

Leible, Stephan; Gücük, Gian-Luca; Simic, Dejan; von Brackel-Schmidt, Constantin;
Lewandowski, Tom

Article — Published Version

Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft

HMD Praxis der Wirtschaftsinformatik

Provided in Cooperation with:

Springer Nature

Suggested Citation: Leible, Stephan; Gücük, Gian-Luca; Simic, Dejan; von Brackel-Schmidt, Constantin; Lewandowski, Tom (2024) : Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft, HMD Praxis der Wirtschaftsinformatik, ISSN 2198-2775, Springer Fachmedien Wiesbaden GmbH, Wiesbaden, Vol. 61, Iss. 2, pp. 344-370,
<https://doi.org/10.1365/s40702-024-01050-x>

This Version is available at:

<https://hdl.handle.net/10419/315880>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



<http://creativecommons.org/licenses/by/4.0/deed.de>



Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft

Stephan Leible · Gian-Luca Gücük · Dejan Simic · Constantin von Brackel-Schmidt · Tom Lewandowski

Eingegangen: 5. November 2023 / Angenommen: 27. Januar 2024 / Online publiziert: 1. März 2024
© The Author(s) 2024

Zusammenfassung Die dynamische Entwicklung und steigende Beliebtheit generativer künstlicher Intelligenz (genKI), besonders durch die Verbreitung und dem Einsatz von ChatGPT, hat das enorme Potenzial dieser Technologie gezeigt, Berufsfelder und Branchen grundlegend transformieren zu können. Die Entscheidung hinsichtlich des Einsatzes von genKI sowie die Identifikation aussichtsreicher Anwendungsszenarien stellen in Anbetracht eines rasch wachsenden und immer komplexeren Marktes erhebliche Herausforderungen dar. Angesichts dieser Gegebenheiten wird mit dem vorliegenden Artikel das Ziel verfolgt, eine Übersicht über die Fähigkeiten und Limitationen von genKI zu präsentieren. Mittels einer systematischen Literaturrecherche wurden vielfältige Anwendungsszenarien eruiert und im Hinblick auf die Ergebnisse des genKI-Einsatzes bewertet, was eine Momentaufnahme der aktuellen Fähigkeiten und Limitationen ermöglichte. Zusätzlich wurde eine Umfrage unter 40 Teilnehmenden durchgeführt, um die Nutzungsgewohnheiten und Erfahrungen im Umgang mit genKI zu erfassen und die Befunde aus der

✉ Stephan Leible · Constantin von Brackel-Schmidt · Tom Lewandowski
Fachbereich Informatik, IT-Management und -Consulting, Universität Hamburg, Hamburg, Deutschland
E-Mail: stephan.leible@uni-hamburg.de

Constantin von Brackel-Schmidt
E-Mail: constantin.schmidt@uni-hamburg.de

Tom Lewandowski
E-Mail: tom.lewandowski@uni-hamburg.de

Gian-Luca Gücük · Dejan Simic
Fachbereich Informatik, Universität Hamburg, Hamburg, Deutschland

Gian-Luca Gücük
E-Mail: gian-luca.guecuk@studium.uni-hamburg.de

Dejan Simic
E-Mail: dejan.simic@studium.uni-hamburg.de

Literatur zu validieren. Die erlangten Einsichten sollen Praktikerinnen und Praktiker bei der Navigation im Bereich genKI unterstützen und eine Entscheidungshilfe bieten, indem die identifizierten Fähigkeiten und Limitationen im Kontext eigener Anwendungsszenarien eingeordnet werden können. Weiterhin liefern die Ergebnisse Anhaltspunkte für die methodische Untersuchung von genKI-Anwendungsszenarien sowie Ausgangspunkte für die wissenschaftliche Vertiefung durch Forscherinnen und Forscher. Mit der Verknüpfung von theoretischer Analyse und praktischer Erhebung bietet der Artikel einen umfassenden Einblick in den aktuellen Stand von genKI.

Schlüsselwörter Generative Künstliche Intelligenz · Große Sprachmodelle · Fähigkeiten · Limitationen · ChatGPT

Between Research and Practice: Capabilities and Limitations of Generative AI and its growing Role in the Technology Landscape

Abstract The dynamic development and increasing popularity of generative artificial intelligence (genAI), especially through the spread and use of ChatGPT, has shown the enormous potential of this technology to fundamentally transform professions and industries. The decision regarding using genAI and identifying promising application scenarios pose considerable challenges in view of a rapidly growing and increasingly complex market. Given these circumstances, this article aims to present an overview of the capabilities and limitations of genAI. A systematic literature review was used to identify a variety of application scenarios and evaluate them in terms of the outcomes of genAI deployment, providing a snapshot of current capabilities and limitations. In addition, a survey was conducted among 40 participants to record usage habits and experiences in dealing with genAI and to validate the findings from the literature. The insights gained should support practitioners in navigating the field of genAI and provide decision-making support by enabling them to classify the identified capabilities and limitations in the context of their own application scenarios. Furthermore, the results provide points of reference for the methodical investigation of genAI application scenarios as well as starting points for in-depth scientific research by researchers. By linking theoretical analysis and a practical survey, the article offers a comprehensive insight into the current state of genAI.

Keywords Generative Artificial Intelligence · Large Language Models · Capabilities · Limitations · ChatGPT

1 Einleitung

Die rasant fortschreitende Entwicklung von künstlicher Intelligenz (KI) hat durch das Aufkommen und die Verbreitung von generativer KI (genKI) eine neue Ära der Technologie eingeläutet, die sowohl die Forschung als auch die Praxis substanzi-

ell beeinflusst (Teubner et al. 2023). Traditionelle KI-Anwendungen fokussierten sich vornehmlich auf Datenanalyse (zum Beispiel für die Mustererkennung (Santos und Qin 2019)), Automatisierung (beispielsweise von Routineaufgaben (Jha et al. 2019)), Empfehlungssysteme (zum Beispiel zur Entscheidungsunterstützung (Nica et al. 2022)) und Bild- und Spracherkennung (beispielsweise für die Prüfung von Anwesenheit (Siddiqui et al. 2020)). Durch genKI erweitert sich das Fähigkeitsenspektrum von KI insbesondere im Bereich der Generierung von Inhalten, basierend auf einem begrenzten Bestand an verarbeiteten Trainingsdaten und einer umfassenden Anzahl an Parametern (Ray 2023). Die Qualität, vornehmlich bei der Textgeneration sowie die universelle Einsetzbarkeit in verschiedenen Szenarien haben ein großes Interesse für genKI geweckt. In einer zunehmend datengetriebenen Welt erkunden sowohl ökonomisch agierende als auch öffentliche Organisationen wie Universitäten und Hochschulen die transformative Kraft von genKI, um den Mehrwert dieser Technologie in verschiedenen Anwendungsszenarien zu erforschen und ihr Potenzial auszuschöpfen (Budhwar et al. 2023; Korzynski et al. 2023). Seit der Einführung des großen Sprachmodells GPT-3.5 in Verbindung mit dem zur Interaktion genutzten Interface ChatGPT im November 2022 (OpenAI 2022) hat die Technologie bereits erhebliche Fortschritte erzielt (beispielsweise Multimodalität (OpenAI 2023a)).

GenKI stellt keine grundlegend neue Technologie dar, jedoch hat die Weiterentwicklung und Verfeinerung dieser Technologie eine höhere Stufe der Funktionalität und Anwendbarkeit erreicht. Modelle wie Generative Adversarial Networks (GANs) haben schon vor ChatGPT die Fähigkeit zur Erzeugung neuer Daten demonstriert, die von realen Daten nicht zu unterscheiden sind (Gui et al. 2023). Beispiele für GAN-basierte Anwendungen umfassen die Hochskalierung von niedrig aufgelösten Bildern, bei der Bildteile (Pixel) generativ auf Basis des Originalmaterials und der Trainingsdaten ergänzt werden (Lee et al. 2019). Ein weiteres bemerkenswertes Beispiel stellen die ethisch umstrittenen DeepFakes dar, welche äußerst realistische Bilder oder Videos generieren, um überzeugende Fälschungen zu erstellen (Yadav und Salmani 2019). Mit ChatGPT wurde jedoch ein neues Niveau der Textgenerierung erreicht, das sowohl hinsichtlich der Qualität als auch der Breite unübertroffen ist. Die Fähigkeit, kontextbezogene und kohärente Texte zu erzeugen, hat die Tür für eine breite Palette von Anwendungen in verschiedenen Sektoren geöffnet (George et al. 2023).

Die rasche Verbreitung von ChatGPT signalisiert ein immenses Interesse und einen substanziellen Bedarf an genKI. Mit über 100 Mio. Nutzerinnen und Nutzer innerhalb von lediglich zwei Monaten avancierte ChatGPT zu einer der am schnellsten adoptierten Konsumenten-Software weltweit (Milmo 2023). Dies unterstreicht den bedeutsamen Einfluss, den genKI auf die aktuelle Technologielandschaft ausübt. Die daraus resultierende Begeisterung hat sowohl die Praxis als auch die Forschung gleichermaßen beeinflusst, was sich durch die zahlreichen neuen Anwendungsszenarien und Forschungsartikel manifestiert. In diesen werden kontinuierlich die Fähigkeiten und Limitationen der Technologie exploriert und erforscht, um innovative Lösungen für bestehende Herausforderungen und Probleme zu entwickeln (Teubner et al. 2023).

Die Dynamik von genKI als sich schnell entwickelndes technologisches Feld begünstigt eine fortlaufende Erweiterung der Möglichkeiten, die im Einklang mit den

vorhandenen Fähigkeiten und Limitationen stehen. Durch regelmäßige Neuerungen, die Einführung neuer Modelle wie beispielsweise GPT-4 sowie Software-Integrationen wie dem auf GPT-4 basierenden Copilot in Microsoft Office 365 erlebt das Funktionalitätsspektrum dieser Technologie eine stetige Expansion (OpenAI 2023b; Spataro 2023). In diesem Kontext bleibt die Fragestellung bezüglich der spezifischen Fähigkeiten und Limitationen von genKI und wie diese praktische Anwendungsszenarien beeinflussen, ein zentraler Punkt für die Erforschung ihres Nutzens sowie für die Entscheidungsfindung hinsichtlich ihrer Verwendung in der realen Welt. Hier knüpft der vorliegende Artikel an und adressiert diesen Untersuchungsbereich mit dem Ziel, ein fundiertes Verständnis der aktuellen technologischen Fähigkeiten und Limitationen von text- wie auch bildausgebender genKI auf Basis einer breiten Palette verschiedener Anwendungsszenarien zu erlangen. Die in unserer Arbeit behandelte Forschungsfrage lautet:

Welche Fähigkeiten und Limitationen weisen aktuelle text- und bildausgebende generative KI-Modelle auf, und wie beeinflussen diese Fähigkeiten und Limitationen ihre praktische Anwendbarkeit in verschiedenen Anwendungsszenarien?

Um die Forschungsfrage zu adressieren, wurde der Multimethodenansatz nach Mingers und Brocklesby (1997) verfolgt. Dabei wurde eine systematische Literaturrecherche nach Xiao und Watson (2019) durchgeführt und durch eine Vorabumfrage nach Groves et al. (2009) mit 40 verwerteten Rückläufern, unterteilt in 22 berufstätige Personen und 18 Studierende, ergänzt. Sowohl die Literaturrecherche als auch die Umfrage zielten darauf ab, Wissen und Erfahrungen zu sammeln, die Aufschluss darüber geben, in welchen Bereichen genKI bereits Anwendung gefunden hat und welche Fähigkeiten sich als wertvoll und nutzenstiftend erwiesen haben. Auch wird untersucht, in welchen Bereichen die Ergebnisse hinter den Erwartungen zurückblieben oder die Technologie an ihre derzeitigen Limitationen stieß. Der Fokus lag auf textausgebenden genKI-Modellen und -Interfaces wie ChatGPT und Gemini (ehemals Bard), als auch auf bildausgebenden genKI-Modellen und -Interfaces wie DALL-E oder Stable Diffusion, die in der Literatur bislang weniger häufig behandelt wurden. Mit den Erkenntnissen dieses Artikels wird ein repräsentativer Orientierungspunkt in Bezug auf die aktuellen Fähigkeiten und Limitationen von genKI angestrebt. Diese sollen Praktikerinnen und Praktikern helfen, Entscheidungen über den Einsatz von genKI in verschiedenen Anwendungsszenarien zu treffen und Forscherinnen und Forschern einen Überblick über bisherige Ergebnisse sowie Ansatzpunkte für weiterführende Untersuchungen bieten.

2 Theoretischer Hintergrund

Im Folgenden wird zuerst genKI als spezielle Form einer KI sowie die zentrale Rolle großer Sprachmodelle beleuchtet und erläutert. Nachfolgend werden diese Modelle, welche den Kern von genKI-Anwendungen bilden, mit ihren charakteristischen Merkmalen beschrieben. Zudem wird eine Übersicht über einige bekannte Modelle präsentiert.

2.1 Grundlagen generativer künstlicher Intelligenz

GenKI markiert einen signifikanten Fortschritt in der modernen KI-Landschaft mit dem Ziel, Algorithmen zu entwickeln, die in der Lage sind, neue Daten zu generieren, die sich ähnlich zu den (realen) Daten verhalten, auf denen sie trainiert wurden (Ray 2023). Hierbei kommen sogenannte große Sprachmodelle zum Einsatz, wie beispielsweise GPT von OpenAI oder Gemini von Google. Diese Modelle zielen grundlegend darauf ab, menschliche Sprache zu verstehen und zu generieren (OpenAI 2022; Anil et al. 2023). GenKI wird meist dazu verwendet, komplexe Textdaten zu erzeugen, die sowohl inhaltlich kohärent als auch strukturell korrekt sind. Die Basis dieser Fähigkeit liegt in der Architektur und den Trainingsverfahren der verwendeten Modelle (Vaswani et al. 2017). Durch die Identifikation von Mustern, Clustern oder Anomalien in den Daten können sie Einsichten gewinnen und eine Grundlage für die Wahrscheinlichkeitsverteilung zur Generierung neuer Daten schaffen. Auf Basis dieser Verteilungen wird geschätzt, wie wahrscheinlich ein Wort oder eine Phrase auf ein anderes folgt. Hierbei wird ein Zufallsfaktor angewendet, der gewährleistet, dass auf dieselbe Eingabe (Prompt) nicht immer der exakt gleiche Text oder das gleiche Ergebnis generiert wird, sondern Variationen davon (Reiss 2023).

GenKI lässt sich in vielfältigen Anwendungsszenarien nutzen, beispielsweise in der Datenrekonstruktion oder Datenimputation. Doch insbesondere die Erzeugung von realistischen neuen Daten stellt ein herausragendes Merkmal von genKI dar, wodurch sie als potenzielles Werkzeug in verschiedenen Bereichen wie Computergrafik, Design und der Synthese von natürlicher Sprache in Frage kommt (Yun et al. 2022; Hutson und Plate 2023a). Diese Fähigkeiten rufen auch ethische Fragestellungen und Herausforderungen hervor (Ray 2023). Beispiele hierfür umfassen die Verbreitung von Desinformation und gefälschten Inhalten, die Erstellung von Leistungsnachweisen im Bildungssektor und datenschutzbedingte Aspekte. Letzteres insbesondere, da die Prompts in genKI-Modellen über (internationale) Server geleitet werden und personenbezogene Daten enthalten können. In diesem Zusammenhang ist es essenziell, einen verantwortungsvollen Umgang mit genKI von Anfang an zu erlernen und zu berücksichtigen. Maßnahmen hierzu sollten unter anderem die Vermittlung entsprechender Datenschutzkompetenzen an die Nutzerinnen und Nutzer umfassen, beispielsweise die Anonymisierung von (personenbezogenen) Daten vor deren Eingabe in genKI-Anwendungen. Zudem ist es aufgrund der Funktionsweise von genKI-Modellen essenziell, dass generierte Aussagen auf Angemessenheit und Korrektheit überprüft werden, um Fehlinformationen oder Verzerrungen zu vermeiden.

2.2 Große Sprachmodelle

Die Wurzeln heutiger großer Sprachmodelle können bis zum Jahr 1966 zurückverfolgt werden, als mit ELIZA eines der ersten Beispiele für ein solches Modell vorgestellt wurde (Weizenbaum 1966). Im Kontext genKI basieren große Sprachmodelle üblicherweise auf der Transformer-Architektur, die sich insbesondere in der natürlichen Sprachverarbeitung durch ihre Fähigkeit auszeichnet, lange Abhängigkeiten in den Daten zu erkennen und zu verarbeiten (Vaswani et al. 2017). Diese

Architektur setzt *selbst-attentive* Mechanismen ein, um die Beziehungen zwischen verschiedenen Teilen eines Textes zu erfassen und auf dieser Grundlage neue Textsegmente zu generieren. Das Training solcher Modelle wird oft in zwei Phasen durchgeführt: *Pretraining* und *Fine-Tuning*. Während der *Pretraining*-Phase wird das Modell mit einem umfangreichen Textkorpus trainiert, um eine breite Palette von Sprachmustern zu erlernen. Im *Fine-Tuning* wird das Modell anschließend mit spezifischen Daten feinjustiert, um bestimmte Aufgaben besser bewältigen zu können. Die Größe des Trainingsdatensatzes, die Qualität der Daten sowie die Anzahl der Parameter sind einige der entscheidenden Faktoren für die Leistungsfähigkeit des resultierenden Modells (Alaswad et al. 2023; García-Peñalvo 2023).

Modelle wie GPT-3 oder dessen Nachfolgeversionen verfügen über Milliarden von Parametern, die im Laufe des Trainings angepasst werden. Diese Parameter erlauben den Modellen den Zugriff auf eine umfangreiche Informationsmenge und die Erkennung komplexer Muster in den Daten (Boscardin et al. 2023), steigern jedoch gleichzeitig die Anforderungen an die Rechenkapazität. Eine hohe Parameteranzahl kann außerdem die Anfälligkeit für *Overfitting* erhöhen. Dies bedeutet, dass ein Modell möglicherweise zu stark an die vorhandenen Trainingsdaten angepasst wird und auf neue, unbekannte Daten entweder nicht oder nur unzureichend reagiert.

In vielen Kontexten können Sprachmodelle auch grundlegendes Faktenwissen auf verschiedenen Niveaus bereitstellen. Hierdurch kann eine interaktive Konversation entstehen, die an typische Chats erinnert, wobei das Sprachmodell kontextbezogene Antworten liefert, die häufig kaum von menschlichen Antworten zu unterscheiden sind (Tlili et al. 2023). Diese Interaktivität ermöglicht den Einsatz großer Sprachmodelle in einer breiten Palette von Anwendungen, von Chatbots bis hin zu assistiven Technologien (Iqbal et al. 2023; Memmert und Tavanapour 2023). So hat sich auch der Begriff *Foundation Modell* etabliert, der Modelle wie GPT oder BERT bezeichnet, die eine große Auswahl an Aufgaben übernehmen können (Waqas et al. 2023). Auf diesem Fundament können wiederum spezifischere Modelle oder Anwendungen entwickelt werden. So lassen sich über ChatGPT seit November 2023 individualisierte GPTs für konkrete Aufgaben erstellen, zum Beispiel als Schreibassistent oder Übersetzer. Zusammenfassend lässt sich feststellen, dass die Kluft

Tab. 1 Exemplarische Übersicht großer Sprachmodelle beschränkt auf aktuelle Versionen

Sprachmodell	Entwickler	Erscheinungsjahr	Besonderheiten
Transformer	Google	2017	Einführung der Transformer-Architektur durch Google
BERT	Google	2018	Bidirektionale Kontextverarbeitung; Vorhergehender und nachfolgender Text wird für die Vorhersage berücksichtigt
T5	Google	2020	Text-zu-Text-Ansatz; Formulierung aller Aufgaben der Sprachverarbeitung als Übersetzungsprobleme
GPT-4	OpenAI	2023	Handhabung von Mehrdeutigkeit; Multimodalität; Mustererkennung; Datenanalysefähigkeiten
PaLM 2	Google	2023	Besondere Fähigkeit, Sprache zu verstehen und mit logischem Zusammenhang zu generieren; Multimodalität
Gemini	Google	2023	Multimodalität, nahtlose Verarbeitung verschiedener Datentypen; Integriert in das Google-Ökosystem

zwischen maschinell generiertem und menschlichem Text durch die Kombination aus leistungsstarker Architektur, umfangreichen Trainingsdaten und hohen Parameterzahlen bei Sprachmodellen wie GPT oder Gemini stetig verringert wird. Sie repräsentieren den aktuellen Stand der Technik im Bereich von genKI und weisen ein großes Potenzial für technologische Fortschritte auf, um noch leistungsfähiger und interaktiver zu werden. In Tab. 1 sind exemplarisch einige große Sprachmodelle in ihrer aktuellen Version aufgeführt. Es ist anzumerken, dass insbesondere im Jahr 2023 eine große Vielfalt neuer Sprachmodelle hinzugekommen sind und es sich um keine vollständige Auflistung handelt.

3 Methodische Vorgehensweise

Um die Forschungsfrage zu adressieren, wurden zwei Forschungsmethoden angewendet: (1) Eine systematische Literaturrecherche, die dem Ansatz von Xiao und Watson (2019) folgt, um einen aktuellen Überblick über die Fähigkeiten und Limitationen von genKI zu erhalten und (2) eine Vorabumfrage zu Nutzungsgewohnheiten sowie Erfahrungen mit genKI.

3.1 Systematische Literaturrecherche

Im Rahmen der systematischen Literaturrecherche lag der Fokus sowohl auf wissenschaftlich begutachteter Literatur als auch auf grauer Literatur. Letztere ist darin begründet, dass durch die Aktualität von genKI von zahlreichen Autorinnen und Autoren ein umfassender Korpus als Vorabversionen (Preprints) veröffentlicht wurde, um die Zeitspanne von Begutachtungsverfahren bis zur Veröffentlichung zu überbrücken. Der von Xiao und Watson (2019) vorgeschlagene Ansatz umfasst die folgenden acht Schritte, denen gefolgt wurde: (1) Formulierung des Problems, (2) Entwicklung und Validierung des Überprüfungsprotokolls, (3) Suche in der Literatur, (4) Screening für die Aufnahme, (5) Bewertung der Qualität, (6) Extraktion der Daten, (7) Analyse und Synthese der Daten und (8) Bericht der Ergebnisse.

Der erste Schritt wurde in den ersten zwei Kapiteln dieses Artikels durch die Darlegung der Problematik rund um die praktische Einschätzung der Fähigkeiten und Limitationen von genKI für verschiedene Anwendungsszenarien sowie der Forschungsfrage umgesetzt. Im zweiten Schritt wurde ein Überprüfungsprotokoll erstellt, welches die Zielsetzung der Literaturrecherche auf Basis der identifizierten Problematik und Forschungsfrage sowie die geplante Suchstrategie umfasst. In dieser Phase wurde beschlossen, aufgrund der größeren Wissensbasis und sprachlichen Konsistenz bei der Analyse der Quellen den Fokus auf englischsprachige Literatur zu legen. Die Zusammenarbeit am Überprüfungsprotokoll erfolgte kollaborativ über ein gemeinsames Dokument, um Informationsasymmetrien zu vermeiden und getroffene Entscheidungen im Prozess zu dokumentieren.

Den dritten Schritt begannen wir mit einer unstrukturierten Literaturrecherche über verschiedene elektronische Datenbanken, um diverse relevante Schlüsselbegriffe und Unterthemen basierend auf der Forschungsfrage zu sammeln. Dieser Schritt wurde von drei Autoren unabhängig voneinander durchgeführt, woraufhin in einer

nachfolgenden Zusammenführung und Abstimmung die Datenbanken AISEL und arXiv ausgewählt wurden, ergänzt durch die ersten 100 nach Relevanz sortierten Suchergebnisse von Google Scholar. AISEL deckt den Bereich der Wirtschaftsinformatik durch einschlägige wissenschaftliche Konferenzen und Journals adäquat ab, während arXiv eine Fülle relevanter (grauer) Literatur wie Preprints zu genKI bereithält, wodurch die Aktualität dieser Auswertung gesteigert wird. Der Einbezug von Google Scholar begründet sich durch die hohe Anzahl an indexierten Datenbanken, sodass komplementär relevante Literatur außerhalb AISEL und arXiv identifiziert und integriert werden kann. Auf Grundlage der Forschungsfrage und der Schlüsselbegriffe, die sich aus der vorläufigen Kartierung ergaben, entwickelten wir den folgenden Suchstring:

("large language models" OR "generative language model" OR "generative artificial intelligence" OR ChatGPT OR "Google Bard" OR "Dall E" OR "Midjourney") AND (capability OR ability OR limitation OR possibility OR feature OR boundary OR constraint OR image) AND ("use case" OR "case study")

Der angewendete Suchstring umfasst übliche englische Terminologien im Kontext von genKI. Als Ergänzung integrierten wir die bekanntesten text- und bildausgebenden genKI-Modelle, die in den meisten Artikeln der vorab durchgeführten unstrukturierten Literaturrecherche referenziert wurden. Die Anwendung des Suchstrings resultierte in insgesamt 601 Suchergebnissen über die drei ausgewählten Datenbanken. Abb. 1 illustriert den Auswertungsprozess der Literatur. Im Auswertungsverfahren wurde ein konservativer Ansatz verfolgt, sodass bei Unsicherheiten die Ergebnisse nicht direkt ausgeschlossen wurden. Zunächst prüften drei Autoren die Titel der Ergebnisse auf Relevanz, wodurch 350 Ergebnisse als irrelevant identifiziert und aussortiert wurden. Somit verblieben 251 Ergebnisse für die weitere Analyse.

In Schritt vier wurden spezifische Ein- und Ausschlusskriterien festgelegt und im Überprüfungsprotokoll verzeichnet. Diese Kriterien vereinfachen den späteren

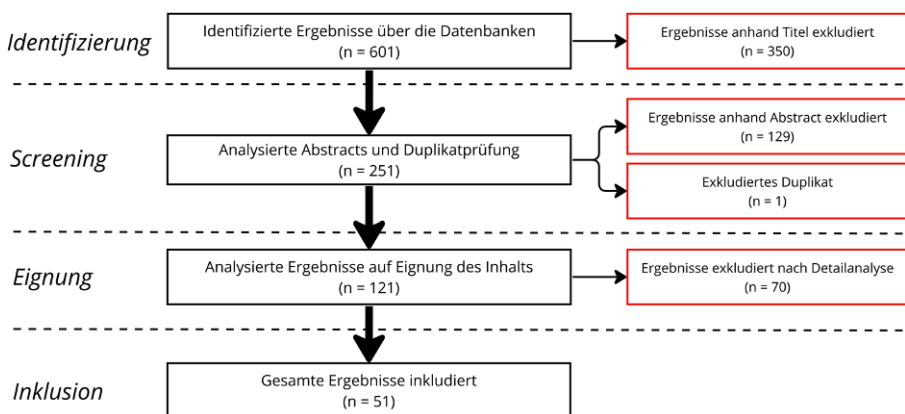


Abb. 1 Auswertungsprozess der systematischen Literaturrecherche

Analyseprozess, indem sie helfen, festzulegen, welche Ergebnisse in den finalen Datensatz aufgenommen werden. Ein Ausschlusskriterium beispielsweise war, dass die Arbeiten in englischer Sprache verfasst und mit den vorhandenen digitalen Zugängen verfügbar sein müssen. Zudem sollen sie relevante Inhalte bereitstellen, die mindestens ein Anwendungsszenario von genKI behandeln und untersuchen, sodass daraus Fähigkeiten und Limitationen extrahiert und synthetisiert werden können, um die Forschungsfrage direkt zu adressieren. In seltenen Fällen wurden konzeptionelle Erkenntnisse oder Präsentationen berücksichtigt, die besondere Anwendungsszenarien oder Ergebnisse beinhalteten. Unter Anwendung der festgelegten Ein- und Ausschlusskriterien überprüften drei Autoren die Abstracts der verbleibenden 251 Ergebnisse. Daraufhin wurden 129 Ergebnisse aufgrund ihrer Inhalte ausgeschlossen. Ein weiteres Ergebnis wurde entfernt, da es sich um ein Duplikat handelte.

Während Schritt fünf wurden die Volltextversionen der verbleibenden 121 Ergebnisse bezogen. Dies ermöglichte es zwei Autoren, Detailanalysen durchzuführen, um die Relevanz der Ergebnisse für die Forschungsfrage final zu bewerten. Bei Uneinigkeiten wurden interne Diskussionen im Autorenteam abgehalten, bis ein Konsens über die Inklusion oder Exklusion eines Ergebnisses erreicht war. Infolgedessen wurden 70 Ergebnisse aufgrund mangelnder Relevanz ausgeschlossen, wodurch der finale Datensatz auf 51 Artikel reduziert wurde.

Im sechsten Schritt kodierten wir mittels Microsoft Excel die qualitativen Aussagen aus dem finalen Datensatz, welche konkrete Fähigkeiten oder Limitationen von genKI in verschiedenen Szenarien adressierten. Dabei wurde induktiv vorgegangen und eine Grundlage für die anschließende Synthese und Verdichtung der Daten gelegt, um übergeordnete Kategorien bezüglich der Fähigkeiten und Limitationen von genKI zu bilden. Im siebten Schritt erfolgte die Analyse der kodierten Daten. Die insgesamt 253 kodierten Aussagen wurden aggregiert und in gemeinsamen Kategorien geordnet. Dieser entscheidende Schritt wurde im Autorenteam gemeinsam umgesetzt, um Einigkeit über die Zuordnungen und Bezeichnungen zu erzielen, welche die strukturelle Grundlage für die Konzeptmatrix nach Webster und Watson (2002) bildete. Im abschließenden achten Schritt gilt es, die Ergebnisse zu präsentieren, was durch diesen Artikel realisiert wird. Hierbei wurde eine Konzeptmatrix gewählt, welche die kuratierte Sammlung der Literatur des finalen Datensatzes sowie die ausgearbeiteten Kategorien visualisiert.

3.2 Vorabumfrage

Zusätzlich zur Literaturrecherche wurde eine ergänzende Umfrage nach dem Vorgehen von Groves et al. (2009) durchgeführt, welche die Ergebnisse einer Vorabumfrage einer größeren geplanten Befragungsinitiative darstellt. Ziel der Umfrage war es, die Nutzungsgewohnheiten und Erfahrungen der Teilnehmenden aus der Praxisperspektive mit genKI zu erheben und auszuwerten. Außerdem sollte ein Verständnis dafür gewonnen werden, inwiefern die in der Literatur identifizierten Fähigkeiten und Limitationen von genKI bestätigt oder erweitert werden können. Insgesamt nahmen 40 Personen an der Umfrage teil, darunter 18 Studierende und 22 Berufstätige, siehe Tab. 2. Durch die Diversifikation der Teilnehmenden auf Studierende und Berufstätige strebten wir an, ein breiteres Spektrum an Perspektiven und Erfahrungen

Tab. 2 Demografische Daten der Umfrageteilnehmerinnen und -teilnehmer

Kategorie	Wert	Anzahl [Σ]	In Prozent [%]
Altersklasse	18–24 Jahre	5	12,5
	25–34 Jahre	21	52,5
	35–44 Jahre	9	22,5
	45–54 Jahre	3	7,5
	55–64 Jahre	2	5
Geschlecht	Männlich	24	60
	Weiblich	15	37,5
	Anderes	1	2,5
Position	Ausbildung oder Trainee	1	2,5
	Studentin oder Student	18	45
	Angestellt (ohne Personalverantwortung)	8	20
	Unteres Management	1	2,5
	Mittleres Management	2	5
	Höheres Management	10	25

abzudecken, da die Anwendungsszenarien im akademischen Umfeld voraussichtlich anders aussehen als in verschiedenen Berufsfeldern. Die Umfrage wurde vom 10. Oktober bis zum 26. Oktober 2023 Online über LimeSurvey durchgeführt. Die Umfrageteilnehmenden wurden über direkte Ansprache im Rahmen verschiedener (Lehr-)Veranstaltungen akquiriert. Die Umfrage umfasste Fragen zu Anwendungsszenarien von genKI und insbesondere zu Situationen, in denen die Nutzerinnen und Nutzer von der Technologie positiv überrascht oder enttäuscht wurden. Alle Rückläufer waren auswertbar und vollständig.

Die Ergebnisse der Vorabumfrage wurden separat ausgewertet und in Bezug zu den Ergebnissen aus der Literaturrecherche gesetzt. Diese Verknüpfung von primären und sekundären Daten ermöglichte eine Vertiefung des Verständnisses für die Fähigkeiten und Limitationen von genKI in verschiedenen Anwendungsszenarien und bot die Möglichkeit, die in der Literatur gefundenen Erkenntnisse exemplarisch zu validieren.

4 Fähigkeiten und Limitationen generativer künstlicher Intelligenz

In Abb. 2 sind die 51 Artikel des finalen Datensatzes in Form einer Konzeptmatrix dargestellt und hinsichtlich der Typen der eingesetzten genKI-Sprachmodelle klassifiziert. Dabei wurde in 47 Artikeln die Verwendung einer textausgebenden genKI identifiziert, in zwei einer bildausgebenden genKI und in weiteren zwei eine Kombination aus beiden. Im Rahmen dieser Analyse wurden diverse Fähigkeiten und Limitationen erkannt und in acht Übergruppen bei den *Fähigkeiten* und sieben bei den *Limitationen* kategorisiert. Es gilt dabei zu beachten, dass sich während der Zeit des Erscheinens von ChatGPT sehr viel bei den Funktionalitäten und dem Reifegrad von genKI in verschiedenen Anwendungsszenarien entwickelt hat. Daher wurden die spezifischen genKI-Sprachmodelle, die in den analysierten Artikeln zum Einsatz

Autorinnen und Autoren	Typ		Fähigkeiten								Limitationen						
	Bildausgebende genKI	Textausgebende genKI	Inhalts-generierung & -arbeit	Muster & Klassifikation	Code & Automatisierung	Sprachverarbeitung	Kreativität	Adaptabilität	Lernunterstützung	Domänenwissen	Ethik & Voreingenommenheit	Datenbedingt & Halluzination	Unterstützungsbedarf	Codenalyse & Berechnungen	Technische Kapazitäten	Textverarbeitung & -ausgabe	Universalität & Originalität
AlAfnan et al. 2023	X	X				X					X	X	X	X			X
Alaswad et al. 2023	X	X				X			X	X	X	X		X	X		X
Ali und OpenAI 2023	X					X		X			X	X					
Alkhaqani 2023	X				X	X					X	X					
Asfour und Murillo 2023	X					X		X			X	X					
Busch und Hausvik 2023	X					X					X	X					X
Campbell 2023	X					X			X								
Cummings et al. 2021	X	X						X									
DeKay 2023	X											X					
dos Santos 2023	X					X		X	X					X		X	X
Fan und Jiang 2023	X					X				X			X			X	X
Feng et al. 2023	X			X							X						X
Frenkel und Emara 2023	X					X								X			
Grodzinsky et al. 2023	X										X						
Guo et al. 2023	X	X	X							X	X		X		X		
Guttupalli et al. 2023	X	X					X										
Hobert et al. 2023	X	X				X									X		X
Hu et al. 2023	X	X				X						X					X
Huang und Sciuchetti 2023	X	X				X						X					
Huang et al. 2023	X	X				X					X				X	X	X
Hutson und Cotroneo 2023	X	X	X				X										
Hutson und Lang 2023	X	X								X					X		X
Hutson und Plate 2023a	X	X				X					X						
Hutson und Plate 2023b	X	X				X						X	X				
Hutson und Plate 2023c	X	X				X	X					X	X			X	
Hutson und Schnellmann 2023	X					X		X					X				
Iqbal et al. 2023	X				X	X		X						X			
Juhasz und Guan 2023	X	X								X				X			X
Kowalczyk et al. 2023	X				X			X			X						
Kumar et al. 2023	X					X										X	
Lappalainen und Narayanan 2023	X					X				X	X	X					
Leadbetter 2023	X	X				X									X	X	X
Memmert und Tavanapour 2023	X	X					X										
Mukherjee und Kamath 2023	X					X											
Panagopoulou et al. 2023	X	X						X			X	X					
Parker et al. 2023	X			X		X						X				X	
Perkins 2023	X	X				X											
Ploennigs und Berger 2023	X	X									X						
Regalia 2023	X	X				X					X						
Robinson 2023	X	X				X		X							X	X	X
Sirithumgul 2023	X	X		X		X							X				X
Soman und Ranjani 2023	X	X											X				
Tang 2023	X															X	
Tlili et al. 2023	X	X				X			X		X	X					
Torres 2023	X									X	X	X					X
Tsai et al. 2023	X						X			X	X	X					
Vasconcelos und dos Santos 2023	X								X	X			X				
Wang et al. 2023	X				X									X			
Weingart et al. 2023	X	X	X			X				X							
Yun et al. 2022	X	X						X				X					
Zeng 2023	X					X		X									
Gesamt [Σ]	4	49	25	4	4	31	5	11	5	10	15	14	11	6	8	9	15
Gesamt [%]	8	96	49	8	8	61	10	22	10	20	29	27	22	12	16	18	29

Abb. 2 Konzeptmatrix der Fähigkeiten und Limitationen von genKI synthetisiert aus der analysierten Literatur

Tab. 3 Anzahl des Vorkommens von genKI-Sprachmodellen oder -Interfaces aus der analysierten Literatur

Genutztes Sprachmodell oder Interface	Häufigkeit der Nutzungen
ChatGPT	35
GPT-4	14
GPT-3.5	6
GPT-3	6
GPT-2	3
Bard	3
Midjourney	3
DALL-E 2	2
Llama	2
Andere	22

kamen, am Ende dieses Kapitels in Tab. 3 identifiziert, gruppiert und quantifiziert sowie einzeln im Onlinematerial aufgelistet, damit die Aussagen über Fähigkeiten und Limitationen den jeweils genutzten Sprachmodellen zugeordnet werden können. Im Folgenden werden die definierten Kategorien von Fähigkeiten und Limitationen näher erläutert¹.

4.1 Fähigkeiten generativer künstlicher Intelligenz

Die erste Kategorie der Fähigkeiten von genKI bezieht sich auf die *Generierung von und Arbeit mit Inhalten*. Insbesondere umfasst diese Kategorie die Erzeugung diverser Inhaltsarten (Hobert et al. 2023; Juhasz und Guan 2023). Beispielsweise sind dies akademische Texte in der Wissenschaft (Perkins 2023), Lehrmaterialien für Bildungszwecke (Tlili et al. 2023), Kunstinhalte durch bildausgebende genKI (Hutson und Cotroneo 2023), Baupläne und Architekturdesigns (Ploennigs und Berger 2023) oder Beiträge in sozialen Medien (Huang und Sciuchetti 2023). Darüber hinaus umfasst diese Kategorie Fähigkeiten wie das Vorschlagen von Ansätzen für verschiedene Vorhaben, zum Beispiel für den Einsatz von Problemlösungsmethoden (Guttupalli et al. 2023), zur Stimulierung zum Finden von Ideen (Guo et al. 2023; Memmert und Tavanapour 2023) oder zur Aufwertung von Ideen (Yun et al. 2022). Auch die Inhaltsarbeit wurde durch die Organisation und Zusammenfassung von Wissen (Robinson 2023) sowie die Evaluierung von Entscheidungen durch Aufzeigen von Vor- und Nachteilen hervorgehoben (Soman und Ranjani 2023). Eine Studie, die sich auf bildausgebende genKI konzentriert, beschreibt zudem die Nutzung dieser Technologie zur Inspirationsfindung in einer Mensch-KI-Kollaboration (Guo et al. 2023). Diese Kategorie wurde in 25 aller untersuchten Artikel (49 %) identifiziert und umfasst erwartungsgemäß und namensgebend für genKI, zahlreiche Anwendungsszenarien.

¹ Die Konzeptmatrix in Abb. 2 enthält die vollständige Auflistung des finalen Literaturdatensatzes. Um die Lesbarkeit im Fließtext zu gewährleisten, werden jedoch lediglich exemplarische Quellen aus dem Datenbestand zitiert.

Die nachfolgende Kategorie innerhalb der Fähigkeiten fokussiert sich auf die Erkennung von *Mustern* sowie die *Klassifikation* dieser. Im Kontext dieser Kategorie wird genKI als (binärer) Klassifikator für Umfragen (Parker et al. 2023), zur Klassifikation von Text (Sirithumgul 2023) oder Nutzerbewertungen (Weingart et al. 2023) eingesetzt. Darüber hinaus wird sie für spezifische Mustererkennungen im Rahmen von Datenanalysen mit anschließender klassenbasierter Strukturierung der Inhalte genutzt, illustriert am Beispiel der Kreditwürdigkeitsbewertung (Feng et al. 2023). Die Einzelfähigkeiten dieser Kategorie wurden in vier Artikeln (8 %) des Datensatzes identifiziert.

Die dritte identifizierte Kategorie umfasst die *Generierung und Verbesserung von Programmiercode* sowie die technische *Automatisierung* von Prozessen. Im Mittelpunkt steht die Generierung von Programmiercode in verschiedenen Programmiersprachen für unterschiedliche Zwecke, beispielsweise zur Unterstützung beim Erlernen der Programmierung (Alkhaqani 2023) oder für möglicherweise missbräuchliche oder Infrastruktur stresstestende Cyberattacken (Iqbal et al. 2023). Weitere Anwendungsszenarien beinhalten die Analyse von Datenflüssen und die Untersuchung von Verhaltensweisen von Software (Wang et al. 2023) sowie die technische Automatisierung von A/B-Tests aus dem Marketingbereich (Kowalczyk et al. 2023). Diese spezifischen Fähigkeiten wurden in vier Artikeln (8 %) des Datensatzes festgestellt.

Die Kategorie der natürlichen *Sprachverarbeitung* stellt die nächste in der Analyse identifizierte Kategorie dar. Eine prominent hervortretende und genannte Fähigkeit liegt in der Verarbeitung und Ausgabe von korrekter, situativ angepasster, relevanter und verständlicher Sprache, die von menschlicher Kommunikation kaum zu unterscheiden ist (Busch und Hausvik 2023; Hutson und Plate 2023a), beispielsweise in Form von Textantworten auf gestellte Fragen (AlAfnan et al. 2023; Zeng 2023). Übersetzungsszenarien sowie Formulierungsmöglichkeiten in diversen Sprachen wurden ebenso positiv in Bezug auf die GPT-Modelle hervorgehoben (Alaswad et al. 2023; Lappalainen und Narayanan 2023). Des Weiteren wurde die gute Qualität und Verständlichkeit der ausgehenden Sprache in Textform betont (Huang et al. 2023), beispielsweise bei der Konsolidierung von Informationen aus verschiedenen Quellen (Regalia 2023) oder bei der Nutzerbewertung eines Produktes (Weingart et al. 2023). Es wurde zudem die effiziente Beantwortung von Testfragen oder Aufgaben durch genKI in Examen zu Informationssystemen genannt (Hobert et al. 2023). In diesem Zusammenhang spielt auch die Verbesserung von Texten sowohl grammatikalisch als auch inhaltlich eine signifikante Rolle (Leadbetter 2023; Regalia 2023), etwa durch das Bereitstellen von konkretem Feedback (Hutson und Plate 2023c). Eine weitere identifizierte Einzelfähigkeit dieser Kategorie umfasst die Kontexterkenkung (Fan und Jiang 2023), beispielhaft illustriert durch individualisierte Kundeninteraktionen (Huang und Sciuchetti 2023) und Sentiment-Analysen (Kumar et al. 2023) sowie die Fähigkeit, Sachverhalte zu vergleichen und auf neue Inhalte anzuwenden (Campbell 2023). Die Kategorie der natürlichen Sprachverarbeitung wurde mit 31 zugehörigen Artikeln (61 %) am häufigsten erkannt.

Die fünfte Kategorie charakterisiert *Kreativität* als eine der Fähigkeiten von genKIs. Im Zentrum stehen hier das Erstellen kreativer Ideen und Lösungen (Tsai et al. 2023) sowie die Unterstützung kreativer Prozesse durch das Vorschlagen von Herangehensweisen und Ansätzen, beispielsweise für wissenschaftliche Texte (Hutson und

Plate 2023a), das Geben von Denkanstößen bei mathematischen Herausforderungen (Guttupalli et al. 2023) oder die Begleitung von Brainstorming-Sitzungen (Memmert und Tavanapour 2023). In einem Artikel wurde die Kombination von text- und bildausgebenden genKIs hervorgehoben, wobei insbesondere die generierten Bildideen als kreativ betont wurden (Hutson und Cotroneo 2023). Kreative Fähigkeiten von genKIs wurden in fünf Artikeln (10 %) des finalen Datensatzes identifiziert.

Die *Adaptabilität* von genKI stellt eine weitere Kategorie der Fähigkeiten dar, die im Rahmen der Literaturanalyse identifiziert wurde. Diese umfasst die Fähigkeit, basierend auf vorliegenden oder vorgegebenen Situationen unterschiedliche Rollen einzunehmen (Asfour und Murillo 2023), beispielsweise die historischer Personen (Hutson und Schnellmann 2023) und sich in verschiedene Szenarien hineinzuversetzen, ob in Diskussionen (Zeng 2023) oder in die Beratung bei Sportereignissen (Robinson 2023). Zudem ist genKI in der Lage, zielgruppenspezifisch, etwa auf physikalische Fragestellungen zu antworten (dos Santos 2023) oder angepasste Marketinginhalte zu erstellen (Kowalczyk et al. 2023). Die konkrete Nennung oder Anwendung dieser Fähigkeit wurde in 11 Artikeln (22 %) des gesamten Datensatzes identifiziert.

Der Bereich *Lernunterstützung* durch genKI wurde in einigen Artikeln positiv hervorgehoben und stellt eine weitere Fähigkeitenkategorie dar. Hierbei geht es insbesondere um die Unterstützung von Lehrenden, etwa durch das Bereitstellen von Feedback oder das Erläutern komplexer Prozesse an Lernende, wodurch die Arbeitslast der Lehrenden verringert werden kann (Tlili et al. 2023). Dies trägt zur Schaffung einer individuelleren personalisierten Lernumgebung bei und kann die Motivation der Lernenden steigern (Campbell 2023; Vasconcelos und dos Santos 2023). Ein Artikel illustriert exemplarisch, wie genKI ein tieferes Verständnis für Schülerinnen und Schüler und ihre Schwierigkeiten mit Inhalten aus dem Mathematik- und Sprachunterricht aufwies als menschliche Lehrerinnen und Lehrer (Alaswad et al. 2023). Anwendungsszenarien dieser Kategorie wurden in fünf Artikeln (10 %) des gesamten Datensatzes identifiziert.

Die letzte Kategorie von Fähigkeiten reflektiert das allgemeine *Domänenwissen* von genKIs, insbesondere ihr spezifisches Wissen in diversen Grundlagendomänen. Dieses fundierte Wissen in Bereichen wie Informationssystemen (Weingart et al. 2023) oder Medizin (Alaswad et al. 2023) ermöglicht es den genKIs beispielsweise Vorschläge für Ernährungspläne zu erstellen (Tsai et al. 2023). Dies wird in den Artikeln als eine herausragende Fähigkeit betont. In einigen Auswertungen wird zudem die Genauigkeit der von genKIs generierten Inhalte genannt (Torres 2023). Ein Artikel im Bereich der Geografie unterstreicht die effektive Handhabung komplexer Workflows durch GPT-4 (Juhász und Guan 2023). Insbesondere bei bildausgebenden genKIs wurden die Designfähigkeiten hervorgehoben (Guo et al. 2023). Fähigkeiten, die dem Domänenwissen zugeordnet sind, wurden in 10 Artikeln (20 %) des gesamten Datensatzes identifiziert.

4.2 Limitationen generativer künstlicher Intelligenz

Eine Kategorie der identifizierten Limitationen umfasst *Ethik und Voreingenommenheit*; letzteres wird häufig Bias genannt. Die ethische Perspektive bezieht sich primär

auf die digitale Ethik, welche durch die zunehmende Verbreitung von KI-Systemen stetig wichtiger wird (D’Onofrio und Portmann 2022). In dieser Kategorie werden insbesondere Themen wie (datenbasierte) Voreingenommenheit beispielsweise bei der Bewertung der Kreditwürdigkeit (Feng et al. 2023) oder der Zusammenfassung von Nachrichtenartikeln (Huang et al. 2023) hervorgehoben. Weitere bedeutsame Aspekte sind Datenschutz- und Urheberrechtsfragen (Kowalczyk et al. 2023; Panagopoulou et al. 2023). Mehrere Artikel berichten von genKI-Anwendungsszenarien, in denen inkonsistente oder nicht reproduzierbare Antworten ausgegeben (Tsai et al. 2023) sowie zufällige Ergebnisse generiert wurden, sowohl in Text- als auch in Bildausgaben (Guo et al. 2023). Dies stellt die Verlässlichkeit der genutzten Sprachmodelle in Frage und kann im Ernstfall zu rechtlichen Konsequenzen (Alaswad et al. 2023) sowie zu unklaren Verantwortungszuschreibungen führen (Ali und OpenAI 2023). Aspekte dieser Limitationskategorie wurden in 15 Artikeln (29 %) des gesamten Datensatzes erwähnt.

Auf *Datenbedingtheit und Halluzinationen* bezieht sich eine weitere Kategorie von Limitationen. Unter Halluzinationen versteht sich im genKI-Kontext, dass plausibel aussehende Ergebnisse generiert werden, die jedoch Fehlinformationen oder erfundene Daten enthalten (Hu et al. 2023; Huang und Sciuchetti 2023). Es kommt teilweise vor, dass Informationen im Dialogverlauf vergessen werden oder die Antworten unvollständig sind (Alaswad et al. 2023), was insbesondere bei Chatbots zu fragwürdigen Ergebnissen führen kann (Lappalainen und Narayanan 2023). Abseits dieser Verhaltensweisen existieren weitere genannte Defizite, welche die Qualität von genKI-Ausgaben negativ beeinflussen. So tendieren diese dazu, keine, falsche oder fiktive Referenzen zu ihren Aussagen anzugeben (Busch und Hausvik 2023), was die Quellenrecherche und -validierung erschwert (DeKay 2023). Zudem können veraltete Daten, basierend auf dem Datenbestand des genutzten Sprachmodells, die Erwartungen der Nutzerinnen und Nutzer unerfüllt lassen (Torres 2023). In 14 Artikeln (27 %) des Datensatzes wurden Limitationen identifiziert, die dieser Kategorie zugeordnet werden konnten.

Darüber hinaus wurde *Unterstützungsbedarf* als eine Kategorie von Limitationen aus den analysierten Artikeln gebildet. Darin sind Aussagen enthalten, dass beispielsweise die Ergebnisse aus der Nutzung von genKI ohne weitere menschliche Unterstützung häufig nicht nutzbar sind, da sie zu generisch und unoriginell sind (AlAfnan et al. 2023; Hutson und Schnellmann 2023). Ebenso wird in einigen Artikeln betont, dass genKIs Mängel in der Transferleistung aufweisen können, etwa in der Führung einer Diskussion (Alkhaqani 2023) oder dem sinnvollen Einbringen von Emotionen (Fan und Jiang 2023). Nicht nur die Ausgaben von genKI-Anwendungen werden erwähnt, sondern auch die Eingaben, die teilweise mehrdeutig sein können, wie beispielsweise sarkastische Texte. Hier muss zusätzliche Unterstützung geleistet werden, um eindeutige Begriffe und Aussagen zu liefern (Soman und Ranjani 2023). Zudem wird die Notwendigkeit hervorgehoben, Expertise im Bereich großer Sprachmodelle zu besitzen. Dies ist für die Auswahl und Implementierung dieser Modelle für den Fall des Eigenbetriebs in Organisationen (Cummings et al. 2021) oder zur Schulung von anderen Mitarbeiterinnen und Mitarbeitern wichtig (Vasconcelos und dos Santos 2023). Im Gegensatz zur bereits aufgeführten Fähigkeit, menschliches Kommunikationsverhalten in Form von Texten gut nachahmen zu

können, gibt es auch Aussagen dazu, dass dies eine Grenze darstellen kann und die Erwartungen für spezifische Anwendungsszenarien nicht erfüllt werden (Asfour und Murillo 2023). In 11 Artikeln (22 %) des Datensatzes wurden Aussagen gefunden, die dieser Kategorie zugeordnet werden konnten.

In Bezug auf die grundlegende Fähigkeit von genKIs, Programmiercode zu generieren und zu analysieren sowie Berechnungen durchzuführen, wurden auch Limitationen für diese Anwendungsszenarien gefunden, die unter der Kategorie *Codeanalyse und Berechnungen* zusammengefasst sind. In den Disziplinen der Mathematik und Physik stießen einige genKI-Modelle auf Schwierigkeiten, korrekte Berechnungen durchzuführen (Juhasz und Guan 2023), Einheiten umzurechnen und Gleichungen zu lösen (dos Santos 2023). Spezifische Aufgaben wie Debugging (Alaswad et al. 2023), die Anwendung von Code-Konzepten (Wang et al. 2023) oder spezielle Fälle wie das testweise Erstellen von Malware (Iqbal et al. 2023) erwiesen sich bei der Umsetzung durch genKIs nicht als vollumfänglich zuverlässig. Relevante Aussagen zu dieser Limitationskategorie wurden in sechs der analysierten Artikel (12 %) des Datensatzes gefunden.

Weitere aus einigen Artikeln erfasste Aussagen wurden in die Limitationskategorie *technische Kapazitäten* eingeordnet. Insbesondere bei chat-basierten Applikationen traten das Token-Limit (Datenmenge, die in einer Anfrage bearbeitet werden kann) und die begrenzte Kontextlänge als Limitationen hervor (Huang et al. 2023; Lappalainen und Narayanan 2023). Des Weiteren wurden einige fehlende Funktionalitäten angeführt, wie beispielsweise die Handhabung multimedialer Eingaben oder die Integration von Echtzeitdaten, zum Beispiel durch den Zugriff auf das Internet (Hobert et al. 2023; Robinson 2023). Zudem wurden bei bildausgebender genKI die technischen Möglichkeiten bei der Erstellung von Bildern als unzureichend für Designer bezeichnet (Guo et al. 2023) sowie eine häufig fehlerhafte Darstellung von Text in generierten Bildern angeführt (Hutson und Lang 2023). Solche technischen Limitationen wurden in insgesamt acht der analysierten Artikel (16 %) des Datensatzes thematisiert.

Im Gegensatz zur Fähigkeit der Sprachverarbeitung wurden in einigen Aussagen die *Textverarbeitung und -ausgabe* in spezifischen Anwendungsszenarien als unzureichend bewertet, weshalb diese als eigenständige Limitationskategorie erfasst wurde. Beispielsweise wurde der Aspekt der Sentiment-Analyse thematisiert, welche in einigen Fällen nicht zufriedenstellend bewältigt werden konnte, wie bei der Untersuchung widersprüchlicher Texte (Tang 2023) oder Hassreden (Kumar et al. 2023). Weitere Limitationen in der textuellen Verarbeitung umfassen das Erkennen der Textstruktur (Fan und Jiang 2023), die Interpretation von besonders langen Texten (Huang et al. 2023) sowie Inkonsistenzen im Ton und Stil von Textausgaben (Robinson 2023). Aussagen, die dieser Limitationskategorie zugeordnet wurden, fanden sich in insgesamt 9 Artikeln (18 %) des Datensatzes.

Die abschließende Kategorie *Universalität und Originalität* umfasst identifizierte Limitationen, bei denen in den Anwendungsszenarien der analysierten Artikel zu generische, daher nicht hilfreiche Ausgaben von genKIs festgestellt wurden. Dies trat insbesondere in Szenarien hervor, in denen spezifisches Domänenwissen erforderlich war (Fan und Jiang 2023), beispielsweise in den Bereichen Physik (dos Santos 2023), Geografie (Juhasz und Guan 2023), Medizin (Sirithumgul 2023) oder bei bestimmten

Methoden wie der SWOT-Analyse (Torres 2023). Dabei tendieren genKI-Modelle zu oberflächlichen Ausgaben, etwa aufgrund fehlender Daten oder fehlendem Bezug zum Anwendungsszenario (Busch und Hausvik 2023; Hu et al. 2023; Robinson 2023). Zudem wird über sich wiederholende und eintönige Antworten berichtet (Alaswad et al. 2023; Leadbetter 2023). Dieser Limitationskategorie zugeordnete Aussagen wurden in 15 aller Artikel (29 %) des Datensatzes gefunden.

4.3 Häufigkeit der genutzten Sprachmodelle und Interfaces

Ergänzend zu den aggregierten Kategorien an Fähigkeiten und Limitationen sind die eingesetzten Sprachmodelle oder Interfaces der identifizierten genKI-Anwendungsszenarien aus den analysierten Artikeln in Tab. 3 zusammengefasst. Hierbei ist erkennbar, dass erwartungskonform vor allem ChatGPT und die GPT-Modelle bevorzugt eingesetzt wurden, was unter anderem ihrer medialen Aufmerksamkeit und Position als Wegbereiter zuschreibbar ist. Darüber hinaus kamen die textausgebenden genKI-Interfaces Bard und Llama zum Einsatz. In vielen Fällen wurden keine spezifischen Angaben zur zugrundeliegenden Version des Sprachmodells bei Nutzung der Interfaces wie ChatGPT und Bard gemacht, sodass anhand des Erscheinungsdatums eines Artikels auf die Version rückgeschlossen werden müsste. Dies erschwerte bei der Analyse die Vergleichbarkeit der Fälle, da unklar ist, ob beispielsweise eine für ChatGPT ausgegebene Limitation für GPT-3 oder GPT-4 gilt. Bei GPT-4 könnte eine für GPT-3 ausgegebene Limitation aufgrund von Erweiterungen und Verbesserungen der Funktionalitäten bereits nicht mehr bestehen. In den identifizierten Artikeln, die sich mit bildausgebender genKI befassen, wurde vornehmlich auf DALL-E 2 und Midjourney zurückgegriffen. Alle Sprachmodelle oder Interfaces, die insgesamt nur einmal explizit erwähnt wurden, beispielsweise PaLM2, sind in Tab. 3 unter der Rubrik „Andere“ zusammengefasst und können im Onlinematerial einzeln eingesehen werden.

5 Vorabumfrage zur Nutzung von generativer künstlicher Intelligenz

Zu Beginn der Umfrage wurde die Bekanntheit und Erfahrung mit einschlägigen text- und bildausgebenden genKI-Interfaces erhoben. Dabei wurden Sprachmodelle nicht gesondert abgefragt, da ihre Nutzung am häufigsten über die genannten Interfaces wie ChatGPT oder Bard erfolgt. Es wurden einige auswählbare Optionen vorgeschlagen und es war möglich, weitere genKI-Interfaces als Freitext anzugeben. Alle Befragten gaben an, bereits Erfahrungen mit ChatGPT gemacht zu haben, wobei 29 Personen (72,5 %) über eine hohe oder sehr hohe Expertise mit diesem Interface berichteten. Die Bekanntheit der anderen Interfaces fiel im Vergleich deutlich geringer aus: 27 Befragte (67,5 %) haben JasperChat, 26 (65 %) Otter, 24 (60 %) Aleph Alpha, 23 (57,5 %) YouChat und 19 (47,5 %) Bard noch nie genutzt. Im Bereich der bildausgebenden genKI hingegen gab es kein Interface, mit dem alle Befragten bereits Erfahrung hatten. Midjourney stellte sich als das populärste Interface heraus, mit 17 Befragten (42,5 %), die eine hohe oder sehr hohe Erfahrung damit angaben. Die anderen Interfaces waren deutlich weniger bekannt: 27 Befragte (67,5 %)

haben Artsmart.ai, 24 (60 %) Firefly, 23 (57,5 %) Leonardo.Ai, 22 (55 %) Canva, 21 (52,5 %) DALL-E und 15 (37,5 %) Stable Diffusion noch nie genutzt.

Weitere genannte Optionen von neun Befragten (22,5 %) umfassten AutoGPT, CLIP, automatic1111, Github Copilot und ChatPDF. Dies deutet darauf hin, dass einigen Befragten nicht klar war, was genau unter genKI zu verstehen ist. CLIP ist ein von OpenAI trainiertes Modell zur Verarbeitung von Texten und Bildern, jedoch nicht speziell zur Generierung von Inhalten bestimmt, auch wenn es einzelne Funktionalitäten wie das Erzeugen von Vorschlägen für Bildbeschreibungen bietet. Die weiteren Nennungen beziehen sich teils auf ChatGPT-Plugins (wie ChatPDF), welche die Funktionalitäten eines Sprachmodells über ein Interface erweitern können, oder teils auf tatsächliche alternative genKI-Interfaces. Die Ergebnisse über die Verteilung der Nutzungshäufigkeit von text- oder bildausgebenden genKI-Interfaces sind in Abb. 3 dargestellt.

Die Umfrageergebnisse zeigen eine weitreichende und regelmäßige Nutzung von genKI-Interfaces unter den Teilnehmenden. In Bezug auf die Nutzung im beruflichen oder privaten Kontext gaben nur zwei Teilnehmende (5 %) an, genKI weder beruflich noch privat zu nutzen. Fünf (12,5 %) nutzen genKI ausschließlich beruflich und vier (10 %) lediglich im privaten Rahmen. Die überwiegende Mehrheit, 29 Teilnehmende (72,5 %), setzt genKI sowohl privat als auch beruflich ein. Für 18 Teilnehmende (45 %) hat genKI bereits einen Einfluss auf ihre berufliche Arbeitsweise beziehungsweise täglichen Aktivitäten. Darüber hinaus gaben 23 (57,5 %) in einer Selbsteinschätzung an, ausreichend über die Funktionsweise von genKI informiert zu sein, um diese sicher und vorteilhaft im beruflichen oder akademischen Kontext nutzen zu können. Zusätzlich sehen 28 Teilnehmende (70 %) eine zentrale Rolle für genKI als digitalen Assistenten für Menschen voraus, während 12 (30 %) dieser Ansicht nicht zustimmen.

Die Befragten äußerten sich detaillierter zur Rolle von genKI als digitalen Assistenten und nannten dabei assistierende Tätigkeiten, die sich auf verschiedene aus der systematischen Literaturrecherche identifizierten Fähigkeiten beziehen lassen. Beispiele hierfür sind Unterstützung bei Informationsbedarf durch Recherchetätigkeiten, Funktionen als Tages- und Kalenderplaner, Ersatz aktueller sprachbasierter

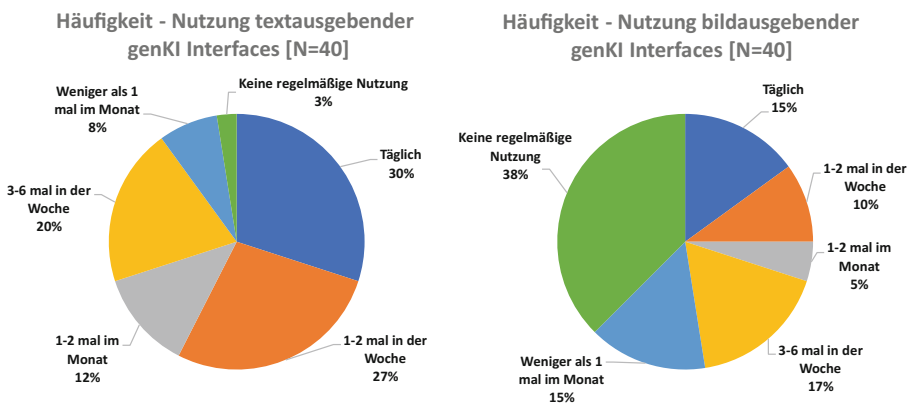


Abb. 3 Häufigkeit der Nutzung text- und bildausgebender genKI-Interfaces

Assistenten wie Alexa (Echo) oder Siri sowie Hilfe bei Aufgaben wie dem Verfassen von E-Mails, Protokollen oder ToDo-Listen. Darüber hinaus wurden Tätigkeiten als Übersetzer, Kommunikationsvermittler und Unterstützer älterer Menschen angeführt. Besondere Erwähnung fand die Fähigkeit von genKI zur Generierung von Bildern für Werbezwecke oder zur Erstellung erster Entwürfe für (Produkt-)Konzepte zur Inspirationsfindung. Angesichts der in der Literatur gefundenen und von den Befragten angegebenen Anwendungsszenarien wird die universelle Einsetzbarkeit dieser Technologie besonders deutlich. Sie kann Menschen erheblich bei zahlreichen privaten und beruflichen Tätigkeiten unterstützen, ihre Effizienz verbessern und ihre Kapazitäten hinsichtlich der bewältigbaren Aufgabenmenge erweitern. In der Umfrage wurden zudem spezifische Anwendungsszenarien erörtert, die in der Literatur nicht erfasst wurden, wie der Einsatz von genKI für die Reiseplanung oder die Auswahl von Geschenken. Insgesamt konnten alle in der Literaturrecherche identifizierten Fähigkeiten durch die praktischen Angaben, die aus den Antworten der 40 Befragten hervorgingen, bestätigt werden.

Gemäß den Rückmeldungen der Befragten weist genKI für sie auch Limitationen auf. Bei der Frage zu Anwendungsszenarien, in denen die Technologie nicht zufriedenstellend eingesetzt werden konnte oder enttäuschte, bestätigten 25 Teilnehmende (62,5 %) solche Erfahrungen gemacht zu haben. Die Gründe hierfür waren unter anderem Limitierungen bei den technischen Kapazitäten und der Sprachverarbeitung, etwa das Übergehen von Kontext bei langen Textverläufen oder nicht berücksichtigte Vorgaben bei der Bildgenerierung. Auch der Einsatz von genKI für die Unterstützung bei Programmiertätigkeiten erfüllte nicht immer die Erwartungen. Weiterhin wurden gesundheitliche und psychologische Fragen als Szenarien genannt, in denen keine zufriedenstellenden Ergebnisse erzielt werden konnten. Plugins, die das Auslesen von Informationen aus PDF-Dateien ermöglichen, wie beispielsweise ChatPDF, konnten aufgrund von Ungenauigkeiten ihre Versprechen nicht immer erfüllen und enttäuschten ebenso wie die Ergebnisse aus Aufforderungen an genKI-Interfaces, Quellen für ihre Ausgaben bereitzustellen, die oftmals halluziniert wurden. Auf die Frage nach Anwendungsszenarien, in denen genKI auch in Zukunft nicht eingesetzt werden würde, umfassten die Antworten haptische Tätigkeiten, Aktivitäten, die eine tiefgreifende soziale und empathische Kommunikation erfordern und spezifische Tätigkeiten, für die es kaum Daten gibt, wodurch Sprachmodelle meist keine nützlichen Antworten generieren können. Insgesamt konnten die in der Literatur identifizierten Limitationen durch die Antworten der 40 Befragten auch aus der Praxis bestätigt werden.

Die Ergebnisse der Vorabumfrage offenbaren auch eine gewisse Ambivalenz der Befragten hinsichtlich genKI. Bei der Frage nach der Einstellung zu Entscheidungen, die vollständig von genKI getroffen werden, äußerten sich 17 Teilnehmende (42,5 %) neutral. Dagegen bekundeten 12 (30 %) geringes oder sehr geringes Vertrauen in die Entscheidungsfindung durch genKI, während 11 (27,5 %) hohes oder sehr hohes Vertrauen mitteilten. Eine ähnliche Divergenz zeigte sich bei der Sorge um ethische Fragestellungen im Zusammenhang mit genKI, insbesondere in den Bereichen Datenschutz und Urheberrecht. Hier äußerten 20 Teilnehmende (50 %) wenig oder keine Besorgnis, im Gegensatz zu 19 (47,5 %), die sich besorgt oder äußerst besorgt zeigten.

6 Diskussion und Einordnung

Die Auswertung und Einordnung der durch die systematische Literaturrecherche und Vorabumfrage erlangten Einsichten beleuchtet die vielschichtige Landschaft von genKI im aktuellen Forschungs- und Anwendungsrahmen. Die Interpretation der Daten zeigt die dynamische Natur der Entwicklung und Anwendung von genKI-Interfaces und der zugrunde liegenden Sprachmodelle, insbesondere im Verlauf des letzten Jahres, wie Tab. 1 verdeutlicht. Erwartungskonform bestätigt die Literatur als auch die Vorabumfrage deutlich, dass die GPT-Sprachmodelle am weitesten verbreitet sind und dass der dahinterstehende Entwickler OpenAI in vielerlei Hinsicht eine Vorreiterrolle einnimmt. Ein wesentlicher Aspekt, der bei der Einordnung der Erkenntnisse dieses Artikels und der Betrachtung anderer Literatur zu berücksichtigen ist, betrifft das jeweils eingesetzte Sprachmodell und dessen Version (Alaswad et al. 2023). Beispielsweise ist die alleinige Angabe eines Interfaces wie ChatGPT unzureichend, da über dieses verschiedene Sprachmodellversionen genutzt werden können, abhängig etwa davon, ob die Nutzerin oder der Nutzer einen kostenpflichtigen Plus-Zugang besitzt. Darüber hinaus könnten einige der in den analysierten Artikeln genannten Limitationen, beispielsweise im Kontext von GPT-3, in späteren Versionen wie GPT-3.5 oder GPT-4 bereits adressiert und abgemildert worden sein, sodass diese möglicherweise obsolet geworden sind (Frenkel und Emara 2023). Die präzise Angabe der eingesetzten Modelle und des Zeitraums der Untersuchungen sind für die Vergleichbarkeit und Beurteilung von genKI unerlässlich und können durch zusätzliche Parameterangaben wie „Temperature“ oder „Frequency“, die im Falle von GPT-Modellen Einfluss auf die Ausgaben haben, ergänzt werden (Soman und Ranjani 2023). In der durchgeführten Literaturrecherche zeigte sich diese Informationslücke in einigen Fällen, was die Analyse erschwerte.

Der Aspekt variierender Sprachmodellversionen bietet eine mögliche Erklärung für aufgedeckte Widersprüche. In der Literatur wurde Domänenwissen als Fähigkeit und Stärke von genKI herausgestellt (Juhasz und Guan 2023; Lappalainen und Narayanan 2023), während diese in den Umfragen teilweise kritisiert wurde. Ähnliches zeigte sich im Bereich der Programmierung, wo die Möglichkeiten mit genKI als weitreichend erachtet, jedoch auch Schwächen notiert wurden (Iqbal et al. 2023; Wang et al. 2023). Ein Erklärungsansatz in diesem Kontext könnte die vorhandene oder fehlende Expertise der Nutzerinnen und Nutzer im Umgang mit genKI-Interfaces und Sprachmodellen sein. Daran anknüpfend stellt die Qualität der Prompts einen entscheidenden Faktor für die Leistungsfähigkeit von genKI und der Kommunikation mit Sprachmodellen in spezifischen Anwendungsszenarien dar (Hutson und Cotroneo 2023). Unpassende oder mehrdeutige Formulierungen sowie fehlender, jedoch erforderlicher Kontext können zu unbefriedigenden Ergebnissen führen (Parker et al. 2023). In vielen Forschungs- und Praxisartikeln bleibt die Qualität der Prompts eine Blackbox, da diese nicht eingesehen werden können. Dies erschwert die Bewertung der Fähigkeiten und Limitationen von genKI, da enttäuschende Ergebnisse eine tatsächliche Limitation der Technologie und des Sprachmodells darstellen können oder möglicherweise auf unklare oder ungeeignet formulierte Prompts zurückzuführen sind. Um einen validen Eindruck zu erhalten, sollte darauf geachtet werden, dass mehrere unabhängige Tests in gleichen Anwendungsszenarien durchgeführt werden,

die zu gleichen oder ähnlichen Ergebnissen sowohl hinsichtlich der Fähigkeiten als auch Limitationen führen.

Auf Grundlage der Ergebnisse der Vorabumfrage zeichnet sich eine bestimmte Ambivalenz hinsichtlich des Vertrauens in genKI als Entscheidungsträger und Technologie ab, die teilweise mit den unterschiedlichen Erfahrungen und Wahrnehmungen in Zusammenhang steht. Personen mit sehr hoher Erfahrung tendieren dazu, ein höheres Vertrauen in die Fähigkeiten der Technologie anzugeben, im Vergleich zu Personen mit geringerer Erfahrung. Diese Art von Diskrepanzen könnten auch Auswirkungen auf die zukünftige Akzeptanz und Anwendung dieser Technologie haben. In diesem Kontext sind auch bestehende ethische Diskussionen zu nennen, die für KI im Allgemeinen und genKI im Besonderen gelten, wie beispielsweise das datengetriebene (rassistische und sexistische) Profiling (Intahchomphoo und Gundersen 2020) oder den Einsatz beim wissenschaftlichen Schreiben (Lund und Ting 2023). Die Autoren des vorliegenden Artikels betrachten eine grundlegende genKI-Bildung an Hochschulen und bereits in Sekundarstufen als wichtig, um jungen Menschen zu ermöglichen, sich auf einer fundierten Wissensbasis eine Meinung zu bilden, genKI-Ausgaben kritisch zu reflektieren und im Falle der Nutzung, die Fähigkeiten der Technologie unter Berücksichtigung der gegenwärtigen Limitationen bestmöglich auszuschöpfen. In Bildungsumgebungen sind bereits Integrationen in die Lehrpraxis in Form angepasster oder neuer Lehrveranstaltungen im Gange (Cooper 2023). Wie auch die Literaturrecherche zeigte, gab es eine feststellbare Mehrheit von Anwendungsszenarien, die in Bildungsumgebungen stattfanden, zum Beispiel von AlAfnan et al. (2023) und Boscardin et al. (2023).

Abseits des Vertrauens fühlen sich viele Berufstätige insbesondere aufgrund der generativen Fähigkeiten in Bezug auf ihre berufliche Stellung bedroht. Ein Beispiel hierfür war der Streik der Gewerkschaft SAG-AFTRA in Hollywood, bei dem klare Regulierungen bezüglich des Einsatzes von KI in der Filmindustrie gefordert wurden (Padtberg 2023). Diese Regulierungen sollen Bereiche wie das Verfassen von Geschichten für künstlerische Werke, Drehbücher für Filmaufnahmen oder auch das Modifizieren von Bild- und Videoinhalten, etwa zum künstlichen Verjüngen von Schauspielerinnen und Schauspielern, betreffen. GenKI als Technologie bietet durch ihre Fähigkeiten dahingehend besonders Produktionsfirmen Vorteile, die damit ihre Prozesse effizienter gestalten, unterstützen und weiterführend wirtschaftlicher machen können. Dieses Beispiel lässt sich auch auf andere Anwendungsszenarien und technologiedurchdrungene oder wissensintensive Branchen übertragen und wird vermutlich noch eine längere Zeit Debatten nach sich ziehen, wie auch ein Bericht von McKinsey & Company über die Implikationen auf dem amerikanischen Arbeitsmarkt andeutet (Ellingrud et al. 2023). Eine Entwicklung in dieser Hinsicht stellt der von der EU verabschiedete AI Act dar, der unter anderem für genKI vorschreibt, dass im Sinne der Transparenz KI-generierte Inhalte gekennzeichnet werden müssen und keine Inhalte für illegale Zwecke erstellt werden dürfen (Europaparlament 2023). Diese Anforderungen werden von zahlreichen in diesem Artikel aufgeführten genKI-Interfaces beziehungsweise Sprachmodellen derzeit nicht erfüllt, sodass weitere Entwicklungen in diesem Kontext zu erwarten sind. Dies nimmt auch Bezug auf eine breit geführte Debatte darüber, wie im Bildungsbereich damit umgegangen wird, dass durch genKI Leistungsnachweise unterstützt oder nahezu vollständig

übernommen werden können und die Eigenständigkeit, der Lerneffekt sowie gegenwärtige Prüfungsformen hinterfragt werden, wie auch Boscardin et al. (2023) in der medizinischen Bildung aufführen.

Als Beispiel herangezogen erfordert der Einsatz von genKI in der Filmindustrie ein grundlegendes Maß an Kreativität und Originalität, insbesondere beim Verfassen einer kohärenten Geschichte. Bei Anwendungsszenarien, in denen es auf Kreativität ankommt, beispielsweise in Bereichen der Problemlösung und Ideengenerierung, kann diese Technologie laut Memmert und Tavanapour (2023) nutzbringend eingesetzt werden. Demgegenüber stehen Aussagen, die Kreativität als eine spezifisch menschliche Fähigkeit definieren, die für eine künstliche Intelligenz herausfordernd und nicht in gleicher Weise realisierbar ist (Chen et al. 2023). Es kommt hinzu, dass Sprachmodelle auf der Grundlage von Datensets vortrainiert sind, was bedeutet, dass sie auf bereits existierenden und teilweise veralteten Daten arbeiten und (ohne Internetzugriff) nur diese nutzen können (Waqas et al. 2023). Dennoch können aus Datenkombinationen neue und vielversprechende Ideen entstehen, die es in dieser Form noch nicht gab und in Rahmen von Mensch-KI-Kollaborationen weiterentwickelbar sind (Guo et al. 2023).

7 Zusammenfassung, Limitationen und Ausblick

Im Rahmen dieses Artikels wurde eine eingehende Untersuchung zu den Fähigkeiten und Limitationen von genKI vorgenommen. Durch die Durchführung einer systematischen Literaturrecherche und die Auswertung einer Vorabumfrage konnten umfassende Einsichten sowohl aus wissenschaftlicher als auch aus anwenderorientierter Perspektive gewonnen werden. Die Analyse offenbart eine dynamische Entwicklungslandschaft, in der genKI bereits in einer breiten Palette von Anwendungsszenarien Verwendung findet, wenngleich bestimmte Limitationen die Qualität in spezifischen Kontexten beeinträchtigen können. Der Markt für neue genKI-Sprachmodelle und -Interfaces expandiert rasant und es werden kontinuierlich neue Alternativen angeboten; auch datenschutzkonforme Lösungen gewinnen zunehmend an Popularität und Akzeptanz. Durch die strukturierte Einordnung und Diskussion der gewonnenen Erkenntnisse bietet der Artikel als Momentaufnahme eine fundierte Einsicht in den aktuellen Stand von genKI und den eingesetzten Sprachmodellen. Darüber hinaus legt er eine Basis für nachfolgende Forschungen in diesem Bereich und bietet eine praxisorientierte Grundlage, um zu bewerten, ob genKI für verschiedene Anwendungsszenarien zweckdienlich sein könnte.

Dieser Artikel weist einige Limitationen auf, die besonders aus den methodischen Ansätzen der systematischen Literaturrecherche und der Vorabumfrage resultieren. Die Auswahl der untersuchten Literatur sowie deren Qualität könnten die Ergebnisse beeinflusst haben. Dies ergibt sich insbesondere daraus, dass aufgrund der Aktualität und Dynamik von genKI auch graue Literatur vor allem von arXiv berücksichtigt wurde, die nicht zwangsweise ein rigoroses wissenschaftliches Begutachtungsverfahren durchlaufen hat. Es besteht zudem die Möglichkeit, dass relevante Artikel nicht einbezogen wurden, falls diese in den ausgewählten elektronischen Datenbanken für die Literatursuche nicht indexiert oder durch den finalen Suchstring nicht er-

fasst wurden. Außerdem ist die Validität der Ergebnisse der Vorabumfrage durch die begrenzte Anzahl von 40 Antworten eingeschränkt. Für eine höhere Repräsentativität und Generalisierbarkeit der Erkenntnisse wäre eine umfangreichere Befragung mit einer größeren Stichprobe erforderlich. Eine bereits geplante Hauptumfrage könnte in Zukunft eine größere Anzahl von Antworten sammeln und das Verständnis der genKI-Nutzungsgewohnheiten vertiefen. Obwohl der Schwerpunkt des Artikels auf der systematischen Literaturrecherche liegt, bieten die Ergebnisse der Vorabumfrage dennoch Einblicke in die Anwendung und Wahrnehmung von genKI aus der Praxisperspektive im Rahmen einer Primärerhebung. Darüber hinaus stellt die dynamische Natur der genKI-Technologie eine weitere Limitation des Artikels dar, da sich die Fähigkeiten und Limitationen mit der Zeit und den Fortschritten in Forschung und Entwicklung rasch ändern können. Diese Faktoren unterstreichen die Notwendigkeit kontinuierlicher Untersuchungen, um das sich entwickelnde Terrain der genKI und ihre Auswirkungen auf verschiedene Anwendungsbereiche besser zu verstehen.

Die durchgeführte Untersuchung eröffnet ein breites Spektrum möglicher Forschungsrichtungen für die Zukunft. Beispielsweise könnten künftige Studien die Auswirkungen von genKI auf spezifische Sektoren oder Berufsgruppen vertieft erforschen. Die kontinuierliche Weiterentwicklung und Verbesserung der zugrunde liegenden Sprachmodelle bietet zudem ein weites Feld für die Exploration neuer Anwendungsszenarien sowie die Überprüfung bisher identifizierter Limitationen durch den Vergleich verschiedener Modelle. Angesichts seiner Bedeutung stellt weiterführend das Thema Prompting ein vielversprechendes Forschungsgebiet dar, um operationalisierbare Prinzipien für die Formulierung effektiver Prompts je nach Anwendungsszenario zu entwickeln.

GenKI hat sich als signifikante Technologie erwiesen, mit dem Potenzial, die Arbeitsweisen von Individuen und Organisationen grundlegend zu transformieren. Die gesellschaftliche Verbreitung von genKI wird voraussichtlich in den kommenden Jahren weiter zunehmen und ihre Fähigkeit, komplexe Aufgaben zu unterstützen oder zu automatisieren sowie in Entscheidungsprozesse eingebunden zu werden, wird sie nach den Meinungen der Autoren dieses Artikels zu einem unverzichtbaren Werkzeug in vielen Bereichen machen. Es ist jedoch von zentraler Bedeutung, ein tiefgreifendes Verständnis der Fähigkeiten und Limitationen dieser Technologie zu erlangen, um ihre sichere und effektive Anwendung zu gewährleisten und mögliche negative Auswirkungen zu antizipieren und zu minimieren. Die hier präsentierten Erkenntnisse leisten einen Beitrag zur fortlaufenden Diskussion über die Rolle und die Implikationen von genKI in unserer Gesellschaft und bieten eine Basis für weitere explorative und evaluative Forschungsarbeiten in diesem sich rasch entwickelnden Feld.

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1365/s40702-024-01050-x>) enthalten.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- AlAfnan MA, Dishari S, Jovic M, Lomidze K (2023) ChatGPT as an Educational Tool: Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses. *J Artif Intell Technol* 3(2):60–68. <https://doi.org/10.37965/jait.2023.0184>
- Alaswad S, Kalganova T, Awad WS (2023) Using ChatGPT and other LLMs in Professional Environments. *Inf Sci Lett* 12(9):2097–2108. <https://doi.org/10.18576/isl/120916>
- Ali F, Open AI (2023) Let the devil speak for itself: Should ChatGPT be allowed or banned in hospitality and tourism schools? *J Glob Hosp Tour* 2(1):1–6. <https://doi.org/10.5038/2771-5957.2.1.1016>
- Alkhaqani AL (2023) Potential Benefits and Challenges of ChatGPT in Future Nursing Education. *Maaen J Med Sci* 2(2):49–53. <https://doi.org/10.55810/2789-9128.1020>
- Anil R, Dai AM, Firat O, Johnson M, Lepikhin D, Passos A et al (2023) PaLM 2 Technical Report. arXiv. <https://doi.org/10.48550/arXiv.2305.10403>
- Asfour M, Murillo J (2023) Harnessing Large Language Models to Simulate Realistic Human Responses to Social Engineering Attacks: A Case Study. *Int J Cybersecur Intell Cybercrime* 6(2):21–49. <https://doi.org/10.52306/2578-3289.1172>
- Boscardin CK, Gin B, Golde PB, Hauer KE (2023) ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity. *J Assoc Am Med Coll* 99(1):22–27. <https://doi.org/10.1097/ACM.00000000000005439>
- Budhwar P, Chowdhury S, Wood G, Aguinis H, Bamber GJ, Beltran JR et al (2023) Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Hum Resour Manag J* 33(3):606–659. <https://doi.org/10.1111/1748-8583.12524>
- Busch PA, Hausvik GI (2023) Too Good to Be True? An Empirical Study of ChatGPT Capabilities for Academic Writing and Implications for Academic Misconduct. In: *Proceedings of the Americas Conference on Information Systems (AMCIS2023)*. Panama City, Panama
- Campbell PF (2023) Using ChatGPT in your Teaching. *Artificial Intelligence Initiative Papers and Presentations*. <https://digitalcommons.salve.edu/aipp/1>. Zugegriffen: 5. Nov. 2023
- Chen L, Sun L, Han J (2023) A Comparison Study of Human and Machine-Generated Creativity. *J Comput Inf Sci Eng* 23(5). <https://doi.org/10.1115/1.4062232>
- Cooper G (2023) Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *J Sci Educ Technol* 32(3):444–452. <https://doi.org/10.1007/s10956-023-10039-y>
- Cummings P, Mullins R, Moquete M, Schurr N (2021) Hello World! I am Charlie, an Artificially Intelligent Conference Panelist. In: *Proceedings of the Hawaii International Conference on System Sciences (HICSS2021)*, S 380–389
- DeKay C (2023) “Source?” “I Made It Up”: The Ethics of Citing ChatGPT in Academia (Final Projects Summer 2023). https://ir.lib.uwo.ca/fims_evolvingtech_finalproj_summer2023/7/. Zugegriffen: 16. Feb. 2024
- D’Onofrio S, Portmann E (2022) HMD 344 Digitale Ethik im KI-Zeitalter. *HMD* 59(2):435–438. <https://doi.org/10.1365/s40702-022-00859-8>
- dos Santos RP (2023) Enhancing Physics Learning with ChatGPT, Bing Chat, and Bard as Agents-to-Think-With: A Comparative Case Study. arXiv. <https://doi.org/10.48550/arXiv.2306.00724>
- Ellingrud K, Sanghvi S, Dandona GS, Madgavkar A, Chui M, White O et al (2023) Generative AI and the future of work in America. <https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america>. Zugegriffen: 5. Nov. 2023
- Europarlament (2023) EU AI Act: first regulation on artificial intelligence. <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>. Zugegriffen: 5. Nov. 2023

- Fan Y, Jiang F (2023) Uncovering the Potential of ChatGPT for Discourse Analysis in Dialogue: An Empirical Study. arXiv. <https://doi.org/10.48550/arXiv.2305.08391>
- Feng D, Dai Y, Huang J, Zhang Y, Xie Q, Han W et al (2023) Empowering Many, Biasing a Few: Generalist Credit Scoring through Large Language Models. arXiv. <https://doi.org/10.48550/arXiv.2310.00566>
- Frenkel ME, Emara H (2023) ChatGPT & Mechanical Engineering: Examining Performance on the FE Mechanical Engineering and Undergraduate Xxams. arXiv. <https://doi.org/10.48550/arXiv.2309.15866>
- García-Peñalvo FJ (2023) The perception of Artificial Intelligence in educational contexts after the launch of ChatGPT: Disruption or Panic? *Educ Knowl Soc* 24. <https://doi.org/10.14201/eks.31279>
- George AS, George ASH, Martin ASG (2023) A Review of ChatGPT AI's Impact on Several Business Sectors. *Partners Univ Int Innov J* 1(1):9–23. <https://doi.org/10.5281/zenodo.7644359>
- Grodzinsky F, Miller K, Wolf M (2023) When AI Moves Downstream. School of Computer Science & Engineering. Faculty Publications. Sacred Heart University
- Groves RM, Fowler FJ, Couper M, Lepkowski JM, Singer E, Tourangeau R (2009) *Survey Methodology*, 2. Aufl. Wiley, Hoboken (496 S)
- Gui J, Sun Z, Wen Y, Tao D, Ye J (2023) A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications. *IEEE Trans Knowl Data Eng* 35(4):3313–3332. <https://doi.org/10.1109/TKDE.2021.3130191>
- Guo X, Xiao Y, Wang J, Ji T (2023) Rethinking designer agency: A case study of co-creation between designers and AI. In: Galluzzo L, Rizzo F, Spallazzo D (Hrsg) *International Association of Societies of Design Research*. <https://doi.org/10.21606/iasdr.2023.478>
- Guttupalli S, Maloy RW, Edwards S (2023) Comparing Teacher-Written and AI-Generated Math Problem Solving Strategies for Elementary School Students: Implications for Classroom Learning. University of Massachusetts Amherst. <https://doi.org/10.7275/8sgx-xj08>
- Hobert S, Groth M, Nießner T, Wilhelmi L (2023) How Today's AI Content Generators Outperform Average Novice Students in Information Systems Exams. In: *Proceedings of the Americas Conference on Information Systems (AMCIS2023)*. Panama City, Panama
- Hu X, Tian Y, Nagato K, Nakao M, Liu A (2023) Opportunities and challenges of ChatGPT for design knowledge management. *Procedia CIRP* 119:21–28. <https://doi.org/10.1016/j.procir.2023.05.001>
- Huang J, Sciuchetti M (2023) ChatGPT Unveiled: Unleashing AI Magic in Online Shopping and Digital Marketing. *Atl Mark J* 12(2):1–7
- Huang K-H, Laban P, Fabbri AR, Choubey PK, Joty S, Xiong C et al (2023) Embrace Divergence for Richer Insights: A Multi-document Summarization Benchmark and a Case Study on Summarizing Diverse Information from News Articles. arXiv. <https://doi.org/10.48550/arXiv.2309.09369>
- Hutson J, Cotroneo P (2023) Generative AI tools in art education: Exploring prompt engineering and iterative processes for enhanced creativity. *Metaverse* 4(1):15–28. <https://doi.org/10.54517/m.v4i1.2164>
- Hutson J, Lang M (2023) Content creation or interpolation: AI generative digital art in the classroom. *Metaverse* 4(1):2–14. <https://doi.org/10.54517/m.v4i1.2158>
- Hutson J, Plate D (2023a) Enhancing Institutional Assessment and Reporting Through Conversational Technologies: Exploring the Potential of AI-Powered Tools and Natural Language Processing. *J Artif Intell Robotics* 1(1):11–22
- Hutson J, Plate D (2023b) Human-AI Collaboration for Smart Education: Reframing Applied Learning to Support Metacognition. In: Soofastaei, A (Hrsg) *Advanced Virtual Assistants - A Window to the Virtual Future*. IntechOpen. <https://doi.org/10.5772/intechopen.1001832>
- Hutson J, Plate D (2023c) Working With (Not Against) the Technology: GPT 3 and Artificial Intelligence (AI) in College Composition. *J Robotics Autom Res* 4(1):330–337
- Hutson J, Schnellmann A (2023) The Poetry of Prompts: The Collaborative Role of Generative Artificial Intelligence in the Creation of Poetry and the Anxiety of Machine Influence. *Glob J Comput Sci Technol D* 23(1):1–14. <https://doi.org/10.34257/GJCSTDVOL23IS1PG1>
- Intahachomphoo C, Gundersen OE (2020) Artificial Intelligence and Race: a Systematic Review. *Leg Inf Manag* 20(2):74–84. <https://doi.org/10.1017/S1472669620000183>
- Iqbal F, Samsom F, Kamoun F, MacDermott Á (2023) When ChatGPT goes rogue: exploring the potential cybersecurity threats of AI-powered conversational chatbots. *Front Commun Networks*. <https://doi.org/10.3389/frcmn.2023.1220243>
- Jha K, Doshi A, Patel P, Shah M (2019) A comprehensive review on automation in agriculture using artificial intelligence. *Artif Intell Agric* 2:1–12. <https://doi.org/10.1016/j.aiaa.2019.05.004>
- Juhasz L, Guan B (2023) Utilizing Large Language Models in geographic contexts - Experiences from the FIU GIS Center. <https://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=1104&context=gis>. Zugegriffen: 15. Feb. 2024

- Korzynski P, Mazurek G, Altmann A, Ejdy J, Kazlauskaitė R, Paliszewicz J et al (2023) Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Cent Eur Manag J* 31(1):3–13. <https://doi.org/10.1108/CEMJ-02-2023-0091>
- Kowalczyk P, Röder M, Thiesse F (2023) Nudging Creativity in Digital Marketing with Generative Artificial Intelligence: Opportunities and Limitations. In: *Proceedings of the European Conference on Information Systems (ECIS2023)*. Kristiansand, Norway
- Kumar D, AbuHashem Y, Durumeric Z (2023) Watch Your Language: Investigating Content Moderation with Large Language Models. *arXiv*. <https://doi.org/10.48550/arXiv.2309.14517>
- Lappalainen Y, Narayanan N (2023) Aisha: A Custom AI Library Chatbot Using the ChatGPT API. *J Web Librariansh* 17(3):37–58. <https://doi.org/10.1080/19322909.2023.2221477>
- Leadbetter L (2023) Maximizing Productivity and Quality in Senior Thesis Writing with Artificial Intelligence and Natural Language Processing Driven Tools. *CMC Senior Theses*. 3285. Claremont McKenna College. (Bachelor thesis)
- Lee D, Lee S, Lee H, Lee K, Lee H-J (2019) Resolution-Preserving Generative Adversarial Networks for Image Enhancement. *IEEE Access* 7:110344–110357. <https://doi.org/10.1109/ACCESS.2019.2934320>
- Lund B, Ting W (2023) Chatting about ChatGPT: how May AI and GPT impact academia and libraries? *Libr Hi Tech News* 40(3):26–29. <https://doi.org/10.1108/LHTN-01-2023-0009>
- Memmert L, Tavanapour N (2023) Towards Human-AI-Collaboration in Brainstorming: Empirical Insights into the Perception of Working with a Generative AI. In: *Proceedings of the European Conference on Information Systems (ECIS2023)*. Kristiansand, Norway
- Milmo D (2023) ChatGPT reaches 100 million users two months after launch. <https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>. Zugegriffen: 5. Nov. 2023
- Mingers J, Brocklesby J (1997) Multimethodology: Towards a framework for mixing methodologies. *Omega* 25(5):489–509. [https://doi.org/10.1016/S0305-0483\(97\)00018-2](https://doi.org/10.1016/S0305-0483(97)00018-2)
- Mukherjee A, Kamath NP (2023) Intelligent Auto-Learning Generative Knowledge Finder for Proactively Aiding a Virtual Agent. *Technical Disclosure Commons*. Defensive Publications Series
- Nica E, Sabie OM, Mascu S, Luțan AG (2022) Artificial Intelligence Decision-Making in Shopping Patterns: Consumer Values, Cognition, and Attitudes. *Econ Manag Financial Mark* 17(1):31–43. <https://doi.org/10.22381/emfm17120222>
- OpenAI (2022) Introducing ChatGPT. <https://openai.com/blog/chatgpt>. Zugegriffen: 5. Nov. 2023
- OpenAI (2023a) ChatGPT can now see, hear, and speak. <https://openai.com/blog/chatgpt-can-now-see-hear-and-speak>. Zugegriffen: 5. Nov. 2023
- OpenAI (2023b) GPT-4 API general availability and deprecation of older models in the Completions API. <https://openai.com/blog/gpt-4-api-general-availability>. Zugegriffen: 5. Nov. 2023
- Padtberg C (2023) Warum künstliche Intelligenz die größte Sorge beim Hollywood-Streik ist. <https://www.spiegel.de/kultur/hollywood-streik-es-geht-um-kuenstliche-intelligenz-a-f5429ce4-ef4c-4b74-a679-12c7c81d0b0f>. Zugegriffen: 5. Nov. 2023
- Panagopoulou F, Parpoula C, Karpouzis K (2023) Legal and ethical considerations regarding the use of ChatGPT in education. *arXiv*. <https://doi.org/10.48550/arXiv.2306.10037>
- Parker MJ, Anderson C, Stone C, Oh Y (2023) A Large Language Model Approach to Educational Survey Feedback Analysis. *arXiv*. <https://doi.org/10.48550/arXiv.2309.17447>
- Perkins M (2023) Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *J Univ Teach Learn Pract* 20(2). <https://doi.org/10.53761/1.20.02.07>
- Ploennigs J, Berger M (2023) AI art in architecture. *Ai Civ Eng* 2(8). <https://doi.org/10.1007/s43503-023-00018-y>
- Ray PP (2023) ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber Physical Syst* 3:121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Regalia J (2023) The Case for Iterative Legal Writing Practice with ChatGPT. https://scholarship.law.wm.edu/incorporating_chatgpt/schedule/fullschedule/2. Zugegriffen: 5. Nov. 2023
- Reiss MV (2023) Testing the Reliability of ChatGPT for Text Annotation and Classification: A Cautionary Remark. *arXiv*. <https://doi.org/10.48550/arXiv.2304.11085>
- Robinson D (2023) Evaluating the Potential of AI in Sports Consulting: Investigating ChatGPT-4's Ability to Consult an MLB Team. *IPHS 300: Artificial Intelligence for the Humanities: Text, Image, and Sound*. Paper 40.

- Santos RS, Qin L (2019) Risk Capital and Emerging Technologies: Innovation and Investment Patterns Based on Artificial Intelligence Patent Data Analysis. *J Risk Financial Manag* 12(4). <https://doi.org/10.3390/jrfm12040189>
- Siddiqui MF, Waqas AS, Ahmedh M, Jumani AK (2020) Face Detection and Recognition System for Enhancing Security Measures Using Artificial Intelligence System. *Indian J Sci Technol* 13(9):1057–1064. <https://doi.org/10.17485/ijst/2020/v013i09/149298>
- Sirithumgul P (2023) Unlocking the Potential of ChatGPT: A Grounded Theory Exploration of its Impact on the Business Landscape. In: *Proceedings of the Americas Conference on Information Systems (AMCIS2023)*. Panama City, Panama
- Soman S, Ranjani HG (2023) Observations on LLMs for Telecom Domain: Capabilities and Limitations. arXiv. <https://doi.org/10.48550/arXiv.2305.13102>
- Spataro J (2023) Introducing Microsoft 365 Copilot - your copilot for work. <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/>. Zugegriffen: 5. Nov. 2023
- Tang Y (2023) Understanding Societal Values of ChatGPT. Washington University, St. Louis <https://doi.org/10.7936/sbad-jm91> (Master thesis)
- Teubner T, Flath CM, Weinhardt C, van der Aalst W, Hinz O (2023) Welcome to the Era of ChatGPT et al. *Bus Inf Syst Eng* 65(2):95–101. <https://doi.org/10.1007/s12599-023-00795-x>
- Tlili A, Shehata B, Adarkwah MA, Bozkurt A, Hickey DT, Huang R et al (2023) What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learn Environ* 10. <https://doi.org/10.1186/s40561-023-00237-x>
- Torres JM (2023) Leveraging AI Platforms for Business Librarians and Information Professionals: Developing Pedagogical Strategies to Disseminate Information and Improve Learning Retention. Rutgers University. <https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1042&context=mbils>. Zugegriffen: 5. Nov. 2023
- Tsai C-H, Kadire S, Sreeramdas T, VanOrmer M, Thoene M, Hanson C et al (2023) Generating Personalized Pregnancy Nutrition Recommendations with GPT-Powered AI Chatbot. In: *Proceedings of the International Information Systems for Crisis Response and Management Conference (ISCRAM2023)*, S 263–271. Omaha, NB, US
- Vasconcelos MAR, dos Santos RP (2023) Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia J Math Sci Technol Educ* 19(7). <https://doi.org/10.29333/ejmste/13313>
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN et al (2017) Attention Is All You Need. arXiv. <https://doi.org/10.48550/arXiv.1706.03762>
- Wang Z, Zhang L, Cao C, Liu P (2023) The Effectiveness of Large Language Models (ChatGPT and CodeBERT) for Security-Oriented Code Analysis. arXiv. <https://doi.org/10.48550/arXiv.2307.12488>
- Waqas A, Bui MM, Glassy EF, El Naqa I, Borkowski P, Borkowski AA et al (2023) Revolutionizing Digital Pathology With the Power of Generative Artificial Intelligence and Foundation Models. *Lab Invest* 103(11). <https://doi.org/10.1016/j.labinv.2023.100255>
- Webster J, Watson RT (2002) Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly* 26(2):13–23
- Weingart P, Wambsganss T, Soellner M (2023) A Taxonomy for Deriving Business Insights from User-Generated Content. In: *Proceedings of the European Conference on Information Systems (ECIS2023)*. Kristiansand, Norway
- Weizenbaum J (1966) ELIZA—a computer program for the study of natural language communication between man and machine. *Commun ACM* 9(1):36–45. <https://doi.org/10.1145/365153.365168>
- Xiao Y, Watson M (2019) Guidance on Conducting a Systematic Literature Review. *J Plan Educ Res* 39(1):93–112. <https://doi.org/10.1177/0739456X17723971>
- Yadav D, Salmami S (2019) Deepfake: A Survey on Facial Forgery Technique Using Generative Adversarial Network. In: *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICCS2019)*, S 852–857. Madurai, India. <https://doi.org/10.1109/ICCS45141.2019.9065881>
- Yun G, Cho K, Jeong Y, Nam T-J (2022) Ideasquares: Utilizing generative text as a source of design inspiration. In: *Proceedings of the Conference of Design Research Society (DRS2022)*. Bilbao, Spain. <https://doi.org/10.21606/drs.2022.484>
- Zeng F (2023) Evaluating the Problem Solving Abilities of ChatGPT. Washington University, St. Louis <https://doi.org/10.7936/vz0-dr08> (Master thesis)

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.