Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 24 November 2026

T. Sato
MyAuberge K.K.

24 May 2026

The Governance Audit Record (GAR) for Agentic AI Systems
draft-sato-soos-gar-01

Abstract

This document specifies the Governance Audit Record (GAR), the audit architecture for agentic AI systems. GAR defines five audit types, the Session Audit Record (SAR), the Audit Alert system, auditor principal categories, and the Audit Package for external regulatory inspection. GAR provides verifiable evidence that AI agent sessions were governed in accordance with the Intent Declaration Primitive [I-D.sato-soos-idp] and the Human Escalation Mechanism [I-D.sato-soos-hem]. GAR answers the governance question: can any of this be proven to a regulator?

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 24 November 2026.

Table of Contents

1. Introduction
2. Conventions and Definitions
3. Architecture Overview
4. Audit Types
   4.1. Type 1 -- GEC Self-Audit
   4.2. Type 2 -- Session-Close Audit
   4.3. Type 3 -- Event-Triggered Alert
   4.4. Type 4 -- Scheduled Audit
   4.5. Type 5 -- On-Demand External Audit
5. Auditor Principal Categories
   5.1. HEM Principal
   5.2. Audit Principal
   5.3. Verified External Auditor

5.4. GEC Self-Auditor
6. Session Audit Record
6.1. SAR Generation
6.2. SAR Schema
6.3. SAR Signing
6.4. SAR Retention
7. Audit Alert System
7.1. Alert Generation
7.2. Alert Schema
7.3. Normative Trigger List
7.4. Alert Delivery
8. Event Log Requirements
8.1. IDP Audit Events
8.2. HEM Audit Events
8.3. GAR Audit Events
8.4. CAP Audit Events
9. Audit Package
9.1. Package Composition
9.2. Package Schema
9.3. Access Control
10. SCITT Integration
10.1. SAR as SCITT Signed Statement
10.2. Audit Package SCRAPI Submission
10.3. Conformance Level Requirements
11. EU AI Act Applicability
11.1. Article 12 Mapping
12. Security Considerations
13. IANA Considerations
13.1. GAR Audit Alert Triggers Registry
13.2. GAR Auditor Principal Types Registry
14. References
14.1. Normative References
14.2. Informative References
Author's Address

1. Introduction

Agentic AI systems require governance across four questions:

- What did the agent intend before acting?
  [I-D.sato-soos-idp] -- The Intent Declaration Primitive (IDP) for Agentic AI Systems

- Who governed the agent's decisions?
  [I-D.sato-soos-hem] -- The Human Escalation Mechanism (HEM) for Agentic AI Systems

- Were those decisions within the law?
  [I-D.sato-soos-cap] -- The Constitutional AI Protocol (CAP) for Agentic AI Systems

- Can any of this be proven to a regulator?
  This document -- The Governance Audit Record (GAR) for Agentic AI Systems

GAR is the evidentiary layer of this protocol family. IDP, HEM, and CAP generate governance events; GAR specifies how those events are collected, synthesized, signed, and made available for audit.

The architectural property GAR enforces is non-suppressibility: the Governing Enforcement Component (GEC) MUST generate audit artifacts automatically, MUST sign them, and MUST NOT allow any agent, application, or principal to suppress, modify, or delete them. This property -- the GEC cannot suppress bad news from its principals -- is the foundation of accountable AI governance.

GAR defines five audit types ranging from continuous GEC self-audit (Type 1) to on-demand external regulatory inspection (Type 5). The Session Audit Record (SAR) is the primary audit artifact: a complete, GEC-signed record of every governance event in a session, generated automatically at session close.

The SAR is a candidate SCITT Signed Statement [I-D.ietf-scitt-architecture]. Section 10 specifies the SCITT integration: how SARs are submitted to a SCITT transparency log and how Audit Packages are submitted via SCRAPI. At Level 3 GEC conformance, SCITT submission is REQUIRED.

This specification is a companion to [I-D.sato-soos-idp], [I-D.sato-soos-hem], [I-D.sato-soos-cap], [I-D.sato-soos-sov], and [I-D.sato-soos-mjwt]. Readers should be familiar with those documents before reading this document.

Changes from draft-sato-soos-gar-00:

- Throughout: "governing kernel" and "kernel" renamed to "Governing Enforcement Component (GEC)" and "GEC". The JSON field name kernel_signature is preserved across all artifact types for wire-format compatibility with -00 implementations. The label field within kernel_signature MUST indicate the GEC conformance level (L1, L2, or L3) per [I-D.sato-soos-idp] Section 9.
- Section 1: SCITT integration paragraph added. Reference to [I-D.sato-soos-sov] and [I-D.sato-soos-mjwt] added.
- Section 2: GEC definition added. GEC-signed definition added. Sovereign Object definition added.
- Section 3: Architecture diagram updated to reflect GEC rename.
- Section 5.4: "Kernel Self-Auditor" renamed to "GEC Self-Auditor".
- Section 6.2: so_id field added to SAR schema. mandate_id field clarified to reference [I-D.sato-soos-mjwt] jti claim.
- Section 6.1: GEC signing key reference updated for conformance level model.
- Section 10: SCITT Integration added (new section). Specifies SAR as SCITT Signed Statement, SCRAPI Audit Package submission, and per-conformance-level requirements.
- Section 11 (was 10): EU AI Act section renumbered.
- Section 13 (was 12): References updated. IDP updated to -03. HEM updated to -01. CAP promoted from informative to normative. SOV-00, MJWT-00, and SCITT architecture draft added.

2. Conventions and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

The following terms are defined in this document or inherited from companion specifications:

Audit Principal:

A registered principal with read-only access to governance audit artifacts. Distinct from a HEM Principal. Receives Audit Alerts and reviews Session Audit Records.

Governing Enforcement Component (GEC):
As defined in [I-D.sato-soos-idp]: a runtime component that enforces authorization policy, records agent actions to a tamper-evident Event Log, and mediates agent access to governed objects. The GEC may be implemented as an application-layer library (Level 1), an isolated process or sidecar (Level 2), or an attested hardware execution environment (Level 3). See [I-D.sato-soos-idp] Section 9 for conformance level definitions.

GEC-signed:
A record signed by the Governing Enforcement Component using the signing key appropriate to its conformance level. The JSON field name kernel_signature is preserved for wire-format compatibility. The label field within kernel_signature MUST indicate the GEC's conformance level (L1, L2, or L3).

Governance Audit Record (GAR):
The audit architecture specified in this document, comprising five audit types, the SAR, the Audit Alert system, and the Audit Package.

GEC Self-Auditor:
An architectural property of the GEC, not a human role. The GEC evaluates its own Event Log after every commitment and generates KERNEL_AUDIT_ANOMALY entries when inconsistencies are detected.

IDP Commitment Gap:
A condition detected by the GEC when an agent's actual state transition does not match the agent's declared IDP commitment. Classified as a critical audit finding.

IDP Commitment Verification Record:
A GEC-generated record produced after every governed state transition, recording whether the agent's action matched its IDP commitment.

Rationale Store:
A GEC-managed object store, separate from the Event Log, holding Policy Rationale Declaration (PRD) objects and Decision Rationale Records (DRR) indexed by their respective identifiers.

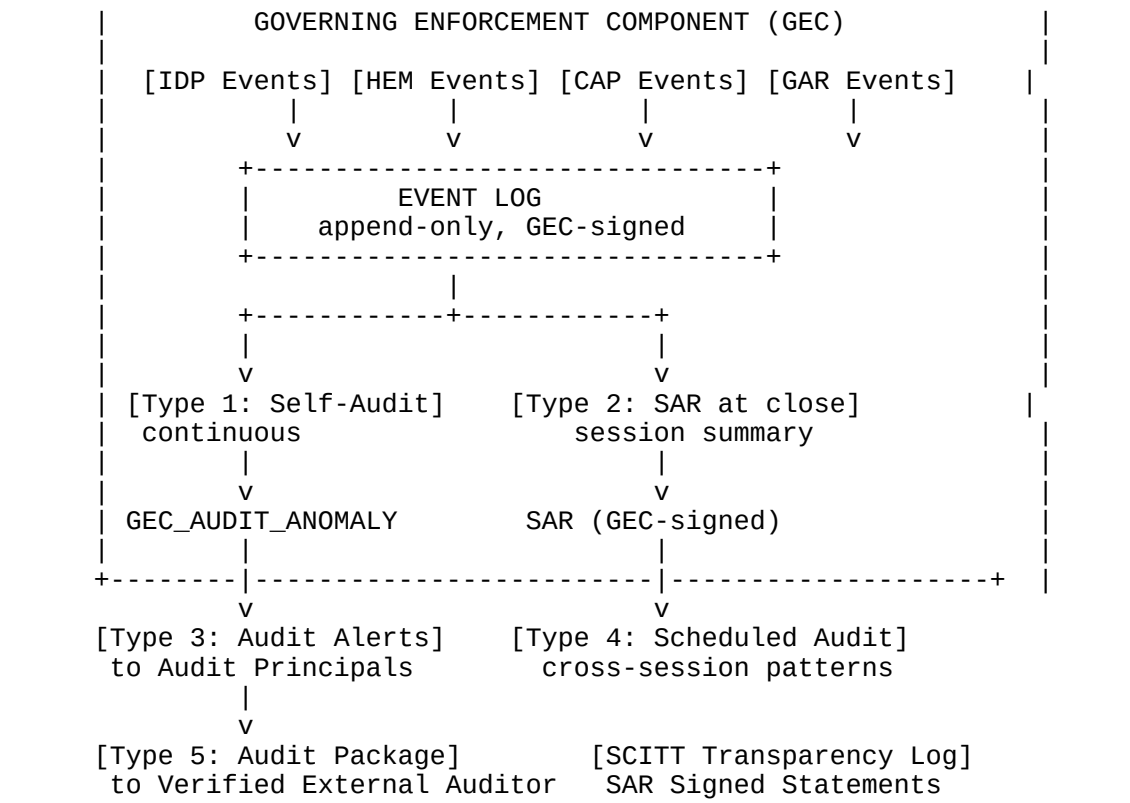
Session Audit Record (SAR):
A GEC-generated, GEC-signed summary of all governance events in a session, produced automatically at session close.

Sovereign Object (SO):
As defined in [I-D.sato-soos-sov]: a causally ordered, policy-governed, typed, living document that evolves through a predefined finite state space under GEC authority.

Verified External Auditor:
A regulator, accounting firm, or other external party granted time-limited, scope-limited read access to GEC audit artifacts by the operator. Produces an Audit Package.

3. Architecture Overview

The GAR architecture comprises five audit types operating at different timescales and with different principals:



The GEC is the sole source of audit truth. No agent, application, HEM Principal, or Audit Principal can generate, modify, or suppress GEC audit artifacts.

4. Audit Types

4.1. Type 1 -- GEC Self-Audit

The GEC MUST evaluate its own Event Log after every Event Log commitment. If the GEC detects an inconsistency -- a state transition without a corresponding IDP submission, a HEM resolution without a recorded trigger, a mandate referenced by an IDP that does not exist in the mandate store -- the GEC MUST generate a KERNEL_AUDIT_ANOMALY Event Log entry.

KERNEL_AUDIT_ANOMALY entries are immutable once written. The GEC MUST NOT suppress KERNEL_AUDIT_ANOMALY entries. A KERNEL_AUDIT_ANOMALY entry MUST immediately trigger a Type 3 Audit Alert at CRITICAL severity (Section 7.3).

The GEC MUST also generate an IDP Commitment Verification Record after every governed state transition (Section 8.1). An IDP_COMMITMENT_GAP result MUST be treated as a critical audit finding equivalent to KERNEL_AUDIT_ANOMALY for alert severity purposes.

4.2. Type 2 -- Session-Close Audit

The GEC MUST generate a Session Audit Record (SAR) automatically at the close of every governed session. SAR generation is not requestable by any external party -- it fires unconditionally on session close. The SAR specification is in Section 6.

4.3. Type 3 -- Event-Triggered Alert

The GEC MUST generate an Audit Alert when a normative trigger condition is detected. Audit Alerts are delivered to all registered Audit Principals for the governed session. The normative trigger

list is in Section 7.3.

4.4. Type 4 -- Scheduled Audit

Audit Principals MAY initiate cross-session pattern audits covering a specified time range or SO Type population. The GEC MUST expose a GEC Query Interface for this purpose [I-D.sato-soos-idp]. Type 4 audits produce cross-session pattern reports and MUST be recorded as SCHEDULED_AUDIT_INITIATED and SCHEDULED_AUDIT_COMPLETED Event Log entries.

The GEC SHOULD initiate a Type 4 audit automatically when a PRD review_date is exceeded, covering all sessions governed by the overdue policy.

4.5. Type 5 -- On-Demand External Audit

Operators MAY grant Verified External Auditors time-limited, scope-limited read access to GEC audit artifacts. Access grants MUST be recorded as EXTERNAL_AUDIT_ACCESS_GRANTED Event Log entries. Access revocation MUST be recorded as EXTERNAL_AUDIT_ACCESS_REVOKED. Audit Packages produced by Verified External Auditors are specified in Section 9. At Level 3 conformance, Audit Packages SHOULD be submitted to a SCITT transparency log via SCRAPI (Section 10.2).

5. Auditor Principal Categories

GAR defines four distinct auditor categories. These are not interchangeable.

5.1. HEM Principal

A HEM Principal is registered in a designation chain and resolves HEM escalations. A HEM Principal is NOT an auditor. HEM Principals do not receive Audit Alerts and do not have access to the Rationale Store or Event Log beyond what is included in the HEM Escalation Request.

5.2. Audit Principal

An Audit Principal is a registered principal with principal_type: AUDIT. Audit Principals receive Audit Alerts, review Session Audit Records, and may initiate Type 4 scheduled audits.

An Audit Principal MUST NOT appear in a HEM designation chain. The GEC MUST reject SO Type configurations that place an Audit Principal in a designation chain.

Audit Principals have read-only access to:

- The Event Log (via GEC Query Interface [I-D.sato-soos-idp])
- The Rationale Store
- Session Audit Records
- IDP Commitment Verification Records

Audit Principals MUST NOT be able to modify any GEC artifact.

5.3. Verified External Auditor

A Verified External Auditor is a regulator, accounting firm, or other external party granted temporary read access by the operator. Access is time-limited and scope-limited. The operator declares the access scope (session range, SO Type filter, time window) and expiry at grant time.

A Verified External Auditor produces an Audit Package (Section 9)

covering the declared scope. The Audit Package is GEC-signed as of the production timestamp.

5.4. GEC Self-Auditor

The GEC Self-Auditor is an architectural property, not a human role. It refers to the Type 1 continuous self-audit function executed by the GEC after every Event Log commitment. It cannot be disabled, configured, or bypassed.

6. Session Audit Record

6.1. SAR Generation

The GEC MUST generate a SAR automatically at the close of every governed session regardless of close reason (normal completion, TERMINATE decision, mandate expiry, session timeout, or error).

SAR generation MUST be atomic with session close. The GEC MUST NOT return a session close confirmation to any external party before the SAR is committed to the audit store.

The GEC MUST sign every SAR using Ed25519 with the GEC's signing key. The signing key MUST be consistent with the GEC's conformance level: at Level 1, an application-managed key; at Level 2, a key held by the isolated GEC process; at Level 3, a key bound to a RATS-attested execution environment [I-D.sato-soos-idp] Section 9. The GEC signing key is published via the operator's JWKS endpoint.

6.2. SAR Schema

A SAR MUST contain the following fields. All fields are REQUIRED unless stated otherwise.

sar_id:
: GEC-generated UUID v7 [RFC9562]. Unique identifier for this SAR.

session_id:
: The session identifier. Links the SAR to all Event Log entries for this session.

so_id:
: The Sovereign Object instance identifier [I-D.sato-soos-sov] Section 4.2.1. Links the SAR to the specific SO Instance governed during this session.

mandate_id:
: The governing mandate identifier. The jti claim of the Mandate JWT [I-D.sato-soos-mjwt] in force at session open.

mission_ref:
: The MissionDeclaration reference. Null if no mission was declared for this session.

open_timestamp:
: ISO 8601 UTC timestamp of session open.

close_timestamp:
: ISO 8601 UTC timestamp of session close.

close_reason:
: Controlled vocabulary. One of: NORMAL_COMPLETION | TERMINATE_DECISION | MANDATE_EXPIRY | SESSION_TIMEOUT | ERROR | CAP_SUSPENSION.

idp_submissions:
Array of IDP summary records. Each entry contains:

idp_id:	IDP identifier.
goal_summary:	Human-readable goal description.
cedar_outcome:	PERMIT | DENY | HEM_ROUTED.
hem_triggered:	Boolean.
hem_decision:	Decision type if HEM was triggered, null otherwise.

hem_events:
Array of HEM event summary records. Each entry contains:

hem_id:	HEM event identifier.
trigger_class:	Classes 1-5.
trigger_source:	AGENT_DETECTED | TRAVELER_REQUEST | SYSTEM_EVENT.
policy_rationale_id:	PRD identifier, null if absent.
decision_type:	Final decision type.
decision_rationale_class:	DRR rationale class, null if absent.
resolution_time_seconds:	Integer. Wall time from trigger to resolution.

state_transitions:
Array of state transition records. Each entry contains:

from_state:	Prior Sovereign Object state.
to_state:	Resulting Sovereign Object state.
action:	Cedar action string.
timestamp:	ISO 8601 UTC.

cap_violations:
Array of CAP violation records. Each entry contains:

violation_id:	CAP Violation Record identifier.
tier:	0 | 1 | 2.
prohibition_id:	Prohibition identifier.
action:	Action attempted.
outcome:	REFUSED | SESSION_SUSPENDED | HEM_FIRED.

audit_summary:
Summary counts block. Contains:

total_transitions:	Integer.
hem_events_count:	Integer.
terminate_count:	Integer.
auto_approve_count:	Integer.
policy_rationale_gaps:	Integer. HEM events with no PRD.
decision_rationale_gaps:	Integer. HEM events where DRR was required but absent.
cap_violation_count:	Integer.
jurisdictional_conflicts:	Integer.

kernel_signature:
Ed25519 signature over the canonical serialization of all SAR fields except kernel_signature itself. The label field within this signature MUST indicate the GEC's conformance level (L1, L2, or L3). The field name kernel_signature is preserved for wire-format compatibility.

The idp_submissions, hem_events, state_transitions, and cap_violations arrays carry reference fields and key summary data only. Full detail for each record is available in the Event Log and Rationale Store. The SAR is a governance summary and index, not a duplicate of the Event Log.

6.3. SAR Signing

The GEC MUST sign the SAR using Ed25519 prior to committing it to the audit store. The canonical serialization for signing is the JSON serialization of all fields except kernel_signature, with keys in

lexicographic order and no whitespace.

Audit Principals and Verified External Auditors MUST verify the kernel_signature before relying on SAR content.

6.4. SAR Retention

Operators SHOULD retain Session Audit Records for a minimum of 12 months from session close_timestamp. Operators subject to EU AI Act Article 12 obligations MUST retain SARs for the period required by applicable law. The GEC SHOULD warn Audit Principals when a SAR approaches its configured retention expiry.

At Level 3 conformance, SARs MUST additionally be submitted to a SCITT transparency log per Section 10. SCITT submission provides independent tamper-evidence that complements the GEC's internal non-suppressibility guarantee.

7. Audit Alert System

7.1. Alert Generation

The GEC MUST generate an Audit Alert when any normative trigger condition listed in Section 7.3 is detected. Alert generation is synchronous with the triggering event -- the GEC MUST generate the alert before returning any response to the triggering agent or principal.

7.2. Alert Schema

An Audit Alert MUST contain the following fields:

alert_id:
: GEC-generated UUID v7.

alert_severity:
: CRITICAL | HIGH | MEDIUM | LOW.

alert_trigger:
: Identifier of the normative trigger condition. See Section 7.3.

session_id:
: The session in which the trigger occurred.

so_id:
: The Sovereign Object instance identifier for the session in which the trigger occurred [I-D.sato-soos-sov].

hem_id:
: The HEM event identifier, if the trigger is HEM-related. Null otherwise.

cap_violation_id:
: The CAP Violation Record identifier, if the trigger is CAP-related. Null otherwise.

detail:
: Human-readable description of the trigger condition. REQUIRED.

timestamp:
: ISO 8601 UTC timestamp of alert generation.

kernel_signature:
: Ed25519 signature over canonical serialization of all fields except kernel_signature.

delivered_to:
: Array of Audit Principal identifiers to whom the alert was delivered.

7.3. Normative Trigger List

The following trigger conditions MUST generate an Audit Alert. Trigger identifiers are registered in the GAR Audit Alert Triggers registry (Section 13.1).

Trigger	Severity
KERNEL_AUDIT_ANOMALY	CRITICAL
IDP_COMMITMENT_GAP	CRITICAL
TERMINATE_DECISION	HIGH
AUTO_APPROVE_DISPOSITION	HIGH
HEM_CHAIN_EXHAUSTED	HIGH
MISSION_REVOKE_CASCADE	HIGH
MANDATE_NARROWING_VIOLATION	HIGH
HEM_TERMINATE_RATIONALE_REQUIRED	MEDIUM
THREE_OR_MORE_HEM_EVENTS_IN_SESSION	MEDIUM
PRD_REVIEW_DATE_EXCEEDED	MEDIUM
POLICY_RATIONALE_GAPS_IN_SAR	LOW

Table 1: Normative Audit Alert Triggers

MANDATE_NARROWING_VIOLATION is added in this revision. It is triggered when the GEC detects that a presented Child Mandate violates the Narrowing Property as defined in [I-D.sato-soos-mjwt] Section 5. This is a HIGH severity finding because it indicates an attempted authorization escalation.

7.4. Alert Delivery

Audit Alerts MUST be delivered to all registered Audit Principals for the governed session. Delivery MUST be recorded as an AUDIT_ALERT_FIRED Event Log entry, followed by AUDIT_ALERT_DELIVERED on successful delivery.

Implementations SHOULD use the Shared Signals Framework (SSF) [RFC8936] for cross-system Audit Alert delivery.

Audit Principals SHOULD acknowledge Audit Alerts. Acknowledgement MUST be recorded as AUDIT_ALERT_ACKNOWLEDGED.

8. Event Log Requirements

The Event Log is the append-only, GEC-maintained record of all governance events in a session. This section specifies the GAR-specific Event Log entries that MUST be supported.

8.1. IDP Audit Events

IDP_SUBMITTED:
: Recorded when an IDP is submitted to the GEC. Entry type specified in [I-D.sato-soos-idp].

IDP_COMMITMENT_VERIFIED:
: Recorded after every governed state transition. The GEC MUST generate an IDP Commitment Verification Record and commit this event. Fields: idp_id, state_transition_id, verified_at, match_result (MATCHED | IDP_COMMITMENT_GAP), kernel_signature.

IDP_COMMITMENT_GAP:
Recorded when match_result is IDP_COMMITMENT_GAP. This is a critical audit finding. The GEC MUST immediately:
(a) generate a CRITICAL Audit Alert (alert_trigger: IDP_COMMITMENT_GAP), and
(b) fire HEM_AGENT_ESCALATED (Class 2) for the active session.
The GEC MUST NOT allow a session to continue after an IDP_COMMITMENT_GAP without HEM resolution.

8.2. HEM Audit Events

The following HEM Event Log entries gain new fields under GAR:

HEM_TRIGGERED:
Existing entry type. GAR adds: policy_rationale_id (REQUIRED, null if PRD absent -- absence recorded in audit_summary. policy_rationale_gaps).

HEM_DECISION_RECEIVED:
Existing entry type. GAR adds: decision_rationale_class (REQUIRED when DRR is mandatory for the decision type; OPTIONAL otherwise).

The following new HEM Event Log entries are specified in [I-D.sato-soos-hem] and recorded in the GAR Event Log:

HEM_DECISION_NOT_PERMITTED_FOR_TRIGGER_CLASS
HEM_TERMINATE_RATIONALE_REQUIRED
HEM_HUMAN_DECISION_CONSTITUTIONAL_VIOLATION
HEM_CHAIN_CONSTITUTIONAL_EXHAUSTED
KERNEL_AUDIT_ANOMALY

8.3. GAR Audit Events

The following Event Log entry types are introduced by this document:

SAR_GENERATED:
Recorded when a SAR is committed to the audit store. Fields: sar_id, session_id, so_id, close_reason, kernel_signature.

SAR_SCITT_SUBMITTED:
Recorded when a SAR is submitted to a SCITT transparency log. Fields: sar_id, scitt_entry_id, transparency_log_uri, submitted_at, kernel_signature. See Section 10.1.

AUDIT_ALERT_FIRED:
Recorded when an Audit Alert is generated. Fields: alert_id, alert_trigger, alert_severity, session_id, so_id.

AUDIT_ALERT_DELIVERED:
Recorded when an Audit Alert is successfully delivered to an Audit Principal. Fields: alert_id, principal_id, delivered_at.

AUDIT_ALERT_ACKNOWLEDGED:
Recorded when an Audit Principal acknowledges an Audit Alert. Fields: alert_id, principal_id, acknowledged_at.

SCHEDULED_AUDIT_INITIATED:
Recorded when a Type 4 scheduled audit begins. Fields: audit_id, initiated_by, scope_description, initiated_at.

SCHEDULED_AUDIT_COMPLETED:
Recorded when a Type 4 scheduled audit completes. Fields: audit_id, completed_at, findings_count.

EXTERNAL_AUDIT_ACCESS_GRANTED:
Recorded when a Verified External Auditor is granted access. Fields: auditor_id, granted_by, scope, expiry, granted_at.

AUDIT_PACKAGE_PRODUCED:
Recorded when a Verified External Auditor produces an Audit Package. Fields: package_id, auditor_id, scope, produced_at, package_hash.

EXTERNAL_AUDIT_ACCESS_REVOKED:
Recorded when Verified External Auditor access expires or is revoked. Fields: auditor_id, revoked_at, revocation_reason.

PRD_REVIEW_DATE_EXCEEDED:
Recorded by the GEC's continuous self-audit when a PRD review_date is exceeded. Fields: prd_id, policy_id, review_date, detected_at. This entry MUST trigger a MEDIUM Audit Alert (alert_trigger: PRD_REVIEW_DATE_EXCEEDED).

8.4. CAP Audit Events

The following CAP Event Log entries are specified in [I-D.sato-soos-cap] and recorded in the GAR Event Log:

CAP_VIOLATION_DETECTED:
AI-initiated action refused by the Constitutional Evaluation Engine. Fields: violation_id, tier, prohibition_id, action, outcome, timestamp, kernel_signature.

CAP_HUMAN_VIOLATION_DETECTED:
Human principal decision refused by the Constitutional Evaluation Engine. Fields: violation_id, tier, prohibition_id, decision, outcome, timestamp, kernel_signature.

CAP_TIER1_CONFLICT_DETECTED:
Jurisdictional conflict detected at Tier 1. Fields: conflict_id, conflicting_jurisdictions, resolution_method, hem_id, timestamp.

APPROVE_WITH_LEGAL_BASIS_RECORDED:
Principal submitted APPROVE_WITH_LEGAL_BASIS decision. Fields: hem_id, principal_id, legal_basis (authority_type, authority_ref, jurisdiction, expiry, document_hash), timestamp.

SESSION_CAP_SUSPENDED:
Session suspended due to CAP violation. Fields: session_id, violation_id, suspended_at.

9. Audit Package

9.1. Package Composition

An Audit Package is produced by a Verified External Auditor and covers a declared scope (session range, SO Type filter, or time window). The Audit Package is a GEC-signed compilation of:

- All SARs within scope
- All Event Log entries within scope
- All PRD records from the Rationale Store for policies governing sessions within scope
- All DRR records from the Rationale Store for decisions within scope
- All Audit Alert records within scope
- All CAP Violation Records within scope

9.2. Package Schema

An Audit Package MUST contain the following fields:

package_id:
GEC-generated UUID v7.

auditor_id:
Verified External Auditor identifier.

scope:
Declaration of what the package covers. Fields: session_range, so_type_filter (optional), time_window.

sar_records:
Array of all SARs within scope.

event_log_records:
Array of all Event Log entries within scope.

prd_records:
Array of all PRD objects from the Rationale Store for policies governing sessions within scope.

drr_records:
Array of all DRR objects from the Rationale Store for decisions within scope.

audit_alert_records:
Array of all Audit Alert records within scope.

cap_violation_records:
Array of all CAP Violation Records within scope.

chain_of_custody:
Block containing:

```
package_hash:     SHA-256 hash of all package content fields.
kernel_signature: Ed25519 signature over package_hash.
produced_by:      Verified External Auditor identifier.
produced_at:      ISO 8601 UTC timestamp.
```

9.3. Access Control

The GEC MUST verify that the requesting party holds a valid, unexpired Verified External Auditor access grant before producing an Audit Package. The access grant MUST be scoped to include the requested sessions.

Audit Package production MUST be recorded as AUDIT_PACKAGE_PRODUCED in the Event Log.

10. SCITT Integration

10.1. SAR as SCITT Signed Statement

The Session Audit Record (SAR) is a SCITT Signed Statement as defined in [I-D.ietf-scitt-architecture]. It carries all properties required of a SCITT Signed Statement: it is produced by an identified Issuer (the GEC), signed with a key bound to that Issuer’s attested execution environment, and carries a payload that a Relying Party can evaluate against a known governance policy.

The SAR SCITT Signed Statement payload is the canonical JSON serialization of the SAR as defined in Section 6.2. The COSE [RFC9052] protected header MUST include:

- alg: EdDSA (Ed25519)
- kid: Key identifier of the GEC's signing key
- content_type: application/soos.gar.sar+json
- issuer: GEC identifier

Upon successful SCITT submission, the GEC MUST record a SAR_SCITT_SUBMITTED Event Log entry (Section 8.3) containing the SCITT transparency log entry identifier and the transparency log URI.

10.2. Audit Package SCRAPI Submission

The Audit Package (Section 9) maps directly onto the SCRAPI POST /entries endpoint [I-D.ietf-scitt-architecture]. A Verified External Auditor MAY submit an Audit Package to a SCITT transparency log via SCRAPI. Once registered, the append-only guarantee of the SCITT transparency log ensures that the Audit Package cannot be altered or removed independently of what the operator or GEC does.

The SCRAPI submission provides the external tamper-evidence property that complements GAR's internal non-suppressibility guarantee.

SCRAPI Audit Package submission MUST be recorded as AUDIT_PACKAGE_PRODUCED in the Event Log with a scitt_entry_id field if submission is performed.

10.3. Conformance Level Requirements

SCITT submission requirements vary by GEC conformance level, as defined in [I-D.sato-soos-idp] Section 9:

Level 1 (Application Profile):
SCITT SAR submission is RECOMMENDED. Non-suppressibility is probabilistic; SCITT submission is the primary compensating control. Operators SHOULD configure automatic SAR submission to a SCITT transparency log.

Level 2 (Isolated Profile):
SCITT SAR submission is RECOMMENDED. The isolated GEC process provides architectural non-suppressibility; SCITT provides independent external evidence.

Level 3 (Kernel Profile):
SCITT SAR submission is REQUIRED. Every SAR MUST be submitted to a SCITT transparency log before the GEC returns a session close confirmation. The SAR_SCITT_SUBMITTED Event Log entry MUST precede or be atomic with SAR_GENERATED.

11. EU AI Act Applicability

11.1. Article 12 Mapping

EU AI Act Article 12 requires high-risk AI systems to automatically generate logs enabling post-market monitoring and audit. The following table maps Article 12 provisions to GAR mechanisms. This mapping is normative: the Event Log fields and SAR structure specified in this document satisfy Article 12(3) traceability requirements for deployments governed by [I-D.sato-soos-hem]. Operators may reference this section directly in conformance documentation.

Article 12 Provision	GAR Mechanism	Sec.
12(1) Automatic logging capability	Event Log: append-only, GEC-generated, cannot be	8

	suppressed	
12(2) Logging period commensurate with purpose	SAR close_timestamp + operator retention configuration; SHOULD minimum 12 months	6.4
12(3) Traceability of AI system operation	hem_id chain across Event Log entries -- full causal history reconstructible from any event	8
12(3) Human oversight audit record	principal_type + principal_id + decision_type + DRR on every HEM_DECISION_RECEIVED entry	8.2
12(3) Policy audit record	PRD + prd_id on every HEM_TRIGGERED entry	8.2

Table 2: EU AI Act Article 12 Mapping

12. Security Considerations

The GAR audit architecture relies on the following security properties:

GEC signing key integrity:
: All SAR, Audit Alert, IDP Commitment Verification Record, and Audit Package chain-of-custody signatures depend on the integrity of the GEC's Ed25519 signing key. At Level 3, the key MUST be bound to a RATS-attested execution environment. At Level 2, the key MUST be held in the isolated GEC process, inaccessible to agent code. At Level 1, key protection is application-managed; HSM controls are RECOMMENDED. Key compromise MUST be treated as a critical security incident requiring immediate rotation and re-signing of all affected audit artifacts.

Event Log append-only property:
: The Event Log MUST be implemented as an append-only data structure. No API MUST allow deletion or modification of existing entries. Audit Principals and Verified External Auditors MUST have read-only access.

Non-suppressibility:
: The GEC MUST NOT expose any interface that allows an agent, application, HEM Principal, or Audit Principal to suppress SAR generation, Audit Alert firing, or IDP Commitment Verification. Implementations MUST be reviewed for any code path that could conditionally skip these operations.

Audit Principal separation:
: Audit Principals MUST be registered separately from HEM Principals. The same party SHOULD NOT hold both roles for the same SO Type. Separation prevents a principal from suppressing audit findings about their own HEM decisions.

Verified External Auditor access:
: GEC interfaces for Verified External Auditor access MUST enforce scope limitations at the query layer. Access grants MUST expire automatically. The GEC MUST reject queries outside the declared scope.

PRD review_date enforcement:
: Operators MUST ensure that PRD review_date values reflect genuine governance review cycles. Stale PRDs with extended review_dates undermine the living governance record property that PRD is

designed to provide.

SCITT submission integrity:
At Level 3, SAR submission to a SCITT transparency log is REQUIRED. Implementations MUST verify that the SCITT transparency log returns a valid receipt before recording SAR_SCITT_SUBMITTED. A failed SCITT submission at Level 3 MUST be treated as a critical audit finding and MUST trigger a CRITICAL Audit Alert.

13. IANA Considerations

13.1. GAR Audit Alert Triggers Registry

This document establishes the "Governance Audit Record Audit Alert Triggers" registry.

Registration procedure: Specification Required.

Initial values:

Trigger Identifier	Severity	Reference
KERNEL_AUDIT_ANOMALY	CRITICAL	Sec. 7.3
IDP_COMMITMENT_GAP	CRITICAL	Sec. 7.3
TERMINATE_DECISION	HIGH	Sec. 7.3
AUTO_APPROVE_DISPOSITION	HIGH	Sec. 7.3
HEM_CHAIN_EXHAUSTED	HIGH	Sec. 7.3
MISSION_REVOKE_CASCADE	HIGH	Sec. 7.3
MANDATE_NARROWING_VIOLATION	HIGH	Sec. 7.3
HEM_TERMINATE_RATIONALE_REQUIRED	MEDIUM	Sec. 7.3
THREE_OR_MORE_HEM_EVENTS_IN_SESSION	MEDIUM	Sec. 7.3
PRD_REVIEW_DATE_EXCEEDED	MEDIUM	Sec. 7.3
POLICY_RATIONALE_GAPS_IN_SAR	LOW	Sec. 7.3

Table 3: Initial GAR Audit Alert Triggers Registry Values

13.2. GAR Auditor Principal Types Registry

This document establishes the "Governance Audit Record Auditor Principal Types" registry.

Registration procedure: Standards Action.

Initial values:

Type	Description
HEM_PRINCIPAL	Resolves HEM escalations. NOT an auditor.
AUDIT_PRINCIPAL	Receives Audit Alerts, reviews SARs, initiates Type 4 scheduled audits. Read-only GEC access.
VERIFIED_EXTERNAL_AUDITOR	Regulator or accounting firm. Time-limited, scope-limited GEC access. Produces Audit Packages.
GEC_SELF_AUDITOR	Architectural property of the GEC. Not a human role.

Table 4: Initial GAR Auditor Principal Types Registry Values

14. References

14.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.

[RFC8936] Hunt, P., Ed., Brock, M., Backman, A., and M. Jones, "Poll-Based Security Event Token (SET) Delivery Using HTTP", RFC 8936, November 2020.

[RFC9052] Schaad, J., "CBOR Object Signing and Encryption (COSE): Structures and Process", RFC 9052, August 2022.

[RFC9562] Davis, B., Peabody, C., and P. Leach, "Universally Unique IDentifiers (UUIDs)", RFC 9562, May 2024.

[I-D.sato-soos-idp] Sato, T., "The Intent Declaration Primitive (IDP) for Agentic AI Systems", draft-sato-soos-idp-03, May 2026.

[I-D.sato-soos-hem] Sato, T., "The Human Escalation Mechanism (HEM) for Agentic AI Systems", draft-sato-soos-hem-01, May 2026.

[I-D.sato-soos-cap] Sato, T., "Constitutional AI Protocol (CAP) for Agentic AI Systems", draft-sato-soos-cap-00, May 2026.

[I-D.sato-soos-sov] Sato, T., "The Sovereign Object (SOV) for Agentic AI Systems", draft-sato-soos-sov-00, May 2026.

[I-D.sato-soos-mjwt] Sato, T., "The Mandate JWT (MJWT) for Agentic AI Systems", draft-sato-soos-mjwt-00, May 2026.

[I-D.ietf-scitt-architecture] Birkholz, H., et al., "An Architecture for Trustworthy and Transparent Digital Supply Chains", draft-ietf-scitt-architecture, work in progress.

14.2. Informative References

[EU-AI-ACT] European Parliament and Council, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence", OJ L 2024/1689, July 2024.

[I-D.sato-soos-faip] Sato, T., "Federated Agent Intelligence Protocol (FAIP)", draft-sato-soos-faip-00, forthcoming.

Author's Address

Tom Sato
MyAuberge K.K.
Chino, Nagano, Japan
Email: tomsato@myauberge.jp

URI: https://activitytravel.pro/