EMPIRISCHER BEITRAG

Reflektierte KI-Nutzung in Lehrveranstaltungen

Befunde aus dem Pretest am Ende des Sommersemesters 2026

Tanja Mühling[1], Sven Warnke[1], Harald Bolsinger[2]

[1] Technische Hochschule Würzburg-Schweinfurt, Fakultät Angewandte Sozialwissenschaften

[2] Technische Hochschule Würzburg-Schweinfurt, THWS Business School

Zusammenfassung

Der Beitrag berichtet Befunde aus dem Pretest einer Studierendenbefragung zur Nutzung generativer KI in Lehrveranstaltungen, in denen der KI-Einsatz didaktisch gerahmt und reflektiert wird. Am Ende des Sommersemesters 2026 wurden 50 Studierende aus drei Studiengängen der Technischen Hochschule Würzburg-Schweinfurt (THWS) befragt. In den Lehrveranstaltungen nutzten viele von ihnen erstmals KI-Anwendungen jenseits von ChatGPT, insbesondere ein lokal betriebenes Sprachmodell (AnythingLLM), Claude und NotebookLM. Am verbreitetsten ist es unter Studierenden, sich Inhalte erklären zu lassen oder die eigenen Inhalte mit KI zu strukturieren; Texte für Prüfungsleistungen lässt sich die Mehrheit höchstens gelegentlich erstellen. Eine Hauptkomponentenanalyse unterscheidet eine produktorientierte, eine reflexiv-dialogische und eine erschließende Nutzung generativer KI. Rund sieben von zehn Befragten geben an, KI am Semesterende bewusster und gezielter einzusetzen als zu Semesterbeginn; Studierende mit geringer Vorerfahrung schätzen ihre Entwicklung deskriptiv am höchsten ein. Den eigenen Umgang mit KI beschreiben die Befragten überwiegend als reflektiert, wobei sie der kritischen Bewertung von KI-Ergebnissen stärker zustimmen als der bewussten Steuerung des KI-Einsatzes. In offenen Antworten nennen die Befragten hinsichtlich ihrer KI-bezogenen Lernfortschritte im Semesterverlauf vor allem gezielteres Prompting, die kritische Prüfung von Ergebnissen und den bewussten Verzicht auf KI bei bestimmten Aufgaben. Da die Ergebnisse auf retrospektiven Selbstauskünften einer kleinen, nicht zufällig gezogenen Stichprobe beruhen, sind sie als explorative Hinweise zu verstehen, die das Erhebungsinstrument für die geplante Längsschnittstudie vorbereiten.

Schlagwörter: generative KI, Hochschullehre, Studierendenbefragung, Nutzungszwecke, reflektierter Umgang mit KI, Pretest

Abstract

Reflective use of AI in university courses: Findings from a pretest at the end of the 2026 summer semester

This article reports on findings from the pretest of a student survey on the use of generative AI in courses where AI use is framed and reflected upon from a pedagogical perspective. At the end of the summer semester of 2026, 50 students from three degree programs at the Würzburg-Schweinfurt University of Applied Sciences (THWS) were surveyed. In these courses, many of them used AI applications beyond ChatGPT for the first time, particularly a locally operated language model (AnythingLLM), Claude, and NotebookLM. The most common use among students is to have content explained to them or to use AI to structure their own content; the majority use AI to generate texts for exams only occasionally, if at all. A principal component analysis distinguishes between a product-oriented, a reflective-dialogic, and an exploratory use of generative AI. About seven out of ten respondents report using AI more consciously and purposefully at the end of the semester than at the beginning; students with little prior experience rate their own development most highly in descriptive terms. Respondents predominantly

Veröffentlicht: 5. November 2026
© 2026 Tanja Mühling, Sven Warnke, Harald Bolsinger · CC BY 4.0

describe their own use of AI as reflective, though they agree more strongly with the critical evaluation of AI results than with the conscious control of AI usage. In open-ended responses, respondents cited, above all, more targeted prompting, the critical examination of results, and the conscious decision not to use AI for certain tasks as key aspects of their AI-related learning progress over the course of the semester. Since the results are based on retrospective self-reports from a small, non-random sample, they should be understood as exploratory findings that help refine the survey instrument for the planned longitudinal study.

Keywords: generative AI, higher education teaching, student survey, purposes of use, reflective use of AI, pretest

Veröffentlicht: 5. November 2026
Verfahren: Eingeladener Beitrag der Herausgeber:innen im Startverfahren der ersten Ausgabe; gemeinsam konzipiert und verantwortet, keine Prüfung durch nicht beteiligte Personen.
KI-Nutzung in der Prüfung: KI-erstellte Gutachten (Claude, Anthropic) dienten als Sammlung möglicher Einwände. Alle Entscheidungen haben die Herausgeber:innen getroffen.

Zitierempfehlung: Mühling, T., Warnke, S., & Bolsinger, H. J. (2026). Reflektierte KI-Nutzung in Lehrveranstaltungen: Befunde aus dem Pretest am Ende des Sommersemesters 2026. *Practice of Artificial Intelligence in Higher Education, 1*(1), Artikel 3. https://doi.org/10.83397/PAIHE.2026.03

DOI: https://doi.org/10.83397/PAIHE.2026.03 · **Lizenz:** © 2026 Tanja Mühling, Sven Warnke, Harald Bolsinger. Dieser Beitrag steht unter der Lizenz CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/deed.de).

1 Einleitung

Generative KI ist innerhalb kurzer Zeit zu einem selbstverständlichen Bestandteil des Studiums geworden. In einer bundesweiten Befragung von 6.311 Studierenden gaben 2023 noch 63,2 % an, KI-basierte Werkzeuge für ihr Studium zu nutzen (von Garrel & Mayer, 2023); in der Folgeerhebung 2025 mit 4.910 Studierenden waren es bereits 91,6 % (von Garrel & Mayer, 2025). Im CHE-Hochschulranking des Wintersemesters 2025/26 nutzten 81 % der 28.584 befragten Studierenden wirtschafts- und rechtswissenschaftlicher Fächer sowie der Sozialen Arbeit KI mindestens wöchentlich, 46,5 % sogar täglich (Horstmann et al., 2026).

Im Vordergrund stehen dabei Zwecke, die das eigene Verstehen und die Informationssuche unterstützen. Verständnisfragen und Erklärungen fachlicher Konzepte nannten 2025 zwei Drittel der Studierenden (66,7 %), gefolgt von der Analyse, Verarbeitung und Erstellung von Texten (51,7 %; von Garrel & Mayer, 2025). Regelmäßig nutzen Studierende KI für Recherche und Überblick (68 %), zur Ideenfindung (60 %) und als Tutor oder Lernpartner (50 %), wobei die Unterschiede zwischen den Fächern erheblich sind (Horstmann et al., 2026).

Mit der Verbreitung hat die Vorbereitung auf einen kompetenten Umgang nicht Schritt gehalten. In der Panelstudie der Dualen Hochschule Baden-Württemberg fühlten sich 75,6 % der Studierenden durch ihr Studium nicht gut oder gar nicht gut auf KI vorbereitet; den Umgang mit KI hatten sie sich überwiegend selbstständig angeeignet (Ehlers & Rauch, 2024). Das Angebot zum Erwerb von KI-Kompetenzen bewerteten Studierende im CHE-Hochschulranking mit durchschnittlich 3,3 von 5 Sternen; 58 % wünschten sich mehr Lernangebote zu KI-Kompetenzen und 70 % klare Richtlinien zu Ethik und Datenschutz (Horstmann et al., 2026).

Darüber hinaus ist die Nutzung sozial ungleich verteilt. In der DHBW-Studie erreichten 49 % der befragten Männer, aber nur 30 % der befragten Frauen hohe oder mittlere Werte in einem Index der KI-Aktivität (Ehlers & Rauch, 2024). Deutliche Bildungsunterschiede zeigen sich auch im Erwerbsleben: Nach Daten des Nationalen Bildungspanels aus der

Befragungswelle 2023/24 nutzen rund drei Viertel der 37- bis 60-jährigen Beschäftigten mit Hochschulabschluss textgenerierende KI-Systeme wie ChatGPT oder Claude zumindest selten, unter beruflich Gebildeten weniger als die Hälfte; oft nutzen solche Systeme 10 % der akademisch und 3 % der beruflich Gebildeten (Dicks et al., 2026, S. 12). Für die Hochschullehre stellt sich damit die Frage, inwiefern eine angeleitete KI-Nutzung solche Unterschiede in den Ausgangsvoraussetzungen verringern kann oder ob sie sie eher verstärkt.

Die vorliegenden Studien beschreiben vor allem die Verbreitung und die Zwecke der Nutzung. Wenig bekannt ist dagegen, wie Studierende KI nutzen, wenn ihr Einsatz in Lehrveranstaltungen ausdrücklich vorgesehen, angeleitet und reflektiert wird, und welche Studierenden davon besonders profitieren. Beitrag 1 (Warnke et al., 2026) begründet hierfür einen Minimalkonsens mit drei Kernfunktionen: eigene Vorleistung vor der KI-Nutzung, KI als Feedback zu eigenen Zwischenständen sowie KI-Nutzung begründet wählen, dokumentieren und reflektieren. Der vorliegende Beitrag berichtet Befunde aus dem Pretest einer Studierendenbefragung in Lehrveranstaltungen, die einer früheren Fassung dieses Konsenses folgten. Er beschreibt die Vorerfahrung der Studierenden mit generativer KI (Abschnitt 3), die erstmalige Nutzung von KI-Anwendungen (Abschnitt 4), die Nutzungszwecke (Abschnitt 5), die wahrgenommene Entwicklung im Semesterverlauf (Abschnitt 6) und die studentischen Selbsteinschätzungen zum reflektierten Umgang mit generativer KI (Abschnitt 7).

2 Methodische Vorbemerkungen

Am Pretest haben 50 Studierende aus drei Studiengängen der THWS-Fakultäten Business School (FBUS) und Angewandte Sozialwissenschaften (FAS) teilgenommen. Ziel ist es, eine Längsschnittstudie aufzubauen, bei der in ausgewählten Lehrveranstaltungen, in denen eine reflektierte KI-Nutzung unter klar definierten Rahmenbedingungen stattfindet, jeweils am Semesterbeginn und am Semesterende der subjektive Kenntnisstand und das Nutzungsverhalten der Studierenden erhoben werden. Diese Paneldaten erlauben Analysen dazu, welche Studierendengruppen im Semesterverlauf besonders starke Kompetenzfortschritte im Bereich der Nutzung generativer KI durch die betreffenden Lehrformate und -inhalte erzielen. Darüber hinaus sollen die Daten semesterübergreifend für Trendanalysen genutzt werden; diese sollen zeigen, inwiefern sich Studierendenkohorten in den kommenden Jahren in ihren Vorkenntnissen, Einstellungen und ihrem Nutzungsverhalten bezüglich generativer KI unterscheiden. Dabei werden auch Aspekte sozialer Ungleichheit (im Sinne von Geschlechterunterschieden oder Unterschieden nach dem Herkunftsland der Studierenden) bei den Auswertungen berücksichtigt.

Die Befragung wurde am Ende des Sommersemesters 2026 als standardisierte Online-Befragung (mit EvaSys) durchgeführt; die Fragebögen wurden zwischen dem 24. Juni und dem 9. Juli 2026 ausgefüllt. Erfasst wurden Kontextmerkmale, die genutzten KI-Anwendungen und Modelle, die Nutzungszwecke, die wahrgenommene Entwicklung im Umgang mit generativer KI seit Semesterbeginn und der reflektierte Umgang mit generativer KI; eine offene Frage schloss den Fragebogen ab. Von den Befragten ordneten sich 40 dem weiblichen und 8 dem männlichen Geschlecht zu, 2 machten keine Angabe. Im 5. oder 6. Studiensemester befanden sich 36 Studierende, im 3. oder 4. Studiensemester 6 und im 7. oder einem höheren Studiensemester 8. Von den 50 Befragten bearbeiteten 37 den Fragebogen in derselben Lehrveranstaltung; 43 gaben an, dass es in ihrer Lehrveranstaltung

verbindliche Regelungen oder Absprachen zur Nutzung generativer KI gab. Alle Befragten sind in Deutschland geboren, sodass sich Unterschiede nach dem Herkunftsland im Pretest nicht untersuchen lassen.

Forschungsethik. Die Teilnahme an der Befragung war freiwillig; Namen wurden nicht erhoben. Für diese Befragung ohne direkte personenidentifizierende Angaben wurde kein formales Votum einer Ethikkommission eingeholt; ein solches war nicht erforderlich.

3 Subjektive Erfahrung mit generativer KI vor der Lehrveranstaltung

Von den befragten Studierenden attestierten sich rückwirkend 62 % selbst nur eine geringe oder mittlere Erfahrung mit generativer KI zu Beginn des Sommersemesters 2026. Die übrigen 38 % stuften die eigene Erfahrung als hoch oder sehr hoch ein (Tabelle 1).

Tabelle 1
Selbst eingeschätzte Erfahrung mit generativer KI vor der Lehrveranstaltung nach Geschlecht

Erfahrung	männlich (*n* = 8)	weiblich (*n* = 40)	keine Angabe (*n* = 2)	Gesamt (*N* = 50)
gering	0 (0,0)	4 (10,0)	1 (50,0)	5 (10,0)
mittel	6 (75,0)	19 (47,5)	1 (50,0)	26 (52,0)
hoch	1 (12,5)	14 (35,0)	0 (0,0)	15 (30,0)
sehr hoch	1 (12,5)	3 (7,5)	0 (0,0)	4 (8,0)

Anmerkung. Angaben als Anzahl (Spaltenprozent). Frage: „Wie schätzen Sie Ihre Erfahrung mit generativer KI vor dieser Lehrveranstaltung ein?“ Die Kategorie „keine Erfahrung“ wurde nicht gewählt. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026.

Studenten stuften die eigene Vorerfahrung mit generativer KI in keinem Fall als gering ein, wohingegen dies 10 % der Studentinnen angaben. Während nur 7,5 % der weiblichen Befragten sich selbst eine sehr hohe Erfahrung zuschrieben, gaben dies 12,5 % der teilnehmenden Männer an. Wegen der geringen Zahl männlicher Befragter (*n* = 8) sind diese Unterschiede nur als Tendenz zu lesen.

Neben dem Geschlechterunterschied lässt sich auch ein Unterschied nach der Zahl der Studiensemester feststellen (Tabelle 2).

Tabelle 2
Selbst eingeschätzte Erfahrung mit generativer KI vor der Lehrveranstaltung nach Studiensemester

Erfahrung	3.–4. Semester (*n* = 6)	5.–6. Semester (*n* = 36)	7. Semester oder höher (*n* = 8)	Gesamt (*N* = 50)
gering	2 (33,3)	3 (8,3)	0 (0,0)	5 (10,0)
mittel	3 (50,0)	18 (50,0)	5 (62,5)	26 (52,0)
hoch	1 (16,7)	13 (36,1)	1 (12,5)	15 (30,0)
sehr hoch	0 (0,0)	2 (5,6)	2 (25,0)	4 (8,0)

Anmerkung. Angaben als Anzahl (Spaltenprozent). Studierende im 1. oder 2. Semester haben nicht teilgenommen. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026.

Je höher das Studiensemester, desto höher wurden rückwirkend die eigenen Erfahrungen mit generativer KI eingestuft. Der Zusammenhang ist allerdings schwach und statistisch nicht abgesichert (Spearmans ρ = ,22, p = ,13).

4 Erstmalige Nutzung generativer KI aufgrund der Lehrveranstaltung

Für 16 vorgegebene KI-Anwendungen gaben die Studierenden an, wie häufig sie diese im Rahmen der Lehrveranstaltung genutzt haben, und für jede genutzte Anwendung, ob sie diese dabei zum ersten Mal verwendet haben. Insgesamt nutzten 40 der 50 Befragten (80 %) mindestens eine Anwendung erstmals, im Mittel sogar 2,2 Anwendungen (*SD* = 2,2; Median = 2). Insbesondere AnythingLLM, Claude und NotebookLM wurden im Rahmen der Lehrveranstaltungen erstmals genutzt. Dagegen war die Erstnutzung von ChatGPT erwartungsgemäß die Ausnahme (Tabelle 3).

Tabelle 3
Nutzung und erstmalige Nutzung von KI-Anwendungen im Rahmen der Lehrveranstaltung

KI-Anwendung	genutzt (*n*)	erstmals genutzt (*n*)	erstmals in % aller Befragten	erstmals in % der Nutzenden
AnythingLLM (lokaler Offlinebetrieb auf privatem Endgerät)	19	17	34	89
Claude	38	16	32	42
NotebookLM	28	14	28	50
Qwen/Qwen Chat	12	12	24	100
Gemini	38	10	20	26
Perplexity	21	8	16	38
Mistral/Le Chat	5	5	10	100
iCanpreneur	7	5	10	71
DeepSeek	5	4	8	80

KI-Anwendung	genutzt (*n*)	erstmals genutzt (*n*)	erstmals in % aller Befragten	erstmals in % der Nutzenden
Kimi/Kimi Chat	4	4	8	100
Ollama (lokaler Offlinebetrieb auf privatem Endgerät)	4	4	8	100
ChatGPT	47	3	6	6
Microsoft Copilot	11	3	6	27
Meta AI	7	3	6	43
GLM/Z.ai	3	3	6	100
Grok	2	0	0	0

Anmerkung. „Genutzt“ = mindestens „selten“ im Rahmen der Lehrveranstaltung. Mehrfachnennungen möglich; Prozentangaben in der vorletzten Spalte bezogen auf alle Befragten (*N* = 50). Sortiert nach der Zahl der Erstnutzungen. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026.

ChatGPT nutzten nahezu alle Befragten (47 von 49 mit gültiger Angabe), aber nur drei erstmals. Umgekehrt war die Lehrveranstaltung für fast alle Nutzenden von AnythingLLM (17 von 19) und für alle Nutzenden von Qwen der erste Kontakt mit der jeweiligen Anwendung; bei Claude und NotebookLM traf dies auf 42 % bzw. 50 % der Nutzenden zu. Die Lehrveranstaltungen haben das Werkzeugrepertoire der Studierenden damit deutlich über die verbreiteten Chatbots hinaus erweitert, insbesondere um lokal betriebene Lösungen (AnythingLLM, Ollama), bei denen keine Daten an Anbieter übertragen werden. Welche Anwendungen erstmals genutzt wurden, spiegelt dabei unmittelbar die didaktische Gestaltung der einzelnen Veranstaltungen wider: So stammen 16 der 17 Erstnutzungen von AnythingLLM aus derselben Lehrveranstaltung (siehe hierzu Beitrag 2; Bolsinger et al., 2026). Mit der Vorerfahrung hängt die Zahl erstmals genutzter Anwendungen nicht zusammen (Spearmans ρ = −,03); auch Studierende, die ihre Erfahrung als hoch einschätzten, lernten also neue Anwendungen kennen.

5 Aktuelle Nutzung von KI im Kontext der Lehrveranstaltung

Die Studierenden wurden im Rahmen der Online-Befragung außerdem gebeten darzulegen, für welche Zwecke sie KI im Kontext der Lehrveranstaltung bzw. der damit verbundenen Prüfungsleistung genutzt haben. Über alle elf erhobenen Nutzungszwecke hinweg sind „Ich nutze generative KI, um mir Inhalte erklären zu lassen“ (mit 86 % häufiger oder sehr häufiger Nutzung) und „Ich nutze generative KI, um meine Inhalte zu strukturieren“ (mit 80 % häufiger oder sehr häufiger Nutzung) aktuell die verbreitetsten Einsatzzwecke. Der Zweck „Texte für Prüfungsleistungen erstellen lassen“ ist dagegen der einzige, bei dem der Modus (36 %) bei der Mittelkategorie „gelegentlich“ liegt; zudem hat über ein Drittel der Studierenden (36 %) „nie“ oder „selten“ angegeben.

Eine Hauptkomponentenanalyse mit Varimax-Rotation ergab nach dem Kaiser-Kriterium drei Komponenten, die zusammen 62,6 % der Varianz erklären. Sie lassen sich als „produktorientierte Nutzung“, „reflexiv-dialogische Nutzung“ und „erschließende Nutzung“ interpretieren (Abbildung 1). Bei der produktorientierten Nutzung werden Aufgaben des Arbeitsprozesses an KI ausgelagert bzw. KI liefert sogar Teile des Arbeitsergebnisses. Dagegen zeichnet sich die reflexiv-dialogische Nutzung dadurch aus, dass KI selbst erstellte

Arbeiten überprüft und Verbesserungsvorschläge macht. Die erschließende Nutzung umfasst KI-gestützte Recherchen und Erklärungen für ein besseres inhaltliches Verständnis der betreffenden Fragestellung. Abbildung 1 zeigt die Antwortverteilung der einzelnen Nutzungszwecke, geordnet nach diesen drei Komponenten.

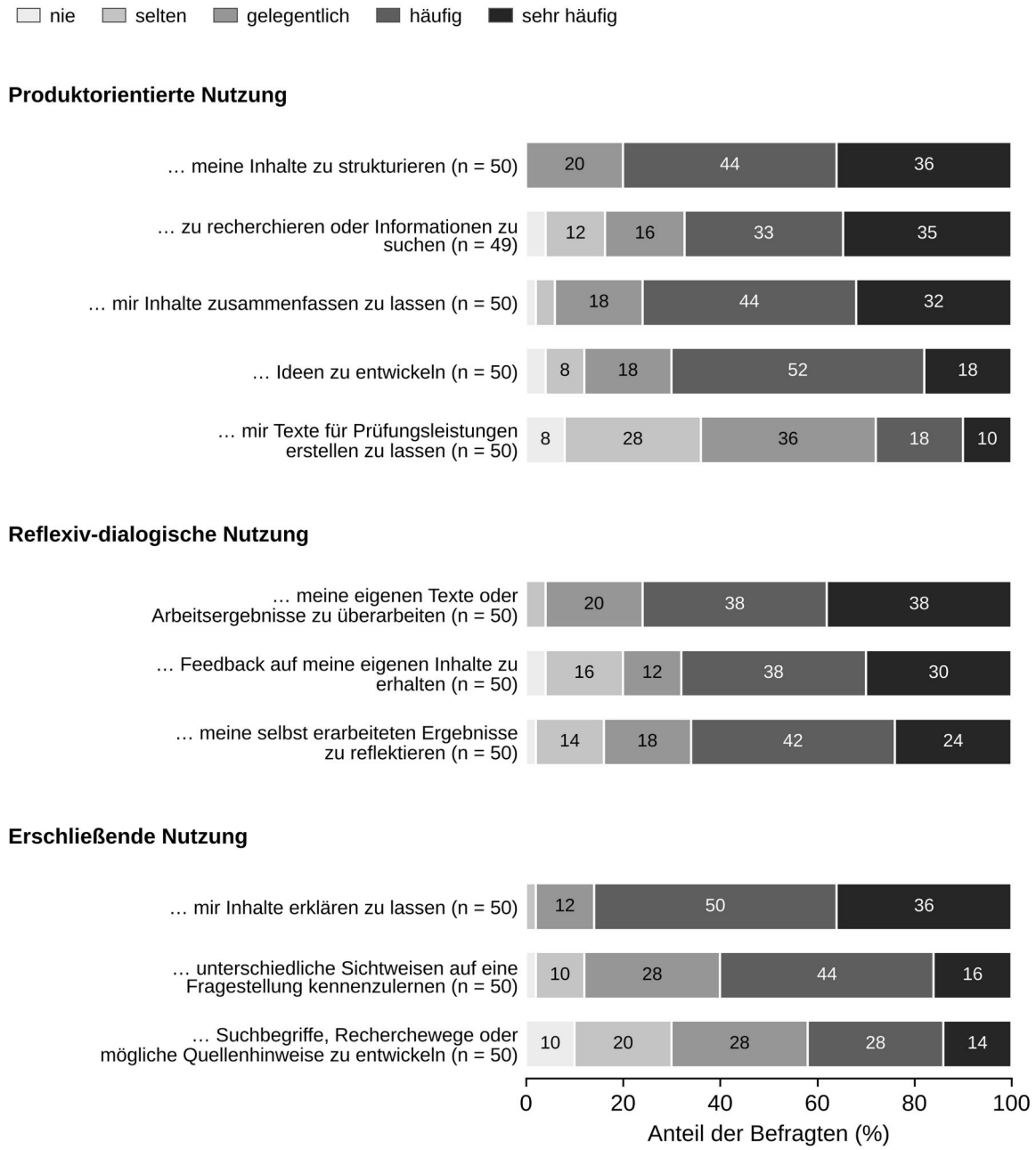
Innerhalb der produktorientierten Nutzung zeigt sich ein deutliches Gefälle: Inhalte zu strukturieren (80 % häufig oder sehr häufig), zusammenfassen zu lassen (76 %) und Ideen zu entwickeln (70 %) ist weit verbreitet, sich Texte für Prüfungsleistungen erstellen zu lassen, liegt dagegen nur bei 28 %. Die Komponente bündelt damit unterstützende Tätigkeiten im Arbeitsprozess und die weitgehende Auslagerung von Prüfungsprodukten, die normativ sehr unterschiedlich zu bewerten sind.

Auch die reflexiv-dialogische Nutzung ist verbreitet: Rund zwei Drittel holen häufig oder sehr häufig Feedback auf eigene Inhalte ein (68 %) oder reflektieren selbst erarbeitete Ergebnisse mit KI (66 %), drei Viertel überarbeiten mit KI eigene Texte (76 %). Zugleich nutzt jede fünfte befragte Person KI nie oder selten, um Feedback zu erhalten (20 %).

In der erschließenden Nutzung dominiert das Erklären von Inhalten (86 %), während KI deutlich seltener eingesetzt wird, um Suchbegriffe, Recherchewege oder Quellenhinweise zu entwickeln (42 %; nie oder selten: 30 %).

Abbildung 1

Nutzungszwecke generativer KI nach Komponenten



Anmerkung. Itemstamm: „Ich nutze generative KI, um …“. Antwortskala von 1 (nie) bis 5 (sehr häufig). Angaben in gültigen Prozent; Werte unter 7 % sind nicht beschriftet, Summen können rundungsbedingt von 100 % abweichen. Zuordnung der Items nach der höchsten Ladung in einer Hauptkomponentenanalyse mit Varimax-Rotation, innerhalb der Komponenten absteigend nach dem Anteil „sehr häufig“ sortiert. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026, N = 50.

Auf Indexebene liegen die drei Nutzungsformen nahe beieinander (produktorientiert: M = 3,73, SD = 0,68; reflexiv-dialogisch: M = 3,85, SD = 0,87; erschließend: M = 3,66, SD = 0,70). Die interne Konsistenz ist für die reflexiv-dialogische (α = ,79) und die produktorientierte Nutzung (α = ,74) zufriedenstellend, für die erschließende Nutzung (α = ,52) jedoch gering; diese Komponente ist daher – zumindest vorerst – auf der Ebene der Einzelitems zu interpretieren. Mit der Vorerfahrung hängen die drei Nutzungsformen nicht zusammen ($|\rho| \leq$,16).

Dass viele Studierende KI als Feedbackinstanz für eigene Arbeiten nutzen, entspricht der zweiten Kernfunktion des Minimalkonsenses aus Beitrag 1 (Warnke et al., 2026). Bemerkenswert ist der Abstand zwischen der allgemeinen Informationssuche (67 % häufig oder sehr häufig) und der Unterstützung eigener Recherchestrategien durch Suchbegriffe und Quellenhinweise (42 %): KI wird demnach deutlich häufiger als direkter Lieferant von Informationen denn als Wegweiser zu eigener Recherche genutzt. Hier könnte ein Ansatzpunkt für die künftige Lehre liegen.

6 Wahrgenommene Entwicklung im Umgang mit generativer KI

Fünf Aussagen erfassten, wie die Studierenden ihren Umgang mit generativer KI am Semesterende im Vergleich zum Semesterbeginn einschätzen. Von den befragten 50 Studierenden stimmte jeweils der folgende Anteil den Aussagen „eher“ oder „voll“ zu:

– 72 %: Ich nutze generative KI heute bewusster als zu Beginn dieses Semesters.
– 72 %: Ich setze generative KI heute gezielter ein als zu Beginn dieses Semesters.
– 70 %: Ich habe heute ein besseres Verständnis für die Möglichkeiten und Grenzen generativer KI als zu Beginn dieses Semesters.
– 68 %: Ich kann heute besser einschätzen, welche Formen der KI-Nutzung meinen Lern- und Arbeitsprozess sinnvoll unterstützen.
– 61,2 % (*n* = 49): Ich kann den Nutzen generativer KI für meinen zukünftigen beruflichen Kontext heute besser einschätzen als zu Beginn dieses Semesters.

Bei allen fünf Aussagen überwiegt die Zustimmung deutlich; ablehnend äußern sich je nach Aussage 6 % bis 18,4 %, die übrigen Befragten antworten mit „teils/teils“. Die geringste Zustimmung und zugleich die meiste Ablehnung erhält die Einschätzung des beruflichen Nutzens (61,2 % bzw. 18,4 %). Möglicherweise liegt der berufliche Transfer für viele Studierende noch außerhalb ihres Erfahrungshorizonts, während sich die Lehrveranstaltungen auf Aufgaben im Studium beziehen.

Aus diesen fünf Einzelitems wurde ein Summenscore im Wertebereich von 5 bis 25 gebildet, der die wahrgenommene Gesamtentwicklung quantifiziert (Cronbachs α = ,86; M = 19,27, SD = 4,02). Anhand der Mittelwertunterschiede zeigt sich deskriptiv, dass Studierende umso stärkere Kompetenzsteigerungen bei sich selbst wahrnehmen, je geringer ihre Erfahrung mit generativer KI vor der Lehrveranstaltung war (Tabelle 4).

Tabelle 4
Wahrgenommene Gesamtentwicklung im Umgang mit generativer KI nach Vorerfahrung

Erfahrung vor der Lehrveranstaltung	*n*	*M*	*SD*
gering	5	21,60	2,70
mittel	26	19,35	3,36
hoch	14	19,07	3,60
sehr hoch	4	16,50	8,89
Gesamt	49	19,27	4,02

Anmerkung. Summenscore aus fünf Items (Wertebereich 5–25); höhere Werte stehen für eine stärker wahrgenommene Entwicklung seit Semesterbeginn. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026.

Studierende mit geringer Vorerfahrung erreichen im Mittel 21,60 Punkte, Studierende mit sehr hoher Vorerfahrung 16,50 Punkte. Das Muster ist plausibel: Wer mit wenig Erfahrung beginnt, hat mehr Raum für wahrgenommene Fortschritte, während erfahrenere Studierende eher an eine Obergrenze stoßen. Die Unterschiede sind jedoch statistisch nicht abgesichert (Spearmans $\rho = -,17$, $p = ,25$). Zudem geht der niedrige Mittelwert der Gruppe mit sehr hoher Vorerfahrung wesentlich auf eine einzelne Person zurück, die keinerlei Entwicklung angab (Summenscore 5); die übrigen drei Personen dieser Gruppe liegen zwischen 14 und 24 Punkten. Mit den Nutzungsformen aus Abschnitt 5 hängt die wahrgenommene Entwicklung ebenfalls nicht signifikant zusammen (ρ zwischen ,15 und ,22). Für die Frage, welche Studierenden von angeleiteter KI-Nutzung besonders profitieren, liefert der Pretest damit einen Hinweis auf einen möglichen Ausgleichseffekt zugunsten weniger erfahrener Studierender, der mit künftigen Messungen zu Beginn und am Ende des Semesters zu prüfen ist.

7 Erkenntnisse zum verantwortlichen Umgang mit generativer KI

Eine wichtige Zielsetzung ist es, Studierende zu einem reflektierten Umgang mit generativer KI zu befähigen. Am Semesterende äußerten sich die Studierenden durch ihren Zustimmungsgrad in geschlossenen Fragen sowie in einer offenen Abschlussfrage zu ihrem kritischen Umgang mit den Ergebnissen und zu ihrem bewussten Einsatz generativer KI.

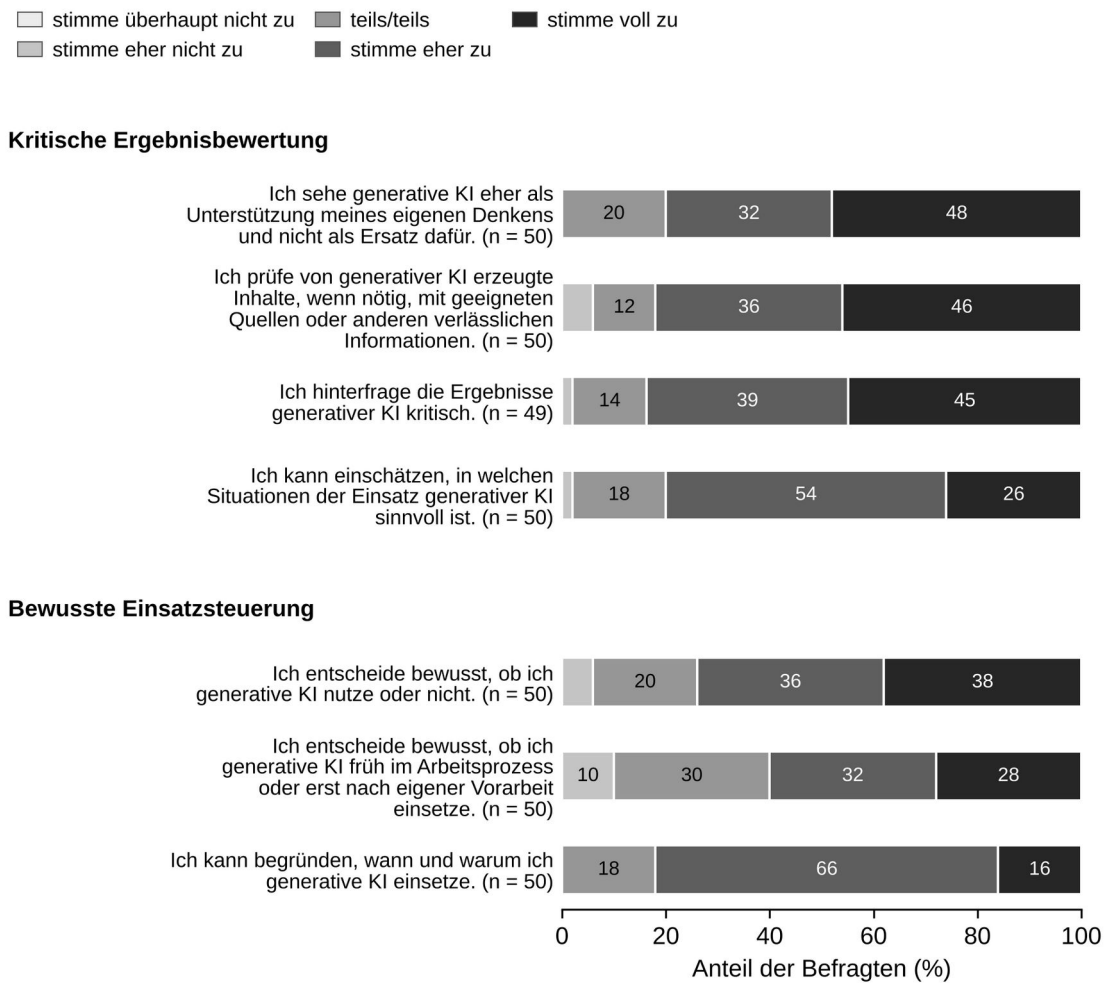
7.1 Kritische Ergebnisbewertung und bewusste Einsatzsteuerung

Der reflektierte Umgang wurde mit sieben Aussagen erfasst. Eine Hauptkomponentenanalyse mit Varimax-Rotation ergab zwei Komponenten, die zusammen 63,0 % der Varianz erklären. Die „kritische Ergebnisbewertung“ umfasst das Hinterfragen und Prüfen von KI-Ergebnissen, die Einschätzung, in welchen Situationen KI sinnvoll ist, und das Verständnis von KI als Unterstützung und nicht als Ersatz des eigenen Denkens. Die Komponente „bewusste Einsatzsteuerung“ bündelt Entscheidungen darüber, ob, wann und warum KI eingesetzt wird. Inhaltlich trennen die Komponenten damit die Bewertung nach der Nutzung von der Entscheidung vor der Nutzung. Abbildung 2 zeigt die Antwortverteilungen.

Den Aussagen zur kritischen Ergebnisbewertung stimmen jeweils 80 % bis 84 % eher oder voll zu. Die Aussagen zur Einsatzsteuerung streuen stärker: 74 % entscheiden nach eigener Angabe bewusst, ob sie KI überhaupt nutzen, und 82 % können ihren KI-Einsatz begründen, wobei hier die eingeschränkte Zustimmung („stimme eher zu“: 66 %) überwiegt. Am wenigsten Zustimmung erhält die Aussage, bewusst zu entscheiden, ob KI früh im Arbeitsprozess oder erst nach eigener Vorarbeit eingesetzt wird (60 %); 40 % antworten hier mit „teils/teils“ oder ablehnend.

Abbildung 2

Reflektierter Umgang mit generativer KI nach Komponenten



Anmerkung. Antwortskala von 1 (stimme überhaupt nicht zu) bis 5 (stimme voll zu); die Kategorie „stimme überhaupt nicht zu“ wurde bei keinem Item gewählt. Angaben in gültigen Prozent; Werte unter 7 % sind nicht beschriftet, Summen können rundungsbedingt von 100 % abweichen. Zuordnung der Items nach der höchsten Ladung in einer Hauptkomponentenanalyse mit Varimax-Rotation (63,0 % erklärte Varianz), innerhalb der Komponenten absteigend nach dem Anteil „stimme voll zu“ sortiert. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026, N = 50.

Auf Indexebene liegt die kritische Ergebnisbewertung (M = 4,20, SD = 0,60; α = ,76) signifikant über der bewussten Einsatzsteuerung (M = 3,94, SD = 0,64; α = ,62), $t(49) = 2{,}91$, p = ,005, d_z = 0,41; beide Indizes korrelieren mit ρ = ,52. Die Befragten sehen sich also eher in der Lage, KI-Ergebnisse kritisch zu bewerten, als den Einsatz von KI im eigenen Arbeitsprozess zu planen. Die geringe interne Konsistenz der Einsatzsteuerung mahnt allerdings zur Vorsicht bei der Interpretation. Zudem bewegen sich die Selbsteinschätzungen nahe der Skalenobergrenze und dürften auch sozial erwünschtes Antwortverhalten widerspiegeln. Mit der Vorerfahrung hängen auch diese beiden Indizes nicht zusammen (ρ = ,08 bzw. ,09).

7.2 Offene Schlussfrage

Die Online-Umfrage endete mit einer offenen Abschlussfrage: „Wenn Sie auf dieses Semester zurückblicken: Was war im Umgang mit generativer KI für Sie besonders

bedeutsam? Gehen Sie dabei gern darauf ein, was sich verändert hat, was Sie gelernt haben, in welchen Situationen Sie sich bewusst gegen einen KI-Einsatz entschieden haben oder welche Aspekte Ihres Umgangs mit KI durch die bisherigen Fragen noch nicht erfasst wurden."

Diese Möglichkeit zur freien Antwort wurde von 20 der 50 Befragten (40 %) genutzt, 15 davon aus derselben Lehrveranstaltung. Die Antworten wurden in Anlehnung an die inhaltlich strukturierende qualitative Inhaltsanalyse induktiv kategorisiert (Kuckartz & Rädiker, 2022); eine Antwort konnte mehreren Kategorien zugeordnet werden. Zur Überprüfung der Kategorienbildung wurde Claude Opus 5.5 (Anthropic) eingesetzt (siehe Hinweis zur KI-Nutzung). Zitate sind mit der laufenden Nummer der befragten Person im Datensatz gekennzeichnet (z. B. B20) und werden unverändert wiedergegeben. Tabelle 5 gibt einen Überblick.

Tabelle 5

Kategorien der Antworten auf die offene Schlussfrage

Kategorie	Inhalt	*n*	%
Gezielter Einsatz und Prompting	genaueres Formulieren von Prompts, Zuweisen von Rollen, Kenntnis verschiedener Anwendungen und ihrer Eignung	11	55
Kritische Prüfung von KI-Ergebnissen	Hinterfragen von Antworten, Gegenprüfen mit anderen Anwendungen oder Quellen, Grenzen der Ergebnisse	8	40
Bewusster Verzicht und eigene Vorleistung	KI erst nach eigener Vorarbeit oder gar nicht einsetzen, etwa bei eigener Reflexion, gestalterischen Aufgaben oder Literaturarbeit	7	35
Rechtliche, ethische und institutionelle Rahmen	Rechtslage, Datenschutz, Klimafolgen, Transparenz, uneinheitliche Haltungen von Lehrenden, Wunsch nach klareren Regeln	7	35
Ambivalenz und Belastung	Mehraufwand durch Prüfen, Überforderung, ungleiche Arbeitsteilung in Gruppen	3	15

Anmerkung. n = Anzahl der Antworten mit mindestens einer Zuordnung zur Kategorie; Prozentangaben bezogen auf die 20 Antwortenden. Mehrfachzuordnungen möglich. Studierendenbefragung zur Nutzung generativer KI, SoSe 2026.

Am häufigsten beschreiben die Studierenden einen gezielteren Einsatz von KI. Sie nennen genaueres Prompting, das Zuweisen von Rollen und die Kenntnis, „wofür welche KI geeignet ist" (B20). Mehrere Befragte heben die Vielfalt der kennengelernten Anwendungen hervor, darunter „Offline-Möglichkeiten" (B12) und Spezialanwendungen, die „wesentlich sinnvoller sind als ChatGPT allein" (B41). Eine Person fasst ihren Lernertrag in drei Worten zusammen: „Struktur, Präzision und Rollenverständnis" (B36).

An zweiter Stelle steht die kritische Prüfung. Eine befragte Person hat gelernt, „die Antworten von KI zu hinterfragen, nicht alles direkt zu glauben" (B9); andere prüfen Ergebnisse mit weiteren KI-Anwendungen oder eigener Internetrecherche gegen (B31, B45). Besonders deutlich wird dies bei wissenschaftlichen Aussagen: „Ich verlasse mich nicht auf literaturbasierte, wissenschaftliche Aussagen von der KI" (B44), weil sich bei der Überprüfung gezeigt habe, dass von dieser Kernaussagen der Literatur ohne entsprechende Anweisung weitergedacht worden seien.

Sieben Antworten beschreiben Situationen, in denen die Studierenden bewusst auf KI verzichten oder sie erst nach eigener Vorarbeit einsetzen. Eine Person hat sich vorgenommen, „KI erst nach meinem eigenen Brainstorming einzusetzen und die Gründe der Nutzung zu hinterfragen“ (B15), was unmittelbar der Kernfunktion der eigenen Vorleistung entspricht. Als Gründe für den Verzicht auf KI nennen die Befragten Aufgaben, „in denen es auf meine eigene, tiefgründige Reflexion ankam“ (B36), die Freude an gestalterischen Tätigkeiten, die eine Person „bewusst NICHT an eine KI abgeben will“ (B20), und die Kontrolle über die eigene Prüfungsvorbereitung (B45). Eine Antwort zeigt eine differenzierte Arbeitsteilung: „Für eher lästige oder sehr routinemäßige Aufgaben nutze ich KI inzwischen häufiger als zu Beginn des Semesters“, bei Themen mit eigenen Schwerpunkten dagegen „eher ergänzend“ (B33).

Ebenso häufig sprechen die Studierenden Rahmenbedingungen an, die der standardisierte Teil des Fragebogens nicht erfasst: rechtliche Fragen (B12, B30), den Wunsch nach Klarheit darüber, „wann ist es okay wann nicht“ (B32), den Schutz eigener Geschäftsideen (B41) und die Klimafolgen von KI (B35). Positiv hervorgehoben wird die Transparenz in Prüfungsleistungen: „Man kann ehrlich sagen wie man zum Ergebnis gekommen ist“ (B42). Kritisch sieht eine Person die uneinheitliche Haltung der Lehrenden: „Da aber viele Dozent:innen KI noch strikt abwehren, ist es schwierig, sich hier darauf einzulassen und wo anders komplett davon weg zu kommen“ (B31).

Einige Antworten sind ausdrücklich ambivalent. Eine Person ist „mittlerweile kein Fan mehr von KI, weil man alles 100 mal überprüfen muss“ (B28), eine andere fühlte sich durch das Thema „sehr überfordert und verunsichert“ (B31). In Gruppenarbeiten wird es als „ernüchternd“ erlebt, „wenn man merkt, dass Mitglieder Ergebnisse nicht hinterfragen“ (B35), weil die Prüfarbeit dann an Einzelnen hängen bleibt.

Die offenen Antworten stützen die standardisierten Befunde: Die kritische Prüfung von KI-Ergebnissen ist für die Studierenden zentral, während die eigene Vorleistung nur selten ausdrücklich als Strategie benannt wird. Darüber hinaus zeigen sie, dass bewusster Verzicht an konkrete Aufgaben gebunden ist und dass rechtliche, ökologische und institutionelle Fragen den Umgang mit KI mitprägen. Da sich vermutlich vor allem besonders engagierte Studierende geäußert haben und drei Viertel der Antworten aus einer Lehrveranstaltung stammen, sind die Kategorien nicht als repräsentative Verteilung für die Gesamtstichprobe zu lesen, sondern als Spektrum von Erfahrungen.

8 Schlussbemerkungen

Der Pretest zeigt, dass Studierende in didaktisch gerahmten Lehrveranstaltungen KI-Anwendungen jenseits von ChatGPT kennenlernen und KI vor allem nutzen, um sich Inhalte zu erschließen und eigene Arbeiten zu strukturieren. Das Muster entspricht den bundesweiten Befunden, nach denen Verständnisfragen und Erklärungen fachlicher Konzepte der häufigste Nutzungszweck sind (von Garrel & Mayer, 2025). Texte für Prüfungsleistungen lässt sich die Mehrheit der Befragten dagegen höchstens gelegentlich erstellen. Ob dies auf die Rahmung der Lehrveranstaltungen zurückgeht oder auf sozial erwünschtes Antwortverhalten, lässt sich mit einer einmaligen Befragung ohne Vergleichsgruppe nicht entscheiden.

Die gefundenen Nutzungsmuster lassen sich auf die Kernfunktionen aus Beitrag 1 (Warnke et al., 2026) beziehen. Die reflexiv-dialogische Nutzung entspricht der Funktion „KI als

Feedback zu eigenen Zwischenständen“; die bewusste Einsatzsteuerung, insbesondere die Entscheidung, KI erst nach eigener Vorarbeit einzusetzen, entspricht der „eigenen Vorleistung vor der KI-Nutzung“. Gerade diese Aussage erhielt im Pretest die geringste Zustimmung (60 %), und auch in den offenen Antworten wird die eigene Vorleistung nur selten als Strategie genannt. Die Befragten bewerten KI-Ergebnisse nach eigenen Angaben also eher kritisch, als dass sie den Zeitpunkt des KI-Einsatzes im Arbeitsprozess planen. Für die Lehre spricht das dafür, die eigene Vorleistung ausdrücklich einzufordern und sichtbar zu machen.

Studierende mit geringer Vorerfahrung berichten deskriptiv die stärkste Entwicklung. Sollte sich dieser Befund im Längsschnitt bestätigen, könnte eine angeleitete KI-Nutzung dazu beitragen, Unterschiede in den Ausgangsvoraussetzungen zu verringern, wie sie für Geschlecht und Bildung berichtet werden (Dicks et al., 2026; Ehlers & Rauch, 2024). Dicks et al. (2026, S. 14) halten es für denkbar, dass sich die Bildungsunterschiede bei der KI-Nutzung am Arbeitsplatz verfestigen, weil akademische Bildung mit breitem Wissen und Urteilsvermögen Grundlagen für die Nutzung von KI vermittelt. Ein urteilsfähiger Umgang mit KI entsteht jedoch nicht von selbst mit einem Studium; er muss in der Lehre gezielt aufgebaut werden.

Die Befunde unterliegen deutlichen Einschränkungen. Die Stichprobe ist klein und nicht zufällig gezogen, knapp drei Viertel der Befragten stammen aus einer Lehrveranstaltung, und Vorerfahrung wie Entwicklung wurden rückblickend am Semesterende erfragt. Gruppenvergleiche, etwa nach Geschlecht, sind deshalb nur als Tendenzen zu lesen. Alle Angaben sind Selbstauskünfte; der hohe Anteil zustimmender Antworten zum reflektierten Umgang deutet zudem auf einen Deckeneffekt hin. Die offene Frage wurde nur von 40 % der Befragten beantwortet.

Für die Weiterentwicklung der Studierendenbefragung ergeben sich daraus mehrere Ansatzpunkte: Erstens erlaubt die Befragung zu Beginn und am Ende des Semesters, die über einen persönlichen Code verknüpft wird, Veränderungen direkt zu erfassen, statt sie rückblickend zu erfragen. Zweitens sollten weitere Lehrveranstaltungen und Studiengänge einbezogen werden, damit sich Lehrformate vergleichen und Gruppenunterschiede belastbarer prüfen lassen; für Analysen nach dem Herkunftsland ist eine heterogenere Stichprobe erforderlich. Nicht zuletzt könnten die Items zum reflektierten Umgang so weiterentwickelt werden, dass sie stärker zwischen den Befragten differenzieren, etwa durch verhaltensnahe Formulierungen, und um Aspekte ergänzt werden, die die offenen Antworten aufzeigen, etwa rechtliche und ökologische Fragen.

Verfügbarkeit von Instrumenten und Daten. Der Fragebogen steht als Zusatzdatei zu diesem Beitrag zur Verfügung. Die Rohdaten werden nicht öffentlich bereitgestellt, weil die Stichprobe klein ist, die meisten Befragten aus einer Lehrveranstaltung stammen und die Freitextantworten Rückschlüsse auf Einzelpersonen zulassen könnten. Aggregierte Auswertungen stellen die Autor:innen auf begründete Anfrage zur Verfügung.

Hinweis zur KI-Nutzung

Bei der Erarbeitung dieses Beitrags haben wir generative KI eingesetzt.

Übersetzen. DeepL.com (Webdienst): Übersetzung des Abstracts.

Datenerhebung und Auswertung. Claude Opus 5.5 (Anthropic), Oktober 2026: Überprüfung der Kategorienbildung auf Basis der offenen Abschlussfrage.

Abbildungen und Tabellen. Claude Opus 5.5, Oktober 2026: Erstellung von Abbildungen und Tabellen.

Weitere Aufgaben. Claude Opus 5.5, Oktober 2026: Formatierung des Beitrags.

Alle Inhalte haben wir geprüft und verantworten sie.

Förderung und Interessenkonflikte

Für diesen Beitrag wurden keine Fördermittel in Anspruch genommen. Die Autor:innen sind Herausgeber:innen von PAIHE. Darin liegt ein Interessenkonflikt, denn im Startverfahren der ersten Ausgabe haben keine unbeteiligten Personen den Beitrag geprüft. Die Befragung fand in Lehrveranstaltungen von Sven Warnke und Harald Bolsinger statt; der Beitrag bewertet damit auch die eigene Lehre der Autor:innen. Weitere Interessenkonflikte bestehen nicht.

Beiträge der Autor:innen (CRediT): Tanja Mühling: Konzeption, Datenkuratierung, Formale Analyse, Methodik, Projektverwaltung, Visualisierung, Schreiben – Erstentwurf, Schreiben – Überarbeitung. Sven Warnke: Konzeption, Durchführung der Untersuchung, Methodik. Harald Bolsinger: Konzeption, Durchführung der Untersuchung, Methodik.

Literatur

Bolsinger, H. J., Warnke, S., & Mühling, T. (2026). The practice of reflective teaching with AI: Future skills development with offline LLMs. *Practice of Artificial Intelligence in Higher Education, 1*(1), Artikel 2. https://doi.org/10.83397/PAIHE.2026.02

Dicks, A., Ehlert, M., Schulz, W., & Solga, H. (2026). KI-Nutzung am Arbeitsplatz: Neue Daten zeigen deutliche Bildungsunterschiede. *WZB Mitteilungen*, (193), 10–14. https://bibliothek.wzb.eu/artikel/2026/f-27779.pdf

Ehlers, U.-D., & Rauch, E. (2024). *KI im Studium aus Studierendensicht: Nutzung, Fähigkeiten und Einstellungen Studierender zu KI. Ergebnisse aus Daten der DHBW-Studierenden-Panelstudie 2024*. NextEducation; Duale Hochschule Baden-Württemberg. https://next-education.org/downloads/2024-11-13_Forschungsbericht_KI-Nutzungsverhalten_Studierender_der_DHBW.pdf

Horstmann, N., Hüsch, M., & Breiter, A. (2026). *Künstliche Intelligenz im Studium – die Sicht von Studierenden im Wintersemester 2025/26* (CHE DatenCHECK 9/2026). CHE Centrum für Hochschulentwicklung. https://hochschuldaten.che.de/kuenstliche-intelligenz-im-studium-die-sicht-von-studierenden-im-wintersemester-2025-26/

Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (5. Aufl.). Beltz Juventa.

von Garrel, J., & Mayer, J. (2023). Artificial intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. *Humanities and Social Sciences Communications, 10*, Artikel 799. https://doi.org/10.1057/s41599-023-02304-7

von Garrel, J., & Mayer, J. (2025). *Künstliche Intelligenz im Studium: Eine quantitative Längsschnittstudie zur Nutzung KI-basierter Tools durch Studierende (2023 & 2025)*. Hochschule Darmstadt. https://doi.org/10.48444/h_docs-pub-533

Warnke, S., Bolsinger, H. J., & Mühling, T. (2026). Leistung ist nicht Lernen: Ein Minimalkonsens für die Gestaltung von Hochschullehre mit generativer KI. *Practice of Artificial Intelligence in Higher Education, 1*(1), Artikel 1. https://doi.org/10.83397/PAIHE.2026.01