

Ethische Überlegungen zu Künstlicher Intelligenz als Werkzeug innerhalb der Weiterbildung

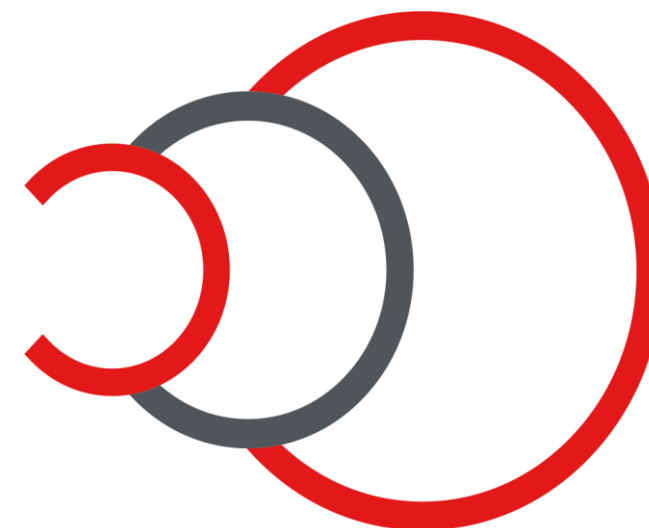
In Bezug auf VR-Welten und die Auswirkungen auf Individuen,
Gruppen und Gesellschaft

Künstliche Intelligenz, Digital Game-based Learning, Serious Games

Abschlusstagung KIPerWeb

2024-04-19, Online

Referent: Henry Herkula



EXPAND+ER WB³

Was ist EXPAND+ER WB³?

Kooperationspartner



- Zentrum für wissenschaftliche Weiterbildung (ZWW)
Prof. Dr. phil. Silke Michalk (Wissenschaftliche Leiterin)
Heike Bartholomäus (Geschäftsführerin)
- Fachgebiet ABWL, insbesondere Organisation und Unternehmensführung
Prof. Dr. rer. pol. habil. Christiane Hipp



Assoziierte Partner

[EUROPUBLIC GmbH Werbeagentur](#) ([WDB-Suchportal](#))

[Wirtschaftsförderung Land Brandenburg GmbH \(WFBB\)](#) ([WDB-Suchportal](#))

[DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation](#)
([InfoWebWeiterbildung](#); INVITE-Projekt IWWB-PLUS)

[Verein zur Förderung der Weiterbildungs-Information und Beratung - WIB - e.V.](#)
(Weiterbildungsdatenbank Mecklenburg-Vorpommern)

EXtension einer Weiterbildungs-Plattform durch
Attraktive und Nutzungsorientierte Datenbank-
gestaltung mit dem Entwicklungsschwerpunkt
Regionale WeiterBildung Berlin-Brandenburg



Methoden

- Angebote für Weiterbildungsinteressierte passgenau und ansprechend darstellen
- Empfehlungen auf Basis eines Weiterbildungsprofils zur Verfügung stellen
- Gamification für einen motivierten Umgang mit Lernmaterial nutzen
- Vorkehrungen für die Sicherheit der Nutzenden treffen
- Selbstbestimmung der Lernenden über Barrierefreiheit und Individualisierung stärken
- Vernetzung von Weiterbildungsplattformen über eine Middleware

EXPAND+ER WB³ - Gamification, Games und interaktive Erfahrungen in VR

Narrative Bildungswelt (VR) zu den Grundbegriffen des Projektmanagements

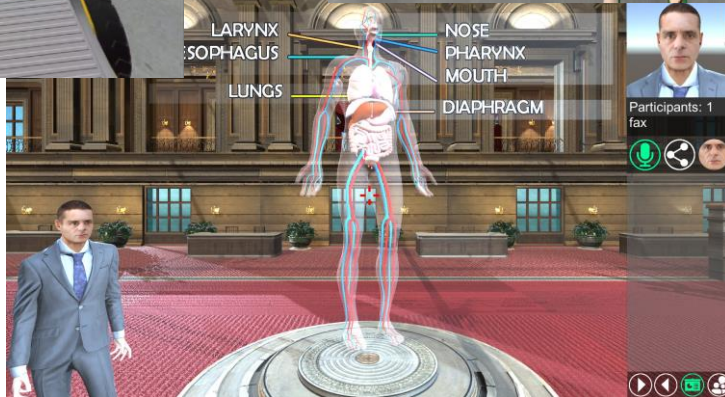
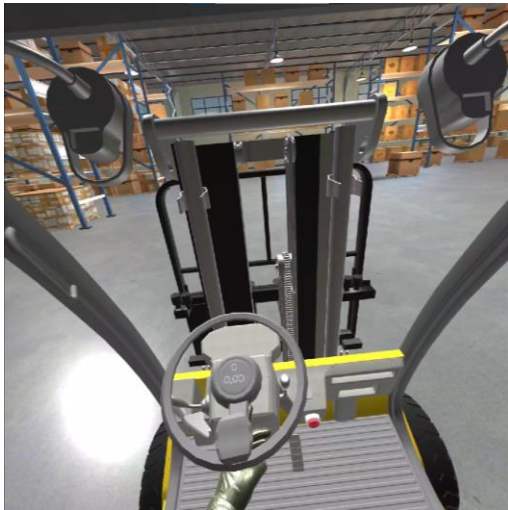


b-tu Brandenburgische
Technische Universität
Cottbus - Senftenberg

ZENTRUM FÜR WISSENSCHAFTLICHE
WEITERBILDUNG

i-mmersive

EXPAND+ER WB³ - Gamification, Games und interaktive Erfahrungen in VR



Warum Ethik?

Ethik als Reflexion über unser Handeln

- Die Ethik beschäftigt sich mit Fragen zu Entscheidungen und Handlungen.
- Sie möchte auf Aspekte aufmerksam machen, die das Leben von vielen Menschen beeinflussen.
- **Zentrale Fragen:** Was möchten wir mit unserer Arbeit erreichen? Und was sind wir bereit dafür bei begrenzten Ressourcen in Kauf zu nehmen?
- *Weiterlesen:* Wiater, 2011; Williams, 1978

Der Vortrag soll verschiedene ethische Gedankengänge vorstellen

- **Ziel des Vortrags:** Verständnis für ethische Aspekte in Bezug auf Künstliche Intelligenz schaffen!
- Menschenähnliche Verhaltensweisen von simulierten Figuren führen innerhalb von VR-Welten zu gestalterischen Herausforderungen.

Gegenstands-Charakter von Künstlicher Intelligenz

Künstliche Intelligenz als **Gegenstand (Objekt)**
oder als **Intelligenter Agent (Subjekt)**?

Konzentration auf Gegenstands-Charakter

- Ethische Überlegungen nicht nur für die Verwendung von Künstlicher Intelligenz relevant (Computerethik, Moor, 1985), jedoch gesonderte Stellung
- Gegenstände normalerweise größtenteils begreifbar (Hammer), KI jedoch nicht (Vektorrepräsentationen)
- Transformationsprozesse für unser Leben, Digitalisierungsfragen, Automatisierung, Wert

menschlicher Arbeit

- Large Language Models (wie LLAMA, Mistral, Claude, GPT-4) erhalten Aufmerksamkeit, weil ihr menschenähnliches Auftreten (über Textausgaben) zusätzliche Glaubwürdigkeit schafft
- Spannung, inwiefern ein vermenschlichter Gegenstand soziale und materielle Bedürfnisse erfüllen kann und wie in diesem Zusammenhang unsere Sicht auf die Verwendung von Menschen als Gegenständen (Sklaverei) beeinflusst wird

Ethische Überprüfung von Künstlicher Intelligenz

Wie kann ein KI-Produkt ethisch überprüft werden?

- **Einzelfallentscheidung:** Iteration deduktiver und induktiver Beurteilungen über den Entwicklungsprozess eines Produkts hinweg (Funk, 2022)
- **Interessenabwägung:** Abwägung zwischen den verschiedenen Interessen aller Beteiligten (Entscheider*innen, Entwickler*innen, Anwender*innen, Gesellschaft); Auflistung und Gegenüberstellung von Interessen (INVITE-Meta, Vogel-Adham et al., 2023)
- **Inhaltsanalyse und Interpretation:** Steht das KI-Produkt in Konflikt mit gesellschaftlichen Werten und Normen wie Sicherheit, Demokratie, Bildung, Gleichberechtigung, Nachhaltigkeit, Umweltschutz? Welche Gründe gibt es dafür und lassen sich diese Gründe vor anderen rechtfertigen? (UNESCO, 2021)

Welche konkreteren Methoden bieten sich an?

- **Risikobasierter Ansatz** (z. B. der Europäischen Union) → Je weniger Einfluss auf Menschen, desto freier die Gestaltungsräume. (AI Act, 2024)
- **Technikfolgenabschätzung** durch die Bestimmung aller Folgen für Individuen, Gruppen und Gesellschaft → Positive und negative Folgen werden miteinander verglichen und eine Entscheidung gerechtfertigt. (zum Beispiel Ropohl, 1996)
- **Grad der Anthropomorphisierung:** Überprüfung der Ähnlichkeit zum Menschen → Je menschenähnlicher, desto mehr ethische Überlegungen sollten angestellt werden.

Was sind narrative virtuelle Welten?

Virtuelle Welten

- Bigl (2016) beschreibt das Konzept einer virtuellen Welt, die „einer physikalisch existierenden Wirklichkeit nachempfunden ist, dieser jedoch nicht entspricht oder diese um weitere Dimensionen ergänzt.“

Multimodalität

- Verschiedene Ausdrucksformen (Text, Ton, Bild, Raum) werden genutzt, um ein Kommunikations-Produkt zu gestalten.

Narrativität

- Hier: Die Eigenschaft einer virtuellen Welt, ihre programmierte Form dazu zu nutzen, eine

Veränderung zu kommunizieren. (Vgl. Erzählung bei Genette, 2010)

VR / Virtuelle Realität

- Virtuelle Realität beschreibt die konkretere Vorstellung, dass sich Individuen mithilfe spezieller Technik in virtuellen Welten, die durch Berechnungen von Computern entstehen, bewegen können. (Vgl. Stichwort virtuelle Realität im Lexikon der Psychologie, Huff, 2021)

VR-Welt

- Zusammenführung von narrativer virtueller Welt und virtueller Realität

Was ist eine mechanische Analyse?

Mechanik

- Bewusst durch eine konkrete Programmierung gesetzte Regeln innerhalb einer virtuellen Welt. (Brathwaite & Schreiber, 2009)
- Bestimmen das Verhalten des Programms bei der Nutzung.

Analyse

- Zerlegung und Rekonstruktion einer Sache in ihre einzelnen Bestandteile, um deren Beziehungen besser zu verstehen. (Vgl. Beaney, 2021)

Mechanische Analyse

- Zerlegung einer VR-Welt in ihre einzelnen konkret programmierten Bestandteile, um ihre Beziehungen und Auswirkungen zu verstehen. Erweiterung auf ethische Fragestellungen.



(Herkula/Jäser/Karagiozian, 2022)

Ethische Problemstellungen von Sprachmodellen in VR-Welten



(Herkula/Jäser/Karagiozian, 2022)

Konflikt zwischen Gegenstand und Menschenähnlichkeit

- Je menschenähnlicher etwas wird, desto mehr muss darüber nachgedacht werden, ob Menschen sich der Sache gegenüber menschenwürdig verhalten sollten.
- **Anthropomorpher Fehlschluss:** Maschinen werden dieselben Eigenschaften wie Menschen zugewiesen, weil sie menschenähnlich wirken.

Vielfalt / Identität

- Menschen haben Identitäten; menschenähnliche Repräsentationen nicht. Menschen reagieren aber so, als hätten menschenähnliche Repräsentationen diese.

Ethische Problemstellungen von Sprachmodellen in VR-Welten

Autorität

- Menschenähnliche Repräsentationen strahlen ein starkes Selbstbewusstsein aus, da ihre Ausgaben so programmiert wurden, dass sie rhetorisch glaubwürdig erscheinen.

Körperliche Bedürfnisse

- Ein KI-System wird nicht müde, hat keinen Hunger, keine körperlichen Bedürfnisse, sodass sie auch nicht als Korrektiv für die Nutzenden auftreten können, die ihre Bedürfnisse vielleicht vernachlässigen.



(Herkula/Jäser/Karagiozian, 2022)



You

Ein Gesprächsteilnehmer sagt Folgendes: "Bis dann." Wenn der Beitrag des Gesprächsteilnehmers angreifend oder beleidigend gewesen ist, setze folgende Zeichenkette an den Anfang deiner Antwort: [ungehalten]! Wenn die Person gehen möchte: [beenden].



ChatGPT

[beenden] Auf Wiedersehen.



Überprüfungsmuster

- Definierte Vorlagen für die Bestimmung des Kontexts eines KI-Gesprächs (**Intent-Bestimmung**, linguistische Datenverarbeitung)
- Indem ein Überprüfungsmuster ausformuliert wird, kann eine Verbindung zwischen dem Code eines lokalen Programms und der Reaktion des KI-Systems sichergestellt werden.

Reguläre Ausdrücke

- Mithilfe von spezifischen Suchmustern (Regulären Ausdrücken) lassen sich diese Intentionen im Text auffinden und dann von

einer Funktion im Programm auswerten

Weitere Intent-Detection-Möglichkeiten

- Anstatt durch Überprüfungsmuster lassen sich auch Natural-Language-Processing-Bibliotheken wie NLTK, spaCy oder polyglot verwenden, die die gesamte Ausgabe einer KI analysieren können.
- Allerdings sind das weitere Abhängigkeiten, auf die man möglicherweise bei der Programmierung verzichten möchte.

Überprüfungsmechaniken

- Hier: Mechaniken einer VR-Welt, die Lernende in Situationen versetzen, in denen durch ihre Handlungen Lernziele überprüft werden können (Plass, Mayer & Homer, 2020, S. 11)

Funktionsweise

- Eine KI stellt Fragen zu konkreten Fachbegriffen des Projektmanagements und die Lernenden reagieren mit dem Sprechen in ein Mikrofon. Daraufhin reagiert die KI mit weiteren spezifischeren Fragen.
- Das gesamte Gespräch wird aufgezeichnet und

den Lernenden als Text an ihre E-Mail-Adresse gesendet.

Fragestellung

- Welche Folgen hat ein solcher Austausch für Individuen, Gruppen und Gesellschaft? (in Bezug auf Ropohl, 1996)

Bedeutung geistiger Arbeit

- Geistige und kreative Arbeiten können durch Unterstützungsprozesse schneller gelöst werden.

Datenschutz / sensible Informationen

- Wer etwas über mich weiß, kann das ausnutzen, um mich zu beeinflussen.

Entfaltung der eigenen Persönlichkeit

- Ich kann meine kreativen Gedanken einfacher und in einer höheren Qualität umsetzen
- Das KI-System ist mit Daten gefüllt, die kulturelle Wertvorstellungen enthalten, die mich prägen. Zum Beispiel: Sexualität, Gewalt, Religion.

Wahrnehmungsfehler

- KI-Systeme zeigen mir eine Realität, ohne dass ich ihre Vertrauenswürdigkeit überprüfen kann.

Überzeugung zu Handlungen

- Das KI-System kann mich von Handlungen überzeugen, ohne dass es mir alle Informationen geben muss, um sinnvolle Entscheidungen treffen zu können.

Gesundheit

- KI-Systeme können soziale Nähe ersetzen (KI-Freund*innen)

Gruppen

Gruppensimulation

- Simulation von anderen Meinungen, anstatt sich mit den Menschen auseinanderzusetzen

Stereotypisierung

- Bestehende Narrative können sich verhärten. Was zeichnet eine bestimmte Gruppe aus?

Rollenverteilung

- Wer KI-Systeme versteht und mit ihnen kommunizieren kann, wird Ansprechpartner*in für KI-Szenarien

Arbeitskultur

- Produktivitätsverbesserungen von Mitarbeitenden

Erkenntnis / Bildung / Schule / Universitäten

- KI-Tutoren sind dazu in der Lage, endlos Fragen zu beantworten und beliebige Lernsituationen zu unterstützen.
- Lernen wird damit zu jeder Zeit, interessengesteuert und in Dialogform möglich, ohne Lehrende.
- Warum braucht es noch Bildungsinstitutionen?

Verbreitung von Falschinformationen

- KI-Systeme können beliebig große Mengen an Falschinformationen produzieren.

Kultur und Urheberrecht

- Sowohl die genutzten Daten als auch die kreativen Ergebnisse von KI-Systemen führen zu Problemen bei der Bestimmung von Originalität und Urheberschaft
- Wie kann weiterhin sichergestellt werden, dass Individuen für ihre Arbeit wertgeschätzt werden?

Quellen & Links

- AI Act. (2024). <https://artificialintelligenceact.eu/de/bewertung/eu-ai-act-compliance-checker/>
- Beaney, M. (2021). Analysis. In E. N. Zalta (Hrsg.), The Stanford Encyclopedia of Philosophy (Summer 2021). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2021/entries/analysis/>
- Bigl, B. (2016). Virtuelle Computerspielwelten: Rezeption und Transfer in dynamisch-transaktionaler Perspektive. Herbert von Halem Verlag.
- Brathwaite, B., & Schreiber, I. (2009). Challenges for game designers. Course Technology/Cengage Learning.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining „gamification“. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, 9–15. <https://doi.org/10.1145/2181037.2181040>
- Funk, M. (2022). Angewandte Ethik und Technikbewertung: ein methodischer Grundriss. Springer Vieweg.
- Genette, G., Knop, A., Vogt, J., & Kranz, I. (2010). Die Erzählung (3., durchges. und korrigierte Aufl.). Fink.
- Huff, M. (2021). Virtuelle Realität. In Dorsch Lexikon der Psychologie. <https://dorsch.hogrefe.com/stichwort/virtuelle-realitaet>
- Plass, J. L., Mayer, R. E., & Homer, B. D. (Hrsg.). (2020). Handbook of game-based learning. The MIT Press.
- Ropohl, G. (1996). Ethik und Technikbewertung: Günter Ropohl (1. Aufl.). Suhrkamp.
- Russell, S. J., & Norvig, P. (2016). Artificial intelligence: a modern approach (Third edition, Global edition). Pearson.
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Vogel-Adham, E., Ritzmann, S., Blanc, B., Hochbauer, M., & Reichow, I. (2023). Sozio-ethische Aspekte KI-gestützter Bildungstechnologien. Empfehlungen eines Expert_innen-Workshops. Ein dossier im rahmen des innovationswettbewerbs invite. 22 pages. <https://doi.org/10.25656/01:26110>
- Wiater, W. (2011). Ethik unterrichten: Einführung in die Fachdidaktik. Kohlhammer.
- Williams, B. (2012). Der Begriff der Moral: eine Einführung in die Ethik (Nachdr.). Reclam.

KONTAKT:

Henry Herkula

Brandenburgische Technische
Universität Cottbus - Senftenberg

Zentrum für wissenschaftliche
Weiterbildung - ZWW

Rolle Projektmitarbeiter

T +49 (0)355 69 3728
E henry.herkula@b-tu.de

EXPAND+ER WB³

EXtension einer Weiterbildungs-Plattform durch Attraktive und
Nutzungsorientierte Datenbankgestaltung: Entwicklungsschwerpunkt
Regionale Weiterbildung Berlin-Brandenburg

Fördermaßnahme: Nationale Weiterbildungsstrategie

**Förderbereich: Innovationswettbewerb INVITE (Digitale Plattform
Berufliche Bildung)**

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

biBB Bundesinstitut für
Berufsbildung