

Bericht zum Teilprojekt
Automatische Textproduktion (ATP)
Fokus: Best Practice & Offene Fragen

1	Einleitung.....	2
2	Methodik..... <i>von Julia Steube, Andrea Hunziker Heeb und Maren Runte</i>	3
2.1	Literaturrecherche.....	3
2.2	Kodiersystem.....	3
2.2.1	Deduktive Herleitung des Systems.....	3
2.2.2	Induktive Überarbeitung des Systems.....	4
2.3	Fokusgruppengespräche.....	4
2.4	Transkription und Datenaufbereitung.....	6
2.5	Kodierprozess.....	6
3	Auswertung.....	8
3.1	Quantitative Auswertung aller Codes und Fragebögen..... <i>von Maren Runte, Julia Steube und Andrea Hunziker Heeb</i>	9
3.1.1	Ergebnisse der Fragebogen-Auswertung.....	9
3.1.2	Ergebnisse der quantitativen Auswertung der Codes.....	10
3.2	Best Practice-Aussagen..... <i>von Maren Runte, Julia Steube und Andrea Hunziker Heeb</i>	13
3.2.1	Vorgehen bei der Auswertung.....	13
3.2.2	Erklärung der Annotationen.....	13
3.2.3	Resultate.....	14
3.2.4	Schlussfolgerungen aus der Auswertung der Best-Practice-Beispiele.....	20
3.3	Offene Fragen..... <i>von Reinhard Riedl und Nicole Küffer</i>	22
3.3.1	Vorgehen bei der Auswertung.....	22
3.3.2	Analyse der Art der offenen Fragen.....	22
3.3.3	Analyse der Ebenen Mikro, Meso, Makro.....	27
3.3.4	Analyse der übergreifenden Themen.....	30
3.3.5	Schlussfolgerungen aus der Auswertung der offenen Fragen.....	37
4	Fazit und Ausblick.....	39
5	Anhang.....	43

1 Einleitung

Das hochschulübergreifende Teilprojekt «Automatische Textproduktion» des Projekts «Digital Literacy in University Contexts» der Berner Fachhochschule (BFH), Université de Neuchâtel (UniNe), Pädagogischen Hochschule Zürich (PHZH) und der Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) wurde kurz nach den ersten medialen Berichten zum Thema ChatGPT Anfang 2023 gestartet. Anhand von Fokusgruppengesprächen mit Hochschulangehörigen wurde der Frage nachgegangen, **wie mit dem plötzlichen Aufkommen von Generativer Künstlicher Intelligenz (GenKI), insbesondere von ChatGPT, in der Hochschullehre und vor allem in Bezug auf das akademische Schreiben tatsächlich umgegangen wird.**

Die 18 Gespräche mit insgesamt 66 Personen geben einen Einblick, wie im Sommer 2023 insbesondere Dozierende den Umgang mit ChatGPT in ihrem Unterricht erproben, sie zeigen erste Tendenzen, Ideen und auch Praktiken im Umgang mit ChatGPT in der Hochschullehre auf und auch, welche Fragen sich in dieser Zeit und für die Zukunft stellten.

Um die unterschiedlichen Ebenen der Aussagen von Mitarbeitenden an Schweizer Hochschulen zu berücksichtigen, wurden diese bei der Auswertung aufgegriffen: Auf der Mikro-Ebene (Lehre) steht der individuellen Umgang der Dozierenden mit den neuen Technologien im Fokus, auf der Meso-Ebene (Hochschule) die institutionellen Anpassungen in der Lehre und die Ausrichtung der Prüfungen und auf der Makro-Ebene (Gesellschaft) die gesellschaftlichen und wirtschaftlichen Auswirkungen der KI-Technologie. Dabei wurde explizit die Frage der Prüfungsformen ausgeklammert, dennoch taucht diese Thematik in den Fokusgruppengesprächen und damit in dieser Auswertung immer wieder auf.

Ziel des Projekts und dieses Berichts ist es, Hochschuldozierenden Einblicke zu geben in die Seminarräume und Überlegungen anderer, die Mut machen, Berührungsängste nehmen und inspirieren. Daher stehen bei der Auswertung besonders die Praxis-Beispiele der Dozierenden und die bei den Hochschulangehörigen offenen Fragen im Fokus, da diese ggf. in spezifischen Weiterbildungsangeboten weiterentwickelt und geklärt werden sollten.

Der vorliegende Bericht dokumentiert **in Kapitel 2** die Methodik, beginnend mit der Literaturrecherche, um eine theoretische Grundlage für das zu entwickelnde Kodiersystem zu schaffen. Im Anschluss wird das Kodiersystem erläutert, das in einem iterativen Prozess sowohl deduktiv als auch induktiv erarbeitet und angepasst wurde. Sowohl die Datengrundlage als auch der Kodierprozess werden im Anschluss dargestellt.

Kapitel 3 gibt einen Überblick über die Ergebnisse der kodierten Gespräche und Fragebögen, die einer quantitativen und qualitativen Analyse unterzogen wurden. Nach einer allgemeinen Auswertung der Code-Frequenzen fokussiert der Bericht auf die Auswertung der Best-Practice-Beispiele sowie auf die Auswertung der offenen Fragen, die zentrale Themen und Unsicherheiten im Umgang mit GenKI an Hochschulen aufzeigen.

In **Kapitel 4** wird eine Zusammenfassung der Ergebnisse sowie Empfehlungen für die Integration von GenKI in die Hochschullehre gegeben.

2 Methodik

von Julia Steube, Andrea Hunziker Heeb und Maren Runte

2.1 Literaturrecherche

Im April und Mai 2023 wurde in einem ersten Schritt eine Literaturrecherche in nationalen und internationalen Bibliotheksdatenbanken durchgeführt, um einerseits einen Überblick über die Forschungslage und andererseits eine erste theoretische Grundlage für das zu erarbeitende Kodiersystem zu erhalten, mit dem die Gespräche kodiert und ausgewertet werden sollten. Kriterium bei der Suche war es, möglichst aktuelle Publikationen zusammenzutragen, die eine Peer-Review durchlaufen hatten, mit dem Fokus auf (Generativer) KI und ChatGPT im akademischen Lehren, Lernen und Schreiben. Als Suchwörter wurden *KI/Künstliche Intelligenz*, *ChatGPT*, *Hochschulbildung*, *Universität*, *studentisches Schreiben*, *wissenschaftliches/akademisches Schreiben*, *Hochschullehre* (und englische Entsprechungen) verwendet. Insgesamt 40 Publikationen aus den Jahren 2020 bis 2023 und aus verschiedensten Fachbereichen wurden so ermittelt.

2.2 Kodiersystem

2.2.1 Deduktive Herleitung des Systems

Für die Erarbeitung einer literaturbasierten Grundlage des Kodiersystems wurde der Fokus auf insgesamt 13 Publikationen aus den Bildungs-, Erziehungs- und Geisteswissenschaften sowie Informatik/IT verengt, die eine deutliche thematische Relevanz aufwiesen.¹ Aus diesen Artikeln wurden zentrale Aussagen als Zitate und Paraphasen zusammengetragen und in eine erste thematisch-hierarchische Struktur als Überblick der in der Literatur identifizierten Themen gebracht.² Dabei wurde bereits deutlich, dass es viele thematische Überschneidungen und Mehrfachzuordnungen gibt und nur selten eine eindeutige Kategorisierung der Aussagen möglich war. In einem nächsten Schritt fand eine Reduktion der Belege auf das Wesentliche rund um das Thema GenKI/ChatGPT im *akademischen Lehren, Lernen und Schreiben* statt.

Beim Blick auf die thematisch-hierarchische Struktur wurden quer zu verschiedenen Themenfeldern (z. B. Lernprozess, Schreibprozess, Ethik & Rechtliches), bestimmte Meta- bzw. Querthemen (z. B. Paradigmenwechsel, Desiderate) sowie Anwendungsbereiche für KI (z. B. Unterrichtsvorbereitung, Prüfungen) ermittelt. Diese drei Kategorien – *Einsatzbereich & Nutzer:innen*, *Themen* und *Querthemen* – wurden als ‚Dimensionen‘ und damit erste Organisationsebene des Kodiersystems gewählt. Hinzugefügt wurde zur Schärfung der Analysen hinsichtlich der Projektziele eine vierte Dimension, die das besondere Erkenntnisinteresse an guten Praxis-Beispielen aus der Lehre und Fragen erfassen sollte, die sich die Teilnehmenden in Bezug auf den Umgang mit GenKI im Allgemeinen stellten.

Zur weiteren Ausdifferenzierung des Kodiersystems wurden im nächsten Schritt sämtliche Belege aus der Literatur mit elementaren Codes versehen, die im Stil von Schlagwörtern bzw. kurzen Überschriften aus den Belegen abgeleitet wurden. Diese Codes wurden anschliessend tabellarisch in den vier Dimensionen festgehalten, in Codes und Subcodes systematisiert und mit kurzen

¹ Liste der Publikationen siehe [Anhang A](#).

² Überblick der in der Literatur identifizierten Themen siehe [Anhang B](#).

Beschreibungen versehen. Weitere Codes und Subcodes wurden dort a-priori hinzugefügt, d. h. vor einem ersten Einsatz an den Gesprächen, wo logische Lücken im Kodiersystem ersichtlich wurden.^{[1][SEP]}

So entstand ein bereits umfassendes deduktives Grundgerüst des Kodiersystems, das auf induktive Offenheit angelegt war und eine Erweiterung (bzw. Reduktion) auf Grundlage der Fokusgruppen-Daten ermöglichte, um dem eigentlichen Einsatz an den Gesprächen gerecht zu werden.

2.2.2 Induktive Überarbeitung des Systems

Die induktive Überarbeitung des Kodiersystems erfolgte in mehreren Schritten:

Beim Korrekturlesen der KI-gestützten Transkriptionen der Fokusgruppengespräche (mehr dazu s. Kapitel 3.2) durch die kodierenden Personen wurden Lücken im Kodiersystem auf Ebene der (Haupt-)Codes identifiziert und diese nach zuvor festgelegten Regeln durch das Hinzufügen eines entsprechenden Codes geschlossen. Zudem wurden Beschreibungen und teilweise Benennungen der Codes angepasst. Der gleiche Arbeitsschritt fand für die Ebene der Subcodes parallel zur ersten Runde des Kodierens mit Hauptcodes statt.

Auf diese Weise entstand ein umfangreiches und ausdifferenziertes Kodiersystem, das abschliessend einer korrigierenden Reduktion unterzogen wurde. Dabei wurden Codes zusammengelegt, verschoben oder aus dem System entfernt, wenn sie für die Projektziele wenig oder keine Relevanz aufwiesen.³

Grundsätzliche Herausforderungen beim Erarbeiten des Systems lagen einerseits in der unterschiedlichen ‚Granularität‘ der Subcodes, z.B. *Lernen* vs. *Schreibprozess* oder *Ideenfindung* (im Schreibprozess) und andererseits, Themen von Querthemen zu differenzieren.

Das finale Kodiersystem gliederte sich schliesslich in folgende vier Dimensionen: (1) *Einsatzbereich & Nutzer:innen*, (2) *Themen*, (3) *Querthemen* und (4) *Erkenntnisinteresse des Projekts*. Es enthielt insgesamt 21 Hauptcodes und zunächst 61 darin enthaltene Subcodes und ihre Beschreibungen.⁴ Für die Fokusverengung im Rahmen der Auswertung wurden für die Codes *Offene Fragen* und *Best Practice* (Dimension «Erkenntnisinteresse des Projekts») weitere Subcodes formuliert (s. Kapitel 3.1.2).

2.3 Fokusgruppengespräche

Im Sommer und Herbst 2023 wurden 13 Gruppen- und 5 Einzelgespräche mit insgesamt 66 Teilnehmenden an den vier Projekt-Hochschulen durchgeführt. Ziel der Fokusgruppengespräche war es, verschiedenartige Erfahrungen zum Umgang mit GenKI pro Hochschule zu erheben und dabei durch ähnliche Anspruchsgruppen und die Verwendung eines strukturierten Gesprächsleitfadens über die Hochschulen hinweg vergleichbare Aussagen zu erhalten. Die Gespräche wurden im Herbst 2023 transkribiert und anschliessend korrigiert (s. Kapitel 2.4 und 2.5). Alle Teilnehmenden wurden über die wissenschaftlichen Zwecke informiert und gaben ihr Einverständnis zur Auswertung der Gespräche. Es wurde zugesichert, dass die Daten nur anonymisiert ausgewertet

³ Überblick zur **Überarbeitung des Kodiersystems** s. Ordner [Anhänge Dokumentation](#).

⁴ **Überblick über die Hauptcodes** s. [Anhang C](#), vollständiges **Kodiersystem und Kodierregeln** s. Ordner [Anhänge Dokumentation](#).

werden. Die Teilnehmenden erklärten sich damit einverstanden, dass einzelne kurze Aussagen für Publikationen oder Vorträge pseudonomisiert verwendet werden dürfen. Diese Einverständniserklärung war für alle Fokusgruppen identisch und vorab durch die Rechtsdienste der beteiligten Hochschulen geprüft worden.

Je nach Hochschule setzten sich die Fokusgruppen aus verschiedenen Personengruppen mit zwei bis sieben Teilnehmenden zusammen: Mindestens eine Gruppe pro Hochschule wurde mit Dozierenden durchgeführt, die Sprachunterricht geben (für Übersetzer:innen und Linguist:innen) oder Sprachdidaktik unterrichten (für angehende Lehrer:innen). Eine weitere Personengruppe bestand aus Dozierenden, welche ein technisches Fach unterrichten. Die dritte Personengruppe wurde aus übergeordneten Abteilungen rekrutiert, die meist formell dem Rektorat zugeordnet waren (Bibliothek, Ressort Lehre, Rechtsdienst etc.).

Die Fokusgruppengespräche hatten in der Regel eine Dauer von 60 bis 90 Minuten und wurden mit Einverständnis der Teilnehmenden audiovisuell aufgezeichnet. Diese Aufnahmen waren die Grundlage für die anschliessende Transkription. Aufgrund der Grösse einiger Fokusgruppengespräche waren Videoaufzeichnungen notwendig, um Aussagen eindeutig zuordnen zu können. Das visuelle Material wurde darüber hinaus nicht verwertet.

Die Fokusgruppengespräche wurden von jeweils zwei Mitarbeitenden des Teilprojekts durchgeführt: Eine Person fungierte als Moderator:in, während die zweite Person den Gesprächsverlauf beobachten und Notizen zu Besonderheiten des Gesprächs machen konnte und dadurch die Möglichkeit hatte, vertiefende Fragen zu stellen. Der von den Moderator:innen verwendete Gesprächsleitfaden⁵ sollte sicherstellen, dass die interviewten Gruppen inhaltlich vergleichbare Fragen und Impulse von den Moderator:innen erhielten.

Das Fokusgruppengespräch bestand in der Regel aus vier Teilen. Der erste Teil umfasste die Begrüssung, die Vorstellung der Teilnehmenden untereinander (hieraus wurden die Metadaten zu den einzelnen Personen gewonnen) und eine Erhebung der individuellen Einstellungen gegenüber und Erfahrungen mit ChatGPT mittels eines kurzen Fragebogens⁶, den die Teilnehmenden vor Ort ausfüllten. Ziel dieser Fragebogenerhebung war es, die Teilnehmenden noch ohne Einflussnahme durch den Gesprächsverlauf die Diskussionsschwerpunkte der verschiedenen Institutionen nennen sowie die eigenen drängenden Fragen zum Umgang mit GenKI aufschreiben zu lassen. Im Anschluss wurde durch Multiple-Choice-Fragen erhoben, wie die Teilnehmenden GenKI einsetzen – selbst und im Unterricht.

Für den zweiten Teil wurden zwei inhaltliche Schwerpunkte festgelegt: die Teilnehmenden sollten über die Auswirkungen von GenKI auf die gegenwärtige und zukünftige Unterrichtspraxis (bei Dozierenden) bzw. auf die Hochschule (bei Personen aus dem Rektorat) Stellung nehmen. In einem dritten Gesprächsteil lag der Fokus auf den Risiken und Chancen, die sich aus der Nutzung von GenKI ergeben. Im abschliessenden vierten Teil konnten ergänzende Bemerkungen und Kommentare geäussert werden.

Eine Besonderheit der Fokusgruppengespräche an der UniNE bestand darin, dass die Gespräche nur einen Gesprächsfokus hatten (gegenwärtige Unterrichtspraxis) und mit Mitarbeitenden der UniNE sowie mit eingeladenen Dozierenden in drei Sprachen (DE, EN, FR) geführt wurden. Diese Gespräche sind in der jeweiligen Sprache transkribiert worden. Um mehr Daten von UniNE-Dozierenden selbst zu erhalten, sind im Anschluss noch 5 Einzelgespräche mit ausgewählten Kolleg:innen durchgeführt worden.

⁵ **Gesprächsleitfäden Dozierende und Nicht-Dozierende** s. Ordner Anhänge Dokumentation.

⁶ **Fragebogen Dozierende und Nicht-Dozierende** s. Ordner Anhänge Dokumentation.

2.4 Transkription und Datenaufbereitung

Insgesamt lagen schliesslich 12:21 Stunden Audioaufnahmen von Gruppengesprächen und 1:27 Stunden Audioaufnahmen von Einzelgesprächen von insgesamt 66 Teilnehmenden an 4 Hochschulen vor. Beteiligt waren 14 moderierende Personen. Für 54 Teilnehmende wurden vollständige Fragebögen erhoben mit handschriftlich erfassten Freitextantworten. Für die Einzelinterviews sowie eine Fokusgruppe konnten diese Fragebogendaten nicht erhoben werden.

Nach Durchführung der Fokusgruppengespräche lagen 13 Gespräche mit mehreren Personen (jeweils mindestens eine Stunde Audiomaterial) und 5 weitere Einzelgespräche zur Transkription vor. Aufgrund der Datenmenge sowie der Tatsache, dass bei der Transkription nonverbale und paraverbale Besonderheiten nicht berücksichtigt werden sollten, entschieden wir uns für die automatische Transkription durch das Tool f4x7 mit Hilfe von KI-gestützter automatischer Transkription, bei der zwar die Aussagen, nicht aber etwa prosodische Aspekte der Äusserungen, transkribiert werden. Für die Transkription lag das Einverständnis der Teilnehmenden vor, die vorab in einem Informationsschreiben über die Gesprächsbedingungen informiert worden waren.⁸

Zur Anonymisierung der Daten sind die Gesprächsanfänge (Teil 1) eliminiert worden, in denen sich die Teilnehmenden mit personenbezogenen Informationen vorstellten (wie Name, Rolle, institutionelle Zugehörigkeit, Dienstjahre). Die Hauptteile der Fokusgruppengespräche wurden über die Plattform des Anbieters f4x transkribiert. Die Transkripte wurden im Anschluss von den Projektmitwirkenden, die auch die Kodierungen (s. Kapitel 2.5) vornahmen, auf die korrekte Transkription des Gesagten sowie die Zuordnung der Äusserungen zu den Teilnehmenden überprüft und wo nötig korrigiert. Parallel zur Korrektur der Transkripte identifizierten sie erste thematische Cluster und fassten ihre Beobachtungen in zwei bis drei Sätzen als erste Eindrücke zu den Gesprächsinhalten zusammen.⁹

Die Fragebögen wurden abgetippt, um die Freitextantworten und die Multiple-Choice-Fragen statistisch auswerten zu können.

Die Personendaten aus den Fragebögen wurden in einer gesonderten Datei gesammelt, in der den Teilnehmenden eine Sigle zugewiesen wurde, die in allen weiteren Arbeiten an den Daten verwendet wurde. Die personenbezogenen Daten wurden bei der Kodierung in MaxQDA als Metadaten hinterlegt.

2.5 Kodierprozess

Für die Kodierung der Fokusgruppengespräche wurde mit MaxQDA 2022/24 gearbeitet. Pro Hochschule waren ein bis drei kodierende Personen am Kodierprozess beteiligt. Die ersten Phasen des Kodierens und die induktive Weiterentwicklung des Kodiersystems (s. Kapitel 2.2.2) liefen als parallele Arbeitsschritte.

Folgende Schritte wurden im Kodierprozess durchgeführt:

⁷ Transkribiert wurde mit f4x von der Firma audiotranskription (<https://www.audiotranskription.de>); für die Transkription mit KI wurde die Audiodatei ohne Personendaten auf einen Server geladen, nach Erhalt der Transkription wurde die Datei nach den Datenschutzrichtlinien der Firma von dem Server gelöscht.

⁸ **Informationsblatt** und **Einverständniserklärung** für Gesprächsteilnehmende s. Ordner Anhänge Dokumentation.

⁹ **Themencluster und Gesprächsschwerpunkte** s. Ordner Anhänge Dokumentation.

1. Kodierung der Gespräche nur mit den Hauptcodes (s. Anhang C), dabei blinde Flecken auf Hauptcode-Ebene identifizieren und fehlende Subcodes vorschlagen
2. Besprechen der möglichen Ergänzungen, Streichungen und Zusammenlegungen von Codes und Subcodes im Kodierteam
3. Probe-Kodierung mit Kodiersystem durch dreiköpfige Testgruppe an Gesprächsauszügen und anschliessende Reduktion und Vereinfachung des Systems
4. Vollständige Kodierung aller Gespräche mit aktualisiertem System (Haupt- und Subcodes)
5. Austausch der kodierten Gespräche mit anderen Kodierenden zur Qualitätssicherung.¹⁰

Das hochschulübergreifende Kodierteam befand sich während des gesamten Prozesses in regelmässigem Austausch.

Die Auswertung der Fragebögen erfolgte für die Multiple-Choice-Fragen statistisch, so dass ein für den Sommer/Herbst 2023 aktueller Stand im Umgang mit GenKI an Schweizer Hochschulen erhoben werden konnte. Die Antworten der Teilnehmenden auf die offenen Fragen sind nach Thema der Aussagen geclustert worden, so dass thematische Schwerpunkte gebildet und die Aussagen diesen zugeordnet werden konnten. Eine kurze Darstellung folgt in Kapitel 3.1.1.

¹⁰ Überblick zu den **Kodierenden und Austauschpartner:innen** s. Ordner Anhänge Dokumentation.

3 Auswertung

Aus pragmatischen Überlegungen und mit einem gezielten Fokus auf anwendungsorientierte Resultate für die Zielgruppe (Dozierende und administratives Personal an Schweizer Hochschulen) wurde entschieden, sich im Rahmen dieses Projektes auf die Auswertung der Gesprächssequenzen zu «Best Practice» mit innovativen Beispielen für die Lehre und «Offene Fragen» für allfälligen zukünftigen Handlungs- oder Klärungsbedarf zu konzentrieren. Als «Best Practice» wurden jene Gesprächsabschnitte kodiert, die Schilderungen Dozierender über den Einsatz von GenKI/ChatGPT in der Lehre enthalten, welche später als Best-Practice bzw. konkrete Anwendungsbeispiele genutzt werden könnten. Als «Offene Fragen» waren die Äusserungen von Fragen der Fokusgruppenteilnehmenden zum Umgang oder Einsatz von GenKI/ChatGPT markiert worden.

Für beide Kategorien «Best Practice» und «Offene Fragen» wurden nachträglich Subcodes¹¹ am Material zusammengestellt, um diese differenziert auswerten zu können. In den 68 Textbelegen «Best Practice» wurden 156 Stellen feinkodiert und in den 142 Belegen «Offene Fragen» 208 Stellen. Die Kodierungen wurden einer Qualitätssicherung durch eine zweite kodierende Person unterzogen.

Bei der folgenden Darstellung der Ergebnisse werden Aussagen bzw. gekürzte Aussagen aus den Fokusgruppengesprächen in die Analyse eingebunden, um dadurch die Haltung der Teilnehmenden zu verdeutlichen. In den Beispielen werden Kürzungen in eckigen Klammern und die Dauer von Pausen durch die Anzahl an Punkten in runden Klammern angegeben (1 Punkt entspricht ca. 1 Sekunde Pause).

¹¹ Subcodes zu *Offene Fragen* und *Best Practice* s. Ordner Anhänge Dokumentation.

3.1 Quantitative Auswertung aller Codes und Fragebögen

von Maren Runte, Julia Steube und Andrea Hunziker Heeb

Im Folgenden werden ausgewählte Ergebnisse aus der Fragebogenerhebung sowie aus der Auswertung aller Codes in den Fokusgruppengesprächen vorgestellt.

3.1.1 Ergebnisse der Fragebogen-Auswertung

36 von den insgesamt 54 Personen, welche den Fragebogen zur Nutzung von GenKI an der Hochschule zu Beginn der Fokusgruppengespräche ausgefüllt hatten, gaben an, GenKI bereits in ihre Lehre zu integrieren oder in naher Zukunft in ihre Lehre integrieren zu wollen.

Hierzu muss ergänzt werden, dass unter den 54 Befragten auch Personen waren, welche im administrativen Bereich arbeiten und nicht unterrichten, d. h. dass der Anteil der Dozierenden mit positiver Antwort auf die Frage unter den Befragten vergleichsweise hoch ist. Allerdings haben die befragten Dozierenden an den Fokusgruppengesprächen teilgenommen, weil sie sich grundsätzlich für das Thema interessieren, so dass bei allen Teilnehmenden von einer gewissen Affinität oder zumindest Offenheit dem Thema gegenüber auszugehen ist.

39 der Befragten gaben an, dass sie vom Einsatz von GenKI durch Studierende wissen, entweder weil Studierende im Zusammenhang mit Übungen dazu aufgefordert werden oder weil Dozierende dies aufgrund der schriftlichen Leistungen vermuten.

Bezüglich der Einsatzbereiche gaben insgesamt 18 der Befragten an, einen eigenen Text entweder allein oder gemeinsam mit den Studierenden durch GenKI überarbeiten zu lassen. 17 Personen nutzen zudem die Funktion der KI, um Zusammenfassungen erstellen zu lassen. 16 Befragte gaben an, dass sie GenKI zur Erstellung von Fliesstexten aus Stichpunkten verwendet haben.

Im Gegensatz dazu war die Nutzung von Funktionen wie Predictive Texting oder automatischer Textergänzung weniger verbreitet: Nur zwei Personen gaben an, Predictive Texting zu nutzen, bei dem während des Schreibens Vorschläge für das nächste Wort gemacht werden. Lediglich eine Person verwendete GenKI zur automatischen Textergänzung, d. h. die Funktion, mit der Vorschläge zur Vervollständigung von Sätzen oder Absätzen gemacht werden.

Die Fragebogen-Auswertung zeigt also zusammengefasst, dass GenKI vor allem in der Überarbeitung und Zusammenfassung von Texten sowie in der Umwandlung von Stichpunkten in Fliesstexte genutzt wird. Die geringere Nutzung der Predictive-Texting- bzw. Textergänzungsfunktionen könnte darauf hindeuten, dass diese Funktionen in der Wahrnehmung der Nutzer:innen möglicherweise weniger bekannt und präsent sind.

Eine explizite Frage an die Fokusgruppenteilnehmenden war, was sie für die dringendste Frage bezüglich ChatGPT und den damit produzierten Texten an Hochschulen halten.

Klassifiziert in 5 Themen-Cluster, die immer wieder genannt werden (Bedarf nach Richtlinien, Frage der Deklaration KI-generierter Texte, Fragen zur Nutzung von Tools bzw. zu Veränderungen des Lernprozesses sowie generell zur Leistungsüberprüfung), wird deutlich, dass zum Zeitpunkt der Erhebung im Sommer und Herbst 2023 neben der Frage der Deklaration vor allem die Frage nach den Auswirkungen von GenKI auf den Lernprozess im Fokus steht (vgl. Abbildung 1).

Dringendste Fragen an Nutzung von ChatGPT an Hochschulen

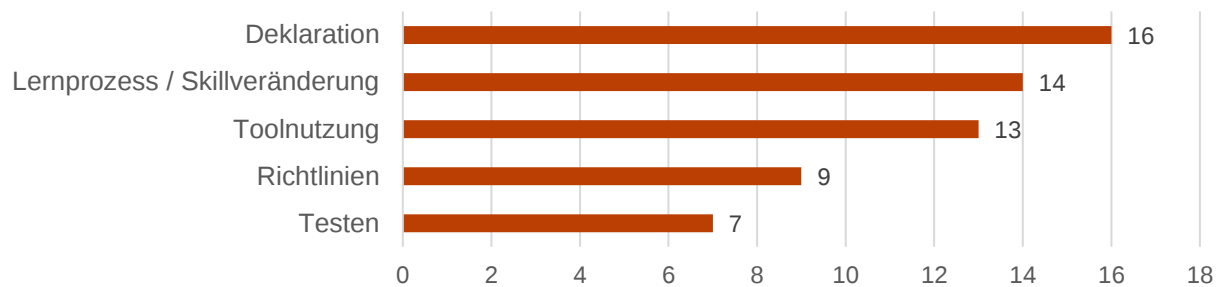


Abbildung 1: Genannte Themen-Cluster (mit absoluter Frequenz) auf die Bitte, die dringendste Frage zur Nutzung von ChatGPT & Co. zu nennen.

3.1.2 Ergebnisse der quantitativen Auswertung der Codes

Wie unter 2.2.2. beschrieben, sind die Fokusgruppengespräche in folgenden vier Dimensionen kodiert worden: *Einsatzbereich & Nutzer:innen (Einsatzbereich)*, *Themen*, *Querthemen* und *Erkenntnisinteressen des Projekts*. Bei der Kodierung wurde darauf geachtet, dass jede Aussage mit möglichst jeder der Dimensionen kodiert wurde – mit Ausnahme der Dimension *Erkenntnisinteresse*. Mit dieser Dimension sind vergleichsweise wenig Aussagen kodiert worden (s. Abbildung 2, 449 von insgesamt 4103 kodierten Aussagen). Insgesamt wurde mit so vielen Codes wie nötig und so präzise wie möglich kodiert, Überlappungen bei der Kodierung waren zudem erlaubt.

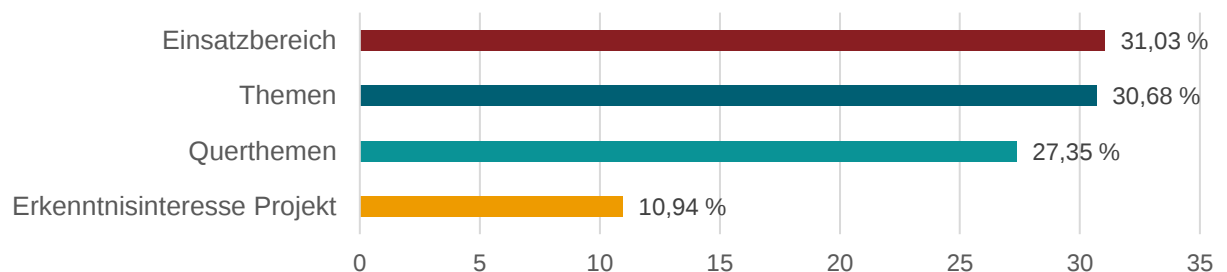


Abbildung 2: Verteilung der kodierten Aussagen auf die 4 Dimensionen (Angaben in Prozent)

Bei der Auswertung der Hauptcodes fällt auf, dass GenKI in der Dimension des **Einsatzbereichs (EN)** besonders im Zusammenhang mit den intendierten Nutzer:innen erwähnt wird, gefolgt vom Einsatz der KI in den Institutionen oder ganz allgemein in der Lehre (vgl. Abbildung 3). Dieses Ergebnis resultiert aus dem Beschluss während des Kodierprozesses, bei der Kodierung immer die Dimension des Einsatzbereichs mitzuberücksichtigen, um so die Aussagen der Dozierendem einem spezifischen Einsatzbereich bzw. einer Nutzer:innen-Gruppe zuordnen zu können.

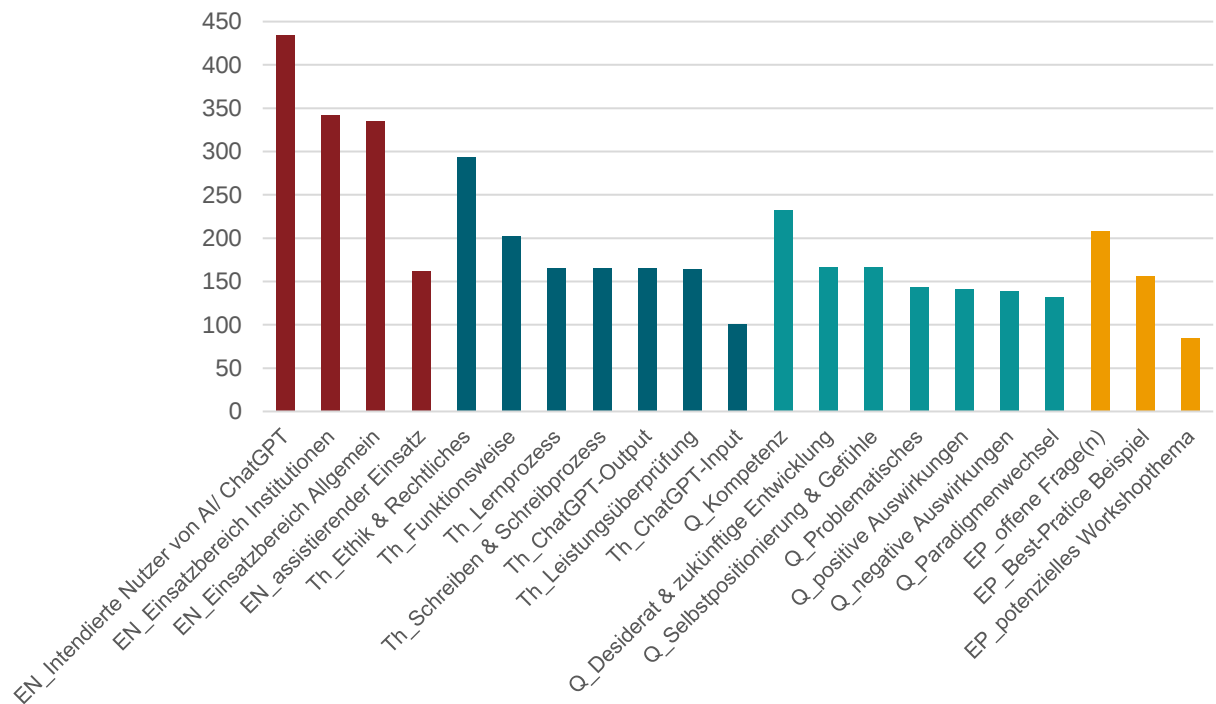


Abbildung 3: Anzahl an kodierten Aussagen pro Hauptcodes gruppiert nach den 4 Dimensionen

In der Dimension **Themen (Th)** sticht in der Auswertung besonders der Hauptcode *Ethik und Rechtliches* heraus, gefolgt von dem Code, welcher die Aussagen in den Fokusgruppengesprächen zur *Funktionsweise* von GenKI auszeichnet. Bei letzterem gibt es wiederum zwei Subcodes, welche häufig angewandt worden sind: Einerseits sind unter *Funktionsweise* die aktuellen Möglichkeiten von GenKI diskutiert worden, aber noch häufiger die *Grenzen der KI*. Auch noch wichtig scheint die Frage, wie transparent die Nutzung von GenKI deklariert werden sollte.

Am stärksten unter den Hauptcodes in der thematischen Dimension ist allerdings die Diskussion der ethischen und rechtlichen Fragen vertreten, welche die Nutzung von GenKI aufwirft. Hier stechen wiederum zwei Subcodes hervor (vgl. Abbildung 4): einerseits das Bedürfnis nach *Schulungen und Aufklärung* über die Funktionsweise bei der Nutzung durch Dozierende, andererseits die Forderung nach *Guidelines* im Umgang mit GenKI.

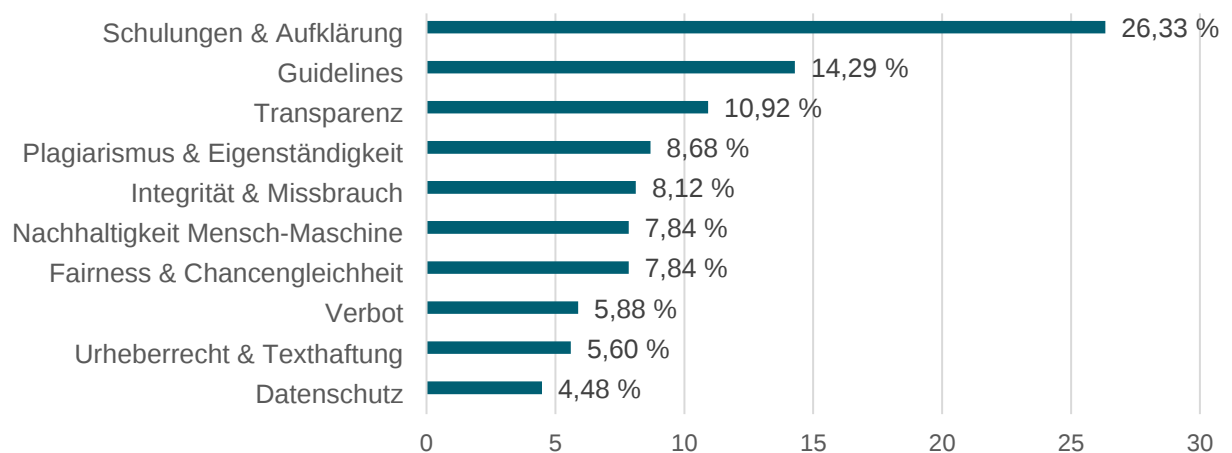


Abbildung 4: Hauptcode Ethik und Rechtliches mit Subcodes und deren Häufigkeit (in Prozent)

In der Dimension der **Querthemen (Q)** wiederum sticht besonders der Hauptcode *Kompetenz* heraus, gefolgt von dem Code *Desiderate und zukünftige Entwicklung* sowie die *Selbstpositionierung* der Befragten bezüglich der Nutzung von GenKI. Interessant an der weiteren Subkodierung des letztgenannten Hauptcodes ist, dass sich bei den Befragten die Selbstpositionierung der Befragten die Waage hält – positive wie negative Aussagen kommen in etwa gleich häufig vor (67 negative Aussagen und 70 positive Aussagen). Diese Ausgewogenheit zeigt sich nicht nur in der Selbstpositionierung, auch kommen Aussagen zu negativen und positiven Auswirkungen von GenKI gleich häufig vor.

Und schliesslich wurden in der Dimension **Erkenntnisinteressen des Projekts (EP)** drei Subcodes verwendet, welche Beiträge zu *offenen Fragen* in Bezug auf GenKI, *Best-Practice-Beispiele* und schliesslich *potenzielle Workshopthemen* umfassen. Die Auswertung der Aussagen zu den beiden ersten Subcodes werden in den Kapiteln 3.2. und 3.3. genauer vorgestellt.

Aus den quantitativen Auswertungen lassen sich bereits wichtige Schlussfolgerungen zur Nutzung von GenKI in der Lehre ziehen:

1. Insgesamt grosse Bereitschaft zur Integration in die Lehre: Trotz der Erwähnung negativer wie positiver Auswirkungen, welche die Nutzung von GenKI nach sich zieht, zeigen die Befragten eine eher hohe Bereitschaft, GenKI in ihre Lehre zu integrieren. Gleichzeitig wird immer wieder erwähnt, dass es ethische und rechtliche Bedenken und einen Bedarf an Guidelines gibt.

2. Fokus auf Überarbeitung von Texten und Bedürfnis nach Schulungen: Eine erste quantitative Auswertung der Codeverteilung zeigt, dass KI-Tools im Befragungszeitraum besonders für die Generierung und die Überarbeitung vorhandener Texte sowie die Erstellung von Zusammenfassungen genutzt werden, dass aber auch das Bedürfnis nach Schulungen über die Funktionsweise von KI-Tools gross ist. Das Kapitel zur Auswertung der Aussagen zu konkreten (Best Practice) Beispielen wird einen vertieften Einblick in Anwendungsbereiche und -praktiken geben.

3.2 Best-Practice-Aussagen

von Maren Runte, Julia Steube und Andrea Hunziker Heeb

3.2.1 Vorgehen bei der Auswertung

Wurde in den Fokusgruppengesprächen über den konkreten Einsatz von GenKI in der Lehre gesprochen, sind diese Abschnitte mit dem Code «Best Practice» kodiert worden (Code-Beschreibung: «Schilderung von Positiv-Beispielen im Einsatz von ChatGPT in der Lehre, die Best Practice veranschaulichen könnten»). Insgesamt wurden 68 Ausschnitte in den 13 Fokusgruppengesprächen und 5 Einzelgesprächen so kodiert. Diese Aussagen stammen von 35 Personen (d. h. nicht alle 66 Teilnehmenden der Gespräche haben solche Aussagen gemacht) und sind aus allen aufgenommenen Gesprächen, so dass Aussagen über die konkrete Einbindung von GenKI in die Lehre als zentrales Thema der Gespräche identifiziert werden kann.

Bei der weiteren Auswertung sind die Aussagen (ggf. in einer Übersetzung für Gespräche in Englisch und Französisch) den Codes, welche die Sprechenden und die verschiedenen Fokusgruppen identifizierten, zugewiesen worden. Aus den Metadaten konnten zusätzlich Angaben zu den Sprechenden (wie Rolle und Funktion und – wenn die Information vorlag – das Dienstalter) ergänzt werden. Die 35 Personen haben ein durchschnittliches Dienstalter von 14 Jahren, wobei die Spanne mit 2 bis 30 Jahren gross ist.

Basierend auf diesen Variablen ist mit den Aussagen weitergearbeitet worden: Einerseits wurden die Codes noch einmal mit spezifischen Subcodes versehen, so dass die Inhalte der Aussagen besser erfasst werden konnten. Andererseits wurde die Sicherheit im Umgang mit GenKI bewertet und entsprechend kategorisiert.

3.2.2 Erklärung der Annotationen

Im iterativen Kodierprozess wurden induktiv Themencluster gebildet, die zu Best-Practice-spezifischen Hauptcodes, Subcodes und Subsubcodes führten. Das Ziel war, die geäusserten Best-Practice- bzw. Anwendungsbeispiele zum Umgang mit KI in der Lehre detailliert erfassen und anschliessend gruppieren zu können.

Der Hauptcode «Unterricht» fokussiert auf die Methoden und Ansätze, die in der Lehre verwendet werden. Die dazugehörigen Subcodes umfassen:

- **Methodenkompetenz/kritische Haltung entwickeln:** Entwicklung unterschiedlicher Einsatzmöglichkeiten von GenKI durch die Studierenden und Prüfung von sinnvollen Szenarien.
- **Offenlegung Richtlinien und Bewertungskriterien:** Transparente Kommunikation der Bewertungsmassstäbe und Richtlinien an die Lernenden.
- **Alternative Unterrichtselemente und Bewertungskriterien:** Einsatz innovativer Unterrichtsformate unter Verwendung von GenKI und dazugehöriger Bewertungsmethoden.

Der Hauptcode «Vor-, Nachbereitung, Prüfungen, Inspiration» befasst sich mit der Vorbereitung und Nachbereitung von Lehre sowie der Inspiration für Dozierende.

Die Subcodes sind:

- **Austausch bezgl. KI-Gebrauch/gemeinsame Kriterien, Vereinheitlichung:** Vorschläge zum Austausch über den Einsatz von GenKI in der Lehre zwischen Dozierenden.
- **Inspiration:** Anregungen und Ideen durch GenKI für Unterrichtsgestaltung
- **Methodenkompetenz/Prompting:** Erwerb von Techniken im Bereich Prompting.
- **Prüfung/Übung erstellen:** Erstellung und Gestaltung von Prüfungen und Übungsaufgaben.
- **Vergleich eigener Wissensstand (Material) mit Output von GenKI:** Analyse und Vergleich von selbst erstelltem Unterrichtsmaterial und -planung bzw. eigenem Wissensstand mit den Ergebnissen, die durch GenKI erzielt werden.

Bei der Auswertung der Aussagen wurde rasch klar, dass die Teilnehmenden bei der Schilderung ihrer Erfahrung mit GenKI inner- und ausserhalb der Lehre unterschiedliche Grade an Sicherheit im Umgang mit GenKI zeigten, was durch den Inhalt selbst, aber auch durch Formulierungen (z. B. Konjunktiv oder Heckenausdrücken) deutlich wird.

Daher wurden die Aussagen in drei Kategorien eingeteilt.

- In **Kategorie 1** machen die Sprechenden ihre Unsicherheit im Umgang mit GenKI sehr deutlich, indem sie z. B. viele Konjunktive verwenden, oder ihren Redebeitrag einleiten mit «Ich könnte mir vorstellen...». Der thematisierte Einbezug von GenKI in die Lehre kann bereits erfolgt sein, es können aber auch Auswirkungen auf die anstehende Lehre formuliert werden.
- In **Kategorie 2** wird von bereits gemachten Erfahrungen berichtet – sowohl in der Lehre wie auch bei der Vorbereitung von Lehre (und in seltenen Fällen auch bei anderen Aufgaben).
- **Kategorie 3** geht darüber hinaus: Hier berichten die Sprechenden von einem bewussten und gezielten (statt experimentellen) Einsatz von GenKI in der Lehre mit konkreten Aufträgen an die Studierenden.

3.2.3 Resultate

3.2.3.1 Thematische Kodierung der Aussagen

45 der 68 Aussagen konnten mit dem Hauptcode «Unterricht», 10 Aussagen mit dem Hauptcode «Vor-, Nachbereitung, Prüfung, Inspiration» versehen werden. Bei 13 Aussagen sind beide Hauptcodes zugeordnet worden (vgl. Abbildung 5). Da es bei vielen Aussagen sinnvoll war, die Inhalte mit unterschiedlichen Subcodes mehrfach zu kodieren, übersteigt die Anzahl der Subcodes die der Hauptcodes bei der folgenden Analyse.

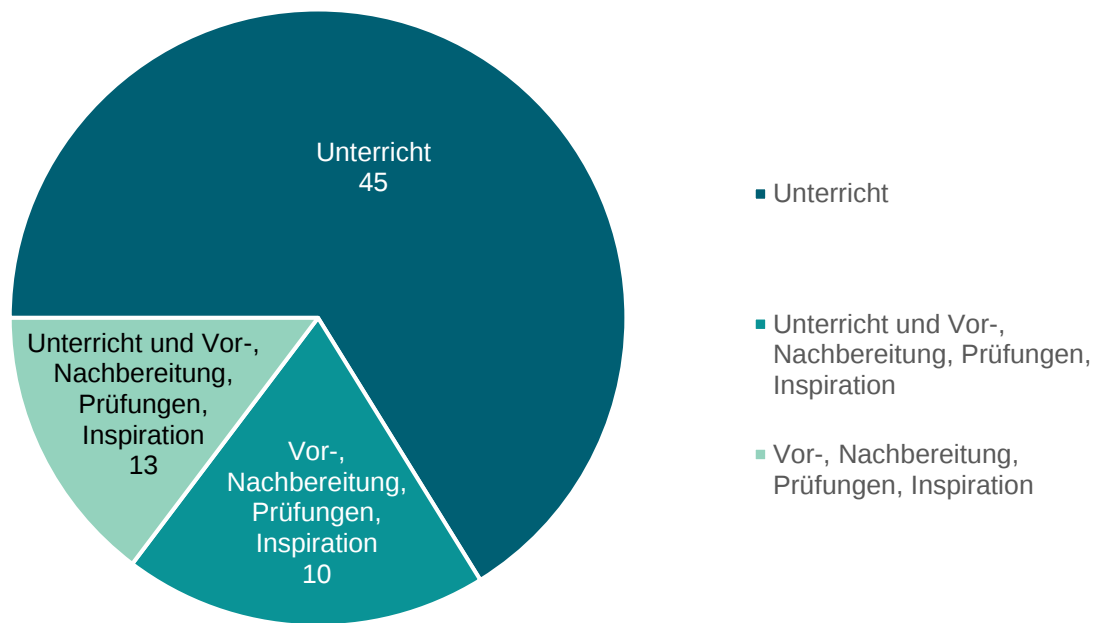


Abbildung 5: Verteilung der 68 Aussagen auf die zwei Subcodes bzw. deren Kombination (absolute Frequenz)

Vor-, Nachbereitung, Prüfungen, Inspiration: Subcodes

Betrachtet man die Subcodes zu «Vor-, Nachbereitung, Prüfungen, Inspiration», fällt auf, dass die GenKI besonders für zwei Tätigkeiten eingesetzt wird (vgl. Abbildung 6): Erstens für die Erstellung von Übungen oder von Prüfungen (10 von 26 Aussagen) und zweitens für den eigenen Aufbau von Wissen im Bereich GenKI (9 von 26 Aussagen), z. B. durch Übungen zum Schreiben von Prompts. Nur zu einem geringen Prozentsatz wird die GenKI hier eingesetzt, um sich Inspiration zu holen.

Mit dem Subcode «Prüfung/Übung erstellen» findet man Aussagen, welche fordern, dass Aufgaben durch die Entwicklung von GenKI anwendungsorientierter sein müssten. Andere Dozierende berichten davon, dass GenKI ihnen hilft, Übungsthemen mit wenig Zeitaufwand zu aktualisieren, fiktive Fallbeispiele oder Vorschläge für die Konkretisierung komplexer Aufgaben zu generieren: «Also ich glaube, was wir nicht mehr machen können, ist ja die generische Aufgabenstellung und [...] was wir ja unsere Studierenden vermitteln wollen, dass sie konkret etwas anwenden können. (.) Ich glaube, (.) da hilft uns ChatGPT fast ein bisschen, weil wir es in Zukunft noch viel mehr machen müssen (.) und dann kann man es nachher wirklich (.) beantworten.»

Ebenso wird GenKI genutzt, um Techniken im Bereich Prompting zu erwerben (Subcode «Methodenkompetenz/Prompting»). Einige Dozierende schildern, dass sie (teilweise gemeinsam mit den Studierenden) Erfahrungen im Prompting sammeln, häufig durch Trial-and-Error-Vorgehen. In diesem Zusammenhang wird auch die Notwendigkeit von Weiterbildungen in diesem Bereich thematisiert.

Anknüpfend an das Prompting wird GenKI auch für Anregungen und zum Brainstorming verwendet, z. B. im konzeptionellen Bereich für neue Lösungswege oder neue Aufgabenstellungen.

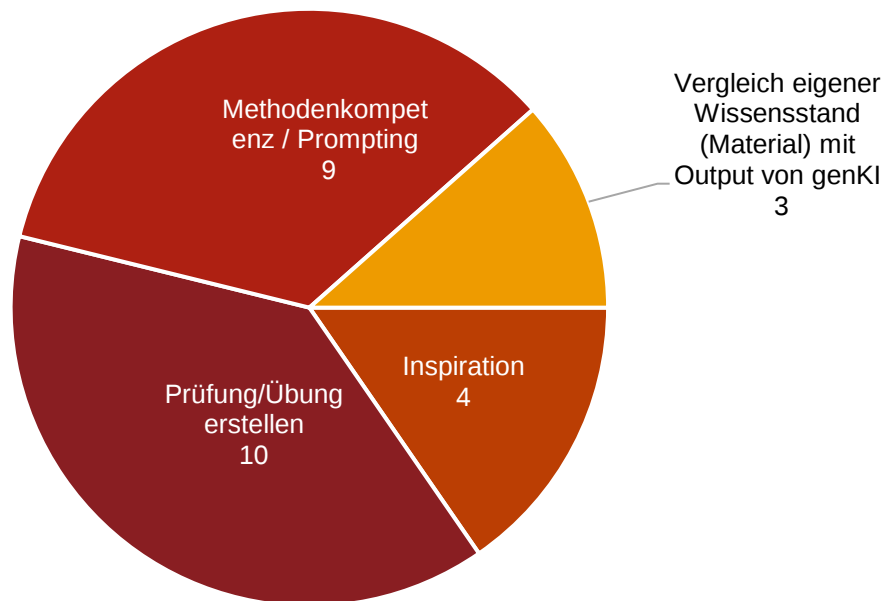


Abbildung 6: Subcodes zu «Vor-, Nachbereitung, Prüfungen, Inspiration» (absolute Frequenz)

Unterricht: Subcodes

Die meisten Aussagen zum Einsatz von GenKI in der Lehre beziehen sich auf den Hauptcode «Unterricht» und hier auf den Subcode «Alternative Unterrichtselemente und Bewertungskriterien» (82 % bzw. 105 von 128 Aussagen, vgl. Abbildung 7), wobei 65 der 105 Aussagen mit weiteren Subsubcodes kodiert sind. Bei all diesen Aussagen geht es um den Einsatz von GenKI in neuen Unterrichtsformaten bzw. den Weg dorthin.

Bei den Subsubcodes treten besonders drei hervor:

- Die Nutzung von GenKI für spezifische Textsorten und Einsatzgebiete wird insgesamt in 21 von 65 Aussagen zum Unterricht erwähnt. So wird GenKI z. B. in der Weiterbildung bei bestimmten Aufgaben als Entlastung beim Schreiben eingesetzt, oder es werden Übungen für Studierende, Feedback für geschriebene Texte oder Zusammenfassung von fertigen Texten generiert.
- In 11 von 65 Aussagen zum Unterricht wird dagegen geschildert, KI zu verwenden, um Feedback zu erhalten: Hier fällt auf, dass die alternativen Unterrichtselemente häufig darin bestehen, dass Texte einmal selbst geschrieben und dann mit KI generiert werden und die Unterschiede im Anschluss kritisch auf verschiedene Kriterien hin überprüft und diskutiert werden sollen (z. B. Stil oder Qualität einer Zusammenfassung). Dies scheint sich als eine häufig verwendete Übung herauszukristallisieren.
- Mit dem Subsubcode «Fakten/Textlogik überprüfen», der in insgesamt in 9 von 65 Aussagen zu Unterricht vorkommt, werden Aufgaben beschrieben, in denen die Studierenden Text mit Hilfe von GenKI generieren, diesen aber durch eigene Faktenrecherche überprüfen müssen: «Ich gebe im Unterricht eine Aufgabe zum Arbeitsrecht und sage ihnen explizit: Schaut mal, was ChatGPT dazu meint. (.) Und dann kommen ganz viele falsche Antworten, weil er nämlich Deutschland, das deutsche Arbeitsrecht und nicht das Schweizer Arbeitsrecht findet. Und dann, an dem zeige ich beispielsweise ein (.) bisschen die kritische Haltung, die man entwickeln sollte.»

Einige Aussagen (22) thematisieren, dass die Ergebnisse von GenKI gemeinsam mit den Studierenden kritisch betrachtet werden müssen (Subcode «Methodenkompetenz/kritische Haltung entwickeln»), was offenbar häufig im Zusammenhang mit der Überprüfung des generierten Inhalts und dessen Passung auf die Aufgabenstellung passiert.

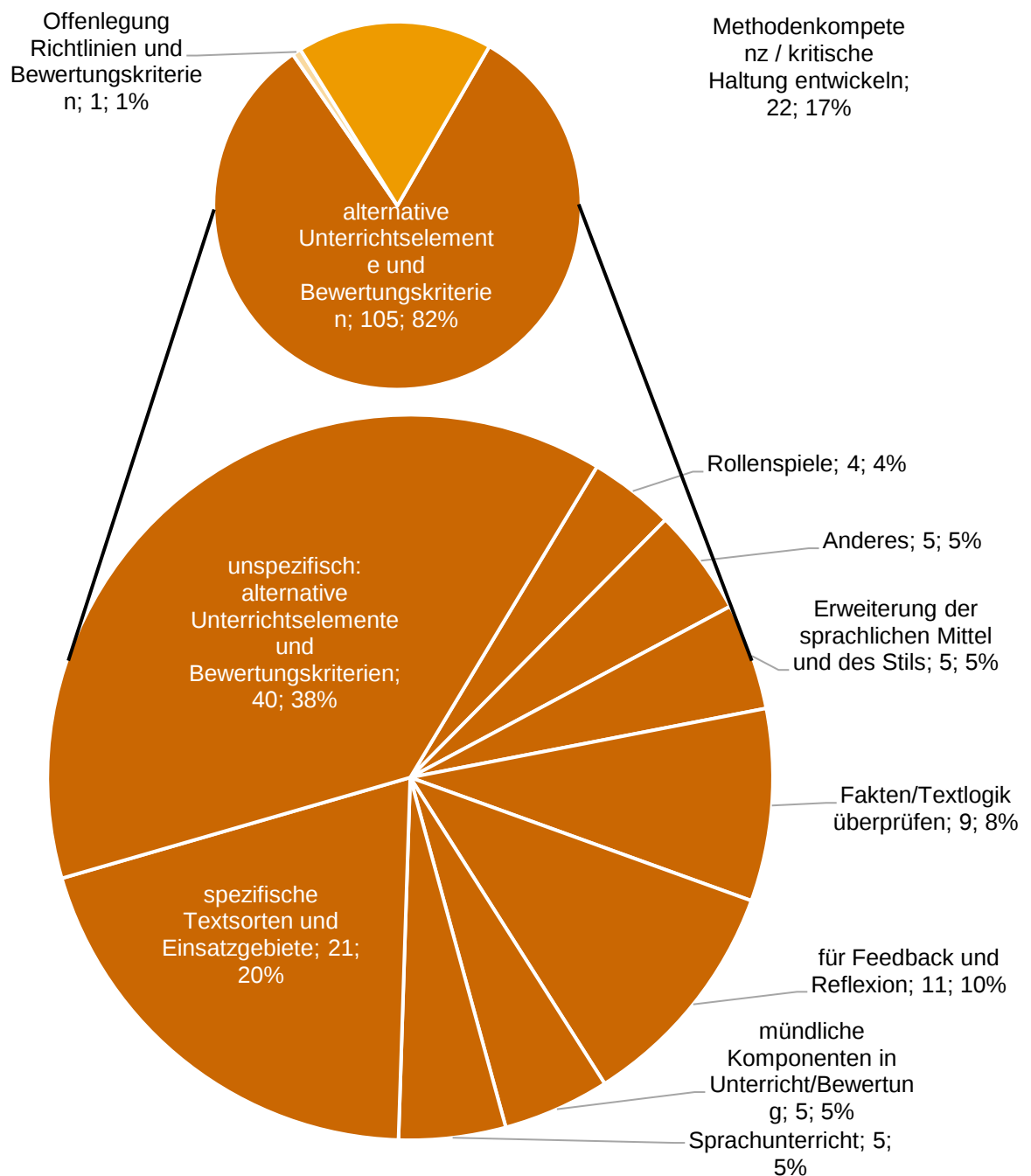


Abbildung 7: Subcodes zu «Unterricht» (Angaben in absoluten Frequenzen und Prozent)

3.2.3.2 Sicherheit im Umgang mit generativer KI

Insgesamt 37 der 68 lassen sich der Kategorie 2 an Sicherheit im Umgang mit GenKI zuordnen, 19 der Kategorie 1 und nur 12 Aussagen der Kategorie 3 (vgl. Abbildung 8), d. h. dass die befragten Dozierenden GenKI vielfach in der Lehre einsetzen, aber noch Unsicherheiten zeigen oder erste Überlegungen anstellen im Umgang mit GenKI.

Zieht man die Metadaten hinzu, ist auffällig, dass Kategorien 1 und 2 Aussagen von Dozierenden enthalten, die ein ähnliches durchschnittliches Dienstalter haben (ca.15 Jahre), wobei in beiden Fällen die Spanne sehr weit ist (2 bis 30 Jahre). Dozierende, die sich im Umgang mit GenKI eher sicherer fühlen und diese gezielt einsetzen (Kategorie 3), haben dagegen ein niedrigeres Dienstalter (durchschnittlich 6 Jahre, ebenfalls mit einer grossen Spanne von 2 bis 22 Jahren). Das Fachgebiet – unterschieden wurde zwischen Sprachdozierenden und Dozierenden aus anderen Fächern – scheint dagegen kaum eine Rolle zu spielen.

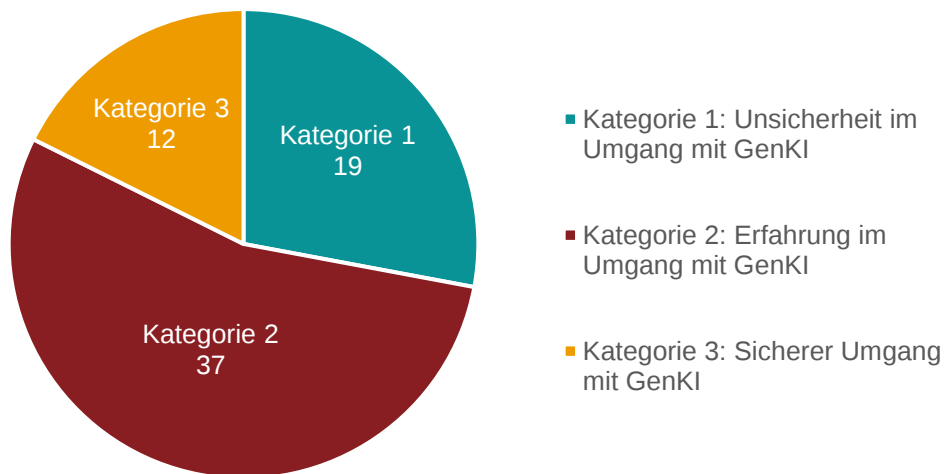


Abbildung 8: Sicherheit im Umgang mit GenKI (absolute Frequenz)

Kategorie 1: Unsicherheit im Umgang mit GenKI

In 19 von 68 Aussagen wird die Unsicherheit der Dozierenden im Umgang mit GenKI deutlich.

Die typischen Aktivitäten, die hier mit allen Unsicherheitsmarkern wie Heckenausdrücken oder Konjunktiven beschrieben werden, bestehen aus dem Experimentieren mit Prompts, um die Funktionalität von KI besser zu verstehen. Häufig testen Dozierende auch den Einsatz von GenKI, um Aufgabenstellungen zu erstellen. Es wird beschrieben, wie z. B. Text- und Fallbeispiele erstellt werden, die dann den Unterricht um praxisnahe Anwendungsszenarien ergänzen könnten: «Ja, vielleicht könnte man sich da auch Fallbeispiele, Textbeispiele (.) kreieren lassen, die ganz bestimmte Probleme beinhalten oder und das(.) quasi als als Übungsmaterial nutzen [...] Auch hier (..) musst du natürlich noch ausprobieren.»

Ebenso wird in dieser Kategorie thematisiert, dass die Integration von GenKI eine Anpassung von Aufgabentypen, Prüfungsformen und Unterrichtsfokus erfordert. Es wird diskutiert, ob man nicht verstärkt wieder auf mündliche Überprüfung setzen sollte, bei der Argumentationsfähigkeit und das Verständnis geprüft werden können: «Vielleicht wird die mündliche Prüfungen jetzt wieder, wieder wichtiger werden an der Stelle, das stimmt. Du kannst dann halt unmittelbar auf Gesagtes eingehen und Rückfragen stellen».

In Präsenzseminaren sollen Studierende Texte ohne KI-Nutzung verfassen, um ihre Kompetenzen unabhängig von automatisierten Tools zu entwickeln.

Ausserdem soll der Fokus auf kritisches Denken und die Überprüfung sachlicher Richtigkeit gelegt werden, während Grammatik und formale Aspekte in Prüfungen weniger gewichtet werden könnte: «So what I'm I'm looking at is changing my assessment to reducing the at the moment

it's a quarter of the marks are for range and accuracy of grammar and vocabulary and I'm going to reduce that and introduce more mediation (.) points.»

In all diesen Aussagen wird der Umgang mit KI als wichtige Fähigkeit angesehen. Es sei aber notwendig, die Nutzung von GenKI mit anderen Seminarinhalten zu koordinieren, um eine umfassende Kompetenzentwicklung zu gewährleisten.

Kategorie 2: Erfahrung im Umgang mit GenKI

Mehr als die Hälfte der mit dem Code «Best Practice» versehenen Aussagen (37 von 68) sind der Kategorie 2 zuzuordnen, d. h. in diesen Aussagen berichten die Dozierenden über bereits gemachte Erfahrungen in der Lehre bzw. bei der Vorbereitung ihrer Lehre.

So nutzen die Dozierenden GenKI, um Unterrichtsmaterialien zu verbessern, Brainstormings durchzuführen, fiktive Debatten für Seminare zu erstellen oder niveaugerechte Hörtexte für L2-Lernende generieren zu lassen.

In der Lehre hat GenKI verschiedene Anwendungsszenarien: Die Dozierenden berichten häufig davon, dass sie Studierende Texte zunächst eigenständig und dann mit GenKI schreiben lassen, um dann mit ihnen die Ergebnisse zu vergleichen (vgl. oben): «Moi de temps en temps je pose des questions, à ChatGPT pour avoir des idées ou oui et puis j'ai fait aussi, enfin je fais produire des textes par exemple, là on a fait (.) encore une lettre de revendication, les étudiants l'ont écrit eux-mêmes et après ils ont donné le même devoir à ChatGPT, ils ont comparé ce genre de choses. Oui, (.) et c'est même amusant de le faire, je trouve. (.) Et puis les étudiants, aiment bien.» Ziel dieser Aufgabe sei es, das kritische Denken zu fördern und die Reflexion über die sprachliche und inhaltliche Qualität der Texte zu ermöglichen.

Eine weitere typische Aufgabe ist es, die Studierenden Fehler in Texten oder Recherchen suchen bzw. die Qualität von Zusammenfassungen bewerten zu lassen. Auch werde zusammen mit den Studierenden das Prompting geübt, um die Ergebnisse bzw. das mit GenKI erstellte Textprodukt zu verbessern.

Bei Leistungsüberprüfungen lassen sich Dozierende mit dieser Sicherheit im Umgang mit GenKI auch Prüfungsfragen erstellen oder holen sich Inspiration für Aufgaben.

Aber auch diese Dozierenden haben Bedenken: So solle der Einfluss von KI bei Prüfungen begrenzt bzw. der Umgang mit KI vereinheitlicht werden.

Kategorie 3: Sicherer Umgang mit GenKI

12 der 68 Aussagen lassen sich der Kategorie 3 zuordnen, also dem bewussten und durchdachten Einsatz von GenKI in der Lehre. Die Dozierenden berichten davon, wie sie GenKI gezielt in Schreibprozesse integrieren lassen, indem Studierende KI-generierte Texte als Ausgangspunkte verwenden und deren Stärken sowie Schwächen erkennen sollen. Dadurch lernen sie, KI bewusst in ihren Workflow einzubinden, etwa um erste Entwürfe zu erstellen oder Schreibideen zu entwickeln.

Um realitätsnahe Aufgabenstellungen in die Lehre zu integrieren, regen die Dozierenden die Studierenden auch dazu an, spezifische Textsorten (z. B. Werkverträge oder E-Mails) durch GenKI verfassen zu lassen.

Einige Fremdsprachen-Dozierende berichten, dass sie GenKI im Bereich des Zweitspracherwerbs bei der Wortschatzarbeit einsetzen. Es helfe, Wortschatzlisten zu generieren oder Texte an das Sprachniveau der Lernenden anzupassen: «Ein Tipp der Woche ist immer mit ChatGPT verbunden und dem Thema verbunden. Zum Beispiel (.) ähm die erste Woche geht es um allgemeine Deskriptoren, von dem common (.) Gemeinsamen Europäischen Referenzrahmen. Wie kann man einen Text ganz einfach adaptieren für differenzierte/ Das ist derselbe Text auf A1 oder A2 oder B1.» Diese Funktionen unterstützen Dozierende dabei, das Lernmaterial zielgerichtet auf die Bedürfnisse der Studierenden abzustimmen.

GenKI wird ausserdem genutzt, um Themen aus unterschiedlichen Perspektiven beleuchten zu lassen, indem die Studierenden das KI-Tool so prompten, dass es bei Antworten unterschiedliche Rollen einnimmt. Auf diese Weise werden komplexe Inhalte abwechslungsreicher gestaltet und die Analysekompetenz der Studierenden gefördert.

Bei der Leistungsüberprüfung wird aber auch hier thematisiert, dass die Inhalte schriftlicher Arbeiten in mündlichen Prüfungen vertieft und die Bewertungskriterien stärker auf systematische und evidenzbasierte Argumentation ausgerichtet werden sollen.

Dozierende binden GenKI auch auf einer Metaebene ein, indem sie die Studierenden dazu anregen, die Funktionsweise und die Ergebnisse von GenKI kritisch zu hinterfragen. Sie stellen verschiedene KI-Tools vor und diskutieren deren Anwendungsmöglichkeiten, um ein Verständnis für den verantwortungsbewussten und reflektierten Einsatz zu fördern.

3.2.4 Schlussfolgerungen aus der Auswertung der Best-Practice-Beispiele

Aus den Ergebnissen lassen sich verschiedene Schlussfolgerungen für den weiteren Umgang mit GenKI in der Lehre ziehen:

1. Zunehmende Integration von GenKI in die Lehre und Bedarf nach Richtlinien

Die hohe Anzahl an Aussagen unter den Befragten über den Einsatz von GenKI in der Lehre zeigt, dass diese Technologie zunehmend als wichtiges Werkzeug betrachtet wird, um Lehr- und Lernprozesse zu bereichern und zu modernisieren.

Gleichzeitig zeigen die Ergebnisse, dass es Bedenken seitens der Dozierenden gibt und es notwendig ist, einheitliche Richtlinien und Standards für den Umgang mit GenKI zu entwickeln, um sicherzustellen, dass akademische und wissenschaftliche Qualitätskriterien gewahrt bleiben.

2. Aktuelle Unsicherheit und wachsender Bedarf an Methodenkompetenz

Die Mehrheit der Dozierenden zeigt Unsicherheiten im Umgang mit GenKI. Viele Aussagen deuten darauf hin, dass eine signifikante Anzahl an Dozierenden Frühlings- und Herbstsemester 2023 GenKI eher experimentell nutzt. Gerade beim Prompting (und daraus schlussfolgernd vielleicht auch zu Funktionsweisen der Tools, um ein vertieftes Verstehen bei Dozierenden und Studierenden zu erreichen) scheint es einen Weiterbildungsbedarf zu geben. Dieser Schritt scheint wichtig, um das volle Potenzial von GenKI im Bildungsbereich zu nutzen.

3. Fokus auf anwendungsorientierte Aufgaben

Der Einsatz von GenKI zur Erstellung von Prüfungen und Übungen mit Blick auf aktuelle und vielschichtige Aufgabenstellungen deutet darauf hin, dass Dozierende auf diese Weise anwendungsorientierte Lerninhalte bieten können. Die Studierenden profitieren so von realistischen Aufgabenstellungen, was ihre Problemlösungsfähigkeiten fördert.

4. Kritische Reflexion und Feedbackprozesse

Die Verwendung von GenKI zur Erstellung und Überprüfung von Texten fördert eine kritische Auseinandersetzung mit dem Text. Besonders in Modellen, die GenKI in den Schreibprozess einbinden und diese dann von den Studierenden kritisch reflektieren lassen, wird nicht nur die Qualität der Texte erhöht, sondern auch die Analyse- und Reflexionsfähigkeiten der Studierenden gefördert. Insgesamt kann das zur Entwicklung einer verantwortungsvollen Nutzung von Technologie im Bildungsbereich beitragen.

5. Vielfältige Einsatzmöglichkeiten und innovative Lehrformate

Die Ergebnisse deuten darauf hin, dass GenKI in verschiedenen Kontexten eingesetzt wird, von der Erstellung spezifischer Textsorten, neueren oder anwendungsorientierteren Aufgabenformen bis hin zu neuen Bewertungsmethoden. Dies fördert die Entwicklung innovativer Lehrformate, die Studierende wiederum auf zukünftige Arbeitsfelder und -formen, die mit hoher Wahrscheinlichkeit auch den Einsatz von GenKI umfassen, vorbereiten können.

3.3 Offene Fragen

von Reinhard Riedl und Nicole Küffer

3.3.1 Vorgehen bei der Auswertung

In den Fokusgruppengesprächen und Einzelgesprächen wurden nicht nur Aussagen über KI in der Lehre oder im wissenschaftlichen Schreiben getroffen, sondern auch ungeklärte Fragen von den Teilnehmenden zum Thema formuliert. Diese Ausschnitte aus den Gesprächen wurden mit dem Code «offene Frage(n)» kodiert. (Code-Beschreibung: Äusserung einer Frage zum Umgang oder Einsatz von AI/ChatGPT). Insgesamt wurden 142 Ausschnitte in 18 Gesprächen so kodiert. Diese Aussagen stammen von 35 Personen und stammen aus 17 von 18 aufgenommenen Fokusgruppengesprächen, so dass offene Fragen zu KI in der Lehre als zentraler Gegenstand der Fokusgruppengespräche identifiziert werden konnte.

Im iterativen Kodierprozess wurden induktiv Cluster gebildet, die den Hauptcode «offene Frage(n)» in 13 Subcodes unterteilen. Diese Subcodes decken verschiedene Arten von offenen Fragen ab und beleuchten spezifische Aspekte und Perspektiven mit dem Ziel, die geäusserten Fragen zum Umgang mit KI in der Lehre detailliert erfassen und gruppieren zu können.

Für die weitere Auswertung wurden die als offenen Fragen kodierten Segmente zusätzlich den Ebenen Mikro, Meso und Makro zugeordnet und grafisch dargestellt. Es wurde deutlich, dass bestimmte Themen übergreifend in mehreren Subcodes auftreten. Um diese übergreifenden Themen zu erfassen und ihre Bedeutungsvielfalt darzustellen, wurden die Aussagen zusätzlich in sechs Themenfelder eingeordnet.

Alle Auswertungen in diesem Teil beziehen sich ausschliesslich auf Aussagen, die im Kodierprozess als offene Fragen kodiert wurden. Das heisst, dass alle Aussagen aus einem Fragekontext stammen, unabhängig davon, ob sie fragend oder affirmativ getätigt werden.

3.3.2 Analyse der Art der offenen Fragen

Der Hauptcode «offene Frage(n)» bestehend nach der ersten induktiven Kodierung aus 13 Subcodes schlüsselt die geäusserten Fragen nach deren Art auf. Die 142 Aussagen wurden mehrfach kodiert und resultierten in 174 Kodierungen (vgl. Abbildung 9). Die Unterteilung in 13 Subcodes deutet darauf hin, dass die Fragen zur KI in der Lehre ein breites Spektrum an Aspekten erfassen und unterschiedliche Perspektiven berücksichtigen.

Die 13 Subcodes sind wie folgt definiert:

- **Ansprüche an Dozierende:** Ansprüche im Umgang mit und Einsatz von GenKI an Dozierende.
- **Ansprüche an Studierende, z. B. Eintritts- und Austrittskompetenzen:** Ansprüche im Umgang mit und Einsatz von GenKI an Studierende sowie veränderten Ansprüchen an ihre Kompetenzen fürs Studium und Beruf.
- **Bewertungskriterien/Leistungsnachweise:** Einfluss von GenKI auf die (bisherigen) Bewertungskriterien und Leistungsnachweise.

- **Change Management:** Offene Fragen zur Gestaltung des Change Managements.
- **Einsatzbereiche:** Offene Fragen zum sinnvollen Einsatz von GenKI.
- **Ethik / Recht:** Offene Fragen in Bezug auf Ethik und Recht beim Einsatz von GenKI.
- **Gesellschaftlicher Umgang, Wert der Schriftlichkeit und menschliche Kommunikation:** Offene Fragen dazu, was in der Gesellschaft diskutiert werden muss, Umgang mit GenKI, Wert der Schriftlichkeit und menschlicher Kommunikation im KI-Zeitalter.
- **Literacy / Future Skills:** Offene Fragen zum Erwerb der nötigen KI-Literacy durch Dozierende und Studierende sowie offene Fragen zu den gefragten / nötigen Kompetenzen in der Zukunft.
- **Richtlinien:** Offene Fragen zum Thema Richtlinien.
- **Rolle der Dozierenden:** Offene Fragen zur Rolle der Dozierenden in der Zukunft.
- **Rolle / Wert des Studiums:** Offene Fragen zur Rolle und zum Wert eines Studiums in der Zukunft
- **Technische Kenntnisse:** Offene Fragen zum technischen Hintergrund von GenKI.
- **Zukünftige Entwicklung:** Arbeitsprozess / Auswirkungen: Offene Fragen zur Eingliederung von GenKI-Nutzung in Arbeitsprozesse und zu den generellen Auswirkungen der GenKI-Nutzung im Studium.

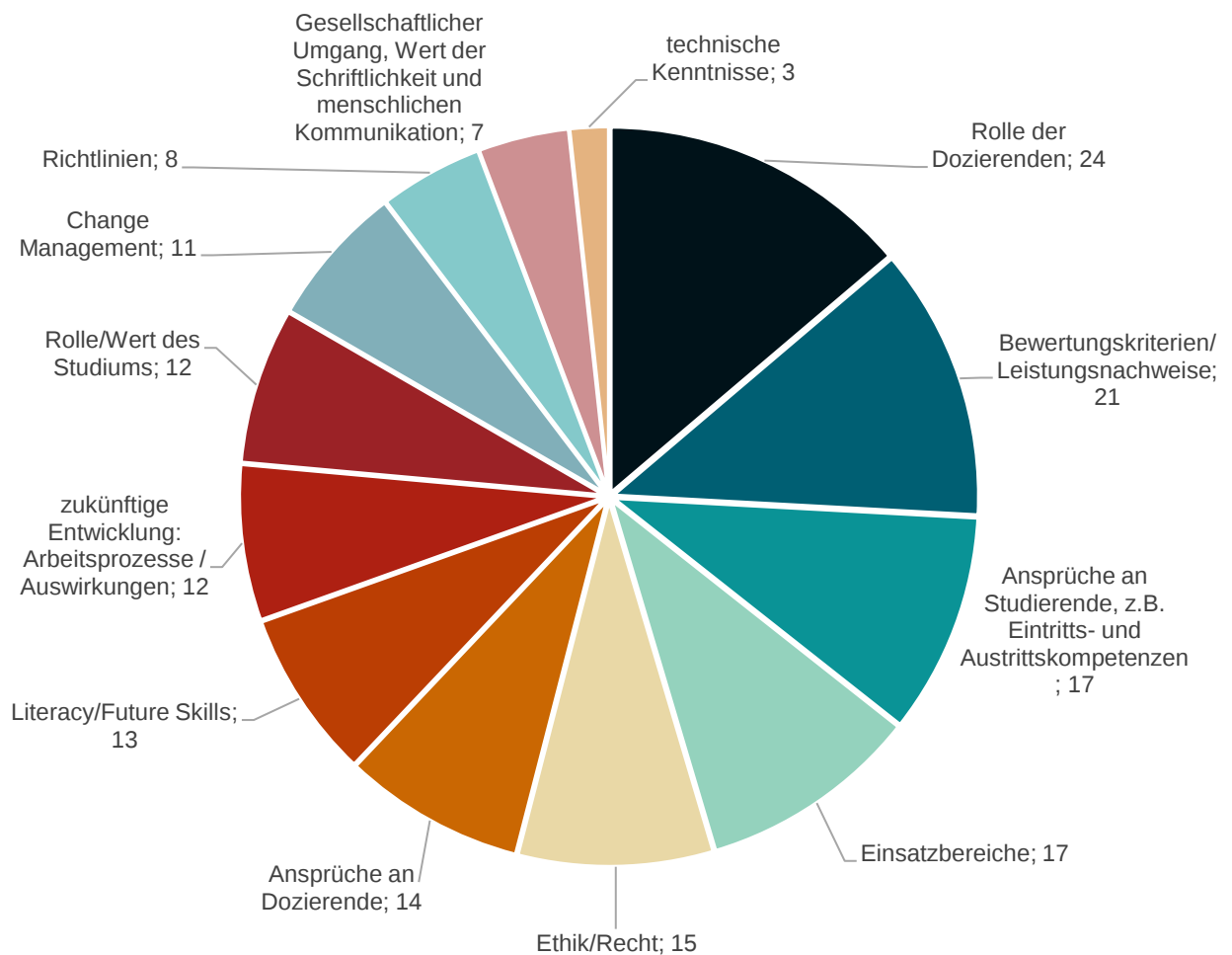


Abbildung 9: Verteilung der 142 Aussagen auf die 13 Subcodes mit Mehrfachkodierung

3.3.2.1 Dozierende im Fokus

Am häufigsten (24 Aussagen) äusserten die Sprechenden Fragen zu ihrer eigenen Rolle. Der Code «Rolle der Dozierenden» umfasst dabei vor allem Aussagen zur mittelfristigen Entwicklung und bietet Überlegungen dazu, wie sich die Rolle der Dozierenden in Zusammenarbeit mit KI in der Lehre potenziell entwickeln könnte. Allerdings beschäftigt die Dozierenden ihre Rolle auch schon in der Gegenwart. Oft wird überdies die Rolle in der gegenwärtigen und unmittelbaren Zukunft implizit thematisiert. Tatsächlich reflektieren die Teilnehmenden in den Fokusgruppengesprächen sehr breit, welche Rolle und Verantwortung sie in den verschiedenen Bereichen tragen. Indirekte Aussagen dazu tauchen dementsprechend bei allen 13 Subcodes auf.

Eng mit dem Subcode «Rolle der Dozierenden» verbunden ist der Subcode «Ansprüche an Dozierende» (14 Aussagen), in dem viele Aussagen den Eindruck vermitteln, dass die Ansprüche an Dozierende unklar sind, und insbesondere auf unklare Anforderungen seitens der Arbeitgeber (d. h. der Hochschulen) hingewiesen wird. Aufgrund der Vielzahl an Aspekten und Perspektiven zur Rolle der Dozierenden wurde dieser in der erneuten induktiven Kodierung für die themenübergreifenden Subcodes berücksichtigt und aufgenommen. Die «Rolle der Dozierenden» wird daher ausführlich in Kapitel 3.3.4 behandelt.

3.3.2.2 Schriftlichkeit und Leistung im Fokus

Fragen zu «Bewertungskriterien/Leistungsnachweisen» wurden mit 21 Aussagen am zweithäufigsten thematisiert. Ein Grossteil dieser Fragen bezieht sich darauf, wie Kompetenzen künftig geprüft werden können und wie mit dem Einsatz von ChatGPT bei Leistungsnachweisen umgegangen werden soll. Das deutet auf eine beträchtliche Unsicherheit in diesem Bereich hin.

Besonders häufig wurden Befürchtungen geäussert, dass nicht mehr die eigentlichen Kompetenzen, sondern lediglich die Fähigkeit zur Nutzung von ChatGPT überprüft werden könnte. Als mögliche Alternativen wurden vermehrte mündliche Prüfungen oder eine Anhebung der Anforderungen diskutiert. Beides wurde jedoch bis zum Zeitpunkt der Gespräche nicht praktisch umgesetzt.

Die Frage nach der Veränderung des Anforderungsniveaus findet sich auch im Subcode «Ansprüche an Studierende, z. B. Eintritts- und Austrittskompetenzen» (17 Aussagen) wieder. Einige Aussagen hinterfragen, welche Kenntnisse und Fähigkeiten künftig in der Wirtschaft benötigt und nachgefragt werden und wie sich daraus resultierend die Anforderungen im Studium und an Studierende verändern könnten. Insgesamt entsteht bei den Aussagen aus diesem Subcode der Eindruck, dass die Sprechenden eine Anhebung der Anforderungen für Studierende erwarten. Dieser Gedanke wird durch den Subcode «Rolle/Wert des Studiums» (12 Aussagen) ergänzt, der Fragen dazu enthält, welchen Mehrwert eine Hochschule bietet, wenn mit ChatGPT Informationen jederzeit abrufbar sind und Prüfungen mit dessen Unterstützung zumindest ausreichend gelöst werden können. Die Anhebung des Niveaus erscheint also als eine Möglichkeit oder auch Notwendigkeit, der Entwertung – respektive dem Sinnverlust – entgegenzutreten.

Zentral ist in diesen Überlegungen die Fortführbarkeit der etablierten Prüfungspraktiken. Neben den konkreten Fragen aus dem Subcode «Bewertungskriterien/Leistungsnachweis» gehen die Überlegungen und Fragen teils weit darüber hinaus und hinterfragen insbesondere bei schriftlichen Leistungsnachweisen den Wert der Schriftlichkeit an der Hochschule und in der Gesellschaft. Dieser Aspekt wird auch im Subcode «Gesellschaftlicher Umgang, Wert der Schriftlichkeit und menschliche Kommunikation» (7 Aussagen) aufgegriffen. Aussagen aus diesem Subcode lassen vermuten, dass der Wert der Schriftlichkeit vor allem im Verhältnis zum Wert der mündlichen Kommunikation gesehen wird. Da der «Wert der Schriftlichkeit» aus unterschiedlichen Perspektiven hinterfragt wird und implizit in verschiedenen Subcodes angesprochen wird, wurde dieser Aspekt in die Themenfelder aufgenommen und wird in Kapitel 3.3.4 diskutiert.

3.3.2.3 KI-Einsatz und Kompetenzen im Fokus

Die meisten konkreten Fragen zur Nutzung von KI, insbesondere zur Frage des «Wie», stammen aus dem Subcode «Einsatzbereiche» (17 Aussagen). Die Fragen lassen darauf schliessen, dass anwendbare Beispiele für den sinnvollen Einsatz von KI in der Lehre noch nicht etabliert sind und nach praktikablen Ansätzen gesucht wird, die einen bedeutenden Mehrwert entweder für Studierende oder Dozierende schaffen. Zudem werden in diesem Subcode auch Fragen aufgeworfen, die sich auf die zukünftigen Fähigkeiten von Dozierenden und Studierenden beziehen.

Detaillierte Fragen zu diesen zukünftigen Fähigkeiten finden sich in den Subcodes «Literacy/Future Skills» (13 Aussagen) und «technische Kenntnisse» (3 Aussagen). Aus den Fragen zu Future Skills lässt sich ableiten, dass Dozierende einerseits darüber nachdenken, welche Kompetenzen sie selbst künftig benötigen, und andererseits, welche Fähigkeiten an Studierende vermittelt werden sollten. Das Thema wird also von zwei Seiten betrachtet.

Das kritische Denken als Kompetenz spielt dabei eine wichtige Rolle, wobei dieser Aspekt eher einseitig als Herausforderung für die Studierenden, respektive für deren Ausbildung durch die Dozierenden, gesehen wird. Es kommt auch in anderen Subcodes vor und beinhaltet verschiedene Perspektiven. Die Aussagen der Sprechenden lassen darauf schließen, dass kritisches Denken als zentrale Kompetenz in der Zukunft gesehen wird. Aussagen wie «Was passiert mit dem kritischen Denken?» oder «Wie kann den Studierenden kritisches Denken beigebracht werden im Umgang mit ChatGPT?» weisen auf die Befürchtung hin, dass die Entwicklung dieser Kompetenz durch KI gestört werden könnte.

Aufgrund der vielen Aspekte und Perspektiven, die um das kritische Denken auftauchen, wurde «Kritisches Denken» als themenübergreifendes Thema in die Themenfelder aufgenommen und wird ebenfalls in Kapitel 3.3.4 breit besprochen. Im Gegensatz zum Subcode «Literacy/Future Skills» befassen sich die Fragen zu den technischen Kenntnissen eher mit grundlegenden Aspekten der Technologie, insbesondere mit dem Funktionsumfang von ChatGPT und der Funktionsweise von GenKI im Allgemeinen.

Eine tiefer gehende inhaltliche Analyse der «Funktionsweise & Nutzungspraxis» erfolgt in Kapitel 3.3.4 im Rahmen der übergreifenden Themenfeldern.

3.3.2.4 Ethik und Richtlinien im Fokus

Nicht zuletzt wurden auch Fragen zu Ethik und Recht aufgeworfen, die im Subcode «Ethik/Recht» (15 Aussagen) zusammengefasst sind. Die Vielfalt der Fragen in diesem Bereich deutet auf erhebliche Unsicherheiten hin. Beispiele hierfür sind: «Wer trägt die Kosten für die Nutzung von ChatGPT – bezahlen wir alle mit unseren Daten?», «Was geschieht, wenn bei Prüfungen nicht nachgewiesen werden kann, dass eine Arbeit mithilfe von ChatGPT erstellt wurde?», «Können Dozierende rechtlich belangt werden?», «Ist eine von ChatGPT kopierte Arbeit nicht einfach ein Plagiat?», «Wird durch den Einsatz von ChatGPT die akademische Integrität gefährdet?», «Dürfen Studierende verpflichtet werden, ChatGPT zu nutzen? Und wie verhält es sich mit dem Datenschutz?», etc. Aus dem Kontext der Aussagen, aus denen die Fragen stammen, lässt sich ein Wunsch nach Richtlinien ableiten. Konkrete Anforderungen zu Richtlinien bleiben allerdings in den Aussagen aus. Dieser Wunsch wird auch in den Aussagen des Subcodes «Richtlinien» (8 Aussagen) widerspiegelt.

3.3.2.5 Zukunft und Wandel im Fokus

Der Subcode «zukünftige Entwicklung: Arbeitsprozesse/Auswirkungen» (12 Aussagen) umfasst Fragen und Überlegungen zur künftigen Integration von ChatGPT und KI in die Arbeitsprozesse von Hochschulen sowie den Studienalltag. Wiederholt wird auch der Wert der Schriftlichkeit thematisiert, insbesondere die Frage, ob Schreiben in nicht-akademischen Berufen lediglich als Mittel zum Zweck ausreichend ist und ob ein grundlegendes Mass an Schriftlichkeit genügen könnte. Im Subcode «Change Management» (11 Aussagen) wird zudem diskutiert, wie dieser Wandel gestaltet und unterstützt werden kann. Die meisten Aussagen weisen darauf hin, dass eine breite Akzeptanz besteht: ChatGPT und KI werden als Technologien angesehen, die bleiben werden und daher eine Weiterentwicklung sowohl des eigenen Unterrichts als auch der Hochschule erforderlich machen. Für den Prozess der Anpassung wird nach Möglichkeiten zum Experimentieren gefragt, um neue Ansätze zu erproben. Gleichzeitig wird die Einbindung der Studierenden als

eher gewünscht angesehen, da vermutet wird, dass sich auch ihr Studienalltag durch diese Entwicklungen grundlegend verändern könnte.

Insgesamt lassen die Aussagen darauf schliessen, dass Antworten auf Fragen zu KI in der Lehre auf individueller Ebene für den eigenen Unterricht, auf Hochschulebene, sowie auf gesellschaftlicher Ebene gesucht werden. Die Einteilung in Ebenen wird im nächsten Kapitel analysiert.

3.3.3 Analyse der Ebenen Mikro, Meso, Makro

In einem weiteren Schritt wurden die 142 Aussagen in die drei Ebenen Mikro, Meso, Makro eingeteilt, mit dem Ziel, differenzierte Einblicke in die unterschiedlichen Perspektiven auf Unterrichts- und Kursebene (Mikro), in Bezug auf die Rolle der Dozierenden und die Hochschule (Meso) sowie auf gesellschaftlicher oder bildungspolitischer Ebene (Makro) zu gewinnen.

Bei der Zuteilung der 142 kodierten Aussagen auf die drei verschiedenen Ebenen wurden 37 Aussagen der Mikro-Ebene, 60 Aussagen der Meso-Ebene, und 45 Aussagen der Makro-Ebene zugeordnet (vgl. Abbildung 10).

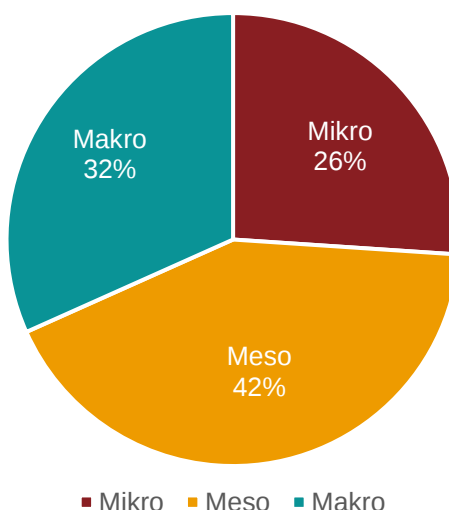


Abbildung 10: Verteilung der 142 Aussagen auf die drei Ebenen Mikro, Meso, Makro

Um zu verstehen, welche offenen Fragen auf den verschiedenen Ebenen geäußert wurden, wurden aus den 142 Aussagen konkrete Fragen abgeleitet und zusammengefasst. Daraus resultierten 45 Fragen, die grafisch im Kreis der drei Ebenen Mikro, Meso, Makro abgebildet werden können (vgl. Abbildung 11). Abbildung 11 enthält sechs Abschnitte, drei repräsentieren Fragen auf einer konkreten Ebene, die drei weiteren Abschnitte repräsentieren Fragen, die von zwei verschiedenen Ebenen betrachtet und gestellt wurden.



Abbildung 11: Abgeleitete konkrete Fragen im Kreis der drei Ebenen Mikro, Meso, Makro und drei Zwischenebenen.

3.3.3.1 Häufige Fragen im Überblick

Die Fragen sind stark geprägt nach dem WIE die KI genutzt werden soll (20 Aussagen) und werden ergänzt mit Fragen nach dem WOFÜR sie genutzt werden, WELCHE Form genutzt werden soll, WER sie nutzen soll und WARUM sie genutzt werden soll. Die gleichzeitig weitgehende Absenz der Frage OB sie genutzt werden soll, deutet darauf hin, dass Dozierende KI in der Lehre als etwas sehen, was bleiben wird – als eine Veränderung, mit der ein Umgang gefunden werden muss. Eben darum dominieren die Fragen nach dem WIE.

Auf Mikro-Ebene sind die meistgestellten Fragen:

- Wie kann ich als Dozierende/r ChatGPT im Unterricht einsetzen? (19 Aussagen)
- Wie halten Dozierende mit dem technologischen Wandel mit? (16 Aussagen)

Einige oft gestellte Fragen beziehen auf Aspekte sowohl der Mikro-Ebene und der Meso-Ebene. Während sie auf der Mikro-Ebene sich auf die Dozierenden persönlich und ihren Unterricht beziehen, wird auf der Meso-Ebene nach Antworten auf der Hochschulebene gesucht:

- Bezüglich Kompetenznachweisen wurden auf der Mikro-Ebene konkrete Überlegungen angestellt und Fragen formuliert, wie diese künftig im eigenen Unterricht aussehen könnten, während auf der Meso-Ebene nach grundsätzlich neuen Methoden und Ideen zur Bewertung gesucht wird: Wie sollen Kompetenznachweise gestaltet werden, so dass die relevante Kompetenz geprüft wird? (10 Aussagen)
- Die Frage nach der Rolle wird auf der Meso-Ebene gestellt aber auch ganz konkret, auf den Unterricht sowie dessen Vor- und Nachbearbeitung bezogen: Wie sieht die Rolle von künftigen Dozierenden aus? (15 Aussagen)

Die meistgestellten Fragen auf Meso-Ebene sind:

- Wer an einer Hochschule entscheidet über Richtlinien im Umgang mit ChatGPT, die Veränderungen im Studienalltag zur Folge haben? (5 Aussagen)
- Wie kann ChatGPT & KI bei der Personalisierung des Unterrichts und Lernens helfen? (5 Aussagen)

Einige oft gestellte Fragen beziehen sich auf Aspekte sowohl der Meso-Ebene als auch der Makro-Ebene. Auf der Meso-Ebene geht es dabei um Umgang mit den Herausforderungen durch die Hochschule oder die Rolle der Dozierenden während auf der Makro-Ebene der grössere gesellschaftlichen Kontext adressiert wird:

- In Bezug auf die Hochschule und die Arbeitswelt: Was sind die relevanten Inhalte und Kompetenzen, die in Zukunft vermittelt werden müssen? (10 Aussagen)
- In Bezug auf die Nutzung von KI im Studium aber auch ausserhalb: Wie kann Studierenden kritisches Denken beigebracht werden im Umgang mit Outputs von KI? (7 Aussagen)
- In Bezug auf die Hochschule und im grösseren akademischen Kontext: Wie wird die akademische Integrität beeinflusst durch die Nutzung von ChatGPT und KI? (6 Aussagen)

Die meistgestellten Fragen auf Makro-Ebene sind:

- Was ist der Wert der Schriftlichkeit? (12 Aussagen)
- Warum zwingen wir Leute zum Schreiben? (8 Aussagen, u. a. auch die Frage explizit so formuliert)

Weiter wurden grundsätzliche Fragen zur Funktionalität von ChatGPT gestellt – Fragen danach, wie ChatGPT funktioniert und was es alles kann. Dies weist darauf hin, dass zumindest bei einem Teil der Dozierenden zum Zeitpunkt der Gespräche ein Grundverständnis für ChatGPT und KI fehlte.

3.3.3.2 Grad der Dringlichkeit

Häufigkeit der Fragen ist ein Aspekt, Dringlichkeit ein ebenso wichtiger anderer. Aus der Perspektive der Dringlichkeit ergeben sich vier Gruppen von Fragen:

1. **Fragen von hoher Dringlichkeit**, welche den Fortbestand der Hochschullehre direkt betreffen: Welche Art von Studierenden wird in Zukunft an Hochschulen studieren? Wie verändert sich das Kompetenzlevel, das in Zukunft gefordert wird? Was sind die relevanten Inhalte und Erkenntnisse, die in Zukunft vermittelt werden müssen? Konkret: Wie kann Studierenden kritisches Denken beigebracht werden im Umgang mit ChatGPT? Wie messen wir in Zukunft Leistung? Wie sollen Kompetenznachweise gestaltet werden? Eine besondere Untergruppe sind Fragen, welche die zukünftige Rolle der Dozierenden im Hochschulunterricht betreffen und deshalb für diese existenziell sind.
2. **Fragen von mittlerer Dringlichkeit**, welche die institutionelle Umsetzung betreffen: Sie gehen vom eher Grundsätzlichen – z. B. Wie vermittelt die Hochschule die relevanten Inhalte? – bis zum sehr Konkreten – z. B. Welche Daten dürfen bei ChatGPT hochgeladen werden.
3. **Fragen von jeweils individuell unterschiedlicher Dringlichkeit**, welche die persönliche Umsetzung betreffen: Dabei stehen der Umgang mit Arbeiten, welche mit Hilfe von ChatGPT erstellt wurden und die Anpassung des Unterrichts im Zentrum.
4. **Fragen, welche die Auswirkungen auf die Gesellschaft betreffen und die wichtig, aber nicht dringlich sind**, weil sie keine direkten Auswirkungen auf das Handeln der Involvierten haben. Beispielsweise Fragen danach, was mit unserem Schreiben (d. h. dessen kultureller Relevanz) und unserem Denken (d. h. unseren kognitiven Praktiken und Kompetenzen) passiert.

Viele dieser Fragen sind auch über ein Jahr nach der Durchführung der Fokusgruppengespräche noch aktuell. Zur ersten und zweiten Gruppe von Fragen Antworten gibt es bislang bestenfalls partielle Antworten. Die Entwicklung des Erkenntnisstands ist langsam. Bezüglich Fragen der dritten Gruppe ist seit dem Zeitpunkt der Datenerhebung die Diskussion weitgehend verschwunden. Es scheint als würde die Nutzung von ChatGPT im Hochschulunterricht weniger als Problem wahrgenommen oder thematisiert werden. Ob die zum Zeitpunkt der Datenerhebung wahrgenommenen Problem zur Zufriedenheit gelöst wurden, müsste durch Nachbefragung beantwortet werden.

Die vierte Gruppe von Fragen zerfällt in zwei Teilgruppen:

- Fragen, über deren Antworten vorerst nur spekuliert werden kann, da es noch wenig oder keine empirische Forschung dazu gibt. Sie zeigen, dass einige grundsätzliche Aspekte durch die Forschung noch kaum aufgegriffen wurden.
- Fragen, welche nach dem aktuellen Wissensstand klare Teilantworten zu haben scheinen – z. B. nach den Unterschieden der Eignung von ChatGPT in Abhängigkeit vom Kulturkreis: nahezuliegen scheint, das Ausmass der Schriftlichkeit als entscheidenden Aspekt zu betrachten – aber trotzdem gestellt werden, ohne im Detail begründet zu werden: Sie zeigen, dass einige Personen komplexere unerkannte Zusammenhänge vermuten.

3.3.4 Analyse der übergreifenden Themen

Bei der Analyse der Subcodes wurde klar, dass es weitere übergreifende Themen gibt, welche sich in den meisten Codes wiederfinden und über fast alle Arten der Fragen in einer Form gestellt werden. Aus diesem Grund wurde die Zuteilung der Subcodes in Themenfelder vorgenommen. Vier der Themenfelder wurden abgeleitet aus bestehenden Subcodes der ersten Kodierrunde,

welche als themenübergreifend identifiziert wurden. Zwei der Themenfeldern, «Heterogenität» und «Unsicherheit», konnten über alle Subcodes hinweg beobachtet werden und wurden deshalb als übergreifende Themenfelder hinzugefügt. Diese Themenfelder werden im folgenden Kapitel inhaltlich analysiert und durch relevante Zitate aus den Aussagen ergänzt. Das Ziel der inhaltlichen Analyse ist es, die Themen, welche die Sprechenden zum Thema KI in der Lehre bewegen, sichtbar zu machen und deren Gedanken zu diskutieren. Auf eine quantitative Analyse wird deshalb verzichtet.

Die 6 übergreifenden Themenfelder sind:

- **Heterogenität:** Unterschiedliche Aussagen und Fragen zu verschiedenen Aspekten (Sichtweisen/Perspektiven, Erfahrungen, Wissensstand, ...)
- **Funktionsweise & Nutzungspraxis:** Aussagen und Fragen zu einer sinnvollen Nutzung von ChatGPT und GenKI aus verschiedenen Perspektiven
- **Kritisches Denken:** Aussagen und Fragen zur Kompetenz des kritischen Denkens aus über verschiedene Aspekte und aus unterschiedlichen Perspektiven.
- **Unsicherheit:** Aussagen und Fragen, die auf Unsicherheit in verschiedenen Aspekten hinweisen.
- **Wert der Schriftlichkeit:** Aussagen und Fragen, die den Wert der Schriftlichkeit aus unterschiedlichen Perspektiven hinterfragen.
- **Rolle der Dozierenden:** Aussagen und Fragen zur Rolle der Dozierenden in unterschiedlichen Zeiträumen und über verschiedene Aufgaben.

3.3.4.1 Heterogenität

Die Gespräche zeigten, dass es notwendig ist, zwischen Fakten und Wahrnehmungen zu unterscheiden. Am klarsten sichtbar wurde dies durch die Heterogenität der Aussagen in vielen Aspekten. Eine rein faktische Auseinandersetzung mit der Frage, was ChatGPT kann und was es nicht kann, geht deshalb ebenso am Thema vorbei, wie eine Beschränkung der Diskussion auf subjektive Wahrnehmungen.

Die Heterogenität betrifft insbesondere folgende Aspekte

- Sichtweisen/Perspektiven
- Erfahrungen
- Wissensstand
- Grad der Verunsicherung
- Praktiken im Umgang mit ChatGPT
- Praktiken bei der reaktiven Adaptierung des Unterrichts
- Grundsätzliche Einschätzung, ob sich denn grosse Fragen stellen oder nicht

So wird beispielsweise von den einen die Adaptierung als persönliche Aufgabe aufgefasst, während andere die Entwicklung der Adaptierung ausserhalb der Lehre sehen – «Ist das nicht eine

Forschungs- und Entwicklungsaufgabe?» – und es zum Teil gerade deshalb von sich weisen – «Dafür fehlt mir das Standing [in Bezug auf Forschung].»

Die gemachten Aussagen legen grosse Unterschiede bei den bisher gemachten Erfahrungen nahe, weniger im Sinne von positiven und negativen Erfahrungen als im Sinne von viel oder wenig Erfahrungen. Die Unterschiede im Abstraktionsgrad der Fragen weisen wiederum auf grosse Unterschiede im Wissensstand hin. Einige wissen noch wenig über die Wirkungsmechanismen, andere haben diese bereits weitgehend reflektiert. Bei einigen ist zudem eine grosse Verunsicherung zu spüren, während andere den Eindruck erwecken, frei von Verunsicherung zu sein. Dem allem entsprechend gibt es auch grosse Unterschiede beim Einsatz von ChatGPT – teils wird es selbst als Werkzeug genutzt, teils nur ausgetestet (Stand Sommer 2023) – und bei der Adaptierung der Unterrichtspraktiken – von gar nicht über substanziell veränderte Praktiken bis zu radikalen Spekulationen. Letzteres betrifft beispielsweise noch weitergehendere Anpassungen, welche zu einer grossen Erhöhung des Unterrichtsniveaus führen würden und viele der heutigen Studierenden überfordern würden.

Selbst in Bezug auf grundsätzliche Haltungen gibt es Unterschiede. Es besteht zwar ein weitgehender Grundkonsens, dass es ein aktives Umgehen mit dem Thema ChatGPT – sei es Problem oder Chance – geben muss. Doch das Ausmass des Damit-Umgehens wird sehr unterschiedlich gesehen. Während die einen sich fragen, wie sie Studierenden den erfolgreichen, reflektierten Einsatz von ChatGPT beibringen sollen, stehen andere dem Werkzeugeinsatz an sich eher ablehnend gegenüber. Explizit wertende Aussagen kommen jedoch selten vor. Die positive Wahrnehmung drückt sich meist über die Formulierung von Chancen aus, die negative Meinung eher als Sentiment der Formulierung.

Allerdings werden Probleme anderer Dozierender explizit angesprochen. So nehmen einige ihre Dozierendenkolleg:innen als rückständiger als die Mehrheit der Studierenden wahr: «Les profs ils vont pas voir parce qu'ils sont en retard par rapport de, je veux dire qu'une partie du corps enseignant est en retard par rapport aux étudiants.».

Unterschiede gibt es des weiteren in der Thematisierung der Dynamik. Oft werden die Herausforderungen der Schnelligkeit – «Die schnelle Entwicklung beschäftigt mich (...), weil das ist ja nicht unser Kernthema.» – und der Beschleunigung – «Wir sind in einem in einem extrem evolutionären und sich beschleunigenden Prozess drin.» – angesprochen, sowie mehr oder weniger direkt die faktische Disruptivität. Zum Teil sieht man sich aber auch am Anfang einer längeren Veränderung: «Wahrscheinlich gibt's vieles, vieles nebeneinander, bevor sich so ein Common Sense oder so ein Mainstream herausgebildet hat und und dann werden wir ganz andere Studierende bei uns haben und und die sind dann schon ein Stück weit sozialisiert und die haben dann Erfahrungen. (...) Ich habe das Gefühl, all die Wegleitungen, all die Richtlinien, das wird dauernd überarbeitet werden müssen, je nachdem, was gerade ansteht.»

Die hier beschriebenen Unterschiede der Wahrnehmung deuten auf situative Unterschiede der Nutzungsszenarien hin, die eine universelle Bewertung von ChatGPT für das wissenschaftliche Schreiben in Frage stellen. Darüber hinaus können die Unterschiede auch dann nicht durch eine faktische Analyse – «was ist richtig, was ist falsch» – adressiert werden, wenn die Wahrnehmung sich auf sehr ähnliche Kontexte beziehen. Denn die Wahrnehmungen prägen das Handeln und können zu kollektiven Phänomenen führen wie einer mehrfach gespaltenen Organisationskultur.

3.3.4.2 Funktionsweise und Nutzungspraxis

Viele Fragen und Diskussionen drehen sich darum, wie ChatGPT funktioniert, worin seine Stärken und auch seine Schwächen und Grenzen liegen, und was alles damit in der Praxis gemacht werden kann – nicht nur durch Studierende, sondern auch durch Dozierende. Dabei sind bereits viele Ideen vorhanden zu den Einsatzmöglichkeiten, die aber grossteils noch nicht ausprobiert wurden, wie diese Äusserung beispielhaft zeigt: «Ich kann mir jetzt vorstellen, habe ich [aber] noch nicht gemacht, dass ich als Lehrer, als Dozent mir Übungen erstellen kann.»

Vier Aspekte stehen bei den Fragen im Zentrum:

- Funktionsweise von ChatGPT: z.B. «Was ich mir noch nicht so vorstellen kann: Wie wird dieses ChatGPT gespiesen? Wie kommen zum Beispiel sinnvolle Hörverstehensaufgaben in diesen Speicher, sodass ChatGPT überhaupt solche Testaufgaben entwickeln könnte? Das übersteigt meine Vorstellung.»
- Einsatzmöglichkeiten im Unterricht: z.B. «Was uns abteilungsmässig beschäftigt hat, [war] die Frage, welche Einsatzmöglichkeiten gibt es? Man hört von verschiedenen Kolleginnen und Kollegen [dass sie] am Überlegen sind, wie sie es einbinden können.»
- Sicherstellung des Lernerfolgs: z.B. «Wie stellen wir sicher, dass sie das trotzdem irgendwie können. Es wird auch für uns nicht ganz anspruchslos.»
- Weiterführung der Prüfungspraxis: z.B. «Was machen wir jetzt, eben mit den Bachelorarbeiten, Masterarbeiten? Genau so kann es wahrscheinlich nicht weitergehen. Was machen wir?»

Einige Fragen gehen sehr tief, etwa in Bezug auf die Möglichkeiten, ChatGPT mit genauen Spezifikationen seines Herangehens an eine Aufgabe zu konfigurieren, und den Grad seiner proaktiven Sensitivität für den kulturellen Kontext. Dabei ist zu beobachten, dass die Handhabung von ChatGPT unterschätzt wird, konkret das «Konfigurieren» der Antworten.

Insgesamt ist die Bandbreite der Hoffnungen gross – von realistischen über mit sehr hohem Aufwand realisierbare bis zu nach aktuellem Stand unrealistischen. So möchten Dozierende etwa Studierenden Werkzeuge bieten für das Schreiben von Zusammenfassungen und Protokollen, digitale Trainingsassistenten zum Spracherwerb bauen, d. h. «interaktiven Tools, um die Sprache zu lernen» oder «Hörverstehensaufgaben» für Übungen und Tests konstruieren.

Die Nutzung von ChatGPT wird meist als eher negativ wahrgenommen und Fragen drehen sich um die möglichen negativen Auswirkungen oder wie diese verhindert werden können. Einige Chancen werden aber trotzdem identifiziert:

- Chancen für Zeiteinsparungen im Unterricht: «L'écrit aura de moins en moins de place dans notre enseignement. On va passer plus de temps avec des choses dont on aura aussi besoin après dans la vie professionnelle ou la vie de tous les jours.»
- Chancen für neue Kompetenzvermittlungen im Unterricht: «Ich glaube, wir müssen noch ein bisschen herausfinden, für was [wir ChatGPT nutzen können]. Was ist der konkrete Lerneffekt vom Schreibprozess mit ChatGPT.»
- Chancen für mehr mündliche Kommunikation: «Es gibt auch die Chance, dass die zwischenmenschliche Kommunikation viel an Wert und Gewicht gewinnt, dass wir mehr interessante Gespräche miteinander haben, um Wissen zu produzieren und weniger allein im

stillen Kämmerchen [arbeiten], dass wir viel weniger uns mit Texten beschäftigen und viel mehr mit etwas, das eine menschliche Bedeutung hat.»

- Chance für die Neudefinition der Dozierendenrolle: «[Es] ist eine Chance, aber auch das Risiko, dass man sich quasi neu erfindet als Lehrperson.»
- Chancen für die Beschleunigung des notwendigen Wandels: «Und ich glaube, da ist schon eine Chance, dass man da jetzt endlich das notwendige Gas gibt.»
- Chancen für mehr Inklusion: «Vielleicht könnte ChatGPT in Zukunft uns auch helfen, eine grössere Diversität von Studierenden zu haben.»

Trotzdem gibt es Ängste vor einem Deskillung, respektive ist dieses Risiko im Bewusstsein der Dozierenden. Sie betrifft einerseits das Handwerkliche der Texterstellung – «Die Frage ist, wenn ich das [Texte verfassen] nie [selbst] gemacht habe, wenn ich das nie gelernt habe, wie sinnvoll ich das Tool nutzen nachher kann, wie zielführend ich das Tool nachher einsetzen kann.» – und andererseits das Inhaltliche, respektive die inhaltlichen Qualitätsstandards: «Vieles, was die KIs auswerfen, ist okay bis gut. So zu 80 % [sind die Qualitätsanforderungen] vielleicht erfüllt und die 20 % bis zur Sechsen fehlen noch. Wenn man sich dann mit den 80 % zufriedengibt und das zum allgemeinen neuen Standard wird, dann werden die 80 % die neuen 100 %. Nicht in ein, zwei Jahren, aber über eine längere Periode.»

Vergleiche mit anderen Technologien kommen zu unterschiedlichen Schlüssen. Einerseits betonen sie das Deskillung-Risiko: «Ich finde zum Beispiel, dass die Orientierungsfähigkeit von jungen Menschen massiv abgegeben hat seit GPS. (...) Was ich daraus schliesse, ist, dass die Kompetenz abnimmt, wenn du sie nicht förderst.» Andererseits hinterfragen sie es auch und vermuten sogar noch unbekannte Chancen: «[Es kann] vielleicht zu einer Art Inflation von bestimmten Fähigkeiten kommen. Als der Taschenrechner eingeführt wurde, haben auch alle gesagt, das ist der Untergang der Menschheit. Jetzt können wir bald alle nicht mehr rechnen, weil alle dieses Ding benutzen. [Das] stimmte nicht. Es ist nicht so gekommen.»

3.3.4.3 Kritisches Denken

Im Folgenden werden die Aussagen zu kritischem Denken zusammengefasst, welche als offene Fragen kodiert wurden. Dabei zeigt sich: kritisches Denken wird implizit an vielen Stellen als notwendige Voraussetzung für ein verantwortungsbewusstes Nutzen von ChatGPT dargestellt. Doch viele teilen die Sorge, dass es verloren gehen könnte: «Sind die Schüler in Zukunft noch in der Lage, selbstständig zu denken? Ich habe den Eindruck, dass die Schüler nicht mehr unbedingt verstehen, was ihnen von ChatGPT vorgeschlagen wird.» Bemerkenswert aber auch repräsentativ an dieser Aussage ist die implizite Annahme, dass kritisches Denken bislang vorhanden ist. Es wird hier nicht argumentiert, dass kritisches Denken in Zukunft wichtiger wird, um mit diesem Phänomen umzugehen, sondern es wird umgekehrt befürchtet, dass das Werkzeug kritisches Denken zum Verschwinden bringen könnte.

Als Grund für die pessimistischen Zukunftserwartungen wird bequemes Anpassungsverhalten angeführt: «Der Mensch ist bequem: Ich glaube nicht, dass du dich [mit ChatGPT] dem aussetzt: [der Suche nach] analytische Lücken, fehlender Stringenz, argumentative Schwäche eines Textes.» Diese durchaus repräsentative Äusserung zeigt, dass beim Thema kritisches Denken der Fokus auf dem Vergleich mit der vor-ChatGPT-Welt dominiert und die Frage der Rolle des kritischen Denkens in einer Welt mit ChatGPT dadurch in den Hintergrund tritt. Es wird im Übrigen

auch gefragt, ob der Verlust wirklich etwas Negatives ist: «Warum muss man das [kritische Denken] noch haben, wenn du eigentlich ganz glücklich damit bist, dass ein Text einigermaßen [lesbar ist].»

Bei Einordnung dieser Diskursperspektiven sollte man allerdings vorsichtig sein. In der Diskussion ist der adressierte Aspekt nicht immer klar: Handelt es sich um die konkrete Sorge, dass die Werkzeugnutzung Studierende überfordert, oder handelt es sich um die grundsätzliche Sorge über den drohenden Verlust kritischer Reflexionen. In vielen Stellungnahmen wird das Thema jeweils sehr breit und zugleich sehr vorsichtig und unpräzise angesprochen. Das vermittelt den Eindruck einer Art Sprachlosigkeit angesichts des neuen Werkzeugs, welches die Spielregeln verändert und den Hochschulunterricht grundsätzlich in Frage zu stellen scheint.

Während es in der wahrgenommenen Gemengelage weitgehend als notwendig wahrgenommen wird, die Fähigkeit zum kritischen Denken zu fördern (oder genauer: seinen Verfall zu verhindern) ist das WIE unklar: «Je me pose clairement la question sur ce critical thinking en fait, comment est-ce qu'on peut le développer maintenant ?» Es werden zwar einige konkreten Aspekte gesehen, die in der Ausbildung adressiert werden sollten «Moi ce que je vois comme solution par rapport à ça, de travailler en fait autour des thématiques, par exemple autour des cultural biases ou des gender biases.», aber eine Auslegeordnung zur Zukunft des kritischen Denkens wird nicht vorgebracht. In der Formulierung wird kritisches Denken oft wie eine Eigenschaft dargestellt, welche sich nicht in unterschiedliche Aspekte differenziert.

3.3.4.4 Unsicherheit

Das Vorhandensein von Unsicherheit wird in viele Aussagen spürbar, oft genug aber auch mehr oder weniger direkt angesprochen. Wesentliche Gründe für die Unsicherheit sind die hohe Geschwindigkeit der Entwicklung und deren unklarer Endpunkt. Einige haben das Gefühl, den aktuellen Stand nicht zu kennen und halten dies für ein Problem. Andere fühlen sich sehr unsicher, wie sie ihren Unterricht und ihre Prüfungspraxis anpassen sollen: «Ich werde auf jeden Fall versuchen, das zu adressieren. Ich bin mir aber [noch] nicht so ganz klar, wie.» Es gibt auch viel reflektierte Unsicherheit: «Mais tu crois pas qu'il faudrait aussi redéfinir en fait plagiat ? Parce que si tu vas, si tu sais, si tu utilises en fait ChatGPT par exemple le prompt c'est le tien.» Und es gibt einiges an Misstrauen: «Wer bezahlt das Ganze? Womit? Ist es Geld, sind es Daten? Wenn es Daten sind, was darf alles in diese Blackbox hineingespielt werden und was nicht?»

Die grösste Unsicherheit betrifft aber die Folgen – von der Anpassung des eigenen Unterrichts und Prüfens bis zur Anpassung der Unterrichtsziele: «Welche Kompetenzen brauchen die Abgängerinnen, Abgänger in fünf Jahren noch, mit dem Studienabschluss im Berufsleben?» Darüber hinaus machen sich einige Personen Sorgen, wie man alle mitnehmen kann in dieser stattfindenden Transformation, konkret fragen sie sich insbesondere, welche Kommunikation es braucht: «Wie informieren wir Mitarbeitende, integrieren, involvieren sie?»

Es werden aber auch Befürchtungen in Bezug auf die emotionalen Auswirkungen der Nutzung von ChatGPT geäußert: «Was macht das mit den Studierenden, wenn sie das Gefühl haben, ChatGPT kann das ohnehin besser, als ich es selbst hinbringen würde?»

3.3.4.5 Wert der Schriftlichkeit

Der Grossteil der Aussagen zum Wert der Schriftlichkeit adressiert die operative Nutzung des Textschreibens für Prüfungen, respektive Leistungsnachweise: «Welche schriftlichen Arbeiten machen Sinn in einem Studium, wenn KI einen Teil dieser Schriftlichkeit abnimmt? Und was brauchen wir an Stelle diesen schriftlichen Arbeiten, damit das Denken trotzdem passiert?» Dies ist ein grosses Thema für viele der Dozierenden, da sie gemeinhin der Schriftlichkeit hohen Wert zumessen. Die Diskussion wird aber nicht auf die Schriftlichkeit beschränkt, denn es wird auch die ganz grundsätzliche Frage gestellt, wie in Zukunft überhaupt noch individuelle Leistung bewertet werden können, etwa bei der Durchführung einer Recherche. Die meisten Diskussionen drehen sich aber um die Bewertung von Texten. Bisweilen wird sogar die zukünftige Bedeutung von Texten in unserer Gesellschaft in Frage gestellt, respektive die aktuelle Unklarheit diskutiert, welche Textproduktionskompetenzen Absolventen der Hochschulen in Zukunft benötigen werden.

Grundsätzlich kann man zwei Sichtweisen unterscheiden, die bisweilen auch in den Aussagen einzelner verschränkt werden:

1. Infragestellung der schriftlichen Arbeiten: Die Rolle der schriftlichen Arbeiten im Studium und insbesondere bei Prüfungen wird hinterfragt und die Frage nach Abschaffung oder auch (wie im Zitat oben) nach Ersetzung wird gestellt: «Was ist der Zweck der Bachelorarbeit? Was trägt (sie) zur Professionalisierung bei?»
2. Infragestellung der Studiungsgestaltung als Ganzes: Dozierende gehen zum Teil davon aus, dass ChatGPT so umfassend im Alltag der Menschen Einzug halten wird, dass das Studium in seiner bisherigen Form in Zukunft keinen Mehrwert mehr bieten wird und deshalb grundsätzlich neu ausgerichtet werden muss: «Wir werden viel anforderungsreicher in unserem Unterricht werden. (...) Die Bloomschen Taxonomiestufen 1-3, können wir wegschmeissen. Wir können bei vier und höher einsteigen. (...) Also ich glaube, es wird für die Studierenden anspruchsvoller.»

Teils werden die Fragen sehr konkret oder rhetorisch gestellt, teils werden sie sehr grundsätzlich gestellt: «Was ist Leistung? Wie misst man Leistung, die eigentlich später verwendet wird in der [Arbeit]?»

Über die Antworten werden unterschiedliche Spekulationen angestellt, wobei aber klar scheint, dass man auf Übergangslösungen setzen muss. Eine sehr konkrete Spekulation lautet: «Wir könnten es uns sehr einfach machen: Wir lassen die Bachelorarbeit, wie sie ist und sagen am Schluss gibt es ein Kolloquium. Nach Zufallsprinzip werden drei Quellen zugelost und [die Studierenden] werden gebeten, uns die grössere Einbettung dieser Quellen in ihren Bereich darzustellen. (...) Ich meine, das wäre eine Anforderung um 1000 % höher, als sie bis jetzt war an die Bachelorarbeit. Sie würden also nicht mehr ins Schreiben investieren, [sondern] ins Lesen und Verarbeiten.» Aber die Heterogenität der Spekulationen ist hoch. So wird beispielsweise auch die Abschaffung der Noten antizipiert, respektive halb gefordert.

Es gibt aber auch noch die Haltung, dass die Vorgabe, ChatGPT nicht zu nutzen eigentlich die beste Problemlösung wäre, die aber am fehlenden ethischen Bewusstsein scheitert: «Après j'ai horreur de dire ça. Mais mais la triche quoi ! (.) La difficulté c'est que tout le monde utilise ça à toutes les sauces pour tout faire et (.) de pas comprendre que ça ne peut pas être, ça peut atteindre l'intégrité académique et du coup ça compte comme de la triche.» Dies ist eine der wenigen Gelegenheiten, in denen tatsächlich das OB in Frage gestellt aber. Aber die Formulierung zeigt, dass dieses In-Frage-Stellen als keine realistische praktische Option angesehen wird.

3.3.4.6 Rolle der Dozierenden

Viele Aussagen vermitteln den Eindruck, dass die Ansprüche an die Dozierenden diesen nicht klar sind. Konkret wird auf die unklaren Erwartungen der Arbeitgeber verwiesen: «[Sollen wir] die Hochschule nicht isoliert betrachten, weil sich die Welt ausserhalb, auf die wir ja ausbilden wollen, verändert durch diese Tools?», respektive «Sodass wir nicht nur sagen, wie ändert sich unsere Lehre im Sinn von Didaktik, sondern auch wie ändern sich die Fachinhalte, die die Skills, die Studierenden und später haben müssen?». Konkret ist auch unklar, ob die Vermittlung von ChatGPT-Nutzungskompetenzen neu zum Aufgabenbereich von Dozierenden gehören: «Est-ce qu'on sera le passeur de compétences techniques ?» und was aus der Rolle der Dozierenden macht: ob es eine zusätzliche Aufgabe oder eine Rollenveränderung darstellt.

Sorgen wegen der Schnelligkeit werden in verschiedenster Form geäußert: «Wir haben ja nicht die Ressourcen oder nicht die Zeit oder vielleicht auch nicht immer das Interesse, wirklich da an der vordersten Front zu [stehen und zu fragen]: Was ist jetzt der Stand?» Und dort, wo Aufgaben als klar wahrgenommen werden, gibt es Unklarheit in Bezug auf die Umsetzung: «Par contre, on a un devoir et ça je sais pas comment on va le faire, mais un devoir de réflexion critique sur le sujet, de leur apprendre à être critique avec cet outil, de leur apprendre à l'utiliser judicieusement, (...) Je ne sais pas comment, dans les différents niveaux de responsabilité qui sont les miens je vais mettre ça en place ?»

Hier werden aber auch explizit Chancen gesehen, dass ChatGPT als ein Enabler für bessere, vor allem adressatengerechtere Lehrtätigkeit dienen könnte: «Also meine Idee war eher, dass die KI sie bei ihrem Vorwissen abholt und ihnen dann eigentlich den Lernpfad, vielleicht auch die Art, wie sie am besten lernen, zusammenstellt.» Hier wird insbesondere die Möglichkeit gesehen, von einem optimalen Unterricht für eine Klasse zu einem individuell optimalen Unterricht für einzelne Studierende zu kommen. Es wird aber auch als Chance wahrgenommen, mittels dieser radikalen Personalisierung der Lernunterstützung die Inklusion zu fördern: «Vielleicht könnte ChatGPT in Zukunft uns auch helfen, eine grössere Diversität von Studierenden zu haben.» Letzteres betrifft insbesondere Studierende, welche aufgrund spezifischer Schwächen ein Studium bislang nicht schafften. Mit gezielter Unterstützung in den Bereichen ihrer Schwächen – konkret wird hier etwa die Mathematik genannt – könnten sie neu zum erfolgreichen Studieren befähigt werden, so die Hoffnung. Diese Spekulation zeigen, dass die Diskussion nicht immer klar trennte zwischen ChatGPT als konkretem Werkzeug und der Nutzung von KI ganz allgemein.

3.3.5 Schlussfolgerungen aus der Auswertung der offenen Fragen

Das Thema ChatGPT für wissenschaftliches Schreiben und in der Hochschullehre erweist sich trotz der Überschaubarkeit des Anwendungsbereiches als weites Feld. Es stellen sich exemplarisch grundsätzliche Zukunftsfragen in Bezug auf die Veränderung der Gesellschaft durch die Verfügbarkeit von hochentwickelter Künstlicher Intelligenz. Die Gemengelage ist dabei komplex und die Heterogenität ist für das inhaltlich überschaubare Feld überwältigend. Hinter allen geäußerten Fragen stehen aber zwei Grundannahmen und zwei Grundfragen.

Die Grundannahmen sind, dass (1) ChatGPT in Zukunft eine wesentliche Rolle im Hochschulstudium spielen wird und dass (2) dies eine Anpassung vieler Hochschulpraktiken, insbesondere (aber nicht nur) bezogen auf Prüfungen, verlangt.

Die Grundfragen sind, wie sich (1) die Hochschulausbildung verändern wird und wie sich (2) dies auf die Rolle der Dozierenden auswirken wird.

Die offenen Fragen ziehen sich von der Mikro-Ebene, d. h. dem individuellen Umgang mit Chancen und Herausforderungen durch die Dozierenden über die Meso-Ebene mit ihren institutionellen Aspekten bis zur Makro-Ebene mit den induzierten gesellschaftlichen Veränderungen. Die Spekulationen zu den längerfristigen Folgen reichen von einer Trivialisierung oder Elitarisierung des Hochschulstudiums – das heisst, dass das Hochschulstudium viel mehr oder viel weniger Menschen in Zukunft zugänglich sein wird – bis zu einer Ermöglichung einer personalisierten Ausbildung, welche es für Menschen mit spezifischen Lerndefiziten öffnen und könnte. Einerseits stellt ChatGPT für einige der Teilnehmenden bisher Gegebenes infrage – insbesondere den Wert der Schriftlichkeit und/oder den Wert des bisherigen Hochschulstudiums – und andererseits besitzt es eventuell das Potenzial, einen Aufbruch zu neuen Ausbildungsformen zu ermöglichen.

Dabei werden zahlreiche mögliche Probleme adressiert: Sie reichen von der Benennung aktueller Probleme bis zu verschiedensten Zukunftsängsten. Viele Fragen sind entweder versteckte Klagen oder beinhalten Klagen in einem Nebensatz. Sorgen und Befürchtung betreffen dabei Mikro-, Meso- und Makro-Ebene. Diesen stehen zahlenmässig weniger Zukunftshoffnungen entgegen. Oft werden Aussagen – auch optimistische – so formuliert, dass eine hohe Unsicherheit sichtbar wird.

Im Fokus der Diskussionen auf der Mikro- und teilweise auf der Meso-Ebene stehen die Rollen der Dozierenden und die Erwartungen an sie. Im Fokus der Diskussionen auf der Meso- und der Makro-Ebene steht der Wert der Schriftlichkeit. Darüber hinaus ist kritisches Denken ein zwar nicht sehr häufig, aber immer wieder auftauchendes Thema, das allerdings auf der Meso-Ebene nur indirekt angesprochen wird. Weitgehende Einigkeit herrscht dabei, dass kritisches Denken für die Nutzung von ChatGPT sehr wichtig ist. Aus Sicht der Dozierenden macht ChatGPT sichtbar, dass es teilweise den Studierenden heute schon fehlt, es besteht aber auch die Gefahr, dass es durch die Nutzung von ChatGPT abhanden kommen könnte. Dagegen fehlen konzeptionelle Ideen für den Umgang mit der Kompetenz *Kritisches Denken* auf institutioneller Ebene, wonach in den Fokusgruppengesprächen jedoch auch nicht explizit gefragt wurde. Insbesondere fehlt in den Diskussionen zu offenen Fragen die methodische Frage, wie kritisches Denken durch eine reflektierte Nutzung von ChatGPT gefördert werden könnte.

4 Fazit und Ausblick

Im Sommer und Herbst 2023 haben die Berner Fachhochschule (BFH), die Université de Neuchâtel (UniNe), die Pädagogische Hochschule Zürich (PHZH) und die Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) 18 Fokusgruppengespräche mit insgesamt 66 Teilnehmenden durchgeführt. Im Zentrum der Gespräche stand die übergeordnete Frage, wie mit dem Aufkommen von GenKI in der Hochschullehre umgegangen wird. In der Auswertung haben wir gezielt die Aspekte der Best Practices und offene Fragen in den Fokus genommen.

Besonders die Aspekte, die in angesprochen und später mit dem Code «Offene Fragen» kodiert worden sind, verdeutlichen, dass die Veränderungen in der Hochschullehre durch generative künstliche Intelligenz (GenKI) weitreichend sind. Auf der Mikro-Ebene (Lehre) betrifft es vor allem den individuellen Umgang der Dozierenden mit den neuen Technologien, auf der Meso-Ebene (Hochschule) die institutionellen Anpassungen in der Lehre und die Ausrichtung der Prüfungen und auf der Makro-Ebene (Gesellschaft) die gesellschaftlichen und wirtschaftlichen Auswirkungen der KI-Technologie.

Obwohl in den Fokusgruppengesprächen gezielt nach Praxis-Beispielen der befragten Dozierenden gefragt sowie offene Fragen bei allen Hochschulangehörigen angesprochen worden sind, wurde bei beiden Themengebieten immer wieder die Unsicherheit der Dozierenden im Umgang mit GenKI thematisiert. Während bei den offenen Fragen der Aspekt der Unsicherheit noch erwartbar war, überraschen entsprechende Aussagen bei der Darstellung von positiven Beispielen der Dozierenden im konkreten Umgang mit GenKI in der Lehre. Daher werden im Folgenden nicht nur die Chancen und Herausforderungen von GenKI in der Hochschullehre zusammengefasst, sondern Handlungsempfehlungen für Hochschulen gegeben, um nicht nur Praxis-Beispiele zu teilen, sondern auch diese Unsicherheiten zu adressieren.

Chancen und Herausforderungen

Insgesamt zeigt die Untersuchung zur Nutzung von GenKI, insbesondere ChatGPT als ein spezifisches Tool, in der Hochschullehre ein differenziertes Bild, das sowohl Chancen als auch Herausforderungen widerspiegelt. Die Ergebnisse aus den quantitativen Auswertungen, den Best-Practice-Beispielen und den offenen Fragen bieten zusammengefasst wichtige Erkenntnisse für eine erfolgreiche Integration in den Bildungsbereich.

1. Hohe Bereitschaft zur Integration in die Lehre und Bedarf an Richtlinien

Die Bereitschaft, GenKI in die Lehre zu integrieren, ist trotz der Vielzahl an Bedenken und an offenen Fragen hoch. Dozierende erkennen zunehmend das Potenzial von KI-Tools, um Lehr- und Lernprozesse zu ergänzen. Besonders in der Textgenerierung und -überarbeitung sowie der Erstellung von Zusammenfassungen zeigt sich eine breite Nutzung. Gleichzeitig werden ethische Richtlinien und klare Vorgaben für den Umgang mit GenKI in der Hochschullehre gefordert.

2. Weiterbildungsbedarf und Kompetenz-Ausbau bei Dozierenden

Ein zentrales Thema ist der grosse Bedarf an Schulungen und Weiterbildungen. Viele Dozierende geben an, dass sie GenKI noch experimentell nutzen und sich unsicher im Umgang damit fühlen. Insbesondere Fähigkeiten im Bereich Prompting sowie der Umgang mit verschiedenen KI-Tools sind Bereiche, in denen gezielte Weiterbildungsmaßnahmen für Dozierende gefordert werden.

Von den Weiterbildungen der Dozierenden profitieren schliesslich auch Studierende. Durch aktuelle Weiterbildungen ist es den Dozierenden möglich, gezielt Vorteile und Probleme von KI-Tools zu thematisieren, wodurch die Studierenden die Technologie reflektierter nutzen können und in ihrem kritischen Denken gefördert werden.

3. Anwendungsorientierte Lehrmethoden

Die Integration von GenKI in praxisorientierte Aufgabenstellungen bietet grosses Potenzial. Dozierende nutzen GenKI auch für die Erstellung realistischer Übungs- und Prüfungsaufgaben, welche die Problemlösungsfähigkeiten der Studierenden fördern. Hierzu müssen auch neue Bewertungsmethoden entwickelt werden, was insgesamt die Entwicklung innovativer Lehrformate mit praktischer Ausrichtung auf die zukünftigen Arbeitsfelder und -formen fördert.

4. Ausbau von Fähigkeiten zur kritischen Reflexion bei den Studierenden

In den untersuchten Fokusgruppengesprächen wird die Rolle der Studierenden im Schreibprozess betont: Texte sollen nicht nur KI-gestützt erstellt, sondern die Ergebnisse kritisch hinterfragt und der Umgang mit den Tools reflektiert werden. Diese Herangehensweise könnte langfristig die Entwicklung einer verantwortungsvollen und reflektierten Haltung gegenüber technologischen Werkzeugen allgemein fördern. Besonders in diesem Aspekt ist es wichtig, von einem positiven Studierendenbild auszugehen, um zu verhindern, dass die breite Diskussion rund um GenKI in ein negatives Narrativ über Kompetenzverluste und Betrugsversuche umschlägt.

Zusammenfassend lässt sich sagen, dass GenKI grosses Potenzial für die Zukunft der Hochschulbildung hat, sowohl in Bezug auf innovative Lehrmethoden als auch auf die Entwicklung neuer Kompetenzen. Dabei sollten aber die aktuellen Bedenken und offenen Fragen ernst genommen werden, insbesondere im Hinblick auf ethische, rechtliche und gesellschaftliche Aspekte.

Handlungsempfehlungen

Aus den dargestellten Ergebnissen lassen sich eine Reihe von Handlungsempfehlungen für Hochschulen ableiten, die sowohl die unmittelbare Integration von GenKI in die Lehre als auch langfristige strategische Überlegungen berücksichtigen.

1. Entwicklung klarer Richtlinien und Standards für den Einsatz von KI in der Lehre

- **Ziel:** Sicherstellung einer verantwortungsvollen und transparenten Nutzung von KI-Tools und Vermeidung von Missbrauch, etwa bei Prüfungsleistungen oder in wissenschaftlichen Arbeiten.
- **Empfehlung:** Bildungseinrichtungen sollten klare **institutionelle Richtlinien** und **Standards** für den Umgang mit GenKI entwickeln, die ethische, rechtliche und akademische Kriterien berücksichtigen. Diese sollten sowohl die Nutzung durch Dozierende als auch durch Studierende und damit die Nutzung auf individueller Ebene – um eine disziplinäre Professionalität zu etablieren – wie auch auf institutioneller Ebene umfassen.

2. Förderung von Weiterbildungen und Kompetenzaufbau für Dozierende

- **Ziel:** Dozierende sollen die Funktionsweise von KI-Tools verstehen und in der Lage sein, diese verantwortungsvoll in ihre Lehre zu integrieren.

- **Empfehlung:** Um Unsicherheiten im Umgang mit GenKI abzubauen und das Potenzial dieser Technologien zu nutzen, sollte die **Methodenkompetenz** von Dozierenden gezielt gefördert werden. Dies sollte sowohl (kontinuierliche) technische als auch didaktische Schulungen wie auch den Austausch von Dozierenden untereinander umfassen.

3. Integration von GenKI in praxisorientierte Lehr- und Lernformate

- **Ziel:** Studierende sollen lernen, KI als Unterstützung in komplexen, realitätsnahen Aufgaben zu nutzen und die Ergebnisse der KI eigenständig zu reflektieren und zu hinterfragen.
- **Empfehlung:** Dozierende sollten **anwendungsorientierte Aufgaben** entwickeln, die die Nutzung von KI als Werkzeug integrieren und Studierende zu einer aktiven Auseinandersetzung mit der Technologie anregen. Der Fokus sollte auf der Anwendung von KI in realen Szenarien liegen, um Problemlösungsfähigkeiten und kritisches Denken zu fördern.

4. Anpassung der Prüfungsformate und Förderung des kritischen Denkens

- **Ziel:** Sicherstellen, dass Studierende weiterhin die erforderlichen fachlichen und überfachlichen Kompetenzen – auch im Umgang mit KI («KI-Literacy») entwickeln. Gleichzeitig sollen sie in die Lage versetzt werden, KI nicht nur als Werkzeug zu verwenden, sondern auch die Ergebnisse kritisch zu hinterfragen und eigenständig zu verbessern. Dadurch wird ein verantwortungsvoller und reflektierter Umgang mit KI-Technologien gefördert.
- **Empfehlung:** Die Hochschulbildung sollte die Prüfungsformate und Kompetenzmodelle an die Möglichkeiten und Herausforderungen der GenKI anpassen. Insbesondere der Stellenwert schriftlicher Arbeiten ist zu prüfen. Ebenfalls wichtig ist die **Entwicklung neuer Prüfungsformate**, die den kritischen Umgang mit KI fördern und gleichzeitig sicherstellen, dass Studierende selbstständig und eigenverantwortlich denken und handeln.

Im Lehr-/Lernprozess sollte der Einsatz von KI nicht nur als technisches Werkzeug, sondern auch als Anlass für eine vertiefte Auseinandersetzung mit den **gesellschaftlichen und ethischen Implikationen** der Technologie dienen. Besonders wichtig ist es, dass Studierende z. B. durch **Feedbackprozesse** die von KI erzeugten Inhalte reflektieren, deren Qualität evaluieren und verbessern.

Es hat sich in den Fokusgruppengesprächen gezeigt, dass Dozierende davon ausgehen, dass GenKI so umfassend im Alltag etabliert wird, dass dies eine umfassende Reform der Studieninhalte und -formen erfordert. Von konkretem Interesse in Bezug auf Kompetenzen ist die Beantwortung folgender Fragen:

- Welche Kompetenzen müssen neu oder vermehrt vermittelt werden?
- Welche Kompetenzen müssen nicht mehr vermittelt werden?
- Wie muss sich die Kompetenzvermittlung ändern?

Die Handlungsempfehlungen bieten einen Ansatz, um die Chancen von GenKI in der Lehre zu nutzen und gleichzeitig die damit verbundenen Herausforderungen zu berücksichtigen. Ziel der Handlungsempfehlungen ist eine langfristige, verantwortungsvolle Integration von KI in Bildungsprozesse.

Notwendige Schritte in der Forschung

Aus methodologischer Perspektive zeigen die Fokusgruppengespräche, dass es derzeit noch an einem kohärenten Forschungsansatz fehlt, der die Auswirkungen der Nutzung von ChatGPT und ähnlichen Technologien auf die Gesellschaft, die Hochschulbildung im Allgemeinen und das Schreiben im Spezifischen umfassend untersucht. Insbesondere ein Dialog zwischen Forschung, Praxis und Politik ist notwendig, um eine verantwortungsvolle Nutzung dieser Technologie zu sichern.

Über diese konkreten Empfehlungen für Hochschulen hinaus lassen sich auch weiterführende Handlungsempfehlungen für den wissenschaftlichen und gesellschaftlichen Umgang mit GenKI ableiten. Ausgehend von unseren Analysen empfehlen wir folgende Schritte:

- A. Eine Einordnung des Diskurses über GenKI in den **allgemeineren und älteren Diskurs zur digitalen Transformation**: Welche Resultate sind auf GenKI übertragbar? Welche spezifischen Fragestellungen rücken durch GenKI in den Vordergrund?
- B. **Transdisziplinäre Forschung** zu den gesellschaftlichen Auswirkungen, welche die Grundlagen für ein besseres Verständnis aus der Perspektive schafft, dass wir uns als Gesellschaft in einem Lernprozess befinden – und nicht aus der Perspektive schneller regulatorischer Reaktionen oder unhinterfragter Technologiegläubigkeit.

5 Anhang

Umfangreiche Anhänge siehe Ordner Anhänge Dokumentation.

A. Liste verwendeter Publikationen

- AlAfnan, M. A., Dishari, S., Jovic, M. & Lomidze, K. (2023). ChatGPT as an Educational Tool: Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses. *Journal of Artificial Intelligence and Technology*, 3(2), 60–68. <https://doi.org/10.37965/jait.2023.0184>
- Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (610-623). ACM. <https://doi.org/10.1145/3442188.3445922>
- Cotton, D. R. E., Cotton, P. A. & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. <https://doi.org/10.1080/14703297.2023.2190148>
- Cox, C. & Tzoc, E. (2023). ChatGPT: Implications for academic libraries. *College & Research Libraries News*, 84(3). <https://doi.org/10.5860/crln.84.3.99>
- Crompton, H. & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00392-8>
- Dergaa, I., Chamari, K., Zmijewski, P. & Ben Saad, H. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of sport*, 40(2), 615–622. <https://doi.org/10.5114/biolsport.2023.125623>
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S. & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), Artikel asi.24750, 570–581. <https://doi.org/10.1002/asi.24750>
- Nazari, N., Shabbir, M. S. & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*, 7(5), e07014. <https://doi.org/10.1016/j.heliyon.2021.e07014>
- Pagel, P. & Ludwig, T. (2023). Künstlich intelligent? *Informatik Spektrum*, 46(2), 63–64. <https://doi.org/10.1007/s00287-023-01536-w>
- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays - should professors worry? *Nature* (London).

‘Graue Literatur’

- Döbeli Honegger, B. (2023). info 7, Ausgabe 1/2023. Das Magazin für Medien, Archive und Information. <https://beat.doebe.li/publicaGons/2023-doebeli-honegger-chatgpt-info7.pdf>
- Mohr, G/ Reinmann, G./ Blüthmann, N. / Lübcke, E. / Kreinsen, E. (2023). Übersicht zu ChatGPT im Kontext Hochschullehre. Hamburg: Universität Hamburg. <https://www.hul.uni-hamburg.de/selbstlernmaterialien/dokumente/hul-chatgpt-im-kontext-lehre-2023-01-20.pdf>

Rickert, A. (2022). «Nehmen uns Maschinen das Schreiben ab?» akzente4 (November).
<https://blog.phzh.ch/akzente/2022/11/25/nehmen-uns-maschinen-dasschreiben-ab/#more-8065>

B. In der Literatur identifizierte Themen

Technisches

- Funktionsweise von LLMs und Chatbots
- Aktuelle Möglichkeiten von ChatGPT
- Blick in die Zukunft

Akademisches Lehren, studentisches Lernen und Schreiben

- Generelles
- Chancen

- Assistierender Einsatz von ChatGPT beim Lehren und Lernen
- (Schreib)didaktischer Einsatz
- Einsatz für konkrete Textsorten
- (Text)Beurteilung

- Risiken

- Fachlich-sachliche Korrektheit und Aktualität
- Wissenschaftliche Glaubwürdigkeit
- (Text)Beurteilung

- Blick in die Zukunft – Potenzielle Entwicklungen und Vorhersagen
- Neue Prüfungsformen

- Empfehlungen & Desiderate für die Hochschule
- Aufklärung und Entwicklung von (ethischen) Guidelines
- Konkrete Nutzung

- Ethik

- Plagiarismus
- Probleme bei der Beurteilung
- Copyright

Wissenschaftliches Arbeiten & Schreiben

- Chancen und Möglichkeiten
- Risiken und Nachteile
- Ethik
- Empfehlungen

Qualität AI-generierter Texte

- Sachlich-logische Korrektheit
- Diversität bezgl. Datengrundlage

Ethik

- Urheberrecht/ Copyright
- Wissenschaftlichen Textproduktion und Zitationspraxis
- Bezgl. Kosten vs. Nutzen für User
- Bezgl. Fairness in Lehre und Beurteilung

Empfehlungen

Potenzielle Entwicklungen

Desiderate

C. Dimensionen und Hauptcodes des Kodiersystems

Dimension «Einsatzbereich/ Nutzer:innen»

Einsatzbereich Institutionen

Einsatzbereich Allgemein

assistierender Einsatz

Intendierte Nutzer:innen von AI

Dimension «Themen»

Funktionsweise

ChatGPT Input

ChatGPT Output

Lernprozess

Schreibprozess

Leistungsüberprüfung

Ethik & Rechtliches

Dimension «Querthemen»

positive Auswirkung

Negative Auswirkungen

Problematisches

Paradigmenwechsel

Desiderate

Selbstpositionierung & Gefühle

Kompetenz(en)

Dimension «Erkenntnisinteresse Projekt»

potenzielles Workshopthema

Best-Practice-Beispiel

Offene Frage(n)