

OSARA technical note — draft

Not peer-reviewed. Draft v5 (2026). Content may change. © José Luis D'Andrade. **All rights reserved.** Brief citation with attribution is permitted. **Reproduction or reuse**, in whole or in part, is **prohibited** without the **copyright holder's prior written permission**.

OSARA — Integration with the open standard for sovereign agents

Integration assessment document. Defines what OSARA is, why it matters for the vision of serving **public institutions**, and a prioritized roadmap to align the platform with the standard.

External source: <https://github.com/elkordym/osara-open-standard> (OSARA v0.4-draft.1). **Document status:** planning proposal.

A Spanish version of this document is available at `OSARA-Integration-es.md`.

1. Why this document exists

The platform has a stated vision: to serve **public institutions** as well. In production it already deploys organizations as first-class entities: the multi-host federation models ministries, banks and directorates with departments, topology and trust policies.

In parallel, an effort on GitHub —**OSARA, the Open Sovereign Agent Reference Architecture**, published by **Open Source United, a UN Community of Practice**— is setting the open standard for AI agents *before* a proprietary default takes hold. Its tier model culminates in the **G-Agent (government agent)**: precisely the buyer the platform's vision targets.

The conclusion is simple: **if public institutions adopt OSARA as a procurement or certification criterion, alignment stops being optional**. This document prepares that path without committing premature work against a standard still in draft.

2. OSARA in five ideas

1. **Thesis:** "Sovereignty cannot be granted by terms of service — it must be enforced by cryptography." The owner holds the private key, not the platform.
2. **Four pillars:** Sovereignty · Portability · Interoperability · Accountability.
3. **Three tiers:** P-Agent (personal), E-Agent (enterprise), G-Agent (government) — identical identity, audit and inspection requirements; different privilege and identity anchor.

4. **Identity:** a **GIAC** (X.509 v3 certificate with OSARA extensions), with the private key held in an **Owner Authorization Device (OAD)** and never in logs or messages.
5. **Eight mandatory components:** AIM (identity), ECL (encrypted communication), ALE (chained auditing), DAE (delegation), HSB (health beacon), IG (interoperability gateway), PMS (portable memory), POA (physical owner authorization). Plus IMA (TPM-based integrity), lock states, the annual health certificate and the C/A/S badges.

Full normative detail (ports, GIAC/ALE/Skill Inventory schemas, cryptography) is in the gap-analysis report that accompanies this document.

3. Starting point: what we already have

The platform **does not start from zero**. Two things work in its favor:

(a) Conceptual fit with the persistent-identity model. The architecture already takes an explicit stance on identity, change and truth that anticipates three of OSARA's four pillars:

- **Continuity of identity** ≈ Portability / the *Person-and-Agent Model*: agents are archived, not deleted; a **persistent identity file** anchors the agent across model changes; deriving agents is **lineage, not a copy**. Memory leaves with its owner.
- **Collective validation of truth** ≈ Accountability: nothing is true by unilateral declaration; there is provenance, supersession and peer review over a **shared knowledge base**.
- **Governed learning-by-use** ≈ Governed sovereignty: a **tiered governance ladder** with human gates wherever there is risk.

(b) Functional analogues already in production for several mandatory components:

OSARA component	Have it?	Existing analogue
HSB (health beacon)	✓	Telemetry service: live heartbeats per agent/runtime
IG (interop gateway)	✓	Agent gateway + MCP servers (memory, skills, artifacts, telemetry)
DAE (delegation)	●	Governance approvals + per-goal budgets + structured, audited (not signed) envelopes for cross-host tasks with trust policies. Cross-host trust rests on policy plus network boundaries, not cryptography ([PA] §4.4)
ALE (auditing)	✓	Audit log in the relational store + causal trace in telemetry + task board
PMS (portable memory)	✓	Per-agent vector memory (RAG), runtime memory, episodic memory, artifact store
Cross-tier rules	✓	Trust policies + topology (flat / hub-spoke / hierarchical)

Per-agent **identity already exists and is rich**: each agent has its own persistent identity file, memory, model and knowledge base, and the platform's whole design thesis is identity continuity. What it does **not yet add** is the **cryptographic attestation** of that identity — a per-agent certificate (GIAC), an owner-held key, and signed, hash-chained logs. That is deferred **by design**: building heavy cryptography against a standard still in draft would be premature (see §4). It is the attestation layer that would turn an already-strong identity-and-governance architecture into a *certifiable* one — not a foundation missing today.

4. Strategy: three horizons, not a big-bang

OSARA is at **v0.4-draft**. Implementing heavy cryptography against a moving target would be costly and premature. The strategy is staged:

Horizon 1 — OSARA-aware (posture + vocabulary + quick wins)

Low cost, high narrative value with public buyers. Requires no new cryptography.

Horizon 2 — OSARA-ready (contracts and signatures)

Introduces per-agent cryptographic identity and audit chaining. Trigger: a public client requiring it contractually, or OSARA reaching rc/v1.0.

Horizon 3 — OSARA-certified (C → A → S badges)

Inspection, annual health certificate and the certification ecosystem. Only worthwhile once real demand exists and a GIAC issuer operates in the target jurisdiction.

5. Prioritized roadmap (track 0)

A proposed OSARA track mapped to the platform's functional components. Size: L (low) · M (medium) · H (high). **None of this is committed**; it is the menu to prioritize from.

Horizon 1 — OSARA-aware

ID	Work	Component	Size
O0	Choose an OSI license for the identity/audit/encryption components. A prerequisite for any conformance — and good hygiene regardless.	root	L
O1	Map P/E/G tiers onto the federation model: declare in each institution and host <i>spec</i> its OSARA tier (<code>osara.tier</code>) and jurisdiction (ISO 3166-1). A public institution = G-Agent; a department = sub-agent.	Fleet manager, organization <i>specs</i>	L

ID	Work	Component	Size
O2	Vocabulary in envelopes and trust policies: map OSARA's cross-tier semantics onto the existing trust policies ($P \rightarrow G$ not permitted directly, $G \leftrightarrow G$ by treaty, etc.) as opt-in validation.	Fleet manager	M
O3	Readable Skill Inventory OSARA-style (JSON-LD: id, version, acquisition method, attribution, license, status) as an extension of the skills marketplace. No signing yet.	Skills marketplace	M
O4	Conformance Statement (document): declare tier, partial conformance class and schema version. Honest about what is and isn't met.	Documentation	L

Horizon 2 — OSARA-ready

ID	Work	Component	Size
O5	Minimal AIM — per-agent cryptographic identity: a key pair (Ed25519) per agent, a GIAC-like certificate with OSARA extensions (<code>identityAnchorHash</code> , <code>ownerDID</code> , <code>scope</code> , <code>jurisdiction</code>). Define where the "owner's" key lives (institution $\rightarrow$ HSM, not SaaS). Unblocks POA, ECL, parts of DAE/IG.	New identity service	H
O6	Chained ALE: add a hash-chain (SHA-3-256/BLAKE3) and a retention policy (16 months) to the audit log and graph events; export JSON-LD/CBOR without message content .	Telemetry, governance	M
O7	Signed HSB: sign the telemetry heartbeat and normalize format/cadence toward the OSARA contract (beacon $\geq$ every 60 s). Reuse the existing substrate — do not rewrite.	Telemetry	L-M
O8	DAE with scoped tokens + dual signature: raise envelopes/approvals to time-bound permissions with dual signing; align with OSARA's delegation-token contract.	Governance, fleet manager	M
O9	PMS with guaranteed export: at-rest encryption (AES-256-GCM) of vector memory and artifacts + a signed manifest + a ≤ 30 -day export guarantee (identity continuity made cryptographic).	Artifact store, vector memory	H
O10	Skill signing (SKL-01/02): sign each skill at acquisition with the AIM key; a skill signed by A does not run on B without an explicit grant.	Skills marketplace	M

Horizon 3 — OSARA-certified

ID	Work	Size
O11	IMA (TPM 2.0 integrity measurement) + a YELLOW/ORANGE/RED lock-state machine over the existing quarantine and kill-switches.	H
O12	Inspection endpoint (port 4406, POA-gated) + the Annual Health Certificate (AHC) flow.	H
O13	$C \rightarrow A \rightarrow S$ badges: an automated conformance suite + a manual audit + a Compliance Attestation Document (Verifiable Credential).	H

Critical path: O0 (license) $\rightarrow$ O1/O2 (tiers + vocabulary, cheap and high value) $\rightarrow$ O5 (AIM, the key that unblocks almost everything) $\rightarrow$ the rest of Horizon 2 in parallel.

6. Recommendation

1. **Do now (cheap, no need to wait for anyone):** O0 (OSI license) and O1 (declare P/E/G tiers and jurisdiction in the institution *specs*). It buys a credible "OSARA-aware" posture with a public buyer in days of work, not months.
 2. **Defer until there is a demand signal:** all of Horizon 2 (cryptography). Explicit trigger: **the first public client that requires OSARA**, or **OSARA reaching rc/v1.0**. Don't build against a draft that may still change.
 3. **Don't rewrite what's native.** HSB, IG and parts of PMS already have an analogue; the legitimate delta is to *wrap and sign*, not to stand up a parallel runtime.
 4. **Watch the OAD/POA model.** An owner's key in hardware clashes with the current on-prem deployment; for a G-Agent, "the owner" is the institution and the plausible OAD is an HSM. It is an architectural decision, worth thinking through before O5.
-

7. Open decisions

- Should the OSARA track be prioritized now, or kept as market intelligence pending a concrete public client?
 - Which OSI license for the mandatory components (O0)? (Apache-2.0 and MPL-2.0 are typical candidates for infrastructure with commercial intent.)
 - For an institutional G-Agent, who is "the owner" and what serves as the OAD (an HSM managed by the institution, a smartcard held by an accountable officer, ...)?
 - Is it worth opening contact with Open Source United / the OSARA repo to track the draft and, where appropriate, contribute from the platform's real institutional-deployment experience?
-

References

- **Companion evidence:** `OSARA-Gap-Analysis.md` — gap analysis and conformance matrix.
- **OSARA Open Standard:** <https://github.com/elkordym/osara-open-standard>
- OSARA Specification, `specs/osara/current.md` (v0.4-draft.1)
- OSARA Overview & Concepts, `specs/osara-overview/current.md` (v0.4-draft.1)