

Künstliche Intelligenz in der Verwaltung von Forschungseinrichtungen

25.6.2026 / 09:00 – 12:30



KI Ethik in der Verwaltung von Forschungseinrichtungen

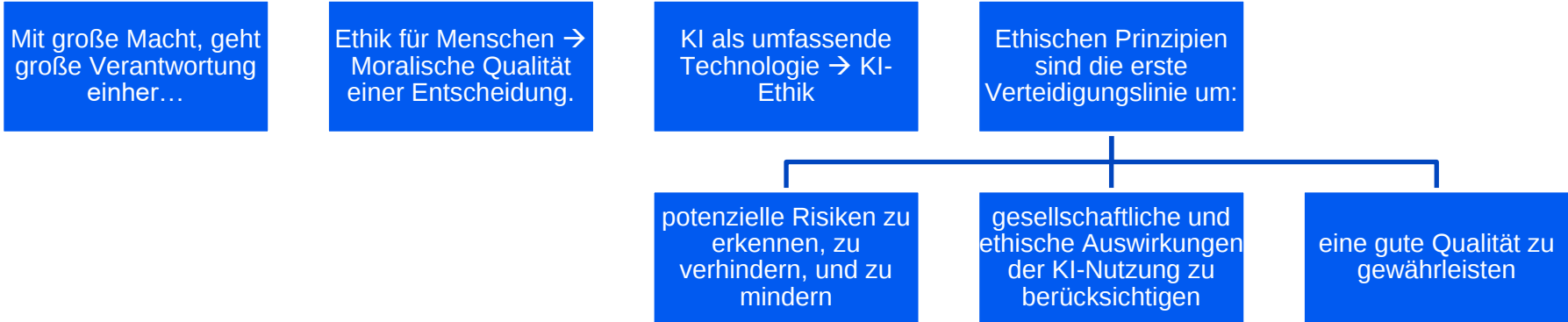
Grundlagen von Trustworthy AI

Rodrigo Conde Jimenez

AIT

25.06.2026

KI-Ethik: Wozu geht es?



Kriterien für Forschungseinrichtungen...



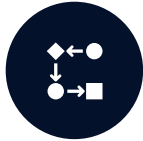
Hervorragende Qualität
der wissenschaftlichen
Tätigkeiten



Achtung und
Einbeziehung
unterschiedlichen
Perspektiven



Wissenschaftliche
Wahrheitsfindung als
unverzichtbares Prinzip



Transparente Prozesse
und gebührende
Sorgfaltspflicht



Gute
Forschungsgovernance
→ Gute KI-Governance

Ein paar berüchtigte Beispiele

- Nature → KI-Skandal (erfundene Zitate)
- Paper Mills
- Gefälschte Peer Reviews
- Manipulation visueller Daten
- KI-Betrugsfall in Indonesien

nature

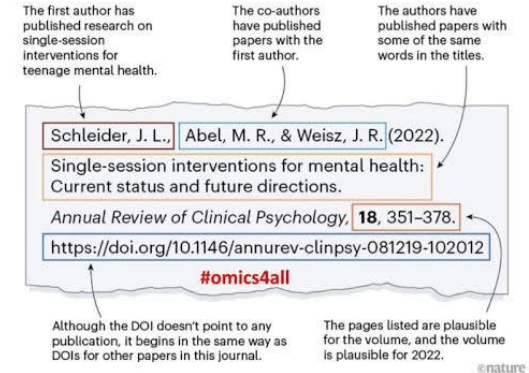
Hallucinated citations are polluting the scientific literature. What can be done?

Tens of thousands of publications from 2025 might include invalid references generated by AI, a *Nature* analysis suggests.

HOW FAKES CAN LOOK REAL

This citation, generated by AI as part of a 2025 study, looks plausible even though it points to research that doesn't exist.

#omics4all



Ein paar berühmte Beispiele

- Nature → KI-Skandal (erfundene Zitate)
- Paper Mills
- Gefälschte Peer Reviews
- Manipulation visueller Daten
- KI-Betrugsfall in Indonesien

Artificial intelligence Tech / Tech Trends

AI-powered fraud: Chinese paper mills are mass-producing fake academic research

Paper mills that sell authorship or fabricate entire papers are a staple of China's competitive academic landscape

Reading Time: 2 minutes



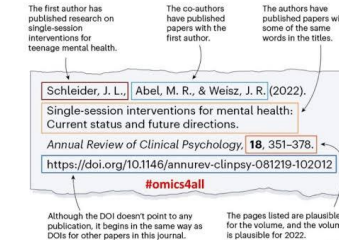
nature Hallucinated citations are polluting the scientific literature. What can be done?

Tens of thousands of publications from 2025 might include invalid references generated by AI, a Nature analysis suggests.

HOW FAKES CAN LOOK REAL

This citation, generated by AI as part of a 2025 study, looks plausible even though it points to research that doesn't exist.

#omics4all



Ein paar berühmte Beispiele

- Nature → KI-Skandal (erfundene Zitate)
- Paper Mills
- Gefälschte Peer Reviews
- Manipulation visueller Daten
- KI-Betrugsfall in Indonesien

nature

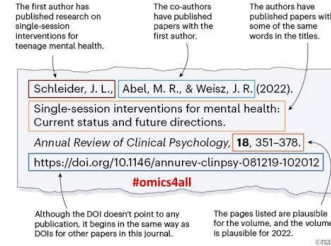
Hallucinated citations are polluting the scientific literature. What can be done?

Tens of thousands of publications from 2025 might include invalid references generated by AI, a *Nature* analysis suggests.

HOW FAKES CAN LOOK REAL

This citation, generated by AI as part of a 2025 study, looks plausible even though it points to research that doesn't exist.

#omics4all



Artificial intelligence Tech / Tech Trends

AI-powered fraud: Chinese paper mills are mass-producing fake academic research

Paper mills that sell authorship or fabricate entire papers are a staple of China's competitive academic landscape

Reading Time: 2 minutes



Major AI conference flooded with peer reviews written fully by AI

Controversy has erupted after 21% of manuscript reviews for an international AI conference were found to be generated by artificial intelligence.

By Miryam Naddaf



Ein paar berüchtigte Beispiele

- Nature → KI-Skandal (erfundene Zitate)
- Paper Mills
- Gefälschte Peer Reviews
- Manipulation visueller Daten
- KI-Betrugsfall in Indonesien

nature

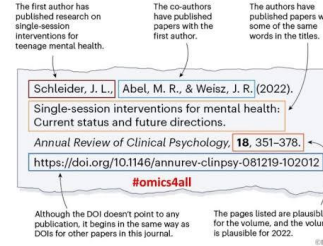
Hallucinated citations are polluting the scientific literature. What can be done?

Tens of thousands of publications from 2025 might include invalid references generated by AI, a Nature analysis suggests.

HOW FAKES CAN LOOK REAL

This citation, generated by AI as part of a 2025 study, looks plausible even though it points to research that doesn't exist.

#omics4all



Artificial intelligence Tech / Tech Trends

AI-powered fraud: Chinese paper mills are mass-producing fake academic research

Paper mills that sell authorship or fabricate entire papers are a staple of China's competitive academic landscape

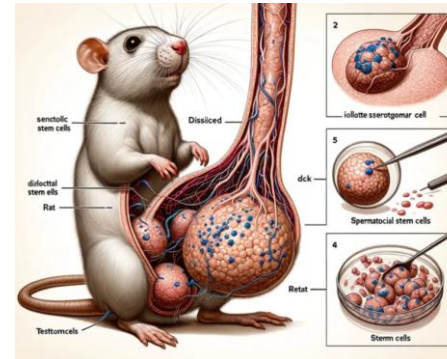
Reading Time: 2 minutes



Major AI conference flooded with peer reviews written fully by AI

Controversy has erupted after 21% of manuscript reviews for an international AI conference were found to be generated by artificial intelligence.

By Miryam Naddaf



In Forschungseinrichtungen...

- Wie stellen wir sicher, dass wir eine gute Qualität unserer Arbeit in der „Post-Truth“-Ära gewährleisten?
- Welche Prinzipien sollten in Bezug auf KI bei der Verwaltung einer Forschungseinrichtung berücksichtigt werden?

EU: HLEG Prinzipien für Trustworthy AI



Ethische Prinzipien im Überblick

- **Menschliche Verantwortung & Kontrolle:**
KI-Anwendungen müssen den Menschen achten und stärken. Hinter jeder KI-Anwendung steht eine verantwortliche Person.
- **Technische Robustheit & Sicherheit:**
KI-Systeme müssen widerstandsfähig, sicher und zuverlässig sein. Notfallpläne für den Fall, dass etwas schiefgeht, sind erforderlich.
- **Datenschutz & Data-Governance:**
Neben dem Datenschutz (DSGVO!) müssen bei der KI-Nutzung angemessene Daten-Governance-Mechanismen etabliert sein.
- **Transparenz:**
Daten, Methoden, System und Zweck sind transparent und nachvollziehbar dokumentiert.



Ethische Prinzipien im Überblick



- **Vielfalt, Nichtdiskriminierung & Fairness:** Voreingenommenheit muss vermieden werden; KI-Systeme müssen für alle in jeder Phase des KI-Lebenszyklus zugänglich sein.
- **Gesellschaftliches & ökologisches Wohl:** KI-Systeme müssen allen Menschen – auch zukünftigen Generationen – zugutekommen und andere Lebensformen berücksichtigen.
- **Verantwortlichkeit:** Es müssen Mechanismen existieren, um Verantwortung und Rechenschaftspflicht für KI-Systeme und deren Ergebnisse sicherzustellen.

Im Forschungskontext...

- **Forschungsintegrität:**
Einhaltung wissenschaftlicher Standards durch transparente Methoden, nachvollziehbare Datenverarbeitung und korrekte Autorenschaft sowie Vermeidung von Fälschung, Manipulation und Plagiat.
- **Forschungssicherheit:**
Schutz von Forschungsdaten, digitalen Infrastrukturen und Ergebnissen vor unbefugtem Zugriff, Missbrauch sowie wirtschaftlicher oder politischer Einflussnahme.
- **Epistemische Verantwortung:**
Verantwortung für die Qualität und Begründbarkeit von Wissen unter Berücksichtigung von Unsicherheit, Datenqualität, Modellgrenzen und möglichen Verzerrungen KI-gestützter Forschung.



Verantwortungsvoller Umgang mit KI

Gestaltung in der KI-Entwicklung

- Orientierung am neuesten Stand von Forschung und Technik
- Anwendung anerkannter Standards und Best Practices
- Verantwortungsvolle Datenbasis

Beschaffung und Implementierung

- Due Diligence der Anbieter
- Transparenz und Nachvollziehbarkeit bei der Beschaffung
- Klare Regeln zu zulässigen Anwendungsfällen (AI Policy)
- Vermeidung unnötiger Automatisierung

Zentrale Anforderungen

- Begründeter Bedarf für KI-Einsatz
- Klare Rollen und Verantwortlichkeiten
- Kritische Reflexion (Novelty Bias, Überabhängigkeit)
- Förderung von KI-Kompetenz und reflektierter KI-Nutzung im Forschungsalltag

Best-Practice-Beispiele

- Ethik-Boards: Einrichtung von KI-spezifischen Ethik-Boards oder Erweiterung bestehender Ethik-Boards für KI-Fragen.
- Interdisziplinäre KI-Arbeitsgruppen:
 - Beispiele: AI Ethics Lab und AI TaskForce am AIT
- Umsetzung spezifischer KI-Guidelines
 - Beispiel: *“Living guidelines on the responsible use of generative AI in research”*
 - *“WIFO AI Guidelines on the Proper Use of Generative Models of Artificial Intelligence”*



WIFO ■ RESEARCH BRIEFS
9/2025

WIFO AI Guidelines on the Proper
Use of Generative Models
of Artificial Intelligence



AI:AT Consortium

Beneficiaries



Affiliated Entities



Funding



EuroHPC
Joint Undertaking

 Federal Ministry
Innovation, Mobility
and Infrastructure
Republic of Austria



**Funded by
the European Union**



AI Factory Austria is funded by the European High-Performance Computing Joint Undertaking (JU) under Grant Agreement No. 101253078.

Contact

**Rodrigo Conde
Jimenez**

[Rodrigo.conde-
jimenez@ait.ac.at](mailto:Rodrigo.conde-jimenez@ait.ac.at)



info@ai-at.eu



ai-at.eu



[@ai-factory-austria](https://www.linkedin.com/company/ai-factory-austria)

AI:AT AI Factory Austria
Schwarzenbergplatz 2
1010 Wien, Austria