

**AI**  
week

# Future Talk: Prüfungen im KI-Zeitalter

Dr. Tobias Halbherr  
ETH-Zürich

Alexander Schulz  
Universität Potsdam

VERANSTALTET VON:



IM RAHMEN EINES PROJEKTES VON:



GEFÖRDERT VON:



# Future Talk: Prüfungen im KI-Zeitalter

# Referenten

## Dr. Tobias Halbherr

- Leiter der Gruppe „Digital Education Services“ der Sektion EduIT an der ETH-Zürich
- Seit 2014 Leiter der SIG „E-Assessment“ von eduhub.ch (SWITCH)
- Seit 2012 verantwortlich für technisch-didaktische und organisatorische Architektur, Betrieb sowie Weiterentwicklung der digitalen Prüfungen an der ETH.
- Mitglied des Product-Owner-Boards des Safe Exam Browsers (SEB)
- Studium lic.phil. kognitive Psychologie und Informatik, Universität Zürich
- Promotion in Lernwissenschaften mit Dissertation „Resource-Rich Assessment in Higher Education“, ETH Zürich
- CAS in Betriebswirtschaft – Business Innovation, Universität Bern & University of Rochester, NY

## Alexander Schulz

- Seit 2025 Direktor des Zentrums für IT und Medienmanagement der Universität Potsdam
- 2024 Aufbau des Schwerpunktbereichs „KI@FU – Lehren, Lernen und Prüfen mit KI“
- 2018 bis 2024 Leitung der BMBF-KI-Verbundprojekte der FU-Berlin: „IMPACT“ (2021) und „tech4comp“ (2018)
- Seit 2017 Koordination des Arbeitsbereichs E-Learning und E-Examinations (EEE) am CeDiS
- Konzeption und Betrieb der beiden E-Examination Center (EEC und EEC<sup>2</sup>)

# Agenda

**#1. Rückschau 2024:** KI und die Zukunft von Prüfungen

**#2. Gegenwart 2025:** Prüfungen im KI-Zeitalter

**#3. Fragen und Diskussion**

# Rückschau 2024: KI und die Zukunft von Prüfungen



# Rückschau 2024: Thesen

Tobias Halbherr	Alexander Schulz
1. Erweiterte Kognition: Das Potential reiner KI wird überschätzt und das Potential der Verbindung menschlicher Intelligenz mit KI wird unterschätzt.	4. Prüfungen können wegen GenKI schon bald nicht mehr wie gewohnt durchgeführt werden.
2. Klausuren: Leistungskontrollen in Form von Klausuren gewinnen angesichts generativer KI an Bedeutung	5. Durch GenKI ist der Aufwand („Produktionskosten“) für die Erstellung von Hausarbeiten so stark gesunken, dass Academic Misconduct schon bald der Normalfall sein wird.
3. Ressourcenreiches Prüfen - auch mit KI: Es wird einen zunehmenden Bedarf nach Leistungskontrollen mit sicherem Zugang zu KI-Werkzeugen geben	6. In Zeiten von GenKI sind mündliche Prüfungen keine Alternative.
	7. GenKI durchdringt bereits jetzt alle gesellschaftlichen Bereiche, daher wird AI Literacy erheblich relevanter werden als viele fachdisziplinäre Kompetenzen.



# Rückschau 2024: Möglichkeiten für Prüfungen I

## 1. Hausarbeiten

Anpassungen	Nachteile
Zusätzliche mündliche Prüfung	Aufwand in der Bewertungsphase steigt.
Wählen einer sehr spezifischen wiss. Fragestellung, bei der qua Spezialisierung eine hohe Wahrscheinlichkeit für Fehler besteht	Aufwand in der Vorbereitungs- und Bewertungsphase steigt. (Für die Vorbereitung könnte man vlt. sogar ChatGPT einsetzen)
Wählen einer Fragestellung, bei der gezielt die KI-Kompetenzen der Studierenden geprüft werden.	Der fachdisziplinäre Inhalt könnte zu weit hinter der Messung der KI-Kompetenz zurücktreten.
Beaufsichtigte klausurähnliche Arbeiten statt unbeaufsichtigter Hausarbeiten	Wissenschaftliches Arbeiten ist so nur teilweise möglich.

# Rückschau 2024: Möglichkeiten für Prüfungen II

## 2. Klausuren (on-campus oder fernbeaufsichtigt)

Anpassungen	Nachteile
Digitale Klausuren (z. B. in PC-Hall, E-Examination Center)	Erfassung mathematisch-naturwissenschaftlicher Formeln ist bislang technisch nicht ausgereift.
Einbezug von KI als Hilfsmittel in digitale Klausuren	Aufwand kann in der Bewertungsphase steigen.
Rückkehr zu klassisch schriftlichen Klausuren	Geringe Nähe zur Arbeitsrealität des 21. Jahrhunderts.

# Rückschau 2024: Möglichkeiten für Prüfungen II

## 3. Mündliche Prüfungen

Anpassungen	Nachteile
Ersetzen von unbeaufsichtigten schriftlichen Arbeiten durch mündliche Prüfungen	<p>Mündliche Prüfungen weisen hinsichtlich der Messung („Reliabilität“) diverse Herausforderungen auf:</p> <ul style="list-style-type: none"><li>• Subjektivität der Prüfer:innen</li><li>• Geringe Standardisierung</li><li>• Kommunikationssituation</li></ul> <p>Dazu steigen die Aufwände insbes. bei hohen Kohortenstärken.</p>

# Gegenwart 2025: Prüfungen im KI-Zeitalter



# Gegenwart 2025:

## Tobias Halbherr

### Hausarbeiten und projektbasierte Lehre (bzw. projektbasiertes Lernen, ‚PBL‘)

- Der Trend zu mehr PBL hält an und wird durch genAI nicht beeinträchtigt
- Hausarbeiten werden sich wandeln:
  - vermehrt formativer Fokus und/oder
  - Nutzung von genAI ist ausdrücklich erlaubt/erwünscht

### Projektarbeit und Klausuren nähern sich einander an

- Es entsteht ein Bedarf für Klausuren, welche nahtlos und authentisch an Projektarbeit anschliessen können

### Offenheit und Betrugssicherheit

- Es wird einen Bedarf nach offenen aber betrugssicheren Prüfungssituationen geben
- Wirksame Beaufsichtigung, technisch sowie vor Ort, gewinnt an bedeutung

### Mehr digitale Klausuren

- Authentische, innovative, digitale Klausurformate gewinnen weiter an Bedeutung
- ...aber die meisten Dozierenden (und Hochschulen sind noch nicht so weit)

# Gegenwart 2025: Schulen

**Alexander Schulz:** Handlungsempfehlungen der BMK vom 10.10.2024

1. Anpassung der Prüfungskultur an Digitalität: Qua Einzugs von KI-Anwendungen in Bildungsprozesse muss die Prüfungskultur dies widerspiegeln. Die Zukunftskompetenzen (4K: Kommunikation, Kollaboration, Kreativität und Kritisches Denken) müssen in Prüfungsprozesse einbezogen werden.
2. Kein allgemeines Verbot von KI-Technologie in Bildungsbereichen, da dies qua Verfügbarkeit von KI-Technologien in der Lebensrealität des 21. Jahrhunderts nicht zielführend sein kann.
3. Gefährdete Prüfungsformate, bei denen die Leistungen nicht eindeutig dem/der Prüfungsteilnehmer:in zugeordnet werden können, sollen abgeschafft oder grundlegend weiterentwickelt werden.
4. Bewertung von Prüfungen kann mithilfe von KI-Anwendungen unterstützt werden. Prüfungsrechtlich muss die Letztentscheidung der Bewertung aber bei den Lehrkräften liegen. Um Transparenz und Nachvollziehbarkeit im Entscheidungsprozess bei einer KI-unterstützten Prüfungsbewertung zu ermöglichen, bedarf es rechtlicher und ethischer Regelungen für den Einsatz dieser Entscheidungsunterstützungssysteme. Bias, der in allen Sprachmodellen mehr oder weniger stark vorhanden ist, stellt eine nicht nur technisch zu lösende Herausforderung dar.

Quelle: [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2024/2024\\_10\\_10-Handlungsempfehlung-KI.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf)

# Gegenwart 2025: Hochschulen I

## Alexander Schulz: KI und Prüfungen an Hochschulen

1. Beaufsichtigte Paper-Pencil-Klausuren bleiben weiterhin bestehen (oftmals um Faktenwissen zu prüfen).
2. Mündliche Prüfungen werden trotz ihrer Schwächen in den Gütekriterien („Reliabilität“) als Ergänzung zu unbeaufsichtigten schriftlichen Prüfungsformaten eingestuft. Obgleich eine Zunahme erkennbar ist, kann von massenhafter Verbreitung bislang m.E. noch nicht gesprochen werden.
3. Haus- und Abschlussarbeiten sind weiterhin als gefährdet einzustufen. Es wird daher empfohlen, bei diesen Formaten, den Reflexionsprozess bei der Erstellung der Arbeit zu dokumentieren. Dies passiert bereits. Aber auch scheint die massenhafte Verbreitung m.E. noch nicht erreicht zu sein.

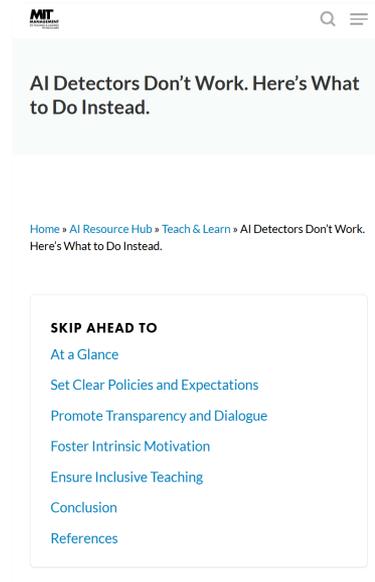
# Gegenwart 2025: Hochschulen II

## Alexander Schulz: KI und Prüfungen an Hochschulen

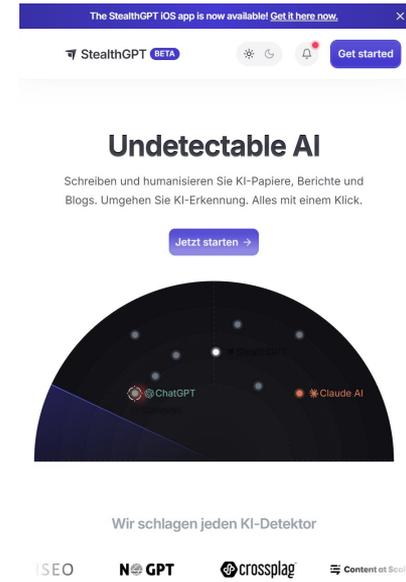
4. Tools zur AI Detection sind aufgrund der hohen Entropie und der stetig ansteigenden Kontextfenstergrößen der KI-Sprachmodelle ohnehin als unzuverlässig und damit vermutlich prüfungsrechtlich nicht belastbar einzustufen. Dies wird dennoch praktiziert. Fälle, die auf

Anscheinsbeweisen basieren, sind bekannt aus D und GB (München: <https://www.gesetze-bayern.de/Content/Document/Y-300-Z-BECKRS-B-2023-N-42327>, Guardian: <https://www.theguardian.com/technology/2024/dec/15/i-received-a-first-but-it-felt-tainted-and-undeserved-inside-the-university-ai-cheating-crisis>)

5. AI Humanizer Tools werben damit, GenAI Texte so zu vermenschlichen, dass sie nicht mehr erkannt werden können (z. B. „StealthGPT“).



mitsloanedtech.mit.edu/ai/teach/ai-detectors-dont-work/



www.stealthgpt.ai/de

# Fragen und Diskussion



# Vielen Dank!

Tobias Halbherr, ETH-Zürich  
[tobias.halbherr@id.ethz.ch](mailto:tobias.halbherr@id.ethz.ch)

Alexander Schulz, Uni Potsdam  
[zim-leitung@uni-potsdam.de](mailto:zim-leitung@uni-potsdam.de)

Freie Universität  
Lehren, Lernen und Prüfen mit  
Künstlicher Intelligenz  
[ki@fu-berlin.de](mailto:ki@fu-berlin.de)

AI Week: Organisatorisches  
[e-teaching@cedis.fu-berlin.de](mailto:e-teaching@cedis.fu-berlin.de)

# AIQualify - Qualifizierung zu Künstlicher Intelligenz in der Hochschullehre

Koordination und Kontakt  
Freie Universität Berlin  
FUB-IT. Lehre, Studium, Forschung (ex-CeDiS)  
E-Learning und E-Examinations (EEE)  
Ihnestraße 24, D-14195 Berlin  
e-teaching@cedis.fu-berlin.de  
<https://www.cedis.fu-berlin.de/nel-ki>

VERANSTALTET VON:



IM RAHMEN EINES PROJEKTES VON:



GEFÖRDERT VON:

