

**Douglas MacKevett<sup>1</sup>, Patricia Feubli<sup>2</sup> & Vinzenz Rast<sup>3</sup> (Luzern)**

# **Digitale Leistungsnachweise im Zeitalter von generativer künstlicher Intelligenz**

## **Zusammenfassung**

Generative künstliche Intelligenz (GKI) könnte bald fähig sein, Prüfungen auf allen Taxonomiestufen erfolgreich zu meistern. Diese These wird gestützt durch eine empirische Untersuchung, die in der Prüfungspraxis an einer Fachhochschule in den Bereichen Wirtschaft und Informatik durchgeführt wurde. Aktuelle Handlungsempfehlungen werden betrachtet und es wird anhand von drei Beispielen gezeigt, wie Hochschulen mit Prüfungen unter Einsatz von GKI umgehen könnten. Abschließend wird dargelegt, wie eine künftige, KI-unterstützte Ausbildung zu gestalten wäre, um den Herausforderungen und Chancen dieser Technologie gerecht zu werden.

## **Schlüsselwörter**

generative künstliche Intelligenz, Leistungsnachweise, Kompetenzorientierung, Prüfungspraxis, Hochschulentwicklung

- 
- 1 Corresponding author; Hochschule Luzern; [douglas.mackevett@hslu.ch](mailto:douglas.mackevett@hslu.ch); <https://www.hslu.ch/de-ch/wirtschaft/ueber-uns/personensuche/person-detail-site/?pid=178>; ORCID 0000-0002-7378-1982
  - 2 Hochschule Luzern; [patricia.feubli@hslu.ch](mailto:patricia.feubli@hslu.ch); <https://www.hslu.ch/de-ch/wirtschaft/ueber-uns/personensuche/person-detail-site/?pid=4645>; ORCID 0000-0002-7522-176X
  - 3 Hochschule Luzern; [vinzenz.rast@hslu.ch](mailto:vinzenz.rast@hslu.ch); <https://www.hslu.ch/de-ch/wirtschaft/ueber-uns/personensuche/person-detail-site/?pid=177>; ORCID 0009-0007-7585-4595

Dieser Beitrag wurde unter der Creative-Commons-Lizenz 4.0 Attribution (BY) veröffentlicht.

<https://doi.org/10.21240/zfhe/19-2/09>

## **Digital exams in the age of generative artificial intelligence**

### **Abstract**

Generative artificial intelligence (generative AI) could soon be capable of successfully mastering all types of examinations at all taxonomy levels. To test this thesis, an empirical study of examination practice at a university of applied sciences in the fields of economics and computer science was conducted. This paper presents current action recommendations and uses three examples to demonstrate how universities could deal with examinations using generative AI. It then describes how a future AI-supported education could be designed to meet the challenges and opportunities of this technology.

### **Keywords**

generative AI, assessment, competence-based assessment, exam practice, higher-education development

# 1 Künstliche Intelligenz schreibt Klausuren

Die rasante Entwicklung der digitalen Technologien, insbesondere der generativen künstlichen Intelligenz (GKI), stellt die traditionelle Prüfungspraxis an Hochschulen vor erhebliche Herausforderungen. GKI ist ein Bereich der künstlichen Intelligenz, der Algorithmen und Modelle entwickelt, die in der Lage sind, neue und originalgetreue Inhalte wie Texte, Bilder, Musik oder Videos zu erzeugen, die auf bereits vorhandenen Daten basieren. Entsprechend ist GKI auch in der Lage, für eine Vielzahl von bestehenden digitalen Prüfungen Lösungen zu liefern. Dadurch verschärft sich die Frage nach der Validität und Reliabilität bestehender Prüfungsverfahren. Eine kompetenzorientierte Leistungsüberprüfung hingegen könnte GKI zu einem Hilfsmittel bei der Lösungserstellung machen, was aber eine grundlegende Veränderung der Prüfungskultur erfordern würde.

Die Forderung nach kompetenzorientiertem Prüfen an Hochschulen ist nicht neu (Long, Dowsing & Craven, 2003; Mouthaan, Olthuis & Vos, 2003; Ilahi, Belcadhi & Braham, 2014). Solche Leistungsnachweise konzentrieren sich auf die Anwendung von Fähigkeiten und Wissen. Aufgaben werden in einen praktischen Kontext eingebettet, in dem Studierende Probleme lösen, die kritisches Denken und Verantwortung für das Ergebnis erfordern. Kompetenzen sind „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“ (Weinert, 2002, S. 27f.).

Noch wird verbreitet an rein wissensbasierten Prüfungsformen festgehalten, um die Reliabilität einfach sicherzustellen (MacKevett & Gutmann, 2023). Mit der Verfügbarkeit der GKI werden aber solche Prüfungsformen insbesondere bei digitalen Prüfungsformaten obsolet. Die aktuelle Prüfungspraxis funktioniert nicht mehr (Buck et al., 2023).

In Anbetracht dieser Entwicklungen steht die Hochschulwelt vor einem entscheidenden Wendepunkt: der Einführung von Kompetenznachweisen, die nicht nur das Wis-

sen, sondern auch die motivationalen, volitionalen und sozialen Fähigkeiten zur verantwortungsvollen und erfolgreichen Problemlösung prüfen. In diesem Entwicklungsbeitrag zeigen wir Möglichkeiten auf, warum diese Wende notwendig ist und wie sie in der Prüfungspraxis zu schaffen sein könnte.

## 2 GKI in der Hochschulausbildung

Bereits heute liefert GKI gute Ergebnisse bei den Fähigkeiten, die aktuell an den meisten Hochschulen bei der Leistungsüberprüfung im Mittelpunkt stehen: Schreiben, Textanalyse, Kreieren und Problemlösung in einem breiten Spektrum von Disziplinen, von den Naturwissenschaften bis zu den Sozialwissenschaften und der Kunst.

### 2.1 GKI im Unterricht

GKI-Applikationen erreichen immer mehr Einsatzgebiete in der hochschulischen Praxis. Öffentlich zugängliche Forschungspapiere können nach Schlüsselbegriffen durchsucht, relevante Beiträge herausgefiltert, zusammengefasst und schnell gesichtet werden. Große Datenmengen können aufbereitet, analysiert und visualisiert werden, samt zugehörigem Python-Code für die detaillierte Dokumentation. Software beschreibt Bilder und beantwortet Fragen dazu. Tools erlauben die Nutzung von GKI-Modellen auf dem eigenen Computer, ohne dass sensitive Daten preisgegeben werden müssen. Vortrainierte Large Language Models (LLM) stehen für das Entwickeln fachspezifischer Anwendungen zur Verfügung. Online-Lernplattformen integrieren GKI in ihre Angebote. Zu den bekanntesten Beispielen zählt die Khan Academy „Khanmigo“ (Khan, 2023), die ein bekanntes und leistungsfähiges LLM als individuellen Tutor und Teaching Assistant einsetzt, was die Effektivität des Unterrichts erheblich zu steigern vermögen könnte, wie Bloom (1984) ausführt. Chatbots beantworten Fragen und können die Interaktion im Lernprozess verstärken (Dutta, 2017). Sie erhöhen das Engagement und unterstützen rund um die Uhr (Gimpel et al., 2023; Tu et al., 2023; Jungherr, 2023).

## 2.2 GKI für Leistungsnachweise

ChatGPT heißt das bekannte GKI-Sprachmodell des US-amerikanischen Softwareunternehmens OpenAI. Es kann scheinbar von Menschen verfasste Texte generieren und in natürlicher Sprache auf Fragen und Anweisungen reagieren. Open AI dokumentiert selbst die Leistung der verschiedenen ChatGPT-Versionen in gängigen berufsorientierten und akademischen Leistungsnachweisen in den USA (vgl. Abbildung 1). ChatGPT 3.5 wurde im November 2022 veröffentlicht, und ChatGPT 4 wurde im März 2023 eingeführt.

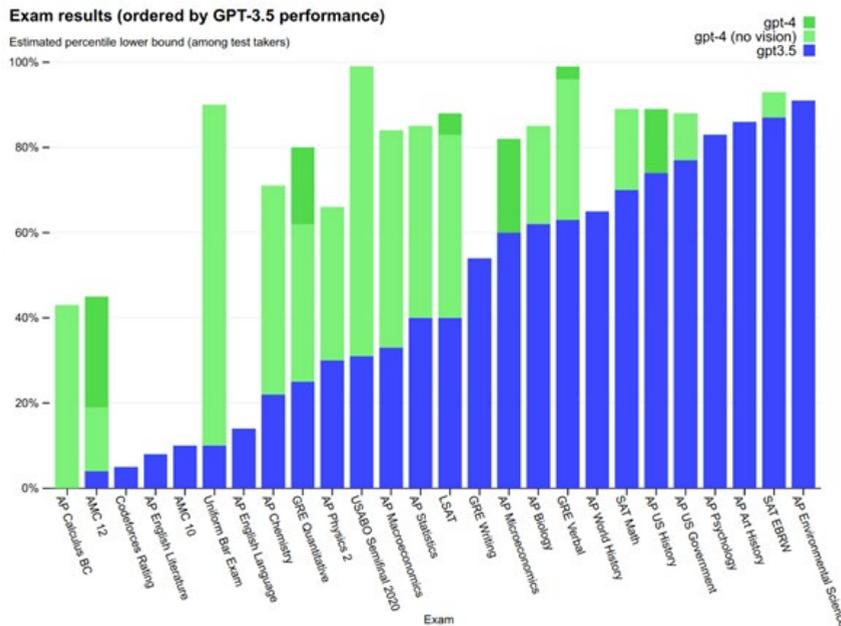


Abb. 1: Prüfungsergebnisse unter Einsatz von ChatGPT 3.5 und 4.0 (Open AI, 2024)

Die von Open AI getesteten Leistungsnachweise umfassen vor allem Übertrittsprüfungen für High-School-Schüler:innen (z.B. Advanced Placement Test, Scholarly Aptitude Tests usw.) vor Studienstart sowie die Zulassungsprüfung für Jurist:innen (Bar) und Themen wie Mathematik, Statistik, Medizin, Geschichte, Biologie, Chemie, Physik, Umweltwissenschaften, Ökonomie, Literatur und Psychologie.

Aus den Resultaten schließt Open AI: “GPT-4 exhibits human-level performance on the majority of these professional and academic exams. Notably, it passes a simulated version of the Uniform Bar Examination with a score in the top 10% of test takers. [...] GPT-4 considerably outperforms existing language models [...]“ (Open AI, 2024). Neue ChatGPT-Versionen erreichen also mittlerweile mehrheitlich die Prüfungsleistung von Menschen.

### **3 GKI in Leistungsnachweisen einer Fachhochschule**

Um zu verstehen, wie stark die Prüfungspraxis an der Hochschule Luzern betroffen ist, haben wir ein „Exam Busting“ mit ChatGPT durchgeführt. Zu diesem Zweck wurden reale Prüfungen aus zwei Departementen jeweils auch von verschiedenen Versionen von ChatGPT gelöst. Im Anschluss wurden die Antworten von den Dozierenden geprüft und benotet oder mit fail/pass bewertet. Die Eingabeaufforderungen oder Fragestellungen, die sogenannten „Prompts“, die ChatGPT zur Bearbeitung übergeben worden sind, haben meist den Fragestellungen in den Prüfungen entsprochen.

#### **3.1 Datensatz von 14 Leistungsnachweisen**

Das „Exam-Busting“ fand zwischen Januar und April 2023 zusammen mit Dozierenden der Departemente Informatik und Wirtschaft der Hochschule Luzern statt. Insgesamt wurden 14 Leistungsnachweise in Modulen von Bachelor- und Master-

studiengängen überprüft. Thematisch reichen die Module von der Unternehmensethik bis zum Natural Language Processing und die Art der Leistungsnachweise von Multiple-Choice-Tests bis zu Fallbearbeitungen (vgl. Tabelle Übersicht „Exam Busting“ unter 10.5281/zenodo.13210101).

### **3.2 Leistung Anfang 2023 zum Teil noch nicht genügend**

In unserem „Exam-Busting“ kamen ausschließlich ChatGPT 3.5 und 4.0 zum Einsatz. Es gibt jedoch noch zahlreiche weitere GKI-Anwendungen, die in Leistungsnachweisen eine Rolle spielen und die Prüfungspraxis infrage stellen könnten (z.B. Google Gemini, Microsoft Copilot, Claude).

Die Leistungen der verschiedenen ChatGPT-Versionen bis April 2023 reichten in den von uns untersuchten Leistungsnachweisen von „nicht bestanden“ bis hin zu „sehr gut“. Im Durchschnitt waren die Leistungen jedoch eher mangelhaft. Festzustellen war aber, dass die Art und Weise der Fragestellungen einen massiven Einfluss auf die Leistung von ChatGPT hat. Ergänzen ausgeklügelte Prompts die eigentlichen Prüfungsfragen, um das Verhalten und den Kontext solcher Fragen dem KI-Modell vorzugeben, ist die Leistung erheblich besser. Solche „System Messages“ geben Anweisungen vor und definieren die Rahmenbedingungen, wie das Modell auf nachfolgende Prompts reagieren soll.

Aufgrund der raschen Entwicklung der GKI dürften einige Resultate des „Exam Busting“ in der Zwischenzeit bereits überholt sein (OpenAI, 2024). Im Spätherbst 2023 hatte ChatGPT noch Mühe mit der Interpretation eines Bildes und der korrekten Beantwortung der Fragen. Die aktuelle Version von ChatGPT, Version 4o („o“ für „omni“), kann Bilder „sehen“. Die Leistungen in Prüfungsaufgaben mit Bildern dürften sich also erheblich verbessert haben.

### 3.3 Mit guten Prompts höhere Taxonomiestufen erreichbar

Im Rahmen eines Leistungsnachweises im Modul „Natural Language Processing“ am Departement Informatik wurde zusätzlich geprüft, wie gut ChatGPT 3.5 mit den verschiedenen Taxonomiestufen von Bloom (1973) zurechtkommt. Das Bloom'sche Taxonomiemodell bildet in der didaktischen Ausbildung der Dozierenden an der Hochschule Luzern einen zentralen Orientierungspunkt und wird für die Formulierung der Lernziele und die Erstellung der Leistungsüberprüfung genutzt.

Tabelle 1 zeigt die Leistung von ChatGPT 3.5. Die dritte Spalte nennt die Fehlerquote in Prozentwerten, also wie groß der Anteil von Fehlern pro Taxonomiestufe, geordnet nach Komplexitätsgraden von unten nach oben, ausgefallen ist.

<b>Taxonomiestufe nach Bloom</b>	<b>Beschreibung</b>	<b>Fehlerquote</b>
Create	Teile zu einem neuen Ganzen bilden	24%
Evaluate	Beurteilung des Wertes von Informationen oder Ideen	-
Analyze	Zerlegen von Informationen in ihre Bestandteile	13%
Apply	Anwendung von Fakten, Regeln, Konzepten und Ideen	33%
Understand	Die Bedeutung der Fakten verstehen	10%
Remember	Erkennen und Erinnern von Fakten	0%

Tab. 1: Fehlerquote von ChatGPT nach Bloom'schen Taxonomiestufen (Marfurt, 2023)

In dieser Versuchsanordnung hat ChatGPT beim Einsatz von ausgefeilten Prompts noch einmal wesentlich besser abgeschnitten. Diese erste Untersuchung lässt also vermuten, dass ChatGPT auch mit höheren Taxonomiestufen gut zurechtkommt.

Selbst auf der höchsten Taxonomiestufe „Create“ erreicht ChatGPT in der Aufgabe im Modul „Natural Language Processing“ in unserem Test 76% der Punkte. Da die Untersuchung jedoch nur auf wenigen Prüfungsfragen beruht, ist eine Verallgemeinerung nicht zulässig.

## **4 Kompetenzorientiert prüfen – exemplarisch**

Das „Exam Busting“ dokumentiert, dass sich GKI schon jetzt als leistungsfähiges Hilfsmittel für viele Formen digitaler Leistungsnachweise abzeichnet. Für Studierende ist es daher naheliegend, GKI bei Leistungsnachweisen zu verwenden. Die entsprechenden Lösungen belegen dann aber eher die Fähigkeiten der GKI als die der Studierenden oder – im besseren Fall – deren Fähigkeiten beim Nutzen von GKI.

Die bisherige Praxis der Leistungsnachweise ist somit überholt, auch weil sich GKI nicht mehr sinnvoll von digitalen Leistungsnachweisen ausschließen lässt. Geeignete digitale Leistungsnachweise sind zu entwickeln, die GKI zulassen, möglichst alle Taxonomiestufen abdecken und die Kompetenzen der Studierenden in den Mittelpunkt rücken.

Hanke (2023) schlägt ein Entscheidungsmodell vor, wie GKI bei der Leistungsüberprüfung berücksichtigt werden kann (vgl. Abb. 2).

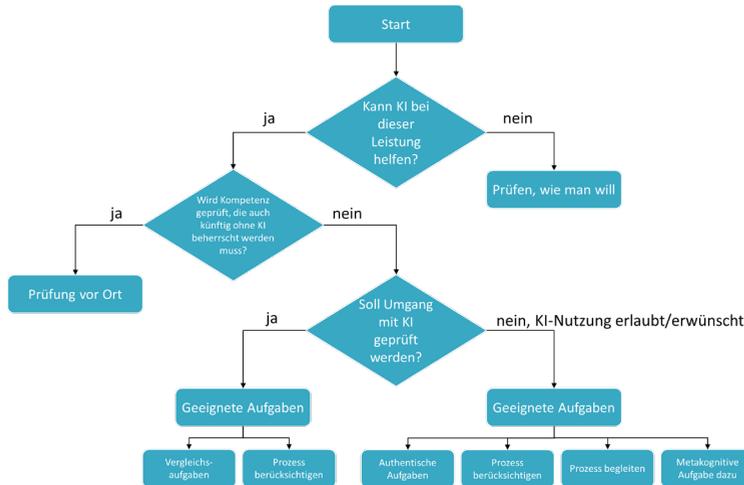


Abb. 2: Entscheidungsmodell für den Einsatz von GKI bei Prüfungen (Hanke, 2023)

Bei digitalen Kompetenznachweisen dürfte insbesondere jenes Szenario häufig eintreten, in dem GKI als Hilfsmittel zugelassen ist, der Nachweis jedoch nicht den Umgang mit GKI prüfen soll. Vielmehr soll er die fachbezogenen Kompetenzen der Studierenden prüfen (rechter, unterer Strang in Abb. 2).

Hierzu hat Hanke (2023) Aufgabentypen zusammengestellt, bei denen der Einsatz von GKI erlaubt oder sogar erwünscht ist: authentische Aufgaben, Prozessdokumentationen, Prozessbegleitung, metakognitive Aufgaben. Diese Aufgabentypen berücksichtigen alle Taxonomiestufen des Bloom'schen Modells mit einem Schwerpunkt auf den Stufen „Apply“, „Analyze“, „Evaluate“ und „Create“.

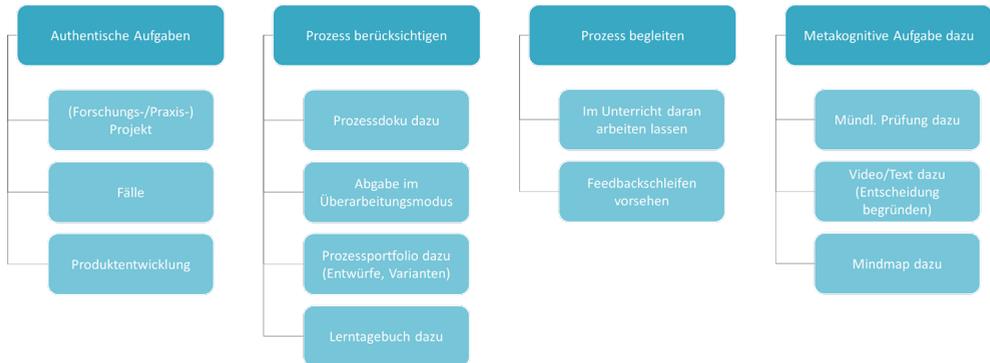


Abb. 3: Aufgabentypen, die auch mit Einsatz von GKI sinnvoll sind (Hanke, 2023)

GKI kann inzwischen Powerpoint-Präsentationen, Videos, Mindmaps usw. generieren. Darum dienen solche Anwendungen allenfalls zur Dokumentation eines Lernprozesses. Ob die Studierenden über eine Kompetenz verfügen, muss nach wie vor durch einen Nachweis in Echtzeit erbracht werden.

Im Folgenden werden nun drei Beispiele der Hochschule Luzern betrachtet, wie digitale Kompetenznachweise umgesetzt werden können. Die drei Beispiele decken die von Weinert (2002), Walzik (2012) und Kaiser (2019) definierten Kompetenzbereiche ab:

1. *Kognitive Kompetenz* ist die Fähigkeit, Informationen zu verarbeiten, zu verstehen und anzuwenden. Sie umfasst mentale Prozesse wie Denken, Wahrnehmen, Erinnern und Problemlösen, um Konzepte zu verstehen, Entscheidungen zu treffen und aus Erfahrungen zu lernen.
2. *Anwendungskompetenz* bezieht sich auf die Fähigkeit, erlerntes Wissen und Fähigkeiten in praktischen oder realen Situationen anzuwenden. Theoretisches Wissen wird in praktischen Handlungen umgesetzt, um spezifische Aufgaben oder Probleme zu lösen.

3. *Situative Kompetenz* ist die Fähigkeit, sich an verschiedene Situationen anzupassen und in diesen angemessen zu handeln. Sie umfasst das Verständnis für den Kontext einer Situation und die Anwendung von Wissen und Fähigkeiten so, dass sie den spezifischen Anforderungen dieser Situation gerecht wird.

#### **4.1 Beispiel 1: Concept-Map zum „Luxury Marketing“**

Das Beispiel zeigt einen Midterm-Leistungsnachweis aus dem Wahlmodul „Luxury Marketing and Digital Challenges“ (Masterausbildung, Betreuungsverhältnis 1:20), das im Frühjahr 2022 unterrichtet wurde.

Die Studierenden erstellten im Team eine Concept Map auf einem „Miro Board“. Beim „Miro Board“ handelt es sich um ein digitales Kollaborationstool, das es Teams ermöglicht, in Echtzeit gemeinsam an visuellen Projekten zu arbeiten.

Die während des Kurses präsentierten Themen und Modelle wurden darauf abgebildet und mit persönlichen Erkenntnissen und Reflexionen angereichert. Die Concept Map dokumentierte drei Teilaufträge:

1. auf Theorien und Ideen aus dem Kursmaterial verweisen
2. im Kurs präsentierte Anwendungsfälle analysieren
3. eine mögliche Abschlussarbeit skizzieren

Ziel des Nachweises war, sowohl das Wissen der Studierenden über Luxusmarketing als auch ihr Verständnis für die Herausforderungen der Digitalisierung im Luxusmarkt zu überprüfen. Zudem sollten sie ein mögliches Thema für ihre Abschlussarbeit vorstellen. Da es sich um einen einwöchigen Intensivkurs handelte, erhielten die Studierenden die Aufgabe zu Beginn der Woche. Sie mussten die Concept Map am

Ende der Woche digital einreichen. Die Ergebnisse wurden anschließend live präsentiert. Die Studierenden erhielten sowohl eine summative Note für die Präsentation als auch ein formatives Feedback zu ihrer Abschlussarbeit.

Bei diesem ersten Beispiel geht es vor allem darum, die kognitive Kompetenz zu prüfen. Zwei Aufgabentypen aus der Systematik von Hanke (2023) werden kombiniert: Prozessbegleitung im Unterricht und metakognitive Aufgaben in der Präsentation und Dokumentation der Reflexion auf der Concept Map. Geprüft wird also auf den Taxonomiestufen „Remember“, „Understand“ und vor allem „Apply“ und „Analyze“.

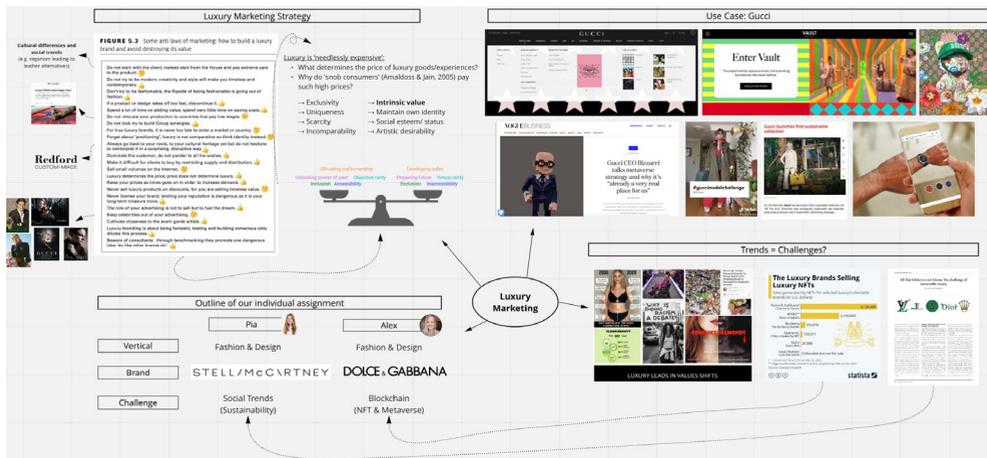


Abb. 4: Beispiel einer Concept Map (mit freundlicher Genehmigung von Alexandra Flimm und Pia Rotkamp, CBS International Business School Cologne, 2022).

## 4.2 Beispiel 2: Projekt zu „Sustainable Re-Commerce“

Projektbasiertes Lernen kommt an der Hochschule Luzern in verschiedenen Studiengängen zum Einsatz. Studierende arbeiten über längere Zeit an einem Projekt, das auf realen Fragen, Problemen oder Herausforderungen basiert.

Hattie und Yates (2013) zeigen, dass projektbasiertes Lernen eine moderate bis hohe Effektgröße haben kann: Es hat also das Potenzial, das Lernen signifikant zu verbessern. Allerdings hängt der Effekt stark von der Art und Weise ab, wie das projektbasierte Lernen gestaltet und umgesetzt wird.

Im Modul „Online Business Strategy“ (Masterausbildung; Betreuungsverhältnis 1:30) werden zum Thema „Sustainable Re-Commerce“ während des ganzen Semesters forschungsbasierte und praxistaugliche Aufgaben zusammen mit Auftraggeber:innen aus der Industrie gelöst. Die Aufgabe umfasst fünf teils mündlich, teils schriftlich realisierte Teilaufträge:

1. Pitch zu Idee, Absicht und Vorhaben
2. Proposal zu Disposition und Entwurf mit Beispielen
3. Defense, um die kognitive Kompetenz unter Beweis zu stellen
4. Project mit Dokumentation von Inhalt und Lernprozess
5. Proof of Concept zur Validierung des Projektes mit Empirie

Durch diesen strukturierten Ansatz können die Studierenden nicht nur fachliche Kompetenzen, sondern auch wichtige Soft Skills wie Präsentationsfähigkeiten, kritisches Denken und Problemlösungsfähigkeiten entwickeln. Die Kombination aus formativen und summativen Assessments ermöglicht eine kontinuierliche Reflexion und die Anpassung des Lernprozesses.

Dieses Beispiel rückt die Anwendungskompetenz in den Mittelpunkt der Leistungsüberprüfung und deckt sämtliche Aufgabentypen ab, die Hanke (2023) zur Überprüfung der Kompetenzen mit GKI als zugelassenes Hilfsmittel vorschlägt. Entsprechend werden auch alle Taxonomiestufen nach Bloom berücksichtigt.

### **4.3 Beispiel 3: Content-Kampagne zu eigenem Thema**

Das dritte Beispiel beschreibt einen Leistungsnachweis, der im Modul „Web Literacy“ (Masterausbildung: Betreuungsverhältnis 1:55) verlangt wird. Die Studierenden erstellen und verwalten eine Content-Kampagne zu einem Thema ihrer Wahl. Begleitet werden sie von Expert:innen in Content-Erstellung, datenbasiertem Marketing und Video für soziale Medien. Dokumentiert wird das Ergebnis in einem Portfolio mit

1. auf der Website der Hochschule veröffentlichten Kampagnen-Beiträgen inkl. Videos
2. einem Poster mit konzeptuellen Überlegungen sowie Datenauswertungen
3. einer Sprachaufnahme mit Erläuterungen zu Strategie, Umsetzung und Anpassungen

In der Mitte des Semesters und somit der Kampagnen-Führung präsentieren die Studierenden den aktuellen Stand der Kampagne und das provisorische Poster. Sie erhalten von den Dozierenden Feedback zu allen Aspekten der Kampagne.

GKI ist als Hilfsmittel für die Konzeption sowie die Inhaltserstellung und -optimierung erlaubt und erwünscht. Drei Anforderungen stellen jedoch sicher, dass die Anwendungs- und situativen Kompetenzen der Studierenden geprüft werden können. Zum einen wird mit der Erstellung des Posters der Prozess dokumentiert und berücksichtigt. Zum anderen wird der Prozess begleitet, weil die Studierenden in der Mitte des Semesters den aktuellen Stand ihrer Kampagne mündlich präsentieren und eine Feedbackschleife durchlaufen. Schließlich bewegen sich die Studierenden mit der

Sprachaufnahme und den darin enthaltenen Erläuterungen und Begründungen auf der metakognitiven Ebene (vgl. Hanke, 2023). Auch hier wird auf allen Taxonomie-stufen geprüft.



Abb. 5: Beispiel eines Kampagnen-Beitrags (mit freundlicher Genehmigung von Joel Koch, Hochschule Luzern, 2022).

## 5 Hochschule der Zukunft mit GKI gestalten

Aktuell befinden sich viele Lehrende und Lernende noch in einer Lernumgebung, die stark von Learning-Management-Systemen, Präsentationen sowie Talking-Head-Videomaterial und den entsprechenden Leistungsüberprüfungen geprägt ist. Diese Lehrmethoden sind zwar bewährt, jedoch wenig interaktiv und kaum angepasst an die individuellen Bedürfnisse der Studierenden.

Aus Kapazitätsgründen haben zudem viele Hochschulen seit Jahren ihre Prüfungsformate vereinfacht und automatisiert. Vielerorts werden Multiple-Choice-Prüfungen eingesetzt, die meist nur auf tiefen Taxonomiestufen Leistungen und kaum Kompetenzen überprüfen lassen. Der Ersatz solcher Prüfungsformate durch mündliche Prüfungsformate kommt wegen der knappen Ressourcen kaum infrage.

Problematisch ist die unreflektierte Digitalisierung analoger Prüfungsprozesse. Der Versuch, traditionelle Prüfungen mit Stift und Papier, Aufsicht und Maßnahmen zur Verhinderung von Betrug in digitale Formate zu übertragen, führt zu nicht-validen Prüfungen. Sie spiegeln nicht die realen Anforderungen und Fähigkeiten wider, die in der heutigen Welt benötigt werden, und sie sind eben nicht kompetenzorientiert.

Unsere Erkenntnisse aus dem „Exam Busting“ zeigen, dass herkömmliche digitale Leistungsüberprüfungen nicht mehr sinnvoll sind bzw. sein werden, auch weil der Einsatz von GKI kaum verhindert werden kann. Unsere drei Beispiele illustrieren aber, wie eine kompetenzorientierte Leistungsüberprüfung erfolgreich eingesetzt und wie GKI dafür auch genutzt werden könnte. Daraus lassen sich drei Anregungen ableiten, wie GKI in der Hochschule der Zukunft in die Ausbildung integriert werden könnte.

## **5.1 Anregung 1: GKI im Unterricht, in der Forschung und in Kompetenznachweisen einsetzen**

Auch wenn ethische und rechtliche Fragen noch ungeklärt sind, sollte der verantwortungsvolle Einsatz von GKI deswegen nicht verhindert werden. Sie wird integriert in die Lehre: GKI und deren themenspezifische Anwendung werden gezeigt und geübt. Der Einsatz von GKI in der Forschung macht den Prozess nicht nur effizienter, sondern erlaubt den Forschenden, auf dem aktuellen Stand zu bleiben. Neue Bewertungskriterien und Prüfungsformate sind zu entwickeln, die GKI als Hilfsmittel und Gestaltungselement berücksichtigen.

GKI kann auch für die Gestaltung von Leistungsnachweisen eingesetzt werden, die sich exemplarisch durch folgende Ideenskizzen veranschaulichen lassen:

- KI-basierte Simulationen und virtuelle Umgebungen erlauben es, Problemlösungs- und Entscheidungsfähigkeiten der Studierenden in realistischen Szenarien zu beurteilen.

- KI-basierte Beurteilungssysteme lassen es zu, den Schwierigkeitsgrad der Fragen an die Leistung der Studierenden anzupassen und so eine individuellere Bewertung zu ermöglichen.
- KI-gestützte Lernanalysen verfolgen die Fortschritte der Studierenden beim Erreichen bestimmter Lernergebnisse, um rechtzeitig gezielte Unterstützung anbieten zu können.

## **5.2 Anregung 2: Kompetenznachweise als „Assessment Center“ denken**

Mit Anregung 2 wird empfohlen, nicht nur Inhalt und Form, sondern auch Zahl und Funktion von Leistungsnachweisen zu überdenken. Die Handlungsempfehlungen der Hochschule RheinMain (Buck et al., 2023) bieten einen groben Überblick, wie eine Prüfungspraxis im Zeitalter von GKI umgebaut werden könnte:

1. Reduktion der Gewichtung von Abschlussarbeiten: ECTS-Punkte werden vom Zeitaufwand entkoppelt.
2. Neue formale und inhaltliche Anforderungen: Textstruktur, sprachliche Korrektheit usw. sind mit GKI gut umsetzbar. Mangelnde Leistungen in diesem Bereich werden nicht geduldet.
3. Ausweitung mündlicher Anteile: Gespräche über konzeptionelle Arbeiten ersetzen schriftliche Korrekturen von Vorstudien.
4. Prozessbegleitung: Der Lernprozess bei Abschlussarbeiten und Leistungsnachweisen tritt in den Fokus. Während der Prozessbegleitung wird die Leistung bewertet.
5. Praktische Prüfungen: Beim projektbasierten Lernen ergänzen sich Studierende, Dozierende und GKI.

Diese hier verkürzt skizzierten Handlungsempfehlungen lassen sich nicht nur in den Disziplinen „Informatik“ und „Betriebswirtschaft“ anwenden, auch wenn diese im Fokus des vorliegenden Beitrags liegen. Sicher sind aber immer die personellen Kapazitäten im Auge zu behalten.

### **5.3 Anregung 3: Modulprüfungsorientierung durch Kompetenzorientierung ersetzen**

Anregung 3 schließlich geht noch einen Schritt weiter. Mit konsequenter Kompetenzorientierung werden die Curricula neu gedacht und Leistungsüberprüfungen individualisiert.

Anstelle standardisierter Stundenpläne und Modulendprüfungen treten Eintritts-Assessments und individualisierte Entwicklungspläne. Für jeden Studiengang werden Kompetenzen definiert, die bei Abschluss des Studiums zu beherrschen sind (Ausgangskompetenzen). Modulspezifische Leistungsnachweise und Testate finden zu Beginn einer Lerneinheit oder gar des Studiums statt. Sie dienen der Bestimmung der Eingangskompetenzen und dem Erkennen von individuellen Lücken der Studierenden. Die Lerneinheiten dienen dazu, diese Lücken entwicklungsorientiert zu schließen. Besucht werden nur diejenigen Module, die für die Erreichung der Ausgangskompetenzen nötig sind, womit ein standardisierter Stundenplan wegfällt.

Wie ein solcher Studiengang aussehen könnte, zeigt die Technische Hochschule Köln mit dem Bachelor-Studiengang „Product Engineering & Context“ (vgl. Abb. 6), in dem rein projektbasiert gelehrt und gelernt wird. Das ganze Semester dient als „Abschlussarbeit“ und nicht die einzelnen Module, die nur zu einer Zerstückelung des Lernens führen (Wampfler, 2023).

Es ist an der Zeit, das Augenmerk auf das individuelle Lernen zu verlagern und eine kompetenzorientierte Prüfungskultur zu schaffen. Wir sollten endlich „überprüfen, was wichtig ist, nicht das, was leicht zu überprüfen ist!“ (Stern, 2010).

### M4 | Modellstudiengang: Product Engineering & Context

➤ Der Studiengang wurde von einem Expert\*innen-Team aus Maschinenbau, Elektrotechnik, Informatik, Design und Betriebswirtschaftslehre entwickelt und im Wintersemester 2022/23 gestartet.

	Groundwork			Consolidation		Synthesis	
	1	2	3	4	5	6	7
<b>Future Engineering</b>	Engineering Office I Case Study zur Entwicklung eines digital vernetzten mechatronischen Produkts • System Engineering • Project Management	Engineering Office II Case Study zur Entwicklung eines digital vernetzten mechatronischen Produkts • System Engineering • Project Management	Engineering Office III Case Study zur Entwicklung eines digital vernetzten mechatronischen Produkts • System Engineering • Innovationen • Change Management	Consolidation Projektbasiertes Arbeiten in den domänenspezifischen Vertiefungsrichtungen (z. B. Smart Systems)		Engineering Lab Projektbasiertes Arbeiten mit Industriepartnern	
<b>Skills &amp; Principles</b>	Anwendungsorientierte Inhalte der Disziplinen • Mathematik • Werkstoffkunde • Informatik	Anwendungsorientierte Inhalte der Disziplinen • Mechanik • Informatik • Werkstoffkunde • Elektrotechnik • Fertigungstechnik	Anwendungsorientierte Inhalte der Disziplinen • Data Science • Betriebslehre • Thermodynamik	Skills & Principles (Smart Systems) Professionelle Vertiefungskurse und übergreifende Pflichtkurse  Production Systems   Social Robotics   Smart Automation • Industrial Eng.   • MM Interact.   • Automatisierung • Messtechnik und Logistik   • Sensorik und Vernetzung   • Mechatronik & Signalverarbeitung		Systemische und disziplinspezifische Fertigkeiten • Forschungsmethoden	
<b>Reflection</b>	• Reflecting Journal • Expert Talks • Mentoring	• Reflecting Journal • Expert Talks • Mentoring	• Reflecting Journal • Expert Talks • Mentoring	Reflection IV • Reflecting Journal • Hochschulweite interdisziplinäre Projektberichte, Wettbewerbskurse in Projektarbeit		Reflection V • Reflecting Journal • Exkursions-, Portfolio / Bewerbungsportfolio im Projektbereich	
							Proposal • Vorstudie  Thesis   Colloquium   Symposium (Paper & Conference)

Prof. Dr. Sylvia Heuchemer | Vizepräsidentin für Lehre und Studium  
08.03.2023

Seite: 18

Technology  
Arts Sciences  
TH Köln

Abb. 6: Modellstudiengang „Product Engineering & Context“ an der Technische Hochschule Köln (TH Köln, online)

## 6 Die Zeit drängt

GKI ist nun Teil der digitalen Welt. Sie zu ignorieren, wird nicht mehr möglich sein, wie wir nachgewiesen haben mit unserem „Exam Busting“ unter Einsatz einer GKI-Software, die mittlerweile schon wesentlich leistungsfähiger geworden ist. Wir haben aber gezeigt, dass mit einer Neuausrichtung der Leistungsnachweise auf die Überprüfung von Kompetenzen zeitgemäße Prüfungsformate denkbar sind. Impulse könnten eine konsequente Integration von GKI in die Lehre und Forschung, aber auch ein Wandel der Einstellung bei Dozierenden und Studierenden setzen. Vielleicht sind teils neue Curricula zu entwerfen, die eine Prüfungspraxis zulassen, die begrenzte personelle Kapazitäten berücksichtigt.

Die technologische Entwicklung wird weitergehen. Neue, bessere und entsprechend zuverlässigere Sprachmodelle kommen auf den Markt und werden teils integriert in unsere täglichen Arbeitsabläufe. Aufgabe der Dozierenden, aber auch der strategischen Gremien von Hochschulen wird sein, diese Entwicklungen zu verfolgen, zu verstehen und idealerweise sogar mitzuprägen. Doch die entsprechenden Institutionen wirken im Vergleich zu privaten Unternehmen, die schnell neue Produkte entwickeln und den Ausbildungsmarkt verändern, eher träge. So konstatiert ein Bericht von der Learntec 2024 in Karlsruhe, der nach eigenen Angaben größten Veranstaltung für digitale Bildung in Europa, dass „sich die Hochschulen im deutschsprachigen Raum auf die Hinterbeine stellen müssen, wenn sie den digitalen Anschluss nicht verpassen möchten.“ (Jörissen, 2024). Verfasst wurde der Bericht von einem Delegierten der Hochschule Luzern.

## 7 Literaturverzeichnis

Bloom, B.S. (Hrsg.) (1973). *Taxonomie von Lernzielen im kognitiven Bereich* (3. Aufl.). Beltz.

Bloom, B.S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational researcher*, 13(6), 4–16. <https://facultycenter.ischool.syr.edu/wp-content/uploads/2012/02/2-sigma.pdf>

Buck, I., Jost, C., Kreis-Hoyer, P., & Limburg, A. (2023). *KI-induzierte Transformation an Hochschulen*. Hochschulforum Digitalisierung. <https://hochschulforumdigitalisierung.de/wp-content/uploads/2023/11/HSRM-Diskussionspapier-Nr.-26-KI-induzierte-Transformation-an-Hochschulen-1.pdf>

Dutta, D. (2017). *Developing an Intelligent Chat-bot Tool to Assist High School Students for Learning General Knowledge Subjects*. Georgia Institute of Technology, repository.gatech.edu. <http://hdl.handle.net/1853/59088>

Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mäde, A. et al. (2023). Unlocking the power of generative AI models and systems such as GPT-4 and ChatGPT for higher education: A guide for students and lecturers. *Hohenheim Discussion Papers in Business, Economics and Social Sciences*, 2023(02). <https://hdl.handle.net/10419/270970>

- Hanke, U. (2023). Selbstlernkurs „Prüfen in einer Welt mit generativen KI-Tools wie Chat-GPT“. <https://hochschuldidaktik-akademie.de/31-2/ki-und-pruefen-kurs-kaufen/>
- Hattie, J., & Yates, G.C. (2013). *Visible learning and the science of how we learn*. Routledge.
- Ilahi, M., Belcadhi, L.C., & Braham, R. (2014). Semantic models for competence-based assessment. *International Journal of Human Capital and Information Technology Professionals*, 5(3), 33–46. <http://dx.doi.org/10.4018/ijhcitp.2014070103>
- Jörissen, St. (2024, 27. Juni). *Wo die Musik spielt – Eindrücke von der Learntec-Messe für digitale Bildung*. ZLLFblog. Zentrum für Lernen, Lehren und Forschen. <https://blog.hslu.ch/blog/2024/06/27/wo-die-musik-spielt-eindrucke-von-der-learntec-messe-fuer-digitale-bildung/>
- Jungherr, A. (2023). *Using ChatGPT and Other Large Language Model (LLM) Applications for Academic Paper Assignments*. Otto-Friedrich-Universität. <https://nbn-resolving.org/urn:nbn:de:bvb:473-irb-589507>
- Kaiser, H. (2019). *Situationsdidaktik konkret. Unterrichtsrezepte, Beispiele, Grundlagen*. Hep Verlag.
- Khan, S. (2023, 14. März). *Harnessing GPT-4 so that all students benefit. A nonprofit approach for equal access Khan Academy*. <https://blog.khanacademy.org/harnessing-ai-so-that-all-students-benefit-a-nonprofit-approach-for-equal-access/>
- Long, S., Dowsing, R., & Craven, P. (2003). Knowledge-Based Systems for Marking Professional IT Skills Examinations. In A. Macintosh, R. Ellis & F. Coenen (Hrsg.), *Applications and Innovations in Intelligent Systems X*. Springer. [https://doi.org/10.1007/978-1-4471-0649-4\\_1](https://doi.org/10.1007/978-1-4471-0649-4_1)
- MacKevett, D., & Gutmann, M. (2023). High-Stakes Online Exams. Faculty Perceptions on Forced Digitization of Assessment During Corona at a Swiss Business School. *International Journal of Emerging Technologies in Learning*, 18(13), 193–208. <https://doi.org/10.3991/ijet.v18i13.38379>
- Marfurt, A. (2023). *Prüfen mit ChatGPT*. Unveröffentlichte Präsentation an der Hochschule Luzern vom 04.05.2023.
- Mouthaan, T.J., Olthuis, W., & Vos, H. (2003). Competence-based EE-learning: (how) can we implement it? Paper presented at *IEEE International Conference on Microelectronic Systems Education* (S. 33–34). <https://research.utwente.nl/files/6151228/01205242.pdf>

Open AI (2024). *GPT-4 Technical Report Version 6*. arXiv. <https://doi.org/10.48550/arXiv.2303.08774>

Stern, Th. (2010). *Förderliche Leistungsbewertung*. Österreichisches Zentrum für Persönlichkeitsbildung und soziales Lernen. [https://www.oezepts.at/wp-content/uploads/2011/07/Leistungsbewertung\\_Onlineversion\\_Neu.pdf](https://www.oezepts.at/wp-content/uploads/2011/07/Leistungsbewertung_Onlineversion_Neu.pdf)

TH Köln (online). *Studienverlaufsplan zur Prüfungsordnung*. [https://www.th-koeln.de/studium/maschinenbau--product-engineering-and-context-bachelor--ordnungen-und-formulare\\_92178.php](https://www.th-koeln.de/studium/maschinenbau--product-engineering-and-context-bachelor--ordnungen-und-formulare_92178.php). Stand vom 10.07.2024

Tu, X., Zou, J., Su, W., & Zhang, L. (2023). *What Should Data Science Education Do with Large Language Models?* arXiv. <https://doi.org/10.48550/arXiv.2307.02792>

Walzik, S. (2012). *Kompetenzorientiert prüfen. Leistungsbewertung an der Hochschule in Theorie und Praxis*. Verlag Barbara Budrich.

Wampfler, P. (2023). *Das Fächerproblem*. <https://schulesocialmedia.com/2022/02/09/das-facherproblem-nicht-nur-am-gymnasium/>

Weinert, F.E. (2002). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. Aufl.) (S. 17–31). Beltz.