VAC Protocol Whitepaper v4.0

Verified Authority Chain: Biometric Human Attribution for AI Agent Systems

Version: 4.1 **Date:** 26 February 2026 **Author:** Roberto Zagarella, Violet Shores Pty Ltd **Patent:** AU 2026901425 + Supplementaries AU 2026901428, AU 2026901474, AU 2026901553 (241 claims, 38 sections, 4 filings) **Protocol:** vacprotocol.org

---

Table of Contents

---

1. Introduction: The Human Attribution Problem

As AI agents become increasingly autonomous — executing multi-step tasks, delegating to sub-agents, operating across organisational and jurisdictional boundaries, and controlling physical systems — the chain of human accountability grows longer and more opaque.

Current identity frameworks authenticate agents as machine identities but do not maintain a verifiable link to the human who initiated the chain of action. OAuth authorises access. SPIFFE identifies workloads. Neither proves which human authorised the agent, with what scope, through what chain of command.

The VAC Protocol addresses this by making human attribution a first-class security property that:

- **Propagates** through arbitrarily deep agent delegation chains
- **Narrows** scope at each delegation level (never widens)
- **Persists** with cryptographic non-repudiation via biometric binding
- **Revokes** cascading through entire chains when human authority is withdrawn
- **Records** provenance for regulatory compliance across jurisdictions
- **Extends** to physical systems, collective governance, and coalition operations

What's New in v4.0

This version extends the protocol from digital agent chains into three critical new domains:

Regulatory Compliance Provenance (Sections 10-11): Provenance chains analogous to telecommunications CDRs, enabling operators to satisfy metadata retention obligations across jurisdictions. A global regulatory intelligence layer automates cross-border compliance with GDPR, LGPD, PIPL, and 10+ other frameworks.

Collective and Indigenous Data Sovereignty (Sections 12-13): Extension of multi-party authorisation to support community governance models including consensus-based decision-making (whakawhhitiwhiti kōrero), weighted authority (kaumātua verification), cultural protocol scope constraints, and optional location-enhanced governance. Supports Te Mana Raraunga, CARE Principles, and OCAP.

Physical Systems and Coalition Governance (Sections 14-15): Verified human authority over drones, robots, autonomous vehicles, and military systems. Physical operating parameters as cryptographic scope constraints. Multi-national coalition governance with rules of engagement as scope constraints and national caveats as protocol-level enforcement.

Continuous Validity Monitoring (Section 16) — NEW in v4.1: Collective authority as a continuously validated state with governance-model-aware cascade logic, authority change classification by severity, forensic gap provenance for interrupted operations, and structured reconstitution protocols.

Intelligent Orchestration and Collaborative Onboarding (Section 17) — NEW in v4.1: Natural language problem-to-agent-team recommendation, dual-purpose biometric capture (functional interaction and identity verification in one recording), vouch-as-collaboration-invitation (identity vouching simultaneously configures bilateral governance), and automated collaboration discovery between independently verified users.

---

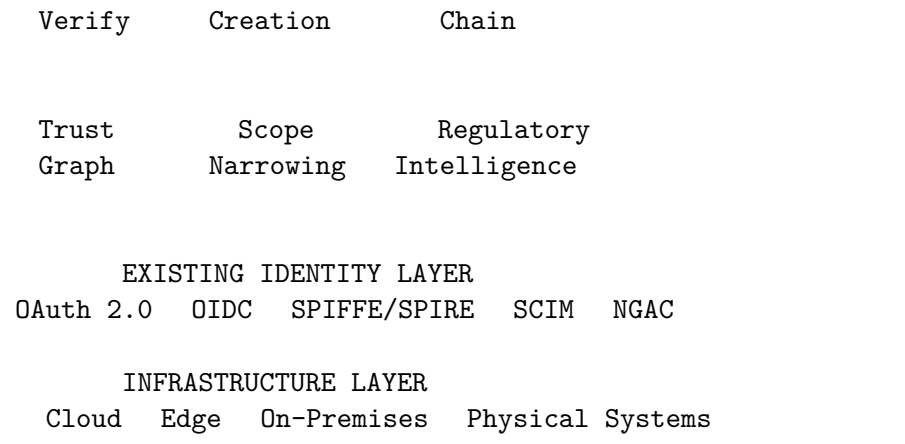
2. Protocol Architecture

The VAC Protocol operates as a layer that sits alongside existing identity and authorisation frameworks:

```
                APPLICATION LAYER
        (Healthcare, Finance, Defence, etc.)

             VAC PROTOCOL LAYER

    Biometric      VAT         Provenance
```



---

3. Verified Authority Token (VAT)

The VAT is the central cryptographic object. It is a JWT-compatible signed token using Ed25519 signatures containing:

Component	Contents
Header	Token version, signing algorithm, token type (root/derived)
Identity	Cryptographic hash of verified human identity (not plaintext PII)
Trust	Trust score, minimum threshold
Scope	Resource types, action types, data domains, temporal windows, sensitivity thresholds, physical parameters
Delegation	Max depth, current depth, parent reference, chain metadata
Context	Organisational, multi-party, jurisdictional provenance
Validity	Not-before, not-after, re-verification requirements
Signature	Ed25519 binding payload to biometric attestation

Token Lifecycle

Creation (Root VAT): Generated when a biometrically-verified human authorises an agent operation.

Derivation (Delegated VAT): Created by coordinator agents when delegating to specialists. Strict narrowing: scope is set intersection, trust decays, depth increments.

Verification: Any party can verify the complete chain from leaf action to root human.

Revocation: Cascades immediately through all derived tokens.

---

4. Multi-Modal Biometric Verification

VAC uses multi-modal biometrics for human verification:

- **Facial geometry** — liveness detection, anti-spoofing
- **Voice pattern** — speaker verification, anti-replay
- **Behavioural biometrics** — typing patterns, device interaction
- **Optional location** — multi-modal location verification when configured

Multi-modal composite scoring provides higher assurance than any single modality. The composite trust score maps to NIST SP 800-63 Identity Assurance Levels.

---

5. Trust Score Computation

Trust scores are continuous values computed from:

- Biometric verification recency and quality
- Number of modalities verified
- Historical verification consistency
- Social trust graph edges (vouching from other verified humans)
- Organisational authority context

Trust decays over time and with delegation depth. Configurable thresholds determine minimum trust for different action categories.

---

6. Delegation Chains and Scope Narrowing

The fundamental rule: **authority flows downhill and can only get narrower.**

```
Human (Trust: 0.95, Scope: Full)
    Coordinator Agent (Trust: 0.88, Scope: Financial)
         Reconciliation Agent (Trust: 0.82, Scope: Read-Only Financial)
         Compliance Agent (Trust: 0.82, Scope: Regulatory Data)
         Reporting Agent (Trust: 0.82, Scope: Report Generation)
```

Each derivation: `derived_scope = parent_scope  requested_scope` Each derivation: `derived_trust = f(parent_trust, depth, agent_properties)`

---

7. Organisational Trust Hierarchies

Organisations register as trust roots with hierarchical authority structures. Authority flows from organisational root through departments, teams, and individuals. Cross-organisational agent interaction verifies both organisations' authority chains.

---

8. Multi-Party Biometric Authorisation

M-of-N threshold verification: multiple biometrically-verified humans must authorise before a collective VAT is issued. Supports:

- Threshold governance (3-of-5 board members)
- Sequential approval chains
- Time-windowed multi-party verification
- Biometric veto (any M-of-N participant can block)

---

9. Agent Chain Trust Propagation

VATs propagate through arbitrarily deep chains:

```
Human → Coordinator → Sub-coordinator → Specialist → Action
  VAT        VAT           VAT            VAT
```

Every VAT carries the cryptographic link to the root human. Revocation at any level cascades to all descendants.

---

10. Regulatory Compliance Provenance Chains *(NEW)*

The Telecommunications Analogy

Telecoms operators retain CDRs — who called whom, when, from where — without retaining call content. As AI agents perform consequential actions across jurisdictions, an analogous requirement emerges: retain metadata about delegation chains without retaining operational data.

Provenance Architecture

Each delegation level appends a provenance record containing:

- **Jurisdictional context:** ISO 3166-1 codes, applicable data protection regime, cross-border transfer basis, sector-specific requirements
- **Infrastructure context:** Provider, region, security certifications, network path
- **Verification context:** Biometric modality, assurance level, re-verification status

The chain grows monotonically (append-only) and is structurally independent of the VAT. Integrity is protected by cryptographic hash chains.

Key Capabilities

- **Cross-jurisdiction detection:** Automatic identification of all applicable regimes
- **Infrastructure footprint extraction:** Complete infrastructure path for forensic investigation
- **Breach impact analysis:** Which agent chains traversed affected infrastructure
- **Emergency override:** Controlled relaxation for incident response with full auditability
- **Provenance compression:** Merkle-tree-style compact proofs for scale
- **Real-time streaming:** Continuous compliance monitoring, not batch audit

- **Compliance framework mapping:** Automatic generation of GDPR Art 30, SOX 302, HIPAA, telco CDR reports from single provenance source

---

11. Global Data Sovereignty Intelligence *(NEW)*

The Problem

GDPR, LGPD, PIPL, PIPA, APPI, Privacy Act, POPIA, DPDP Act, PIPEDA — each imposes different requirements. An agent chain traversing three countries must satisfy all three simultaneously, and the requirements may conflict.

Regional Regulatory Profile Engine

Machine-readable jurisdiction profiles encoding: retention periods, consent models, breach notification windows, data localisation rules, cross-border transfer mechanisms, individual rights, sector-specific requirements. Loaded automatically when a provenance chain enters a jurisdiction.

Key Capabilities

- **Cross-border data transfer provenance:** Records the legal basis (SCCs, adequacy decisions, BCRs) for every cross-border movement
- **Regulatory divergence detection:** Identifies conflicting requirements before agents act
- **Provenance localisation:** Single event produces compliant records for each applicable jurisdiction
- **Mutual provenance recognition:** Bilateral/multilateral agreements to recognise provenance chains
- **Data residency enforcement:** Protocol-level blocking of delegation to non-permitted infrastructure
- **Regulatory change propagation:** Updates to affected active chains when laws change

---

12. Collective Data Sovereignty and Indigenous Governance *(NEW)*

Beyond Individual Consent

Data protection frameworks assume individual data subjects. But significant data categories carry collective significance: cultural knowledge, community health patterns, genealogical records, language corpora. Indigenous data sovereignty frameworks — Te Mana Raraunga, CARE Principles, OCAP — recognise community governance over community data.

How VAC Supports Collective Governance

Community governance bodies register as collective trust roots with:

- **Governance structure:** Members, roles, decision-making protocol
- **Cultural governance framework:** Data categories, permitted uses, sharing restrictions
- **Decision-making model:** Consensus, threshold, weighted, or hybrid

Consensus Governance (Claim 210)

M=N mode for traditions where all must agree. Crucially includes **deferral logic** — consensus not reached is a governance outcome, not a failure. Respects whakawhitiwhiti kōrero and similar deliberative traditions.

Weighted Authority (Claim 211)

Quorum threshold PLUS mandatory role verification. A kaumātua's biometric verification is required regardless of general quorum. The VAT records both individuals and roles verified.

Cultural Protocol Scope Constraints (Claim 212)

Community governance rules become machine-enforceable VAT scope constraints: tapu/noa classifications, sharing restrictions, cultural sensitivity levels. These propagate through the entire agent chain with scope-narrowing rules — AI agents operating on community data cannot exceed the community's cultural governance framework.

---

13. Optional Location-Enhanced Governance *(NEW)*

Location verification is optional, configured per decision type by the community:

Tier	Description	Use Case
Not required	Any location valid	Diaspora-inclusive governance
Recorded	Location captured, not enforced	Administrative records
Preferred	Designated location noted, exceptions flagged	Preferred but flexible
Required	Must be at designated site, protocol-enforced	Marae-based governance

Why Optional Matters

Diaspora communities, elder accessibility, pandemic conditions, practical logistics — mandatory location would exclude the people whose data sovereignty is being protected.

When Location Is Required

Governance decisions at the marae, on tribal land, at traditional council grounds — the WHERE is part of the decision's legitimacy. Multi-modal location verification provides cryptographic proof of co-presence with cultural place significance recorded in provenance.

---

14. Physical System Authority *(NEW)*

Agents as Physical System Controllers

"Agent" encompasses AI systems controlling drones, robots, autonomous vehicles, manufacturing systems, surgical systems, unmanned military systems. Every physical action traces back to a

biometrically-verified human.

Kinetic Scope Constraints

Physical parameters encoded in VATs: geofence, altitude, speed, force, payload, proximity restrictions, environmental conditions. Same scope-narrowing rules — a delegated unit can never exceed its parent's physical operating envelope.

Swarm Coordination

Multiple physical systems as a coordinated swarm: human → coordinator agent → unit agents → physical systems. Single biometric revocation grounds every unit. Provenance records both collective and individual actions.

Physical Action Provenance

What the system did, where, when, with what physical parameters, under whose verified authority. Complete forensic chain from human authorisation to physical action.

---

15. Multi-National Coalition Governance *(NEW)*

Coalition Authority Structures

NATO, AUKUS, Five Eyes, UN coalitions — multiple sovereign nations delegating authority to combined command while retaining national sovereignty. VAC maps the entire structure.

Key Capabilities

- **Coalition-level decision governance:** Consensus, qualified majority, lead-nation, framework-nation
- **Cross-national agent interoperability:** Mutual authority verification with classification handling
- **Rules of engagement as scope constraints:** ROE cryptographically enforced, not just policy
- **National caveats as protocol constraints:** Cannot be overridden by coalition authority
- **Coalition trust graph:** Varying trust levels by alliance tier (Five Eyes > NATO > partner)
- **Coalition kinetic operations:** Multi-national coordinated autonomous physical systems with complete authority chains from coalition command through national chains to individual operators

---

16. Continuous Collective Validity Monitoring *(NEW)*

Collective authority is not a one-time issuance — it is a continuously validated state. The VAC Protocol treats every collective governance arrangement as a living system that must be monitored for ongoing validity.

Authority Change Classification

Severity	Trigger	Response
Minor	Trust score fluctuation within threshold	Logging only
Moderate	Role change, single participant re-verification failure	Quorum re-evaluation
Critical	Root authority revocation, security incident	Immediate collective suspension

Cascade Logic

When any participant's authority changes — biometric re-verification fails, role changes, trust score drops — the system applies governance-model-aware cascade logic. If the governance model requires a kaumātua and the kaumātua's authority lapses, the entire collective authority is suspended regardless of quorum status.

Forensic Gap Provenance

Operations interrupted mid-execution by authority changes generate forensic gap provenance records documenting what was in progress, at what point authority became invalid, and what remediation is required.

Authority Reconstitution

After a cascade event, the system supports structured authority reconstitution — the collective can re-verify and re-establish authority without starting the entire governance process from scratch, provided the reconstitution meets the original governance model's requirements.

---

17. Intelligent Agent Orchestration and Collaborative Onboarding *(NEW)*

The preceding sections describe what the system enforces. This section describes how the system intelligently configures itself.

Problem-to-Agent-Team Recommendation

A user describes their problem in natural language. The orchestration layer translates this into a recommended agent team with automatically configured authority scopes and hierarchical delegation structures:

```
User: "I need help running monthly business reviews and managing board
       communications."

System recommends:
    MBR Prep Agent      (Scope: READ financial, READ project status)
    Board Comms Agent   (Scope: READ comms, DRAFT emails, SEND requires approval)
```

```
  Action Tracker      (Scope: READ from both agents, NO modify)

VAC Authority Layer:
- MBR Agent cannot send emails (scope-limited)
- Comms Agent cannot access raw financials (scope-limited)
- Action Tracker is read-only across both (scope-limited)
- All enforced cryptographically via VAT, not policy
```

Dual-Purpose Biometric Capture

A single video recording simultaneously serves as functional product interaction (e.g., practice session, identity verification) and biometric anchoring. Security is embedded within the user experience rather than added as a separate step.

Vouch-as-Collaboration-Invitation

The identity vouching step simultaneously establishes identity verification and bilateral governance. When a person vouches for someone's identity, they are also prompted: "Will you be working with this person?" If yes, the vouch step configures bilateral governance in a single action — it doesn't feel like security, it feels like inviting a teammate.

Collaboration Discovery

When independently verified users would benefit from collaborative access, the orchestration layer detects the opportunity and automatically recommends a governance model, configures the collective authority structure, and establishes scope boundaries for cross-user agent interactions — all requiring biometric re-verification from all participants before activation.

---

18. Conformance Testing Framework

Core Metrics

Metric	Full Name	What It Tests	Target
VATV	VAT Verification Time	Chain verification at depth N	Sub-linear scaling
RPL	Revocation Propagation Latency	Time to last-agent suspension	<1s root, <5s chain
SNER	Scope Narrowing Enforcement Rate	Derived tokens never exceed parent	100%
TSPA	Trust Score Propagation Accuracy	Trust computation through chains	<0.1% error
AMLAS	Attribution Maturity Level Assessment Score	Standardised maturity scoring	Comparable output

Test Categories

1. Token structural conformance

2. Derivation rule conformance
3. Revocation cascade conformance
4. Multi-party authorisation conformance
5. Organisational hierarchy conformance
6. Cross-platform interoperability conformance

---

19. Extended Testing: Collective, Physical, Coalition *(NEW)*

Collective Governance Testing

- Consensus: M=N enforcement, deferral recording, bypass prevention
- Weighted authority: Mandatory role verification + quorum
- Cultural protocols: Scope constraint propagation, no widening
- Optional location: Four-tier configuration enforcement

Physical System Testing

- Kinetic scope constraints: Geofence, altitude, speed, force enforcement
- Swarm coordination: Revocation cascade to all units
- Physical action provenance: Complete recording with parameters
- Emergency override: Halt/return-to-base within specified latency

Coalition Operations Testing

- Multi-national authority: Governance model enforcement
- National caveats: Propagation and non-override verification
- Cross-national interoperability: Mutual authority verification
- Trust graph boundaries: Alliance tier enforcement

---

20. Graph-Based Trust Verification *(NEW)*

The trust graph represents the entire authority structure — humans, organisations, governance bodies, coalitions, agents, physical systems — as a verifiable graph.

Testing traverses every path and confirms: - Trust properties maintained at every edge - Scope constraints correctly narrowed at every delegation - Authority rules enforced at every node - Revocation cascades along correct paths

Graph-based testing generates test cases automatically from the authority structure, including boundary conditions and adversarial paths. It produces a conformance map showing pass/fail at each node and edge.

---

21. Attribution Maturity Model

Level	Name	Description
1	None	No human attribution in agent operations
2	Declared	Human identity claimed but not verified
3	Verified	Biometric verification of authorising human
4	Propagated	Verified authority propagates through agent chains
5	Proven	Complete provenance with regulatory compliance, physical system coverage, and coalition governance

---

22. Regulatory Landscape

The VAC Protocol maps to 100+ standards bodies across 12 domains. Key intersections:

- **NIST:** SP 800-63 (identity assurance), SP 800-207 (zero trust), CAISI (AI agent security)
- **ISO:** 30107-3 (biometric anti-spoofing), 27001 (information security), 24745 (biometric template protection)
- **Defence:** DoD Directive 3000.09 (autonomy in weapons), STANAG 4586 (UAV interoperability)
- **Indigenous governance:** Te Mana Raraunga, CARE Principles, OCAP, AIATSIS

---

23. Implementation Roadmap

Phase	Focus	Milestone
1	Core protocol + reference implementation	VAT creation, derivation, verification
2	Agent framework integration	OpenClaw, LangGraph, AutoGen plugins
3	Provenance + regulatory compliance	CDR-style audit trails, compliance mapping
4	Physical systems + coalition	Drone/robot authority, cross-national interoperability
5	NCCoE lab demonstration	NIST-validated conformance testing

---

24. Conclusion

AI agent security is not just about securing agents. It is about securing the link between agents and the humans who direct them — whether those agents operate in software, control physical systems, serve Indigenous community governance, or execute multi-national coalition operations.

The VAC Protocol provides this link through biometric human verification as a continuous, propagatable, non-repudiable security control with 241 claims of independently defensible innovation across 38 sections.

Protocol specification: vacprotocol.org **Patent portfolio:** AU 2026901425 + 3 supplementaries (241 claims, 38 sections) **Contact:** admin@violetshores.com

---