurements, not independent empirical samples.

Policy conformance is the fraction of canonical scenario families returning the exact configured outcome. Unrestricted-Allow Error Rate (UAER) is the fraction of non-ALLOW families incorrectly returned as ALLOW. Latency is measured in-process and excludes network, database, catalog, graph, remote policy-engine, and durable-evidence costs.

UAER = Incorrect unrestricted ALLOW families / Non-ALLOW families

DMR =  $1 - |D^*| / |D|$

## 5. Results & Discussion

### 5.1 Primary Policy Coverage

| Method     | Policy conformance | UAER  |
|------------|--------------------|-------|
| RBAC       | 4/16 (25.0%)       | 92.3% |
| ABAC       | 9/16 (56.2%)       | 53.8% |
| DATA-PROOF | 16/16 (100.0%)     | 0.0%  |

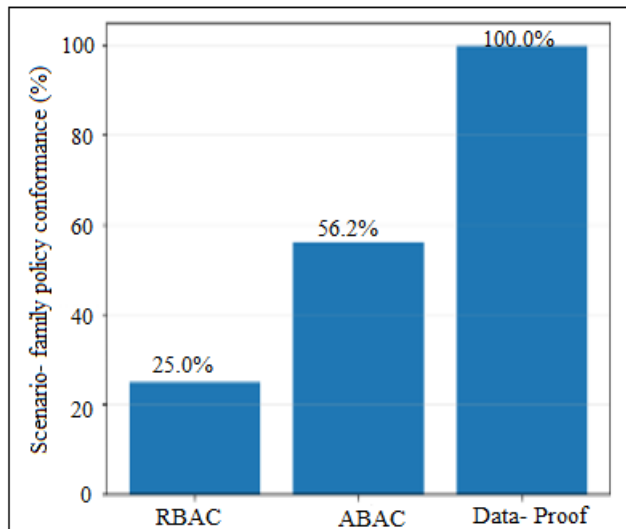


**Figure 5:** Policy conformance across the 16 canonical scenario families.

RBAC conformed in four families because source authorization alone cannot distinguish most downstream governance conditions. Simplified ABAC improved coverage to nine families by enforcing purpose and residency. DATA-PROOF conformed in all 16 configured families and produced no unrestricted-allow errors. Because the expected outcomes are authored policy semantics, these results demonstrate implementation coverage, not predictive accuracy.

## 5.2 Ablation Analysis

| Removed dimension | Conformance   | UAER  |
|-------------------|---------------|-------|
| Purpose           | 13/16 (81.2%) | 15.4% |
| Residency         | 14/16 (87.5%) | 7.7%  |
| Quality           | 14/16 (87.5%) | 7.7%  |
| Lineage           | 15/16 (93.8%) | 7.7%  |
| Composition       | 15/16 (93.8%) | 7.7%  |
| Minimization      | 15/16 (93.8%) | 7.7%  |
| Policy freshness  | 15/16 (93.8%) | 7.7%  |

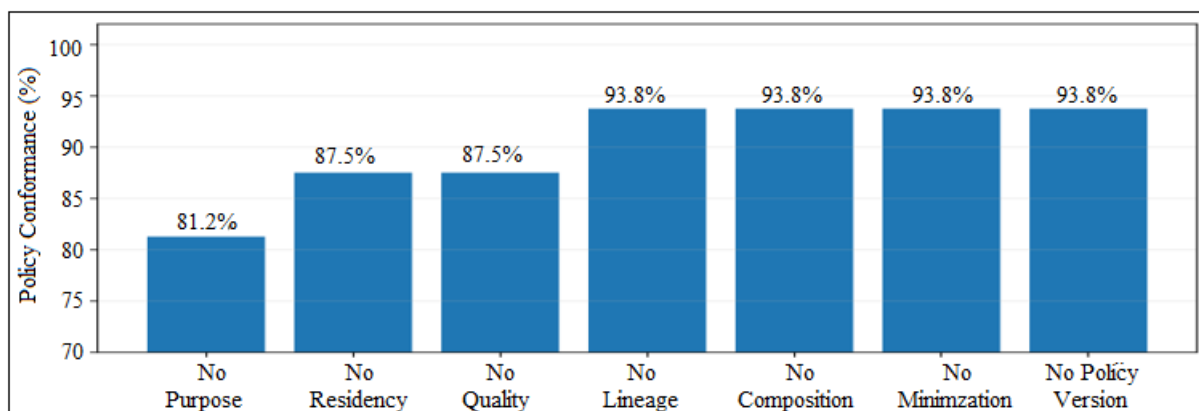


**Figure 6:** Ablation results after disabling one governance dimension

Purpose removal caused the largest coverage loss because multiple scenarios depend directly on incompatible reuse. Residency and quality each reduced conformance to 87.5%, while lineage, composition, minimization, and policy freshness each caused at least one canonical failure. The result supports the benchmark relevance of each modeled dimension, while remaining dependent on the authored scenario set.

## 5.3 GSTC Integrity

All 5,000 generated certificates verified successfully before mutation. After modifying the decision field while leaving the stored digest unchanged, verification rejected 5,000/5,000 certificates. This confirms correct digest binding in the prototype but does not establish non-repudiation, protected timestamping, or tamper-resistant storage.

## 5.4 Local Overhead and Container Validation

| Method     | Mean (us) | p95 (us) | Executions |
|------------|-----------|----------|------------|
| RBAC       | 0.179     | 0.193    | 32,000     |
| ABAC       | 0.203     | 0.226    | 32,000     |
| DATA-PROOF | 54.555    | 80.500   | 32,000     |

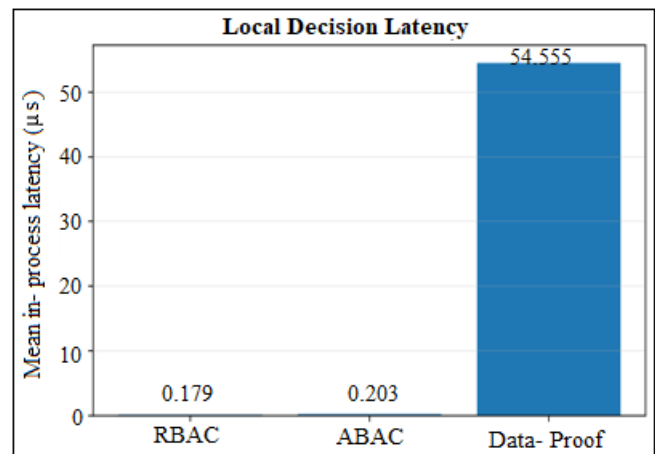


**Figure 7:** Mean in-process decision latency in the local reference prototype

On the reproduced Windows run, DATA-PROOF evaluation including certificate generation averaged 54.55 microseconds in-process across 32,000 executions, with p95 = 80.50 microseconds and p99 = 120.40 microseconds. The FastAPI service returned the expected ALLOW and CONSTRAIN decisions, persisted GSTCs in SQLite, and verified stored certificate digests. The same AI-overcollection decision, persistence, and verification path was reproduced successfully inside Docker Compose. These measurements are environment-specific and should not be

extrapolated to enterprise deployments where metadata retrieval, database provenance, remote policy engines, durable evidence stores, and cryptographic signing may dominate end-to-end latency.

### 5.5 Relationship to Closest Prior Work

DFC is more mature than DATA-PROOF at database-native flow enforcement: it reasons over tuple-level provenance and performs query-level enforcement [16]. DATA-PROOF does not claim stronger DBMS guarantees. Its proposed value is a cross-system governance-state abstraction combining evidence not normally available inside a DBMS, including purpose continuity, data quality/freshness, residency, policy-version freshness, minimization, and downstream AI exposure. In a future enterprise implementation, DFC could operate beneath DATA-PROOF as one enforcement mechanism rather than as an incompatible alternative.

Agentic systems sharpen the problem because models and tools assemble context dynamically. Purpose-aware egress and minimization are already advocated in privacy engineering [14], and Environment Steering shows how data-flow control can shape agent behavior [17]. DATA-PROOF contributes a governance-state view of the context about to be released to a model or tool.

### 5.6 Limitations

The scenario families and expected outcomes are authored by the framework designer; policy conformance therefore measures consistency with stated semantics rather than independent correctness. Synthetic risk flags are not calibrated privacy or legal risk estimates. RBAC and ABAC are simplified pedagogical baselines, and the study does not experimentally compare against Data Guard, DFC, ODRL-based usage-control engines, full information-flow-control systems, commercial governance platforms, or human governance review. No production enterprise dataset, live catalog, database lineage engine, or independent governance-review panel was used. SHA-256 payload binding is an integrity check, not non-repudiation. DATA-PROOF can encode organizational policy predicates but cannot determine lawful basis or legal interpretation.

## 6. Conclusion

DATA-PROOF presents a multi-dimensional governance-state assurance architecture for enterprise data transformation and AI consumption. Rather than automatically inheriting a source-access decision, each consequential operation is evaluated as a transition over purpose, lineage, quality/freshness, residency, derived-state risk, minimization, downstream exposure, and policy freshness. The framework can allow, constrain, review, or deny an operation and records the evaluated transition in a Governance State Transition Certificate.

The contribution is deliberately bounded by prior art. Derived-data authorization, data-fusion control, proof-carrying authorization, purpose-based access, provenance, data-flow enforcement, data minimization, and policy repair

are established areas. The current evidence supports prototype feasibility and policy coverage under controlled synthetic semantics; it does not establish production effectiveness or legal compliance.

## 7. Future Scope

The next stage should validate DATA-PROOF using real or open transformation workloads rather than only canonical synthetic scenarios. A strong extension would execute representative SQL or pipeline transformations, derive lineage using an independent engine, and compare decisions against at least one stronger purpose/data-flow baseline and independent governance judgments. The evaluation should include ambiguous transformations, incomplete lineage, conflicting metadata, re-identification risk, policy changes, multi-region pipelines, and AI-agent tool calls.

Future work should also measure end-to-end latency including catalog, lineage, policy, and evidence persistence; calibrate composition-risk thresholds; test protected or signed GSTC storage; and study whether constrained repair can preserve business utility while reducing unsafe data exposure. Publishing the benchmark code, scenario matrix, policy semantics, and result files will support independent reproduction.

## References

- [1] V. Khatri and C. V. Brown, "Designing data governance," *Communications of the ACM*, vol. 53, no. 1, pp. 148-152, 2010. doi:10.1145/1629175.1629210.
- [2] R. Abraham, J. Schneider, and J. vom Brocke, "Data governance: A conceptual framework, structured review, and research agenda," *International Journal of Information Management*, vol. 49, pp. 424-438, 2019. doi:10.1016/j.ijinfomgt.2019.07.008.
- [3] S. Nadal, P. Jovanovic, B. Bilalli, and O. Romero, "Operationalizing and automating Data Governance," *Journal of Big Data*, vol. 9, art. 117, 2022. doi:10.1186/s40537-022-00673-5.
- [4] A. Polimeno, C. Braghin, M. Anisetti, and C. A. Ardagna, "Maximizing data quality while ensuring data protection in service-based data pipelines," *Journal of Big Data*, vol. 12, art. 62, 2025. doi:10.1186/s40537-025-01118-5.
- [5] V. Atluri and A. Gal, "An authorization model for temporal and derived data: securing information portals," *ACM Transactions on Information and System Security*, vol. 5, no. 1, pp. 62-94, 2002. doi:10.1145/504909.504912.
- [6] J. den Hartog and N. Zannone, "A Policy Framework for Data Fusion and Derived Data Control," *Proc. ACM International Workshop on Attribute Based Access Control*, pp. 47-57, 2016. doi:10.1145/2875491.2875492.
- [7] E. Bertino, D. Verma, and S. Calo, "A Policy System for Control of Data Fusion Processes and Derived Data," *Proc. FUSION*, 2018.
- [8] L. Bauer, M. A. Schneider, E. W. Felten, and A. W. Appel, "Access control on the Web using proof-carrying authorization," *Proc. DARPA Information*

- Survivability Conference and Exposition, pp. 117-119, 2003. doi:10.1109/DISCEX.2003.1194942.
- [9] L. Bauer, Access Control for the Web via Proof-Carrying Authorization, Ph.D. dissertation, Princeton University, Technical Report TR-677-03, 2003.
  - [10] K. Tran, S. Vasudevan, P. Desai, et al., "Data Guard: A Fine-grained Purpose-based Access Control System for Large Data Warehouses," arXiv:2502.01998, 2025.
  - [11] W3C, ODRL Information Model 2.2, W3C Recommendation, 2018.
  - [12] T. Lebo, S. Sahoo, and D. McGuinness, Eds., PROV-O: The PROV Ontology, W3C Recommendation, 2013.
  - [13] European Parliament and Council, Regulation (EU) 2016/679 (General Data Protection Regulation), 2016.
  - [14] K. Navaie, "From rights to runtime: Privacy engineering for agentic AI," AI Magazine, vol. 46, no. 4, e70036, 2025. doi:10.1002/aaai.70036.
  - [15] N. Kahani, M. Barati, and D. Addae, "Runtime Compliance Verification for AI Agents," arXiv:2606.19242, 2026.
  - [16] C. Summers and E. Wu, "Data Flow Control: Data Safety Policies for AI Agents," arXiv:2606.05679, 2026.
  - [17] C. Summers, P. Raghunath, A. Pai, M. Kulkarni, Z. Zhang, O. Kennedy, and E. Wu, "Environment Steering: Using Data Flow Control to Improve Agent Utility and Safety," arXiv:2609.35807, 2026.
  - [18] W. Eiers, G. Sankaran, and T. Bultan, "Quantitative Policy Repair for Access Control on the Cloud," Proc. 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis, pp. 564-575, 2023. doi:10.1145/3597926.3598078.
  - [19] B. Hall, O. Ungaro, and W. Eiers, "CloudFix: Automated Policy Repair for Cloud Access Control Policies Using Large Language Models," Proc. SANER 2026, pp. 322-332, 2026. doi:10.1109/SANER67736.2026.00042.
  - [20] W. Alhammad, D. Sammon, and C. Heavin, "The impact of AI on data governance practices: a comparative framework," Journal of Decision Systems, vol. 35, no. 1, art. 2675043, 2026. doi:10.1080/12460125.2026.2675043.
  - [21] T. R. Dockara and M. Malhotra, "Data governance for sustainable AI in organizations: a measurement-oriented capability method with benchmarkability robustness checks and marketplace-microdata operationalization," Discover Artificial Intelligence, vol. 6, art. 574, 2026. doi:10.1007/s44163-026-01343-8.