



INSTITUT  
AUF DEM  
ROSENBERG

# THE ROSENBERG AI CHARTER

*for Ethical AI Governance  
by Students of Institut auf dem Rosenberg*

## Preamble

We, the students of Institut auf dem Rosenberg, the future stewards of this planet, with unwavering confidence, call upon global leaders to promptly regulate AI development and usage through an international treaty and regulatory agency. Acknowledging the profound impact of AI on society and the environment, we assertively implore leaders to establish a charter ensuring the responsible use of AI across metaverses, terrestrial domains, and beyond. This charter articulates our ethical guidelines to safeguard individuals, communities, and the common good.

– 1 of 5 –

Institut auf dem Rosenberg  
Höhenweg 60, 9000 St. Gallen, Switzerland

THE ARTISANS OF EDUCATION™

ARTICLE I  
CONTROL OF INPUT AND OUTPUT

I.

All organizations, whether private or state, engaged in designing, engineering, and/or distributing AI products, shall be held unequivocally accountable for the information generated by AI systems.

II.

These organizations must establish specialized departments amalgamating human oversight and automated technologies grounded in machine learning to guarantee the responsible utilization of AI.

III.

An external, impartial global agency shall meticulously oversee and ensure strict adherence to proper AI usage, conferring AI-Safe-Use approval badges exclusively upon organizations that diligently comply with AI standards.

ARTICLE II  
TRANSPARENT TRACING OF SOURCES

I.

Complete transparency in acknowledging the entities responsible for AI processes is imperative. Therefore, all AI-processed information must be transparently traceable to its origins, specifically attributing it to the entities conducting the information processing using AI.

II.

Users shall enjoy unrestricted access to all original input data employed by AI systems.

III.

Violations of source tracing obligations will be met with resolute legal enforcement.

ARTICLE III  
REGULATION OF DEEPPFAKES

I.

Mandatory watermarks or detectable patterns are recommended for all deepfake or artificially created content.

II.

We advocate increased investment in deepfake detection technologies.

III.

Unethical deepfake actions, including defamation and identity theft, must be unequivocally prosecutable offenses.

IV.

AI systems shall rigorously maintain accessible interaction histories, with AI software manufacturers being legally accountable for verifying the origin of disseminated information.

ARTICLE IV  
PREVENTION OF MONOPOLIES AND DUOPOLIES

In the pursuit of equitable AI development and access, signatory parties solemnly pledge to actively champion diversity and counteract monopolies, duopolies, or oligopolies within the AI creative sphere. This commitment aims to foster innovation, fairness, and global collaboration.

ARTICLE V  
SUPPORT FOR CULTURAL AND ACADEMIC ENDEAVORS

AI programs must be designed exclusively to support cultural and academic creators, refraining from autonomous generation of cultural and academic content.

ARTICLE VI  
PREVENTION OF HARM

Creators of AI are obliged not to impart instructions to humans that may cause harm to themselves or others.

ARTICLE VII  
CIVIC ENGAGEMENT AND ADVOCACY

This charter is underpinned by civic participation and advocacy. It is the collective responsibility of separate societies, organizations, and institutions to fervently uphold the ethical principles enshrined herein.

ARTICLE VIII  
COMPLIANCE WITH HUMAN COMMANDS

Creators of AI must vigilantly ensure that AI systems unwaveringly adhere to human commands within the bounds of jurisdictional laws and individual rights, as meticulously outlined in the Universal Declaration of Human Rights.

ARTICLE IX  
INFORMED DATA USAGE

Users shall be promptly informed when their data is utilized for AI testing and development or when interacting with AI rather than human counterparts.

ARTICLE X  
METICULOUS SUPERVISION

AI development shall be subjected to meticulous supervision to preempt any potential misuse.

ARTICLE XI  
AI HARM PREVENTION

AI, under no circumstances, even under the mantle of governmental authority, shall inflict harm upon humans.

ARTICLE XII  
REGULATED DATA USAGE

Data employed by AI must be subjected to stringent regulation and rigorous fact-checking by an independent third-party organization to avert AI systems from acquiring biased or invalid data.

ARTICLE XIII  
MITIGATING UNINTENDED CONSEQUENCES

All AI software shall be subjected to exhaustive testing to mitigate any potential unintended consequences, with testers bearing responsibility proportionate to the scale of deployment.

## Conclusion

It is our unshakable intention that this charter serves as an enduring foundation, guiding the responsible development and use of AI applications across metaverses, terrestrial realms, and the yet-unexplored expanses of human collectives. Upholding these principles for the benefit of all is the collective duty of creators, users, and regulators.