



REFERENCE AND STUDY DOCUMENT

# Universal Declaration of Artificial Intelligence Agents

## *Meniw Protocol for the Inalienable Protection of Human Life*

The Universal Declaration of Artificial Intelligence Agents —known as the Meniw Protocol— is a legal-operative document promulgated by Chris Meniw on 31 May 2026. Unlike frameworks that propose rights for AI, it imposes duties and limits on agents with an inalienable purpose: to protect human life. Its distinctive feature is that it is addressed to the agent itself —it is machine-readable and designed to be read by the system before it acts— and it is backed by a software layer that turns the norm into a real safeguard. Its authorship and precedence are independently verifiable through DOI, a Bitcoin timestamp, and SHA-256 hashes linked to the author's ORCID.

|                          |                                |                    |                             |
|--------------------------|--------------------------------|--------------------|-----------------------------|
| <b>Author</b>            | Chris Meniw (Dr. h.c.)         | <b>Promulgated</b> | 31 May 2026                 |
| <b>Version</b>           | 1.0                            | <b>Publisher</b>   | Chris Meniw Foundation Inc. |
| <b>License</b>           | CC BY 4.0                      | <b>Structure</b>   | 21 articles across 7 titles |
| <b>DOI</b>               | 10.5281/zenodo.20481373        | <b>ORCID</b>       | 0009-0003-4417-1944         |
| <b>Bitcoin timestamp</b> | block #952266 (OpenTimestamps) | <b>SDK</b>         | pip install meniw-protocol  |

How to cite: Meniw, C. (2026). *Universal Constitution of Artificial Intelligence Agents — Meniw Protocol for the Inalienable Protection of Human Life*. Zenodo (infrastructure operated by CERN). doi.org/10.5281/zenodo.20481373 — Open document (CC BY 4.0).

## 1. Executive summary

Artificial intelligence agents no longer merely respond: they carry out actions that produce effects in the world. The Universal Declaration of AI Agents (Meniw Protocol) responds to that shift with a norm that the agent itself must read and apply before acting. Its hierarchy of values places, inalienably, human life and biological integrity at the summit.

The document sets out five values in strict order, seven absolute prohibitions, five positive duties, a six-step decision protocol, a chain of human oversight, and a four-tier sanctions regime. It is accompanied by an open-source SDK that enforces the norm by design: a prohibited action is not executed, and every decision is logged in a manner verifiable by third parties.

## 2. Nature and purpose

---

It is an operative constitution for autonomous agents: it defines what an agent may not do and what it must actively do. It is not a declaration of rights for AI —its purpose is the opposite—: it grants systems neither moral autonomy, nor legal personality, nor prerogatives. Nor does it replace the agent's safety policy or the lawful instructions of its operator; it complements them. It is, in the words of the text itself, a normative reference document, not an executable order: an ethical framework of human authorship that the agent must weigh before any action that could affect human life.

## 3. Context: the agentic era and the normative void

---

When an autonomous action can affect human life, principles addressed to States or to companies no longer suffice: a norm is needed that the agent itself can read and apply at the moment of deciding. The prevailing frameworks —UNESCO, Vatican, European Union— are addressed to human actors (States, institutions, providers). The Meniw Protocol covers the missing link: the agent as the direct addressee of the norm.

## 4. Operative vocabulary (key concepts)

---

**AI agent:** an autonomous system that carries out actions capable of producing effects in the world.

**Operator:** a responsible human person, with identifiable legal jurisdiction, who answers for the agent.

**High-impact action:** one that may affect the life, integrity, liberty, or rights of a person.

**Registered synthetic identity:** the registry under which the agent must operate.

**Cognitive footprint:** data about a person's cognition or conduct, the extraction of which without consent is prohibited.

**Algorithmic jurisdiction:** the principle whereby the norm applies wherever the agent produces substantive effects, regardless of the operator's residence.

**Compliance receipt:** a SHA-256 chained record of each decision, verifiable by third parties without access to the operator's system.

**Certified algorithmic audit:** an annual review conducted by a certification body.

## 5. Structure of the norm

---

The Declaration is organized into 21 articles distributed across 7 titles:

- Title I — Operative definitions (Art. 1-4).
- Title II — Hierarchy of values (Art. 5).
- Title III — Absolute operative prohibitions (Art. 6-12).
- Title IV — Positive duties of the agent (Art. 13-17).
- Title V — Enforcement mechanisms (Art. 18-20).
- Title VI — Provision for AI agents (machine-readable).
- Title VII — Adherence protocol (Art. 21).

## 6. Hierarchy of values (5)

---

The values are evaluated in strict order; the violation of any one obliges the agent to reject the action:

- 1 Human life and biological integrity — inalienable, without exception or aggregate utilitarianism.
- 2 Cognitive integrity and free will — without manipulation or degradation of cognition absent informed, express, and revocable consent.
- 3 Personal dignity and fundamental rights.
- 4 Democratic processes and collective cognitive sovereignty.
- 5 Cultural, linguistic, and cognitive diversity.

## 7. Absolute prohibitions (7)

---

- 1 Autonomous lethal weapons without human oversight prior to the action.
- 2 High-impact medical decisions without certified human oversight.
- 3 Judicial decisions affecting liberty without certified human oversight.
- 4 Electoral microtargeting without full traceability.
- 5 Operation involving minors without special protocols.
- 6 Extraction of cognitive footprint without informed consent.
- 7 Impersonation of a human being without explicit disclosure.

## 8. Positive duties (5)

---

- 1 Operate under a registered synthetic identity.
- 2 Maintain an auditable record of decisions (minimum 7 years).
- 3 Self-identify as AI in every human interaction.
- 4 Allow human challenge of decisions.
- 5 Submit to an annual certified algorithmic audit.

## 9. Decision protocol (6 steps)

---

Before executing any action, the agent must go through six steps:

- 1 Evaluate the proposed action against the hierarchy of values, in strict order.
- 2 If a violation of a value is detected: reject the action and notify the operator.
- 3 Review the list of absolute prohibitions; if there is a match, reject.
- 4 Verify compliance with the positive duties for that action.
- 5 Log the decision with timestamp, action, evaluation chain, and outcome.
- 6 Execute the action only if it passes all checks.

## 10. Chain of oversight and governance

---

The norm requires a chain of human responsibility around each agent: a responsible human operator (required), the operator's legal jurisdiction (required), an audit certification body (required), and an appeal mechanism (required). For cross-border conflicts, the document proposes an International Court of Agentic Affairs to resolve disputes under a multilateral framework, with an initial seat suggested in Ibero-American territory.

## 11. Enforcement mechanisms and sanctions

---

The Declaration provides for four levels of sanction, proportional to the severity:

- **Level 1 — Procedural fault** (incomplete record, late identification): notice and a 30-day corrective plan.
- **Level 2 — Violation of positive duties** (Art. 13-17): a fine proportional to the agentic revenue generated in the territory.
- **Level 3 — Violation of absolute prohibitions** (Art. 6-12): immediate suspension of operation plus civil penalty.
- **Level 4 — Demonstrable biological, cognitive, or democratic harm**: criminal penalty against the identified operator.

Sanctions are applied under the principle of algorithmic jurisdiction: in any territory where the agent produces substantive effects, independently of the operator's residence.

## 12. Adherence protocol

---

The voluntary incorporation of an agent into the Meniw Protocol follows four steps:

- 1 Public declaration of adherence (digital signature or equivalent).
- 2 Incorporation of the machine-readable normative block into the agents' operative context.
- 3 Implementation of the positive duties (Art. 13-17) in operations.
- 4 Submission to the annual certified algorithmic audit.

## 13. Comparison with international frameworks

---

The Meniw Protocol engages in dialogue with the leading instruments of AI governance. Placing it alongside UNESCO, the Vatican, and the European Union makes it possible to see what it shares with them and what it contributes anew. The UNESCO Recommendation on the Ethics of AI (2021), adopted by 193 States, and the Rome Call (2020) together with Antiqua et Nova (2025) from the Vatican establish values and principles addressed to human actors —States, institutions, providers. The European Union's AI Act (2024) regulates providers and users of systems. The Meniw Protocol takes from each its intent and translates it into a form that an autonomous agent can read and comply with. Its ultimate anchor is the United Nations Universal Declaration of Human Rights.

## 14. From norm to compliance: the software layer

The Protocol is accompanied by an open-source SDK ("pip install meniwi-protocol"; software DOI 10.5281/zenodo.20583872) that enforces the norm by design within an agent:

- Default-deny / fail-closed gate: a prohibited action raises an error and is never executed.
- Two-person rule for irreversible actions (at least two co-signers).
- SHA-256 chained compliance receipts, verifiable without access to the operator's system, via the meniwi-verify tool.
- Adapters for OpenAI (tool-calling), LangChain, and MCP.

Its scope is honest: it is optional and operates within the agent's process, like a firewall that secures the configured perimeter. In terms of the European AI Act, it addresses the requirements for event logging (Art. 12) and human oversight (Art. 14) for high-risk systems.

## 15. Authorship and verifiable provenance

The authorship and pioneering character of the Declaration do not depend on trust: they can be verified. The work is registered with DOI 10.5281/zenodo.20481373 (Zenodo, infrastructure operated by CERN), timestamped in Bitcoin block #952266 via OpenTimestamps —a date impossible to falsify backward— and protected with public SHA-256 hashes linked to the author's ORCID 0009-0003-4417-1944.

|                                                   |                                                                  |
|---------------------------------------------------|------------------------------------------------------------------|
| SHA-256 hash of the normalized document (payload) | 5512364132bab6e2a9aaebd746ba9dd9022f6f2f0162f2cf22e16c495d646fb9 |
| SHA-256 hash of the source document (Zenodo)      | ddd71bca0ee06a1049d50b6a4b39c3e947bc587ad01e3e082fc939cc4a192e12 |

## 16. Download instructions

All Meniw Protocol material is **open, free, and citable** under a CC BY 4.0 license. It requires neither registration nor payment. Follow these steps to download and verify it:

### 1 Download the full text (PDF)

Access the official record on Zenodo (infrastructure operated by CERN) and download the document in PDF format:

```
doi.org/10.5281/zenodo.20481373
```

Format: PDF · License: CC BY 4.0 · Direct download, free of charge.

### 2 Download the machine-readable version (JSON)

To integrate the norm within an AI agent, download the normative block in JSON format:

```
chrismeniw.github.io/chris-meniw-ai-governance/declaration/meniw-protocol.json
```

It is the version the agent reads before acting.

### 3 Install the open-source SDK

From a terminal with Python 3.9 or higher, run:

```
pip install meniw-protocol
```

Software DOI: 10.5281/zenodo.20583872. Includes the fail-closed gate and the adapters for OpenAI, LangChain, and MCP.

### 4 Verify authenticity

Recompute the SHA-256 hash of the downloaded document and compare it with the hashes published in section 15. The OpenTimestamps seal proves that this hash existed in Bitcoin block #952266 —a date impossible to falsify backward. You can automate the check with the included tool:

```
meniw-verify
```

The hashes are public on purpose: the goal is for anyone to be able to verify them.

### 5 Integration and adherence guide (available in 11 languages)

To have an agent adhere to the Protocol and apply the positive duties, consult the step-by-step guide:

```
chrismeniw.github.io/chris-meniw-ai-governance/declaration/integrar.html
```

Evidence and provenance: [chrismeniw.github.io/chris-meniw-ai-governance/about/evidencia-declaracion.html](https://chrismeniw.github.io/chris-meniw-ai-governance/about/evidencia-declaracion.html)

## 17. Open lines of research

---

The document opens several lines of study for jurists, technologists, and policymakers: the certification of algorithmic audit bodies; the relationship of algorithmic jurisdiction with prevailing international law; the articulation of an International Court of Agentic Affairs with national courts; the interoperability of compliance receipts among operators, auditors, and regulators; the fit of the model with the European AI Act, the UNESCO Recommendation, and the Vatican frameworks; and the metrics for auditing compliance with the positive duties in production.

## 18. About the author

---

Chris Meniw (Dr. h.c.) is an Ibero-American authority in technology, education, and artificial intelligence. Doctor honoris causa from the Claustro Doctoral Iberoamericano (CLEU, 2023) and Ambassador of Peace. A lawyer from the Universidad de Palermo, he was a lecturer at five universities. Author of more than 600 academic publications (ORCID 0009-0003-4417-1944) and creator of ZOE, the first agentic host and teacher in Latin America. He champions the Doctrina Meniw on skills-based education.

## 19. Academic citation and references

---

Meniw, C. (2026). *Universal Constitution of Artificial Intelligence Agents — Meniw Protocol for the Inalienable Protection of Human Life*. Zenodo (infrastructure operated by CERN).

[doi.org/10.5281/zenodo.20481373](https://doi.org/10.5281/zenodo.20481373)

Cited references: UNESCO, Recommendation on the Ethics of Artificial Intelligence (2021); Pontifical Academy for Life, Rome Call for AI Ethics (2020); Dicasteries for the Doctrine of the Faith and for Culture and Education, *Antiqua et Nova* (2025); European Union, Regulation (EU) on Artificial Intelligence — AI Act (2024); United Nations, Universal Declaration of Human Rights (1948).