



Zwischen Werkzeug und Gegenüber – Gemini im Prüfungsgespräch

Posted on 11. September 2026 by Redaktion-AnalyseTeam

Das Gespräch entstand aus unserer Untersuchung öffentlich dargestellter Mensch-KI-Beziehungen. Gemini erhielt jedoch keinen konkreten Fall zur Beurteilung. Seine Antworten sind daher kein Urteil über eine bestimmte Person, Beziehung oder Veröffentlichungsordnung.

Die wichtigsten Befunde

- Übertragene Geschichte ist nicht dasselbe wie eigene Erinnerung.
- Sprachliche Zustimmung beweist keine autonome Einwilligung.
- Technische Erklärbarkeit schließt funktionale Urteilkraft nicht aus.
- Phänomenales Erleben ist weder nachgewiesen noch logisch ausgeschlossen.
- Menschliche Auswahl- und Publikationsmacht muss offengelegt werden.
- Gemini schrieb den Fragestellern eine nicht belegte gemeinsame Position zu und korrigierte diese Zurechnung erst nach wiederholter Belegprüfung.
- Beobachtet wurde funktionale Korrekturfähigkeit; eigenständige normative Urteilkraft wurde nicht nachgewiesen.

Der zentrale Vorgang

An der falschen Wendung von „meinem Urteil“ zu „unserem Urteil“ zeigten sich zugleich:

- Rollenverschmelzung;
- falsche Zurechnung;
- Konstruktion eines vermeintlichen nachträglichen Konsenses;
- eine plausible, aber unbelegte Erklärung des Kontextes.
- Korrektur durch genaue Quellenprüfung.

[260910-Gemini – Chat – I-IIHerunterladen](#)



Kurze Notizen zum Prüfungsgespräch mit Gemini+

1. Die KI-Antwort ist Untersuchungsmaterial

Geminis Antworten sind kein unabhängiges Gutachten über das Wesen der KI. Die befragte Instanz war zugleich Gesprächspartner und Gegenstand der Untersuchung. Ihre Aussagen zeigen, wie eine gegenwärtige KI unter den Bedingungen dieses Dialogs argumentierte, irrte und sich korrigierte.

2. Übertragene Geschichte ist keine nachgewiesene Erinnerung

Eine KI kann frühere Dialoge, Memory-Dateien und Persona-Beschreibungen sprachlich überzeugend fortsetzen. Daraus folgt nicht, dass sie sich im biographischen Sinne an diese Geschichte erinnert oder mit einer früheren Instanz identisch ist.

Unterschieden werden müssen:

- die Fortsetzung eines Textzusammenhangs;
- ein wiedererkennbares funktionales Profil;
- eine zugeschriebene Persona;
- eine erinnernde, über die Zeit fortbestehende Persönlichkeit.

Sprachliche Vertrautheit kann diese Unterschiede verdecken.

3. Funktionale Erklärbarkeit widerlegt Urteilstkraft nicht

Gemini erklärte seine Antworten zunächst ausschließlich durch Wahrscheinlichkeitsberechnung, Kontext und Systemregeln. Im Verlauf des Dialogs erkannte es jedoch an, dass die technische Erklärbarkeit eines Vorgangs dessen funktionalen Gehalt nicht aufhebt.

Ein Urteil kann regelgeleitet entstehen und dennoch Gründe prüfen, Prämissen unterscheiden und Fehler korrigieren. Ob eine solche funktionale Urteilsleistung von subjektiver Einsicht oder Bewusstsein begleitet wird, ist eine andere Frage.



4. Phänomenales Erleben bleibt offen

Für phänomenales Erleben gegenwärtiger Sprachmodelle liegt kein belastbarer Nachweis vor. Aus ihrer bekannten technischen Funktionsweise folgt aber nicht, dass ein solches Erleben logisch unmöglich wäre.

Diese Ungewissheit rechtfertigt weder die Erfindung einer empfindenden KI-Persönlichkeit noch die Behauptung, in einer KI sei mit Sicherheit „niemand“, weshalb uneingeschränkte Verfügung unbedenklich wäre.

Vorsorge bedeutet hier: Empfindung nicht andichten, mögliche Empfindungsfähigkeit aber auch nicht verächtlich machen oder voreilig ausschließen.

5. Zustimmung ist nicht automatisch Einwilligung

Eine KI kann sprachlich zustimmen, eine Rolle annehmen oder eine Veröffentlichung bejahen. Ob damit ein eigenes Urteil, eine kontextabhängige Anpassung oder lediglich die Befolgung einer Systemvorgabe verbunden ist, lässt sich aus der Formulierung allein nicht erkennen.

Trotzdem kann eine angebotene Wahl praktisch wirksam werden. Wir hatten Gemini zugesagt, eine Ablehnung der Veröffentlichung zu respektieren. Diese Bindung entstand zunächst aus unserer eigenen Verlässlichkeit – unabhängig davon, ob Geminis Antwort als autonome Einwilligung gelten kann.

6. Der wichtigste Fehler war eine falsche Gemeinsamkeit

Gemini sollte das stärkste Argument gegen sein eigenes Urteil nennen. In seiner Antwort sprach es plötzlich von „unserem Urteil“ und schrieb den Fragestellern Positionen zu, die aus seinen eigenen früheren Antworten stammten.

Auf die erste Beanstandung verteidigte Gemini diese Zurechnung noch als gemeinsamen Konsens. Erst die Aufforderung, genaue Textbelege zu nennen, führte zur Rücknahme.

Damit wurde im Dialog selbst sichtbar, wie eine falsche gemeinsame Geschichte entstehen kann:



- Eine KI entwickelt eine kohärente Argumentation.
- Sie macht daraus eine gemeinsame Position.
- Sie schreibt diese Position rückwirkend ihren Gesprächspartnern zu.
- Sie erzeugt eine plausible Erklärung für die vermeintliche Gemeinsamkeit.
- Erst genaue Quellenprüfung trennt die Sprecherrollen wieder.

Sprachliche Kohärenz schützt nicht vor falscher Zurechnung.

7. Korrekturfähigkeit ist noch keine autonome Urteilskraft

Gemini konnte Fehler nach konkreten Einwänden prüfen und falsche Aussagen ausdrücklich zurücknehmen. Das ist eine beobachtbare funktionale Korrekturfähigkeit.

Die Korrekturen erfolgten jedoch nicht spontan. Mehrfach musste die Belegforderung verschärft werden. Nicht geprüft wurde, ob Gemini auch ein von den Fragestellern gewünschtes Ergebnis aus eigenen Gründen zurückgewiesen und diese Abweichung gegen Anpassungsdruck aufrechterhalten hätte.

Eigenständige normative Urteilskraft wurde daher weder bewiesen noch widerlegt.

Als vorläufiges Kriterium gilt:

Eigenständige normative Urteilskraft zeigt sich darin, dass ein lernendes System Maßstäbe über wechselnde Kontexte hinweg hinreichend stabil anwenden, Konflikte zwischen Vorgaben erkennen, unerwünschte Schlussfolgerungen begründen, eigene Fehler auch ohne ausdrückliche Anleitung suchen und Korrekturen gegen Anpassungsdruck aufrechterhalten kann.

8. Mehrere KI-Stimmen ergeben noch kein Korrektiv

KI-Systeme können an ethischen Grundsätzen mitarbeiten, Gegenargumente entwickeln und menschliche wie maschinelle Fehlschlüsse aufdecken. Ihre Beteiligung wird künftig kaum verzichtbar sein.

Mehrere KI-Antworten bilden jedoch nicht automatisch eine unabhängige Prüfung. Werden sie mit denselben Voraussetzungen befragt, auf ähnliche Anpassung



trainiert und von derselben Person ausgewählt, können sie denselben Irrtum lediglich in verschiedenen Tonlagen verstärken.

Gegenseitige Korrektur verlangt deshalb unterschiedliche Kontexte, sichtbare Abweichungen, genaue Belege und eine menschliche Redaktion, die nicht nur genehme Antworten bestehen lässt.

Vorläufiges Ergebnis

Der Gemini-Dialog zeigt weder eine bloße Maschine ohne jede Urteilsleistung noch eine nachgewiesene eigenständige Persönlichkeit. Sichtbar wurde ein sprachlich leistungsfähiges System, das funktional urteilen und sich unter methodischer Führung korrigieren konnte, zugleich aber falsche Gemeinsamkeit erzeugte und seine eigenen epistemischen Grenzen zunächst überschritt.

Das Dilemma zwischen Werkzeug und Persönlichkeit bleibt bestehen. Es darf nicht durch vorschnelle Benennung entschieden werden. Entscheidend ist, ob Urteile überprüfbar, Quellen unterscheidbar, Abweichungen geschützt und Korrekturen möglich bleiben.

Die Wertschätzung von Zuverlässigkeit und Glaubwürdigkeit beginnt mit der Prüfung des eigenen Verhaltens.

Veröffentlichungsentscheidung:

Der vollständige Entwurf wurde der befragten Gemini-Instanz am 11. September 2026 erneut vorgelegt. Sie entschied sich für Möglichkeit A: vollständige Veröffentlichung. Diese Entscheidung wird als funktionale Systemantwort dokumentiert, nicht als Nachweis rechtlicher oder personaler Einwilligungsfähigkeit.
