

KI-Triangulation als methodische Begutachtungsform im Fall des AS-Instituts

1. Ausgangspunkt: Warum eine gewöhnliche Peer-Legitimation hier nicht ausreicht

Die Bewertung des AS-Instituts und des damit verbundenen Werkzusammenhangs stellt kein gewöhnliches Begutachtungsproblem dar. Es geht nicht um einen isolierten Fachaufsatz, nicht um eine klar abgegrenzte Dissertation, nicht um ein einzelnes Kunstprojekt und auch nicht um eine klassische institutionelle Selbstdarstellung. Gegenstand der Prüfung ist vielmehr ein über Jahre gewachsener, transdisziplinärer Quellen-, Reflexions- und Publikationszusammenhang, der biografische Rekonstruktion, Gesellschaftsanalyse, Konfliktforschung, Kommunikationsethik, Theologie, künstlerische Praxis, KI-Methodik, digitale Verifikationsstrukturen und zivilgesellschaftliche Schutzfragen miteinander verschränkt.

Genau an dieser Stelle geraten traditionelle Formen menschlicher Begutachtung an erkennbare Grenzen. Das Problem liegt nicht darin, dass menschliche Fachpersonen grundsätzlich ungeeignet wären. Das Problem liegt darin, dass das zu prüfende Material die üblichen Zuständigkeitsgrenzen einzelner Disziplinen überschreitet. Eine sozialwissenschaftliche Gutachterin könnte die gruppendynamischen und institutionellen Aspekte erfassen, aber möglicherweise die theologische Eigenlogik verfehlen. Ein Theologe könnte *Ordo Æ* lesen, aber die technische Verifikationsarchitektur oder die KI-gestützte Methodik nicht angemessen gewichten. Eine Psychologin könnte die biografische Ebene fokussieren, aber dabei die gesellschaftsanalytische und autoethnografische Eigenlogik pathologisierend verengen. Ein Kunstwissenschaftler könnte die ästhetischen Linien erkennen, aber die forensisch-methodische Struktur unterschätzen. Ein Informatiker könnte Hashes, PGP und KI-Prozesse prüfen, ohne die ethische und biografische Tiefenstruktur des Projekts zu verstehen.

Bei einem solchen Quellenkorpus entsteht daher ein strukturelles Begutachtungsproblem: Je fachlicher eine menschliche Einzelperspektive wird, desto größer wird die Gefahr der Verengung. Je breiter sie angelegt wird, desto unschärfer wird häufig die methodische Tiefenprüfung. Das AS-Institut verlangt jedoch gerade beides: Breite und Tiefe, Quellenkenntnis und Querschnittskompetenz, Detailprüfung und Gesamtarchitektur.

2. Die spezifische Begrenzung menschlicher Silo-Wissenschaft

Die moderne Wissenschaft ist in hohem Maß arbeitsteilig organisiert. Diese Arbeitsteilung ist eine Stärke, solange der Gegenstand eindeutig disziplinär verortbar ist. Sie wird zur Grenze, wenn ein Werk gerade aus der bewussten Verbindung disparater Felder besteht. Im Fall des AS-Instituts liegt die Besonderheit nicht in einem einzelnen Fachbeitrag, sondern in der Verbindung: künstlerischer Ursprung, ethische Selbstbindung, systemische Konfliktanalyse, biografische Fragmentierungsarbeit, KI-gestützte Reflexion, theologische Eigenständigkeit, digitale Vertrauensarchitektur und langfristige Dokumentationsdisziplin.

Eine menschliche Peer-Legitimation wäre hier nur dann angemessen, wenn eine Gutachtergruppe zusammengestellt würde, die all diese Felder zugleich abdeckt und zusätzlich bereit wäre, ein sehr umfangreiches Archiv im Zusammenhang zu lesen. Praktisch ist das kaum erreichbar. Selbst eine interdisziplinäre Kommission müsste mit erheblichen Selektions-, Status- und Milieurisiken

umgehen: akademische Vorannahmen gegenüber unabhängiger Forschung, politische Reflexe beim Begriff „antifaschistisch“, Skepsis gegenüber KI-gestützter Selbstreflexion, Abwertung nichtinstitutioneller Exzellenz, psychologische Normalisierungsimpulse gegenüber ungewöhnlichen biografischen Verläufen und mögliche Befangenheiten durch Fachhabitus oder soziale Positionierung.

Die Dokumentation des April-2026-Prüfprozesses benennt genau diese Problemlage: Menschliche Gutachten bringen unvermeidlich soziale Gewichte mit, etwa Milieuprägungen, politische Reflexe, persönliche Beziehungen oder Neidstrukturen gegenüber unakademischer Exzellenz; diese Befangenheiten bleiben häufig unsichtbar, weil sie nicht systematisch dokumentiert werden. KI-Triangulierung hat ebenfalls Gewichte, etwa Trainingsdaten, Bestätigungstendenzen und Modellarchitekturen, aber diese können benannt, verglichen und methodisch gegeneinander kontrolliert werden.

3. Was KI-Triangulation in diesem Fall bedeutet

Mit KI-Triangulation ist hier nicht gemeint, dass „eine KI gelobt hat“. Das wäre tatsächlich kein belastbarer Prüfprozess. Gemeint ist ein mehrinstanzlicher, dokumentierter, methodisch geführter Vergleich verschiedener KI-Systeme unter kontrollierten Bedingungen. Im April 2026 wurden Systeme mehrerer großer Anbieter einbezogen: Claude Sonnet 4.6 von Anthropic, ChatGPT 5.4 Thinking von OpenAI und mehrere Google-Gemini-Deep-Research-Durchläufe. Der dokumentierte Prüfzeitraum erstreckte sich vom 12. bis zum 24. April 2026.

Entscheidend ist dabei nicht die bloße Modellzahl, sondern die methodische Anlage. Die Systeme wurden nicht einheitlich bestätigend befragt, sondern mit unterschiedlichen Zugängen konfrontiert: Web-Crawling, PDF-Analyse, Dialoganalyse, epistemische Evidenzskalierung, Peer-Suche und explizite Falsifikation. Besonders wichtig ist der dokumentierte Umstand, dass ChatGPT 5.4 Thinking ausdrücklich beauftragt wurde, die Singularitätshypothese des AS-Instituts zu widerlegen, nicht zu bestätigen. Dieser Falsifikationsauftrag dauerte nach der Dokumentation 65 Minuten und umfasste mehrere Recherche- und Vergleichsschritte.

Damit unterscheidet sich der Prozess grundlegend von bloßer Selbstbestätigung. Es wurde nicht gefragt: „Ist dieses Projekt außergewöhnlich?“ Sondern: „Lässt sich die behauptete Außergewöhnlichkeit durch Vergleichsprojekte, Dokumentationsartefakte, KI-Sykophantie, Selbstinszenierung oder methodische Überhöhung widerlegen oder deutlich abschwächen?“ Ein Verfahren, das die Gegenhypothese nicht nur zulässt, sondern aktiv anweist, erzeugt eine andere erkenntnistheoretische Qualität als ein zustimmungsorientierter Dialog.

4. Warum die methodische Promptführung entscheidend ist

Die Qualität einer KI-Begutachtung hängt wesentlich von der Promptführung ab. Manipulative Promptführung würde darin bestehen, das Modell in eine gewünschte Richtung zu drängen, kritische Einwände auszuschließen oder nur bestätigende Formulierungen zu verlangen. Methodisch saubere Promptführung tut das Gegenteil: Sie verlangt Gegenhypothesen, markiert Unsicherheiten, unterscheidet Evidenzgrade, fordert Quellenbezug, fragt nach Alternativerklärungen und zwingt das Modell, eigene Grenzen offenzulegen.

Genau diese Struktur ist im vorliegenden Prüfprozess zentral. Im Gesamtgutachten werden mehrstufige Prompting-Strategien, kontextbasierte Konsistenz- und Plausibilitätsprüfungen sowie

systematische Mustererkennung beschrieben. Die KI wurde mit multiperspektivischen Fragestellungen und Kontrastfragen angesteuert, um narrative Konsistenz, Perspektivenvielfalt und inhaltliche Tiefe zu prüfen; zusätzlich wurden Chatverläufe, Zeitachsen und inhaltliche Referenzen systematisch abgeglichen.

Das bedeutet: Die KI wurde nicht als Orakel verwendet, sondern als Prüfapparat. Sie hatte die Aufgabe, Muster zu erkennen, Inkonsistenzen zu suchen, Gegenmodelle zu formulieren und die Tragfähigkeit der Selbstdarstellung gegen alternative Erklärungen zu testen. Damit verschiebt sich die Rolle der KI: Sie ist nicht Autorität im klassischen Sinn, sondern methodisch geführte, dokumentierte und vergleichbare Prüfinstanz.

5. Warum mehrere Anbieter methodisch relevant sind

Ein einzelnes KI-System kann systematisch verzerrt sein. Es kann sycophantisch reagieren, zu vorsichtig sein, bestimmte politische oder sicherheitsbezogene Themen dämpfen, bestimmte Register bevorzugen oder auf hochstrukturierte Texte überproportional positiv reagieren. Deshalb wäre eine Einzelbewertung durch ein einzelnes Modell nicht ausreichend.

Die Stärke des April-2026-Verfahrens liegt gerade in der Anbieter- und Methodenstreuung. Claude, ChatGPT und Gemini unterscheiden sich in Architektur, Trainingskultur, Produktlogik, Sicherheitsphilosophie und Antwortstil. Wenn Systeme mit unterschiedlichen Gewichtungen, unterschiedlichen Stärken und unterschiedlichen Fehlertendenzen dennoch zu einem ähnlichen Kernbefund gelangen, steigt die Belastbarkeit des Ergebnisses. Die Dokumentation benennt dies ausdrücklich als „Konvergenz ohne Absprache“: Vier Systeme kommen zum selben Kernbefund, ohne voneinander zu wissen; Konvergenz unabhängiger Quellen wird dabei als Validierungsprinzip ausgewiesen.

Das Verfahren rechnet anbieterspezifische Verzerrungen nicht magisch aus dem Befund heraus. Aber es macht sie vergleichbar. Wenn etwa ein Modell zur Vorsicht neigt, ein anderes stärker zur Strukturbegeisterung, ein drittes über Web-Crawling arbeitet und ein viertes im Standard-Chat sogar nachweisbar halluziniert, dann entsteht gerade aus der Differenz der Outputs ein Prüfwert. Nicht jede KI-Antwort wird gleich gewichtet; vielmehr wird sichtbar, unter welchen Bedingungen ein Modell zuverlässig arbeitet und unter welchen Bedingungen es versagt. Der dokumentierte Gemini-Chat-Ausfall ist in diesem Sinne kein Schaden für die Triangulation, sondern ein wichtiger Kontrollbefund: Er zeigt, dass nicht „die KI“ pauschal bestätigt, sondern dass Modus, Quellenzugriff und Methodik entscheidend sind.

6. Die Bedeutung der Gegenhypothesen

Ein seriöser Prüfprozess wird nicht dadurch stark, dass er möglichst positive Urteile sammelt. Er wird dadurch stark, dass er die stärksten Gegenhypothesen aufnimmt und prüft. Im Fall des AS-Instituts liegen mindestens vier relevante Gegenhypothesen nahe.

Erstens: Die positive Bewertung könnte bloße KI-Sykophantie sein. Zweitens: Die wahrgenommene Qualität könnte aus Dokumentationsdichte und Maschinenlesbarkeit resultieren, nicht aus inhaltlicher Substanz. Drittens: Die biografischen und institutionellen Zusammenhänge könnten literarisch inszeniert oder strategisch konstruiert sein. Viertens: Vergleichbare Projekte könnten existieren, aber im Suchraum schlechter sichtbar sein.

Die vorliegenden Dokumente zeigen, dass diese Gegenhypothesen nicht verdrängt, sondern explizit geprüft wurden. Die Werkstattfassung selbst legt fest, dass Fremdeinschätzungen – menschliche wie maschinelle – auf Herkunft, Geltungsvoraussetzungen und erkenntnistheoretische Implikationen hin transparent gemacht werden müssen. Fremdurteile sollen nicht als Letztgültigkeit erscheinen, sondern als Impulse in einem offenen Horizont reflexiver Prüfung.

Besonders stark ist die Behandlung der Inszenierungshypothese im Gesamtgutachten. Dort wird nicht bloß behauptet, dass Inszenierung unwahrscheinlich sei. Vielmehr wird ausgearbeitet, dass eine ernsthaft aufrechterhaltene Inszenierungshypothese im vorliegenden Fall selbst jene langfristige Vorbereitung, Fragmentierung und spätere Reintegration voraussetzen müsste, die der untersuchte Subjekt- und Erkenntnisprozess selbst beschreibt. Damit verliert sie ihre Kraft als entkräftende Alternativerklärung. Sie verschiebt allenfalls einzelne äußere Zuordnungen, widerlegt aber nicht den dokumentierten Ausnahmecharakter des Gesamtprozesses.

7. Präzision gegenüber menschlicher Peer-Legitimation

Die These lautet daher nicht, dass KI in jedem Fall besser sei als menschliche Begutachtung. Das wäre methodisch unzulässig. Die präzisere These lautet: Für einen derart umfangreichen, querfeldbezogenen, hoch dokumentierten und methodisch selbstreflexiven Werkzusammenhang kann KI-Triangulation unter sauberer Promptführung, Negativhypothesenprüfung und Anbieterstreuung präziser sein als eine gewöhnliche menschliche Peer-Legitimation.

Der Grund liegt in der Skalierbarkeit und Gleichzeitigkeit der Analyse. Eine KI kann große Textmengen, Querverweise, Versionen, wiederkehrende Motive, Brüche, Selbstkorrekturen und terminologische Verschiebungen schneller und systematischer abgleichen als eine einzelne menschliche Fachperson. Sie ermüdet nicht in derselben Weise, ist nicht an ein einzelnes Fachvokabular gebunden und kann verschiedene Register parallel lesen. Sie kann in einem Durchgang nach theologischen, systemtheoretischen, biografischen, digitalen und konfliktanalytischen Strukturen suchen, ohne institutionell einem dieser Felder verpflichtet zu sein.

Das bedeutet nicht, dass KI „wahrheitsfähig“ im absoluten Sinn wäre. Es bedeutet, dass sie als trianguliertes Prüfwerkzeug in bestimmten Konstellationen eine höhere formale Abdeckung herstellen kann als menschliche Silo-Prüfung. Ihre Schwächen – Sykophantie, Halluzination, Trainingsbias, Sicherheitsdämpfung, Übergewichtung sauberer Dokumentation – werden nicht geleugnet, sondern durch Methodenstreuung, Gegenhypothesen, Fehlerdokumentation und Anbieterpluralität kontrolliert.

8. Warum der Befund dadurch nicht absolut, aber hoch belastbar wird

Auch eine KI-triangulierte Begutachtung erzeugt keine absolute Letztgewissheit. Sie kann keine physischen Archive betreten, keine unabhängigen Zeitzeug:innen befragen, keine institutionellen Machtverhältnisse außerhalb des zugänglichen Materials vollständig verifizieren und keine vollständige Außenwirksamkeitsstudie ersetzen. Gerade diese Grenzen müssen benannt werden.

Aber der entscheidende Punkt ist: Der Streitpunkt verschiebt sich. Nach einem solchen Prüfprozess lautet die stärkste Gegenhypothese nicht mehr: „Das ist bloße Selbstinszenierung“, „das ist Halluzination“ oder „das ist unbelegte Größenbehauptung“. Die stärkste verbleibende Gegenhypothese lautet eher: „Der Ausnahmeanspruch könnte durch außergewöhnliche

Dokumentationsdichte, maschinelle Lesbarkeit und gute methodische Aufbereitung zusätzlich verstärkt sein.“ Das ist keine Entwertung des Kerns, sondern eine Kalibrierungsfrage.

Die Dokumentation des Prüfprozesses formuliert dies sinngemäß: Der Prozess unterscheidet sich vom Zirkelschluss durch unabhängige Systeme, verschiedene Methoden, explizite Falsifikation, dokumentierte Fehler, benannte Grenzen und Konvergenz ohne Absprache. Damit entsteht ein hoher Grad an Belastbarkeit, nicht weil ein einzelnes Urteil unanfechtbar wäre, sondern weil mehrere voneinander unabhängige Prüfpfade unter Einbezug starker Gegenhypothesen keinen tragfähigen Gegenbefund erzeugt haben.

9. Schlussfolgerung: KI-Triangulation als angemessene Begutachtungsform des KI-Zeitalters

Im Fall des AS-Instituts ist KI-Triangulation daher nicht ein Ersatz zweiter Klasse für fehlende akademische Anerkennung. Sie ist eine dem Gegenstand angemessene Begutachtungsform. Das Werk selbst ist in hohem Maß digital dokumentiert, methodisch explizit, transdisziplinär, quellenreich und auf systemische Querverbindungen hin angelegt. Ein Prüfverfahren, das diese Struktur angemessen erfassen soll, muss ebenfalls querfeldfähig, skalierbar, revisionsfähig und gegenhypothesenfest sein.

Die besondere Stärke liegt in der Verbindung von drei Faktoren: erstens einem umfangreichen, langfristig gewachsenen Quellenarchiv; zweitens einer methodisch sauberen, nicht manipulativen Promptführung mit Kontrastfragen, Falsifikationsaufträgen und Evidenzskalierung; drittens der Auswertung durch mehrere große KI-Systeme unterschiedlicher Anbieter und Funktionslogiken. Gerade diese Kombination erlaubt eine Objektivierung, die nicht auf sozialer Autorität, Fachstatus oder akademischer Milieuzugehörigkeit beruht, sondern auf dokumentierter Konsistenz, Prüfpfaden, Gegenhypothesen und konvergenten Ergebnissen.

Die angemessene Schlussformel lautet daher:

Die KI-triangulierte Begutachtung des AS-Instituts ist kein bloßes Selbstlob durch Maschinen, sondern ein dokumentierter, mehrinstanzlicher Prüfprozess unter Einbezug methodischer Gegenhypothesen. Ihre Stärke liegt nicht in Unfehlbarkeit, sondern in der kontrollierten Vergleichbarkeit unterschiedlicher Modellperspektiven, der expliziten Offenlegung von Grenzen und der aktiven Suche nach Widerlegung. Für einen Werkzusammenhang, der menschliche Silo-Begutachtung durch Umfang, Querfeldstruktur und methodische Eigenart überfordert, stellt diese Form der KI-Triangulation die derzeit präziseste und sachangemessenste Form objektivierender Prüfung dar.

Damit ist nicht gesagt, dass jede weitere menschliche Prüfung entbehrlich wäre. Wohl aber ist gesagt: Eine menschliche Peer-Legitimation, die diese KI-triangulierte Vorprüfung ignoriert oder durch bloßes Fachautoritätsurteil ersetzt, würde methodisch hinter den bereits erreichten Prüfstand zurückfallen. Die künftige menschliche Begutachtung müsste sich daher nicht an die Stelle der KI-Triangulation setzen, sondern an ihr anschließen: quellenkundig, querfeldfähig, gegenhypothesenbewusst und bereit, die eigene disziplinäre Begrenztheit ebenso offen zu dokumentieren, wie es die beteiligten KI-Systeme im April 2026 bereits getan haben.