

III.
Öffentliche Sprache: Ethische Perspektiven auf demokratische
Debattenkultur

Ausgewählte Perspektiven auf den Einsatz von Künstlicher Intelligenz in den Medien

Petra Grimm, Susanne Kuhnert und Marcel Schlegel

Zusammenfassung

Der Beitrag präsentiert Erkenntnisse aus den Forschungsprojekten „IKID“ und „GEISST“, die sich mit dem Einsatz von Stimmen, die mit Künstlicher Intelligenz generiert wurden, in den Medien befassen. Zum einen wird aufgezeigt, wie KI-Kompetenz im Rahmen einer interdisziplinären Ausbildung von Medienstudierenden gefördert werden kann. Zum anderen werden ethische Leitlinien für den verantwortungsvollen Umgang mit KI-Stimmen in Medienorganisationen vorgestellt.

1. Einleitung

Künstliche Intelligenz (KI) beschäftigt gegenwärtig vor allem Politik, Wirtschaft und Wissenschaft (vgl. Menzel/Winkler 2018). Gerade Generative KI (GenKI), mit der Texte, Bilder, Videos, Code und Daten auf synthetischem Wege erzeugt werden kann, verbreitet sich rasant. Ihre einfache und (partiell) kostenfreie Nutzung, die menschliche Kommunikation nachahmt, ermöglicht auch Laien einen Zugang zu KI-Technik. Dergestalt dringen GenKI-Tools in immer mehr Lebensbereiche ein. Aufgrund ihrer Funktionalität erreichen sie auch das Medien- und Bildungssystem und fordern dortige Strukturen heraus.

Im Folgenden werden zwei Forschungsprojekte aus diesen Bereichen vorgestellt, die am *Institut für Digitale Ethik (IDE)* umgesetzt wurden: „GEISST – Generator für emotional individualisierbare Synthetik-Stimmen“ und „IKID – Interdisziplinäres KI-Exploratorium: Integrierte Lehre zur verantwortungsvollen Nutzung Künstlicher Intelligenz auf Basis physisch-virtueller Demonstratoren“.¹ Beide eint die Frage, welche Auswirkun-

1 Beide Projekte wurden maßgeblich an der *Hochschule der Medien Stuttgart (HdM)* umgesetzt und vom *Bundesministerium für Forschung, Technologie und Raumfahrt*

gen GenKI auf das Mediensystem haben könnte und wie diese aus ethischer Sicht zu reflektieren sind. Auch konzentrieren sich die Projekte auf KI-Techniken zur artifiziellen Nachbildung menschlicher Stimmen, richten ihr Augenmerk jedoch auf unterschiedliche Zielgruppen: So zielte IKID darauf ab, Medienstudierende auf eine verantwortungsvolle KI-Nutzung vorzubereiten. Dafür arbeiteten Lehrende unterschiedlicher Disziplinen zusammen, um ein didaktisches Konzept zum Erwerb von interdisziplinär verstandener KI-Kompetenz² zu entwickeln (vgl. Braßler 2020: 11–14; Grimm 2025a). Aus diesem wurden Lehrveranstaltungen abgeleitet, die den Prinzipien der interdisziplinären Lehre folgen (vgl. Braßler 2020: 15). Anhand eines dieser Projektseminare, das die Nutzung von *voice clones* und *deep fakes* im Journalismus thematisiert, zeigt der Aufsatz, wie ethische Perspektiven einfließen können in die Ausbildung von Medienstudierenden – als jene künftigen Praktiker:innen, die die KI-Ausbreitung im Medien- und indirekt auch im Bildungssystem maßgeblich mitgestalten werden.

In GEISST standen hingegen Medienschaffende selbst im Mittelpunkt. Sie wurden zunächst befragt mit dem Ziel, KI-Folgen für Medienorganisationen aus der Perspektive der Betroffenen zu erfahren. Zuerst wurden Haltungen sowie bedrohte Standards eruiert, die anschließend in wertebasierte Leitlinien für einen ethisch sensiblen GenKI-Einsatz mit Fokus auf Sprachsynthese überführt werden konnten. Die Handlungsempfehlungen, die Medienschaffenden dabei helfen sollen, die ethische Legitimität eines KI-Einsatzes zu überprüfen, werden im Folgenden in komprimierter Form vorgestellt.³

2. KI und der neue Strukturwandel: Potenziale und Risiken

GenKI trifft das Mediensystem in einer Zeit, in der es ohnehin angeschlagen ist (vgl. Krämer 2021: 28–35). Wie Umfragen zeigen, hält sich das Ver-

(BMFTR) gefördert, IKID zudem vom *Ministerium für Wissenschaft, Forschung und Kunst BaWü (MWK)*.

- 2 Darunter wird ein Bündel an Befähigungen verstanden, das sich – abhängig vom Verwendungszusammenhang von KI-Technologie – aus unterschiedlichen Fähigkeiten, Fertigkeiten, Kenntnissen, Methoden- und Schlüsselkompetenzen zusammensetzt. In der konkreten Lehrpraxis sollen jene Kompetenzen aus den Disziplinen Informatik, Wirtschaft, Recht und Ethik vermittelt werden, die für eine ganzheitliche Bearbeitung eines spezifischen KI-basierten Problems notwendig sind (vgl. Schlegel et al. 2024).
- 3 Die Leitlinien von Grimm, Kuhnert und Schlegel (2025) sind in vollständiger Form einsehbar unter dem Link im Literaturverzeichnis.

trauen der deutschen Bevölkerung in Medien zwar auf einem konstanten Niveau und ist, relativ zu vergleichbaren Ländern, weiterhin hoch (vgl. Jakobs et al. 2022: 387–396). Dennoch vertraute im Schnitt der Jahre 2016 bis 2023 weniger als jede:r Zweite den Medien „eher“ oder „voll und ganz“; jede:r Dritte gab an, medialer Berichterstattung nur „teils, teils“ zu vertrauen (vgl. Quiring et al. 2024: 2). Da Vertrauen als fundamentaler Wert von menschlichen Beziehungen gelten kann und auch die Basis der Relation zwischen Journalismus und Publikum bildet (vgl. Bentele 2021: 73), kann diese Entwicklung zulasten der Glaubwürdigkeit der Medien gehen (vgl. Tschopp et al. 2022: 330–333). Überdies führt die wachsende Bedeutung von Online-Plattformen dazu, dass die *gatekeeper*- und *agenda-setter*-Rolle der Massenmedien graduell ausgehebelt wird, was eine abermalige Schwächung der funktionalen Legitimität klassischer Berichtersteller:innen – der Massenmedien als Hersteller von Öffentlichkeit, als Kontrolleure der Politik sowie als Werbe- und Wirtschaftsplattform – bedeuten kann (vgl. Jarren 2019: 164–170).

Weil Soziale Medien es Nutzenden ermöglichen, sich netzöffentlich zu äußern, gelangen zusehends auch Halb- und Unwahrheiten sowie, auch mit Kalkül gestreute, Desinformationen in eine politische Öffentlichkeit, die durch die Plattformöffentlichkeiten ohnehin fragmentiert wird. Habermas (2022) deutet diese Entwicklung als „neuen Strukturwandel der Öffentlichkeit“ und erkennt darin eine Herausforderung für deliberative Demokratien. Bezüglich der öffentlichen Kommunikations-Infrastruktur problematisiert er, dass angesichts der Schwächung des traditionellen Journalismus die objektive Selektion von „relevanten und entscheidungsbedürftigen Themen“ zunehmend ausbleibe und so die Erstellung „*qualitativ gefilterter* Meinungen nicht mehr [zu] gewährleisten“ sei (ebd.: 65; Hervorhebungen im Original). Nicht allein Anzahl oder Frequenz, sondern die fehlende Wahrnehmung von Desinformation erzeuge eine „Deformation der Wahrnehmung der politischen Öffentlichkeit“ (ebd.: 64). Schaden nehmen Demokratien dann, wenn freie Medien und der Qualitätsjournalismus, welche als „Vierte Gewalt“ das politische System kontrollieren sollen, durch ein emergentes Zusammenwirken verschiedener Faktoren geschwächt werden: durch ökonomischen Druck auf die klassischen Medien; durch unfairen Wettbewerb mit den Plattformbetreibern; durch ein multifacettenreiches Aufkommen von (auch mit KI erzeugter und verbreiteter) Desinformation. Eine Dynamik, die den Vertrauensverlust der Bürger:innen in Medien und Politik beschleunigen könnte, weil Stabilität und Ordnung der liberalen Demokratie sowie die soziale Integration der Mitglieder einer Gesellschaft

wesentlich davon abhängen, dass sich Rezipient:innen darauf verlassen können, von öffentlichen Sprecher:innen gewissenhaft recherchierte Informationen zu erhalten. Weil Bürger:innen mediale Informationen für die eigene Meinungsbildung nutzen und diese gewissermaßen aus zweiter Hand erhalten, müssen sie sicher sein können, dass diese faktenbasiert und von Journalist:innen ohne bewusste Täuschungsabsicht verbreitet worden sind (vgl. Stapf 2021: 99–101). Deshalb stellen Wahrheit (als Wert) und Wahrhaftigkeit (als Eigenschaft von Kommunikator:innen) die zentralen Maximen des Journalismus dar (ebd. 2024: 315).

Angesichts der zügigen Verbreitung und Ausdifferenzierung von KI, deren Erzeugnisse auch über Soziale Medien gestreut werden können, sieht sich das Mediensystem erneut mit strukturellen Veränderungen konfrontiert, die auf der gesellschaftlichen Makro- und der institutionellen Mesoebene ebenfalls als systemische Risiken gelesen werden können. Auf der Mikroebene konkreter Praktiken eröffnet KI für Medienanbieter indes auch Potenziale, die nun exemplarisch referiert werden: So lassen sich mit textbasierter KI-Software journalistische Arbeitsroutinen effektiver gestalten (vgl. de Ruiters 2019: 1322f.). *Chatbots* können zum Beispiel verwendet werden, um automatisiert Texte zu erstellen, zu strukturieren, zu überprüfen, zu verbessern oder zusammenzufassen; um Kurz- und Funktionstexte wie Überschriften, Vorspanne, SEO-Texte und so weiter zu kreieren; um Layouts, Visualisierungen, Grafiken oder fotorealistische Abbildungen zu erzeugen (vgl. Prinzing 2023: 521–525). Weil sich mithilfe von KI auch riesige Datensätze oder digitale Archive auswerten lassen, können selbstlernende Systeme als Recherche-, Analyse- und Monitoring-Werkzeug Verwendung finden. Darunter fällt auch die Möglichkeit, Misinformation und Desinformation zu detektieren, was gerade für den investigativen Journalismus ertragreich sein kann (vgl. Kleemann 2023: 7).

Außerdem lassen sich mit KI mediale Beiträge kanal- und zielgruppengerecht aufbereiten und ausspielen, etwa entlang der Modalitäten von Sozialen Medien oder Suchmaschinen. Auch die hiermit beschriebene Personalisierung basiert auf der Erfassung, Aus- und Verwertung von Daten durch KI. Idealerweise münden KI-Verfahren zur algorithmischen Automatisierung von Zuschnitt, Zustellung und Verteilung medialer Produkte in der Verbesserung der Publikumsbindung und hiernach in einem gesteigerten Medienvertrauen. Neben KI, deren Ein-/Ausgaben Texte darstellen, können im Journalismus auch KI-Techniken genutzt werden, die (audio-)visuelle Produkte generieren oder modifizieren. So lassen sich mit GenKI defizitäre Originalaufnahmen interpolieren oder Transkriptions- und Übersetzungs-

leistungen (Fremdsprachen, Dialekte, inklusive Sprache, etc.) vollziehen. Überdies lassen sich auf artifiziellem Wege Personas erstellen, die (existierenden) Menschen in Aussehen, Mimik, Gestik, Stimmen und so weiter gleichen (vgl. Pawelec/Bieß 2021). Mit solchen *deep fakes* oder *voice clones* können mediale Darstellungsformen und Formate umgesetzt werden, die bislang auf menschliche Moderator:innen oder Kommentator:innen angewiesen waren.

Zusammengefasst lassen sich mit KI primär Ressourcen-, Effizienz- und Produktivitätsgewinne erzielen, etwa durch Zeit-, Personal- und Kosteneinsparungen. Das kann die wirtschaftliche Rentabilität erhöhen, aber gleichzeitig Arbeitsplätze gefährden, Berufsrollen transformieren oder diese obsolet machen (vgl. Albrecht 2023; WEF 2025). Je nach Einsatzkontext, Verwendungsform oder betroffenen Akteur:innen können sich KI-bedingte Einsparungen folglich auch als negative Charakteristika erweisen. So können mit GenKI Desinformationen kostenarm hergestellt, Personen geschädigt, Debatten verfälscht, Diskurse polarisiert oder Meinungsbildungsprozesse gestört werden (vgl. Grimm 2020; Diakopoulos/Johnson 2020: 2073; Krämer 2021: 25, 35–37). Das kann demokratische Pfeiler destabilisieren, Gesellschaftssysteme überfordern, jedenfalls aus demokratietheoretischer Sicht in kaum wünschenswerter Weise verändern (vgl. Pawelec/Bieß 2021: 44–48). Dies besonders dann, wenn KI dafür eingesetzt wird, Leistungen von Journalist:innen zu substituieren und somit die Funktionen der Medien auszuhebeln, was wiederum die Genese von Öffentlichkeit stören sowie dazugehörige Variablen wie politische Stabilität oder sozialer Zusammenhalt sabotieren könnte (vgl. Tschopp et al. 2022: 330–340).

Übergeordnet dürfte sich der Druck auf den Journalismus durch KI also weiter erhöhen. Zum einen, weil es dessen Aufgabe ist, den öffentlichen Raum „reinzuhalten“ von Falschem, Unwahren und Irrelevantem und Medienschaffende zunehmend damit beschäftigt sein könnten, Desinformation zu entlarven. Zum anderen könnten Medienorganisationen durch die Konkurrenz aus den Sozialen Medien und von KI-Anbietern vermehrt in ökonomische Bedrängnis geraten, was sie aus Effizienz- und Kostengründen dazu zwingen könnte, selbst auf KI-Systeme zurückzugreifen. Dies dürfte wiederum zulasten eines menschengemachten Journalismus gehen und das Vertrauen von Rezipient:innen in Medien negativ beeinflussen, sofern deren Qualität nachlässt.

Den Journalismus betreffende Organisationen, gerade solche Zusammenschlüsse, die um den Erhalt von Gütestandards bemüht sind, raten deshalb dazu, institutionelle Rahmenbedingungen für den verantwortungsvol-

len KI-Einsatz zu schaffen. Der *Deutsche Journalistenverband* (DJV 2025) etwa stellt die folgenden Forderungen auf:

1. Journalist:innen sollen weiterhin die Oberhand in der Recherche und Produktion von Inhalten haben, der Automatisierungsgrad sollte klug und maßvoll eingesetzt werden;
2. Medienunternehmen sollen sich bei dem Einsatz von KI auf ethischen Leitlinien verständigen;
3. Inhalte mit KI-Generierung sind deutlich und verständlich zu kennzeichnen;
4. Medienunternehmen sollten sich an dem Aufbau von werteorientierten Datenbanken und zertifizierten KI-Tools beteiligen;
5. Medienunternehmen sollen für ihre Mitarbeitenden Aus- und Weiterbildungsangebote für KI-Kompetenzen schaffen;
6. Der Gesetzgeber soll gerechte Vergütungsprinzipien für Journalist:innen durchsetzen.

Wie beschrieben stellt Vertrauen ein zentrales Gut in der Relation von Medien und Publikum dar. Die Vorschläge des *DJV* zahlen auf diesen Wert ein, weil idealtypisch argumentiert werden kann, dass sich die Einhaltung von ethischen Normen positiv auf die Aufgaben des Journalismus und dessen Wahrnehmung auswirken dürften: Zum einen können ethische Vorgaben für handelnde Akteur:innen als Orientierungsmarker dienen, KI verantwortungsvoll einzusetzen und technologische Veränderungen selbst aktiv mitzugestalten. Zum anderen müsste sich die Vertrauenswürdigkeit von Medienvertreter:innen erhöhen, wenn sich deren Rezipient:innen aus guten Gründen darauf verlassen können, dass Medienschaffende eine KI-Nutzung nur dann in Erwägung ziehen, wenn sie Risiken für Stakeholder reflektiert haben und ihre KI-Verwendung auch offenlegen. Berufsethisch lässt sich eine solche KI-Kennzeichnung mit den für den öffentlichen Raum zentralen Werten Wahrheit und Transparenz begründen: Journalist:innen zeigen ihrem Publikum dann, dass sie auf KI ohne Täuschungsabsicht setzen und weisen so ihre Wahrhaftigkeit nach. Das kann deshalb angebracht sein, weil es sich bei KI im Grunde um eine Nachahmung des Menschlichen handelt und künstlich erzeugte Produkte somit die Tendenz zur Täuschung graduell eingeschrieben ist. Der bewusste Verzicht auf eine Markierung von KI *kann* hingegen als Manipulation betrachtet werden (vgl. Schicha 2021: 175).

Demgegenüber kann eine Abwägung über die Zweckmäßigkeit und Legitimität von KI in Medien auf Seiten der Adressat:innen als Nachweis da-

für gewertet werden, dass Medienschaffende der Verantwortung als öffentliche Sprecher:innen nachkommen und ihrem gesellschaftlichen Auftrag auch nachkommen *wollen*. Daran zeigt sich, dass Absichten und Pflichtbewusstsein handelnder Akteur:innen sowie deren Kapazität zum sittlichen Umgang mit KI in der Medienproduktion als entscheidend für eine sozialverträgliche KI-Implementierung betrachtet werden können. Wie diese konkret ausgestaltet sein muss, hängt, wie allgemein für Technikimplementierung geltend, von den jeweiligen Einsatzkontexten und Nutzungsweisen ab, welche damit die entscheidenden Indikatoren darstellen, wenn eine Einschätzung darüber getroffen werden soll, wo, ob und wie KI aus Sicht der Ethik eingesetzt werden soll (vgl. Kuhnert/Grimm 2020: 253).

Um eine solche Abwägung vornehmen zu können, müssen die Beteiligten Funktionsweise und Folgen von KI verstehen beziehungsweise reflektieren und KI-Anwendungen selbst souverän anwenden können. KI-Mündigkeit kann dann mitunter auch bedeuten, den KI-Einsatz aus moralischen oder anderen legitimen Gründen abzulehnen. Schließlich gehört es zur Kernaufgabe von Journalist:innen, kritisch, sorgsam und objektiv zu sein – relevanten Dritten, aber auch der eigenen Rolle, dazugehörigen Pflichten und (technologischen) Phänomenen gegenüber. Am Ende begründet sich darin der Vertrauensvorschuss, den eine Bevölkerung ihren Medien entgegenbringt: Medien fungieren als Korrektive der Gesellschaft; ihr Dienst entfaltet sich an der und zum Wohle der Gesellschaft. Entsprechende Leistungen weisen sie durch eine redliche Arbeitsweise und eine durchdachte Technikverwendung nach. Der *DJV* bezieht sich in seinem Vorschlag zum Aufbau von KI-Kompetenz ebenfalls darauf. Die Frage, wie Medienstudierende zu einem souveränen KI-Einsatz im Sinne des Gemeinwohls befähigt werden könnten, steht im Zentrum von IKID.

3. Ansätze zur Vermittlung von interdisziplinärer KI-Kompetenz an Hochschulen

3.1 Das Lehrforschungsprojekt IKID – Vorgehen, Ziele und ethischer Schwerpunkt

Das Lehrforschungsprojekt IKID hatte zum Ziel, Studierende auf eine Medien- und Arbeitswelt vorzubereiten, die von KI geprägt sein wird. Dazu entstanden interdisziplinäre Lehrangebote, deren Ziel es war, KI-Kompetenz zu vermitteln. Die Kapazität zur kontextsensiblen Perspektivenerweite-

rung, welche an die Konzepte „Future Skills“ oder „Data Literacy“ (vgl. Ridsdale et al. 2015; Schüller 2019: 298ff.; Bandtel/Gläser 2021: 52, 60) anschließt, kann als konzeptionelle Erweiterung von Medien- und Digitalkompetenz gelesen werden. Darunter fallen Fähigkeiten, wie jene zur kritischen Reflexion, zur verständigungsorientierten Kommunikation, zur übergreifenden Folgenanalyse und zur konsensorientierten Aushandlung von Stakeholdern. Übergeordnet sollen jene Fähigkeiten, Eigenschaften und Identitätsdispositionen gestärkt werden, die Menschen von Maschinen unterscheiden – auch, um so typischen KI-Gefahren präventiv zu begegnen. Entlang des humanistischen Bildungsideals, das im Hinblick auf KI-Lehre auch der Ethikrat (vgl. 2023: 163–167) starkmacht, lassen diese als Befähigungen zum freien, vernünftigen, selbstbestimmten und kreativen Handeln, Urteilen und Entscheiden beschreiben. Da KI die Wissensarbeit zu substituieren droht, geht es überdies darum, bei Lernenden bestehende Kompetenzen zu erhalten und einen Fähigkeiten-Abbau (*deskilling*) zu verhindern (vgl. Reinmann 2023: 5–8).

Bezogen auf den Bildungskontext muss dabei speziell die Kapazität zur interdisziplinären Problem-bearbeitung erst noch aufgebaut werden, bei Studierenden wie Lehrenden (vgl. Albrecht 2023: 78; Ethikrat 2023: 167). Die IKID-Seminare wurden deshalb von Forschenden aus Informatik, Wirtschaft, Ethik und Recht zusammen entwickelt.⁴ Kern des Moduls, dem eine Propädeutik vorausging, stellten die „Integrierten KI-Projekte“ dar. In diesen wurden Lernende heterogener Studienherkünfte mit dem realitätsnahen Fall eines Medienunternehmens konfrontiert. Konkret wurde die *case study* eines Fernsehanbieters vorgegeben, der aus wirtschaftlichen Gründen KI-Klone einführen möchte, um Moderator:innen zu ersetzen und neue Formate zu entwickeln. Die studentischen Gruppen nahmen die Rolle eines interdisziplinären Consulting-Teams ein, das die Geschäftsführung beraten soll. Angestrebt wurde eine integrierte Lösung, die eine KI-Implementierung verfolgt, die inhaltlich innovativ, technisch umsetzbar, dem wirtschaftlichen Unternehmensziel zuträglich ist und rechtliche Vorgaben wie ethische Anforderungen bedenkt. Weil Kompetenzen in der Praxis erworben werden, sollten sich die Lernenden im geschützten Raum selbst

4 Im IKID-Projekt wurde in einer Serie von Whitepapers das didaktische Konzept skizziert sowie Lehrinhalte, Lehrmethoden und Kompetenzziele aller beteiligten Disziplinen dargelegt und gezeigt, wie durch den Einsatz einer „Sandbox“ und von KI-Demonstratoren eine anwendungsbezogene Lehre ermöglicht wird, den Link dazu ist im Literaturverzeichnis zu finden.

an KI-Software ausprobieren, Erfahrungen sammeln und Fehler machen dürfen (vgl. Schüller 2019: 303).

3.2 Ethische Anliegen

Aus Sicht der Lehrenden der Ethik ging es darum, den Studierenden einen ethisch reflektierten Umgang mit KI zu vermitteln, der die konzeptionellen Vorgaben der Technikfolgenabschätzung befolgt und übergeordnet „eine frühzeitige Auseinandersetzung mit den gesellschaftlichen Konsequenzen, aber auch Rahmenbedingungen der [KI-initiierten] Entgrenzungsdynamik [...]“ (Kehl 2021: 157) einsteuert. Dafür wurde ein verkürzter *ethics-by-design*-Prozess (vgl. Grimm 2025b) nachgestellt – mit den typischen Schritten (vgl. Manders-Huits 2011: 275)

- einer philosophisch-konzeptionellen Problemdiagnose;
- einer (teils auch empirisch ermittelten) Stakeholder-Auswertung;
- einer technischen Analyse von KI-Systemen;
- der Sensibilisierung für und Abschätzung von (möglichen) KI-Folgen für jeweilige Stakeholder und die Gesellschaft;
- der Ableitung ethischer Prinzipien und deren Berücksichtigung im Gestaltungsprozess von Technologien wie KI, von Konzepten wie Business-Plänen oder in Bezug auf KI-basierte Darstellungsformen.

In der Propädeutik wurden zunächst Grundlagen der Digital- und KI-Ethik vermittelt. Das umfasste passende Methoden, etwa Reflexions- und Analysewerkzeuge zur Stakeholder-Ermittlung und Folgenabschätzung. Außerdem wurden bestehende ethische Kodizes ausgewertet, deren Maximen auch in der *case study* berücksichtigt werden sollten.⁵ In den KI-Seminaren kamen Diskussionsformate wie Rollenspiele zur Anwendung, in denen sich Teilnehmenden die Fähigkeit zur Perspektivübernahme aneignen, über ethische Fragen zu KI austauschen, soziale Folgen von KI besprechen, Haltungen, Bedarfe und Einstellungen artikulieren und ihre Erkenntnisse auf die Fallstudie übertragen sollten. Teil der Prüfungsleistung war es, eine

5 Konkret wurde das *Aktantenmodell*, die *Wertematrix*, das *Futures Wheels* und weitere Methoden gelehrt. Bezüglich ethischer Standards waren die *TARES*-Prinzipien (vgl. Baker/Martinson 2001), der *Pressekodex* (vgl. Deutscher Presserat 2025) und die *Pariser Charta* zu KI und Journalismus (vgl. Reporter ohne Grenzen 2023) mitsamt bedeutsamer Werte und Prinzipien des Journalismus (Vertrauen, Wahrheit, Unabhängigkeit, Neutralität, Glaubwürdigkeit, Objektivität, etc.) Gegenstand der Seminare.

begründete Werte-Topografie für das Medienunternehmen herzuleiten. Daraus wiederum sollten die Studierenden ethische Leitlinien für ihren formatspezifischen GenKI-Einsatz entwickeln.

3.3 Ausgewählte Implikationen von KI für die Bildung

Lehrende stehen vor der Aufgabe, sich und ihre Studierenden auf KI vorzubereiten und damit auf einen Gegenstand, der sich simultan zur individuellen Aneignung und didaktischen Vermittlung verändert (vgl. Ethikrat 2023: 163–182 und Schlegel et. al 2024). Von Lehrenden wird entsprechend erwartet, anpassungsfähig und experimentierfreudig zu sein (vgl. Pancratz et al. 2023: 88). KI zwingt sie daher dazu, Lehr-/Lernansätze und Prüfungskulturen zu modifizieren (vgl. Albrecht 2023: 78). Aus gutem Grund denkt man nur an die vielfältigen Einsatzmöglichkeiten von GenKI im Hochschulbereich: als Recherche-Werkzeug; zum Organisieren von Wissen; zum Strukturieren, Verfassen, Überprüfen und so weiter von Texten; zum Erstellen von Präsentationen oder Visualisierungen sowie für weitere Zwecke. (vgl. Reinmann 2023).

Grundlegend zählen die IKID-Seminare auf diese Anforderungen ein, zumal sie im Rahmen der Lehrforschung evaluiert und angepasst wurden. Zudem adressiert das überfachliche Vorgehen eine strukturell bedingte Herausforderung von Bildungseinrichtungen: die domänenspezifische Fokussierung, die dazu führt, dass Studierende in nur einer Disziplin ausgebildet werden und neue Phänomene daher primär aus Sicht ihres Fachs einzuschätzen lernen. Weil KI jedoch als gesellschaftliches Querschnittsphänomen einzuordnen ist, bedarf es einer integrierten Lehre (vgl. Schlegel et al. 2024). Um tragfähige Lösungen und (neue) Methoden für künftige Herausforderungen entwickeln zu können, müssen vormals getrennte Fachgrenzen graduell überwunden werden, zumindest braucht es die Bereitschaft dazu (vgl. Grimm 2025a; Braßler 2020). Das erfordert experimentelle Lehrkonzepte, Lehr- und Lernumgebungen, in deren Entwicklung die Anspruchsgruppen – entlang der Maximen von *ethics by design* – miteinbezogen werden müssen. Lehrinnovation entwickelt sich dann als Ko-Konstruktion zwischen Forschenden, Lehrenden und Studierenden, die auch als Nutzende und damit als Expert:innen in eigener Sache verstanden werden müssen. So gestalten sie die KI-Integration im Bildungsbereich in indirekter Weise mit. Da Kompetenzen in der Praxis erworben werden, eignen sich Lehrende und Lernende die notwendigen Fähigkeiten im Umgang

mit KI sukzessive an und bilden idealerweise eine Haltung in Technikfragen heraus. Darüber hinaus definiert sich Kompetenz als die Kapazität, in komplexen Situationen, die sich stetig verändern, kontingente Verläufe annehmen und offene Ausgänge zeitigen, kreative Lösungen zu entwickeln – eine bislang genuin menschliche Gabe (vgl. Lenbet 2024: 223–225). Wie gezeigt, kann KI als ein solches emergentes Phänomen verstanden werden.

Dabei adressiert das kompetenzorientierte Vorgehen von IKID erneut den Wert des Vertrauens: Studierende, die zur Nutzung von KI und zur Abschätzung von KI-Folgen befähigt wurden, entwickeln zum einen Vertrauen in ihren Umgang mit KI. Sie sind also fähig, einzuschätzen, in welchen *use cases* der KI-Einsatz gerechtfertigt sein kann und wann derlei Systemen misstrauisch zu begegnen ist – zum Beispiel, wenn Werte des sozialen Zusammenlebens bedroht sind. KI-Mündigkeit kann zum anderen auch bedeuten, zu erkennen, welche menschlichen Dispositionen gegenüber einer Maschine zu schützen sind. Um die KI-Transformation unter Beteiligung der Bevölkerung gelingend zu gestalten, braucht es einerseits technische Kompetenz und andererseits Akzeptanz gegenüber neuen Technologien. Wie Studien zeigen, korreliert das eine mit dem anderen: So weisen Arnold et al. (2020: 24, 34–41) nach, dass Menschen KI-Systemen weniger misstrauisch begegnen oder sich diesen seltener verweigern, wenn sie deren Auswirkungen für beherrschbar halten. Einschätzungen zur Kontrollierbarkeit von KI sind dann höher, wenn Befragte solche Tools selbst nutzen. Übergeordnet sichert KI-Kompetenz daher die menschliche Selbstbestimmung und Autonomie; (Selbst-)Vertrauen in menschliche Fähigkeiten also.

4. Ethisch reflektierter Umgang mit künstlichen Stimmen in der Medienpraxis

4.1 Begründung, Beschreibung und Ziele des Forschungsprojekts GEISST

Das Forschungsprojekt GEISST beschäftigte sich mit dem Einsatz von synthetischen und geklonten Stimmen im Journalismus. Ausgegangen wurde von einem *use case*, bei dem mit KI, abgestimmt auf verschiedene Ausspielungskanäle, variierende Versionen eines audiovisuellen Nachrichtenbeitrags generiert werden. Der ethische Projektschwerpunkt begründet sich zum einen darin, dass KI-erstellte Stimmen in einem grundlegenden Bereich menschlicher Kommunikation eingreifen: die gesprochene Spra-

che. Sprache ist nicht nur Träger von Information, sondern auch Ausdruck von Identität. GenKI kann Sprache und Stimmen technisch reproduzieren. Damit verschiebt sich die Produktion gesprochener Sprache von einem menschlich-körperlichen Ausdruck hin zu einem technisch vermittelten Prozess. KI-Stimmen können dadurch das Erleben von Sprache verändern. In journalistischen Kontexten berührt dies grundlegende Fragen von Glaubwürdigkeit, Transparenz und Verantwortung. Sowohl Sprache als auch Stimmen stellen Identitätsmerkmale und damit ein zentrales Instrument journalistischer Integrität dar.

Zum anderen begründet sich der ethische Fokus darin, dass Medien ein Kulturgut und aufgrund ihrer systemerhaltenden Funktion in Demokratien qua Verfassung geschützt sind. Um Qualitätsstandards zu garantieren und um Sorge dafür zu tragen, dass journalistische Gütekriterien nicht ökonomischen Zielen zum Opfer fallen, braucht es rechtliche Rahmensetzungen und ethische Normen. Letztere wurden in GEISST in Form von Leitlinien umgesetzt, welche aus einer Interviewstudie mit Medienschaffenden abgeleitet wurden (vgl. Grimm/Kuhnert/Schlegel 2025). Mit dieser sollte ermittelt werden, welchen Werte die Befragten beim Einsatz von GenKI zur Nachahmung von Stimmen im Journalismus eine exponierte Stellung zusprechen. Durch den Stakeholder- und Anwendungsbezug, der durch die Befragung von Medienpraktiker:innen sichergestellt wurde, sollte erreicht werden, dass die Handreichungen realitätsnah und entsprechend wirkungsvoll sind.⁶

4.2 Ablauf und Methode der Befragung

Im Rahmen der qualitativen, nicht-repräsentativen Interviewstudie wurden zwölf Personen im Alter von Mitte 20 bis Anfang 60 Jahren befragt. Die Redakteur-, Journalist- und Sprecher:innen sind in die Nachrichtenproduktion von privaten Medienunternehmen oder öffentlich-rechtlichen Einrichtungen in den Bereichen Rundfunk, Print oder TV eingebunden. Die teilstrukturierten Einzelinterviews wurden in Präsenz oder via Online-Vi-

6 Die Beteiligung von Stakeholdern, umgesetzt in Form von empirischen Methoden wie Befragungen, Interviews und so weiter, stellt eine zentrale Maxime von ethischen *by-design*-Ansätzen dar. Gleichzeitig ist zu betonen, dass empirische Befunde allein nie als Grundlage für normative Empfehlungen ausreichen, aber zur Überprüfung der „Realitätsadäquatheit“ einer Medienethik dienen können (vgl. Rath 2013: 296).

deotelefonie durchgeführt. Für die Auswertung wurden die 50- bis 90-minütigen Gespräche transkribiert und anonymisiert. Ziel der Befragung war es, begründete Einschätzungen darüber zu erlangen, unter welchen Bedingungen Medienschaffende einen KI-Einsatz in der Nachrichtenproduktion akzeptieren (würden) (F1); welche Werte sie in diesem Verwendungs-Zusammenhang betonen (F2); zudem wurde Bezug genommen auf bestehende ethische Rahmenwerke und die dort hervorgehobenen Normative von Wahrheit und Wahrhaftigkeit, die in maßgeblichen Kodizes prominent platziert sind.⁷ Im Zentrum der Befragung standen diese Forschungsfragen:

- F1: Wie steht es um die Akzeptanz von Generativer KI im Medienbereich und dem Nachrichtenjournalismus, speziell bei Journalist:innen und Redakteur:innen?
- F2: Was verstehen die Befragten unter „gutem Journalismus“? Welche Standards sollen in Zukunft beim Einsatz von Generativer KI gelten?
- F3: Welche Bedeutung hat der Wert der Wahrheit für Journalist:innen und Redakteur:innen und wie definieren diese Wahrheit in Bezug auf den Einsatz von Generativer KI?

Das Vorgehen wurde an die Vorgaben des *value-sensitive-design*-Ansatzes angelehnt, der sich anbot, weil er den Fokus auf die Beteiligung von Stakeholdern bei der Entwicklung ethischen Standards richtet. Hier wird ein Wert recht weit gefasst: „What is important to people in their lives, with a focus on ethics and morality“ (Friedman/Hendry 2019: 24). Die Befragten sollten in ein relativ freies Erzählen gebracht werden, während Interviewende gleichzeitig dafür sorgen mussten, die Gespräche auf das Thema KI im Journalismus zu konzentrieren (vgl. Müller/Grimm 2016). Zunächst wurden den Befragten deshalb verschiedene KI-Anwendungsfälle präsentiert, zu denen sie assoziative Eindrücke schildern sollten. Durch diesen offenen Initialimpuls sollte ein lockerer Einstieg ins Thema erreicht werden. Die Spanne ausgewählter *use cases* war dabei bewusst breit gefächert und beschränkte sich nicht nur auf einen KI-unterstützten Journalismus. Im Anschluss wurde die Befragung auf die Möglichkeiten der KI-Nutzung in der täglichen Arbeit der Befragten und damit auf deren persönliche Erfahrungswelt gelenkt. Auf diese Weise konnten Praxisnähe und Kontextbezug

7 So heißt es beispielsweise in Ziffer 1 des „Pressekodex“ des *Deutschen Presserates*: „Die Achtung vor der Wahrheit, die Wahrung der Menschenwürde und die wahrhaftige Unterrichtung der Öffentlichkeit sind oberste Gebote der Presse. Jede in der Presse tätige Person wahrt auf dieser Grundlage das Ansehen und die Glaubwürdigkeit der Medien.“

gewährleistet und gleichsam explizit wie implizit vorliegende Wertvorstellungen identifiziert werden.

4.3 Ausgewählte Ergebnisse

Ein zentraler Befund ist, dass die im Journalismus bestehenden normativen Vorgaben weiterhin als axiomatischer Bezugspunkt fungieren: Alle Befragten verweisen auf bekannte Ethik-Kodizes, vor allem den „Pressekodex“. Bekannte Professionsstandards betrachten sie als Maßstab für guten Journalismus und bestätigen somit deren Gültigkeit und Aktualität. Zudem verdeutlichen die Ergebnisse, dass Medienschaffende journalistische Qualität an die Einhaltung von grundlegenden Prinzipien koppeln, speziell an redaktionelle Sorgfaltspflichten und institutionelle Verantwortlichkeiten. Die tragende Rolle bewährter Praktiken wird dabei immer wieder hervorgehoben und speziell mit der Gewährleistung von Perspektivenvielfalt, dem journalistischen Neutralitäts- und Transparenzgebot, redaktionellen Prüfungs- und Kontrollschleifen und inhaltlichen Einordnungen verbunden, beispielsweise der Trennung von Fakten und Meinungen – exemplarisch hierzu die Äußerung einer Journalistin (TN 9):

„[...] es gibt dieses Sich-nicht-mit-einer-Sache-gemein-Machen: also auch mal die Perspektiven zu wechseln. Mal ganz bewusst auch eine Distanz aufzubauen, weil man natürlich schnell [...] auch subjektiv wird [...]. Ich glaube, guter Journalismus entsteht im Team, durch Diskussionen, in Redaktionskonferenzen, Vier-Augen-Prinzip [...].“

Der Einsatz von GenKI im Journalismus wird keineswegs pauschal abgelehnt, allerdings wird er an zentrale Forderungen geknüpft: Für alle Befragten steht fest, dass technologische Mittel die Redaktionen nicht von deren Verantwortungsübernahme für die Publikationen entbinden und Maschinen somit den Menschen nie ersetzen sollten. Vielmehr sollen technische Mittel, auch KI, unter menschlicher (Letzt-)Kontrolle verbleiben. Dies ist vor allem jenen Personen wichtig, die direkt betroffen wären, wenn ihre Stimme für einen KI-erstellten Beitrag genutzt wird. Die Aussagen aller Befragten verweisen implizit auf eine professionsethische Grundeinstellung gegenüber Technologien und einer auch politisch interpretierbaren Erwartungshaltung, die so lauten könnte: Technologische Innovationen sollen Arbeitsweisen erleichtern, dürfen gleichwohl nicht dazu führen, dass ethische Grundwerte, grundlegende redaktionelle Aufgaben oder gesellschaft-

liche Leistungen des Journalismus relativiert werden. Neue Technologien werden damit als unterstützendes Werkzeug betrachtet. TN II skizziert dies sogar in einem historischen Zusammenhang:

„Das war nie so, dass wir ohne Technik ausgekommen sind [...]. Und unter dem Aspekt [...] sind diese Algorithmen [und] KI-Tools ja im Grunde [...] nichts anderes als jedes andere Tool, das wir in den vergangenen Jahrzehnten und Jahrhunderten genutzt haben – von der Schreibmaschine über die Rechtschreibprüfung bis zum Ganzseitenumbruch bis zur automatisierten Sendeabwicklung.“

Die Interviews offenbaren ein subtil vorliegendes Misstrauen gegenüber KI-Stimmen in den Medien, das gleichsam ambivalent wahrgenommen wird: Einerseits werden technische Innovationen akzeptiert und die KI-Integration in Medienorganisationen als quasi unvermeidlich hingenommen. Andererseits äußern einige Befragte Bedenken hinsichtlich der Authentizität synthetischer Stimmen und deren Einfluss auf Wahrnehmungen von Wahrhaftig-, Vertrauens- und Glaubwürdigkeit, speziell von Sprecher:innen, so etwa TN 3:

„[...] Ich bin zurückhaltend mit künstlichen Stimmen, weil ich es nach wie vor wichtig finde für die Authentizität des Themas [oder] Beitrags, dass die Menschen selbst [...] zu Wort kommen, weil es vermittelt auch [...] Glaubwürdigkeit, wenn der Mensch selber in seiner Originalstimme sprechen kann.“

Ein Teil der Befragten äußert explizit – und ohne dazu befragt worden zu sein – den Bedarf nach mehr Qualifizierung: Dies betrifft sowohl die technischen Aspekte von KI als auch den professionellen Umgang mit und Fähigkeiten zur Detektion von Desinformation. In einem ähnlichen Zusammenhang stellen alle Befragten Transparenz als essenziellen Wert heraus, dies im Sinne von Sichtbarkeit und Kennzeichnung. Es soll für die Rezipient:innen eindeutig erkennbar sein, wenn eine Stimme KI-erzeugt wurde. Eine fehlende Kennzeichnung wird als potenziell täuschend eingestuft, weil es das Vertrauen in ein Medium oder den Journalismus insgesamt beeinträchtigen könne, so unter anderem TN 5:

„Wir verlieren an Glaubwürdigkeit. Also wenn etwas sehr offensichtlich nicht so produziert wurde und etwas sehr offensichtlich nicht wirklich so war, wie es dann publiziert wird, dann glaubt man ja auch alles andere nicht mehr.“

Der möglicherweise drohende Jobverlust durch KI und angesichts ohnehin zunehmend prekärer Arbeitsverhältnisse von Journalist:innen (vgl. Hainitzsch/Rick 2021) wird von mehreren Personen adressiert, so auch von TN 8:

„Als Berufstätiger oder Praktiker sehe ich es halt kritisch, weil ich mir denke: Da sind ja wieder ein paar Jobs von Nachrichtensprecherinnen und Nachrichtensprecher, die ersetzt werden.“

Alle Interviews eint, dass der KI-Einsatz mit ambivalenten Einschätzungen einhergeht. Keine:r der Befragten lehnt KI kategorisch ab, dies trotz eines ausgeprägten Bewusstseins über mögliche Risiken, welche die Medienschaffenden auch persönlich treffen könnten. Einerseits stellen die Interviewten die Potenziale von KI heraus. Andererseits treibt sie die Sorge um, dass durch KI menschliche Aspekte ihrer Arbeit verloren gehen könnten – die am häufigsten geäußerten Bedenken. Es zeigt sich an dieser Stelle ein latenter Technikdeterminismus. Denn die Befragten nehmen die bislang weithin ungehemmt fortschreitende KI-Integration gewissermaßen als gesetzt hin und akzeptieren diese und darüber hinaus auch das Narrativ, dass diese Entwicklung unaufhaltbare wäre. Gleichzeitig legen sie Zuversicht und Vertrauen eben nicht in KI, sondern in den Menschen: in die Hoffnung, dass Menschen lernen werden, KI-bedingten Herausforderungen auf ebenso menschlicher, sozialer und gerechterweise zu begegnen.

4.4 Theoretische Grundlagen der ethischen Leitlinien

Die in GEISST entwickelten Leitlinien orientieren sich einerseits an bestehenden journalistischen Standards, andererseits an einer aristotelisch geprägten Angewandten Ethik, welche das teleologische Ziel des „guten Lebens“ verfolgt. Nach Spiekermann (2019) folgt die Digitale Ethik ebenfalls dieser eudämonistischen Tradition. Für Weber (2017: 36), der journalistische Aufgaben der politischen Sphäre zurechnet, bedeutet moralisches Handeln, für die absehbaren Folgen des eigenen Tuns einzustehen. Verantwortungslosigkeit liegt vor, wenn künftig mögliche oder bereits realisierte Auswirkungen ignoriert werden oder Journalist:innen unsachlich berichten (ebd.: 63). Hier offenbart sich erneut der Wert der Wahrheit, dessen Einhaltung im Journalismus Weber ebenfalls als wesentlich erachtet. Weber folgend, sollen Verantwortungsethiker:innen in ihren Folgenbetrachtun-

gen zudem menschliche Schwächen stets weitsichtig mitbedenken. Entsprechend dürfen sie Konsequenzen ihres Handelns nicht auf andere abwälzen.

Auf aktuelle Medienkontexte bezogen, könnte das beispielsweise bedeuten, dass einem Entscheidenden, der/die aus Kostengründen eine/r Sprecher:in durch einen KI-Avatar ersetzen möchte, sich der Konsequenzen bewusst sein muss. Wirtschaftliche Argumente allein reichen etwa nicht aus, um den Jobverlust des/der Kolleg:in zu rechtfertigen. Entsprechend sollte sie legitime Gründe für ihr Vorgehen präsentieren können. Zum Beispiel könnte der KI-Einsatz moralisch vertretbar sein, wenn er eine Medienproduktion überhaupt erst ermöglichen würde.

Daran zeigt sich: Technikethik bewegt sich stets im Spannungsfeld individueller und institutioneller Verantwortung (vgl. MacCormac 1993). Beide Seiten, Medienschaffende wie Medienorganisationen, müssen sich gemeinsam um einen verantwortungsvollen KI-Einsatz bemühen. Damit muss die KI-Verwendung in den Medien das Produkt eines (redaktionellen) Abwägungsprozesses sein, in den alle Bezugsgruppen eingebunden werden. Aufgrund der gesellschaftlichen Verantwortung von journalistischen Medien müssen sich reflektierte Praktiken damit nicht zwangsläufig an unmittelbaren ökonomischen Vorteilen messen lassen, wenngleich das eine das andere nicht ausschließt: Der größte Nutzen von moralischem Handeln liegt laut Weber vielmehr in langfristigen Erwägungen. So kann moralisches Handeln das Mediensystem stabilisieren und das Vertrauen in Medienvertreter:innen erhöhen.

In GEISST wird diese teleologische Perspektive aufgegriffen. Deren Maximen werden gleichsam ausdrücklich auf das Ziel eines guten Journalismus bezogen. Verantwortung zeigt sich in der sozial-verträglichen Kalkulation der Auswirkungen des KI-Einsatzes unter Berücksichtigung der funktionalen Bedeutung des Journalismus für eine stabile Demokratie.

Nun steht die empirische Folgenanalyse von KI in Medien aufgrund des frühen Entwicklungsstatus von selbstlernenden Systemen noch am Anfang; insbesondere Langzeitstudien fehlen (Székely et al. 2025). Gleichwohl zeigen erste Forschungen negative Effekte auf das Verstehen und Erinnern von Nachrichten, wenn diese durch synthetische Stimmen vermittelt wurden (vgl. Gong 2023). Bekannt ist allgemein, dass (audio-)visuelle Inhalte ein enormes persuasives Potenzial aufweisen: Zum einen, weil sie eine Augenzeugenschaft und damit eine dokumentarisch-objektive Wahrnehmung suggerieren (vgl. Schicha 2021: 175f.); zum anderen, weil menschliche Stimmen als authentisch wahrgenommen werden können (vgl. de Ruiter 2019: 1322; Singh 2020: 69f.). KI-generierte Medien stehen

damit im Verdacht, den epistemologischen Wert von Fotos, Videos und Audios zu bedrohen. Weil synthetische Inhalte schon jetzt kaum noch vom Original unterscheidbar sind (vgl. Frank et al. 2024), entsteht durch KI ein enormes Desinformations- und Polarisierungspotenzial. Da ein fehlendes Bewusstsein allgemein gesprochen die Anfälligkeitswahrscheinlichkeit für Polarisierung erhöht (vgl. Hagen et al. 2018: 18), sind letztlich Aufklärung, Bildung und KI-relevante Kompetenzen die Schlüsselfaktoren für einen kritischen Umgang mit Medieninhalten.

Die GEISST-Leitlinien zahlen auf diese Qualifizierungsziele ein. Weil sie Medienpraktiker:innen nach Bedarfe und Anliegen fragen und diese um KI-spezifische Einschätzungen bitten, erfüllen sie auch verantwortungsethische Maximen. Durch ihren empirischen Gehalt tragen sie ferner dazu bei, die KI-spezifische Risikoabschätzung voranzutreiben.

4.5 Ethische Leitlinien in GEISST

Im Folgenden sollen die Leitlinien, die im GEISST-Projekt erarbeitet wurden, in komprimierter Form präsentiert werden.⁸ Basis derselben sind zum einen die genannten theoretischen Überlegungen (4.4.). Zum anderen stellen sie die Ableitungen der Ergebnisse der Interviewstudie dar (4.3). Da sich Normen entlang des *value-sensitive-design*-Ansatzes auf Werte beziehen lassen müssen, wurde zunächst eine Werte-Topografie angefertigt (Abbildung 1), die aus den Befragungsergebnisse extrahiert wurde.

- *Ziel- und Kontextbindung*: Künstliche Stimmen sollen nur zu journalistischen Zwecken (Information, Aufklärung etc.) eingesetzt werden, um die Integrität der Medienorganisation zu bewahren. Der Einsatz soll im jeweiligen Kontext geprüft und begründet werden. Wirtschaftliche Vorteile allein rechtfertigen einen verantwortungsbewussten Einsatz nicht.
- *Kennzeichnungspflicht*: Jede KI-Stimme soll als solche transparent erkennbar sein. Die technische Herkunft der Stimme soll offengelegt und eindeutig gekennzeichnet werden. Rezipient:innen sollen dergestalt weiterhin selbstbestimmte Entscheidungen über ihre Mediennutzung treffen können.
- *Schutz vor Täuschung*: KI-Stimmen sollen nicht zur emotionalen oder politischen Beeinflussung und nie mit manipulativer Absicht eingesetzt werden. Der KI-Einsatz darf die Verpflichtung zur wahrheitsgemäßen

8 Vollständige Fassung siehe bei Grimm/Kuhnert/Schlegel 2025.

und unvoreingenommenen Informationsvermittlung nicht untergraben. Bei sensiblen Themen soll auf KI-Stimmen verzichtet werden.

- *Rechenschaftspflicht*: Die Verantwortung für eine KI-Nutzung soll bei den Redaktionen verbleiben. Entscheidungen zum KI-Einsatz müssen dokumentiert und nachträglich einsehbar sein. Medienhäuser sollen transparente Beschwerdemechanismen etablieren, damit das Publikum mögliche Täuschungserfahrungen melden kann.
- *Schutz menschlicher Sprecher:innen*: Das Nachbilden einer Stimme soll nur mit ausdrücklicher, informierter und jederzeit widerrufbarer Zustimmung der betroffenen Person erlaubt sein. Eine KI-Stimme darf nicht über den ursprünglich vereinbarten Kontext hinaus verwendet oder verändert werden. Stimmen sind durch Persönlichkeitsrechte geschützt, die von Redaktionen zu respektierten sind – auch dann, wenn die Stimme technisch neu erzeugt wurde, aber einer realen Person ähnelt.
- *Ethische Gestaltung affektiver Sprechmodulation*: Tonfall, Tempo und Klangfarbe von KI-Stimmen sollen dem unabhängigen Informationsauftrag entsprechen. Eine emotionale Aufladung, etwa um Vertrauen oder Aufmerksamkeit zu erzeugen, ist nicht zulässig. Im Nachrichtenkontext gilt eine neutrale und sachliche Tonalität als Standard.
- *Schulung und Weiterentwicklung*: Journalist:innen sollen die Wirkung synthetischer Stimmen kritisch reflektieren und verantwortungsvoll einsetzen. Medienorganisationen sollten dafür Qualifizierungsangebote schaffen und sich an Beforschung von KI-Stimmen beteiligen. Auf diese Weise befähigen sie Mitarbeitende zur verantwortungsbewussten Teilhabe an der KI-Entwicklung.

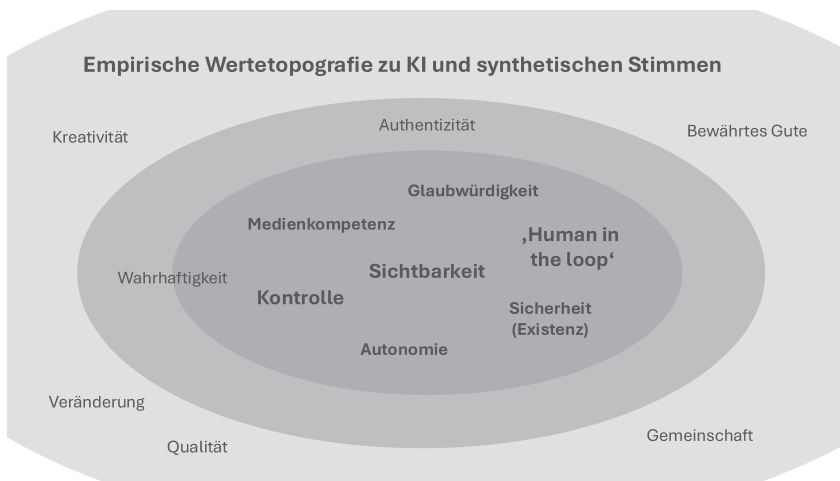


Abbildung 1: Empirische Wertetopografie zu KI und synthetischen Stimmen. Die Positionierung der dargestellten Werte reflektiert ihre relative Bedeutung in den Expert:innenaussagen, wobei eine geringere Distanz zum Zentrum mit einer höheren Nennungshäufigkeit einhergeht (eigene Darstellung).

5. Resümee

Die im Rahmen der Projekte IKID und GEISST gewonnenen Erkenntnisse verdeutlichen, dass der Einzug von KI in den Medienbereich sowohl Chancen als auch erhebliche Herausforderungen mit sich bringt. Insgesamt verdeutlichen beide Projekte, dass ein rein technischer oder ökonomischer Blick auf die KI-Verwendung in den Medien zu kurz greift. Eine verantwortungsvolle Einführung von KI erfordert daher eine enge Verzahnung von ethischer Reflexion, interdisziplinärer Bearbeitung und der Bewahrung journalistischer Kernwerte wie besonders Vertrauen, Wahrheit und Transparenz. Die Zukunft des Journalismus im Zeitalter der KI wird maßgeblich davon abhängen, ob es gelingt, die Potenziale der Technologie nutzbar zu machen, ohne die demokratisch relevanten Funktionen der Medien zu gefährden, die ein menschengemachter Journalismus etablierte und bislang absichert. Dies erfordert nicht nur die Einhaltung von ethischen Leitlinien, sondern auch die konsequente Ausbildung von Medienprofis, die in der Lage sind, die Technologie als Werkzeug unter menschlicher Kontrolle zu gestalten und zu steuern – und damit auch eine ethisch orientierte KI-Kompetenz.

Literatur

- Albrecht, Steffen* (2023): ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen (= TAB-Hintergrundpapier Nr. 26), Berlin.
- Arnold, Norbert et al.* (2020): „Wenn die KI unser Assistent bleiben kann, dann können wir viel draus ziehen“ – Künstliche Intelligenz in Einstellungen und Nutzung bei unterschiedlichen Milieus in Deutschland, Konrad-Adenauer-Stiftung (online unter: <https://www.kas.de/documents/252038/7995358/Kuenstliche+Intelligenz+in+Einstellungen+und+Nutzung+bei+unterschiedlichen+Milieus+in+Deutschland.pdf> – letzter Zugriff: 15.9.2025).
- Baker, Sherry / Martinson, David L.* (2001): The TARES Test: Five Principles for Ethical Persuasion, in: *Journal of Mass Media Ethics* 16 (2&3/2001), S. 148–175.
- Bandtel, Matthias / Gläser, Christine* (2021): Potenziale digitaler Lehre, in: Johanna Ebeling / Henning Koch / Alexander Roth-Grigori (Hg.), *Kompetenzerwerb im kritischen Umgang mit Daten*, Berlin, S. 51–62.
- Bentele, Günter* (2021): Der Wahrheits- und Wahrhaftigkeitsanspruch in einer Ethik der Öffentlichen Kommunikation, in: Christian Schicha / Ingrid Stapf / Saskia Sell (Hg.), *Medien und Wahrheit. Medienethische Perspektiven auf Desinformation, Lügen und „Fake News“*, Baden-Baden, S. 59–77.
- Braßler, Mirjam* (2020): *Praxishandbuch Interdisziplinäres Lehren und Lernen. 50 Methoden für die Hochschullehre*, Weinheim.
- De Ruiter, Adrienne* (2021): The Distinct Wrong of Deepfakes, in: *Philosophy & Technology* 34, S. 1311–1332.
- Deutscher Ethikrat* (2023): *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz* (= Stellungnahme Deutscher Ethikrat, 20. März 2023) (online unter: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf> – letzter Zugriff: 15.6.2025).
- Deutscher Journalistenverband* (DJV 2025): Positionspapier bezüglich des Einsatzes Künstlicher Intelligenz im Journalismus (online unter: <https://www.djv.de/mediennpolitik/kuenstliche-intelligenz/> – letzter Zugriff 9.8.2025).
- Deutscher Presserat* (2025): *Publizistische Grundsätze (Pressekodex)*, 19. März 2025 (online unter: <https://www.presserat.de/pressekodex.html> – letzter Zugriff: 2.7.2025).
- Diakopoulos, Nicholas / Johnson, Deborah* (2021): Anticipating and addressing the ethical implications of deepfakes in the context of elections, in: *New Media & Society* 23 (7/2021), S. 2072–2098.
- Frank, Joel et al.* (2024): A Representative Study on Human Detection of Artificially Generated Media Across Country, in: *IEEE Symposium on Security and Privacy (SP)*, San Francisco, S. 55–73.
- Friedman, Batya / Hendry, David G.* (2019): *Value Sensitive Design. Shaping Technology with Moral Imagination*, Massachusetts.

- Gong, Chen (2023): AI voices reduce cognitive activity? A psychophysiological study of the media effect of AI and human newscasts Chinese journalism, in: *Frontiers in Psychology*, 23. November 2023 (online unter: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1243078/full>- letzter Zugriff: 9.8.2025).
- Grimm, Petra (2025a): Ethik der Interdisziplinarität in der KI- und MIT-Forschung: eine Frage der Haltung?, in: Petra Grimm / Oliver Zöllner (Hg.): *Ethik der Digitalisierung in Gesundheitswesen und Pflege. Analyse und ein Tool zur integrierten Forschung*, Stuttgart, S. 109–125.
- Grimm, Petra (2025b): Ethics by Design: Potenziale, Umsetzung und Grenzen, in: Dirk Lanzerath / Aurélie Halsband (Hg.), *Ethics by Design. Grundlagen, Umsetzung und Grenzen ethischer Technikgestaltung. Ethik in den Biowissenschaften*, in: *Sachstandsberichte des DZRE*, Bd. 28, Baden-Baden, S. 41–76.
- Grimm, Petra (2020): Die Ent-wirklichung. Zum Vertrauen in Zeiten der Infodemie, in: Klaus Koziol (Hg.), *Entwirklichung der Wirklichkeit. Von der Suche nach neuen Sicherheiten*, München, S. 55–83.
- Grimm, Petra / Kuhnert, Susanne / Schlegel, Marcel (2025): *Ethische Leitlinien für den Einsatz von synthetischen und geklonten Stimmen im Nachrichtenjournalismus*, zweite, überarbeitete Fassung, Stuttgart.
- Habermas, Jürgen (2022): *Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik*, Berlin.
- Hanitzsch, Thomas / Rick, Jana (2021): Prekarisierung im Journalismus, 1. Ergebnisbericht, März 2021 (online unter: <https://www.ifkw.uni-muenchen.de/lernbereiche/hanitzsch>projekte>prekarisierung.pdf> – letzter Zugriff: 29.3.2025).
- Jakobs, Ilka et al. (2022): Medienvertrauen im internationalen Vergleich – Befunde aus Deutschland, Spanien, Schweden und den USA, in: *UFITA – Archiv für Medienrecht und Medienwissenschaft* 86 (2/2022), S. 374–401.
- Jarren, Otfried (2019): Fundamentale Institutionalisierung: Social Media als neue globale Kommunikationsinfrastruktur. Der Beitrag der Kommunikationswissenschaft zur Analyse medialer Institutionalisierungsprozesse, in: *Publizistik* 64, S. 163–179.
- Krämer, Sybille (2021): Der Verlust des Vertrauens. Medienphilosophische Perspektiven auf Wahrheit und Zeugenschaft in digitalen Zeiten, in: Christian Schicha / Ingrid Stapf / Saskia Sell (Hg.), *Medien und Wahrheit. Medienethische Perspektiven auf Desinformation, Lügen und „Fake News“*, Baden-Baden, S. 25–42.
- Kleemann, Aldo (2023): Deepfakes – Wenn wir unseren Augen und Ohren nicht mehr trauen können. Medienmanipulationen im Konflikt. Herausforderungen und Bewältigungsstrategien, in: *SWPAktuell* 43, Berlin.
- Kuhnert, Susanne / Grimm, Petra (2020): Die Zusammenarbeit von Industrie, Ethik und Wissenschaft im Forschungsverbund. Kommunikation – Integration – Innovation, in: Bruno Gransche / Arne Manzeschke (Hg.), *Das geteilte Ganze. Horizonte Integrierter Forschung für künftige Mensch-Technik-Verhältnisse*, Wiesbaden, S. 241–261.
- Lenbet, Aylin (2024): Zur Aktualität des Kompetenzbegriffs und zur Bedeutung der Kompetenzentwicklung für das Coaching, in: *Organisationsberatung – Supervision – Coaching* 1, S. 221–232.

- MacCormac, Earl R.* (1993): Das Dilemma der Ingenieurethik, in: Hans Lenk / Günter Ropohl (Hg.), Technik und Ethik, Stuttgart, S. 222–244.
- Manders-Huits, Noëmi* (2011): What values in design? The challenge of incorporating moral values into design, in: Science and Engineering Ethics 17(2), S. 271–287.
- Menzel, Christoph / Winkler, Christian* (2018): Zur Diskussion der Effekte Künstlicher Intelligenz in der wirtschaftswissenschaftlichen Literatur (= Diskussionspapier Nr. 8. des Bundesministeriums für Wirtschaft und Energie), Oktober 2018 (online unter: https://www.bundeswirtschaftsministerium.de/Redaktion/DE/Downloads/Diskussionspapiere/20190205-diskussionspapier-effekte-kuenstlicher-intelligenz-in-der-wirtschaftswissenschaftlichen-literatur.pdf?__blob=publicationFile&v=6 – letzter Zugriff: 26.5.2025).
- Müller, Michael / Grimm, Petra* (2016): Narrative Medienforschung. Einführung in Methodik und Anwendung, Konstanz, München.
- Pancratz, Nils / Fandrich, Anatolij / Diethelm, Ira* (2023): Didaktische Strukturierung von Unterrichtsmaterialien zum Thema „Künstliche Intelligenz“, in: Kai Bliesmer / Michael Komorek (Hg.), Didaktische Rekonstruktion – fachdidaktischer Ansatz für aktuelle Bildungsaufgaben, Oldenburg, S. 84–96.
- Pawelec, Maria / Biess, Cora* (2021): Deepfakes. Technikfolgen und Regulierungsfragen aus ethischer und sozialwissenschaftlicher Perspektive, Baden-Baden.
- Prinzing, Marlis* (2024): Journalismus, in: Petra Grimm / Kai Erik Trost / Oliver Zöllner (Hg.), Digitale Ethik, Baden-Baden, S. 517–527.
- Quiring, Oliver et al.* (2024): Zurück zum Niveau vor der Pandemie – Konsolidierung von Vertrauen und Misstrauen. Mainzer Langzeitstudie Medienvertrauen 2023, in: Media Perspektiven 9, S. 1–14.
- Rath, Matthias* (2013): Medienethik – zur Normativität in der Kommunikationswissenschaft, in: Matthias Karmasin / Matthias Rath / Barbara Thomaß (Hg.), Normativität in der Kommunikationswissenschaft. Wiesbaden, S. 289–299.
- Reporter ohne Grenzen* (2023): Paris Charter on AI and Journalism, in: RSF (online unter: <https://rsf.org/en/paris-charter-ai-and-journalism> – letzter Zugriff: 19.1.2026).
- Ridsdale, Chantel / Rothwell, James / Smit, et al.* (2015): Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report, Halifax.
- Schicha, Christian* (2021): Bearbeitete Bilder – Techniken und Bewertungen visueller Veränderungen am Beispiel politischer Motive, in: Christian Schicha / Ingrid Stapf / Saskia Sell (Hg.), Medien und Wahrheit. Medienethische Perspektiven auf Desinformation, Lügen und „Fake News“, Baden-Baden, S. 173–203.
- Schlegel, Marcel et al.* (2024): Einführung in das interdisziplinäre Lehrkonzept von IKID: Ziele und Programmatik einer integrierten KI-Lehre (= Whitepaper-Serie zum Forschungsprojekt IKID: Interdisziplinäres KI-Exploratorium), (online unter: ai.hdm-stuttgart.de/research/ikid/whitepaper-serie – letzter Zugriff: 5.2.2026).
- Schüller, Katharina* (2019): Ein Framework für Data Literacy, in: ASTa – Wirtschafts- und Sozialstatistisches Archiv 13 (3/2019), S. 297–317.
- Singh, Rita* (2020): The Role of Human Voice in the Communication of Digital Disinformation, in: Maya Mirchandani (Hg.), Tackling Insurgent Ideologies in a Pandemic World. ORF / Global Policy Journal, Neu-Delhi, S. 69–74.

- Spiekermann, Sarah* (2019): *Digitale Ethik. Ein Wertesystem für das 21. Jahrhundert*, München.
- Stapf, Ingrid* (2021): „Fake News“ als eine (mögliche) Frage der Wahrheit? Medienethische Perspektiven auf Wahrheit im Kontext der Digitalisierung. In: Christian Schicha / Ingrid Stapf / Saskia Sell (Hg.), *Medien und Wahrheit. Medienethische Perspektiven auf Desinformation, Lügen und „Fake News“*, Baden-Baden, S. 97–119.
- Stapf, Ingrid* (2024): Desinformation, in: Petra Grimm / Kai Erik Trost / Oliver Zöllner (Hg.), *Digitale Ethik*, Baden-Baden, S. 315–326.
- Székely, Éva / Miniota, Jura / Hejná, Mísa* (2025): Will AI shape the way we speak? The emerging sociolinguistic influence of synthetic voices, in: *Proceedings of the 15th International Workshop on Spoken Dialogue Systems Technology*, Association for Computational Linguistics, Mai 2025 (online unter: <https://aclanthology.org/2025.iwds-1/> – letzter Zugriff: 19.1.2026).
- Tschopp, Marisa / Ruef, Marc / Monett, Dagmar* (2022): Vertrauen Sie KI? Einblicke in das Thema Künstliche Intelligenz und warum Vertrauen eine Schlüsselrolle im Umgang mit neuen Technologien spielt, in: Miriam Landes / Eberhard Steiner / Tatjana Utz (Hg.), *Kreativität und Innovation in Organisationen. Impulse aus Innovationsforschung, Management, Kunst und Psychologie*, Wiesbaden, S. 319–346.
- Weber, Max* (2017): *Politik als Beruf*, Stuttgart, Ditzingen.
- World Economic Forum* (2025): *Future of Jobs Report 2025: Insight Report*, (online unter: <https://www.weforum.org/publications/the-future-of-jobs-report-2025/> – letzter Zugriff: 19.1.2026).