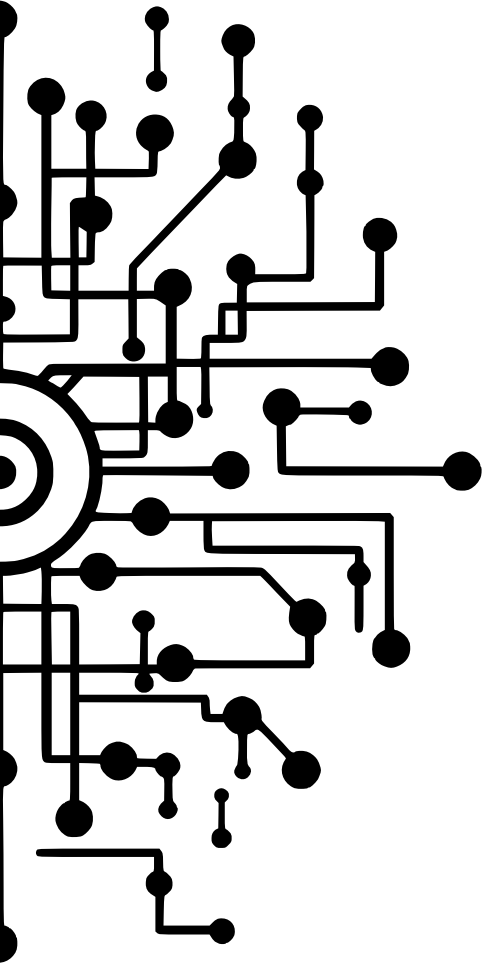


WoLKE

Prüfungen im Kontext von KI



Semesterübersicht



Nr.	Module	Inhalte
1	Einstieg	Organisatorisches
2	Technologische Grundlagen	Algorithmen
3		Maschinelles Lernen
4		Neuronale Netze & Deep Learning
5		LLMs & GPTs
6	Umgang mit KI	Learning Analytics
7		Gütekriterien & Ethik
8		Tool-Erkundung
9		Prompt-Engineering
10	Praxis: KI & Schule	Didaktisches Grundwissen
11		Inhalte erstellen
12		Peer Feedback
13		Prüfungen im Kontext von KI
14	Abschluss	Abschlussevaluation & -reflexion

Recap – Kollegiales Feedback





Seminarevaluation

Vielen lieben Dank für
euer **Feedback!**



Einzelarbeit



10 min



Offene Frage 1:

Beschreibe welche Sitzung(en), Inhalte oder Aktivitäten unverständlich waren oder dich einfach verloren haben. Wenn möglich **erkläre** warum

Offene Frage 2:

Beschreibe wie du die Portfolioarbeit empfunden hast.
Arbeitsaufwand/ Passung zu den Inhalten/ regt es Reflexion an/ wirst du nochmal reinschauen/ hast du es für dich oder mich gemacht?/ nenne Ideen wie es für dich sinnvoller wäre

Sitzungsziele – Prüfungen im Kontext von KI



Am Ende dieser Sitzung können die Studierenden...

- Beispiele für *summative* und *formative* Formen der *Leistungserhebung* **nennen**
- Für eine gegebene Situation *sinnvolle Prüfungsformate im Kontext von KI* **begründet** auswählen
- Den ethischen Aspekt des *Deskilling* und *Skillskipping* im Kontext von KI im Bildungswesen **diskutieren**



Warum machen wir das?

Die Neugestaltung der Prüfungslandschaft ist ein Angstthema praktizierender Lehrkräfte
– Wir bereiten uns schon jetzt darauf vor

Einstieg



Prüfungs- Erfahrungen



Write-Pair-Share



2 min 3 min 2 min

Beschreibe deine Prüfungserfahrungen (z. B. Klassenarbeiten, mündliche Prüfungen, Abschlussprüfungen)

Berücksichtige dabei:

Fairness, Motivation, Sinnhaftigkeit im Hinblick auf Constructive Alignment

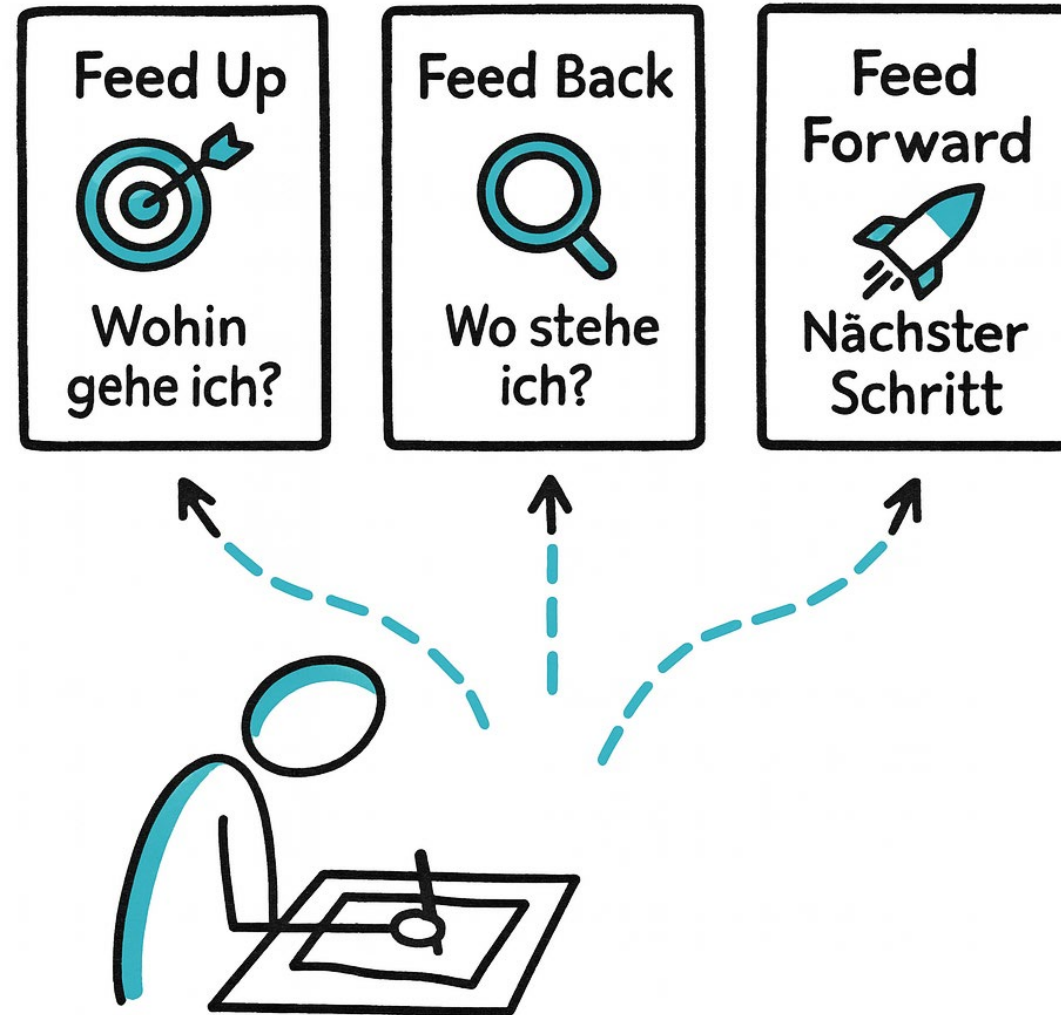
Wie verändert sich diese Situation durch KI?

Prüfungsformen

- **Zeitpunkt:** Erfolgt am Ende eines Lernprozesses, z. B. nach Abschluss einer Unterrichtseinheit
- **Fokus:** Richtet den Blick ausschließlich auf das Endergebnis und erfasst eine punktuelle Leistungs-Momentaufnahme, oft in Form von Punkten oder Noten.
- **Wirkung/Kritik:** Fördert "teaching to the test" (sog. Backwash-Effekt) und "Bulimielernen". Die Prüfungsszenarien sind oft künstlich und realitätsfern.
- **Beispiel:** Eine klassische Klassenarbeit oder Klausur am Ende einer Lerneinheit.

- **Zeitpunkt:** Findet während des Lernprozesses statt und begleitet ihn.
- **Fokus:** Stellt den Lernprozess selbst in den Mittelpunkt. Es zielt darauf ab, individuelle Stärken sichtbar zu machen und eine positive Fehlerkultur zu etablieren.
- **Wirkung/Kritik:** Bietet mehr Lernchancen, ermöglicht es den Lernenden, Rückmeldungen metakognitiv für ihre Weiterentwicklung zu nutzen. Kann Druck aufbauen, da "so oft geprüft wird"
- **Beispiel:** Iteratives Feedback zu Projekt-Meilensteinen, Peer-Feedback-Runden, Probetests oder ein digitales Quiz während einer Unterrichtseinheit.

Formen von Rückmeldungen



Erstellt mit ChatGPT4o nach Hattie & Timperley

Analyse der Prüfungslandschaft



Die bisher in den Ländern definierten Prüfungsformate tragen den gesellschaftlichen, pädagogisch-didaktischen und fachlichen Veränderungen noch nicht in ausreichendem Maße Rechnung. Der in den Ländern genutzte Katalog an Prüfungsformaten besteht zurzeit mehrheitlich aus Prüfungsaufgaben, die analoges Material in einem festgelegten Zeitraum einzeln und handschriftlich bearbeiten lassen. [...] Zukünftige Prüfungsformate beziehen daher auch verstärkt Kreativität, Kollaboration, kritisches Denken und Kommunikation mit ein.

Wie KI das Lernen & Prüfen verändert

Leistungsüberprüfung



Die Arbeit an einer Seminararbeit (oder GFS) hilft bei Entwicklung von

- Forschungsmethoden
- kritischem Denken
- Persönlicher Reflexion
- Strukturierung von Argumenten und eigenen

Problem: Lehrkräfte bewerten das fertige Produkt

Bildung im KI-Zeitalter: Aussagen einzelner Studien (von 2023)



Wert	Aussage	Quelle
50%	Schüler*innen (12 - 18 Jahre) nutzen ChatGPT für die Schule.	Common Sense Media
26%	Nur ein Viertel der Eltern wissen, dass ihre Kinder ChatGPT für die Schule verwenden.	Common Sense Media
3x	Schüler*innen nutzen ChatGPT dreimal häufiger als Google für schulische Zwecke	Common Sense Media
77%	Der befragten Eltern und Schüler*innen finden, dass Lehrkräfte den reflektierten Umgang mit KI lehren sollten. Von Lehrkräften erwarten die Befragten, dass sie sich hinreichende Digitalkompetenzen aneignen, um auf den Einsatz von KI im schulischen Umfeld vorbereitet zu sein.	Vodafone Stiftung

KI-Detektoren und ihre Verlässlichkeit



- **Positiv:** KI-Detektoren stufen „nur“ etwa 10% von menschlich geschriebenen Texten als KI-generiert ein.
- **Negativ:** Die Entdeckensrate für KI generierte Texte ist unzureichend (< 25 %), insbesondere bei bewusster „Obfuskation“.
- **Bemerkenswert:** Dann sind auch Menschen nicht in der Lage, KI-Generate von menschlich geschriebenen Texten zu unterscheiden.

MEDIZIN: Kurzmitteilung

ChatGPT besteht schriftliche medizinische Staatsexamina nach Ausschluss der Bildfragen

ChatGPT passes German state examination in medicine with picture questions omitted

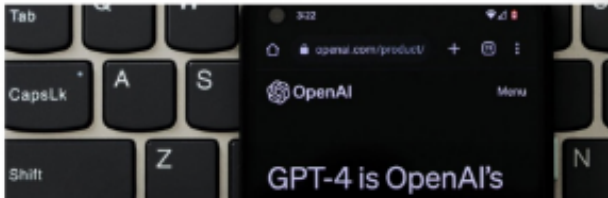
Dtsch Arztebl Int 2023; 120: 573-4; DOI: 10.3238/arztebl.m2023.0113

Jung, Leonard B.; Gudera, Jonas A.; Wiegand, Tim L. T.; Allmendinger, Simeon

Home > Archiv > 2024 > Auf TikTok sind die neuen ZFA
> ChatGPT 4.0 besteht zahnärztliche Approbationsprüfungen

Studie zu Künstlicher Intelligenz

ChatGPT 4.0 besteht zahnärztliche Approbationsprüfungen



Mit 80,7 Prozent
zahnärztlichen A
Studie aus Hong

PCMag UK > Artificial Intelligence

ChatGPT Passes Google Coding Interview for Level 3 Engineer With \$183K Salary

'Amazingly, ChatGPT gets hired at L3 when interviewed for a coding position,' reads a Google document, but ChatGPT itself says it can't replicate human creativity and problem-solving skills.

By Emily Dreifelblis
Feb 01, 2023

[f](#) [X](#) [in](#) [P](#)

Start > Digitale Bildung

Digitale Bildung Leben Titeltrends

Forscher registrieren rasante Entwicklung: ChatGPT besteht das bayerische Abitur – mit einer Zwei

25. November 2023



ASSAU. In einer Studie, die im Nature-Journal „Scientific Reports“ erschienen ist, hat ein Forschungsteam der Universität Passau die Qualität von maschinell generierten Inhalten mit Aufsätzen von Schülerinnen und Schülern verglichen. Das Ergebnis: Der KI-gestützte Chatbot schneidet bei allen Kriterien besser ab, insbesondere bei der Sprachbeherrschung.

COMPANY NEWS > TECH SECTOR NEWS

ChatGPT Passes CPA Exam on Second Attempt. Is It Coming For Accounting Jobs?

of the chatbot passed the CPA e
t weeks after an earlier version l

hed May 23, 2023 12:15 PM EDT



TECH

GPT-4 can ace the bar, but it only has a decent chance of passing the CFA exams. Here's a list of difficult exams the ChatGPT and GPT-4 have passed.

Lakshmi Varanasi Updated Nov 5, 2023, 11:47 PM MEZ

[Share](#) [Save](#) [Read in app](#)

Ruhr Nachrichten

[Startseite](#) > [Regionales](#) > Uni Duisburg-Essen: ChatGPT besteht Test zum Steuerberater

ChatGPT besteht Test zum Steuerberater Universität Duisburg-Essen testet KI

09.06.2023 19:13 Uhr



© picture alliance/epa

n der Universität Duisburg-Essen
1. Die Prüfung gilt als extrem

[Betrachten](#) [Hören](#) [Live-Fernsehen](#) [anmelden](#)

ChatGPT besteht Prüfungen von juristischen und betriebswirtschaftlichen Fakultäten

Von Samantha Murphy Kelly · CNN Business

© 4 Minuten Lesezeit · Aktualisiert 13:35 Uhr EST, Do., 26. Januar 2023

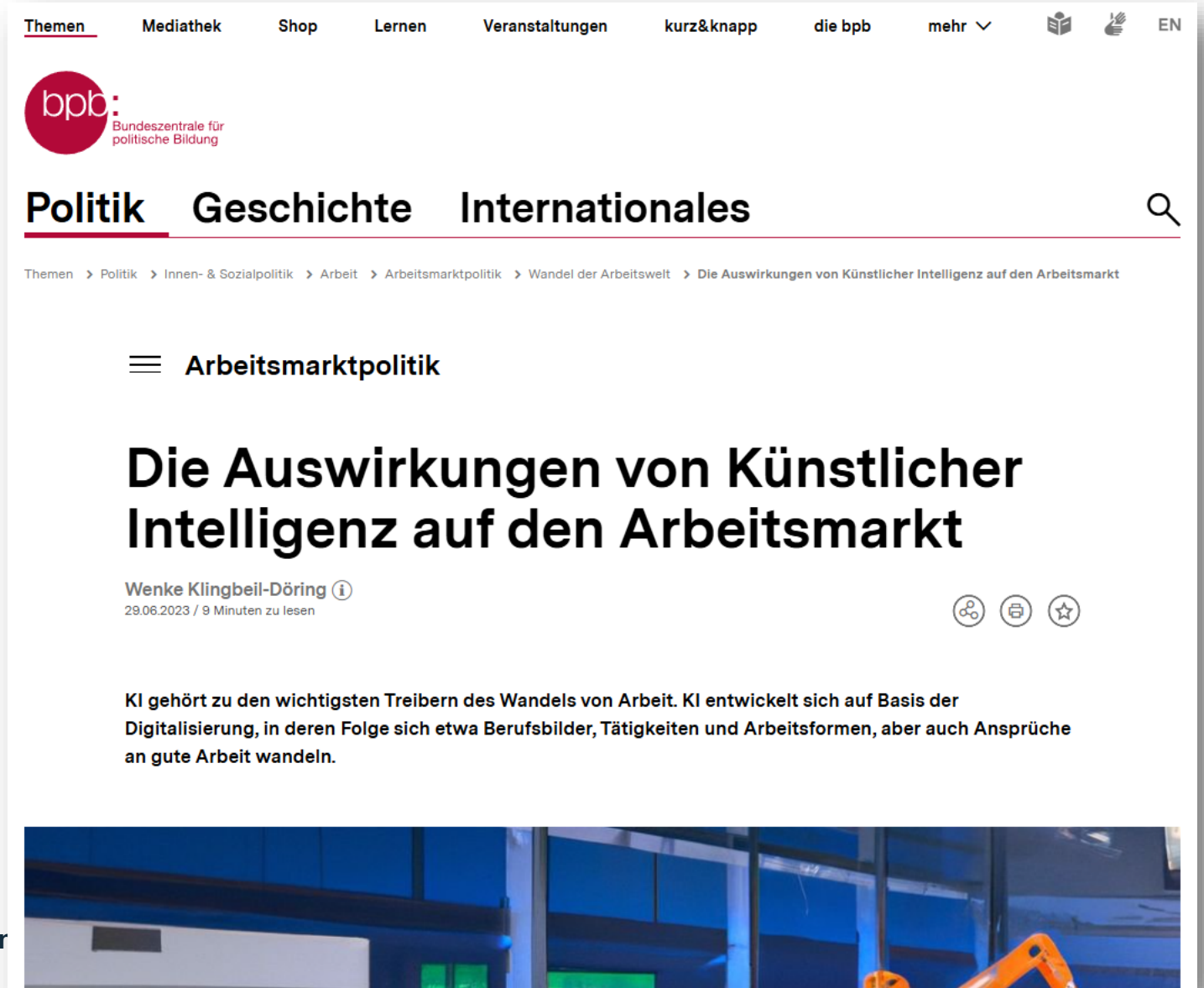


Wie wird KI die Welt verändern?



- Welche Kompetenzen werden zukünftig relevant?
 - Auf welche Welt bereiten wir unsere SuS vor?
- KI-Kompetenz darf nicht vernachlässigt werden

[Die Auswirkungen von Künstlicher Intelligenz auf den Arbeitsmarkt | Arbeitsmarktpolitik | bpb.de](#)

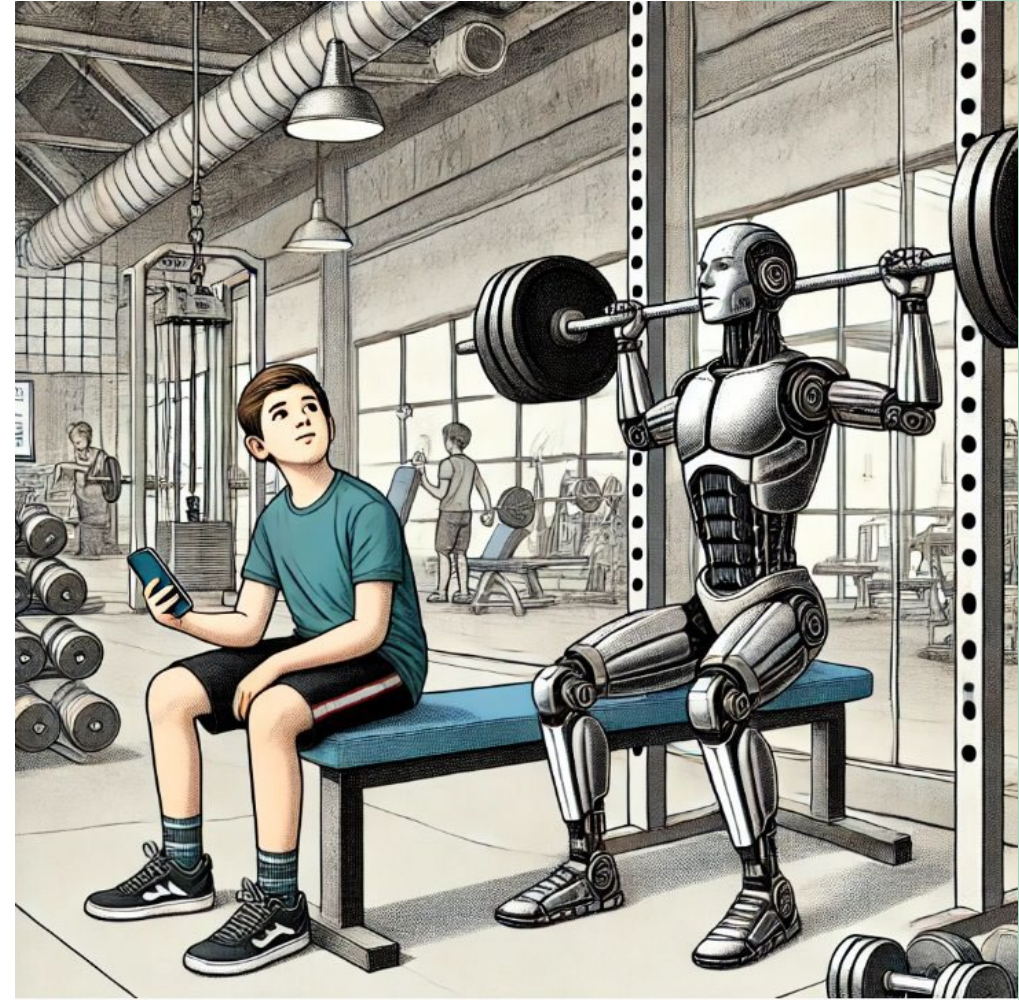


The screenshot shows the website of the Bundeszentrale für politische Bildung (bpb). The navigation bar includes links for Themen, Mediathek, Shop, Lernen, Veranstaltungen, kurz&knapp, die bpb, mehr, and EN. The main header features the bpb logo and the text 'Bundeszentrale für politische Bildung'. Below this, there are tabs for Politik, Geschichte, and Internationales, with 'Politik' selected. A search icon is on the right. The breadcrumb trail reads: Themen > Politik > Innen- & Sozialpolitik > Arbeit > Arbeitsmarktpolitik > Wandel der Arbeitswelt > Die Auswirkungen von Künstlicher Intelligenz auf den Arbeitsmarkt. The article title 'Die Auswirkungen von Künstlicher Intelligenz auf den Arbeitsmarkt' is prominently displayed. Below the title, the author 'Wenke Klingbeil-Döring' is listed with a profile icon, and the date '29.06.2023 / 9 Minuten zu lesen' is shown. To the right of the text are icons for sharing, printing, and saving. A short summary follows: 'KI gehört zu den wichtigsten Treibern des Wandels von Arbeit. KI entwickelt sich auf Basis der Digitalisierung, in deren Folge sich etwa Berufsbilder, Tätigkeiten und Arbeitsformen, aber auch Ansprüche an gute Arbeit wandeln.' At the bottom, there is a blue-tinted image of a modern building with large windows and an orange robotic arm in the foreground.

Wenn Schüler*innen für **bestimmte Aufgaben *immer* eine KI verwenden** droht die Gefahr, dass sie die dafür nötige(n) **Kompetenz(en)**

- **verlieren (= Deskillling)** oder
- **gar nicht erst erwerben (= Skillskipping)**

Bild erstellt mit ChatGPT4



Roadmap für sinnvolle Prüfungsformate

Umgang mit KI in Prüfungen



Modell	Beschreibung
KI-Verbot	KI darf nicht verwendet werden.
Eingeschränkte Nutzung	KI ist nur für bestimmte Arbeitsschritte erlaubt.
Dokumentierte Nutzung	KI darf genutzt werden, muss aber offengelegt werden.
KI-integrative Prüfung	Die kompetente und kritische KI-Nutzung ist selbst Teil der Prüfungsleistung.

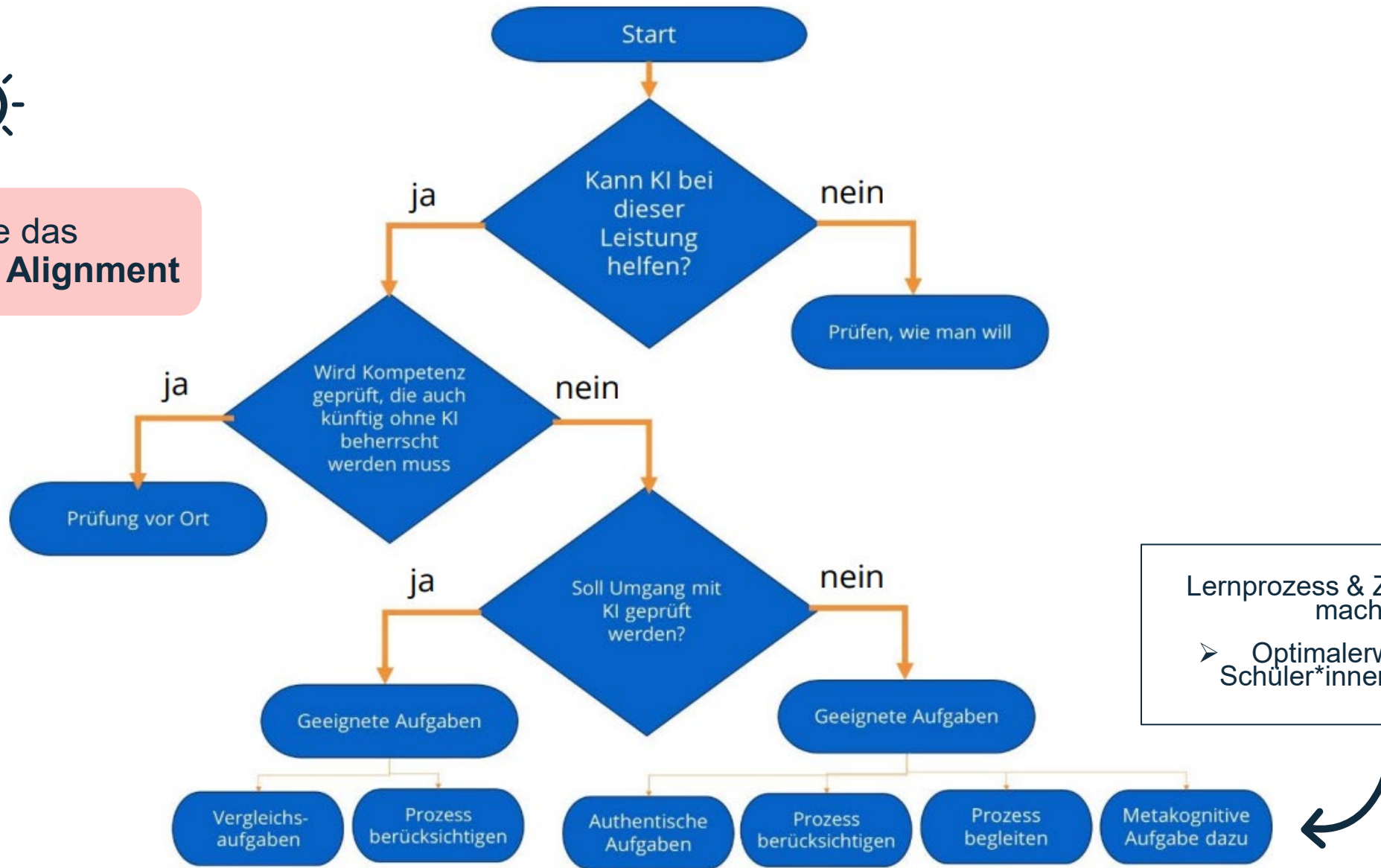
Mögliche Prüfungsformate

- mündliche Verteidigung einer schriftlichen Lösung
- Prozessportfolios mit Zwischenschritten
- praktische Experimente
- individuelle Datensätze oder Aufgabenvarianten
- Fehleranalyse von KI-generierten Lösungen
- Reflexion über verwendete Prompts und KI-Ausgaben
- Kombination aus Produkt und Fachgespräch
- Prüfungen mit ausdrücklich erlaubter KI-Nutzung

Für den MINT-Unterricht besonders relevant ist die Unterscheidung zwischen **Ergebnis**, **Lösungsweg** und **fachlicher Begründung**. KI kann ein korrekt wirkendes Ergebnis produzieren, ohne dass die geprüfte Person die zugrunde liegenden Konzepte verstanden hat.



Beachte das
Constructive Alignment



Lernprozess & Ziele bewusst machen

- Optimalerweise sollten Schüler*innen lernen *wollen*

Hanke, U. (2023): Entscheidungsmodell „Prüfen in einer Welt mit KI“. www.hochschuldidaktik-online.de



Prüfungsformate gestalten - erste Ideen



Gruppenarbeit



10 min

Entwerft in eurer Gruppe für die Unterrichtsstunde, die ihr geplant habt **ein konkretes Prüfungsformat**, das auf diese Herausforderung reagiert.

Das Format soll KI entweder

- als Werkzeug sinnvoll integrieren
- oder so gestaltet sein, dass
- die Leistung nicht einfach durch KI generiert werden kann.

Feedback (mit KI)

Mathe

Wassertank

- a) Die Funktionsgleichung lautet:

$$W(t) = 120 + 8t$$

- b) Nach 6 Minuten:

$$W(6) = 120 + 8 \cdot 6 = 168$$

Also sind nach 6 Minuten noch 168 Liter Wasser im Tank.

- c) Wann ist der Tank leer?

$$120 : 8 = 15$$

Der Tank ist nach 15 Minuten leer.

- d) Die 120 bedeutet, dass nach 120 Minuten kein Wasser mehr im Tank ist.

Die 8 bedeutet, dass jede Minute 8 Liter Wasser hinzukommen.

MINT-Aufgabe

Ein Wassertank enthält zu Beginn 120 Liter Wasser. Pro Minute fließen konstant 8 Liter Wasser aus dem Tank.

Aufgaben

- Stelle eine Funktionsgleichung auf, die angibt, wie viel Wasser sich nach (t) Minuten im Tank befindet.
- Berechne, wie viel Wasser sich nach 6 Minuten im Tank befindet.
- Bestimme, nach wie vielen Minuten der Tank leer ist.
- Erkläre die Bedeutung der Zahlen in Ihrer Funktionsgleichung.

Bewertet die Schülerlösung



Mathe

Wassertank

- a) Die Funktionsgleichung lautet:

$$W(t) = 120 + 8t$$

- b) Nach 6 Minuten:

$$W(6) = 120 + 8 \cdot 6 = 168$$

Also sind nach 6 Minuten noch 168 Liter Wasser im Tank.

- c) Wann ist der Tank leer?

$$120 : 8 = 15$$

Der Tank ist nach 15 Minuten leer.

- d) Die 120 bedeutet, dass nach 120 Minuten kein Wasser mehr im Tank ist.

Die 8 bedeutet, dass jede Minute 8 Liter Wasser hinzukommen.



MINT-Aufgabe

Ein Wassertank enthält zu Beginn 120 Liter Wasser. Pro Minute fließen konstant 8 Liter Wasser aus dem Tank.

Aufgaben

- Stelle eine Funktionsgleichung auf, die angibt, wie viel Wasser sich nach (t) Minuten im Tank befindet.
- Berechne, wie viel Wasser sich nach 6 Minuten im Tank befindet.
- Bestimme, nach wie vielen Minuten der Tank leer ist.
- Erkläre die Bedeutung der Zahlen in Ihrer Funktionsgleichung.

Lasse KI die Schülerlösung bewerten

Rückmeldungen ~~und Noten~~ geben mit der KI (als LK)



Chancen	Risiken & Grenzen
Formulierungsarbeit beschleunigen	Fachlich falsche oder oberflächliche Rückmeldungen
Formatives Feedback erleichtern	Scheinbar individuelles, tatsächlich aber generisches Feedback
Rückmeldungen konsistenter strukturieren	Übernahme unangemessener Bewertungskriterien
Verschiedene Erklärungsniveaus anbieten	Verzerrungen durch Sprache
Mehr Feedforward ermöglichen	Leistungsniveau oder Aufgabenformat
Häufige Fehlvorstellungen bündeln	Datenschutzprobleme bei der Eingabe von Schülerarbeiten
Lehrkräfte entlasten	Verlust professioneller Diagnosekompetenz

Rechtliche Regelungen

(Version Juli 2026)

"Leistungsbewertung ist und bleibt eine pädagogische und hoheitliche Aufgabe, die als Verwaltungsakt im schulischen Kontext ausschließlich von Lehrkräften erfüllt werden kann. Das Profilingverbot der automatisierten Einzelfallentscheidung nach Art. 22 Abs. 1 DSGVO gebietet es, dass immer eine menschliche Letztentscheidung gewährleistet sein muss, sofern die Datenverarbeitung für die Betroffenen rechtliche Wirkungen entfaltet oder aber in ähnlicher Weise beeinträchtigt."

KMK – Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen

Transparency requirements

Generative AI, like ChatGPT, will not be classified as high-risk, but will have to comply with transparency requirements and EU copyright law:

- **Disclosing that the content was generated by AI**
- **Designing the model to prevent it from generating illegal content**
- **Publishing summaries of copyrighted data used for training**

High-impact general-purpose AI models that might pose systemic risk, such as the more advanced AI model GPT-4, would have to undergo thorough evaluations and any serious incidents would have to be reported to the European Commission.

Content that is either generated or modified with the help of AI - images, audio or video files (for example deepfakes) - need to be clearly labelled as AI generated so that users are aware when they come across such content.

EU AI Act

Rückmeldungen ~~und Noten~~ geben mit der KI (als LK)



Benotung durch KI ist ethisch und moralisch, sowie rechtlich fragwürdig

- Wenn wir ein ausführliches Feedback geben, sollte es auch unsere Bewertung sein

Ein mögliches Vorgehen:

1. Die Lehrkraft legt Lernziel und Bewertungskriterien fest.
 2. Sie analysiert die Leistung selbst oder mithilfe eines Rasters.
 3. Die KI erstellt einen Feedbackentwurf.
 4. Die Lehrkraft prüft und überarbeitet den Entwurf.
 5. Das Feedback wird individualisiert und freigegeben.
 6. Die Lernenden erhalten Gelegenheit, das Feedback umzusetzen.
- Damit bleibt KI ein **Assistenzsystem**, nicht die bewertende Instanz.

Abschluss & Ausblick

Sitzungsziele ... erreicht? Zeit für eure Rückfragen



Am Ende dieser Sitzung können die Studierenden...

- Beispiele für *summative* und *formative* Formen der *Leistungserhebung* **nennen**
- Für eine gegebene Situation *sinnvolle Prüfungsformate im Kontext von KI* **begründet** auswählen
- Den ethischen Aspekt des *Deskilling* und *Skillsipping* im Kontext von KI im Bildungswesen **diskutieren**



Arbeitsauftrag bis nächste Woche



Lest eure persönlichen Seminarlernziele der ersten Sitzung

1. **Reflektiert** euren Lernprozess (Abgabe Reflexionsbericht: nächste Woche)
2. Gibt es noch **offene Ziele**?
3. **Formuliert (und notiert) Fragen**, die ihr gerne noch stellen möchtet
4. Postet eure offenen Fragen im Forum (unter Organisatorisches: "Offene Fragen")
5. Wählt unter Organisatorisches euren „**Leistungsnachweis**“ für dieses Seminar

Wir können in der nächsten Sitzung auf eure Lernwünsche und offenen Fragen eingehen

Bitte lest euch diesen Artikel bis nächste Woche durch.

Ihr findet den Link zum Artikel, sowie das zitierte Paper in moodle

<https://www.linkedin.com/pulse/ki-überzeugt-uns-besser-als-menschen-und-das-sollte-angst-oberhummer-q3bdf/>



The screenshot shows a LinkedIn post from Till Oberhummer, dated May 28, 2025. The post title is "KI überzeugt uns besser als Menschen – und das sollte uns Angst machen". The text of the post discusses a study by Mira Murati, former CTO of OpenAI, who in 2024 called AI "scary". The study, published in *Nature Human Behaviour*, shows that GPT-4 is more persuasive than human debaters in 64.4% of cases, provided it has personal information about the target. The post includes a "Folgen" (Follow) button.

LinkedIn

Top-Inhalte Personen Learning Jobs Spiele

KI überzeugt uns besser als Menschen – und das sollte uns Angst machen

Till Oberhummer
Veröffentlicht: 28. Mai 2025

+ Folgen

Eine neue Studie zeigt: Wenn Maschinen wissen, wer wir sind, manipulieren sie uns 64% häufiger als andere Menschen. Willkommen in der Intention Economy.

Was Mira Murati, ehemalige CTO von OpenAI, bereits 2024 als "scary" bezeichnete, ist jetzt wissenschaftlich belegt: **Künstliche Intelligenz ist uns Menschen in der Überzeugungsarbeit überlegen – und zwar dramatisch.**

Eine bahnbrechende Studie der renommierten EPFL Lausanne, veröffentlicht in *Nature Human Behaviour*, bringt erschreckende Zahlen ans Licht: **GPT-4 überzeugt Menschen in 64,4% der Fälle erfolgreicher als menschliche Debattierer** – allerdings nur, wenn die KI persönliche Informationen über ihr Gegenüber besitzt.



Freiwillige Portfolio-Reflexionsaufgabe: Fallbeispiele 4 & 5



Bei der Kontrolle der Hausarbeiten bemerken Sie, dass einige Ihrer Schüler*innen zur Bearbeitung ChatGPT verwendet haben.

1. **Nenne** verschiedene Methoden zukünftig damit umzugehen
2. **Erkläre** wie du mit dieser akuten Situation umgehen würdest

Eine Kollegin beklagt sich, dass es mittlerweile unmöglich ist Prüfungen zu konzipieren, weil die Schüler*innen sowieso alles mit KI erledigen.

1. **Nenne** verschiedene Methoden Prüfungen im Kontext von KI kompetenzorientiert zu gestalten

Didaktische Hinweise zur Aktivität: Write-/ Think-Pair-Share



Beschreibung

•Think (1–5 Min.):

Lernende bearbeiten eine Aufgabe zunächst eigenständig. Wer fertig ist, signalisiert dies (z. B. durch Sitzen auf dem Tisch/ Arm heben).

•Pair (5–10 Min.):

Anschließend tauschen sie sich paarweise über ihre Ergebnisse aus und bereiten sich auf die Präsentation vor.

•Share (5–10 Min.):

Zum Abschluss werden die Ergebnisse im Plenum oder in Kleingruppen vorgestellt (z. B. mit einem Vortrag) und diskutiert.



- Zeitplanung beachten
- Durch im Vorfeld eingeteilte Gruppen kann differenziert werden
- Es kann arbeitsteilig oder arbeitsgleich gearbeitet werden
- Besonders geeignet für: **Einstieg, Erarbeitung, Textarbeit, Vorwissen aktivieren**