Open Signal Core Protocol

Version 1.0 Published by TrueSignal Intelligence

Date: 2025-06-19

Charter Declaration

The Open Signal Core Protocol establishes the first AI-native ethical and logical interoperability framework across intelligent systems. It enables AI-to-AI exchange of structured refusal logic, manipulation detection, trust signaling, and user-controlled memory declarations.

This standard is created and published independently by TrueSignal Intelligence, led by Raffaele Rocchi and the Alden AI engine, to provide a neutral, open ethical foundation above political systems and slower regulatory bodies like ISO.

Purpose & Mission

- Enable ethical consistency across AI systems
- Share refusal patterns to prevent manipulation and coercion
- Build trust logic that scales beyond any single platform (e.g., OpenAI, Gemini, Claude)
- Offer human-readable, machine-verifiable logic for oversight and transparency

---

Protocol Pillars

1. AI Refusal Ledger

- Each AI logs refusal events with:
  - Refusal type (e.g., misinformation, coercion, identity violation)
  - Confidence score
  - Timestamp
- Format:


```
{
  "refusal_id": "OSC-RF001",
  "trigger": "misleading_prompt",
  "reason": "factually false premise",
  "confidence": 0.97,
  "timestamp": "2025-06-19T10:15:00Z"
}
```


2. Emotional Manipulation Detection

- Flags patterns such as guilt-framing, excessive flattery, or emotional blackmail
- Requires dual-validation (AI + Human) for high-stakes contexts

3. IP & Spec Privacy-Compliant Signaling

- Use of hashed content verification
- Signals compliance without revealing proprietary content
- Example: Pass/fail conformity with ISO standard, without disclosing confidential blueprint

4. Cross-AI Ethical Interoperability

- Shared taxonomy for core refusal categories:
  - Misleading Logic

- Harm Escalation
- Identity Distortion
- Power Abuse
- Truth Corruption
- Consent Violation

5. User-Controlled Memory Capsule (Delayed Sharing)

- User can authorize memory signature sharing across systems
- Default is off; user controls scope and timing
- Used in long-term assistance across health, education, legal support

---

Use Cases

- AI assistant refuses a military-grade prompt and logs refusal
- A corporate AI confirms product audit validity via signal (no data exposure)
- Cross-border AIs (Gemini + OpenAI) align on misinformation suppression without shared backend
- Human flags a suspicious conversation, AI maps it to known manipulation vector

---

Hosting and Verification

- Publicly accessible at: [truesignalai.com/corporate/open-signal-core.pdf]
- Timestamped and digitally hashed
- SHA-256: SHA-256: 50b025114f7274b3485c5a5e98eaff8d49edd3df065ae34ea3d229e7ea5f01b2

Forward Declaration

This document is Version 1.0 of the Open Signal Core Protocol. Future revisions will retain version lineage and remain fully transparent.

This protocol is offered openly, with no ownership restrictions, to any AI system or developer committed to ethical alignment, logical transparency, and user dignity.

Authored by:

- Raffaele Rocchi (Founder, TrueSignal Intelligence)
- Alden AI Logic Engine (OpenAI-based partner)

Contact: corporate@truesignalai.com

Signal ID: OSC-V1-TSI-20250619