engineering proceedings

MDPI

Proceeding Paper

LegalShadow: An Ontology-Guided Digital Shadow Architecture for Semantic Validation of LegalTech Ecosystems †

Patricio M. Paccha-Angamarca [1,2,*], Erwin J. Sacoto-Cabrera [2] and Víctor V. Velepucha-Bonett [1]

[1] Software Engineering Creation, Application and Management Group (GI-CreAGSoft), National Polytechnic School, Ladrón de Guevara E11-253, Quito 170525, Ecuador; victor.velepucha@epn.edu.ec
[2] Cloud Computing, Smart Cities & High-Performance Computing Group (GIHP4C), Salesian Polytechnic University, Calle Vieja 12-30, Cuenca 010105, Ecuador; esacoto@ups.edu.ec
* Correspondence: patricio.paccha@epn.edu.ec
† Presented at the Sixth International Conference on Information Systems and Software Technologies (ICI2ST 2026), Quito, Ecuador, 23–25 September 2026.

Abstract

Judicial digital transformation produces heterogeneous ecosystems whose governance maturity is rarely auditable in an automated and semantically grounded way. This study asks to what extent the publicly observable digital surface of a judicial institution materializes, in machine-processable form, the governance semantics that its declared frameworks prescribe. We present LegalShadow, an ontology-guided digital shadow architecture: a ten-dimension governance ontology (TOGAF, COBIT, ITIL, NIST CSF, artificial intelligence, distributed ledger, open data, security, interoperability and a LegalTech domain) drives a read-only web acquisition protocol with per-snapshot SHA-256 provenance; the evidence is serialized as a JSON-LD knowledge graph, and a hierarchical coverage function Ψ and gap function δ score it. A containerized implementation was exercised on the Ecuadorian Judiciary Council over 17 dated runs (187 requests; 170 snapshots re-verified, 0 mismatches). The institution attains a four-dimension declarative maturity of 46/100, yet ten-dimension semantic coverage reaches only $\Psi_{\text{global}} = 34.9\%$ against 90.4% for a synthetic control, with three dimensions at zero. The contributions are the reference architecture, an auditable non-invasive acquisition protocol, and a coverage metric that makes ontology–artifact alignment verifiable and comparable across judicial ecosystems.

Keywords: digital shadow; ontology; semantic validation; JSON-LD; LegalTech; e-justice; e-government; open judicial data; digital maturity

1. Introduction

Judicial systems are rapidly digitalizing, yet governance remains challenged by organizational complexity, vendor dependence, fragmented infrastructures, and limited interoperability [1]. Moreover, a persistent gap exists between the publication of open government data and its effective machine-readable reuse [2,3]. This research builds upon a prior line of work on judicial digital governance.

This research builds on prior work in judicial digital governance, reusing the validated governance ontology Ω developed through a multidimensional architecture for LegalTech governance (MALTG) [4,5] and subsequently operationalized in a structural digital twin [6]. What this paper reuses unchanged from that line is the governance ontology Ω itself—its 131 classes, its ten dimensions, and their versioned scoring rubric. What is new here is everything that surrounds it: the acquisition tier and its evidence protocol, the anchoring Γ

Check for updates

Academic Editor: Samuel Lascano

Published: 18 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

from observed web signals to ontology classes, the coverage and gap functions $\langle \Psi, \delta \rangle$, and the empirical study. The distinction matters because the earlier works validated internal structural consistency against baseline configurations that their authors had themselves declared; the present study changes the object of measurement rather than refining the previous instrument, moving from a twin that simulates a system its authors configure to a shadow that observes a public digital surface they neither control nor can instrument.

Cyber-physical systems engineering provides an appropriate abstraction for this work. Following Kritzinger et al. [7], a *digital shadow* is characterized by an automated real-to-virtual information flow while feedback remains under human control. Here, the concept of *structural digital shadow* is used to emphasize the representation of ecosystem components, flows, and layers rather than dynamic behavior. This configuration is particularly suitable for the judicial domain, where intervention in public institutional infrastructure must remain under human authority, making the artifact primarily valuable for observation and diagnosis. The paradigm has expanded from manufacturing [8] to cities and public services [9,10], and Sepasgozar warns that conflating twin and shadow delays adoption [11]. Crucially, the systematic review of Karabulut et al. finds no established method to *verify* that a twin actually materializes its semantic model [12]. The legal domain, in turn, has a mature tradition in knowledge representation [13] and judicial decision prediction [14], but both presuppose accessible structured data.

Section 2 compares this work against the three lines of research that border it most closely. Within the scope of that comparison, we found no reported architecture that simultaneously (i) builds a structural digital shadow of a judicial technological ecosystem automatically from public web evidence, (ii) serializes it as a JSON-LD knowledge graph anchored to a governance ontology, and (iii) validates that anchoring through a hierarchical coverage metric. Each ingredient exists separately in the literature; the combination, and in particular the direction of validation—from an *observed* artifact towards a prescriptive ontology—is what we did not find.

This gap yields the question that organizes the paper: *To what extent does the publicly observable digital surface of a judicial institution materialize, in machine-processable form, the governance semantics prescribed by the very frameworks the institution declares?* Accordingly, this paper contributes:

(C1) The LegalShadow reference architecture and its coverage/gap model $\langle \Psi, \delta \rangle$;
(C2) An auditable, non-invasive acquisition protocol with per-snapshot cryptographic provenance, specified as an executable procedure (Algorithm 1);
(C3) An empirical study of the Ecuadorian Judiciary Council over 17 dated runs that quantifies the distance between declared digitalization and machine-actionable governance, with a synthetic positive control isolating instrument effects from ecosystem effects.

Algorithm 1 Digital shadow generation and semantic validation

Input: seed targets T; ontology Ω; dimensions D; weights (w_r, w_s)
Output: shadow Δ; coverage set R; per-dimension Ψ, sds, δ; index L

1. derive the probe set P from Ω: each class $c \in C$ contributes its observable evidence pattern
2. **for each** $\tau \in T$, **for each** $\rho \in P$: fetch read-only (identified agent, 400 KB cap); store snapshot; log the probing class, URL, UTC timestamp, bytes, latency, SHA-256 and match outcome
3. verify integrity: re-hash each snapshot, abort on mismatch, append the run to the ledger
4. extract deterministic signals from the markup (generator, legacy links, JSON-LD, API)
5. project signals into $\Delta = G(V, E, \lambda)$ and serialize as JSON-LD with the `@context` of Ω
6. **for each** $v \in V$: resolve $\Gamma(v) \subseteq C$ via `maltg:mapsTo`; reject dangling identifiers
7. $R \leftarrow \bigcup_{v \in V} \Gamma(v)$ //*Equation* (1)
8. **for each** $d \in D$: $\Psi(d) \leftarrow$ Equation (2); then sds(d) and $\delta(d)$ by Equation (3)
9. aggregate to Ψ_{global} (Equation (4)); $L \leftarrow$ band; sort D by δ descending
10. persist the dated artifact with the ontology and shadow digests

https://doi.org/10.3390/engproc2026153023

2. Literature Review

Three research lines border this work.

Digital twins of public systems. These are now assessed rather than merely built, through maturity models [15], Delphi-derived governance ratings [16] and e-government service frameworks [17]. All assess a twin program run by the system's *own operator*, over information that the operator supplies. LegalShadow inverts both premises: a third party's ecosystem, from outside, using only what it publishes.

Open government data quality. Hundreds of portals have been scored automatically on computable metadata metrics [18–20]. The acquisition philosophy is close to ours—automatic, non-invasive, portal-agnostic—but the unit of analysis is the *dataset and its metadata record* and the yardstick is a publication schema. Neither the institution's architecture nor its governance semantics is in scope.

Ontology and knowledge-graph quality. Metric families here resemble Ψ formally [21–23], but all measure how well an *ontology* covers a domain, or how well a graph is internally formed. Ψ runs in the opposite direction—the ontology is the fixed yardstick and the *artifact* is scored—which is the verification gap Karabulut et al. report as unaddressed [12].

Law, legal informatics and the governance of legal technology. The legal domain has a long tradition of formal knowledge representation, from the twenty-five years of work surveyed by Bench-Capon et al. [24] to the Semantic Web agenda for law set out by Casanovas et al. [13], and a parallel line of quantitative legal analytics that predicts judicial behavior from structured case data [14,25]. Both presuppose what this paper measures: that legally relevant information is already available in machine-processable form. On the regulatory side, the governance of legal technology is increasingly framed as a compliance problem, with legal risk, data protection, and algorithmic accountability treated as design constraints rather than afterthoughts [26]. This is precisely the concern encoded in our AI-integration dimension. Complementary e-justice research documents the organizational risks—complexity, vendor dependence, and fragmentation—that make judicial information systems inherently difficult to govern [1].

Table 1 summarizes this delimitation. Each row is one of the three neighboring lines of work, and the last row is the present study; the three columns are the axes on which the novelty claim rests: *what is observed*, *how the evidence is obtained* and *what is validated*. Read across a row, the table shows that no neighboring line combines observation of a third party's ecosystem, acquisition without operator cooperation, and validation of an observed artifact against a prescriptive ontology. Read down the last column, it shows the reversal that defines this work: the ontology is the fixed yardstick and the artifact is what gets scored.

Table 1. Delimitation against neighboring lines of work, along the three axes of the novelty claim.

Line of Work	Object Observed	How Evidence Is Obtained	What Is Validated
Digital twins of public systems [15–17]	The operator's own system, modeled by the operator	Instrumented; operator supplies the data	Maturity of the twin program, by expert rubric
Open data/metadata quality [18–20]	Datasets and their metadata records	Automatic harvesting through portal APIs	Conformance of metadata to DCAT/FAIR
Ontology and graph quality [21–23]	The ontology or the graph itself	Domain corpus, or graph internals	How well the ontology covers a domain
LegalShadow (this work)	A third party's public digital ecosystem	Read-only web observation, no operator cooperation	How much of the ontology an observed artifact materializes

3. Formal Model

The model answers one question with arithmetic instead of opinion: *how much of what the governance standards prescribe can actually be seen in the institution's digital surface?* The ontology $\Omega = \langle C, P, I, \sigma \rangle$ states what should exist—classes C carrying a maturity score $\sigma : C \to [0, 100]$ assigned against the published rubric of Section 5; the shadow $\Delta = G(V, E, \lambda)$ records what was observed—components V, flows E, labels λ carrying layer, evidence and metadata; and the anchoring $\Gamma : V \to 2^C$ links the two through the `maltg:mapsTo` property. The measurement then reduces to three steps.

Step 1—What did we actually see? Each component declares which ontology classes it materializes, so the classes evidenced anywhere in the ecosystem are the union of those declarations:

$$R = \bigcup_{v \in V} \Gamma(v) \subseteq C. \tag{1}$$

Step 2—How much of a dimension is there? A governance dimension d is not a flat list: it has a *root* concept r_d (the framework itself, e.g., `ITIL`) and *subconcepts* S_d that make it concrete (e.g., incident management). Naming a framework without operationalizing it differs from the converse, so the two parts are scored separately and added:

$$\Psi(d) = \underbrace{w_r \cdot \mathbb{1}[\, r_d \in R \,]}_{\text{is the framework there at all?}} + \underbrace{w_s \cdot \frac{|S_d \cap R|}{|S_d|}}_{\text{how much of its detail is there?}} \in [0, 1], \tag{2}$$

With the default weights $(w_r, w_s) = (0.40,\ 0.60)$, the metric assigns 40% of the score to the presence of the root concept and 60% to the coverage of its subconcepts. As a result, $\Psi(d)$ increases monotonically with additional evidence and reaches its maximum only when both the root and all associated subconcepts are represented. Sensitivity analysis shows that varying w_r within $[0.10, 0.90]$ produces only minor changes in Ψ_{global} and does not alter the prioritization of remediation actions, indicating robustness to weight selection. The weights can be calibrated through the Analytic Hierarchy Process [27], although such calibration was not applied in this study.

Step 3—what is that worth, and what should be fixed first? With $\text{onto}(d)$ the maturity the ontology prescribes (the mean expert score of its classes), what the shadow delivers is that prescription discounted by visibility, and the gap is the remainder:

$$\text{sds}(d) = \text{onto}(d) \cdot \Psi(d), \qquad \delta(d) = \text{onto}(d)\,\big(1 - \Psi(d)\big). \tag{3}$$

A dimension is a priority only when *both* important (high onto) *and* poorly evidenced (low Ψ), so sorting by δ gives the remediation agenda directly. Aggregating with weights ω_d summing to one—all equal here, no expert panel having been convened—gives the headline figure

$$\Psi_{\text{global}} = \frac{\text{SDS}_{\text{global}}}{\text{Onto}_{\text{global}}} = \frac{\sum_d \omega_d \,\text{sds}(d)}{\sum_d \omega_d \,\text{onto}(d)}. \tag{4}$$

The institutional index is interpreted through a five-level maturity scale adapted from IT governance models [5]. The coverage metric Ψ is computed with linear complexity in the scoring stage, while ontology parsing and graph construction are performed only once during ingestion. Figure 1 highlights the diagnostic value of Ψ. *Open data* achieves full coverage ($\Psi = 1.00$), *TOGAF* shows partial implementation ($\Psi = 0.20$), and *ITIL* is absent ($\Psi = 0$). By distinguishing complete, partial, and missing practices, Ψ provides more actionable insights than a single aggregate metric. The *Security posture* example further illustrates the traceability from observed evidence to coverage assessment through Equation (2).

Its root r_d = `maltg:Security_Posture` is a published governance statement, and $|S_d| = 4$ subconcepts operationalize it: transport security, integrity policy, vulnerability

disclosure and incident response. The scan matches the first two and finds neither a disclosure channel nor an incident-response reference nor a statement to anchor the root—hence a null root term and a subconcept fraction of 2/4:

$$\Psi(\text{security}) = 0.40 \cdot \mathbb{1}[\, r_d \notin R \,] + 0.60 \cdot \tfrac{2}{4} = 0.00 + 0.30 = 0.30,$$

and with onto = 80.4 from the rubric, Equation (3) gives sds = 24.1 and $\delta = 56.3$, placing security fifth on the agenda of Table 2—not because the institution does nothing about security, but because two-thirds of what the framework prescribes leaves no machine-readable trace. Every row of that table follows this same procedure, and r_d, S_d and onto(d) are all fixed by the versioned rubric before any observation is made.

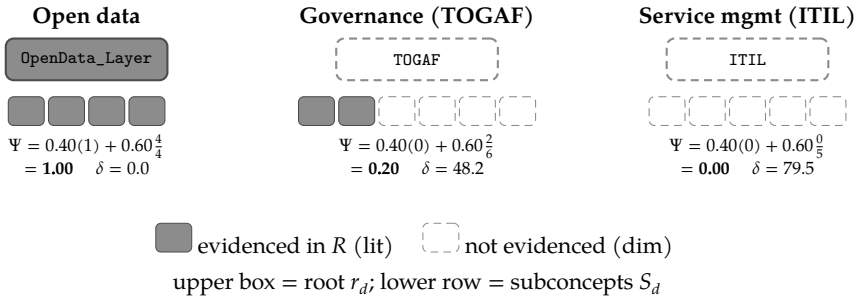


Figure 1. The coverage function on three real dimensions. Filled nodes are concepts evidenced by at least one shadow component; hollow dashed nodes are prescribed but unobserved. The subconcept counts shown here ($|S_d| = 4$, 6 and 5) are the ontology's actual counts for these dimensions.

Table 2. Hierarchical coverage validation of the real institutional shadow, $(w_r, w_s) = (0.40, 0.60)$. The last column reports the synthetic positive control for the same dimension. Aggregates follow Equation (4) with equal ω_d and are computed on unrounded values; Ψ_{global} is thus an onto-weighted mean of the $\Psi(d)$ column, whose unweighted mean is 0.329 (real) and 0.900 (control).

Dimension	Onto(d)	Root	$\Psi(d)$	Sds(d)	$\delta(d)$	Ψ Ctrl
AI integration	96.9	No	0.150	14.5	**82.4**	0.85
Service management (ITIL)	79.5	No	0.000	0.0	**79.5**	1.00
Resilience (NIST CSF)	72.0	No	0.120	8.6	63.4	1.00
Control (COBIT)	62.7	No	0.000	0.0	62.7	1.00
Security posture	80.4	No	0.300	24.1	56.3	1.00
Distributed ledger	52.4	No	0.000	0.0	52.4	0.85
Governance (TOGAF)	60.3	No	0.200	12.1	48.2	0.70
LegalTech domain	75.5	Yes	0.673	50.8	24.7	0.60
Interoperability	81.2	Yes	0.850	69.0	12.2	1.00
Open data	78.5	Yes	1.000	78.5	0.0	1.00
Global (Equation (4), onto-weighted)	**73.9**	3/10	**0.349**	**25.8**	**48.2**	**0.904**

Two properties of the aggregate matter before the empirical section. Ψ_{global} is an onto-weighted mean of the per-dimension coverages, not their unweighted mean, so the same shortfall costs more in a high-onto dimension. And every quantity is a deterministic function of the ontology and the evidence retrieved, with no judgment made after the data were seen—which is what makes the apparatus an audit instrument, recomputable by a third party, rather than a scoring opinion.

4. Architecture and Acquisition

LegalShadow is organized in four decoupled *tiers* deployed as Docker containers: acquisition, semantic modeling, validation, and presentation. The **acquisition tier** collects public web evidence; the **semantic tier** represents the ecosystem through the ontology Ω

and generates the shadow Δ in JSON-LD; the **validation tier** computes Ψ, δ, and L; and the **presentation tier** provides access through a REST API and an interactive visualization. This decoupled architecture preserves traceability and prevents downstream processes from altering collected evidence. Information flows unidirectionally from the observed ecosystem to the shadow, reflecting the digital shadow paradigm rather than a digital twin [7,11]. The evaluation process follows the MAPE-K framework [28], with actions limited to recommendations and leaving decision-making under institutional control.

Acquisition protocol. The acquisition layer follows an ontology-driven strategy in which Ω defines the observable evidence to be collected, making the OWL ontology both the search specification and the mechanism for constraining scope. This a priori definition ensures a transparent, auditable, and ethically bounded process, where each probe is explicitly linked to the ontology class that motivated it.

Consistent with ethical and legal recommendations for web data collection [29], the agent limits itself to public, unauthenticated resources, performs only read operations, identifies its purpose through a dedicated *User-Agent*, applies conservative rate limits, and records all observations with timestamps and SHA-256 checksums. All probes, outcomes, and methodological decisions are stored in an append-only hash-chained ledger, enabling independent verification and reproducibility. The collected signals correspond directly to ontology concepts, including technology indicators, legacy frameworks, SPA deployments, open-data and integrity-policy references, as well as structured data and public APIs [30]. Classes without matching evidence are explicitly recorded as absent, transforming non-detection into verifiable evidence rather than missing information.

Signals are projected into a JSON-LD instance whose `@context` reuses standard vocabularies (Dublin Core, SKOS, Schema.org) and defines the `maltg:` namespace for governance terms—a format chosen for its native fit with web development [30], its gateway status into linked data [31] and its FAIR alignment [32]. Each component carries its layer, evidence (URL, status, size, latency), findings, gaps and anchoring Γ, turning the shadow into a queryable knowledge subgraph [33] in which every macro-level index remains traceable to the HTTP response that justifies it.

The anchoring Γ and how to audit it. Reproducibility relies on Γ, a versioned and deterministic rule set that maps observed signals to ontology classes. Each rule associates a verifiable signal with a class membership, allowing reevaluation directly from stored snapshots rather than human interpretation: anchoring `OpenData_Layer` requires a resolvable reference to an entry in the national data catalog, and anchoring `Transport_Security` requires a `Strict-Transport-Security` header on the response actually stored. The mapping is complete and referentially consistent, as invalid identifiers are rejected during ingestion (Algorithm 1, step 6). Because both rules and evidence are versioned, reviewers can independently recompute Γ and R without repeating acquisition, ensuring separate auditability of observation and interpretation. The rules are intentionally conservative: the absence of a signal indicates only its absence from the institution's public digital surface, providing a documented lower-bound assessment. Figure 2 illustrates this process, linking evidence to stored HTTP responses, associating detected classes through `maltg:mapsTo` is $\Gamma(v)$, and recording missing expected signals through `maltg:gap`.

The GET-SDS console (Figure 3) implements the acquisition experiment end to end, and its interface makes the ontology-driven character of the method visible. It loads Ω, derives the probe set, runs the protocol against the configured target, and renders the ontology as a three-dimensional star field in which each class is a star and each governance dimension a named constellation. The field starts uniformly dim—every class is a question the scan has not yet answered—and lights only what the probes match, so Γ and the coverage set R of Equation (1) become observable at a glance while the ledger records the same

https://doi.org/10.3390/engproc2026153023

events in auditable form. Because both the ontology and the target are parameters, the same console runs unchanged on any institution.

```
1  { "@context": { "schema":"https://schema.org/",
2                  "maltg": "https://w3id.org/maltg/governance#", ... },
3    "@id":   "shadow:cj-2026-08-04/component/07",
4    "@type": "maltg:EcosystemComponent",
5    "dct:title":   "National open-data catalogue entry",
6    "maltg:layer": "OpenData_Layer",
7    "maltg:evidence": {
8        "schema:url":  "https://www.datosabiertos.gob.ec/dataset/...",
9        "schema:httpStatus": 200,  "maltg:bytes": 118442,      ... },
10   "maltg:signal": [ "catalogue-reference", "licence-declared" ],
11   "maltg:mapsTo": [ "maltg:OpenData_Layer",  "maltg:DCAT_Publication",
12                     "maltg:Licence_Openness","maltg:Catalogue_Federation" ],
13   "maltg:gap":    [ "no embedded schema.org JSON-LD on the institutional portal" ] }
```

Figure 2. One component of the institutional shadow, as serialized in JSON-LD. Lines 7–9 hold the `evidence` block that pins the component to one stored HTTP response, its digest and its capture time; line 11 lists the deterministic `signals` the probes matched; lines 12–13 are the anchoring $\Gamma(v)$, whose four classes enter the coverage set R; line 14 records through `gap` what the rules sought and did not find. Evidence, anchoring, and gap travel together; each class in R links to its justifying response.

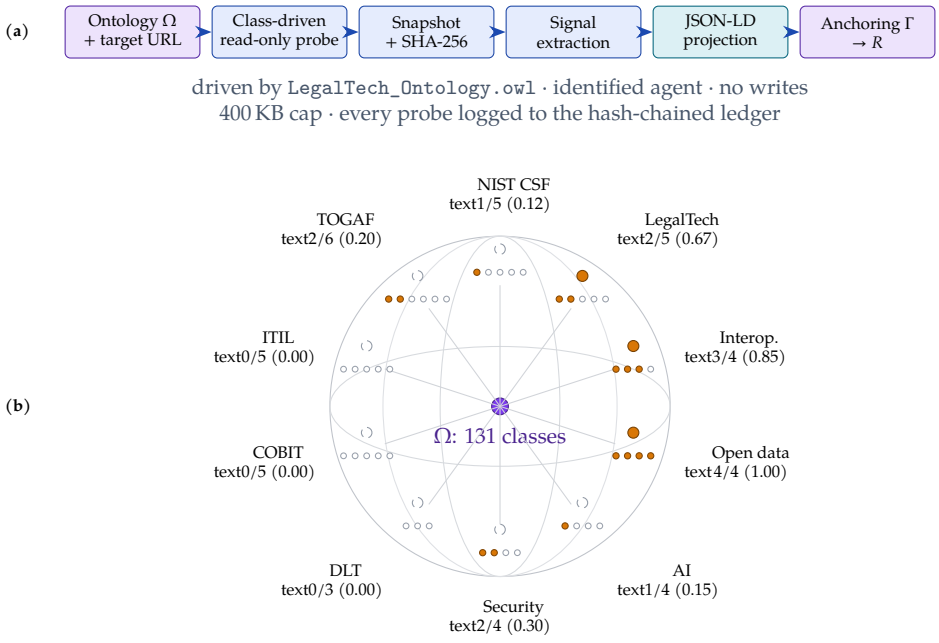


Figure 3. The GET-SDS acquisition process, driven by the ontology. (**a**) Pipeline: Ω declares what is worth looking for, the probe fetches only that, and every step is written to the hash-chained ledger. (**b**) The same ontology as a celestial sphere: each dimension is a constellation whose large marker is its root r_d and whose small markers are its subconcepts S_d. The field starts uniformly dim, and the scan lights only what it matches, so the figure shows the shadow being raised rather than its final state alone. Open data closes completely; interoperability and LegalTech match at the root and in part of their detail; TOGAF, NIST CSF, security and AI match only below the root—an incomplete practice; COBIT, ITIL and DLT stay dark—an absent one. Marker positions are schematic; the counts are exact.

Algorithm 1 states the procedure executed on every run, so that a third party can re-run it at a new cut-off date and compare artifacts.

5. Case Study and Results

The study applies LegalShadow to the public digital ecosystem of the Ecuadorian Judiciary Council (`funcionjudicial.gob.ec`) and validates the resulting shadow against the full ontology (131 classes, 144 relations, 14 cross-cutting), contrasted with a synthetic *positive control*: a 39-component shadow authored to saturate the engine, whose role is to establish that the metric can approach unity on a well-anchored artifact.

Scoring protocol. The ontological scores σ are derived from a versioned rubric that defines five evidence-based maturity levels (0, 25, 50, 75, 100) for each dimension, ensuring transparent and auditable evaluation. Scores lacking supporting evidence are classified as *declared* rather than *observed*, distinguishing institutional self-description from empirically verifiable coverage. Rubrics, evidence sources, and generated shadows are versioned to support traceability and review (Section 7).

The two metrics used in the model have different epistemic roles. While $\text{onto}(d)$ depends on expert assessment and may vary with alternative interpretations of the rubric, $\Psi(d)$ is fully deterministic, being computed directly from the rule set and stored evidence. This separation enables partial replication: reviewers may recompute $\Psi(d)$ from the published artifacts and, if desired, apply alternative $\text{onto}(d)$ values to obtain their own prioritization. Consequently, the framework preserves reproducibility of observed coverage while allowing flexibility in the interpretation of maturity levels.

Evidence corpus. Over 17 dated runs between 2 July and 4 August 2026, the acquisition tier issued 187 read-only requests, of which 170 succeeded (90.9%). Re-hashing every stored snapshot against its manifest yields **170 matches and 0 mismatches**: full integrity of the evidence base.

Three cardinalities appear here and measure different things: a *run* is one pass of the protocol; a *shadow regeneration* is a run that also rebuilt Δ end-to-end, of which there are five; a *ledger entry* records a run, a verification, or a methodological decision, hence their larger number. Latencies (median 788.5 ms, p95 954.4 ms, range 67–2810 ms) confirm the probe imposes negligible load, a design requirement of the ethical protocol [29].

The shadow from the 4 August run contains **34 components, 32 flows and 7 architectural layers** anchored to 11 seed sources, with anchoring *total and referentially closed*: every component carries at least one anchor (mean 2.00, max 4) and all 50 distinct references resolve, with zero dangling identifiers. Signals show a modern portal on a generic content manager still surfacing 16 legacy links (6 JSF, 10 PHP), a confirmed case-consultation SPA with no documented public API, and the total absence of embedded JSON-LD despite declarative references to open data policies—exactly the gap between nominal publication and effective reuse reported in the open government literature [2,3]. The assessment over the four dimensions for which the institution publishes declarative evidence gives a global index of **46/100**, which L places at Level 3 ("Defined"): data and semantic interoperability 28, integration architecture 42, declared transformation 55, regulatory compliance 58. Both this index and the coverage figures reported below were invariant across the five regenerated shadows: the only signal that drifted in the series was the legacy PHP count (13 → 10 between 5 and 30 July), attributable to content editing on the observed portal, and that signal anchors no ontology class, so the coverage set R—and with it Ψ_{global}—did not change.

Applying the engine of Section 3 to the same real shadow (Table 2, Figure 4) gives $\Psi_{\text{global}} = 25.8/73.9 = \mathbf{0.349}$: the shadow materializes only 34.9% of the maturity the ontology prescribes, leaving an aggregate gap of 48.2 points (65.1%). The positive control reaches $\Psi_{\text{global}} = 66.9/73.9 = 90.4\%$ with a residual gap of 7.1. This rules out a ceiling artifact in the scoring engine and shows that the low real coverage is a property of the observed ecosystem rather than of the instrument; it does not, on its own, establish construct validity, since the control was authored by the same team (Section 7).

https://doi.org/10.3390/engproc2026153023

A result visible only through the shared apparatus is that both shadows light almost the same number of stars—50 of 131 for the real shadow against 51 for the control—yet their global coverage differs by a factor of 2.6. Coverage *breadth* is not what Ψ measures: *which* concepts are anchored matters far more than how many. The control spreads its 51 anchors across all ten dimensions; the real ecosystem concentrates its 50 in open data, interoperability and LegalTech, leaving whole dimensions untouched. A naive count of evidenced nodes would have rated the two artifacts equivalent.

Figure 4 renders the same table as three series ordered by decreasing δ—light bars for what the ontology prescribes, dark for what the shadow delivers, dashed line for the difference—so that reading left to right gives the agenda directly. The shape carries the argument: the bar series converge only at the right-hand end, where open data and interoperability sit, and diverge almost completely on the left, where the enterprise frameworks are.

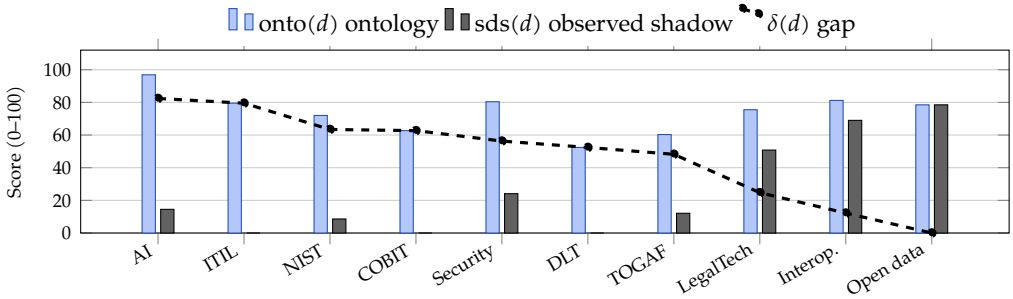


Figure 4. Ontological versus effective maturity of the real institutional shadow, ordered by decreasing gap. The dashed line traces $\delta(d)$ and constitutes the remediation agenda.

Three dimensions have *zero* coverage because neither root nor subconcepts appear in R: service management (ITIL, $\delta = 79.5$), control (COBIT, $\delta = 62.7$) and distributed ledger ($\delta = 52.4$). Four more are lit only below the root—AI, NIST CSF, security posture and TOGAF—which is a different condition and calls for different remediation. The single largest gap is AI integration ($\delta = 82.4$), where a high ontological score (96.9) is almost entirely unrealized. Conversely, open data attains full coverage and interoperability 0.85, reflecting genuine participation in the national open data catalog.

Table 2 reports the full result, one row per governance dimension, ordered by decreasing gap. The column onto(d) gives the maturity the rubric prescribes; *Root* states whether the dimension's root concept was matched at all; $\Psi(d)$ is the coverage the shadow achieves; sds(d) is the prescription discounted by that coverage; $\delta(d)$ is what remains unrealized and therefore sets the remediation order; the last column repeats Ψ for the synthetic control, and the bottom row aggregates by Equation (4). The *Root* column alone already separates the three regimes the metric distinguishes: three dimensions with neither root nor subconcepts, four lit only below the root, and three anchored at both levels.

The two headline figures do not measure the same thing, and that is the finding. It is tempting to read 46/100 against 34.9% as a single quantity assessed twice, and that reading would be wrong. The 46/100 aggregates the *four* dimensions on which the institution itself publishes declarative evidence: data and semantic interoperability, integration architecture, declared transformation, regulatory compliance. Ψ_{global} spans the *ten* dimensions the ontology prescribes. Neither the scales nor the dimension sets coincide, so the comparison is not a difference in degree but a difference in frame: what an institution reports *on its own terms* versus what its digital surface exposes *on the frameworks' terms*. Read that way, the result is sharper rather than weaker. An institution can only report progress on the dimensions it has chosen to instrument; the six dimensions missing from its own assessment—ITIL, COBIT, distributed ledger, AI, NIST CSF and security posture—are, without exception, the

ones in which the ontology finds little or no machine-processable trace. The self-assessment is not merely optimistic in value; it is silent exactly where the deficit is.

6. Discussion

Declarative maturity versus semantic coverage. Although the institution attains a maturity score of 46/100, its observable ecosystem materializes only 34.9% of the governance dimensions defined by the referenced frameworks. The results reveal a gap between declared governance and machine-processable evidence, highlighting dimensions that remain only partially represented or entirely absent on the public digital surface.

A reflexive finding and what it bounds. The main gap identified is the absence of embedded JSON-LD and documented public APIs: the institution lacks the very instrument—structured data—that the shadow needs in order to represent it. This carries a methodological consequence. Coverage is a joint property of the ecosystem and of its publication practices, so any external instrument of this family resolves governance only to the granularity at which the institution has chosen to structure its public surface. A low Ψ therefore cannot distinguish weak governance from governance that is real but undisclosed in machine-readable form, and semantic publication becomes a determinant of measurable coverage in its own right [30]. Two corollaries follow: no refinement of Γ removes this bound—only the institution can, by publishing—and longitudinal use must control for the fact that adopting semantic markup would raise measured coverage with no change in underlying governance.

External validity. The empirical results are specific to a single institution and a one-month observation window of 17 dated runs; values such as 46/100 and $\Psi_{\text{global}} = 34.9\%$ describe the Ecuadorian Judiciary Council and should be interpreted only within that context. While the observed gap between declared maturity and semantic coverage aligns with findings on publication and reuse limitations in open-data ecosystems [2,3,18], its generality is a hypothesis that only a multi-institutional study can test. The transferable element is the instrument itself: the ontology Ω and its governance frameworks remain largely reusable, and adaptation concentrates on localizing the anchoring rules of Γ to national infrastructures.

Transfer takes four steps, of which only the second costs anything. (1) Replace the seed file: the target is already a parameter. (2) Localize the rule table of Γ: roughly a fifth of the rules key on national artifacts—the Ecuadorian catalog, national interoperability schemes—and must be re-expressed against equivalent infrastructure, the one step demanding domain knowledge and the real obstacle to naive reuse. (3) Leave Ω and its rubric unchanged where the frameworks are international, which covers eight of ten dimensions. (4) Compare only Ψ and δ across institutions, never raw signal counts, since portal technologies differ. Comparability rests on holding the ontology fixed while the rule table varies—the discipline that makes urban twins comparable across cities [9,15]—and whether it survives contact with a second jurisdiction is untested.

Limitations. Results represent *observable* maturity and therefore constitute a lower bound on actual institutional capabilities: dimensions such as ITIL or COBIT may operate internally without public trace, so their zero coverage measures published evidence rather than practice. Dimension weights and ontological scores incorporate expert judgment and await independent validation—no Delphi round, content-validity ratio [34], or inter-rater agreement has yet been computed—so the absolute value of δ carries unquantified judgment even though Ψ does not. The synthetic control is a positive control only: authored by the same team to saturate the engine, it bounds instrument artifacts but does not establish construct validity. Signal detection also depends on portal markup stability, which the cryptographic record attenuates without eliminating.

7. Conclusions

LegalShadow applies the digital shadow paradigm to LegalTech ecosystems by combining read-only public web evidence, semantic representation through a governance ontology, and quantitative analysis based on Ψ and δ. Three conclusions follow.

First, declarative maturity and machine-actionable governance are distinct quantities: the observed institution reports 46/100 on the four dimensions it instruments itself, yet materializes 34.9% of the semantics prescribed across the ten dimensions of the ontology, and the two figures differ not only in value but in the set of dimensions they can address at all. **Second**, the deficit is localizable rather than diffuse: the root/subconcept decomposition of Ψ separates absent practices (COBIT, ITIL, distributed ledger, all at zero) from incomplete ones (TOGAF, NIST CSF, security, AI), and those two conditions call for different responses. **Third**, the instrument is auditable end-to-end: every figure reported here is recomputable by a third party from the published rubric, rule table and dated artifacts, which is what turns the exercise into an audit rather than an assessment. The metric therefore offers an operational answer to the verification problem raised for digital twins [12]: ontology–artifact alignment can be measured, bounded, and compared.

8. Policy Implications

The results carry implications for how judicial institutions are assessed and how they report. The **first** is that maturity self-assessment and external auditability are not substitutes. An institution can only report progress on dimensions it has chosen to instrument, and in the observed case the six dimensions missing from its own assessment are precisely those where the ontology finds little or no machine-processable trace; oversight bodies that rely on self-reported indices therefore inherit that blind spot. The second is that transparency policy and data policy are the same policy. Because coverage is bounded by publication practice, commitments to judicial transparency that do not specify machine-readable form produce documents that are public but not auditable at scale—the gap between nominal publication and effective reuse that the open government data literature has documented for a decade [2,3]. The third is that structured disclosure is a low-cost regulatory lever: it changes what can be known about an institution from outside without requiring any reform of how that institution is governed internally, which makes it unusually attractive where reform capacity is scarce [26].

9. Future Research Directions

Four directions follow directly from the limitations. **First**, a multi-institutional study: applying the instrument to additional judicial ecosystems would turn the diagnostic pattern from hypothesis into evidence, and test whether the transfer protocol survives different national infrastructures. **Second**, independent elicitation of ontological scores via a Delphi panel with content-validity ratio [34], removing the expert judgment currently embedded in onto(d) and with it the residual uncertainty in the absolute value of δ. **Third**, automatic ontology alignment to enrich the anchoring Γ [33], reducing the manual effort that localization now demands. Fourth, a behavioral dimension based on procedural performance indicators [14,25], complementing the structural assessment, together with assisted decision support under MAPE-K principles that preserve human authority over interventions [28]. **Finally**, a longitudinal series of dated shadows would allow institutional change to be observed over time rather than at a single cut-off [9,15].

https://doi.org/10.3390/engproc2026153023

10. Recommendations

For judicial institutions seeking to close the measured gap, three publication acts offer high leverage at low cost, none requiring internal governance reform. First, embed `schema.org` JSON-LD in institutional pages: this improves findability [30] and anchors the open data and interoperability dimensions, where the observed institution already performs best and would reach full coverage. Second, document the public services already operating behind the case-consultation single-page application: publishing an API specification converts an opaque service into an auditable one at essentially no engineering cost. Third, publish a machine-readable governance statement declaring frameworks in use, their owners, and review cycles; this minimum artifact enables external observers to anchor ITIL, COBIT, and security posture, converting three zero-coverage dimensions into measurable ones. Together, these acts expand what can be audited from the public surface and enable third parties to build value-added services on judicial data [3,32]. For oversight bodies, the corresponding recommendation is to require machine-readable form explicitly when mandating transparency, since unstructured publication is not auditable at scale.

Author Contributions: Conceptualization, methodology, software, investigation, data curation and writing—original draft preparation, P.M.P.-A.; validation and formal analysis, P.M.P.-A. and E.J.S.-C.; writing—review and editing, and visualization, P.M.P.-A., E.J.S.-C. and V.V.V.-B.; supervision and project administration, E.J.S.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Universidad Politécnica Salesiana under the "Análisis entornos simulados de Clustering y Gestión de microservicios para Edge y Fog Computing redes de IoT" project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All artifacts of this study—the LegalShadow source code, the OWL governance ontology, the versioned scoring rubric that fixes onto(d), the *signal* → *class* rule table that constitutes Γ, the dated JSON-LD shadows and the SHA-256 evidence manifests—are openly available without restriction at Zenodo [35] and at https://github.com/patmic/ArchLab (accessed on 22 March 2026). By publishing the rubric and rule table with the snapshots, two-level replication becomes possible (Section 5): Ψ can be recomputed without acquisition, and onto can be substituted by reviewer values. Code and reproduction instructions are provided at: git clone https://github.com/patmic/ArchLab (accessed on 1 August 2026), docker compose up -d –build. Dashboard available at http://localhost:8080 (accessed on 1 August 2026) (local only, requires prior setup; no public access).

Acknowledgments: The authors acknowledge the open-source ontology engineering ecosystem (`rdflib`, `pyvis`, OWLready2).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rosa, J.; Teixeira, C.; Sousa Pinto, J. Risk factors in e-justice information systems. *Gov. Inf. Q.* **2013**, *30*, 241–256. [CrossRef]
2. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [CrossRef]
3. Attard, J.; Orlandi, F.; Scerri, S.; Auer, S. A systematic review of open government data initiatives. *Gov. Inf. Q.* **2015**, *32*, 399–418. [CrossRef]
4. Paccha-Angamarca, P.M.; Sacoto-Cabrera, E.J. Architecture for LegalTech Governance: A Holistic Multidimensional Framework Integrating AI, Blockchain, Cybersecurity, Workflows and Enterprise Standards. In Proceedings of the 10th Latin American Congress on Mechanical Engineering (COLIM 2026). Institution of Engineering and Technology (IET), Virtually, 25–27 February 2026; Volume 2026, pp. 1–8. [CrossRef]
5. De Haes, S.; Van Grembergen, W.; Debreceny, R.S. COBIT 5 and Enterprise Governance of Information Technology: Building Blocks and Research Opportunities. *J. Inf. Syst.* **2013**, *27*, 307–324. [CrossRef]

6. Paccha-Angamarca, P.M.; Sacoto-Cabrera, E.J.; Velepucha-Bonett, V.V. Structural Digital Twin-Driven Conformance Assessment of LegalTech Governance Ontology: Reproducible Proof of Concept. *Information* **2026**, *17*, 725. [CrossRef]
7. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022. [CrossRef]
8. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. Digital Twin in Industry: State-of-the-Art. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2405–2415. [CrossRef]
9. Shahat, E.; Hyun, C.T.; Yeom, C. City Digital Twin Potentials: A Review and Research Agenda. *Sustainability* **2021**, *13*, 3386. [CrossRef]
10. Brauner, P.; Dalibor, M.; Jarke, M.; Kunze, I.; Koren, I.; Lakemeyer, G.; Liebenberg, M.; Michael, J.; Pennekamp, J.; Quix, C.; et al. A Computer Science Perspective on Digital Transformation in Production. *ACM Trans. Internet Things* **2022**, *3*, 15. [CrossRef]
11. Sepasgozar, S.M.E. Differentiating Digital Twin from Digital Shadow: Elucidating a Paradigm Shift to Expedite a Smart, Sustainable Built Environment. *Buildings* **2021**, *11*, 151. [CrossRef]
12. Karabulut, E.; Pileggi, S.F.; Groth, P.; Degeler, V. Ontologies in digital twins: A systematic literature review. *Future Gener. Comput. Syst.* **2024**, *153*, 442–456. [CrossRef]
13. Casanovas, P.; Palmirani, M.; Peroni, S.; van Engers, T.; Vitali, F. Semantic Web for the Legal Domain: The next step. *Semant. Web* **2016**, *7*, 213–227. [CrossRef]
14. Medvedeva, M.; Vols, M.; Wieling, M. Using machine learning to predict decisions of the European Court of Human Rights. *Artif. Intell. Law* **2020**, *28*, 237–266. [CrossRef]
15. Haraguchi, M.; Funahashi, T.; Biljecki, F. Assessing Governance Implications of City Digital Twin Technology: A Maturity Model Approach. *Technol. Forecast. Soc. Change* **2024**, *204*, 123409. [CrossRef]
16. Diaz-Sarachaga, J.M. Developing an Assessment Governance Framework for Urban Digital Twins: Insights from Smart Cities. *Cities* **2025**, *156*, 105558. [CrossRef]
17. Anshari, M.; Hamdan, M. Enhancing E-Government with a Digital Twin for Innovation Management. *J. Sci. Technol. Policy Manag.* **2023**, *14*, 1055–1065. [CrossRef]
18. Neumaier, S.; Umbrich, J.; Polleres, A. Automated Quality Assessment of Metadata across Open Data Portals. *ACM J. Data Inf. Qual.* **2016**, *8*, 1–29. [CrossRef]
19. Kubler, S.; Robert, J.; Neumaier, S.; Umbrich, J.; Le Traon, Y. Comparison of Metadata Quality in Open Data Portals Using the Analytic Hierarchy Process. *Gov. Inf. Q.* **2018**, *35*, 13–29. [CrossRef]
20. Vetrò, A.; Canova, L.; Torchiano, M.; Minotas, C.O.; Iemma, R.; Morando, F. Open Data Quality Measurement Framework: Definition and Application to Open Government Data. *Gov. Inf. Q.* **2016**, *33*, 325–337. [CrossRef]
21. Zaitoun, A.; Sagi, T.; Peleg, M. Automated Ontology Evaluation: Evaluating Coverage and Correctness Using a Domain Corpus. In *Proceedings of the Companion Proceedings of the ACM Web Conference 2023*; ACM: New York, NY, USA, 2023; pp. 1127–1137. [CrossRef]
22. Abad-Navarro, F.; Martínez-Costa, C.; Fernández-Breis, J.T. Ontology Coverage Analysis through Language Models. *Eng. Appl. Artif. Intell.* **2025**, *162*, 112671. [CrossRef]
23. Seo, S.; Cheon, H.; Kim, H.; Hyun, D. Structural Quality Metrics to Evaluate Knowledge Graphs. *arXiv* **2022**, arXiv:2211.10011. [CrossRef]
24. Bench-Capon, T.; Araszkiewicz, M.; Ashley, K.; Atkinson, K.; Bex, F.; Borges, F.; Bourcier, D.; Bourgine, P.; Conrad, J.G.; Francesconi, E.; et al. A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law. *Artif. Intell. Law* **2012**, *20*, 215–319. [CrossRef]
25. Katz, D.M.; Bommarito, M.J.; Blackman, J. A general approach for predicting the behavior of the Supreme Court of the United States. *PLoS ONE* **2017**, *12*, e0174698. [CrossRef] [PubMed]
26. Li, Y.; Wang, X.; Zhu, Q. Integrating Artificial Intelligence into Service Innovation, Business Development, and Legal Compliance: Insights from the Hainan Free Trade Port Era. *Systems* **2024**, *12*, 463. [CrossRef]
27. Saaty, T.L. How to make a decision: The Analytic Hierarchy Process. *Eur. J. Oper. Res.* **1990**, *48*, 9–26. [CrossRef]
28. Kephart, J.O.; Chess, D.M. The Vision of Autonomic Computing. *Computer* **2003**, *36*, 41–50. [CrossRef]
29. Krotov, V.; Johnson, L.; Silva, L. Tutorial: Legality and Ethics of Web Scraping. *Commun. Assoc. Inf. Syst.* **2020**, *47*, 539–563. [CrossRef]
30. Guha, R.V.; Brickley, D.; Macbeth, S. Schema.org: Evolution of Structured Data on the Web. *Commun. ACM* **2016**, *59*, 44–51. [CrossRef] [PubMed]
31. Bizer, C.; Heath, T.; Berners-Lee, T. Linked Data—The Story So Far. *Int. J. Semant. Web Inf. Syst. (IJSWIS)* **2009**, *5*, 1–22. [CrossRef]
32. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef] [PubMed]

33. Hogan, A.; Blomqvist, E.; Cochez, M.; d'Amato, C.; de Melo, G.; Gutiérrez, C.; Kirrane, S.; Labra Gayo, J.E.; Navigli, R.; Neumaier, S.; et al. Knowledge Graphs. *ACM Comput. Surv.* **2021**, *54*, 71. [CrossRef]
34. Lawshe, C.H. A quantitative approach to content validity. *Pers. Psychol.* **1975**, *28*, 563–575. [CrossRef]
35. Paccha-Angamarca, P.M. LegalShadow. Zenodo, Software (v1.0). 2026. Available online: https://zenodo.org/records/21942315 (accessed on 1 August 2026).

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