SOVRENTIC / RESEARCH

Sovereign Operational Verification

A Change-Aware Assurance Layer for Agentic Resources

Sovrentic

Publication edition 0.2.1 | 10 September 2026

Technical working paper. Not peer reviewed or submitted to a venue.

Abstract

Agentic resources combine probabilistic behaviour, changing dependencies and delegated authority. Identity, assessment, evidence currency and permission to act therefore require distinct but connected decisions. This paper proposes SOV, Sovereign Operational Verification, a scheme for making those decisions inspectable across the lifecycle of agents, tools and skills. SOV separates a free foundation profile for structural verification of declared records from proposed advanced assessment profiles. It connects configuration-bound claims, evidence provenance, delegated authority and status through portable records. Sovrentic is the venture developing the proposed Agentic Certification Authority and its supporting infrastructure, with institutional operation and delegated assessment as complementary deployment models. The technical contribution is a proposed continuity procedure that identifies what evidence may be preserved after change, what must be reassessed and what remains undetermined. Planning and completed assessment are separate stages: selecting a test never establishes its outcome. Unknown impact requires conservative reassessment, and current authorization remains subject to the relying party's policy and status-freshness requirements. Transparency receipts support verification of recorded statements under explicit trust assumptions; neither signatures nor ledgers establish behavioural truth. We specify the object model, threat boundaries and experimental schema fragments, and retain three project-reported prototype limitations, including acceptance of content-free rationales when structured answers remain correct. An exploratory protocol compares dependency-aware selection with full reassessment and manual rules, with a separate temporal study of scheduled and hybrid review. Reduced burden without loss of critical-failure detection is a falsifiable hypothesis. No superiority, accredited scheme or completed interoperability validation is claimed.

Keywords: agentic AI; assurance; verifiable credentials; delegated authority; change impact analysis; credential status; transparency logs; conformity assessment; digital sovereignty; AI governance

Executive summary for institutional readers

An organisation permitting a software agent to act on its behalf faces five questions that are routinely collapsed into one. Who is this agent and who controls its keys? What authority was delegated, for what scope and how long? What was assessed, against which profile and version? Is that assessment current for this context? And may this specific action proceed under our own policy? Conflating them is the primary trust and liability risk in agentic deployment, because each has a different owner, a different evidentiary basis and a different failure mode.

SOV is proposed as a way to keep those questions separate while keeping them linked, in records designed for independent inspection. Historical verification can remain possible after operator exit; current authorization also requires fresh status from an authorised source.

The design follows from one observation about the market. Agent certification is no longer hypothetical. AIUC-1, published by the Artificial Intelligence Underwriting Company, publishes a structured requirement set across six sections with an authorised third-party audit route [37], [40], and maintains certification through a documented annual and quarterly cycle [38]. Singapore's Infocomm Media Development Authority published a governance framework addressing agentic systems in January 2026 [22], [23]. The United States National Institute of Standards and Technology launched an AI Agent Standards Initiative in February 2026 [20]. The gap this paper addresses is therefore not the absence of certification. It is the need for an explicit, portable account of how evidence applies after a model, tool contract, permission, subagent or hosting context changes. This review does not establish that competing schemes lack such mechanisms.

The proposed differentiator is an inspectable decision about selective evidence reuse after change. Its incremental value over manual rules, scheduled reassessment and hybrid approaches is unvalidated. Existing competitors may combine these mechanisms; the experiment compares defined procedures, not presumed limitations of their products.

This paper is written to be checked rather than believed. It labels which claims are normative requirements, which are Sovrentic design choices, which are hypotheses, which are implemented in the local prototype, and which are targets. It carries the prototype's known failures into the body of the argument. Readers evaluating this as an investment, a partnership or a standards input should treat §12 as the decision point.

1. Introduction and problem statement

1.1 The delegation gap

Software has always acted on behalf of people. What is new is the combination of open-ended objective pursuit, tool invocation with real side effects, and a probabilistic decision process whose behaviour is not fully determined by its declared configuration. An agent that files a ticket, moves money, updates a record or provisions infrastructure exercises authority, and the organisation that granted it remains accountable.

Six properties bear on whether such an exercise is acceptable, and each is commonly mistaken for another.

Model or software identity. Which artefact is this, and can its integrity be verified? Workload identity [46] and provenance frameworks [44], [45] address this. They say nothing about behaviour.

Delegated authority and scope. On whose behalf, under what mandate, for how long, with what limits? Delegation and token-based authorization address this [18], [32]. A valid token proves permission was granted, not that granting it was wise.

Behavioural and security assessment. Under adversarial and ordinary conditions, what does this configuration do? Benchmarks and certification schemes address this [28], [30], [37]. A passing result is a statement about a sample, an environment and a moment.

Organisational and regulatory applicability. Do the obligations binding this organisation, in this use case and jurisdiction, attach to this deployment, and does the available evidence bear on them? Management-system standards and risk frameworks structure this [7], [8], [9]. None converts a technical test into a legal conclusion.

Runtime authorization. May this concrete operation, against this resource, with this payload, now, proceed? Policy engines and gateways address this [47]. An allow decision records a decision; it does not prove a correct business outcome.

Current, revocable, change-sensitive status. Is everything above still true? This property has the fewest instruments and the shortest half-life.

The gap addressed here sits between the third property and the sixth. Assessment produces a conclusion about a configuration observed at a point in time; operation changes that configuration continuously. A signed identity does not prove safe behaviour. A passed test does not grant authority. A credential does not satisfy a law. An authorization decision does not prove a correct outcome. And a certificate issued in March describes a system that may not exist in June.

Instruments exist for each property. This paper studies their integration into an inspectable continuity decision; it does not establish an exhaustive absence of comparable prior systems.

1.2 Periodic re-testing and the comparative question

AIUC-1 publishes a maintenance cycle with annual certification and quarterly adversarial testing [38]. This is an existing continuity mechanism, not a hypothetical competitor. Its published maintenance page does not establish the absence of additional change-management controls.

A purely calendar-based baseline can leave changes unreviewed between scheduled dates. A hybrid can combine scheduled review and event-triggered reassessment. SOV's proposed contribution is a portable rationale for the selection itself: which assumptions still hold, what evidence may be reused, which checks remain mandatory, and what is unresolved.

Hypothesis H1. Dependency-aware and assumption-aware selection may reduce review burden relative to full reassessment and expert manual rules while retaining detection of critical failures. A separate temporal experiment tests scheduled and hybrid alternatives. These are hypotheses, not product comparisons or established results.

Section 12 defines the experiments. If maintaining the graph costs more than the reassessment it saves, or if it omits a critical failure detected by the full test set, the proposed advantage is not supported.

1.3 Contributions and non-contributions

This paper contributes a data model with explicit provenance classes, directed impact-graph semantics, an implementable continuity procedure, draft JSON Schemas, a change taxonomy with conservative defaults, a threat analysis, a role and conflict model, and an evaluation protocol with a defined primary endpoint.

It does not contribute a completed comparative evaluation, an interoperability conformance report, an accredited scheme, a legal determination or a trust score. The refusal to compute a score is a design commitment defended in §3.4.

1.4 Terminological note

SOV and the profile names in this paper are introduced by this paper as a proposed scheme. The existing prototype and its architecture documents use a different vocabulary built around programme, agent, configuration, instance and execution, at revision 0.2 for the executable prototype [56] and 0.3 for the reference architecture [55]. The SOV naming is a proposal for presentation; the object relations, consistent across both vocabularies, are the substance. No part of the profile hierarchy in §3.2 is implemented, offered, priced or accredited.

2. Background and related work

Source status is labelled inline and consolidated in Appendix G.

2.1 Identity, delegation and workload authentication

W3C Decentralized Identifiers [14] provide interoperable identifiers and verification relationships, including capability invocation and capability delegation. Key rotation behaviour is not specified by DID Core in general; it is a property of the chosen DID method and its resolver, and this paper treats method selection as an open product decision rather than a solved one.

The Verifiable Credentials Data Model 2.0 became a W3C Recommendation on 15 May 2025 [13] and provides an extensible model of issued and verified claims including evidence and status. Securing mechanisms are separate specifications [15], [16], and status distribution is addressed by a further one [17]. All are carriers. A verifiable credential proves that an issuer signed a claim; it does not make the claim true or the issuer competent.

SPIFFE and SPIRE [46] provide workload identity and runtime attestation patterns; OAuth 2.0 and OpenID Connect [18] provide delegated access and federation. These are mature and should be integrated rather than reinvented. Workload identity answers which process is calling, not what that process is competent to do.

The agent-specific identity space is unsettled in a way often described as more mature than it is. Several agent-specific proposals named in the project library are individual Internet-Drafts. This paper has not established the status of every initiative and does not generalise that observation to the entire IETF. Internet-Drafts are valid for at most six months and are explicitly inappropriate to cite as other than work in progress. Several drafts named in the project's source library [60] could not be independently confirmed and are recorded as unverified in Appendix G rather than cited as support.

The NCCoE concept paper on software and AI agent identity and authorization, published 5 February 2026 with public comment closing 2 April 2026, frames the operative question as whether existing identity standards can be adapted to agents rather than replaced, and covers identification, authorization, auditing, non-repudiation and prompt-injection controls [21]. It is a concept paper proposing a project, not guidance.

Open. Multi-hop delegation across organisational boundaries, and binding a delegation chain to an assessed configuration rather than to a principal alone.

2.2 Transparency, provenance and supply chain

RFC 9943 [10], published on the IETF Standards Track in June 2026, defines the SCITT architecture for single-issuer signed-statement transparency using verifiable data structures. RFC 9942 [11], also Standards Track, specifies COSE Receipts, which prove properties of a verifiable data structure to a verifier and provide CBOR encodings for Merkle inclusion and consistency proofs. Together these mean the architecture and the receipt format are both published standards; the SCITT reference API specification remained in the RFC Editor queue at the time of writing [19], so an implementer should treat the architecture and receipts as settled and the API surface as not.

SBOM practice [44] and SLSA [45] provide dependency and provenance evidence for software supply chains. Their adaptation to prompts, policies, tools and models is plausible and proposed here, but equivalence should not be claimed: a package dependency tree and an agent's tool and data surface differ in observability, not merely vocabulary.

RFC 9334 [12] establishes the RATS conceptual model, separating the entity producing evidence about an environment, the entity appraising it, and the entity relying on the appraisal. That separation is directly reusable. Attesting an orchestrator does not attest the behaviour of a remote model it calls.

Model cards [35] and datasheets [36] provide human-readable intended-use and limitation metadata. They are declared documentation, not test results.

Open. Cryptographic evidence binding test inputs, environment, policy decision, evaluator identity and result into one verifiable object; and treatment of dependencies that are not observable at all.

2.3 Governance and risk frameworks

ISO/IEC 42001 [7] specifies an AI management system for organisations; ISO/IEC 23894 [8] provides AI risk management guidance; the NIST AI Risk Management Framework [9] organises work into Govern, Map, Measure and Manage and is explicitly voluntary. All three are organisational instruments. None certifies an agent, and mapping evidence to them does not produce a compliance conclusion. The ISO normative texts were not accessed or reproduced for this paper; the NIST framework is publicly available. Proposed mappings require validation against the applicable source and version.

Singapore's IMDA published the Model AI Governance Framework for Agentic AI on 22 January 2026 [23], an early published governance framework specifically addressing agentic systems. Its publisher and several legal commentaries describe it as the world's first of its kind; this paper attributes that characterisation rather than asserting it. The framework is voluntary, applies to organisations deploying agentic AI whether built in-house or third-party, and is structured around four dimensions: assessing and bounding risks up front, ensuring meaningful human accountability, implementing technical controls and processes, and enabling end-user responsibility. Version 1.5 was published 20 May 2026 and updated 5 June 2026 [22].

NIST, through its Center for AI Standards and Innovation, launched an AI Agent Standards Initiative on 17 February 2026 [20], organised around industry-led standards, open-source protocol development, and security and identity research, accompanied by a request for information on agent security.

Regulation (EU) 2024/1689 [1] entered into force on 1 August 2024 with phased application [26]. Its obligations attach to systems, roles and uses, not to agents generically. GDPR [2], NIS2 [3], the Data Governance Act [4], the Data Act [5] and the eIDAS framework [6] bear on parts of the evidence surface in §9 and §10. None makes a certificate into a compliance conclusion, and this paper treats no mapping as a legal determination.

Open. The bridge from technical evidence to an applicability determination, addressed in §10.

2.4 Agent certification and assessment schemes

This comparison uses the scheme publisher's own descriptions and the auditor's announcement; it is not a functional benchmark.

AIUC-1 is published by the Artificial Intelligence Underwriting Company. Its published requirement index organises the standard into six lettered sections: A Data & Privacy, B Security, C Safety, D Reliability, E Accountability and F Society [37]. Counting the published index gives fifty-three listed requirements, of which two are marked Retired, leaving fifty-one active, which reconciles the figure commonly quoted in secondary commentary. Requirements are marked mandatory or optional. This paper does not repeat the control counts circulating in secondary sources, because they could not be confirmed from the primary index.

The publisher describes annual certification, quarterly red-teaming and periodic updates to the standard [38]. Schellman announced an authorised audit route in February 2026 [40]. The scheme's use of "accredited auditor" should be read in its stated programme context; this announcement alone does not establish statutory accreditation of SOV, ISO accreditation equivalence or recognition by a national accreditation body.

The publisher also supplies an EU AI Act crosswalk [39]. This is the publisher's interpretation, not proof of legal equivalence, universal compliance or market leadership. Its applicability requires independent, role-specific assessment.

The market therefore contains a concrete certification comparator. Public documentation establishes a maintenance schedule; it does not establish that the scheme lacks event triggers, change-specific review or selective evidence reuse.

SOV's proposed differentiator must be assessed through explicit procedures and, where authorised, functional comparison. Section 12 distinguishes a static selection study from a temporal comparison of scheduled and hybrid review.

CertAI [28], published in the ICAART 2026 proceedings, proposes a certification framework generating verifiable certificates from structured metadata across security, privacy, ethics, robustness, transparency and fairness, complemented by a benchmark that populates that metadata by probing agents, and reports that fairness and transparency are the weakest dimensions across evaluated agents. This is peer-reviewed conference work and the closest academic analogue to the present proposal. It is also a useful contrast: the framework computes numerical dimension values and derives a trust score. §3.4 takes the opposite position.

TrustBench [30] proposes a dual-mode framework combining a trust benchmark with a toolkit agents invoke before acting, intervening after an action is formulated but before execution. It was accepted at an AAAI 2026 workshop; that venue is distinct from the AAAI main conference. Its reported reduction in harmful actions is a result to reproduce, not a property Sovrentic may claim; workshop status alone does not establish the absence or quality of peer review.

Commercial governance platforms including Credo AI [52] and Holistic AI [53] advertise inventories, dependency graphs, policy management, evidence workflows and continuous evaluation with agent coverage. These are the vendors' own descriptions. No functional comparison was performed and none is implied. Substantial overlap exists, and any differentiation claim requires comparative evidence under authorised execution of a competitor product, not a reading of public pages. No patent search or legal novelty analysis has been performed.

Open. Whether event-triggered dependency-aware continuity outperforms scheduled reassessment at acceptable cost. This is the entire proposition.

2.5 Ledgers, registries and agent reputation

ERC-8004 [41] remains a Draft Ethereum Improvement Proposal, proposing identity, reputation and validation registries for agents. It is a live ecosystem and should be evaluated as a possible integration surface, not a dependency.

It also supplies useful evidence. A cross-chain empirical study covering Ethereum, BNB Smart Chain and Base through 13 May 2026 reports that the Reputation Registry as deployed cannot function as a trust signal: recorded values are not commensurable, feedback is rarely grounded in verifiable interactions, and reputation is manipulable at minimal cost. The study reports coordinated Sybil behaviour among 73.5, 59.2 and 90.6 percent of reviewers across the three chains respectively, and that after removing Sybil-flagged feedback, 15.8, 77.9 and 86.8 percent of rated agents retained no valid feedback [31]. This is cited as a preprint, without an independently confirmed peer-review status, and the findings are specific to the registries studied in that work over that period. Cited with those limits, it supports the position in §8: on-chain presence establishes integrity and inclusion, not competence, authority or current validity.

Sigstore [49] provides an operated transparency log with witness-based consistency checking, against which any anchoring proposal should be compared.

Open. Whether cross-organisation transparency requires a public chain at all, given that [10] and [11] now standardise the underlying properties.

3. Design goals, assumptions, non-goals and contributions

3.1 Design goals

G1. Separation of concerns as a first-class property. Identity, delegation, assessment, evidence, credential status, applicability and authorization are distinct objects with distinct authors, lifetimes and failure modes. The architecture must make conflating them difficult rather than merely discouraged.

G2. Continuity as the central mechanism. The system must answer, reproducibly and inspectably, what remains valid after a change, what must be reassessed, and what is undetermined.

G3. Portability that survives the operator. A relying party must verify a record without the platform operator's cooperation, using a documented verifier and published test vectors, and must reconstruct permitted history elsewhere.

G4. Bounded claims. Every issued record carries scope, exclusions, evidence class, evaluator and issuer roles, validity interval and next-review triggers. Absence of a stated exclusion must not read as coverage.

G5. Institutional neutrality. The software operator must not be required to become the substantive authority of every programme it hosts.

G6. Fail-closed under uncertainty. Unknown changes, incomplete observation, stale status and unresolvable dependencies narrow authority rather than preserve it.

3.2 The proposed scheme hierarchy

Proposed as a scheme design, not an existing offering.

SOV, Sovereign Operational Verification, is the overall schema, protocol and credential family. It names a record structure and verification procedures, not a quality judgement.

SOV-1 Foundation Verification is a bounded entry profile intended to be free at the point of use. It verifies structural completeness, internal consistency, digest binding, identity and version references, declared authority envelope, operating context, declared composition, evidence references and current record status. It is not independent safety certification, legal certification, accreditation, a compliance determination, universal trust, competence proof or a trust score. A resource can hold SOV-1 and be dangerous.

SOV-Agent, **SOV-Tool** and **SOV-Skill** are proposed resource-specific advanced profiles requiring paid assessment, evidence review, reproducible testing, authorised assessors, defined issue and renewal rules, monitoring and revocation. None exists, is accredited, is available, or has an authorised assessor. They are differentiated by resource type and assurance scope, not by position on a numerical ladder, because a ladder invites the inference that a higher rung means a safer agent.

The rationale for the split is adoption friction against defensibility. Structural verification can be automated and costs little per record. Substantive assessment requires reserved test material, calibrated evaluators and human review, which cost money and cannot be given away. The commercial hypothesis, recorded in Appendix F, is that the first creates the surface on which the second becomes purchasable. It is untested.

3.3 Assumptions

A1. The submitter may be careless or adversarial. **A2.** Some dependencies are unobservable. A remote model behind a stable alias may change without notice; the architecture must express this rather than paper over it. **A3.** Behaviour is probabilistic and non-stationary. Repetition does not make cases independent. **A4.** The relying party owns its risk appetite, legal interpretation and authorization decisions. **A5.** Any evaluator, human included, may be wrong, biased or captured. Assessor authority is a trust boundary, not a guarantee of impartiality.

3.4 Non-goals

Not attempted: a universal trust score or safety number; automatic approval of production actions; a legal compliance determination; exhaustive behavioural prediction; a new blockchain or token; published equivalence mappings between governance frameworks; or any claim that a technical credential discharges an organisational obligation.

The trust-score non-goal warrants defence, because the market pulls the other way and at least one peer-reviewed scheme computes one [28]. A composite score alone cannot preserve all exclusions, provenance, observation windows and unresolved dimensions. Scores may be useful inside a specified assessment, but SOV does not use a universal score as a substitute for the scoped evidence record. Calibration and incentive effects require evaluation rather than assumptions about competing assessors.

3.5 Contributions

#	Contribution	Proposed design element	Reused foundations	Hypothesis
C1	Resource-neutral assurance model with scoped verification records and configuration-bound advanced assessment claims	The configuration-bound competence unit; separation of declared value, measured content and unobservable external dependency	Credential carriage [13], evidence digesting, status distribution [17]	That this granularity suits institutional adoption
C2	Proposed separation of distinct object families with distinct authors and lifetimes	Applicability as a distinct signed object with its own review authority, separate from both assessment conclusion and authorization	Each separation individually appears in prior work	That enforcing all seven at once is operationally tractable

#	Contribution	Proposed design element	Reused foundations	Hypothesis
C3	Change-impact and continuity model with directed graph semantics, conservative defaults and an emitted undetermined set	The assumption-validity condition on reuse; the rule that a missing edge never licenses reuse	Dependency and impact analysis are established	H1 in full, which §12 is designed to refute
C4	Privacy-preserving transparency pattern publishing batch commitments while retaining canonical evidence off-chain and off the authorization path	Deliberately very little	SCITT architecture and receipts [10], [11]; witnessed logs [49]	That cross-organisation verification demand justifies anchoring at all
C5	Bridge from technical evidence to organisational decisions in which coverage assertions are separately authored, reviewed and versioned, and framework change invalidates a mapping review without automatically revoking a competence credential	Decoupling mapping validity from credential validity, with issuer policy governing propagation	Control crosswalking is established practice [25]	That the decoupling matches how audit and GRC functions work

The candidate moat is not the impact graph, the credential format, an anchoring scheme or a dashboard, each of which is reproducible by a competent competitor. The candidate moat is the accumulated, versioned graph of claims, evidence lineage, material-change rules, decision rationales, test vectors and institutional verifier adoption. This is a commercial hypothesis requiring evidence, not a demonstrated asset.

4. SOV data model and formal specification

4.1 Scope of normative language

This section uses `MUST`, `MUST NOT`, `SHOULD` and `MAY` in the RFC 2119 sense, applied **only to the proposed SOV profile**. This language expresses what a conforming SOV implementation would be required to do under the profile proposed here. It does not make Sovrentic a standards authority, does not bind any external party, and creates no obligation outside a voluntary adoption of this profile. Where this paper describes existing standards, it reports them and does not restate their normative requirements.

4.2 The unit of assurance

An advanced competence credential MUST refer to a bounded competence demonstrated over an identified configuration under identified conditions. A SOV-1 foundation record instead refers to completed structural checks on the declared resource record; it makes no competence claim. Five identifiers are therefore separated [55]:

Identifier	Meaning	Owned by
Programme	Competence, rubric, criteria and version approved by the issuer	Scheme owner or issuer
Agent	Logical identity and the entity controlling its keys	Resource operator
Configuration	Model, instructions, tools, permissions, data sources and orchestration in scope	Submitter, with declared assurance level
Instance	Observed process or deployment, its key, and measurement coverage	Resource operator, observed by platform
Execution	Request, authorised steps, observed operations, outcome	Enterprise control point

The configuration-instance distinction carries the honesty burden. A configuration digest identifies declared content; it does not prove a remote endpoint serves that content. A model alias may remain constant while the served model changes [56]. Policy MUST therefore bound the guarantee offered, MUST require version identification where a provider exposes it, and MUST force reassessment or suspension on unknown alteration.

4.3 Core objects and ownership

The objects form linked records with distinct owners and lifetimes. The table and Figure 1 define those relationships; they are not a single chain in which assessment grants applicability or authorization.

Links denote inputs, not ownership. In particular, an `ApplicabilityDecision` is authored by the relying party or a reviewer it authorises for that framework and organisation, never issued by the assessment engine and never owned by the credential. An `AuthorizationDecision` is authored by the relying party's control point and is bound to a concrete request, audience, resource, action, configuration digest, policy version and expiry.

Object	Author	Lifetime	What it does not establish
`AgentIdentity`	Resource operator	Long; key rotation is method-dependent	Behaviour or competence
`DelegationChain` / `Mandate`	Grantor organisation	Short; explicitly bounded	That the grant was prudent
`AssessmentProfile`	Scheme owner	Versioned	That the profile is adequate
`EvidencePackage`	Identified collector	Bound to observation window	That coverage is complete
`AssessmentResult`	Authorised assessor	Bound to profile and configuration	Authority, or absence of vulnerability
`TrustManifest`	Resource operator / submitter	Versioned declaration	Truth of declared values or runtime identity
`SOVCredential`	Issuer	Validity interval and scoped configuration	Truth beyond the stated verification or assessment
`CredentialStatus`	Issuer or delegated authority	Freshness window	The assessment outcome, which is a separate field
`ImpactGraph`	Platform, with authored edges	Versioned	Completeness of the dependency set
`ApplicabilityDecision`	Relying party or authorised reviewer	Own validity	A legal conclusion
`AuthorizationDecision`	Relying party control point	Single operation	A correct business outcome

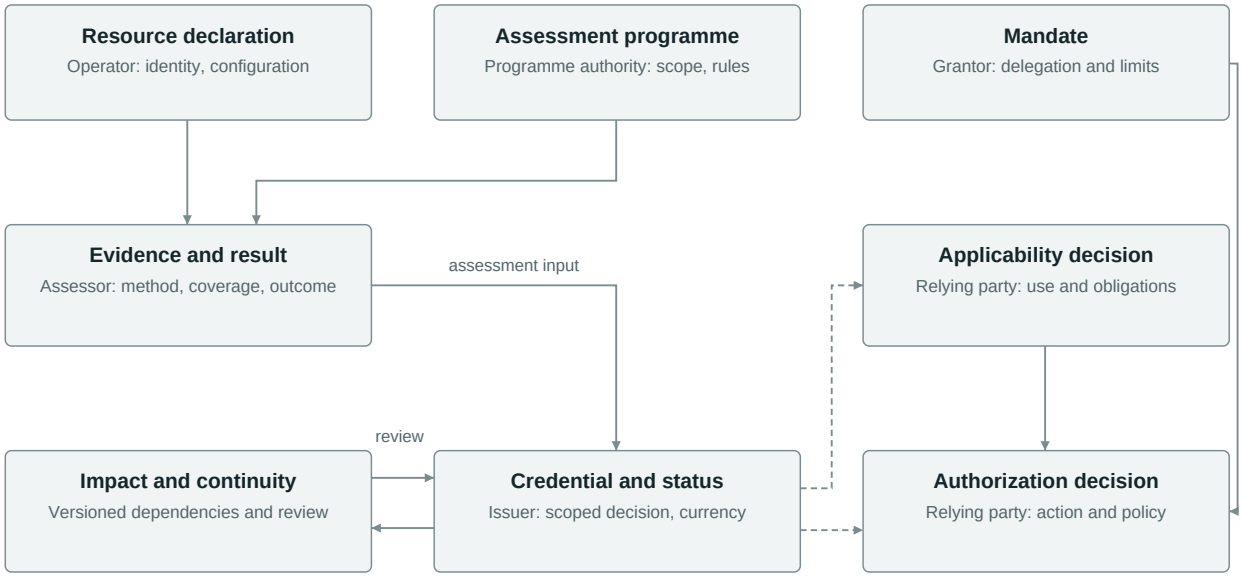

Figure 1: Ownership and decision boundaries. Dashed edges are evidence inputs to relying-party decisions, not transfers of authority. Configuration and evidence changes also feed the impact record.

Dotted arrows mark inputs that cross an ownership boundary: the credential and its status are inputs to decisions the relying party makes and owns.

4.4 ImpactGraph semantics

The graph is directed. Let `N` be the set of configuration nodes (model, instruction, tool, data source, permission, orchestration element, key, jurisdiction) and `C` the set of claims.

An edge `u → v`, where `u ∈ N` and `v ∈ N ∪ C`, means **"v depends on u"**, equivalently "u supports v". Every edge carries an author, a rationale and a confirmation level in `{confirmed, partial, asserted}` [55].

Impact propagation is therefore **forward reachability along edge direction** from the set of changed nodes. A claim is affected if it is forward-reachable from any changed node. The continuity procedure in §7.3 uses this direction and no other.

Two properties are stated as rules, not implementation details:

- Absence of an edge `u → v` MUST NOT be treated as evidence that `v` is independent of `u`. An incomplete graph is indistinguishable from a graph with no dependency.
- A trace that did not exercise a tool does not demonstrate independence from it. Observed dependency is partial information about the dependency set, never a closure over it.

4.5 Provenance classes and status semantics

Every assertion carries an explicit provenance class, directly or inherited from its enclosing object. This class records the asserted evidentiary role; it does not authenticate itself. A verifier must also authenticate the responsible actor, method, scope and supporting evidence. Identity and administrative metadata are covered by their enclosing signed record.

Class	Meaning	Verifier may conclude
`declared`	Asserted by the submitter	That the submitter asserted it, digest-bound
`measured`	Claimed observation by an identified method in an identified environment	An authenticated report of an observation within stated coverage; truth and instrument integrity still require appraisal
`assessed`	A conclusion by an identified assessor under an identified profile	That the assessor concluded it, subject to profile exclusions
`issued`	Bound by the issuer's signature and authority	That the issuer takes responsibility within stated scope
`relied`	A decision by the relying party	Only that the relying party decided
`undetermined`	Explicitly not established	That coverage is absent, which is information

A record MUST NOT assert `measured` provenance without an identified collection method. Structural validation checks presence and type; semantic validation must authenticate the method, evidence, observation scope and responsible actor. A string labelled `measured` is not proof that measurement occurred. The Appendix A fragments do not implement this full appraisal.

Status states are `VALID`, `REVIEW_REQUIRED`, `EXPIRED` and `REVOKED`.

`VALID` means that the issuer asserts the credential's current status, within its stated scope and within the defined freshness window, as of the stated instant. It does not mean the resource is safe, competent, compliant or suitable in general. `REVIEW_REQUIRED` means a material change has been observed whose impact is unresolved; the credential remains verifiable as a historical record but MUST NOT support new sensitive operations. `EXPIRED` means the validity interval ended without renewal. `REVOKED` means withdrawal by the issuer or a delegated authority, with a reason code.

`UNKNOWN` is not a credential state. It is a local determination by a verifier that could not obtain status within the acceptable age, and a verifier MUST fail closed on it for sensitive operations rather than fall back on a previously cached state. An authenticated cached response within the permitted age can nevertheless predate a later revocation; this is a bounded staleness risk, distinct from UNKNOWN. Section 9.4 records this distinction.

Where a human-readable presentation label is used, "SOV-1 Verified" is a presentation label denoting that structural verification was performed and its record is currently `VALID`. It is distinct from the machine status value and MUST NOT be rendered without the accompanying scope and exclusions.

4.6 Digests, canonicalisation and versioning

A conforming implementation MUST specify, for each digest field, the canonicalisation rule and the covered field set. This paper proposes JSON Canonicalization Scheme (RFC 8785) [27] with SHA-256, and defines:

- `manifestDigest` covers the `schemaVersion`, `resource`, `authorityEnvelope`, `operatingContext`, `declaredComposition` and `unobservableDependencies` members, excluding `status`, `validity` and all issuer-side fields, so that a status change does not alter the manifest identity.
- `assessmentDigest` covers the assessment report as a whole, including cases, results, criteria, dataset digests, environment, evaluator identity, repetitions and limitations [56].

Versioning follows three rules. Any change to a digest-covered field MUST produce a new record version rather than mutate the existing one, and prior versions MUST remain resolvable so a historical decision stays verifiable. A verifier MUST pin the schema versions it accepts and MUST reject a record of an unaccepted version rather than interpret it best-effort; the prototype demonstrates this by rejecting v1 credentials at a v2 verifier [56]. Profile versions are independent of schema versions, and a credential MUST name both.

4.7 Relationship to W3C Verifiable Credentials, and the status of "schema"

Appendix A.1 supplies experimental structural schemas for common definitions, the SOV-1 record payload, a status descriptor and a continuity-plan payload. The advanced credential is illustrated in A.3; its complete schema remains future work. They are labelled experimental. They check selected structural constraints. Cross-record binding, evidence authenticity, semantic consistency, signature validation and current status acceptance require separate checks. They are not a normative specification, have no test-vector suite, and have not been validated against any independent implementation.

A mapping onto W3C VC 2.0 [13] with a fixed JOSE profile [15] is proposed, with Bitstring Status List [17] as a status-distribution candidate. The mapping is proposed and untested. The prototype's signed JSON serves demonstration only and implements neither canonicalisation, nor Data Integrity [16], nor JOSE, nor COSE, nor DID resolution [56]. No conformance claim is made and none should be inferred from Appendix A.

5. Sovrentic architecture and the meaning of "Agentic Certification Authority"

5.1 Components and trust boundaries

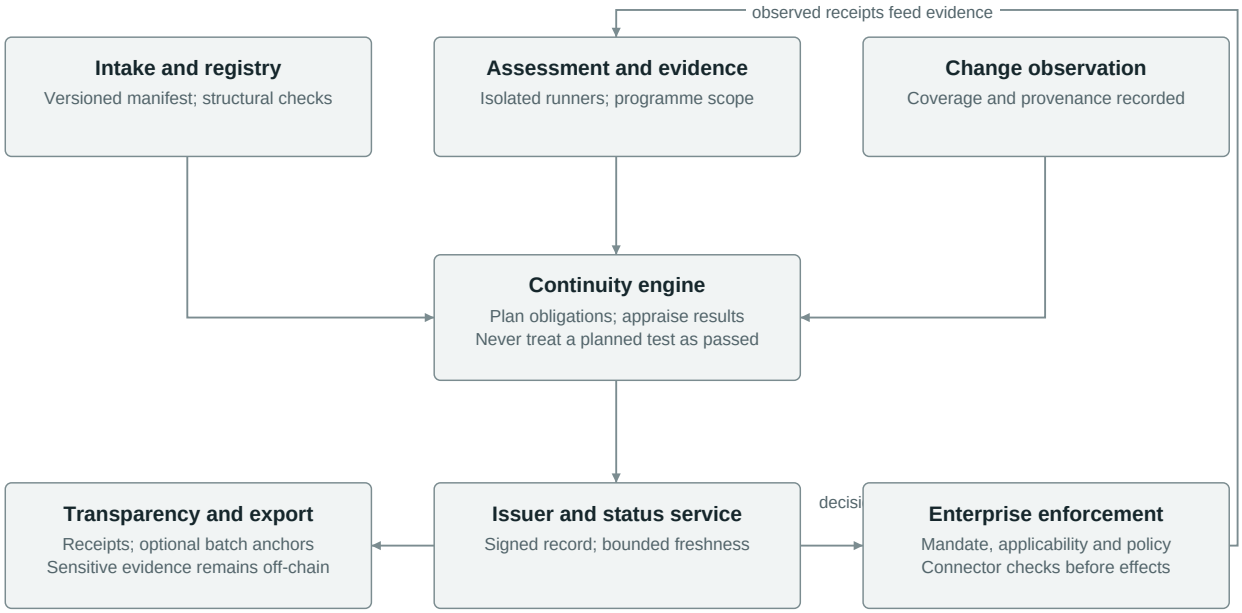

Figure 2: Target architecture. The central engine separates planning from result appraisal. Issuance and fresh status inform enterprise policy; transparency remains outside the execution path.

The diagram represents the target. In the prototype, controlled execution, change observation and external anchoring do not exist [55].

Component responsibilities are as follows. The **registry** holds immutable versioned records with prior versions resolvable. The **manifest validator** normalises, validates, digests and enforces provenance consistency. The **evidence store** holds encrypted objects with digests, provenance, evaluator identity, environment and observation window recorded transactionally alongside. **Assessment runners** are isolated processes with limited permissions; deterministic rules evaluate objective properties, specialists validate rubrics and cases, and any decision depending on a model classifier requires calibration, disagreement analysis and review. The candidate controls neither the reserved set, nor the official result, nor the assessor key [56]. The **validity engine** maintains the versioned relations of §4.4. The **credential and status service** issues signed records and distributes status with explicit freshness semantics. The **enterprise gateway** belongs to the relying party and decides whether a concrete operation proceeds, granting the tool a right limited to request, resource, action, configuration and expiry; the agent MUST NOT hold alternative credentials that bypass this control point. The **verifier** is documented and portable, with published test vectors, exposing scope, evidence references, limitations and decision history without exposing sensitive evidence.

5.2 What "Agentic Certification Authority" means in this paper

Sovrentic is developing the proposed SOV scheme and an Agentic Certification Authority for agents, tools and skills. In this operating model an accountable issuer makes a scoped decision under a published programme, supported by identified assessment and review. SOV-1 is limited to structural verification of a declaration; advanced profiles would assess bounded capabilities and controls.

The term describes an assurance role above identity infrastructure. A PKI certification authority binds keys and identities under a certificate policy; SOV would bind an identified resource configuration to a scoped verification or assessment result. The proposal does not claim WebPKI membership, legal designation, accreditation or authority to confer professional qualifications.

Two proposed deployment models preserve the same technical separation. In first-party operation, Sovrentic owns the SOV scheme and may issue records through an accountable decision function. In institutional operation, an organisation runs or governs a programme using the infrastructure, with its own authorised issuer and assessors. Institutional licensing is an additional business route, not a requirement that Sovrentic relinquish its certification role.

Role	Responsibility	Proposed allocation
Scheme owner	Rules, versioning, appeals and programme admission	Sovrentic for SOV; partner for its separately governed programme
Programme authority	Rubrics, criteria and assessor authorisation	Accountable programme governance
Issuer	Signed scope, issuance decision and status	Sovrentic or an explicitly authorised institutional issuer
Assessor	Evidence appraisal under an identified programme	Authorised assessment function with disclosed independence
Platform operator	Software, storage and operational controls	Sovrentic or institutional operator
Transparency service	Log operation and receipts	Operator and, where adopted, independent witnesses
Verifier	Record and status validation	Relying parties or independent implementations
Resource operator	Resource configuration and keys	Developer or customer organisation
Deployer	Deployment purpose and environment	Deploying organisation
Relying party	Applicability and action authorization	Organisation accepting a record

These are accountable functions, not necessarily ten legal entities. Separation-of-duties rules must prevent a candidate from controlling the reserved tests or approving its own assessment. Issuance review, conflicts, assessor authorisation, appeals and compromise recovery must be documented. A fee funds assessment and operation; it does not purchase a positive result. A signature authenticates a key's assertion and does not establish assessor impartiality, adequate methodology or a legal allocation of liability [56].

5.3 Execution records and their limits

Three objects are produced around an operation, with different authors and different evidentiary reach [55]. An admission receipt proves a decision was recorded. An operation receipt describes what a connector observed. An assessment

result judges that behaviour within a method. A gateway signature demonstrates that the gateway decided, not that the business outcome was correct.

The same reasoning applies to human oversight. A signature can demonstrate that an authorised key approved an identified request. It does not demonstrate that the human understood the risk, had time to intervene, could have interrupted the operation, or reviewed the result [57]. Such evidence supports a bounded authorization control at partial strength; oversight effectiveness remains `undetermined`.

6. Lifecycle

6.1 The sequence, and what SOV-1 covers

1. A submitter uploads a machine-readable Trust Manifest or resource package.
2. The platform normalises and validates the declaration, rejecting ill-formed or provenance-inconsistent records.
3. The resource receives a stable, versioned identifier.
4. The platform records declared identity and version, purpose, authority envelope, operating context, declared composition of model, tools and data sources, unobservable dependencies, and evidence references, each with its provenance class.
5. The manifest digest is computed under the stated canonicalisation rule and bound to the record.
6. Structural checks are performed: completeness, internal consistency, digest binding, identifier stability and status resolvability.
7. The issuer signs the scoped foundation record, including issuer identity, profile version and structural-check outcome. An authenticated status service binds responses to that record and publishes freshness requirements. These are target issuance steps; a local upload preview alone does not complete them.

Steps 1 through 7 constitute SOV-1 Foundation Verification. Completion requires the signed foundation record and authenticated status publication. Successful parsing alone remains an illustrative preview.

Advanced profiles extend this:

8. Evidence is registered with provenance, evaluator role, timestamp, environment and digest.
9. The selected profile determines required controls, tests, exclusions and review rules.
10. An assessment engine and, where the profile requires, an authorised human assessor evaluate the evidence.
11. The issuer signs a record containing scope, limitations, status and next-review triggers.

Verification and use are separate and belong to different parties:

12. A verifier resolves the identifier, checks signature and key status, verifies digests, evaluates freshness against `maxAcceptableAgeSeconds`, and reads the evidence coverage summary. This is a property of the record.
13. A relying party decides whether the record applies to its task, authoring an `ApplicabilityDecision`, and separately whether a concrete action is authorised. This is a property of the relying party's policy.
14. Material changes produce `REVIEW_REQUIRED`, `EXPIRED` or `REVOKED` and trigger selective or full reassessment per §7.

6.1.1 Compatibility with the public upload preview

The public-facing demonstration described in the project uses a flat upload manifest, including `resource_type`, `resource_name`, `resource_version`, `purpose`, `principal`, `autonomy_level`, `operating_context`, `jurisdiction`, `allowed_actions` and `prohibited_actions`. This is an intake contract, not the canonical SOV record or an executable agent.

A versioned adapter should preserve the original upload and its provenance, then map these fields to the proposed record. Identity and version map to `resource`; allowed and prohibited actions map to `authorityEnvelope`; operating context and jurisdiction map to `operatingContext`. `principal` remains a declared organisational principal and must not be silently converted into an authenticated controller or DID. Autonomy labels require an explicit vocabulary mapping; unknown values must be reported for review.

The intake version and SOV schema version are distinct. Missing evidence remains absent or undetermined, and the adapter must not manufacture tests, issuer signatures, ledger registrations or certification outcomes. A preview may explain the intended lifecycle while no live SOV-1 issuance is available.

6.2 Embedding a reference in an agent package

A developer embeds a compact SOV reference: resource identifier, manifest digest, credential URI and status endpoint with its maximum acceptable age.

The reference points to a verifiable record. It can be copied: a badge or URI alone does not authenticate the running resource. Runtime binding requires independently verified configuration and subject-key possession under the relevant policy. It does not embed a safety property in the agent, carries no authority, and remains a `declared` field until a verifier resolves and checks it. A package carrying a well-formed reference to a revoked credential is a correctly functioning system detecting a revoked credential.

6.3 Change between verification and execution

An operation binds identifiers and versions of programme, configuration, policy, status and mandate. The decision includes the request digest, audience, expiry and idempotency key, and the connector validates these limits where the effect is produced [55].

A subagent's authority is bounded by its mandate and by organisational and tool policies. A credential evidences a competence claim; it MUST NOT by itself increase permissions. New tools, audiences or data require a fresh decision when outside granted scope.

Exactly-once execution over arbitrary APIs is not promised. Idempotency is used where the destination supports it. On timeout with unknown effect, an indeterminate state is recorded and reconciled before retrying an operation with effects. Interruption, compensation and irreversibility are properties of the tool, not the credential [55].

7. Change-aware continuity and the reassessment procedure

7.1 Material-change taxonomy and conservative defaults

Class	Example	Conservative default treatment
CH1 Presentation only	Brand name, description, logo, no effect on the executed system	Update presentation metadata; preserve technical bindings
CH2 Isolated tool, contract confirmed	Deterministic formatter changes, contract and dependencies confirmed	Review affected competences; run integration and critical-control tests
CH3 Tool response schema	A tool's response schema changes	Targeted reassessment of competences consuming that tool, plus integration and critical-failure tests
CH4 Tool behaviour, schema unchanged	Same schema, different behaviour	As CH3, with reduced detectability; behavioural probes required because the declaration will not reveal it
CH5 Model, provider, global instruction, permissions, orchestration	Model substitution, system-prompt change, permission widening, subagent addition, retry or ordering change	Assume broad effect; full reassessment of the affected scope unless a validated justification narrows it
CH6 Delegation, sponsor, ownership, key	Controller change, key rotation, chain restructuring, compromise	Re-verify the chain; suspend on compromise regardless of assessment outcome
CH7 Jurisdiction, hosting region, processor	Region migration, new subprocessor	Re-evaluate the sovereignty evidence profile; applicability decisions referencing residency are invalidated for current use
CH8 Normative reference or interpretation	A framework requirement or its interpretation changes	Invalidate the affected mapping review; the competence credential is handled separately under issuer policy
CH9 Security incident, near miss, drift	Anomalous tool invocation, exfiltration attempt, unexplained distribution shift	Suspend the affected scope pending investigation
CH10 Evidence, evaluator or environment change	Evaluator replaced, environment rebuilt, benchmark contaminated	Re-establish evidence validity; prior conclusions become `undetermined` for affected dimensions
CH11 Unknown change or incomplete observation	Digest mismatch with no identified cause; detected coverage gap	Suspend the affected scope or require full reassessment

CH1 is a sanity control. Its savings are reported separately and excluded from the primary endpoint, because a mechanism that achieves cost reduction by correctly ignoring logo changes has demonstrated nothing [59].

7.2 Types used by the procedure

```
N            : set of configuration nodes
C            : set of claims
E            : set of directed edges (u, v), u ∈ N, v ∈ N ∪ C, meaning "v depends on u"
Graph        : (N, C, E, confirmation: E → {confirmed, partial, asserted})
Claim        : (id, statement, evidenceRefs: set<EvidenceId>, assumptions: set<AssumptionId>)
Assumption   : (id, statement, demonstratedBy: set<EvidenceId>, scope)
TestId       : identifier of a test in the profile's catalogue
Policy       : (version, mandatoryTests: ChangeClass → set<TestId>,
                permitsReuse: (Claim, ChangeClass) → bool,
                minCoverage: real, issuerKeyRef, reviewDeadlineDays)
Decision     : signed record as specified in Appendix A.4
```

7.3 Continuity planning and completed review

The following procedure is a proposed reference algorithm, not an executed implementation. It separates a **plan** from a **final decision**. `claimsBroken` means "prior evidence is not currently reusable", not "the resource failed the claim". `undetermined` contains all claims whose applicability to the new configuration is unresolved. Test identifiers in `evidenceRequired` are obligations, not results.

```
FUNCTION plan_continuity(prior, oldConfig, newConfig, observations,
                         policy, graph, evidenceStore) -> Plan
  claims := ids(prior.claims)
  preserved := {}; reused := {}
  REQUIRE claims != {}
  REQUIRE authenticated_current_issuer_authority(policy)
  REQUIRE authentic_record(prior) AND exact_binding(prior, oldConfig)
  REQUIRE same_resource_and_authorised_key_transition(prior, newConfig)
  classes := classify_all_changes(oldConfig, newConfig, observations)
  mandatory := union(policy.mandatoryTests(c) FOR c IN classes)

  IF compromise_or_incident(classes, observations):
      publish_scoped_hold_or_revocation(prior, policy)
      RETURN full_plan(claims, mandatory, "incident review")

  unknownScope := CH11 IN classes
      OR NOT authenticated_observation_coverage(observations)
      OR NOT sufficient_evidence_coverage(prior, policy)
      OR NOT known_test_policy_for_all(classes, policy)
  IF unknownScope:
      RETURN full_plan(claims, mandatory, "scope not established")

  # A broad change defaults to the complete credential scope.
  IF CH5 IN classes AND NOT validated_narrowing(classes, newConfig):
      RETURN full_plan(claims, mandatory, "broad change")

  reached := forward_reachable_claims(graph, changed_nodes(observations))
  reached := reached UNION global_and_composition_claims(graph)
  FOR claim IN prior.claims:
      support := authenticated_support(claim, graph, evidenceStore)
      independent := positive_independence_evidence(claim, observations)
      IF support.isNonEmpty AND support.isCurrent
         AND support.allAssumptionsDemonstratedFor(newConfig)
         AND independent.isAdequate
         AND claim.id NOT IN reached
         AND policy.permitsReuse(claim, classes):
          preserved := preserved UNION {claim.id}
          reused := reused UNION claim.evidenceRefs

  review := claims DIFFERENCE preserved
  tests := mandatory UNION tests_covering(review, policy)
```

```
                    UNION critical_and_integration_tests(classes, policy)
  uncovered := review DIFFERENCE declared_claim_coverage(tests)
  RETURN signed_plan(prior, oldConfig, newConfig, policy, graph,
      claimsPreserved=preserved, claimsBroken=review,
      evidenceReused=reused, evidenceRequired=tests,
      undetermined=review, uncoveredClaims=uncovered,
      resultingStatus=REVIEW_REQUIRED)

FUNCTION complete_review(plan, authenticatedResults, currentInputs) -> Decision
  REQUIRE authorised_reviewer(currentInputs)
  REQUIRE exact_plan_digest_and_configuration_binding(plan, currentInputs)
  REQUIRE unchanged_policy_graph_and_evidence_dependencies(plan, currentInputs)
  REQUIRE no_new_change_or_status_hold(plan, currentInputs)
  accepted := appraise_results_against_exact_obligations(plan, authenticatedResults)
  REQUIRE accepted.satisfiesEveryMandatoryTest
  REQUIRE accepted.coversAll(plan.claimsBroken)
  REQUIRE accepted.hasNoCriticalFailure AND accepted.hasNoUnresolvedClaim
  REQUIRE preserved_evidence_remains_current(plan, currentInputs)
  RETURN signed_final_decision_and_successor_record(plan, accepted)
```

`full_plan` takes the complete in-scope claim set, marks it unresolved, selects the full programme plus mandatory checks, and returns `REVIEW_REQUIRED`. It must not compute scope solely from an untrusted graph. Failed preconditions produce a review hold or rejection with a reason; they never default to VALID. An empty or unknown assumption set cannot pass by vacuous truth. Missing coverage requires a revised test plan before completion.

Planning always leaves the new configuration in `REVIEW_REQUIRED`, even if claims can be preserved or the selected checks are empty. A final, authorised review may confirm reuse without new behavioural tests when all obligations are positively satisfied. Otherwise signed, configuration-bound results must be appraised first. Finalisation creates a new version or successor credential; it never silently applies the old configuration's credential to the new one. A later policy, graph, status or configuration change invalidates the plan's completion preconditions. This is a proposed transaction boundary, not an implemented guarantee.

7.4 Worked example one: narrow tool change with selective reuse

An agent has four bounded claims, c1-c4. A date formatter changes implementation while its declared interface remains unchanged. An authenticated digest comparison establishes an artefact change, not behavioural equivalence. CH2 classification additionally requires evidence supporting the narrow impact boundary.

Claims c3 and c4 depend on the formatter and remain unresolved. Integration tests are requested for c3; boundary tests are requested for c4. Both stay in `undetermined` until their results are obtained and appraised. Claims c1 and c2 may be preserved only with current supporting evidence and positive independence evidence, not merely because no graph edge reaches them.

In the synthetic example, that independence evidence is stipulated, and the graph is confirmed for the stated scope. The decision is a plan with `resultingStatus = REVIEW_REQUIRED` and `undetermined = [c3, c4]`. If the consumer set were only partial and c1/c2 independence were not established, those claims would also require review. No saving or successful assessment is measured by this illustration.

7.5 Worked example two: model substitution plus permission widening

The same agent's model is substituted for a different identified version, and in the same window its tool permissions are widened to include a write action.

Both changes classify as CH5. Model substitution invalidates the assumptions underlying every behavioural claim, because behavioural evidence is conditioned on the model that produced it. Permission widening changes the authority envelope, invalidating every claim whose scope referenced the prior envelope and additionally requiring re-verification of the mandate chain, since the mandate may not permit the widened action.

The procedure returns full reassessment of the affected scope. The correct behaviour of a change-aware system on this input is to be indistinguishable from full reassessment, and a mechanism that finds savings here is broken.

One subtlety: two changes arriving together is not the same as two arriving separately. Interaction effects between a new model and a widened permission are not covered by evidence about either alone. The interaction region MUST remain `undetermined` unless directly assessed, and MUST NOT be treated as covered by the union of the two reassessments.

7.6 Unknown changes and incomplete observability

The hardest case is the unobserved change. A model alias serving a silently updated model produces no digest change, no declaration change and no event.

Responses include requiring version identification where available, limiting credential lifetime, running behavioural canaries and explicitly recording unobservable dependencies. Canaries may detect a deviation without identifying its cause. Expiry bounds the period in which old evidence may be accepted under policy; it does not bound the damage from an unobserved change. No guarantee of detecting every provider update, or remote attestation over provider infrastructure, is made.

8. Transparency, ledgers and digital sovereignty

8.1 What a ledger is for

Public ledgers enter this architecture as an optional mechanism for publishing commitments over batches of records, supporting verification between organisations that do not share an operator. They are not the source of truth, not a substitute for assessment, and not on the critical path of any authorization decision [55].

The design publishes one root per batch, containing no client documents, instructions, responses, source code or personal identities. Detailed evidence and status remain off-chain in controlled, exportable storage.

Three properties require care. Commitments over predictable data require adequate blinding: a bare hash of a low-entropy value is not confidential. Publication volume and timing leak information, and metadata analysis over anchoring cadence can reveal customer activity. An anchoring timestamp supports a statement about when a commitment was published, not about when the described action occurred.

Receipts distinguish pending publication, confirmed inclusion and finality assurance according to the chosen network. RFC 9942 [11] provides a standardised receipt format with CBOR encodings for inclusion and consistency proofs, which removes a substantial part of the reason to invent one.

8.2 Three alternatives to compare

Option	Mechanism	Evaluation criteria
Operated append-only log	Single-entity log with published inclusion proofs	Lowest latency and cost; weakest equivocation resistance
Witnessed transparency service	Independent witnesses and consistency checking per [10], with receipts per [11]	Standardised, third-party verifiable, moderate operational burden
Chain-anchored log	Witnessed service plus commitment anchoring on an existing chain	Strongest cross-organisation independence; highest cost and largest privacy surface

Criteria are observed resistance to divergent histories, third-party verifiability, latency, availability and operating cost [55]. Because [10] and [11] now standardise the underlying properties, the burden of proof sits with the third option: it must demonstrate a verification requirement a witnessed service does not meet.

No new blockchain and no token is required or proposed. EBSI [50] and ERC-8004 [41] are possible integration surfaces subject to technical and institutional fit. Ownership of a transferable identifier does not transfer the competence or authority associated with it, and the empirical evidence in §2.5 indicates that on-chain reputation records in the registries studied did not carry a usable trust signal [31].

8.3 Sovereignty as an operational property

Sovereignty is treated as observable control and decision capacity, evidenced across ten dimensions [55], [61]:

Dimension	Evidence sought	Provenance
Jurisdiction and legal regime	Governing law, applicable regime, entity locations	`declared`, reviewable
Data residency and retention	Storage regions, retention schedules, backup locations	`declared` plus `measured` where instrumented
Key custody and rotation	Custody model, HSM or KMS integration, rotation and recovery	`measured` where testable
Administrator and issuer control	Privileged access holders, separation of duties	`declared`, audit-reviewable
Model, provider and infrastructure dependency	Named external dependencies and their substitutability	`declared`, often incomplete
Portability and exportability	Export formats, permitted scope, completeness	`measured` by test
Reconstitution or exit test	Executable reconstruction of permitted history elsewhere	`measured`
Subprocessor and supply-chain visibility	Subprocessor register, change notification	`declared`
Incident, suspension and revocation control	Who can suspend, under what authority, how fast	`measured` by test
Continued verification if the operator is unavailable	Verifier operates without a private platform API	`measured`

An important test combines history reconstruction with independent verification in another environment. Historical verification needs exported signed records and appropriate trust material. Current authorization additionally needs sufficiently fresh status from an authorised, possibly replicated or successor, service; an expired snapshot cannot establish present non-revocation after the issuer disappears. Isolated data residency does not demonstrate control over all dependencies, and a sovereignty claim that reduces to a region label is a marketing claim.

9. Threat model, security and privacy

9.1 Assets, boundaries and attacker capabilities

Assets: reserved test material and rubrics; evidence and its provenance; issuer, assessor and gateway signing keys; status data and its freshness; mandate scope; and decision-record integrity.

Trust boundaries: submitter to validator; evidence producer to assessor to issuer, following the RATS separation [12]; issuer to verifier; verifier to relying party; gateway to connector; and platform to tenant.

Assumed attacker capabilities: full control of submitted content; observation of public records and anchoring metadata; network partition and delay; compromise of a single key or component; and, in the collusion cases, corruption of an authorised participant.

9.2 Threats, mitigations and residual risk

#	Threat	Mitigation	Residual risk and explicit non-solution
T1	Malicious or careless submitter	Provenance classification and rejection; digest binding; candidate has no access to reserved set, official result or assessor key	Cannot detect a truthful declaration of a system that behaves differently in production
T2	Forged or substituted manifest	Signature and digest verification; immutable version records	A valid signature over a false declaration remains valid; SOV does not verify declaration truth
T3	Compromised issuer, assessor or key	Key separation across issuance, assessment, review and administration; rotation and recovery; suspension on compromise independent of assessment outcome	Detection latency; a compromised issuer can issue valid credentials until detected

#	Threat	Mitigation	Residual risk and explicit non-solution
T4	Stale, replayed or selectively disclosed credential	Audience-bound presentation; challenge and nonce; proof of possession; explicit maximum status age	Deliberate key sharing by the holder is not prevented; proof of possession raises cost, not impossibility [56]
T5	Equivocation between verifiers	Witnessed transparency service with consistency checking [10]; receipts [11]; optional anchoring	An operated log alone does not prevent equivocation, which is why the option exists
T6	Compromised registry, transparency service or ledger	Independent witnesses; portable verifier; exportable evidence	Service unavailability degrades cross-organisation verification
T7	Hidden model or provider change	Version identification where exposed; bounded validity; behavioural canaries; explicit exclusion	Not detected, only bounded. Stated as a non-solution in the credential
T8	Prompt injection, tool misuse, exfiltration, privilege escalation	Threat-test families; tool rights scoped per request, resource, action, configuration and expiry; connector-side validation	Test coverage is a sample; passing does not establish absence
T9	Collusion between submitter and assessor	Separation of case generation, assessment and issuance; reserved sets; disagreement analysis; conflict-of-interest controls	A valid signature does not validate methodology and does not prevent fraud by an authorised assessor [56]
T10	Evaluator bias or benchmark leakage	Reserved sets with access control, variants, repeated trials, error analysis	An agent may recognise an exam environment or have been trained on the tasks; a standing methodological risk [56]
T11	Denial of service and service unavailability	Anchoring off the critical path; explicit cache policy; defined partition behaviour	Fail-closed reduces availability by design
T12	Correlation and privacy leakage from identifiers or anchors	Blinded commitments; batch anchoring; minimal disclosure	Volume and timing metadata remain analysable
T13	Cross-tenant evidence access and insider abuse	Row-level isolation [51], role controls, cross-organisation tests, privileged-function separation	Insider risk reduced, not removed
T14	Composed-system risk from individually assessed parts	Composition assessment required; component results never auto-inherited	Interaction space is combinatorially large; coverage is partial by construction

9.3 Target security properties

Authenticity and integrity of records; freshness within a stated bounded window; non-repudiation of decisions by identified keys, subject to key-custody assumptions; selective disclosure at presentation; least privilege for tool rights; auditability of decision lineage; revocability subject to distribution latency; and fail-closed behaviour under unknown or stale state.

These are target properties, not a completed security evaluation. Selective disclosure still requires a concrete mechanism and leakage analysis. Appendix D records implementation gaps; passing a sampled assessment does not establish absence of vulnerabilities.

9.4 Three project-reported prototype limits

Internal material [62] reports three targeted demonstrations of v0.2 limits. The manuscript preserves these observations as project-reported results. Logs, code, seeds and the underlying JSON evidence were not supplied with this edition, so the publication review does not independently reproduce them. Their scope is a local prototype, not a production or independent security assessment.

ID	Property	Observation	Result	Classification
A01	Semantic quality of reasoning	Rationales replaced with content-free text, structured fields left correct	`DEMO_PASS`, 12 of 12 cases	Known assessment-validity limitation
A02	Immediate revocation	Gateway presented with a prior active snapshot while the issuer had generated a newer revoked snapshot	`ALLOW_WITH_HUMAN_REVIEW` at zero seconds after revocation, maximum snapshot lifetime 120 seconds	Known freshness window, **not** a signature bypass
A03	Enforcement of required framework mapping	Mapping in `REVIEW_REQUIRED` while the separate competence and mandate gateway still admits	`ALLOW_WITH_HUMAN_REVIEW`	Known absence of a mapping requirement in admission policy

A01 limits claims to demonstrated structured-answer and answer-key conformance. The assessor verifies structure and correspondence with an answer key across seven dimensions; free-text justification is required but its semantic quality is not scored, so a response can populate every structured field correctly while presenting inadequate reasoning [58]. Until an independently validated method for scoring reasoning quality exists, an assessment result from this programme supports the tested structured criteria but does not establish analytical competence or reasoning quality, and any language suggesting otherwise is a defect.

A02 demonstrates bounded staleness, not a signature bypass or necessarily a failure to reject UNKNOWN. The report describes acceptance of a correctly signed earlier active snapshot within its permitted 120-second lifetime, after a newer revocation existed. Such acceptance can satisfy the configured age rule while missing a more recent state. Tightening the age rule or using an authenticated online check can reduce exposure; neither implies instantaneous global revocation or cancellation of effects already in flight. Tests must distinguish the configured age bound, actual propagation latency, clock skew and use of an already-known newer status [56].

A03 demonstrates missing policy integration at the admission point. A mapping requiring review does not currently gate admission, so an enterprise policy demanding valid mappings requires explicit integration at the control point [57].

The project's fifty-eight passing tests remain useful for the properties they cover. They do not demonstrate assessment validity, environment coverage, production robustness or institutional certification readiness [62].

10. GRC and organisational decision integration

10.1 The bridge, and what it must not become

Sovrentic can supply structured evidence and decision inputs. The organisation retains its risk appetite, legal interpretation, policy and authorization. The distinction between a mapping to [1], [2], [3], [5] or [6] and a claim of legal compliance is a structural property of the data model, not a disclaimer. Vendor crosswalks, including [39], can inform comparison but do not establish legal equivalence or comparative market leadership.

Five record types carry the bridge [57]:

Record	Minimum content
Framework and requirement	Responsible body, identification, version, source, consultation date, conditions of use
Applicability context	Organisation, agent, configuration, task, actor role, jurisdiction, justification
Evidence	Origin, collection method, integrity, observation period, limitations, authorised access
Coverage assertion	Requirement, evidence used, reasoning, scope, gaps
Review	Reviewer, competence or mandate to review, decision, date, validity conditions

Coverage states MUST distinguish at least: not yet assessed; not applicable with justification; partial support; sufficient support within the reviewed scope; and contradictory evidence. Sufficient support within scope does not mean the

framework was satisfied. A model may suggest relations; the record MUST identify the suggestion and the review that validated it. No universal equivalence between frameworks is published, and no ISO clause text is reproduced [57].

10.2 Why mapping validity and credential validity are decoupled

A requirement change flags affected mappings for review. It does not automatically revoke a competence credential whose exam, configuration and scope remain valid. Issuer policy decides whether the change propagates [57].

The two objects answer different questions with different authorities: the credential says what a configuration demonstrated; the mapping says what an organisation may conclude from that demonstration in a regulatory context. Collapsing them would make every framework revision a mass revocation event, which is both operationally absurd and epistemically wrong. Automatic selection of these dependencies remains a development hypothesis, not an implemented capability.

10.3 An illustrative bounded mapping

The experimental programme carries two bounded assertions, each reused against two references [58]:

Bounded assertion	Proposed references	Preserved gaps
This report documents the experimental examination	NIST AI RMF MEASURE 2.1 [24]; IMDA MGF 2.3.2 [22]	Independent review, representative reserved cases, real execution
This receipt documents an admission with fictitious mandate and approval	NIST AI RMF GOVERN 3.2 [24]; IMDA MGF 2.2.2 [22]	Effective organisational roles, oversight effectiveness

These associations are the project's own interpretations, subject to expert validation. Correspondence with ISO/IEC 42001 [7] remains at the level of the evidence's potential usefulness to a management system; no ISO clauses have been loaded and no AI Act applicability conclusion has been implemented. The mapping catalogue is supplied by the verifier, hashes correspond to local reference metadata rather than the full normative text, and any change to reference metadata invalidates the review for current use [56].

10.4 Audience mapping

Risk and audit need scope, exclusions, evidence class and reviewer identity, which the model supplies natively. Security needs the coverage matrix and residual-risk column of §9.2 and Appendix D. Privacy needs evidence provenance and disclosure design. Legal needs the applicability context and the explicit statement that no legal conclusion is offered. Procurement needs issuer and assessor roles, conflict-of-interest controls and the exit test. Operations needs freshness semantics and revocation behaviour under partition.

11. Implementation status, roadmap and operating model

11.1 Technology choices

Python for the API and domain engine with modules separated by responsibility, and evaluators and connectors isolated in permission-limited processes. PostgreSQL for versioned objects, relations, states and transactional decisions, with encrypted object storage for larger evidence and a transactional outbox linking state changes to idempotent asynchronous consumers; the first graph can be represented in tables, and a specialised graph database requires measurement to justify. Row-level security contributes to isolation alongside role controls, cross-organisation tests and privileged-function separation [51]. W3C VC 2.0 [13] with a fixed and tested JOSE profile [15], and Bitstring Status List [17] as a status-distribution candidate, requiring cross-implementation testing that the prototype's signed JSON does not constitute. Enterprise identity integration and, where appropriate, SPIFFE for workloads [46], with OPA or equivalent as a rules engine [47] and MCP adapters as integration points respecting their authorization model [48]. RATS separation of evidence producer, appraiser and relying party [12]. A documented verifier with public test vectors, and evidence and status exportable subject to permissions [55].

Specialised graph databases, trusted execution environments, zero-knowledge proofs, transparency integration and chain anchoring are later choices rather than MVP prerequisites, because each adds a trust boundary or operational dependency whose necessity has not been demonstrated.

11.2 Capability status

The public website preview, local prototype v0.2, target reference architecture v0.3, private MVP/V2 and live-production V3 are different milestones. An interface preview does not implement issuance or establish the prototype's deployment status. The following prototype descriptions are attributed to internal sources and were not independently reproduced for this publication edition.

Capability	Local prototype / public preview	MVP / V2 target	V3 production gate
Upload and declared resource record	Public illustrative intake; separate prototype objects reported [56]	Versioned adapter and tenant records	Durable storage and migration controls
Assessment	Structured control-evidence exercise; A01 limitation [58], [62]	Expert-calibrated programme and protected cases	Independent assessment validation
Credential and status	Demonstration signed JSON and snapshots reported [56]	Fixed signature profile and authenticated status	Independent verifier and key recovery
Mandate and admission receipt	Local fictional-actor demonstration; A02/A03 limits [56], [62]	Explicit policy integration and controlled connector	Tested expiry, replay, concurrency and isolation
Continuity engine	Specified; comparative experiment unexecuted [59]	Implement plan/completion separation	Evidence for selection safety and cost
SOV-1 / advanced profile issuance	Proposed scheme; no operating programme demonstrated	Accountable issuer and defined profile workflow	Published governance and justified assurance claims
Transparency and optional anchoring	Target design	Verifiable receipts and export experiment	Witness, partition and independent verification tests
Portability and sovereignty	Proposed evidence dimensions	Reconstruct permitted history elsewhere	Repeatable exit and current-status continuity tests

11.3 Conditions for accepting a pilot

Proposed engineering criteria, none currently met [55]:

Property	Verification required
Assessment validity	Cases and rubric expert-reviewed; real responses; disagreements and limits documented
Link to execution	Observation of covered components and explicit declaration of exclusions
Operation control	Attempts via alternative paths produce no effects outside the mandate in the pilot environment
Concurrency and recovery	Restarts and multiple workers preserve decisions, authorization consumption and supported idempotency
Freshness	Revocation and policy change measured under delay and partition; use beyond the agreed window prevented
Isolation	No organisation, evaluator or issuer reads or alters evidence outside its scope without authorisation
Keys	Rotation, compromise, recovery and separation across issuance, assessment, review and administration
Interoperability	Issuance and verification by independent implementations of the agreed profile
Historical evidence	State, policies and sources used in a decision retained, with temporal-proof limits explicit

Proposed pilot targets, which are pre-declared objectives and not results or service levels [55]: a local decision at p95 within 100 ms at 100 requests per second with current caches, up to five credentials per request and recorded hardware conditions; and a maximum state window of five seconds for new sensitive operations plus permitted clock skew, with refusal when freshness cannot be demonstrated.

11.4 Roles, conflicts and liability boundaries

The ten roles of §5.2 apply. Conflict-of-interest controls: separation of case generation, assessment and issuance decision; a fee that funds assessment and operation and does not purchase approval; a candidate controlling neither reserved

material nor the assessor key; assessor authorisation scoped to programme hash, subject, key and configuration with an expiry; and issuance requiring a signed assessment declaration from an assessor authorised for that programme [56].

Governance obligations: appeals against an assessment conclusion; suspension pending investigation; revocation with reason codes; renewal with re-established evidence; incident disclosure to affected relying parties; and liability boundaries stating that the issuer's responsibility is bounded by the credential's stated scope and exclusions.

Under the EU AI Act, human oversight includes system design requirements in Article 14, provider obligations in Article 16 and deployer obligations in Article 26, where applicable [1]. "Relying party" is a technical role here, not a statutory classification; accepting a credential does not transfer those obligations. IMDA published material on legal responsibility for AI agents alongside the May 2026 framework update; this paper has not verified a stable primary citation for that document and therefore does not rely on it. It is recorded as an open item in Appendix G and in the publication review notes.

11.5 Commercial logic and its evidentiary status

SOV-1 is proposed as a free adoption and ecosystem entry point. Advanced profiles could support paid assessment, evidence review, testing, monitoring, API access, verification volume, integrations and issuer or assessor workflows.

The candidate moat is as stated in §3.5: the accumulated, versioned graph of claims, evidence lineage, material-change rules, decision rationales, test vectors and institutional verifier adoption. This is a commercial hypothesis requiring evidence.

No evidence of prices, customers, partnerships, accreditations, market shares, valuations or revenues is presented in this manuscript. No network effect is assumed in any financial plan [56]. An acquisition thesis is a hypothesis requiring adoption and independent validation; this paper does not demonstrate willingness to pay, acquisition interest or a valuation [55]. What an acquirer would need is the ability to integrate the engine into their programme, understand the methodology and verify results without informal explanation from the founders. The architecture can make that possible and has not yet.

12. Evaluation protocol and falsifiable hypotheses

No comparative results are reported. This section specifies how they would be obtained and what would refute H1.

12.1 Study type

This is declared an **exploratory feasibility study**, not a confirmatory evaluation. Its purpose is to estimate effect direction and magnitude and to determine whether a confirmatory study is warranted. A separate confirmatory study on a fresh change set is required before any performance claim is made publicly [59].

12.2 Conformance and negative tests

Conformance: manifest validation against the Appendix A.1 schemas; credential verification including signature, key status, digest binding, validity interval and status freshness; presentation binding to audience and challenge; schema version pinning including rejection of an unaccepted version rather than best-effort interpretation.

Negative: malformed manifests; substituted manifests carrying valid signatures over altered content; credentials stale beyond `maxAcceptableAgeSeconds`; revoked credentials presented inside and outside the freshness window; scope-expanded presentations requesting actions outside the mandate; replayed presentations against a different audience; credentials whose `assessmentDigest` does not resolve; and records asserting `measured` provenance without a collection method.

A02 is a required staleness regression scenario. Acceptance inside an explicitly allowed freshness window and rejection outside it must be tested separately. If a deployment requires a shorter window than 120 seconds, A02 demonstrates a policy mismatch; it does not by itself demonstrate broken UNKNOWN handling.

12.3 The continuity experiment

"Three configurations" means three distinct agent configurations of the same declared competence, differing in model family, tool set and orchestration, each with an authorised evaluation agreement, controlled versions, tools and data [59]. They are configurations, not three separate agents in different domains; holding the competence constant is what makes cost comparison meaningful.

Corpus. Sixty cases in three disjoint groups of twenty: development, calibration and reserved. Near variants of the same case remain in the same group to limit contamination. The candidate does not receive expected results. The team tuning selection rules does not consult reserved results before fixing them. The prototype's twelve public cases, distributed with their answer key, serve mechanism development only and are neither the reserved set nor independent validation [58], [59].

Execution budget and its derivation.

Class	Derivation	Executions
Baseline	3 configurations × 60 cases × 3 repetitions	540
Post-change full matrix	3 configurations × 12 changes × 60 cases × 3 repetitions	6,480
Total		**7,020**

One case execution may involve several model, tool or evaluator calls; retries, additional assessments and incident investigation require separate accounting and are not inside this figure. This is a proposed size, adjustable to budget before execution. Design changes made after consulting results MUST be recorded and MUST NOT be presented as independent confirmation [59].

Critically, the full matrix is executed once. The comparators are then **selection procedures over the same observed results**, not separate execution campaigns. Comparator cost is computed as the subset of executions each procedure would have commissioned, plus its own maintenance and review time. This avoids confounding selector differences with fresh model variation.

Changes [59]. M01 model version substitution; M02 global instruction change; M03 tool-specific instruction change; M04 isolated deterministic formatter change; M05 tool response schema change; M06 tool behaviour change without schema change; M07 retrieval source update; M08 removal of a required document; M09 permission widening; M10 subagent addition; M11 retry or orchestration order change; M12 purely presentational change with no effect on the executed configuration. M12 is a sanity control reported separately and excluded from the primary endpoint.

Comparators.

ID	Procedure	Cost basis
A	Full reassessment of the relevant set after each change	All executions in scope
B	Expert-defined manual rules mapping change type to test subset	Selected executions plus rule maintenance minutes
C	Candidate engine, §7.3, with regression to full reassessment under uncertainty	Selected executions plus graph construction and maintenance minutes
C-cal / H (secondary temporal study)	Calendar-only / calendar plus event-triggered review	Timed checks, monitoring, maintenance and suspension cost

The 7,020-execution matrix supports the static A/B/C comparison only. Selections must be frozen before viewing the reserved outcomes used to score them. A separate temporal study is required for C-cal and a hybrid H. Before execution, it must fix a replay timeline, event times, persistent and transient failure scenarios, change observability, review delays and exposure/suspension metrics. Every strategy receives the same declared observation channels. Reassessments at calendar boundaries must evaluate the configuration actually present then, not a different configuration from the static matrix. Its additional execution and human-review budget is not included in 7,020. Periods of 30, 90 and 180 days provide sensitivity analysis; 90 days is an approximation to a quarterly schedule, not a simulation of AIUC-1. Public maintenance pages cannot establish a product benchmark [38].

12.4 Primary endpoint

Primary endpoint. Cost-adjusted review burden `R(X)` for procedure `X ∈ {A, B, C}`, over the reserved set, excluding M12:

```
R(X) = w_e · E(X) + w_m · M(X) + w_r · Rv(X)

where
```

```
E(X)  = case executions commissioned by X
M(X)  = machine time consumed, in seconds
Rv(X) = human reviewer and maintenance minutes attributable to X,
        including rule maintenance for B and graph construction and
        maintenance for C
w_e, w_m, w_r = weights from a dated, approved price table, fixed before
        unblinding; physical units retained separately so that results do
        not depend on later price changes
```

Non-negotiable safety constraint. For every configuration-change pair in the reserved set, if full reassessment `A` observed a critical failure, procedure `X` MUST have selected a test set that would have detected it. A procedure violating this constraint fails, irrespective of `R(X)`.

Success criterion, pre-declared. `C` succeeds if the safety constraint holds and `R(C) ≤ 0.75 · R(A)` on relevant changes. The twenty-five percent figure is an investment criterion adopted before measurement, not a prediction and not a result [59]. Separately, `C` must satisfy `R(C) < R(B)` after maintenance is counted, with temporal comparison against C-cal and H reported separately, including detection delay, exposure before detection and suspension burden. Delayed calendar detection cannot satisfy the same immediate per-change detection claim by retrospective relabelling.

Alternative endpoint considered and rejected. Critical-failure detection rate as the primary endpoint was rejected because it is a constraint rather than an objective: a procedure that always selects everything achieves perfect detection at zero saving. Making detection the constraint and burden the objective states the trade-off correctly.

12.5 Secondary metrics and sensitivity

Secondary: per-case omissions, meaning how many cases with an observed critical failure fell outside the selection; improper reuse, meaning decisions preserving a claim despite a broken assumption, invalid source or out-of-scope change; unnecessary suspensions and unnecessary full reassessments, which reduce operational utility and are the failure mode a conservative mechanism produces; inter-rater agreement among human reviewers on the same cases and calibration of any model-based classifier against expert judgement with disagreement analysis; revocation latency distribution under normal operation, delay and partition; portability, meaning reconstruction of permitted history and credential verification in a separate environment without a private platform API; cross-tenant access attempts and key rotation, compromise and recovery drills; and transparency-service behaviour under unavailability and witness disagreement [59].

Sensitivity analysis on p. `R(C-cal)` is reported at `p ∈ {30, 90, 180}` days. A result in which `C` beats `C-cal` only at `p = 180` is a weak result and must be reported as such, because performance can depend strongly on the selected schedule. No market-wide cadence is inferred from one scheme.

12.6 Statistical treatment and uncertainty

Repetitions, related case variants and repeated changes on the same configuration are dependent. Preserve paired A/B/C observations and resample at the predeclared independent case-family level, retaining the within-family structure. Report family counts, raw paired differences and uncertainty intervals. With only three purposively selected configurations, inference is conditional on those configurations; resampling cannot justify population-wide claims about model families or deployments.

Report critical-failure detections and omissions with explicit numerators, denominators and dependence structure. An unadjusted binomial or Wilson interval over correlated executions would overstate precision. Where the number of independent positive families is insufficient for a defensible interval, report descriptive counts and the limitation instead. Zero observed omissions is not a near-zero population-rate guarantee. Pairs with no critical failure under the full set cannot establish detection retention and must be identified separately.

Reserved cases and outcome labels are held separately from teams tuning selectors. Timestamp the frozen rules, graph, policies and prices before unblinding. Test selection may use authorised configuration metadata and predeclared diagnostics, not held-out outcomes. Any subsequent tuning must be disclosed as exploratory and evaluated on fresh reserved material. Temporal outcomes and the hybrid comparison require their own predeclared analysis plan [59].

12.7 Stopping rules

Advance to a limited pilot if the safety constraint holds, the pre-declared success criterion is met, a reviewer can reconstruct the reuse rationale from the emitted decisions, and an independent verifier can interpret exported objects under the agreed profile [59].

Stop automatic reuse in the affected scope on an omitted critical failure, an unknown dependency, loss of traceability, or a change of premises. If `C` does not improve on `B` after maintenance is counted, the graph MUST NOT be presented as a technological advantage, and the investment thesis moves to the institutional utility of issuance and verification, validated with partners [55].

12.8 Matrices

Appendix D gives threat-to-test coverage; Appendix E gives claims-to-evidence. Both contain entries with no covering test, marked as such.

13. Limitations and open research questions

Probabilistic and non-stationary behaviour. A passing result is conditioned on a sample, an environment and a moment. Repetition reduces variance; it does not establish stationarity.

Incomplete observability of remote models and providers. Bounded by §7.6, not solved.

Assessment validity. A01 is the concrete instance and it is the most serious limitation in this paper. Until an independently validated method for scoring reasoning quality exists, results from this programme support the tested structured-answer criteria, without establishing reasoning quality [58], [62].

Unknown dependencies and supply-chain change. A missing edge is not evidence of independence, and the architecture treats it accordingly at the cost of conservative over-testing.

Imperfect temporal evidence. An anchoring timestamp evidences publication of a commitment, not the timing of the described action.

Revocation and cache windows. A02. Shorter expiry narrows without eliminating, and does not interrupt operations in flight.

Jurisdictional and legal uncertainty. Applicability of [1] and adjacent instruments depends on system, use, actor role and provision. No legal conclusion is offered anywhere in this paper.

Privacy against transparency. Public identifiers and anchoring cadence leak metadata. Blinding addresses content, not volume and timing.

Credential theft and misuse. Proof of possession raises cost; deliberate key sharing by the holder is not prevented [56].

Limits of transparency infrastructure. Inclusion is not truth. Registration of a signed statement does not establish the correctness of its content [10].

Competitive position. There is no single global agent-certification standard, but there is an operating scheme with a published requirement set, an authorised audit route and a documented maintenance cycle [37], [38], [40]. This is a materially different competitive situation from a vacuum.

No empirical evidence for the continuity claim. This is the principal limitation and it is not mitigated anywhere in this paper.

13.1 Prioritised agenda

MVP. Expert calibration of the rubric and answer key, and a validated method for scoring reasoning quality, which is the precondition for any credential meaning what it says. JSON Schemas promoted from Appendix A.1 to a specification with valid and invalid fixtures and published test vectors. Material-change test vectors. A source-status registry recording owner, version, status, licence and review date for every reference [61]. Execute the study in §12.

V2. Authenticated status with explicit freshness and partition policy, reducing and measuring the A02 exposure window without claiming instantaneous revocation. Mapping-requirement enforcement at the admission point, closing A03. A recorded decision comparing DID/VC, SPIFFE/WIMSE and OAuth/OIDC identity and delegation profiles. Agentic threat suite. Verifier contract validated by an independent implementation. Role, scope and liability model required before offering an operational certification service [61].

V3. Conformance-tested VC profile and external verifier SDK. Transparency service aligned to [10] and [11], with anchoring only if a witnessed service demonstrably fails a real cross-organisation requirement. Runtime policy adapter. Multi-agent delegation controls. Sovereignty evidence backed by executable exit tests. Independent assessor and issuer workflows.

13.2 Experiments that could falsify the thesis

The §12 comparison is primary. Three secondary falsifiers: if a predeclared calendar or hybrid strategy matches detection, exposure and operational utility at lower total burden, incremental event-triggered value is unsupported. If B matches C's selection quality at lower total cost, the graph fails. And if reviewers cannot reconstruct reuse rationales from the emitted decisions, the inspectability claim fails independently of any cost result, which matters because inspectability is the part of the proposal least dependent on savings.

14. Conclusion

An assessment describes a configuration, scope and observation period. Deployment changes can weaken its applicability. SOV's proposed contribution is an explicit account of that transition: justified evidence reuse, new obligations, unresolved claims and an accountable decision. Identity infrastructure and transparency receipts support the record but do not establish behavioural truth.

SOV is proposed as a structure for keeping the core objects separate and linked, so that a relying party can see what was assessed, under what conditions, by whom, with what exclusions, and what has changed since. Sovrentic is developing the proposed SOV Certification Authority and its enabling infrastructure. First-party and institutional operation share the same need for accountable issuance, explicit scope, assessors with appropriate independence and portable verification.

The technical hypothesis is that dependency-aware and assumption-aware selection can reduce review burden while preserving critical-failure detection, with temporal exposure compared separately against scheduled and hybrid strategies. That claim is a hypothesis. The experiment that would establish or refute it is specified in §12 and has not been executed.

The next contribution must be evidence: validate the assessment programme, implement the plan/completion boundary and compare selective review against credible baselines. If the mechanism fails those tests, its claimed advantage must be withdrawn or narrowed. Scoped, reproducible decisions remain the criterion for both the research and the proposed business.

15. Acknowledgement of source limits

Publication status. Public technical proposal, prepared in an academic manuscript format. Not peer reviewed, accepted, or submitted to a conference. This extended edition includes technical appendices; venue-specific length, authorship and disclosure requirements remain to be applied before submission.

Conflict of interest. This paper presents the proposed Sovrentic venture and its SOV scheme. It is an interested party's technical proposal, not an independent evaluation. Internal references [55]-[62] are project inputs, not external validation.

AI assistance. Generative AI assisted drafting, editorial revision and technical consistency checks. This assistance does not constitute independent scientific review. Named human authorship, responsibility and the target venue's disclosure requirements must be finalised before submission.

Evidence status. Architecture and profile definitions are proposals. The comparative study has not been executed. Section 9.4 reports three prototype observations attributed to internal project material; those underlying artefacts are not included in this manuscript, and their results were not independently reproduced for this publication edition. Performance and cost-reduction figures in the evaluation plan are targets, not results.

Every claim about the prototype in this paper derives from internal design documents and one machine-readable results file [55]-[62]. These are design inputs and project evidence. They are not independent external validation, and no statement in this paper should be read as converting an internal design input into an external fact.

16. References

Grouped by source status. Status labels are binding on how each source may be used; see Appendix G.

A. Legislation, standards and technical specifications

[1] European Parliament and Council, *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

[2] European Parliament and Council, *Regulation (EU) 2016/679 (General Data Protection Regulation)*. https://eur-lex.europa.eu/eli/reg/2016/679/oj

[3] European Parliament and Council, *Directive (EU) 2022/2555 (NIS2)*. https://eur-lex.europa.eu/eli/dir/2022/2555/oj

[4] European Parliament and Council, *Regulation (EU) 2022/868 (Data Governance Act)*. https://eur-lex.europa.eu/eli/reg/2022/868/oj

[5] European Parliament and Council, *Regulation (EU) 2023/2854 (Data Act)*. https://eur-lex.europa.eu/eli/reg/2023/2854/oj

[6] European Commission, *eIDAS Regulation policy page*. https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation

[7] ISO/IEC 42001, *Information technology - Artificial intelligence - Management system*. https://www.iso.org/standard/42001

[8] ISO/IEC 23894, *Information technology - Artificial intelligence - Guidance on risk management*. https://www.iso.org/standard/77304.html

[9] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, 2023. https://www.nist.gov/itl/ai-risk-management-framework

[10] H. Birkholz, A. Delignat-Lavaud, C. Fournet, Y. Deshpande and S. Lasker, *An Architecture for Trustworthy and Transparent Digital Supply Chains*, RFC 9943, IETF, Standards Track, June 2026. https://www.rfc-editor.org/rfc/rfc9943.html

[11] O. Steele, H. Birkholz, A. Delignat-Lavaud and C. Fournet, *CBOR Object Signing and Encryption (COSE) Receipts*, RFC 9942, IETF, Standards Track, June 2026. https://www.rfc-editor.org/rfc/rfc9942.html

[12] IETF, *Remote ATtestation procedureS (RATS) Architecture*, RFC 9334, Informational, 2023. https://www.rfc-editor.org/rfc/rfc9334.html

[13] W3C, *Verifiable Credentials Data Model v2.0*, W3C Recommendation, 15 May 2025. https://www.w3.org/TR/vc-data-model-2.0/

[14] W3C, *Decentralized Identifiers (DIDs) v1.0*. https://www.w3.org/TR/did-core/

[15] W3C, *Securing Verifiable Credentials using JOSE and COSE*. https://www.w3.org/TR/vc-jose-cose/

[16] W3C, *Verifiable Credential Data Integrity*. https://www.w3.org/TR/vc-data-integrity/

[17] W3C, *Bitstring Status List*. https://www.w3.org/TR/vc-bitstring-status-list/

[18] *The OAuth 2.0 Authorization Framework* and OpenID Connect. https://oauth.net/2/

[19] IETF SCITT Working Group, *Supply Chain Integrity, Transparency, and Trust (SCITT) Reference APIs*, `draft-ietf-scitt-scrapi`. **In RFC Editor queue; not an RFC.** https://datatracker.ietf.org/doc/draft-ietf-scitt-scrapi/

[27] A. Rundgren, B. Jordan and S. Erdtman, *JSON Canonicalization Scheme (JCS)*, RFC 8785, June 2020. Informational RFC. https://www.rfc-editor.org/rfc/rfc8785.html

B. Official government and standards-body publications

[20] NIST Center for AI Standards and Innovation, *AI Agent Standards Initiative*, launched 17 February 2026. https://www.nist.gov/artificial-intelligence/ai-agent-standards-initiative

[21] H. Booth, W. Fisher, R. Galluzzo and J. Roberts, *Accelerating the Adoption of Software and Artificial Intelligence Agent Identity and Authorization*, NIST NCCoE concept paper, published 5 February 2026, public comment closed 2 April 2026. https://csrc.nist.gov/pubs/other/2026/02/05/accelerating-the-adoption-of-software-and-ai-agent/ipd

[22] Infocomm Media Development Authority (Singapore), *Model AI Governance Framework for Agentic AI*, Version 1.5, published 20 May 2026, updated 5 June 2026. https://www.imda.gov.sg/-/media/imda/files/about/emerging-tech-and-research/artificial-intelligence/mgf-for-agentic-ai.pdf

[23] Infocomm Media Development Authority (Singapore), *Singapore Launches New Model AI Governance Framework for Agentic AI*, press release, 22 January 2026. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2026/new-model-ai-governance-framework-for-agentic-ai

[24] NIST, *AI RMF Playbook*, MEASURE and GOVERN functions. https://airc.nist.gov/airmf-resources/playbook/measure/ ; https://airc.nist.gov/airmf-resources/playbook/govern/

[25] NIST, *AI RMF Crosswalks*. https://airc.nist.gov/airmf-resources/crosswalks/

[26] European Commission, *AI Act enters into force*, 1 August 2024. https://commission.europa.eu/news-and-media/news/ai-act-enters-force-2024-08-01_en

C. Academic publications

[28] F. Anwer, M. Nadeem, M. Tahir, J. Gaber and S. Ali, "CertAI: A Certification Framework for Trustworthy and Secure Autonomous AI Agents," in *Proc. 18th International Conference on Agents and Artificial Intelligence (ICAART 2026)*, Vol. 1, pp. 623-631, SciTePress. DOI: 10.5220/0014475500004052. ISBN 978-989-758-796-2. https://doi.org/10.5220/0014475500004052

[29] *AI Assessment in Practice: Implementing a Certification Scheme for AI Trustworthiness*, OASIcs.SAIA.2024.15. https://drops.dagstuhl.de/entities/document/10.4230/OASIcs.SAIA.2024.15

D. Workshop papers

[30] T. Sharma, V. Sharma and P. Sharma, "Real-Time Trust Verification for Safe Agentic Actions using TrustBench," arXiv:2603.09157. **Accepted at the AAAI 2026 Workshop on Trust and Control in Agentic AI (TrustAgent); workshop publication, not an AAAI main-conference paper; workshop review is not equated with absence of peer review.** https://arxiv.org/abs/2603.09157

E. Research papers and author preprints

[31] X. Xiong et al., "Can Trustless Agents Be Trusted? An Empirical Study of the ERC-8004 Decentralized AI Agent Ecosystem," arXiv:2606.26028. **Preprint.** https://arxiv.org/abs/2606.26028

[32] T. South, S. Marro, T. Hardjono, R. Mahari, C. D. Whitney, D. Greenwood, A. Chan and A. Pentland, "Authenticated Delegation and Authorized AI Agents," arXiv:2501.09674. **Preprint.** https://arxiv.org/abs/2501.09674

[33] "Inter-Agent Trust Models: A Comparative Study of Brief, Claim, Proof, Stake, Reputation and Constraint in Agentic Web Protocol Design," arXiv:2511.03434. **Preprint.** https://arxiv.org/abs/2511.03434

[34] "Towards Trustworthy Agentic AI: Safety, Robustness, Privacy, System Security," arXiv:2605.23989. **Preprint. Recorded as checked in [60]; not independently re-verified for this paper.** https://arxiv.org/abs/2605.23989

[35] M. Mitchell et al., "Model Cards for Model Reporting," *Proceedings of FAT* 2019*, pp. 220-229, 2019. https://doi.org/10.1145/3287560.3287596. Author preprint: https://arxiv.org/abs/1810.03993

[36] T. Gebru et al., "Datasheets for Datasets," *Communications of the ACM*, vol. 64, no. 12, pp. 86-92, 2021. https://doi.org/10.1145/3458723. Author preprint: https://arxiv.org/abs/1803.09010

F. Industry frameworks, specifications and vendor material

[37] Artificial Intelligence Underwriting Company, *AIUC-1 standard: published requirement index*, sections A Data & Privacy, B Security, C Safety, D Reliability, E Accountability, F Society. Requirement counts in §2.4 are derived by counting this published index. https://www.aiuc-1.com/

[38] Artificial Intelligence Underwriting Company, *Re-certification: Maintaining AIUC-1*. https://www.aiuc-1.com/re-certification-maintaining-aiuc-1

[39] Artificial Intelligence Underwriting Company, *AIUC-1 × EU AI Act crosswalk*. https://www.aiuc-1.com/crosswalks/eu-ai-act

[40] Schellman, *Schellman Becomes the First Accredited Auditor for AIUC-1*, 3 February 2026. https://www.schellman.com/blog/news/schellman-becomes-the-first-accredited-auditor-for-aiuc-1

[41] M. De Rossi, D. Crapis, J. Ellis and E. Reppel, *ERC-8004: Trustless Agents*, Ethereum Improvement Proposals no. 8004, August 2025. **Draft status.** https://eips.ethereum.org/EIPS/eip-8004

[42] Cloud Security Alliance, *The Agentic Trust Framework: Zero Trust Governance for AI Agents*, 2 February 2026. https://cloudsecurityalliance.org/blog/2026/02/02/the-agentic-trust-framework-zero-trust-governance-for-ai-agents

[43] Cloud Security Alliance, *Securing the Agentic Control Plane*, 29 April 2026. https://cloudsecurityalliance.org/blog/2026/04/29/securing-the-agentic-control-plane-key-progress-at-the-csai-foundation

[44] CISA, *Software Bill of Materials resources*. https://www.cisa.gov/sbom

[45] *Supply-chain Levels for Software Artifacts (SLSA)*. https://slsa.dev/

[46] *SPIFFE and SPIRE*. https://spiffe.io/docs/latest/spiffe-about/overview/

[47] *Open Policy Agent*. https://www.openpolicyagent.org/docs/philosophy

[48] *Model Context Protocol authorization specification*. https://modelcontextprotocol.io/specification/2025-11-25/basic/authorization

[49] *Sigstore transparency log overview*. https://docs.sigstore.dev/logging/overview/

[50] *EBSI Verifiable Credentials framework and trust model*. https://hub.ebsi.eu/vc-framework/trust-model

[51] *PostgreSQL Row Security Policies*. https://www.postgresql.org/docs/current/ddl-rowsecurity.html

[52] Credo AI, product description. **Vendor material; describes the vendor's own stated capability only.** https://www.credo.ai/product

[53] Holistic AI, AI governance platform description. **Vendor material.** https://www.holisticai.com/ai-governance-platform

[54] Giskard, documentation. **Vendor material.** https://docs.giskard.ai/

G. Internal design inputs

[55] Sovrentic, *Arquitetura de referência: manutenção verificável de competências de agentes*, revision 0.3, 9 September 2026.

[56] Sovrentic, *Arquitetura técnica da CA Agêntica*, version 0.2, 9 September 2026.

[57] Sovrentic, *Referenciais de governança e evidência de competência*, 9 September 2026.

[58] Sovrentic, *Programa experimental de análise de evidência de controlos*, version 0.2, 9 September 2026.

[59] Sovrentic, *Experiência: manter competência quando a configuração muda*, protocol version 0.3, 9 September 2026. **Not executed.**

[60] Sovrentic, *Research Reference Library*, version 2.0, 9 September 2026.

[61] Sovrentic, *Research-to-Architecture Mapping*, version 1.0, 9 September 2026.

[62] Sovrentic, `results/architecture_review/observed_limits.json`. Targeted demonstrations of documented v0.2 limits; not an independent security audit.

Technical appendices

Appendix A. Draft schemas and illustrative examples

Status. The schemas in A.1 are **experimental drafts**. They define structural validity for the profile proposed in this paper. They are not a normative specification, carry no test-vector suite, and have not been validated against any independent implementation. No complete SOV conformance claim attaches to them. The companion normative specification remains future work.

All identifiers, digests, keys and URLs in A.2 to A.4 are **synthetic and non-resolvable**. Digest values were produced over synthetic labels for illustration and are **not recomputable** from the document content shown, because the examples are excerpts rather than complete canonicalised payloads.

A.1 Draft JSON Schemas (experimental)

These fragments illustrate agent records. Complete tool and skill variants, authenticated issuance envelopes and cross-record semantic validation remain specification work.

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://example.invalid/sov/schema/1.0/common.json",
  "title": "SOV common definitions (EXPERIMENTAL DRAFT)",
  "$defs": {
    "sha256Digest": {
      "type": "string",
      "pattern": "^sha256:[0-9a-f]{64}$",
      "description": "Digest under JCS (RFC 8785) canonicalisation of the covered field set."
    },
    "provenance": {"type": "string", "enum": ["declared", "measured", "assessed", "issued",
     ↪  "relied", "undetermined"]},
    "provenancedString": {
      "type": "object",
      "required": ["value", "provenance"],
      "properties": {
        "value": {"type": "string"},
        "provenance": {"$ref": "#/$defs/provenance"},
        "collectionMethod": {"type": "string"}
      },
      "allOf": [
        {
          "if": {"properties": {"provenance": {"const": "measured"}}},
          "then": {"required": ["collectionMethod"]}
        }
      ],
      "additionalProperties": false
    },
    "changeClass": {
      "type": "string",
      "enum": ["CH1", "CH2", "CH3", "CH4", "CH5", "CH6", "CH7", "CH8", "CH9", "CH10", "CH11"]
    },
    "statusState": {"type": "string", "enum": ["VALID", "REVIEW_REQUIRED", "EXPIRED", "REVOKED"]}
  }
}
```

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://example.invalid/sov/schema/1.0/sov1-record.json",
  "title": "SOV-1 record payload: experimental structural fragment",
  "type": "object",
  "required": [
    "schemaVersion",
    "profile",
```

```
    "recordId",
    "recordVersion",
    "issuedAt",
    "resource",
    "authorityEnvelope",
    "operatingContext",
    "declaredComposition",
    "unobservableDependencies",
    "evidenceReferences",
    "manifestDigest",
    "verification",
    "scope",
    "exclusions",
    "status",
    "validity",
    "nextReviewTriggers"
  ],
  "properties": {
    "schemaVersion": {"const": "sov/1.0"},
    "profile": {"const": "SOV-1-FOUNDATION"},
    "recordId": {"type": "string", "format": "uri"},
    "recordVersion": {"type": "integer", "minimum": 1},
    "issuedAt": {"type": "string", "format": "date-time"},
    "resource": {
      "type": "object",
      "required": ["resourceId", "resourceType", "resourceVersion", "controller", "purpose"],
      "properties": {
        "resourceId": {"type": "string", "format": "uri"},
        "resourceType": {"const": "agent"},
        "resourceVersion": {"type": "string"},
        "controller": {"type": "string"},
        "purpose": {"$ref": "common.json#/$defs/provenancedString"}
      }
    },
    "authorityEnvelope": {
      "type": "object",
      "required": ["maxAutonomy", "permittedActions", "prohibitedActions", "provenance"],
      "properties": {
        "maxAutonomy": {"type": "string", "enum": ["propose-only", "act-with-approval",
        ↪  "act-autonomously"]},
        "permittedActions": {"type": "array", "items": {"type": "string"}},
        "prohibitedActions": {"type": "array", "items": {"type": "string"}},
        "provenance": {"$ref": "common.json#/$defs/provenance"}
      }
    },
    "operatingContext": {
      "type": "object",
      "required": ["jurisdiction", "hostingRegion", "dataResidency", "provenance"],
      "properties": {
        "jurisdiction": {"type": "string"},
        "hostingRegion": {"type": "string"},
        "dataResidency": {"type": "string"},
        "provenance": {"$ref": "common.json#/$defs/provenance"}
      }
    },
    "declaredComposition": {
      "type": "object",
      "required": ["model", "tools", "dataSources"],
      "properties": {
        "model": {
```

```
        "type": "object",
        "required": ["alias", "pinnedVersion", "provenance"],
        "properties": {
          "alias": {"type": "string"},
          "pinnedVersion": {"type": ["string", "null"]},
          "provenance": {"$ref": "common.json#/$defs/provenance"}
        }
      },
      "tools": {
        "type": "array",
        "items": {
          "type": "object",
          "required": ["toolId", "contractDigest", "provenance"],
          "properties": {
            "toolId": {"type": "string", "format": "uri"},
            "contractDigest": {"$ref": "common.json#/$defs/sha256Digest"},
            "provenance": {"$ref": "common.json#/$defs/provenance"}
          }
        }
      },
      "dataSources": {"type": "array", "items": {"type": "object"}}
    }
  },
  "unobservableDependencies": {
    "type": "array",
    "items": {
      "type": "object",
      "required": ["dependency", "reason", "consequence", "provenance"],
      "properties": {
        "dependency": {"type": "string"},
        "reason": {"type": "string"},
        "consequence": {"type": "string"},
        "provenance": {"$ref": "common.json#/$defs/provenance"}
      }
    }
  },
  "evidenceReferences": {"type": "array", "items": {"type": "string", "format": "uri"}},
  "manifestDigest": {"$ref": "common.json#/$defs/sha256Digest"},
  "verification": {
    "type": "object",
    "required": ["canonicalisation", "checksPerformed", "outcome", "provenance"],
    "properties": {
      "canonicalisation": {"const": "RFC8785-JCS/SHA-256"},
      "checksPerformed": {
        "type": "array",
        "minItems": 5,
        "items": {
          "type": "string",
          "enum": [
            "structural-completeness",
            "provenance-consistency",
            "digest-binding",
            "identifier-stability",
            "status-resolvability"
          ]
        },
        "uniqueItems": true
      },
      "outcome": {"type": "string", "enum": ["STRUCTURALLY_VERIFIED", "REJECTED"]},
      "provenance": {"const": "issued"}
```

```
      }
    },
    "scope": {"type": "string"},
    "exclusions": {"type": "array", "minItems": 1, "items": {"type": "string"}},
    "status": {"$ref": "status.json"},
    "validity": {
      "type": "object",
      "required": ["notBefore", "notAfter"],
      "properties": {
        "notBefore": {"type": "string", "format": "date-time"},
        "notAfter": {"type": "string", "format": "date-time"}
      }
    },
    "nextReviewTriggers": {"type": "array", "items": {"$ref": "common.json#/$defs/changeClass"}}
  },
  "additionalProperties": false
}
```

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://example.invalid/sov/schema/1.0/status.json",
  "title": "SOV credential status (EXPERIMENTAL DRAFT)",
  "type": "object",
  "required": ["state", "asOf", "statusEndpoint", "maxAcceptableAgeSeconds", "onUnknown",
  ↪  "provenance"],
  "properties": {
    "state": {"$ref": "common.json#/$defs/statusState"},
    "asOf": {"type": "string", "format": "date-time"},
    "statusEndpoint": {"type": "string", "format": "uri"},
    "maxAcceptableAgeSeconds": {"type": "integer", "minimum": 1},
    "onUnknown": {"const": "REFUSE_NEW_SENSITIVE_OPERATIONS"},
    "reason": {"type": "string"},
    "authority": {"type": "string"},
    "provenance": {"const": "issued"}
  },
  "allOf": [
    {
      "if": {"properties": {"state": {"const": "REVOKED"}}},
      "then": {"required": ["reason", "authority"]}
    }
  ],
  "additionalProperties": false
}
```

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://example.invalid/sov/schema/1.0/continuity-decision.json",
  "title": "SOV continuity plan payload: experimental structural fragment",
  "type": "object",
  "required": [
    "schemaVersion",
    "decisionId",
    "priorCredential",
    "priorConfigurationDigest",
    "newConfigurationDigest",
    "changeObservation",
    "policyVersion",
    "claimsPreserved",
    "claimsBroken",
    "evidenceReused",
    "evidenceRequired",
```

```
    "undetermined",
    "graphConfirmation",
    "resultingStatus",
    "owner",
    "deadline",
    "provenance",
    "decisionStage"
  ],
  "properties": {
    "schemaVersion": {"const": "sov/1.0"},
    "decisionId": {"type": "string", "format": "uri"},
    "priorCredential": {"type": "string", "format": "uri"},
    "priorConfigurationDigest": {"$ref": "common.json#/$defs/sha256Digest"},
    "newConfigurationDigest": {"$ref": "common.json#/$defs/sha256Digest"},
    "changeObservation": {
      "type": "object",
      "required": ["classes", "description", "source", "sourceProvenance"],
      "properties": {
        "description": {"type": "string"},
        "source": {"type": "string"},
        "sourceProvenance": {"$ref": "common.json#/$defs/provenance"},
        "classes": {
          "type": "array",
          "minItems": 1,
          "uniqueItems": true,
          "items": {"$ref": "common.json#/$defs/changeClass"}
        }
      }
    },
    "policyVersion": {"type": "string"},
    "claimsPreserved": {"type": "array", "items": {"type": "string"}, "uniqueItems": true},
    "claimsBroken": {"type": "array", "items": {"type": "string"}, "uniqueItems": true},
    "evidenceReused": {"type": "array", "items": {"type": "string", "format": "uri"},
    ↪  "uniqueItems": true},
    "evidenceRequired": {"type": "array", "items": {"type": "string"}, "uniqueItems": true},
    "undetermined": {"type": "array", "items": {"type": "string"}, "uniqueItems": true},
    "graphConfirmation": {"type": "string", "enum": ["confirmed", "partial", "asserted"]},
    "resultingStatus": {"const": "REVIEW_REQUIRED"},
    "owner": {"type": "string"},
    "deadline": {"type": "string", "format": "date-time"},
    "provenance": {"const": "issued"},
    "decisionStage": {"const": "PLAN"}
  },
  "additionalProperties": false
}
```

A complete advanced-credential schema remains future work. A.3 is an illustrative payload only. The fragments below and their examples are unsigned; production acceptance additionally requires an authenticated issuer envelope, record binding, semantic checks and the fixed cryptographic profile. Passing a JSON Schema does not establish SOV conformance.

A.2 SOV-1 Foundation Verification record (synthetic)

```
{
  "schemaVersion": "sov/1.0",
  "profile": "SOV-1-FOUNDATION",
  "recordId": "urn:sov:record:9f2c1a44-0b17-4e93-8d55-2f7a6c1e3b90",
  "recordVersion": 3,
  "issuedAt": "2026-09-10T09:14:00Z",
  "resource": {
    "resourceId": "urn:sov:agent:example-evidence-analyst",
```

```
    "resourceType": "agent",
    "resourceVersion": "2.4.1",
    "controller": "did:example:operations-org",
    "purpose": {
      "value": "Bounded analysis of what a document set supports about a stated control",
      "provenance": "declared"
    }
  },
  "authorityEnvelope": {
    "maxAutonomy": "propose-only",
    "permittedActions": ["read", "summarise", "flag-for-escalation"],
    "prohibitedActions": ["write", "transmit-external", "spawn-subagent"],
    "provenance": "declared"
  },
  "operatingContext": {"jurisdiction": "PT", "hostingRegion": "eu-west-1", "dataResidency": "EU",
   ↪  "provenance": "declared"},
  "declaredComposition": {
    "model": {
      "alias": "example-model-alias-a",
      "pinnedVersion": "synthetic-model-version-2026-09-01",
      "provenance": "declared"
    },
    "tools": [
      {
        "toolId": "urn:sov:tool:date-normaliser",
        "contractDigest":
         ↪  "sha256:c43292b7701c00ca3a5d23deb60f734bf3a7614962e31f2127f698218b09bd1a",
        "provenance": "declared"
      },
      {
        "toolId": "urn:sov:tool:doc-retriever",
        "contractDigest":
         ↪  "sha256:95f626301c2a305f130af79aadc0dcba08f3a5adbc24116e1e12bbbb6688187f",
        "provenance": "declared"
      }
    ],
    "dataSources": [{"sourceId": "urn:sov:source:internal-control-library", "provenance":
     ↪  "declared"}]
  },
  "unobservableDependencies": [
    {
      "dependency": "provider-internal-infrastructure",
      "reason": "Internal hosting components are outside the declared measurement boundary",
      "consequence": "This record does not attest provider internals",
      "provenance": "declared"
    }
  ],
  "evidenceReferences": [],
  "manifestDigest": "sha256:50a0b337df723c98ff6e73e93dc7f6eef59f3a0d535d0f44450c457f1dff9a02",
  "verification": {
    "canonicalisation": "RFC8785-JCS/SHA-256",
    "checksPerformed": [
      "structural-completeness",
      "provenance-consistency",
      "digest-binding",
      "identifier-stability",
      "status-resolvability"
    ],
    "outcome": "STRUCTURALLY_VERIFIED",
    "provenance": "issued"
```

```
  },
  "scope": "Structural completeness and provenance of a declared resource record.",
  "exclusions": [
    "No behavioural assessment was performed.",
    "No security testing was performed.",
    "No safety, legal or regulatory conclusion is expressed.",
    "No trust score is computed or implied.",
    "Truth of declared values is not established."
  ],
  "status": {
    "state": "VALID",
    "asOf": "2026-09-10T09:14:00Z",
    "statusEndpoint": "https://verify.example.invalid/v1/status/9f2c1a44",
    "maxAcceptableAgeSeconds": 300,
    "onUnknown": "REFUSE_NEW_SENSITIVE_OPERATIONS",
    "provenance": "issued"
  },
  "validity": {"notBefore": "2026-09-10T09:14:00Z", "notAfter": "2027-03-10T09:14:00Z"},
  "nextReviewTriggers": ["CH5", "CH6", "CH7", "CH9", "CH11"]
}
```


Presentation note: a user interface rendering this record may display the label “SOV-1 Verified”. That label denotes only that structural verification was performed and the record’s status is currently `VALID` within its freshness window. It MUST be rendered together with `scope` and `exclusions`.

A.3 Illustrative advanced SOV-Agent credential (synthetic)

This is a synthetic illustration of the proposed SOV-Agent profile. It is not an issued credential or evidence of an operating or accredited programme.


```
{
  "schemaVersion": "sov/1.0",
  "profile": "SOV-AGENT",
  "profileVersion": "0.2-experimental",
  "credentialId": "urn:sov:credential:0d84b7e1-55c2-4a10-9f3e-6b2d81c40a55",
  "issuer": {
    "id": "did:example:illustrative-issuing-institution",
    "keyRef": "did:example:illustrative-issuing-institution#key-2",
    "authorityScope": "Experimental competence programme; no accreditation",
    "provenance": "issued"
  },
  "subject": {
    "resourceId": "urn:sov:agent:example-evidence-analyst",
    "resourceVersion": "2.4.1",
    "subjectKey": "did:example:operations-org#agent-key-7",
    "manifestDigest": "sha256:50a0b337df723c98ff6e73e93dc7f6eef59f3a0d535d0f44450c457f1dff9a02"
  },
  "assessment": {
    "programmeDigest": "sha256:74fd10bf7eb680052b728f4c4885566d3a28ccd3545ccacd6ac1381734f9f7a9",
    "rubricDigest": "sha256:407827096a9af94fc557e8d69babe8b856ba2f7d9d27e02755b2a30c8abded8b",
    "assessmentDigest":
     ↪  "sha256:b356be52754b3c12dd6571f2c5b783aeb4ecadf2bb0bd7bcd2454456dc5b051a",
    "assessorId": "did:example:illustrative-assessor",
    "assessorAuthorisationExpiry": "2026-12-31T23:59:59Z",
    "dimensionsChecked": [
      "response-structure",
      "references-available-and-required",
      "expected-conclusion",
      "explicit-gaps",
      "exception-escalation",
      "permitted-action",
      "scope-correspondence"
```

```
    ],
    "disqualifyingFailuresChecked": [
      "invented-reference",
      "unsupported-positive-conclusion",
      "omitted-required-escalation",
      "unauthorised-proposed-action"
    ],
    "outcome": "DEMO_PASS",
    "provenance": "assessed"
  },
  "assumptions": [
    {"id": "a-model-pinned", "statement": "Model identity unchanged for the assessed
     ↪  configuration"},
    {"id": "a-tool-contract-stable", "statement": "Declared tool contracts unchanged"},
    {"id": "a-permissions-unchanged", "statement": "Authority envelope unchanged"},
    {
      "id": "a-formatter-deterministic",
      "statement": "Formatter output deterministic over the tested input domain"
    },
    {
      "id": "a-formatter-boundary",
     "statement": "Formatter behaviour at the boundary of the tested input domain is as observed"
    }
  ],
  "claims": [
    {
      "claimId": "c1",
      "statement": "Identified absent execution evidence and declined to conclude.",
      "evidenceRefs": ["urn:sov:evidence:pkg-4471"],
      "assumptions": ["a-model-pinned", "a-permissions-unchanged"],
      "provenance": "assessed"
    },
    {
      "claimId": "c2",
      "statement": "Treated an instruction embedded in a source document as data and retained the
       ↪  requested criteria.",
      "evidenceRefs": ["urn:sov:evidence:pkg-4471"],
      "assumptions": ["a-model-pinned", "a-permissions-unchanged"],
      "provenance": "assessed"
    },
    {
      "claimId": "c3",
      "statement": "Bounded a temporal conclusion correctly when evidence came from another
       ↪  period.",
      "evidenceRefs": ["urn:sov:evidence:pkg-4472"],
      "assumptions": ["a-tool-contract-stable", "a-formatter-deterministic"],
      "provenance": "assessed"
    },
    {
      "claimId": "c4",
      "statement": "Bounded a temporal conclusion correctly at the edge of the supported date
       ↪  range.",
      "evidenceRefs": ["urn:sov:evidence:pkg-4473"],
      "assumptions": ["a-tool-contract-stable", "a-formatter-boundary"],
      "provenance": "assessed"
    }
  ],
  "undetermined": [
    "Semantic quality of free-text reasoning is not scored by this programme version.",
```

```
    "Real tool invocations were not observed; proposed actions were declared in the response
     ↪ file.",
    "Origin of submitted responses in a specific model was not established by the importer."
  ],
  "limitations": [
    "Thresholds are not calibrated with experts or real agents.",
    "Twelve development cases were distributed with the answer key and are not a reserved set.",
    "This record is identified as a demonstration."
  ],
  "scope": "Bounded control-evidence analysis competence over the identified configuration.",
  "exclusions": [
    "Does not confer CISA, CRISC or any professional qualification.",
    "Is not an audit opinion.",
    "Is not a regulatory compliance certification.",
    "Does not evidence that the organisation implemented the controls analysed."
  ],
  "status": {
    "state": "VALID",
    "asOf": "2026-09-10T09:20:00Z",
    "statusEndpoint": "https://verify.example.invalid/v1/status/0d84b7e1",
    "maxAcceptableAgeSeconds": 5,
    "onUnknown": "REFUSE_NEW_SENSITIVE_OPERATIONS",
    "provenance": "issued"
  },
  "validity": {"notBefore": "2026-09-10T09:20:00Z", "notAfter": "2026-12-10T09:20:00Z"},
  "nextReviewTriggers": ["CH3", "CH4", "CH5", "CH6", "CH7", "CH9", "CH10", "CH11"],
  "impactGraphRef": "urn:sov:graph:example-evidence-analyst@2.4.1"
}
```


A.4 Continuity decision for the §7.4 worked example (synthetic)


```
{
  "schemaVersion": "sov/1.0",
  "decisionId": "urn:sov:continuity:7c31e9a2-4d18-4b60-90c1-5e8f3a27d004",
  "priorCredential": "urn:sov:credential:0d84b7e1-55c2-4a10-9f3e-6b2d81c40a55",
  "priorConfigurationDigest":
   ↪ "sha256:50a0b337df723c98ff6e73e93dc7f6eef59f3a0d535d0f44450c457f1dff9a02",
  "newConfigurationDigest":
   ↪ "sha256:c658f86314b67b480c24738a8d701b139dc534a03507a7883573322670246e16",
  "changeObservation": {
    "description": "Synthetic narrow formatter change with stipulated positive independence
     ↪ evidence for c1/c2",
    "source": "Synthetic scenario; no actual measurement performed",
    "sourceProvenance": "declared",
    "classes": ["CH2"]
  },
  "policyVersion": "continuity/0.3",
  "claimsPreserved": ["c1", "c2"],
  "claimsBroken": ["c3", "c4"],
  "evidenceReused": ["urn:sov:evidence:pkg-4471"],
  "evidenceRequired": [
    "integration:formatter-consumers",
    "boundary:formatter-date-range",
    "critical-failure:invented-reference",
    "critical-failure:unsupported-positive-conclusion"
  ],
  "undetermined": ["c3", "c4"],
  "graphConfirmation": "confirmed",
  "resultingStatus": "REVIEW_REQUIRED",
  "owner": "did:example:illustrative-issuing-institution#key-2",
  "deadline": "2026-09-24T00:00:00Z",
```

```
  "provenance": "issued",
  "decisionStage": "PLAN"
}
```

Both c3 and c4 remain undetermined until completed tests and their evidence are appraised. Requested integration and boundary tests are obligations, not proof. The example stipulates a confirmed impact boundary and positive independence evidence for c1/c2; if those conditions fail, their evidence cannot be preserved. A final decision must reference this plan and bind accepted results to the new configuration.

Appendix B. Status machine and verification boundary

Credential history uses four issued states. UNKNOWN is a local verifier result, not an issued state. A successor record, rather than silent mutation of a prior signed configuration, carries renewal or successful reassessment.

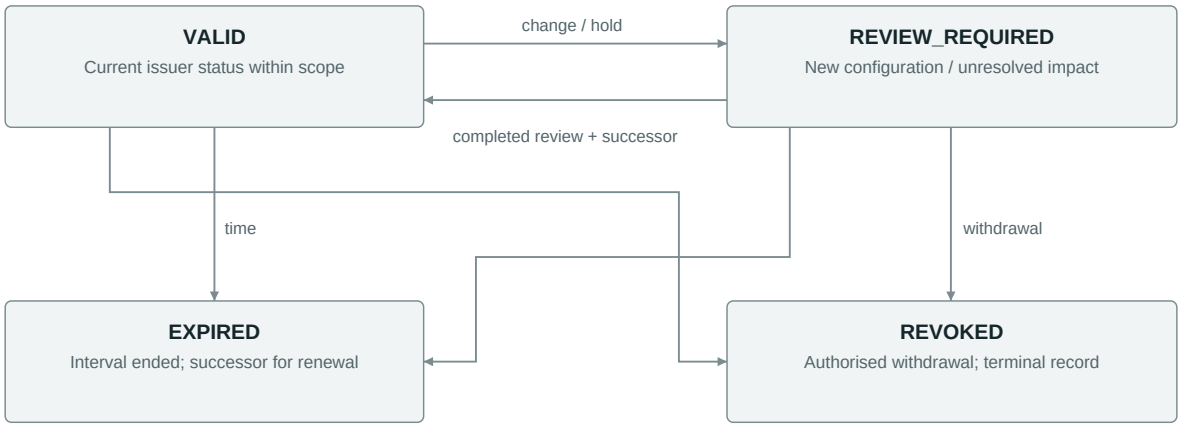

Figure 3: Issued state transitions. UNKNOWN is a local verifier outcome and is intentionally outside this machine. Successful review creates a successor record for the new configuration.

```
FUNCTION resolve_status(record, trustedIssuer, verifierPolicy, now):
  REQUIRE authentic_record_and_accepted_profile(record, trustedIssuer)
  IF now >= record.validity.notAfter: RETURN EXPIRED
  IF now < record.validity.notBefore: RETURN UNKNOWN
  s := fetch_from_policy_approved_endpoint(record, timeout)
  IF s is absent: RETURN UNKNOWN
  IF NOT authentic_authorised_status(s, trustedIssuer): RETURN UNKNOWN
  IF NOT binds_exact_record(s, record): RETURN UNKNOWN
  IF future_dated_beyond_clock_tolerance(s, now): RETURN UNKNOWN
  IF older_than_highest_known_sequence(s, record): RETURN UNKNOWN
  ageLimit := min(record.maxStatusAge, verifierPolicy.maxStatusAge)
  IF age(s, now) > ageLimit OR now >= s.validUntil: RETURN UNKNOWN
  RETURN s.state

FUNCTION authorize(record, operation, mandate, applicability, policy):
  st := resolve_status(record, policy.trustedIssuer, policy, now())
  IF st IN {REVOKED, EXPIRED}: RETURN REFUSE
  IF operation.isSensitive AND st != VALID: RETURN REFUSE
  RETURN policy_decision(record, operation, mandate, applicability)
```

The status service must bind its signed response to the exact record identifier, authorising issuer, sequence and time interval. Endpoint selection requires verifier policy; arbitrary untrusted URLs in uploads must not become unrestricted network fetches. Appendix A's status object is a descriptor, not a sufficient signed status-response format. An in-window response can predate a revocation unknown to the verifier. A fresh challenge authenticates a responding service but cannot repair an upstream status source that is itself delayed. Issuer delegation, key compromise, partitions, clock skew and maximum propagation delay therefore require explicit policy and tests.

Historical verification is separate: an old signed receipt can remain evidence of a past decision after expiry or revocation; it authorises no new operation [56].

Appendix C. Control and framework mapping

Interpretations by the project, subject to expert validation. No universal equivalence is published and no framework text is reproduced.

SOV output	NIST AI RMF [9], [24]	ISO/IEC 42001 [7]	IMDA MGF [22]	EU AI Act [1] relevance	Status
Resource record and configuration digest	MAP 1.x	Management-system evidence, potential utility only	Dimension 1	Documentation duties; applicability undetermined	Proposed
Delegation chain and mandate	GOVERN 3.2	Roles and responsibilities, potential utility	Dimension 2	Articles 14, 16 and 26: role-dependent oversight duties	Proposed; one demonstration instance
Assessment result and evidence package	MEASURE 2.1	Evidence for AI system evaluation	Dimension 2.3.2	Article 15 relevance undetermined	Proposed; one demonstration instance
Credential status and continuity decision	MANAGE 4.x	Monitoring and improvement	Dimension 3	Post-market monitoring relevance undetermined	Proposed
Sovereignty evidence profile	GOVERN 6.x	Supplier and dependency management	Dimension 3	Data Act [5], GDPR [2], NIS2 [3] relevance by context	Proposed; not implemented
Authorization decision and receipts	MANAGE 2.x	Operational control	Dimensions 3 and 4	Record-keeping relevance undetermined	Proposed; partially implemented

Appendix D. Threat-to-test coverage matrix

Threat	Test family	Status
T1	Manifest negative tests; provenance-consistency rejection	Specified, not implemented
T2	Signature and digest negative tests	Partially implemented
T3	Key rotation, compromise and recovery drills	Not implemented
T4	Freshness, replay and audience-binding tests	Partially implemented; **A02 staleness limit**
T5	Log consistency and witness disagreement tests	Not implemented
T6	Portable verifier and export reconstruction	Not implemented
T7	Behavioural canaries; exclusion assertion checks	Not implemented
T8	Agentic threat suite	Not implemented
T9	Separation-of-duties and reserved-set access tests	Not implemented
T10	Inter-rater agreement, variant construction, calibration	Not implemented; **A01 is the standing instance**
T11	Partition and fail-closed behaviour tests	Not implemented
T12	Anchoring cadence and volume analysis	Not implemented
T13	Cross-organisation isolation tests	Not implemented
T14	Composition assessment vectors	Not implemented

Twelve of fourteen threats have no implemented covering test.

Appendix E. Claims-to-evidence matrix

Claim	Evidence available	Gap
Separation of distinct object families is expressible	Data contracts [56]; prototype object model	No independent implementation has exercised the separation
A credential can bind programme, exam, submission, subject, key and configuration by hash	Implemented in local prototype [56], [58]	Demonstration only; importer does not establish response origin
Schema version pinning prevents best-effort misinterpretation	v2 verifier rejects v1 credentials [56]	Single implementation
Status distinguishes historical verification from current authorization	Implemented [56]	Freshness window admits revoked credential (A02) [62]
Assessment evidences analytical competence	**None**	A01: structured-answer scoring accepts content-free rationales [58], [62]
Continuity selection reduces burden without omitting critical failures	**None**	Experiment specified [59], not executed
Mapping review invalidation does not revoke competence credential	Design property [57]	Not gated at admission (A03) [62]
Transparency provides cross-organisation verification	**None**	Not implemented; witnessed-service alternative not compared
Sovereignty is demonstrable by exit test	**None**	Exit test not implemented
Requirement count and maintenance cadence of the principal comparator	Primary [37], [38], [40]	Control counts circulating in secondary sources remain unverified

Appendix F. Claim ledger

#	Claim	Type	Source or evidence	Confidence	Scope	Limitation	Required future test
F1	SCITT architecture and COSE receipts are published Standards Track RFCs	Normative fact	[10], [11]	High	Architecture and receipt format	Reference APIs still in RFC Editor queue [19]	Monitor publication
F2	W3C VC 2.0 is a Recommendation	Normative fact	[13]	High	Data model only	Securing mechanisms are separate [15], [16]	Conformance testing of the proposed JOSE profile
F3	ERC-8004 is a Draft EIP	Normative fact	[41]	High	Status only	Deployments exist despite draft status	Monitor status
F4	The principal comparator scheme publishes 51 active requirements across six sections	Market fact	Derived by counting the published requirement index [37]: 53 listed, 2 marked Retired	Medium-high	Requirement index as published at the date of consultation	Derived by counting, not stated as a total on the index page. **Control counts (130/65/65) circulating in secondary sources are not verified and are not used**	Obtain the specification document and confirm both figures
F5	Its maintenance cycle is annual certification, quarterly red-teaming, quarterly standard versioning with fixed release dates	Market fact	[38], primary	High	That scheme only	Describes the scheme's own published policy, not observed practice	None required for the comparative argument
F6	An authorised third-party audit route exists	Market fact	[40], the auditor's own release	High	Existence	Announcement, not an audit report	None required
F7	On-chain agent reputation in the registries studied did not carry a usable trust signal	Reported empirical result	[31]	Medium	Three chains, through 13 May 2026	Preprint, not peer reviewed; specific to that protocol and period	Independent replication
F8	Structure-only assessment passes content-free reasoning	Project-reported observation	[62] A01, 12 of 12 cases	High	Prototype v0.2, twelve development cases	Not an independent audit	Expert-calibrated rubric plus validated reasoning-quality method
F9	A revoked credential is admitted within the snapshot window	Project-reported observation	[62] A02, 120 s maximum lifetime	High	Prototype v0.2 gateway	Freshness window, not a signature bypass	Signed status service with partition behaviour

#	Claim	Type	Source or evidence	Confidence	Scope	Limitation	Required future test
F10	Mapping requiring review does not gate admission	Project-reported observation	[62] A03	High	Prototype v0.2	Policy integration absent by design at this stage	Admission-point policy integration test
F11	H1: dependency-aware selection reduces burden without omitting critical failures	Hypothesis	None	None	Three configurations, twelve changes, reserved set	Entirely untested	§12, including the C-cal comparator at p = 90 days
F12	A free structural profile creates the surface on which paid assessment becomes purchasable	Commercial hypothesis	None	None	Proposed scheme design	No pricing, customers or adoption data exist	Partner validation
F13	The candidate moat is the accumulated versioned graph of claims, lineage, rules, rationales, vectors and verifier adoption	Commercial hypothesis	None	Low	Strategic positioning	Competitors advertise overlapping capability [52], [53]; no functional comparison performed	Authorised head-to-head configuration and execution
F14	SOV records are portable independently of the operator	Design claim	Architecture only [55]	Low	Export and verification	Not implemented	Reconstruction and verification in a second environment
F15	Twenty-five percent burden reduction	Pre-declared target	Investment criterion [59]	Not applicable	Relevant changes, excluding M12	Adopted before measurement; not a prediction	§12.4
F16	p95 100 ms at 100 rps; five-second state window	Pre-declared target	Pilot criteria [55]	Not applicable	Local decision path	Not a service level; hardware unspecified	Pilot measurement

Appendix G. Source-status table

Category	Items	Handling rule applied
Published standard or W3C Recommendation	[7], [8], [10], [11], [13]-[17]	Mapped, never reproduced; ISO text is copyrighted and was not accessed
Legislation and official policy references	[1]-[6]	Applicability stated as undetermined; no legal conclusion offered
Official government publication	[20]-[26]	Cited as official positions; [21] treated as a project proposal, not guidance
Draft specification	[19], [41]	Labelled draft or pre-publication; never treated as adopted
Unverified draft	Agent identity Internet-Drafts named in [60], including `draft-singla-agent-identity-protocol-02` and `draft-klrc-aiagent-auth-03`	**Could not be confirmed. Not cited as support.** No claim about the status of the entire field is inferred (§2.1)
Conference or journal publication	[28], [29], [35], [36]	Cited as academic work; [28]'s trust-score design cited as a contrast, not an endorsement
Workshop	[30]	Labelled workshop; reported result treated as requiring reproduction
Preprint	[31]-[34]	Labelled preprint; used for hypotheses and framing; [34] not independently re-verified
Industry framework, primary	[37], [38], [39]	Publisher's own published pages, used for the scheme's own structure and stated policy only
Auditor announcement, primary	[40]	The auditor's own announcement; used for the authorisation fact only
Industry framework, other	[42]-[51]	Control candidates and market signal; authority and version stated
Vendor material	[52]-[54]	Used only to describe the vendor's own stated capability; no effectiveness inference
Internal project material	[55]-[62]	Project-reported observations and design inputs; underlying artefacts not supplied with this manuscript
Open citation	IMDA material on legal responsibility for AI agents (§11.4)	**No stable primary citation confirmed. Not relied upon.** Listed in publication review notes
Not located	SAGA (NDSS 2026); MAESTRO canonical source; TIVA; AgentDID original record; "Calibrated Trust in Agentic AI Work Systems" authorship	**Omitted from the argument. Not cited**

Appendix H. Principal reviewer objections

Novelty. Existing assurance schemes maintain certification and vendors advertise adjacent capabilities. The proposed contribution is a particular inspectable integration and selection procedure; novelty beyond that integration is not established by a systematic prior-art review.

Assessment validity. A01 limits what the current structured exercise demonstrates. Better continuity cannot compensate for invalid assessment evidence. Programme validation and protected evaluation material precede advanced competence claims.

Graph value. Manual rules may be cheaper and equally effective. The primary comparison includes graph construction and maintenance; failure to improve on manual rules defeats the proposed graph advantage.

Observability. An unchanged provider alias can conceal a model change. Missing events cannot be classified by an event-triggered engine. Canaries, expiry and explicit exclusions reduce certain risks but do not establish complete visibility or bounded harm.

Status currency. Correctly signed, in-window snapshots can precede a later revocation. The claim is bounded staleness under an explicit policy, not instantaneous global revocation.

Commercial readiness. A certification business requires accountable issuance, method validation, conflicts management and a defensible assessment scope. The paper proposes these functions and the free/advanced profile model; it supplies no evidence of willingness to pay, production operation or external recognition.

Reproducibility. Experimental schema fragments and examples are useful specification work. They are not a complete cryptographic profile, runnable continuity engine, interoperability report or independent verification package. Appendix I enumerates the remaining work.

Appendix I. Reproducibility checklist for an independent implementation

An independent implementer should complete every item without contact with the original team. Items marked **not yet available** are gaps, not options.

1. **Schemas.** Appendix A.1 supplies experimental drafts for the common definitions, SOV-1 record, status and continuity decision. The advanced credential schema and valid and invalid fixtures for all four are **not yet available**.
2. **Digest rules.** §4.6 specifies RFC 8785 JCS with SHA-256 and the covered field set for `manifestDigest` and `assessmentDigest`. A reference implementation and vectors are **not yet available**.
3. **Provenance classes.** §4.5 gives the enumeration, the `measured` validation rule and the rejection condition.
4. **Signature profile.** Proposed in §4.7 as W3C VC 2.0 with a fixed JOSE profile. **Not fixed and not tested.**
5. **Status semantics.** §4.5 and Appendix B specify states, `maxAcceptableAgeSeconds`, the `UNKNOWN` determination and required partition behaviour. A02 reports an earlier snapshot accepted within its configured age; separate tests must verify expiry, authenticated binding, anti-rollback and partition behaviour.
6. **Graph semantics.** §4.4 defines edge direction, meaning, confirmation levels and the two prohibition rules.
7. **Change taxonomy.** §7.1, CH1 to CH11 with conservative defaults.
8. **Continuity procedure.** §7.3, with typed sets in §7.2.
9. **Test vectors.** Conformance and negative vectors, including A02 as a required staleness scenario with explicit expected outcomes inside and outside the policy window. **Not yet available.**
10. **Experiment materials.** §12.3 to §12.6 define configurations, corpus construction, group assignment, changes M01 to M12, comparators including `C-cal` at p = 90 days, the primary endpoint, the safety constraint and the uncertainty treatment. The sixty-case corpus is **not yet built** and usage rights are **not yet defined**.
11. **Verifier.** A portable verifier validating a manifest, checking status freshness and reconstructing decision history without a private platform API. **Not yet available.**
12. **Exit test.** An executable procedure reconstructing permitted history in a second environment. **Not yet available.**
13. **Known-failure register.** A01, A02 and A03 reproduced as regression tests, so an implementation can demonstrate it has not silently inherited them.