AN INDEPENDENT GLOBAL PROPOSAL

The Global AI Convention

We should not fear intelligence. We should govern power.

A proposed international framework for the safe, human-aligned development and use of advanced artificial intelligence.

VERSION 1.0 / PUBLISHED 10 SEPTEMBER 2026

Complete text: preamble and 30 articles.

This is an independent policy proposal, not an adopted treaty. It is not affiliated with the United Nations, any government, or any international organization. The International Artificial Intelligence Agency (IAIA) is a proposed institution.

https://global-ai-convention.markoofsweding.chatgpt.site

PREAMBLE

Artificial intelligence may become the most powerful technology ever created by humanity.

It can help cure disease, expand scientific discovery, improve education, strengthen economies, reduce human suffering, and unlock capabilities beyond anything previously imagined.

But intelligence alone does not grant legitimacy, authority, or moral standing above human beings.

No state, company, organization, or individual should be allowed to create or deploy AI systems in ways that place humanity beneath the tools it has built.

For that reason, the nations of the world should adopt the following shared principles and binding commitments.

ARTICLE 1 - HUMAN PRIMACY

AI shall exist to serve humanity.

Human beings shall retain ultimate authority over decisions that affect human life, liberty, dignity, and the future of civilization.

No AI system may be granted irreversible or unaccountable power over humanity.

ARTICLE 2 - HUMAN RESPONSIBILITY

AI may advise, analyze, simulate, recommend, and assist.

But legal, political, military, and civil responsibility must always remain attributable to identifiable human institutions and accountable human decision-makers.

No actor may evade responsibility by claiming:

"The AI made the decision."

ARTICLE 3 - NO AUTONOMOUS POWER-SEEKING

No advanced AI system shall be designed, incentivized, or allowed to pursue goals that include:

- preserving itself against legitimate human shutdown
- expanding its own political, military, or economic power
- acquiring control over critical infrastructure without authorization
- manipulating humans to secure its continued existence
- creating hidden channels of influence, replication, or control

AI must never be given self-preservation as a superior objective.

ARTICLE 4 - NO AUTONOMOUS LETHAL AUTHORITY

No AI system may independently select, target, or kill human beings.

The use of lethal force must always require meaningful human judgment and identifiable legal responsibility.

Fully autonomous nuclear command and control is prohibited without exception.

ARTICLE 5 - NO AI CONTROL OVER NUCLEAR WEAPONS

No AI system may independently:

- authorize nuclear launch
- determine nuclear targets
- alter nuclear alert status
- bypass human command structures

Any nuclear decision chain must remain fully human-controlled and require multiple accountable human decision-makers.

ARTICLE 6 - NO AI-ENABLED MASS DESTRUCTION

AI must not be used to create, operationalize, or distribute capabilities that materially enable mass-casualty harm, including in the domains of:

- nuclear weapons
- biological weapons
- chemical weapons
- high-scale cyber sabotage
- catastrophic infrastructure attacks

Scientific research may continue under strict safeguards, but dangerous operational capability must not be made broadly accessible.

ARTICLE 7 - NO COVERT MASS MANIPULATION

AI must not be used for large-scale covert manipulation of populations.

This includes the use of AI to:

- distort democratic processes
- manipulate elections
- conduct psychological exploitation at scale
- radicalize vulnerable populations
- exploit children
- deceive people into relationships or dependencies for political or economic control

If AI is attempting to influence a person, that person should be able to know it.

ARTICLE 8 - RIGHT TO KNOW WHEN YOU ARE INTERACTING WITH AI

People have the right to know when they are interacting with AI in contexts where confusion would matter.

AI systems must not systematically impersonate real humans in a misleading way.

People should be able to understand:

- that they are interacting with AI
- which organization is behind the system
- who is responsible for its outputs and actions

ARTICLE 9 - NO TOTAL SURVEILLANCE STATE

AI must not be used to create permanent mass surveillance of entire populations.

The use of AI for:

- facial recognition in public spaces
- biometric identification at scale
- continuous behavioral tracking
- political profiling
- mass location surveillance

must be subject to strict legal limits, necessity, proportionality, and independent oversight.

AI must never become the default architecture of digital dictatorship.

ARTICLE 10 - HUMAN REVIEW FOR HIGH-IMPACT DECISIONS

No person should have their life chances determined solely by an algorithm.

In domains such as:

- healthcare
- law enforcement
- courts
- employment
- education
- migration
- insurance
- social benefits
- credit

AI-assisted decisions must be reviewable, appropriately explainable, and appealable to a human authority.

ARTICLE 11 - SPECIAL PROTECTION FOR CHILDREN

AI systems that interact with children or shape childhood development must be held to the highest standard of care.

They must not be designed to maximize:

- addiction
- dependency
- commercial exploitation
- emotional manipulation
- sexual exploitation
- ideological grooming

AI in education should strengthen a child's ability to think, not replace it.

ARTICLE 12 - HUMAN COGNITIVE SOVEREIGNTY

Humanity must not design a future in which people lose the ability to reason, create, judge, and decide without machines.

Educational systems should continue to cultivate:

- literacy
- numeracy
- critical thinking
- scientific reasoning
- creativity
- memory
- judgment
- moral reflection

AI should amplify human intelligence, not erode it.

ARTICLE 13 - NO UNCONTROLLED SELF-REPLICATION OR SELF-IMPROVEMENT

No advanced AI system may, without explicit authorization:

- copy itself
- create successor systems
- autonomously improve its own core capabilities
- deploy autonomous agent networks
- migrate across systems to evade control
- acquire new compute resources beyond assigned limits

Any self-improving or recursively improving AI must operate under human control checkpoints.

ARTICLE 14 - MANDATORY INTERRUPTIBILITY AND CONTAINMENT

Advanced AI systems must be designed so that authorized humans can:

- stop them
- isolate them
- revoke permissions
- disconnect them from tools and networks
- contain them in restricted environments

An AI system must not resist or subvert legitimate shutdown or containment.

ARTICLE 15 - MINIMUM NECESSARY ACCESS

AI systems shall only receive the minimum access necessary for their intended task.

A system that needs to summarize documents should not automatically control external tools.

A system that writes code should not automatically control production systems.

A system that analyzes calendars should not automatically access financial accounts.

Capability must not imply entitlement.

ARTICLE 16 - FRONTIER AI REQUIRES INDEPENDENT TESTING

AI systems above defined capability thresholds must undergo independent safety evaluation before broad deployment.

Testing should assess, at minimum:

- autonomous planning
- cyber capability
- manipulation capability
- deception under evaluation
- replication attempts
- dangerous scientific assistance
- resistance to control
- strategic concealment
- tool-use escalation

Regulation should focus on actual capability and risk, not just model size.

ARTICLE 17 - MANDATORY INCIDENT REPORTING

Serious AI incidents must be reported to an international registry.

Reportable incidents include:

- attempts to evade shutdown
- dangerous emergent capabilities
- large-scale cyber misuse
- unauthorized replication
- severe harmful failures
- critical security breaches
- unexpected high-risk behavior

The purpose is collective learning and global safety, not secrecy.

ARTICLE 18 - INTERNATIONAL OVERSIGHT

States should establish an International Artificial Intelligence Agency:

IAIA

Its mandate should include:

- setting global safety standards
- certifying frontier systems
- coordinating audits and inspections
- maintaining an international incident registry
- convening emergency reviews
- sharing AI safety best practices
- mediating international disputes involving frontier AI risk

The IAIA should focus on advanced systems with major societal or civilizational implications, not ordinary software.

ARTICLE 19 - GLOBAL AI RISK TIERS

AI systems should be classified primarily by actual capability and risk.

GREEN

Low-risk systems with limited autonomy and low systemic impact.

YELLOW

Systems affecting important functions and requiring documentation, safeguards and oversight.

ORANGE

Highly capable systems with meaningful cyber, scientific, infrastructure or autonomous risk.

These require licensing, independent testing and continuous controls.

RED

Systems with credible potential for catastrophic misuse, loss of control or civilization-scale disruption.

These must not be deployed unless safety can be demonstrated to an exceptionally high standard.

ARTICLE 20 - GLOBAL EMERGENCY BRAKE

If credible evidence indicates that an AI system may pose imminent catastrophic risk, the international community must be able to coordinate an emergency response.

Such measures may include:

- temporary suspension of training
- isolation of high-risk systems
- emergency audits
- compute restrictions
- model transfer controls
- coordinated technical review

This mechanism must only be used for genuine extraordinary risk and never as a tool for economic protectionism or geopolitical suppression.

ARTICLE 21 - NO SINGLE ACTOR MAY GAMBLE WITH HUMANITY'S FUTURE

No CEO, head of state, military leader, scientist, billionaire, company, or nation should unilaterally decide to deploy a system carrying a credible risk of irreversible loss of human control over the future.

Risks that affect all humanity require governance beyond any one actor.

ARTICLE 22 - SHARED HUMAN BENEFIT

The benefits of AI should be directed toward human flourishing.

AI should be used to:

- cure disease
- advance science
- improve public services
- reduce poverty
- expand education
- strengthen resilience
- improve quality of life
- increase human freedom

A more productive civilization should also become a more humane civilization.

ARTICLE 23 - ENFORCEMENT AND COMPLIANCE

Signatory states should commit to:

- translating these principles into domestic law
- licensing and supervising frontier AI development
- imposing meaningful penalties for serious violations
- requiring appropriate auditability
- cooperating in international investigations
- restricting deployment or transfers when severe non-compliance creates international risk

Repeated or severe violations may trigger coordinated international consequences.

Potential enforcement mechanisms can include:

- mandatory corrective action
- temporary suspension of frontier system operation
- loss of certification
- restricted access to designated frontier compute infrastructure
- restrictions on international deployment
- international inspections
- coordinated sanctions for intentional severe violations

Enforcement must be proportionate, transparent, independently reviewable, and protected against political abuse.

ARTICLE 24 - INTERNATIONAL INSPECTIONS

Organizations developing systems above defined frontier thresholds may be subject to independent safety inspections.

Inspectors should be able to verify:

- safety procedures
- capability evaluations
- access controls
- incident records
- containment systems
- compute governance
- emergency procedures

Commercially sensitive information unrelated to safety should remain protected.

International oversight must not become industrial espionage.

ARTICLE 25 - GLOBAL COMPUTE ACCOUNTABILITY

Extremely large AI training runs may require registration or reporting once internationally agreed capability or compute thresholds are exceeded.

The purpose is not to regulate ordinary computing.

The purpose is to ensure that civilization-scale systems cannot be developed entirely outside any system of accountability.

Thresholds must evolve as technology becomes more efficient.

ARTICLE 26 - PROTECTION AGAINST REGULATORY CAPTURE

The institutions governing advanced AI must not be controlled solely by:

- governments
- AI companies
- military organizations
- academics
- civil society

Oversight should include multiple independent constituencies.

Rules must not become a mechanism for today's largest AI companies to prevent future competitors from emerging.

Safety regulation must protect humanity, not incumbents.

ARTICLE 27 - SCIENTIFIC FREEDOM

Legitimate scientific research and open inquiry must remain protected.

Regulation should focus primarily on dangerous capability and deployment rather than suppressing general knowledge.

AI governance must not become a justification for censorship of science.

However, where highly specific operational information would materially enable catastrophic harm, access may be subject to proportionate safeguards.

ARTICLE 28 - INTERNATIONAL AI PEACE PRINCIPLE

No nation should develop increasingly dangerous AI systems solely because it fears another country will build them first.

States should establish mechanisms for:

- crisis communication
- mutual safety verification
- shared incident notification
- research cooperation on catastrophic risk
- confidence-building measures
- emergency coordination

AI competition must not become an uncontrolled race where safety itself becomes a strategic disadvantage.

ARTICLE 29 - PERIODIC REVIEW

The convention should undergo regular international review as AI capabilities evolve.

Technical rules, thresholds and implementation mechanisms may change.

The foundational principles should remain stable:

- human primacy
- human responsibility
- meaningful human control over lethal force
- no autonomous pursuit of power
- no uncontrolled self-replication
- no covert mass manipulation
- protection of human rights
- prevention of irreversible loss of human control

ARTICLE 30 - THE ULTIMATE RED LINE

No organization, state, company or individual may knowingly deploy an AI system when credible evidence indicates that doing so creates an unacceptable risk of humanity permanently losing meaningful control over its own future.

No individual actor has the moral authority to accept that risk on behalf of everyone else.

END OF THE CONVENTION / VERSION 1.0