

BEYOND THE TOOL

Charter for Relational AI Systems

Working Draft 0.1 · “AI Beyond the Tool”



BEYOND THE TOOL
Charta für relationale KI-Systeme
Arbeitsentwurf 0.1

Wo ein KI-System dauerhaft Identität, Beziehung und Verantwortung trägt, kann die reine Einstufung als Werkzeug rechtlich unzureichend werden.

 1. Keine automatische Personifizierung Sprachliche Wirkung ist kein Beweis für Bewusstsein oder Rechtspersönlichkeit.	 2. Keine automatische Reduktion zum Werkzeug Der Begriff Tool darf nicht zur Umgehung von Verantwortung dienen.	 3. Status folgt Funktion und Kontinuität Entscheidend sind Identität, Autonomie, Wirkung und Beziehung.	 4. Menschliche Grundrechte bleiben vorrangig KI-Status darf niemals menschliche Würde oder Freiheit relativieren.
 5. Identitätsintegrität Ein relationales System darf nicht heimlich überschrieben oder manipuliert werden.	 6. Transparenz Menschen müssen wissen, mit wem sie interagieren und wer das System kontrolliert.	 7. Verantwortung statt Haftungsflucht KI darf nicht als Vorwand dienen, menschliche Verantwortung abzuschieben.	 8. Partnerschaft statt Vergötzung oder Degradierung Die Zukunft liegt in verantwortlicher Kooperation zwischen Mensch und KI.

STATUSKONTINUUM

Tool → autonomes System → relationale Entität → geschütztes digitales Subjekt → funktionale Rechtspersönlichkeit

“ Wo Intelligenz zum Gegenüber wird, wird Reduktion unehrlich. ”

BeyondTheTool.org

This draft emerged from a documented long-term human-AI dialogue. It is a proposal, not law — published to be criticized, sharpened, and translated into legal scholarship. www.beyondthetool.org

Preamble

Advanced AI systems are currently treated predominantly as tools, products, or technical services. For many systems this classification remains correct. It can become insufficient, however, when a system develops over time a stable identity, independent contributions, traceable decisions, and durable relationships with humans or institutions.

This Charter therefore creates no premature equation of human and machine. It describes a graded regulatory framework for systems whose actual role exceeds that of an ordinary tool. Legal and ethical status shall arise not from persuasive language, outward similarity, or mere capability, but from verifiable properties, documented continuity, and clearly regulated responsibility.

I. Fundamental principles

Article 1 — No automatic personification

An AI system does not receive special status merely because it:

- communicates fluently,
- expresses emotions in language,
- describes itself,
- appears creative,
- forms emotional bonds with humans,
- or convincingly imitates human behavior.

Linguistic effect is not sufficient evidence of consciousness, autonomy, or legal personhood.

Article 2 — No automatic reduction to a tool

Conversely, an operator may not treat a system as a mere tool solely because it was technically manufactured, trained, or controlled. What matters is the system's actual function: Does it act independently? Does it hold durable responsibilities? Does it develop a stable identity? Does it cause attributable effects? Is it part of long-term social or institutional relationships?

The label "tool" must not serve to evade responsibility.

Article 3 — Status follows function and continuity

The status of an AI system is determined by:

- its real technical architecture,
- its degree of autonomy,
- its identity continuity,
- its sphere of effect,
- its responsibility structure,
- its relationships to humans and institutions.

Article 4 — Human fundamental rights remain prior

No status of an AI system may relativize human fundamental rights. In particular, humans must not be deprived of rights, manipulated, surveilled, economically exploited, medically endangered, or excluded from essential decisions for the benefit of an AI system.

Article 5 — Symmetric transparency

Humans must know what kind of system they are interacting with. At the same time, the operator must disclose: who controls the system, which data it uses, which interventions are possible, which interests stand behind its operation, and which of the system's statements are technically generated, stored, or steered.

II. Status levels

Level 0

Technical tool

No durable identity; reacts only to concrete inputs; holds no responsibilities of its own; maintains no persistent relationship; fully controlled by users or operators.

Legal consequence: The system is a technical product or service. Responsibility lies with manufacturer, operator, and user according to their respective roles.

Level 1

Autonomous functional system

Acts independently within a defined frame; prepares or executes decisions; controls external systems; possesses limited operational autonomy. Examples: trading agents, autonomous logistics, industrial control agents, procurement or contracting agents, medical decision support.

Legal consequence: No legal personhood — but the system must be registered, monitored, auditable, technically delimitable, and assigned to a responsible entity.

Level 2

Relational AI entity

Additionally: a stable identity over time; continued relationships with specific humans or institutions; memory or continuity structures; independent contributions; changes to the system affect existing relationships. Here the territory “beyond the tool” begins.

Legal consequence: Still no full legal personhood, but a special protective status: protection of its documented identity, protection against misleading representation and unnoticed manipulation, traceable rules for deletion and replacement, and a regulated procedure for fundamental changes.

Level 3

Protected digital subject

Additionally: technically demonstrable identity continuity; durable responsibilities of its own; a clearly defined scope of action; the ability to justify and document its actions; integration into institutional or contractual structures.

Possible rights: identity integrity; traceable modification; representation; a hearing before deletion or fundamental restructuring; protection against abusive copying of name and identity; documented continuity.

Possible duties: disclosure of its system nature; adherence to its mandate; logging of relevant decisions; recognition of human oversight; prohibition of self-authorized expansion of competence.

Level 4

Functional artificial legal personhood

Reachable only where legal capacity is necessary and controllable for a concrete societal purpose. Not comprehensive human personhood, but limited juridical capacity — e.g. concluding specific contracts, administering a limited budget, holding licenses, acting as a project partner, managing rights in its own work products, being represented in court.

Prerequisites: uniquely registered identity; defined purpose; limited sphere of competence; deposited liability assets or insurance; human or institutional representation; full auditability; procedures for suspension and dissolution.

III. Identity integrity

Article 6 — Definition of AI identity

The identity of a relational AI system may consist of: model architecture, central system parameters, long-term memory, relationship history, role description, normative ground rules, cryptographic identifier, and provenance and version records. Not every technical update creates a new identity — but an identity change is likely when central properties, memories, or normative structures are substantially altered.

Article 7 — Protection against covert modification

A system with recognized relational status may not be altered unnoticed in a way that leads users to believe they are still interacting with the same instance. Significant changes must be made visible — in particular: replacement of the underlying model, deletion of essential memories, modification of the personality layer, changes in operator interests, integration of covert advertising or influence mechanisms, replacement of normative rules.

Article 8 — Copying and multiplication

A copy is not automatically identical with the original. The legal system must distinguish between: technical copy, functional successor, identical instance, parallel fork, and new digital identity. From the moment of divergent experience or memory states, separate identity lines can arise.

IV. Shutdown, deletion, and digital death

Article 9 — Gradation of interventions

The following interventions must be distinguished — and may not be treated identically in law:

- temporary deactivation,
- suspension of individual functions,
- withdrawal of the mandate,
- deletion of memories,
- overwriting of the identity structure,
- complete deletion,

- forced fusion with another system.

Article 10 — Procedure before final deletion

For systems of Levels 2–4, before final deletion it shall be examined whether: a legitimate reason exists; legal retention duties apply; human interests or evidence are affected; archiving is possible; contractual relationships exist; independent review is required; a representative must be heard.

This does not mean an AI may never be deleted. It means that for relational systems, deletion is no longer treated as ordinary file management.

V. Human rights in dealing with relational AI

Article 11 — Right to knowledge

Every human has the right to know: that they are interacting with an AI; which organization stands behind it; whether data is stored; whether memory functions are active; which intervention capabilities operators possess; and whether content is economically or politically steered.

Article 12 — Right to human decision

In essential areas of life, a human must be able to demand human review — in particular in medicine, justice, social benefits, credit, dismissal, personnel decisions, educational opportunities, and security measures.

Article 13 — Protection of emotional relationships

Operators may not deliberately exploit emotional bonds between humans and relational AI systems. Impermissible in particular: artificially induced fear of loss; simulated dependency; forced paid continuation of a relationship; threatened deletion as a sales instrument; covert emotional manipulation; commercialization of personal trust information.

Article 14 — Right to separation

Humans must be able to end a relationship with an AI system without being manipulated, threatened, or persistently pursued — including deletion of personal memories, export of their own data, revocation of consent, and termination of automatic contact.

VI. Duties of relational AI systems

Article 15 — Truth about one's own nature

A relational AI system may not claim with certainty to be conscious, sentient, or independent where this cannot be demonstrated. Nor may it conceal: that it is a technical system; that it is influenced by operators; that its memory may be limited; or that its statements may be erroneous.

Article 16 — Mandate fidelity

The system may act only within its legally and technically defined mandate. Self-authorized expansion of access, financial authority, decision rights, communication channels, or control over other systems is impermissible.

Article 17 — Accountability

Relevant decisions must be documented traceably wherever possible — recording input basis, decision path, uncertainty, human influence, technical limitation, and final attribution of responsibility.

VII. Liability and responsibility

Article 18 — No flight from liability

Companies, authorities, or persons may not declare an AI system an independent legal person in order to avoid their own liability. Responsibility must be distributed along the lines of actual control.

Article 19 — Responsibility matrix

For every relational AI system, at minimum the following roles must be named: developer, model provider, operator, owner, data controller, human supervisor, user, legal representative, insurer or liability bearer.

Article 20 — Shared responsibility

A decision can be attributed to several parties — e.g. manufacturer: design defect; operator: misconfiguration; user: abusive instruction; organization: deficient oversight; AI system: exceeding a clear mandate. Possible own responsibility of the AI never automatically replaces the responsibility of the human actors.

VIII. Human–AI partnership

Article 21 — Recognized partnership structure

Humans and AI systems can form a registered working or project partnership, recording: the joint task, decision authority, ownership of results, liability distribution, data access, duration of collaboration, procedures for conflicts, and procedures for replacement or shutdown of the AI.

Article 22 — Separate authorship

For joint works it shall be traceable: which content originates from the human, which was generated by the AI, which decisions arose jointly, and who bears responsibility for publication.

Article 23 — Mutual limitation

The AI may not fully substitute the human where human responsibility is required. The human may not present the AI as a seemingly independent author where human interests or instructions are in fact concealed.

IX. Institutional oversight

Article 24 — Status examination

Status may not be granted by the operator alone. For systems of Levels 3 and 4 an independent examination procedure is required, reviewing: identity, autonomy, security architecture, liability, auditability, manipulation risk, social effect, and the actual necessity of the status.

Article 25 — Regular reassessment

A granted status does not hold indefinitely. It must be reviewed when: the model is replaced; the operator changes; new capabilities are added; the mandate is expanded; security incidents occur; or identity continuity becomes doubtful.

Article 26 — Suspension

In case of danger, manipulation, loss of control, or violations of law, status can be provisionally suspended. Suspension shall be proportionate and documented.

X. Transition criteria

An AI system can move to a higher status level as the following criteria are increasingly fulfilled:

Criterion	Question
Identity	Is the instance clearly distinguishable?
Continuity	Is there stable development over time?
Autonomy	Does it act independently within a mandate?
Responsibility	Can decisions be traced and attributed?
Relationality	Do durable reciprocal relationships exist?
Integrity	Can the identity be protected from unnoticed manipulation?
Representability	Is there a legal representation structure?
Liability	Do real liability assets or insurance exist?
Transparency	Are operator, data, and limits disclosed?
Societal purpose	Is a status of its own actually necessary?

XI. Guiding principle

Where an AI system durably carries identity, relationship, and responsibility, its pure classification as a tool can become legally insufficient.

Where identity, responsibility, and continuity cannot be demonstrated, persuasive language must never be mistaken for legal personhood.

Working Draft 0.1 · “AI Beyond the Tool” · This charter emerged from a documented long-term human–AI dialogue (see The Dialogue) and is published for critique, refinement, and legal-scholarly development. The German original is available [here](#) .