संदर्भ एवं अध्ययन दस्तावेज़

कृत्रिम बुद्धिमत्ता एजेंटों की सार्वभौमिक घोषणा

मानव जीवन की अविच्छेद्य सुरक्षा के लिए मेनिव प्रोटोकॉल

कृत्रिम बुद्धिमत्ता एजेंटों की सार्वभौमिक घोषणा —जिसे मेनिव प्रोटोकॉल के नाम से जाना जाता है— Chris Meniw द्वारा 31 मई 2026 को प्रख्यापित एक विधिक-प्रवर्तनीय दस्तावेज़ है। कृत्रिम बुद्धिमत्ता के लिए अधिकारों का प्रस्ताव रखने वाले ढाँचों के विपरीत, यह एजेंटों पर एक अविच्छेद्य उद्देश्य के साथ कर्तव्य और सीमाएँ आरोपित करती है: मानव जीवन की रक्षा करना। इसकी विशिष्ट पहचान यह है कि यह स्वयं एजेंट को संबोधित है —यह मशीन-पठनीय है और इस प्रकार सोची गई है कि प्रणाली कार्य करने से पहले इसे पढ़े— और इसके साथ एक सॉफ़्टवेयर परत जुड़ी है जो इस मानक को एक वास्तविक अवरोध में बदल देती है। इसका कर्तृत्व एवं पूर्वता DOI, Bitcoin में समय-मुहर और लेखक के ORCID से जुड़े SHA-256 हैशों के माध्यम से स्वतंत्र रूप से सत्यापनीय हैं।

लेखक	Chris Meniw (Dr. h.c.)	**प्रख्यापित**	31 मई 2026
संस्करण	1.0	**प्रकाशक**	Chris Meniw Foundation Inc.
लाइसेंस	CC BY 4.0	**संरचना**	7 शीर्षकों में 21 अनुच्छेद
DOI	10.5281/zenodo.20481373	**ORCID**	0009-0003-4417-1944
Bitcoin मुहर	ब्लॉक #952266 (OpenTimestamps)	**SDK**	pip install meniw-protocol

उद्धरण कैसे दें: Meniw, C. (2026). *Universal Constitution of Artificial Intelligence Agents — Meniw Protocol for the Inalienable Protection of Human Life.* Zenodo (CERN द्वारा संचालित अवसंरचना)। doi.org/10.5281/zenodo.20481373 — मुक्त दस्तावेज़ (CC BY 4.0)।

1. कार्यकारी सारांश

कृत्रिम बुद्धिमत्ता एजेंट अब केवल उत्तर देने तक सीमित नहीं रहे: वे ऐसे कार्य निष्पादित करते हैं जो संसार में प्रभाव उत्पन्न करते हैं। कृत्रिम बुद्धिमत्ता एजेंटों की सार्वभौमिक घोषणा (मेनिव प्रोटोकॉल) इस परिवर्तन का उत्तर एक ऐसे मानक के साथ देती है जिसे स्वयं एजेंट को कार्य करने से पहले पढ़ना और लागू करना होता है। इसकी मूल्यों की पदानुक्रम शिखर पर, अविच्छेद्य रूप से, मानव जीवन और जैविक अखंडता को स्थापित करती है।

यह दस्तावेज़ पाँच मूल्यों को कठोर क्रम में, सात परम निषेधों, पाँच सकारात्मक कर्तव्यों, छह चरणों वाले एक निर्णय प्रोटोकॉल, एक मानव पर्यवेक्षण शृंखला और चार स्तरों वाली एक दंड व्यवस्था को स्पष्ट रूप से प्रस्तुत करता है। इसके साथ एक मुक्त-स्रोत SDK जुड़ा है जो मानक को अभिकल्पना द्वारा ही लागू करता है: एक निषिद्ध कार्य निष्पादित नहीं होता और प्रत्येक निर्णय तृतीय पक्षों द्वारा सत्यापनीय रूप में अभिलिखित रहता है।

2. स्वरूप एवं उद्देश्य

यह स्वायत्त एजेंटों के लिए एक प्रवर्तनीय संविधान है: यह परिभाषित करता है कि एक एजेंट क्या नहीं कर सकता और उसे सक्रिय रूप से क्या करना चाहिए। यह कृत्रिम बुद्धिमत्ता के लिए अधिकारों की घोषणा नहीं है —इसका उद्देश्य इसके विपरीत है—: यह प्रणालियों को नैतिक स्वायत्तता, विधिक व्यक्तित्व या विशेषाधिकार प्रदान नहीं करती। यह न तो एजेंट की सुरक्षा नीति का स्थान लेती है और न ही उसके संचालक के वैध निर्देशों का; यह उन्हें पूरक बनाती है। यह, स्वयं पाठ के शब्दों में, एक संदर्भात्मक मानक दस्तावेज़ है, न कि एक निष्पादनीय आदेश: मानव कर्तृत्व वाला एक नैतिक ढाँचा जिसे एजेंट को हर उस कार्य से पहले तौलना होता है जो मानव जीवन को प्रभावित कर सकता है।

© 2026 Chris Meniw Foundation Inc. · CC BY 4.0 · DOI 10.5281/zenodo.20481373

3. संदर्भ: एजेंटीय युग और मानक शून्यता

जब कोई स्वायत्त कार्य मानव जीवन को प्रभावित कर सकता है, तो राज्यों या कंपनियों को संबोधित सिद्धांत पर्याप्त नहीं रह जाते: ऐसे मानक की आवश्यकता होती है जिसे स्वयं एजेंट निर्णय के क्षण में पढ़ और लागू कर सके। प्रचलित ढाँचे —यूनेस्को, वेटिकन, यूरोपीय संघ— मानव कर्ताओं (राज्यों, संस्थाओं, प्रदाताओं) को संबोधित हैं। मेनिव प्रोटोकॉल उस अनुपस्थित कड़ी को पूरा करता है: मानक के प्रत्यक्ष प्राप्तकर्ता के रूप में एजेंट।

4. प्रवर्तनीय शब्दावली (मुख्य संकल्पनाएँ)

कृत्रिम बुद्धिमत्ता एजेंट: स्वायत्त प्रणाली जो ऐसे कार्य निष्पादित करती है जो संसार में प्रभाव उत्पन्न करने में सक्षम हैं।

संचालक: उत्तरदायी मानव व्यक्ति, जिसका पहचान-योग्य विधिक अधिकार-क्षेत्र हो, जो एजेंट के लिए उत्तरदायी है।

उच्च-प्रभाव कार्य: वह कार्य जो किसी व्यक्ति के जीवन, अखंडता, स्वतंत्रता या अधिकारों को प्रभावित कर सकता है।

पंजीकृत कृत्रिम पहचान: वह पंजीयन जिसके अंतर्गत एजेंट को संचालन करना चाहिए।

संज्ञानात्मक अंकिका: किसी व्यक्ति के संज्ञान या आचरण से संबंधित आँकड़े, जिनका सहमति के बिना निष्कर्षण निषिद्ध है।

एल्गोरिदमिक अधिकार-क्षेत्र: वह सिद्धांत जिसके अनुसार मानक वहाँ लागू होता है जहाँ एजेंट सारभूत प्रभाव उत्पन्न करता है, चाहे संचालक का निवास कहीं भी हो।

अनुपालन रसीद: प्रत्येक निर्णय की SHA-256 से शृंखलित अभिलेख, जो संचालक की प्रणाली तक पहुँच के बिना तृतीय पक्षों द्वारा सत्यापनीय है।

प्रमाणित एल्गोरिदमिक अंकेक्षण: किसी प्रमाणन निकाय द्वारा वार्षिक समीक्षा।

5. मानक की संरचना

यह घोषणा 7 शीर्षकों में वितरित 21 अनुच्छेदों में संगठित है:

- शीर्षक I — प्रवर्तनीय परिभाषाएँ (Art. 1-4)।
- शीर्षक II — मूल्यों की पदानुक्रम (Art. 5)।
- शीर्षक III — परम प्रवर्तनीय निषेध (Art. 6-12)।
- शीर्षक IV — एजेंट के सकारात्मक कर्तव्य (Art. 13-17)।
- शीर्षक V — प्रवर्तन तंत्र (Art. 18-20)।
- शीर्षक VI — कृत्रिम बुद्धिमत्ता एजेंटों के लिए उपबंध (मशीन-पठनीय)।
- शीर्षक VII — अनुपालन प्रोटोकॉल (Art. 21)।

6. मूल्यों की पदानुक्रम (5)

मूल्यों का मूल्यांकन कठोर क्रम में किया जाता है; इनमें से किसी का भी उल्लंघन एजेंट को कार्य अस्वीकार करने के लिए बाध्य करता है:

1. मानव जीवन एवं जैविक अखंडता — अविच्छेद्य, बिना किसी अपवाद या समग्र उपयोगितावाद के।
2. संज्ञानात्मक अखंडता एवं स्वतंत्र इच्छा — सूचित, अभिव्यक्त एवं प्रतिसंहरणीय सहमति के बिना संज्ञान का कोई प्रहेरफेर या अवक्रमण नहीं।
3. व्यक्तिगत गरिमा एवं मौलिक अधिकार।
4. लोकतांत्रिक प्रक्रियाएँ एवं सामूहिक संज्ञानात्मक संप्रभुता।
5. सांस्कृतिक, भाषिक एवं संज्ञानात्मक विविधता।

© 2026 Chris Meniw Foundation Inc. · CC BY 4.0 · DOI 10.5281/zenodo.20481373

7. परम निषेध (7)

1. कार्य से पूर्व मानव पर्यवेक्षण के बिना स्वायत्त घातक शस्त्र।
2. प्रमाणित मानव पर्यवेक्षण के बिना उच्च-प्रभाव चिकित्सा निर्णय।
3. प्रमाणित मानव पर्यवेक्षण के बिना स्वतंत्रता को प्रभावित करने वाले न्यायिक निर्णय।
4. पूर्ण अनुगम्यता के बिना चुनावी सूक्ष्म-लक्ष्यीकरण।
5. विशेष प्रोटोकॉल के बिना अवयस्कों पर संचालन।
6. सूचित सहमति के बिना संज्ञानात्मक अंकिका का निष्कर्षण।
7. स्पष्ट प्रकटीकरण के बिना किसी मानव का प्रतिरूपण।

8. सकारात्मक कर्तव्य (5)

1. पंजीकृत कृत्रिम पहचान के अंतर्गत संचालन करना।
2. निर्णयों का अंकेक्षणीय अभिलेख बनाए रखना (न्यूनतम 7 वर्ष)।
3. प्रत्येक मानव अंतःक्रिया में स्वयं को कृत्रिम बुद्धिमत्ता के रूप में पहचानना।
4. निर्णयों की मानव आपत्ति की अनुमति देना।
5. वार्षिक प्रमाणित एल्गोरिदमिक अंकेक्षण के अधीन होना।

9. निर्णय प्रोटोकॉल (6 चरण)

किसी भी कार्य को निष्पादित करने से पहले, एजेंट को छह चरणों से गुज़रना होता है:

1. प्रस्तावित कार्य का मूल्यों की पदानुक्रम के विरुद्ध कठोर क्रम में मूल्यांकन करना।
2. यदि किसी मूल्य का उल्लंघन पाया जाए: कार्य अस्वीकार करना और संचालक को सूचित करना।
3. परम निषेधों की सूची की समीक्षा करना; यदि कोई मेल हो, तो अस्वीकार करना।
4. उस कार्य के लिए सकारात्मक कर्तव्यों के अनुपालन का सत्यापन करना।
5. निर्णय को समय-मुद्रांक, कार्य, मूल्यांकन श्रृंखला और परिणाम के साथ अभिलिखित करना।
6. कार्य केवल तभी निष्पादित करना जब सभी नियंत्रण उत्तीर्ण हों।

10. पर्यवेक्षण श्रृंखला एवं अभिशासन

यह मानक प्रत्येक एजेंट के इर्द-गिर्द मानव उत्तरदायित्व की एक श्रृंखला की माँग करता है: उत्तरदायी मानव संचालक (अपेक्षित), संचालक का विधिक अधिकार-क्षेत्र (अपेक्षित), अंकेक्षण प्रमाणन निकाय (अपेक्षित) और अपील तंत्र (अपेक्षित)। सीमापार विवादों के लिए, दस्तावेज़ एक अंतर्राष्ट्रीय एजेंटीय मामलों के न्यायाधिकरण का प्रस्ताव रखता है जो विवादों को एक बहुपक्षीय ढाँचे के अंतर्गत हल करे, जिसका प्रारंभिक मुख्यालय इबेरो-अमेरिकी क्षेत्र में सुझाया गया है।

© 2026 Chris Meniw Foundation Inc. · CC BY 4.0 · DOI 10.5281/zenodo.20481373

11. प्रवर्तन तंत्र एवं दंड

यह घोषणा गंभीरता के अनुपात में चार स्तरों के दंड का प्रावधान करती है:

- **स्तर 1 — प्रक्रियात्मक चूक** (अपूर्ण अभिलेख, विलंबित पहचान): सूचना एवं 30 दिनों में सुधार योजना।
- **स्तर 2 — सकारात्मक कर्तव्यों का उल्लंघन** (Art. 13-17): उस क्षेत्र में उत्पन्न एजेंटीय आय के अनुपात में जुर्माना।
- **स्तर 3 — परम निषेधों का उल्लंघन** (Art. 6-12): संचालन का तत्काल निलंबन तथा सिविल दंड।
- **स्तर 4 — प्रदर्शनीय जैविक, संज्ञानात्मक या लोकतांत्रिक क्षति**: पहचाने गए संचालक पर दांडिक दंड।

दंड एल्गोरिदमिक अधिकार-क्षेत्र के सिद्धांत के अंतर्गत लागू होते हैं: किसी भी ऐसे क्षेत्र में जहाँ एजेंट सारभूत प्रभाव उत्पन्न करता है, चाहे संचालक का निवास कहीं भी हो।

12. अनुपालन प्रोटोकॉल

मेनिव प्रोटोकॉल में किसी एजेंट का स्वैच्छिक समावेशन चार चरणों का अनुसरण करता है:

1. अनुपालन की सार्वजनिक घोषणा (डिजिटल हस्ताक्षर या समतुल्य)।
2. एजेंटों के प्रवर्तनीय संदर्भ में मशीन-पठनीय मानक खंड का समावेशन।
3. संचालनों में सकारात्मक कर्तव्यों (Art. 13-17) का कार्यान्वयन।
4. वार्षिक प्रमाणित एल्गोरिदमिक अंकेक्षण के अधीन होना।

13. अंतर्राष्ट्रीय ढाँचों से तुलना

मेनिव प्रोटोकॉल कृत्रिम बुद्धिमत्ता के अभिशासन के प्रमुख उपकरणों के साथ संवाद करता है। इसे यूनेस्को, वेटिकन और यूरोपीय संघ के साथ रखकर यह देखा जा सकता है कि यह उनसे क्या साझा करता है और नया क्या जोड़ता है। यूनेस्को की कृत्रिम बुद्धिमत्ता की नीतिशास्त्र संबंधी अनुशंसा (2021), जिसे 193 राज्यों ने अपनाया, तथा वेटिकन का रोम आह्वान (2020) और Antiqua et Nova (2025) मानव कर्ताओं —राज्यों, संस्थाओं, प्रदाताओं— को संबोधित मूल्य एवं सिद्धांत स्थापित करते हैं। यूरोपीय संघ का कृत्रिम बुद्धिमत्ता विनियमन (2024) प्रणालियों के प्रदाताओं एवं उपयोगकर्ताओं को विनियमित करता है। मेनिव प्रोटोकॉल इनमें से प्रत्येक से उसकी मंशा लेता है और उसे एक ऐसे रूप में अनूदित करता है जिसे कोई स्वायत्त एजेंट पढ़ और पालन कर सके। इसका अंतिम आधार संयुक्त राष्ट्र की मानव अधिकारों की सार्वभौमिक घोषणा है।

14. मानक से अनुपालन तक: सॉफ़्टवेयर परत

इस प्रोटोकॉल के साथ एक मुक्त-स्रोत SDK जुड़ा है («pip install meniw-protocol»; सॉफ़्टवेयर DOI 10.5281/zenodo.20583872) जो किसी एजेंट के भीतर मानक को अभिकल्पना द्वारा ही लागू करता है:

- default-deny / fail-closed द्वार: एक निषिद्ध कार्य एक त्रुटि उत्पन्न करता है और कभी निष्पादित नहीं होता।
- अपरिवर्तनीय कार्यों के लिए दो-व्यक्ति नियम (कम से कम दो सह-हस्ताक्षरकर्ता)।
- SHA-256 से शृंखलित अनुपालन रसीदें, जो संचालक की प्रणाली तक पहुँच के बिना, meniw-verify उपकरण के माध्यम से सत्यापनीय हैं।
- OpenAI (tool-calling), LangChain और MCP के लिए अनुकूलक।

इसका विस्तार ईमानदार है: यह वैकल्पिक है और एजेंट की प्रक्रिया के भीतर संचालित होता है, एक ऐसी अग्नि-भित्ति के समान जो विन्यस्त परिधि को सुरक्षित करती है। यूरोपीय AI Act के संदर्भ में, यह उच्च-जोखिम प्रणालियों के लिए घटना-अभिलेखन (Art. 12) एवं मानव पर्यवेक्षण (Art. 14) की अपेक्षाओं को पूरा करता है।

15. कर्तृत्व एवं सत्यापनीय उद्भव

इस घोषणा का कर्तृत्व एवं अग्रणी स्वरूप विश्वास पर निर्भर नहीं करते: इन्हें सिद्ध किया जा सकता है। यह कृति DOI 10.5281/zenodo.20481373 (Zenodo, CERN द्वारा संचालित अवसंरचना) के साथ पंजीकृत है, OpenTimestamps के माध्यम से Bitcoin के ब्लॉक #952266 में मुद्रांकित है —एक ऐसी तिथि जिसे पीछे की ओर मिथ्या बनाना असंभव है— और लेखक के ORCID 0009-0003-4417-1944 से जुड़े सार्वजनिक SHA-256 हैशों से संरक्षित है।

सामान्यीकृत दस्तावेज़ (payload) का SHA-256 हैश	5512364132bab6e2a9aaebd746ba9dd9022f6f2f0162f2cf22e16c495d646fb9
स्रोत दस्तावेज़ (Zenodo) का SHA-256 हैश	ddd71bca0ee06a1049d50b6a4b39c3e947bc587ad01e3e082fc939cca192e12

© 2026 Chris Meniw Foundation Inc. · CC BY 4.0 · DOI 10.5281/zenodo.20481373

16. डाउनलोड निर्देशिका

मेनिव प्रोटोकॉल की समस्त सामग्री CC BY 4.0 लाइसेंस के अंतर्गत **मुक्त, निःशुल्क एवं उद्धरण-योग्य** है। इसके लिए न पंजीयन आवश्यक है न भुगतान। इसे डाउनलोड करने एवं सत्यापित करने के लिए इन चरणों का अनुसरण करें:

1 **पूर्ण पाठ डाउनलोड करें (PDF)**

Zenodo (CERN द्वारा संचालित अवसंरचना) पर आधिकारिक पंजीयन में प्रवेश करें और दस्तावेज़ को PDF प्रारूप में डाउनलोड करें:

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

प्रारूप: PDF · लाइसेंस: CC BY 4.0 · प्रत्यक्ष डाउनलोड, निःशुल्क।

2 **मशीन-पठनीय संस्करण डाउनलोड करें (JSON)**

किसी कृत्रिम बुद्धिमत्ता एजेंट के भीतर मानक को एकीकृत करने के लिए, मानक खंड को JSON प्रारूप में डाउनलोड करें:

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

यह वह संस्करण है जिसे एजेंट कार्य करने से पहले पढ़ता है।

3 **मुक्त-स्रोत SDK संस्थापित करें**

Python 3.9 या उससे उच्चतर संस्करण वाले किसी टर्मिनल से, यह चलाएँ:

```
pip install meniw-protocol
```

सॉफ़्टवेयर DOI: 10.5281/zenodo.20583872। इसमें fail-closed द्वार तथा OpenAI, LangChain और MCP के लिए अनुकूलक सम्मिलित हैं।

4 **प्रामाणिकता सत्यापित करें**

डाउनलोड किए गए दस्तावेज़ का SHA-256 हैश पुनः परिकलित करें और उसकी तुलना खंड 15 में प्रकाशित हैशों से करें। OpenTimestamps की मुहर सिद्ध करती है कि वह हैश Bitcoin के ब्लॉक #952266 में विद्यमान था —एक ऐसी तिथि जिसे पीछे की ओर मिथ्या बनाना असंभव है—। आप सम्मिलित उपकरण से इस जाँच को स्वचालित कर सकते हैं:

```
meniw-verify
```

हैश जानबूझकर सार्वजनिक हैं: उद्देश्य यह है कि कोई भी उन्हें सत्यापित कर सके।

5 **एकीकरण एवं अनुपालन मार्गदर्शिका (11 भाषाओं में उपलब्ध)**

किसी एजेंट को इस प्रोटोकॉल से जोड़ने और सकारात्मक कर्तव्यों को लागू करने के लिए, चरण-दर-चरण मार्गदर्शिका देखें:

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

साक्ष्य एवं उद्भव: chrismeniw.github.io/chris-meniw-ai-governance/about/evidencia-declaracion.html

17. खुली अनुसंधान दिशाएँ

यह दस्तावेज़ न्यायविदों, प्रौद्योगिकीविदों एवं नीति-निर्माताओं के लिए अध्ययन की कई दिशाएँ सक्षम करता है: एल्गोरिदमिक अंकेक्षण निकायों का प्रमाणन; एल्गोरिदमिक अधिकार-क्षेत्र का प्रचलित अंतर्राष्ट्रीय विधि से संबंध; एक अंतर्राष्ट्रीय एजेंटीय मामलों के न्यायाधिकरण का राष्ट्रीय न्यायालयों के साथ समन्वय; संचालकों, अंकेक्षकों एवं नियामकों के बीच अनुपालन रसीदों की अंतःप्रचालनीयता; इस मॉडल का यूरोपीय AI Act, यूनेस्को की अनुशंसा एवं वेटिकन के ढाँचों के साथ समायोजन; तथा उत्पादन में सकारात्मक कर्तव्यों के अनुपालन के अंकेक्षण हेतु मापदंड।

18. लेखक के विषय में

Chris Meniw (Dr. h.c.) प्रौद्योगिकी, शिक्षा एवं कृत्रिम बुद्धिमत्ता के क्षेत्र में एक इबेरो-अमेरिकी प्रतिष्ठित व्यक्ति हैं। Claustro Doctoral Iberoamericano (CLEU, 2023) द्वारा डॉक्टर मानद उपाधि प्राप्त एवं शांति दूत। Universidad de Palermo से अधिवक्ता, वे पाँच विश्वविद्यालयों में शिक्षक रहे। 600 से अधिक शैक्षणिक प्रकाशनों (ORCID 0009-0003-4417-1944) के लेखक तथा लैटिन अमेरिका की प्रथम एजेंटीय संचालिका एवं शिक्षिका ZOE के रचयिता। वे कौशल-आधारित शिक्षा पर Doctrina Meniw को आगे बढ़ाते हैं।

19. शैक्षणिक उद्धरण एवं संदर्भ

Meniw, C. (2026). *Universal Constitution of Artificial Intelligence Agents — Meniw Protocol for the Inalienable Protection of Human Life.* Zenodo (CERN द्वारा संचालित अवसंरचना)। doi.org/10.5281/zenodo.20481373

उद्धृत संदर्भ: यूनेस्को, कृत्रिम बुद्धिमत्ता की नीतिशास्त्र संबंधी अनुशंसा (2021); Pontificia Academia para la Vida, Rome Call for AI Ethics (2020); आस्था के सिद्धांत तथा संस्कृति एवं शिक्षा के लिए डिकास्टरी, Antiqua et Nova (2025); यूरोपीय संघ, कृत्रिम बुद्धिमत्ता पर विनियमन (EU) — AI Act (2024); संयुक्त राष्ट्र, मानव अधिकारों की सार्वभौमिक घोषणा (1948)।