

Dilger, Alexander

Working Paper

Erfahrungen und Planungen zu KI in Lehre und Leistungsbeurteilungen

Diskussionspapier des Instituts für Organisationsökonomik, No. 6/2025

Provided in Cooperation with:

University of Münster, Institute for Organisational Economics

Suggested Citation: Dilger, Alexander (2025) : Erfahrungen und Planungen zu KI in Lehre und Leistungsbeurteilungen, Diskussionspapier des Instituts für Organisationsökonomik, No. 6/2025, Universität Münster, Institut für Organisationsökonomik, Münster, <https://doi.org/10.17879/73948673813>

This Version is available at:

<https://hdl.handle.net/10419/320420>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Diskussionspapier des
Instituts für Organisationsökonomik

6/2025

Erfahrungen und Planungen zu KI
in Lehre und Leistungsbeurteilungen

Alexander Dilger

Discussion Paper of the
Institute for Organisational Economics

**Diskussionspapier des
Instituts für Organisationsökonomik
6/2025**

Juni 2025

ISSN 2750-4476

**Erfahrungen und Planungen zu KI
in Lehre und Leistungsbeurteilungen**

Alexander Dilger

Zusammenfassung

Die Entwicklung der Künstlichen Intelligenz (KI) hat insbesondere bei Sprachmodellen große Fortschritte gemacht, die auch Auswirkungen auf Hochschulen haben. Hier wird über die bislang begrenzten Erfahrungen des Instituts für Organisationsökonomik an der Universität Münster hinsichtlich KI in der Lehre und bei Leistungsbeurteilungen berichtet und werden diesbezügliche Pläne für die Zukunft vorgestellt.

JEL Codes: A22, A23, C88, D83, I23, O33

Stichwörter: ChatGPT, Hochschule, Künstliche Intelligenz, Lehre, Prüfung

Experiences and Plans for AI in Teaching and Performance Assessment

Abstract

The development of artificial intelligence (AI) has made significant progress, particularly in large language models, which also has implications for higher education. This discussion paper reports on the limited experience of the Institute for Organisational Economics at the University of Münster with AI in teaching and performance assessment and presents future plans in this regard.

Im Internet unter:

http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO_06_2025

DOI: 10.17879/73948673813

Universität Münster
Institut für Organisationsökonomik
Scharnhorststraße 100
D-48151 Münster

Tel: +49-251/83-24303 (Sekretariat)
E-Mail: io@uni-muenster.de
Internet: www.wiwi.uni-muenster.de/io

Erfahrungen und Planungen zu KI in Lehre und Leistungsbeurteilungen

1. Einleitung

Die Künstliche Intelligenz (KI) hat im Bereich der Sprachmodelle große Fortschritte gemacht, was durch ChatGPT Ende 2022 der breiten Öffentlichkeit bewusst wurde und nicht zuletzt auf die Lehre an Hochschulen einschließlich Prüfungen großen Einfluss hat. Entsprechend ging es in dem „Interview mit ChatGPT über KI an Hochschulen“ vor gut zwei Jahren (Dilger 2023) auch um die Lehre, jedoch noch sehr allgemein und hypothetisch. Inzwischen liegen eigene konkrete Erfahrungen vor, wozu allerdings auch das Nichterleben von neuen Möglichkeiten gehört. Darüber wird im nachfolgenden zweiten Kapitel berichtet. Interessanter sind vielleicht konkrete Planungen für die Zukunft, weshalb im dritten Kapitel solche Planungen zum Einsatz von KI in der Lehre vorgestellt werden und im vierten Kapitel Planungen zu (veränderten) Leistungsbeurteilungen wegen KI. Der Beitrag schließt mit einem kurzen Fazit und Ausblick im fünften Kapitel.

2. Erfahrungen mit KI in Lehre und Leistungsbeurteilungen

Bislang gibt es am Institut für Organisationsökonomik kaum Erfahrungen zum Einsatz von KI in der Lehre. Von den Professoren und der Hälfte der Mitarbeiter wurde sie im Bereich der Lehre noch gar nicht eingesetzt und auch sonst nur für periphere Aufgaben wie sprachliche Korrekturen oder die Überprüfung und marginale Verbesserung eigener Vorlagen, nicht deren Generierung. Es erfolgte auch keine KI-gestützte Bewertung studentischer Leistungen und bislang auch keine Überprüfung mit KI, ob studentische Texte mit KI generiert wurden.

Die von der Fakultät bereitgestellte Möglichkeit zur Plagiatsprüfung beruht ebenfalls nicht auf KI, sondern weniger anspruchsvollen Textvergleichen, weshalb sie z. B. keine Übersetzungsplagiate und systematische Umformulierungen zur Verschleierung erkennen kann, was jedoch menschlichen Korrektoren gelegentlich gelingt. Dagegen ist am Institut bislang noch keine unerlaubte studentische KI-Nutzung hinreichend nachgewiesen und sanktioniert worden. Allerdings wird jetzt jeweils vorher explizit darauf hingewiesen, dass KI-Nutzung nicht pauschal verboten ist, sie jedoch als Hilfsmittel angegeben werden muss, was gelegentlich geschieht und dann jeweils für zulässige Anwendungen wie sprachliche Korrekturen, Übersetzungen oder Unterstützung bei der Zusammenfassung.

3. Planungen zu KI in der Lehre

Mittelfristig, in den nächsten Semestern, ist geplant, KI für spezifische Aufgaben in der Lehre einzusetzen oder zumindest den Einsatz zu erproben. Gerade am Anfang ist eine starke menschliche Kontrolle nötig, um „Halluzinationen“ der KI (also völlig falsche Tatsachenbehauptungen einschließlich erfundenen Zitaten und Literaturhinweisen sowie grobe Argumentationsfehler) auszuschließen. Der Mehraufwand kann jedoch als Investition angesehen werden, der sich auszahlt, wenn bzw. falls KI später vermehrt eingesetzt wird mit konkreten Erfahrungen hinsichtlich der Schwächen und Risiken und wie diesen wirksam begegnet werden kann.

Zu den spezifischen Aufgaben, für die der Einsatz oder zumindest Test von KI geplant wird, gehört die Überarbeitung von vorhandenen Präsentationsfolien. Von studentischer Seite wurde teilweise gewünscht, dass die eingesetzten Folien ausführlicher oder auch leichter verständlich gestaltet werden, wofür KI grundsätzlich gut geeignet erscheint. Es könnten künftig auch Antworten auf allgemeine studentische Anfragen von einer KI vorformuliert werden.

Interessanter und wohl auch wichtiger ist es, KI und den Umgang damit selbst zum Lehrinhalt zu machen. Was die wirtschaftliche Bedeutung von KI angeht und wie sie Unternehmensprozesse und ganze Märkte verändert, so wurde das Angebot vom Institut für Wirtschaftsinformatik angenommen, künftig einen Vorlesungstermin der Einführung in die Betriebswirtschaftslehre zu übernehmen, um einen entsprechenden Überblick über KI in betriebswirtschaftlichen Themenbereichen zu vermitteln. Außerdem soll demnächst ein Seminar dazu stattfinden. Aspekte davon sind auch gute Themen für Abschlussarbeiten, wobei am Institut für Organisationsökonomik die Studierenden ihre Themen grundsätzlich selbst wählen dürfen, so dass die tatsächliche Themenvergabe von entsprechender Nachfrage abhängt.

Für die eigene Anwendung von KI durch Studierende z. B. beim Erstellen wissenschaftlicher Texte ist demnächst ein Projektseminar geplant, um verschiedene KI-Modelle, deren Möglichkeiten, Risiken und Missbrauchsgefahren auszuprobieren. Die Ergebnisse davon können dann in Hinweise und Empfehlungen für normale Seminare und Abschlussarbeiten einfließen. Bis dahin bleibt es bei dem generellen Hinweis, dass die Nutzung von KI nicht verboten ist, sie jedoch als Hilfsmittel anzugeben ist. Dabei gibt es sinnvolle Nutzungsmöglichkeiten wie die sprachliche Korrektur oder inhaltliche Anregungen, aber auch zu weitgehende wie das Erstellen der gesamten Arbeit durch eine KI. Wenn das ehrlich angegeben wird, ist das kein Täuschungsversuch, aber die Eigenleistung zu gering fürs Bestehen. Im Grunde ist es wie bei

menschlichen Hilfspersonen, die jedoch noch schwerer entdeckt werden könnten, falls sie verschwiegen werden.

4. Planungen zu Leistungsbeurteilungen wegen KI

Eine KI könnte in Zukunft die Vorkorrektur von Seminar- und Abschlussarbeiten unterstützen. Allerdings ist dabei der Nutzen begrenzt, da das vollständige Lesen dieser Arbeiten am meisten Zeit benötigt, worauf zumindest anfangs nicht verzichtet werden kann. Auch bei der Überprüfung von Literaturangaben und Zitaten scheint die Verwendung einer KI bislang weniger hilfreich als herkömmliche Internetrecherchen. Klausuren müssten erst eingegeben oder eingelesen werden, was aufwendiger erscheint als eine Korrektur direkt durch einen Menschen.

Viel wichtiger ist der Punkt, dass für jeden kostenlos oder zumindest für wenig Geld verfügbare KI-Anwendungen inzwischen so leistungsstark sind, dass sie Texte verfassen können, die zumindest besser scheinen als die von durchschnittlichen Studierenden¹ und sogar Promovierenden. Bei der Leistungsbeurteilung darf das nicht ignoriert werden, weil ein Großteil der Studierenden KI tatsächlich nutzt² und diese Nutzung schwer nachzuweisen ist, wenn sie nicht freiwillig angegeben wird. Dies führt dazu, dass sich die Angemessenheit verschiedener Prüfungsformate verschiebt.

Noch vor drei Jahren galten schriftliche Ausarbeitungen wie Seminar- und Abschlussarbeiten als besonders gute Prüfungsformen. Die Studierenden haben hinreichend Zeit, um zu zeigen, was sie können, und das Ergebnis ihrer Bemühungen liegt schriftlich fixiert vor, um später beurteilt werden zu können. Das Anfertigen solcher Texte ähnelt auch eher ihren späteren Arbeitsaufgaben als Klausuren und mündliche Prüfungen, die es in dieser Form nur in Schulen und Hochschulen gibt. Ghostwriter waren schon früher ein potentielles Problem, welches jedoch (wahrscheinlich) praktisch selten auftrat. Die häufigsten Täuschungsversuche waren Plagiate, von denen auch nicht alle entdeckt wurden, deren Nachweis im Entdeckungsfall jedoch relativ einfach und rechtssicher möglich war (und immer noch ist).

¹ Hersh (2025) zeigt, dass KI in vielen Aufgabenstellungen besser ist als durchschnittliche Studierende der Medizin. Studierende sind mit KI außerdem besser als ohne, nach deren Wegnahme allerdings deutlich schlechter.

² Hüsch/Horstmann/Breiter (2025) haben über 23.000 Studierende befragt, von denen nur 6,1 Prozent angeben, noch nie KI genutzt zu haben. 24,9 Prozent nutzen sie täglich und weitere 40,1 Prozent wöchentlich. Nur 34 Prozent geben an, sie noch nie zur Textgenerierung genutzt zu haben.

Der Nachweis einer nicht angegebenen KI-Nutzung ist hingegen schwieriger, insbesondere wenn sie gut gemacht ist. Eine moderne KI wie ChatGPT von OpenAI oder Gemini von Google kann inzwischen kostenlos auf einen Befehl („Prompt“) hin in wenigen Minuten eine komplette Seminar- oder Abschlussarbeit verfassen, die bei oberflächlicher Betrachtung besser erscheint als eine solche Arbeit von normalen Studierenden. Bei genauerer Betrachtung können Fehler auffallen, die für heutige KI-Sprachmodelle typisch sind, allerdings in der Zukunft geringer werden oder sogar ganz verschwinden dürften. So werden bislang relativ häufig Behauptungen, Literaturquellen und auch Zitate einfach erfunden, ohne dass sie stimmen bzw. tatsächlich existieren. Im Grunde ist ein erfundenes Zitat die Umkehrung eines Plagiats, bei dem eine tatsächlich zitierte Stelle nicht korrekt angegeben wird. Eine KI kann auch Plagiate hervorbringen, wobei sie nicht einfach einen fremden Text kopiert, sondern quasi von selbst erneut erzeugt, weil die Wortfolge gegeben ihre Datengrundlage besonders plausibel erscheint.

Wenn die KI-Nutzung korrekt angegeben wird, sind solche Fehler immer noch Mängel, die zu einer schlechteren Note führen können, doch es liegt kein wissenschaftliches Fehlverhalten vor, welches mit Nichtbestehen oder noch strenger sanktioniert werden müsste. Wer sich etwas Mühe gibt, kann außerdem solche Fehler leicht beheben. Insbesondere Literaturangaben und Zitate der KI können mit nur geringem Aufwand überprüft werden, ob sie stimmen, was mit etwas mehr Aufwand auch für faktische Behauptungen gilt. Zukünftige KI-Modelle könnten selbst entsprechende Datenbank- bzw. Internetrecherchen umfassen. Das erschwert dann jedoch zugleich das Aufdecken einer nicht angegebenen KI-Nutzung. Als Prüfer kann man eine KI fragen, ob sie einen Text für KI-generiert hält, aber das ist in beide Richtungen nicht sehr genau, die KI kann also Texte einer KI für menschlich halten und umgekehrt. Ein weiteres Indiz ist eine sprachlich nahezu perfekte Arbeit, die früher ein starkes Indiz für hohe Intelligenz und Fleiß und damit auch eine inhaltlich sehr gute Arbeit gewesen wäre, während sie heute eher auf eine KI hindeutet, aber natürlich kein Beweis ist.

Die Leistungsstärke der KI-Modelle erfordert Anpassungen bei der Aufgabenstellung. Eine Möglichkeit bestünde darin, dass der Einsatz von KI nicht nur erlaubt wird, sondern empfohlen oder sogar erwartet wird. Auch im späteren Arbeitsleben dürfte es zukünftig immer bedeutender werden, KI sinnvoll einzusetzen. Das Entdecken und Korrigieren der oben angeführten typischen KI-Fehler ist dann eine wichtige Aufgabe der den Text zusammen mit der KI generierenden Menschen. Es könnten auch die inhaltlichen und argumentativen Anforderungen an Texte erhöht werden, bei denen KI geholfen hat. Bei anderen Computeranwen-

dungen ist das schon lange so. Gerade bei Abschlussarbeiten erwartet niemand Rechnungen per Hand, sondern Rechner können schon lange viel besser rechnen als Menschen und entsprechend können viel kompliziertere Berechnungen erwartet werden. Das gilt auch für ökonomische Schätzungen oder z. B. die Literaturrecherche mit dem Internet und Datenbank darin im Gegensatz zu Karteikarten früher. Die Prüfer müssen dann jedoch auch eigene Erfahrungen machen, welche Anforderungen realistischerweise gestellt werden können und wie entsprechende Kompetenzen zu vermitteln sind.

Keine Lösung ist das pauschale Verbot von KI-Nutzung, welches nicht nur schwierig rechtsicher umzusetzen wäre, sondern vor allem einfach ignoriert werden könnte. Dagegen ist es möglich, Aufgabenstellungen zu wählen, die sich zumindest deutlich schwieriger mit KI bewältigen lassen als andere. Reine Literaturarbeiten waren schon vor ChatGPT & Co. zweifelhaft, während empirische Arbeiten mehr eigenen Einsatz erfordern, zumindest solange nicht auch gleich die ganze Datengrundlage mit KI erzeugt wird. Das gilt auch für sehr spezielle Themen, bei denen die KI keine hinreichende Datengrundlage hat. Schließlich ist die KI nicht gut bei sehr tiefeschürfenden Argumentationen,³ die allerdings auch den meisten Studierenden sehr schwerfallen, während sie als Kernkompetenz von Professoren zukünftig noch an Bedeutung gewinnen dürften.

Durch KI werden andere Prüfungsformate, bei denen sie sich weniger gut oder gar nicht einsetzen lässt, an Bedeutung gewinnen. Klausuren sind nun besser geeignet für eine faire Leistungsbeurteilung als Seminar- und Abschlussarbeiten. Technische Hilfsmittel, mit denen sich kommunizieren lässt, waren ohnehin schon verboten, was nur durch Corona-Maßnahmen kurzzeitig ausgesetzt oder abgeschwächt wurde mit z. T. dramatischen Folgen. Auch mündliche Leistungen gewinnen an Gewicht. So könnte wieder eine mündliche Verteidigung von Abschlussarbeiten eingeführt werden oder eigenständige mündliche Abschlussprüfungen. Allerdings müssten dazu die Prüfungsordnungen geändert werden und ist der Mehraufwand in Massenstudiengängen nicht zu unterschätzen. Schließlich ist jetzt der Seminarvortrag wichtiger geworden relativ zur schriftlichen Seminararbeit, die mit KI generiert worden sein könnte. Allerdings muss dazu zukünftig auf einen relativ freien Vortrag bestanden werden. Beim reinen Ablesen könnte es sich ebenfalls um KI-generierten Texte handeln. Bei der nachfolgenden Diskussion ist darauf zu achten, dass alle referierenden Studierenden sich daran tatsächlich beteiligen.

³ Siehe z. B. Shojaee et al. (2025), bei deren Tests *Large Language Models* bei relativ einfachen intellektuellen Problemen überlegen sind, *Large Reasoning Models* bei mittelschweren Problemen besser sind und beide Arten von KI bei hochkomplexen Aufgabenstellungen versagen.

5. Fazit und Ausblick

Die Entwicklung der KI hat in den letzten Jahren große Fortschritte gemacht, insbesondere im Bereich der Sprachmodelle, die nun sprachlich nahezu perfekte und auch inhaltlich recht gute Texte verfassen können, auch wenn es noch systematische Fehler gibt. Es handelt sich noch immer nicht um eine künstliche Superintelligenz, die der menschlichen Intelligenz ganz allgemein überlegen wäre, doch die Zahl der konkreten Anwendungen, bei denen Computer z. T. deutlich bessere (und schneller) Ergebnisse produzieren können als Menschen, nimmt weiter zu und umfasst nun auch sprachliche Kompetenzen. Das hat große Auswirkungen auf die menschliche Gesellschaft und Wirtschaft allgemein, aber auch speziell auf die Hochschulen, deren Lehre und Prüfungen.

Davon wurden exemplarisch konkrete Auswirkungen auf die Lehre und Prüfungen am Institut für Organisationsökonomik an der Universität Münster betrachtet, die in den nächsten Semestern zu erwarten bzw. geplant sind. Die langfristigen Auswirkungen werden viel tiefgreifender sein, lassen sich jedoch noch gar nicht abschätzen, da sie nicht nur von noch unbekanntem technischen Entwicklungen, sondern auch von rechtlichen Vorgaben und Reaktionen künftiger Studierender abhängen werden. Es ist schwierig, diese Situation proaktiv zu gestalten, doch zumindest muss auf bereits laufende Entwicklungen jeweils angemessen reagiert werden, um weiterhin relevante Lehre und adäquate sowie faire Prüfungen anbieten zu können.

Literatur

- Dilger, Alexander (2023): „Interview mit ChatGPT über KI an Hochschulen“, Diskussionspapier des Instituts für Organisationsökonomik 3/2023, Münster.
- Hersh, William (2025): „Generative Artificial Intelligence: Implications for Biomedical and Health Professions Education“, arXiv:2501.10186, im Internet unter <https://arxiv.org/abs/2501.10186> (letzter Zugriff am 30.06.2025).
- Hüsch, Marc/Horstmann, Nina/Breiter, Andreas (2025): „Künstliche Intelligenz im Studium – die Sicht von Studierenden im Wintersemester 2024/25“, DatenCHECK 6/2025, CHE Centrum für Hochschulentwicklung, Gütersloh, im Internet unter <https://hochschuldaten.ch.de/kuenstliche-intelligenz-im-studium-die-sicht-von-studierenden-im-wintersemester-2024-25> (letzter Zugriff am 30.06.2025).
- Shojaee, Parshin/Mirzadeh, Iman/Alizadeh, Keivan/Horton, Maxwell/Bengio, Samy/Farajtabar, Mehrdad (2025): „The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity“, arXiv:2506.06941, im Internet unter <https://arxiv.org/abs/2506.06941> (letzter Zugriff am 30.06.2025).

Diskussionspapiere des Instituts für Organisationsökonomik

Seit Institutsgründung im Oktober 2010 erscheint monatlich ein Diskussionspapier. Im Folgenden werden die letzten zwölf aufgeführt. Eine vollständige Liste mit Downloadmöglichkeit findet sich unter <http://www.wiwi.uni-muenster.de/io/de/forschen/diskussionspapiere>.

- DP-IO 6/2025** Erfahrungen und Planungen zu KI in Lehre und Leistungsbeurteilungen
Alexander Dilger
Juni 2025
- DP-IO 5/2025** Home Bias in Different German Football Leagues
Tom Böttger/Lars Vischer
Mai 2025
- DP-IO 4/2025** Prospect Theory and Psychological Factors in Football
Effects of the Video Assistant Referee on Penalty Conversion Rates
in the German Bundesliga
Tom Böttger/Lars Vischer
April 2025
- DP-IO 3/2025** Individuelle Entscheidungen zur Steuerverwendung
Alexander Dilger
März 2025
- DP-IO 2/2025** Finanzparlament
Alexander Dilger
Februar 2025
- DP-IO 1/2025** Going Global, Going Digital
Firm Internationalisation and Digital Resource Use
Pascal Mayer
Januar 2025
- DP-IO 12/2024** Auswirkungen von weniger Studierenden
Alexander Dilger
Dezember 2024
- DP-IO 11/2024** Home Market Competitive Pressure and the Tendency to Internationalise
An Inverse U-shape
Pascal Mayer/Thomas Schäper
November 2024
- DP-IO 10/2024** 14. Jahresbericht des Instituts für Organisationsökonomik
Alexander Dilger
Oktober 2024
- DP-IO 9/2024** Ökonomische Bildung in der Sozialen Marktwirtschaft
Korreferat zu Nils Goldschmidt und Marco Rehm
Alexander Dilger
September 2024
- DP-IO 8/2024** The Impact of Host Country Institutional Factors on International Investments
Miriam Nebenführ
August 2024
- DP-IO 7/2024** Flexibles Arbeiten im Hochschulmanagement
Caroline Werner
Juli 2024



Herausgeber:
Prof. Dr. Alexander Dilger
Universität Münster
Institut für Organisationsökonomik
Scharnhorststr. 100
D-48151 Münster

Tel: +49-251/83-24303
Fax: +49-251/83-28429

www.wiwi.uni-muenster.de/io

