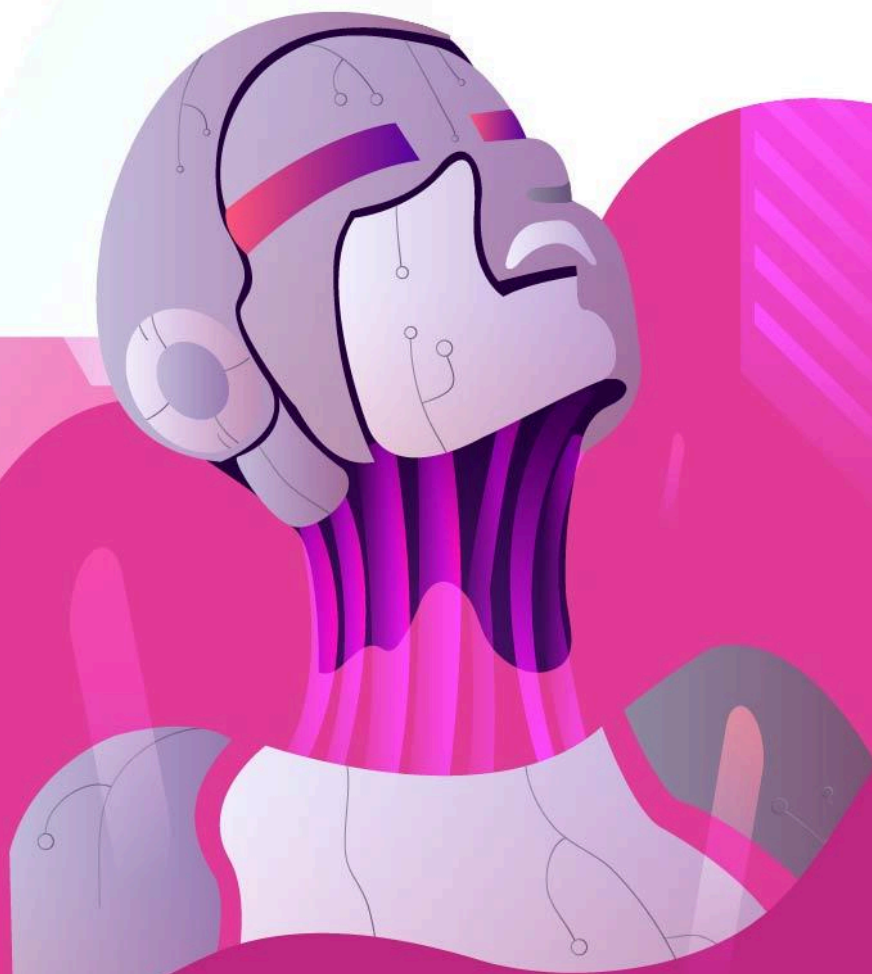




**Digital Diversity:
Crafting Inclusive AI Narratives**

Projektnummer: 2025-1-DE02-KA210-VET-000354956

Handbuch „Barrierefreie Künstliche Intelligenz im Bildungsbereich: Ein Leitfaden für inklusives Design und Lernen“



**Kofinanziert von der
Europäischen Union**

Handbuch

„Barrierefreie Künstliche Intelligenz im Bildungsbereich: Ein Leitfaden für inklusives Design und Lernen“

Digital Diversity: Crafting Inclusive AI Narratives
(D2CIN)

Projektnummer: 2025-1-DE02-KA210-VET-000354956



Kofinanziert von der
Europäischen Union

Von der Europäischen Union finanziert. Die geäußerten Ansichten und Meinungen entsprechen jedoch ausschließlich denen des Autors bzw. der Autoren und spiegeln nicht zwingend die der Europäischen Union oder der Europäischen Exekutivagentur für Bildung und Kultur (EACEA) wider. Weder die Europäische Union noch die EACEA können dafür verantwortlich gemacht werden.

Dokumentname	Handbuch „Barrierefreie Künstliche Intelligenz für die Bildung: Ein umfassender Leitfaden für inklusives Design und Lernen“
Projektaktivität	Aufgabe 2: Barrierefreie Künstliche Intelligenz für die Bildung: Ein umfassender Leitfaden für inklusives Design und Lernen
Revisionsart	Final
Revisionsdatum	30.04.2026
Autoren	Mimic Productions (Deutschland), Budakov Films (Bulgarien)

Urheberrechtserklärung:



Dieses Dokument ist durch die internationale Creative Commons Lizenz Attribution-Non-Commercial-Share-Alike 4.0 geschützt. Sie dürfen:

- Teilen — das Material in jedwedem Format oder Medium vervielfältigen und weiterverbreiten
- Bearbeiten — das Material remixen, verändern und darauf aufbauen, unter folgenden Bedingungen:
- Namensnennung — Sie müssen angemessene Urheber- und Rechteangaben machen, einen Link zur Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Dies darf auf jede zumutbare Art und Weise geschehen, jedoch nicht so, dass der Eindruck entsteht, der Lizenzgeber unterstütze gerade Sie oder Ihre Nutzung.
- Nicht kommerziell — Sie dürfen das Material nicht für kommerzielle Zwecke nutzen.
- Gleiche Bedingungen — Wenn Sie das Material remixen, verändern oder anderweitig direkt darauf aufbauen, dürfen Sie Ihre Beiträge nur unter derselben Lizenz wie das Original verbreiten.

Jede unbefugte Nutzung oder Vervielfältigung der Inhalte dieses Lernmoduls gilt als Verstoß gegen das Urheberrecht und unterliegt rechtlichen Schritten.

Inhaltsverzeichnis

Zusammenfassung.....	5
Einleitung: Zweck und Nutzungsanleitung.....	8
Kapitel 1: Erkennung und Abschwächung von Verzerrungen.....	9
Kapitel 2: Inklusives UX-Design.....	16
Kapitel 3: Barrierefreiheitsstandards in der KI.....	24
Kapitel 4: KI-Ethik und Compliance.....	29
Kapitel 5: Digitales Storytelling für inklusives Lernen.....	36
Fazit.....	42

Zusammenfassung

Das Handbuch „Barrierefreie Künstliche Intelligenz für die Bildung: Ein umfassender Leitfaden für inklusives Design und Lernen“ stellt das zentrale intellektuelle Produkt von Aktivität 2 dar. Es dient als strategische Antwort auf den festgestellten Mangel an digitalen Kompetenzen im europäischen kreativen Berufsbildungssektor und befasst sich insbesondere mit der Schnittstelle von Künstlicher Intelligenz, Ethik und inklusiver Pädagogik.

Kontext und Evidenzgrundlage

Die Entwicklung dieses Handbuchs ist nicht theoretisch; sie basiert auf fundierten Belegen, die im Rahmen von Aktivität 1 erhoben wurden.

- **Interne Lücke:** Unsere anfängliche Bedarfsanalyse ergab, dass 89 % des Zielpersonals (bei Mimic Productions und Budakov Films) keine formale Ausbildung in inklusiven KI-Praktiken besitzen, insbesondere hinsichtlich Bias-Minderung und Barrierefreiheitsstandards [Quelle: Projektantrag].
- **Externe Validierung:** Die Inhalte wurden durch interdisziplinäre Fokusgruppen-Analysen in Deutschland (Experten für inklusive Bildung) und Bulgarien (UX- und Barrierefreiheits-Designer) weiter verfeinert. Diese Sitzungen bestätigten, dass Pädagogen der „pädagogischen Gerechtigkeit“ Vorrang einräumen, während Designer „kognitive Belastung“ und „Vertrauen“ priorisieren – was den doppelten Bedarf an einem Handbuch verdeutlicht, das diese beiden Disziplinen verbindet [Quelle: Analyse- und Empfehlungsbericht].

Das strategische Ziel

Das Hauptziel dieses Dokuments besteht darin, die Forschungsergebnisse aus Aktivität 1 in einen greifbaren „Wie-man-es-macht“-Leitfaden für Berufsbildungsausbilder, kreative Freiberufler und Lehrdesigner zu überführen. Sein Ziel geht über die bloße Einhaltung des EU-KI-Gesetzes hinaus und zielt auf die praktische Anwendung von „Privacy-by-Design“ und „Representational Fairness“ in den täglichen Arbeitsabläufen ab.

Kernaspekte des Handbuchs

Um einen ganzheitlichen Ansatz zur digitalen Inklusion sicherzustellen, ist das Handbuch um fünf kritische Säulen strukturiert, die durch die Ergebnisse der Fokusgruppen bestätigt wurden:

1. Erkennung und Abschwächung von Verzerrungen: Übergang von abstrakten Konzepten zu praktischen „Datensatz-Audits“, um historische und repräsentative Verzerrungen in KI-Avataren zu verhindern [Quelle: Analysebericht, Abschnitt 6].
2. Inklusives UX-Design: Bereitstellung von Heuristiken zur Verwaltung der „kognitiven Belastung“ und zur Vermeidung des „Uncanny-Valley“-Effekts für neurodivergente Lernende [Quelle: Analysebericht, Abschnitt 6].
3. Barrierefreiheitsstandards in der KI: Über WCAG 2.2 AA hinausgehen, um „multimodale Barrierefreiheit“ (Screenreader-Kompatibilität, adaptives Tempo) zu gewährleisten, die spezifisch für Berufsbildungsumgebungen ist [Quelle: Analysebericht, Abschnitt 5.2].
4. KI-Ethik und Compliance: Operationalisierung der „mehrschichtigen Erklärbarkeit“, die sicherstellt, dass Ausbilder den Lernenden erklären können, warum ein KI-Modell eine bestimmte Entscheidung getroffen hat [Quelle: Analysebericht, Abschnitt 5.3].
5. Digitales Storytelling für inklusives Lernen: Einsatz von „narrativer Handlungsfähigkeit“ zur Schaffung vielfältiger, authentischer Szenarien, die Tokenismus vermeiden und emotionale Verbindung fördern.

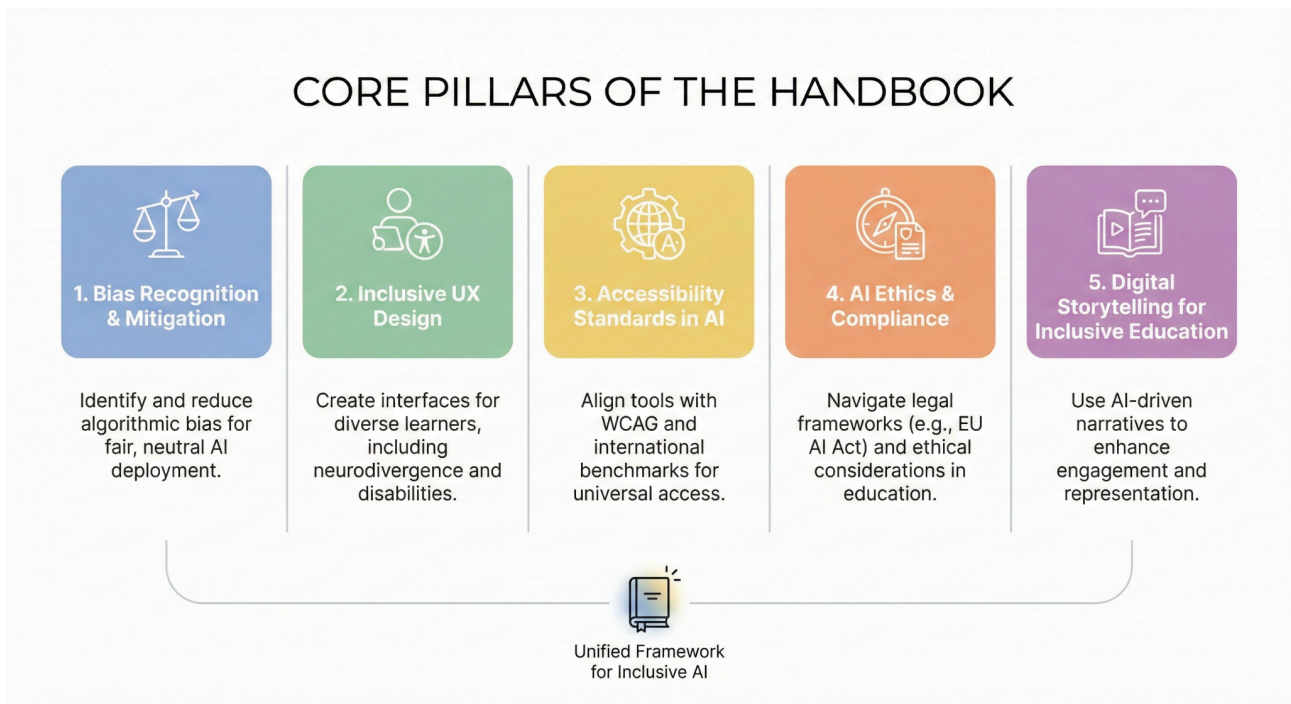


Abbildung: Kernsäulen des Handbuchs

Erwartete Wirkung

Durch die Übernahme der Rahmenwerke in diesem Leitfaden werden Berufsbildungsfachleute befähigt:

- Ihre eigenen KI-Tools vor der Einführung auf Ausschluss und Verzerrungen zu prüfen.
- Lernerfahrungen zu gestalten, die sich an unterschiedliche kognitive und physische Bedürfnisse anpassen.
- Ihre neuen Kompetenzen über den Europass-Rahmen zu zertifizieren und damit zum Priorität „Digitale Bereitschaft und Resilienz“ des Erasmus+-Programms beizutragen.

Einleitung: Zweck und Nutzungsanleitung

Zweck dieses Leitfadens Dieser Leitfaden, „Barrierefreie Künstliche Intelligenz für die Bildung“, ist das zentrale intellektuelle Produkt von Aktivität 2. Er fungiert als strategische Brücke zwischen der technischen Implementierung von KI und pädagogischer Inklusivität.

Dies ist kein rein theoretischer Text, sondern ein Werkzeugkasten, der darauf ausgelegt ist, die spezifischen Kompetenzlücken im kreativen Berufsbildungssektor zu schließen, wo unsere Forschung zeigt, dass 89 % des Personals keine formale Ausbildung in Bias-Minderung und inklusivem Design besitzen. Sein Hauptziel ist es, Berufsbildungsfachleute zu befähigen, KI-Avatare und -Narrative zu erstellen, zu trainieren und einzusetzen, die:

- Ethisch fundiert sind: Die Grundsätze von „Privacy-by-Design“ und Datenminimierung respektieren.
- Rechtskonform sind: Den Transparenzanforderungen des EU-KI-Gesetzes und der DSGVO entsprechen.
- Kulturell repräsentativ sind: „Repräsentationsgerechtigkeit“ nutzen, um Tokenismus und Stereotypen in Bildungsinhalten zu vermeiden.

Wie Sie diesen Leitfaden verwenden Dieses Dokument wurde als aktives Implementierungswerkzeug entwickelt, nicht als passives Lesematerial. Es ist dazu gedacht, gemeinsam mit dem Personalisierten Feedback-Navigator (Aktivität 1) zur Prüfung und Verbesserung realer Projekte eingesetzt zu werden.

Für Entwickler und kreative Technologen:

- Als Prüfprotokoll: Verwenden Sie die „Praxisszenarien“ als verbindliche Qualitätssicherungsprüfungen. Bevor Sie ein KI-Modul einführen, überprüfen Sie es anhand der Bias- und Barrierefreiheits-Checklisten in den Kapiteln 1 und 3.
- Für Design-Spezifikationen: Konsultieren Sie die Heuristiken für „Inklusives Nutzererlebnis“ in Kapitel 2 während der Wireframing-Entwicklungsphase, um den „Uncanny-Valley“-Effekt zu verhindern und die kognitive Belastung für neurodivergente Nutzer zu steuern.

Für Ausbilder und Lehrdesigner:



- Als Entscheidungsunterstützungssystem: Verwenden Sie „Grundlegende theoretische Konzepte“ und „Fallstudien“, um ethische Dilemmata im Unterricht zu bewältigen (z. B. Warum hat die KI diesen Schüler schlecht bewertet?).

Für Führungskräfte in Organisationen:

Als Standardisierungsrahmen: Nutzen Sie das Handbuch, um eine einheitliche „Inklusive KI-Richtlinie“ für Ihre Institution zu etablieren und sicherzustellen, dass alle digitalen Lernmaterialien den hohen Standards des Erasmus+-Programms in Bezug auf Vielfalt und Inklusion entsprechen.

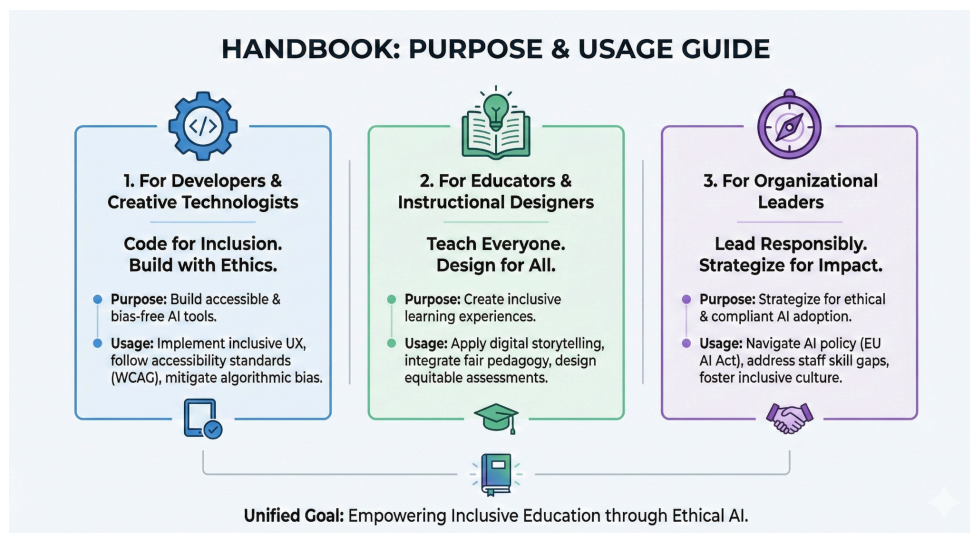


Abbildung: Leitfaden zu Zweck und Verwendung

Kapitel 1: Erkennung und Abschwächung von Verzerrungen

Kontext und Begründung

Das Problem: Die Bedarfsanalyse von D2CIN ergab, dass der Großteil des Zielpersonals Schwierigkeiten hat, unvoreingenommene Konversations-KI zu erstellen. Dies ist nicht nur eine technische Lücke; es ist ein Ausdruck einer tieferen Kompetenzlücke. Wenn sie gebeten wurden, ihre eigenen Tools zu prüfen, hatten Mitarbeiter Schwierigkeiten, die Grundursachen für verzerrte KI-Ergebnisse zu identifizieren. Darüber hinaus betonten die interdisziplinären Fokusgruppen,

insbesondere Experten für inklusive Bildung in Deutschland, dass KI-Systeme unweigerlich die Werte, Annahmen und blinden Flecken ihrer Schöpfer widerspiegeln. Wenn ein KI-Avatar mit historisch nicht repräsentativen Daten trainiert wird, produziert er nicht nur ein „neutrales“ Ergebnis – er verstärkt aktiv alte gesellschaftliche Hierarchien, marginalisiert Lernende aus nicht-dominanten Hintergründen und untergräbt ihr Selbstvertrauen. (Quelle: D2CIN Insights Report).

Das Ziel: Dieses Kapitel befasst sich mit der kritischen Unfähigkeit, Verzerrungen zu erkennen und zu korrigieren, bevor sie den Lernenden erreichen. Es bietet Berufsbildungsausbildern und -entwicklern eine klare Taxonomie der Verzerrungstypen und eine praktische, schrittweise Methodik für die Prüfung von Datensätzen. Am Ende dieses Kapitels wird der Leser in der Lage sein, diskriminierende Muster in KI-Trainingsinhalten proaktiv zu identifizieren und die Grundsätze der repräsentativen Fairness anzuwenden, um eine gerechtere Bildungserfahrung für alle Lernenden zu schaffen.

2. Lernziele (LZ)

LZ1: Identifizieren und kategorisieren Sie häufige Formen algorithmischer Verzerrungen (historisch, repräsentativ und messbedingt) in Berufsbildungsdatensätzen, um die Verstärkung von Stereotypen zu verhindern.

LZ2: Wenden Sie Strategien der „repräsentativen Fairness“ an, um sicherzustellen, dass KI-Avatare und -Stimmen vielfältige demografische Merkmale widerspiegeln, ohne auf Tokenismus zurückzugreifen.

3. Grundlegende theoretische Konzepte

Um KI-Tools effektiv auf Verzerrungen prüfen zu können, müssen Berufsbildungsfachleute zunächst das Vokabular des Problems verstehen. Verzerrungen entstehen nicht aus böswilliger Absicht; sie entstehen aus unkritischer Datenwahl. Die folgenden drei Konzepte bilden den grundlegenden Rahmen dieses Kapitels.

3.1 Datenverzerrung vs. Ausgabeverzerrung: Die Wurzel und die Frucht

Datenverzerrung bezieht sich auf systematische Verzerrungen in den Trainingsdatensätzen, aus denen ein KI-Modell lernt. Wenn ein Datensatz für KI-Filmproduktion ausschließlich aus Oscar-prämierten Hollywood-Filmen zwischen 1980 und 2010 besteht, wurde der KI beigebracht, dass „Qualitätskino“ auf eine sehr spezifische Weise aussieht, klingt und gedreht wird. Dies ist Datenverzerrung – sie ist eingebettet, bevor die KI auch nur eine einzige Antwort generiert.

Ausgabeverzerrung ist das, was der Lernende erlebt: Die KI bewertet konsequent experimentelle, nicht-westliche ästhetische Entscheidungen als „unter dem Standard“ oder ein KI-Avatar für Karriereberatung empfiehlt selbstsicher unterschiedliche Berufswege für Lernende auf Basis ihres Geschlechts. Die Ausgabe ist der Ort, an dem der Schaden entsteht, aber die Ursache liegt immer weiter oben in den Daten.

Im praktischen Kontext bedeutet dies, dass Entwickler ihre Ausgangsmaterialien mit demselben kritischen Blick prüfen müssen, den ein Lektor auf ein Manuskript anwendet. Die Frage, die vor dem Training eines KI-Modells gestellt werden muss, lautet: Wer hat diesen Inhalt erstellt? Für wen ist er gedacht? Und vor allem: Wer fehlt vollständig darin?

3.2 Kulturelle Neutralität: Balance ohne Auslöschung

Kulturelle Neutralität wird oft fälschlicherweise als das Entfernen jeglicher kultureller Spezifität aus KI-Inhalten verstanden. Dies ist sowohl unmöglich als auch unerwünscht. Echte kulturelle Neutralität im KI-Training bedeutet, eine pädagogische Balance zu erreichen: Feedback und Inhalte bereitzustellen, die global anwendbar sind, ohne kulturell vorschreibend zu sein oder zu implizieren, dass ein kultureller Standard das universelle Maß der Exzellenz ist.

In einem Modul für Grafikdesign beispielsweise könnte eine KI, die überwiegend auf westlichem Minimalismus trainiert wurde, die von der japanischen Wabi-Sabi-Ästhetik inspirierte Typografie als „übermäßig gestaltet“ bewerten. Dies ist keine neutrale Bewertung, sondern kulturell voreingenommen. Eine kulturell neutrale KI würde die Arbeit anhand ihrer eigenen erklärten ästhetischen Absicht bewerten. Für Berufsbildungsfachleute, die in einem europäischen, multikulturellen Kontext arbeiten, ist diese Unterscheidung wesentlich für die Entwicklung von Tools, die das gesamte kreative Erbe ihrer Lernenden respektieren.

3.3 Bedrohung durch Stereotypen: Die psychologischen Folgen voreingenommenen Feedbacks

Die Stereotypenbedrohung, ein Konzept, das von den Psychologen Claude Steele und Joshua Aronson entwickelt wurde, beschreibt das psychologische Risiko, bei dem eine Person, die sich eines negativen Stereotyps über ihre Gruppe bewusst ist, zusätzliche Angst erlebt, die ihre Leistung beeinträchtigt. Im Kontext des KI-Trainings – wenn eine weibliche Lernende konsequent Feedback von einer KI erhält, das sie subtil oder offen von technischen Führungsrollen abhält – ist sie nicht einfach nur anderer Meinung mit dem Vorschlag. Die Forschung zeigt, dass sie beginnen kann, die Voreingenommenheit zu internalisieren, was zu Rückzug und vermindertem beruflichen Ehrgeiz führt.

Voreingenommene KI ist kein neutrales Werkzeug, sondern ein aktiver Teilnehmer an der Selbstwahrnehmung des Lernenden als Fachkraft. Dies verleiht der Bias-Minderung eine

moralische Dringlichkeit, die über die bloße Gesetzeskonformität hinausgeht. Es geht um Bildungsgerechtigkeit.

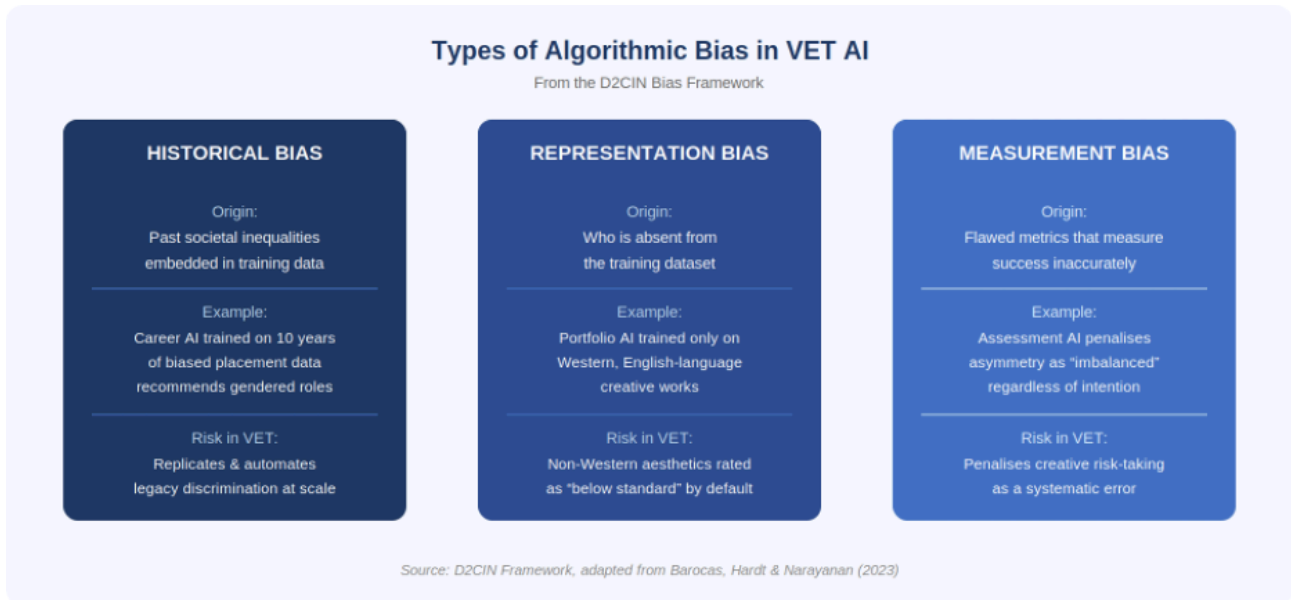


Abbildung 1.1 Die drei Arten algorithmischer Verzerrung in der KI in der Berufsbildung (D2CIN-Rahmenwerk)

4. Fallstudie „Der geschlechtssensible Karriereberater“

Das Szenario

Eine deutsche Berufsbildungseinrichtung, die auf digitale Medienproduktion spezialisiert ist, implementierte einen KI-gestützten Karriereberatungsassistenten, der Studierenden helfen sollte, ihre Berufsnische im kreativen Sektor zu identifizieren. Das Tool wurde mit zehn Jahren erfolgreicher Absolventendaten der Einrichtung trainiert. Im ersten Semester der Nutzung zeichnet sich ein Muster ab: Die KI empfiehlt weiblich gelesenen Studierenden konsequent die Rollen „Projektmanagement“ und „Kundenbeziehungen“, während sie männlich gelesene Studierende zu „Führung“, „Technischer Leitung“ und „Führungskräfteentwicklung“ lenkt. Das System war nie explizit programmiert worden, geschlechtsbezogene Unterscheidungen zu treffen, aber es hatte sie gelernt. Da die historischen Beschäftigungsdaten der Einrichtung ein Jahrzehnt unbewusster Geschlechtervorurteile in ihren eigenen Einstellungsnetzwerken widerspiegeln, wurde die KI tatsächlich darauf trainiert, ein systemisches Problem zu reproduzieren.

Warum es scheiterte

Dieser Fall veranschaulicht historische Verzerrungen in den Trainingsdaten. Die KI generiert keine diskriminierenden Ergebnisse, weil sie schlecht kodiert ist, sondern weil die Daten, aus denen sie lernt, selbst das Produkt eines voreingenommenen beruflichen Umfelds sind. Dies ist das zentrale Paradoxon algorithmischer Verzerrung: Die KI erledigt ihre Arbeit einwandfrei. Sie identifizierte Muster und optimierte Empfehlungen – aber das Muster, das sie zu optimieren gelernt hatte, war die geschlechtsbasierte Karrieresegregation.

Wichtige Erkenntnis

Ohne einen proaktiven Datensatz-Audit vor der Einführung hat die Einrichtung ein Erbe der Ungleichheit in ein Tool eingebettet und automatisiert, das den Lernenden als objektive Karriereberatung präsentiert wurde.

5. Praxisszenario (Anwendung „Danach“)

Datensatz-Audit für Repräsentationsgerechtigkeit

Dieser Audit ist eine strukturierte Qualitätssicherungsübung, die durchgeführt wird, bevor ein KI-Trainingsdatensatz oder Inhaltsskript finalisiert wird. Ziel ist es, unsichtbare Unterrepräsentierungsmuster sichtbar zu machen, damit sie korrigiert werden können, bevor das KI-Modell trainiert wird.

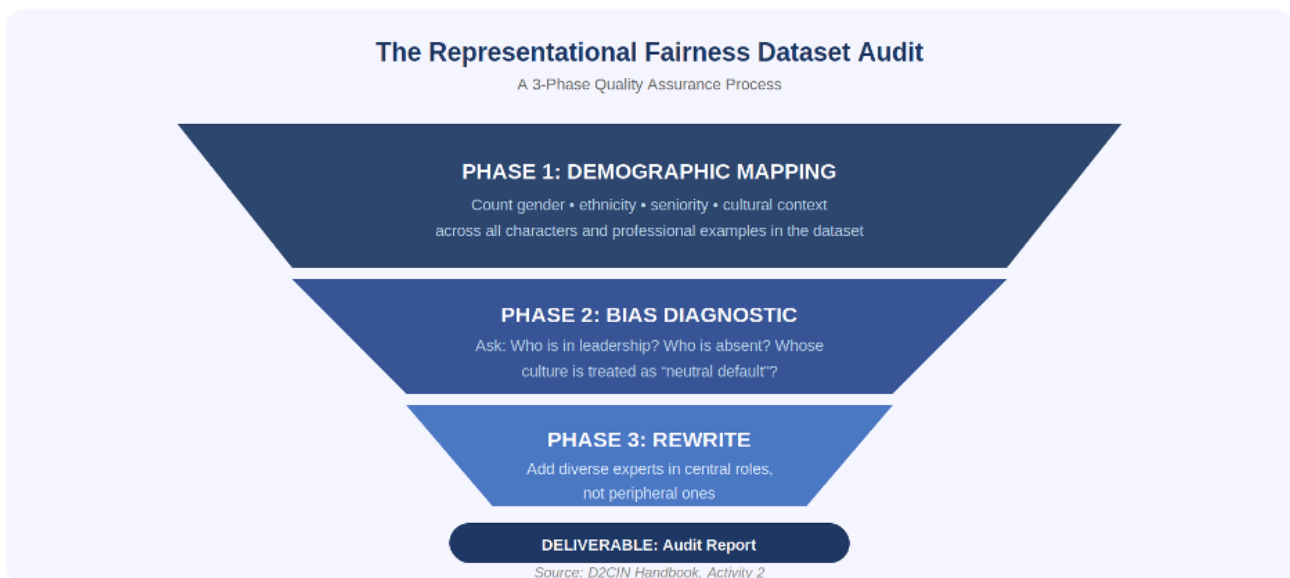


Abbildung 1.2 - 3-Phasen-Audit-Prozess für Repräsentationsgerechtigkeit

- Anweisungen

Phase 1: Demografisches Mapping (Die Bestandsaufnahme)

Überprüfen Sie Ihren Trainingsdatensatz, Ihr Skript oder Ihre Fallstudienbibliothek. Für jede erwähnte Figur, jeden Beispiel-Fachmann oder jedes Vorbild erfassen Sie die folgenden Attribute in einer einfachen Tabelle:

1. Sichtbares Geschlecht (Männlich / Weiblich / Nicht-binär / Nicht angegeben)
2. Sichtbare ethnische Zugehörigkeit oder kultureller Hintergrund (wie aus Namen, Standorten und Kontexten abgeleitet werden kann)
3. Rollen-/Dienstgradniveau (Lernende/r / Junior / Senior / Führung / Experte)
4. Geografischer oder kultureller Kontext (Europäisch / Nicht-europäisch / Universal)

Phase 2: Bias-Diagnose (Die Analyse)

Nachdem Sie das demografische Mapping abgeschlossen haben, beantworten Sie folgende Diagnosefragen:

5. Haben alle Figuren in Führungs- oder Expertenpositionen dasselbe sichtbare Geschlecht oder dieselbe ethnische Herkunft?
6. Werden Figuren aus nicht-westlichen kulturellen Kontexten nur als „zu lösende Probleme“ oder „Misserfolgsbeispiele“ dargestellt, nicht als Quellen von Expertise?
7. Fehlen bestimmte demografische Gruppen vollständig im Datensatz? Wer ist einfach nicht vertreten?
8. Gibt es Disziplinen, Kunstformen oder Berufspraktiken, die implizit als „Standard“ behandelt werden, während andere als „alternativ“ oder „spezialisiert“ behandelt werden?

Phase 3: Umschreiben (Die Korrektur)

Für jede „Ja“-Antwort in Phase 2 haben Sie einen Repräsentationsfehler identifiziert. Ihre Aufgabe besteht darin, den entsprechenden Abschnitt Ihres Datensatzes oder Szenarios zu überarbeiten. Schreiben Sie mindestens zwei Fallstudienfiguren oder Berufsbeispiele um, um eine größere demografische Vielfalt einzuführen. Wichtig zu beachten: Die vielfältigen Figuren, die Sie hinzufügen, müssen Experten- oder Führungsrollen einnehmen, keine unterstützenden oder peripheren. Dies ist der Unterschied zwischen Repräsentationsgerechtigkeit und Tokenismus.

● Produkt

Ein ausgefüllter Repräsentationsgerechtigkeits-Audit-Bericht, der Ihre ursprüngliche demografische Mapping-Tabelle, Ihre Diagnoseanalyse mit den identifizierten Verzerrungen und Ihre

überarbeiteten Inhalte mit einer kurzen schriftlichen Begründung der Entscheidungen zur Verbesserung des Repräsentationsgleichgewichts enthält.

6. Reflexionsfrage

Wenn wir KI ausschließlich auf Portfolios, Absolventendaten und Berufsbeispielen der letzten 10 Jahre trainieren, welche künstlerischen Stile, Berufswege und kulturellen Perspektiven schließen wir dann systematisch und unsichtbar aus der KI-Definition von „Exzellenz“ aus?

7. Referenzen

Steele, C.M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811.

Europäische Kommission. (2021). Vorschlag für eine Verordnung zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (KI-Gesetz). Amt für Veröffentlichungen der EU.

D2CIN-Projektkonsortium. (2025). Analyse- und Empfehlungsbericht (Aktivität 1.2). Mimic Productions und Budakov Films.

Barocas, S., Hardt, M. & Narayanan, A. (2023). *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press.

Kapitel 2: Inklusives UX-Design

Kontext und Begründung

Das Problem: Die Integration von KI in die Berufsbildung (BV) konzentriert sich oft stark auf technische Fähigkeiten und vernachlässigt die psychologische Erfahrung des Lernenden. Die Bedarfsanalyse von D2CIN ergab eine kritische Kompetenzlücke: 89 % des Zielpersonals hat keine formale Ausbildung in inklusiven KI-Praktiken und -Design. Darüber hinaus betonten unsere interdisziplinären Fokusgruppen – insbesondere UX- und Barrierefreiheits-Designer in Deutschland und Bulgarien –, dass Mikro-UX-Entscheidungen (wie der Ton der Avatar-Stimme, das Tempo und die Gesichtssymmetrie) das Vertrauen und den Komfort der Nutzer erheblich beeinflussen.

Das Ziel: Dieses Kapitel befasst sich mit dem Mangel an formaler Ausbildung für das Design von Schnittstellen, die auf neurodivergente Lernende ausgerichtet sind. Es bietet Berufsbildungsausbildern und -entwicklern anwendbare Design-Heuristiken, um die kognitive Bandbreite zu verwalten und psychologische Sicherheit zu fördern, und stellt sicher, dass KI-Schnittstellen das Lernen erleichtern, anstatt Barrieren zu schaffen.

2. Lernziele (LZ)

- LZ1: Entwerfen Sie KI-Schnittstellen, die aktiv die „kognitive Belastung“ verwalten, um eine Überlastung der Lernenden zu verhindern, und stellen Sie sicher, dass Inhalte für verschiedene kognitive Profile verdaulich sind.
- LZ2: Wenden Sie spezifische Design-Heuristiken an, um den „Uncanny-Valley“-Effekt zu minimieren, und stellen Sie sicher, dass KI-Avatare als vertrauenswürdig und unterstützend wahrgenommen werden, nicht als „unheimlich“ oder beunruhigend.

3. Grundlegende theoretische Konzepte

Um wirklich inklusive KI zu entwerfen, müssen Berufsbildungsfachleute über die grundlegende „Benutzerfreundlichkeit“ hinausgehen und die tiefere psychologische Verbindung zwischen dem Lernenden und der Maschine verstehen. Dieses Kapitel stützt sich auf drei theoretische Säulen, die während der Forschungsphase von D2CIN identifiziert wurden, und bietet eine Roadmap für die Entwicklung von Bildungstools, die den kognitiven und emotionalen Bedürfnissen verschiedener Lernender Rechnung tragen.

3.1. Theorie der kognitiven Belastung in der KI: Verwaltung der „mentalen Bandbreite“

Die Theorie der kognitiven Belastung (TKB) bezieht sich auf die begrenzte Menge an Arbeitsgedächtnisressourcen, die dem Lernenden zu einem bestimmten Zeitpunkt zur Verfügung stehen. Im Kontext von KI-Schnittstellen ist es äußerst wichtig, zwischen intrinsischer Belastung – dem Aufwand, der zum Verständnis des eigentlichen Unterrichts erforderlich ist – und extrinsischer Belastung zu unterscheiden, die den verschwendeten Aufwand darstellt, der für die Navigation in einer verwirrenden Schnittstelle oder die Verarbeitung chaotischer Visualisierungen aufgewendet wird. Für neurodivergente Lernende, insbesondere solche mit ADHS oder Legasthenie, kann eine hohe extrinsische Belastung eine sofortige Abkopplung vom Lernprozess auslösen. Eine Schnittstelle, die gleichzeitig einen sich bewegenden Avatar, scrollenden Text und blinkende Benachrichtigungssymbole präsentiert, schafft ein „sensorisches Hindernis“, das das Lernen behindert.

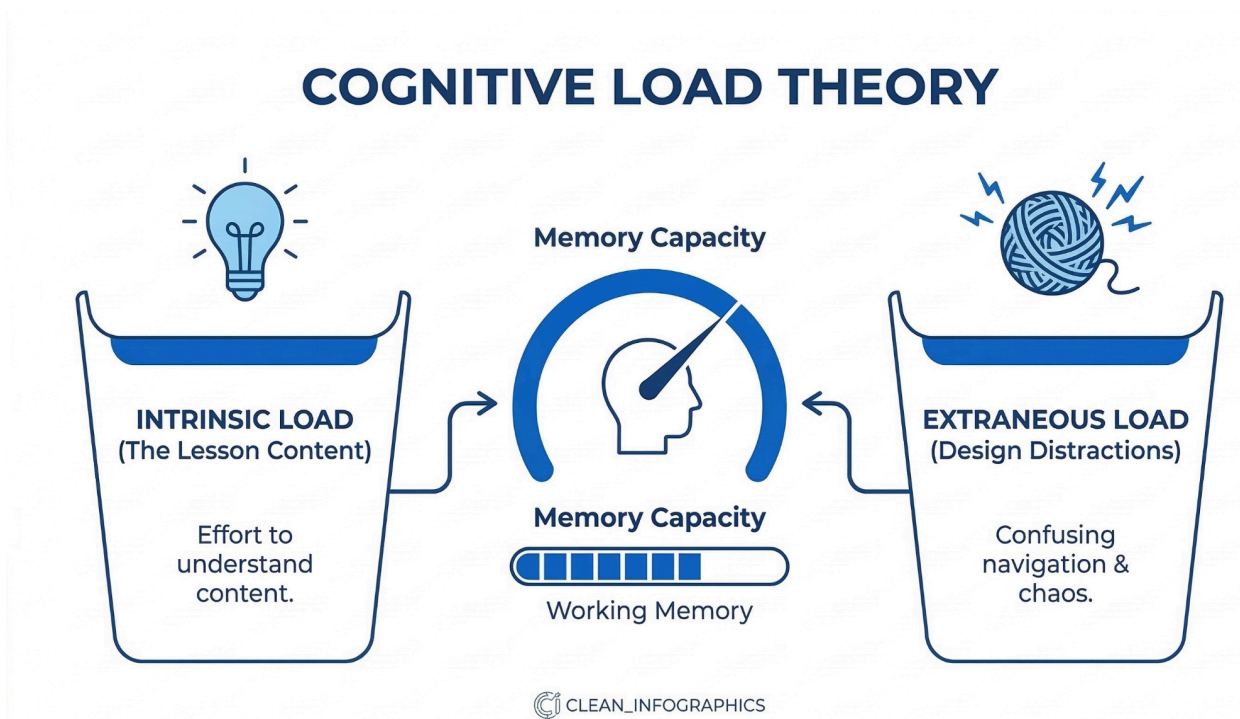


Abbildung: Theorie der kognitiven Belastung. Quelle: Gemini, 2025.

Um dem entgegenzuwirken, unterstützt die D2CIN-Forschung das Gebot einer „stillen Schnittstelle“. Wie von UX-Experten während der Fokusgruppen festgestellt wurde, erfordert gutes Design keine Aufmerksamkeit; es erleichtert die Konzentration. Entwickler sollten die kognitive

Überlastung durch übermäßige Anpassungsoptionen vermeiden. Anstatt eine riesige Reihe von Einstellungen anzubieten, sollte die Schnittstelle eine klare Auswahl zwischen „Rich-Media-Modus“ (Avatar + Stimme + Text) und „Fokusmodus“ (Nur Text + Vereinfachtes Layout) bereitstellen. Dieser Ansatz ermöglicht es dem Lernenden, seine eigene mentale Bandbreite basierend auf seinen unmittelbaren sensorischen Bedürfnissen zu regulieren.

3.2. Das Uncanny Valley: Umgang mit visuellem Unbehagen

Das „Uncanny Valley“ ist ein psychologisches Phänomen, bei dem ein humanoides Objekt, wie z. B. ein KI-Avatar, einem Menschen fast identisch ähnelt, aber auf subtile, nonverbale Weisen versagt. Mängel wie steife Augenbewegungen, verzögerte Lippensynchronisation oder fehlende Mikro-Ausdrücke können beim Benutzer ein Gefühl von Unbehagen, Abscheu oder einem „unheimlichen“ Effekt verursachen. In einem Berufsausbildungsszenario hat Vertrauen oberste Priorität; wenn ein Avatar im Uncanny Valley landet, hört der Lernende auf, den Bildungsinhalt zu verarbeiten, und fixiert sich auf die Künstlichkeit des Avatars. Der „Lehrer“ wird effektiv zu einer Quelle der Ablenkung und nicht der Unterstützung.

Daher müssen Designer die „Realismus-Falle“ navigieren. Obwohl es oft einen Drang nach Hyperrealismus gibt, warnen Experten vor dieser Verfolgung, wenn das Budget keine Perfektion zulässt. Projektergebnisse zeigen, dass stilisierte oder abstrakte Avatare – wie illustrierte Charaktere – oft unvollkommene fotorealistische übertreffen. Indem sie die Erwartungen des Benutzers an streng menschliches Verhalten senken, umgehen diese stilisierten Formen das Uncanny Valley vollständig. Daher schafft es, wenn Ressourcen begrenzt sind, eine komfortablere und effektivere Lernumgebung, wenn man der verhaltensbezogenen Konsistenz gegenüber visuellem Realismus Vorrang einräumt.

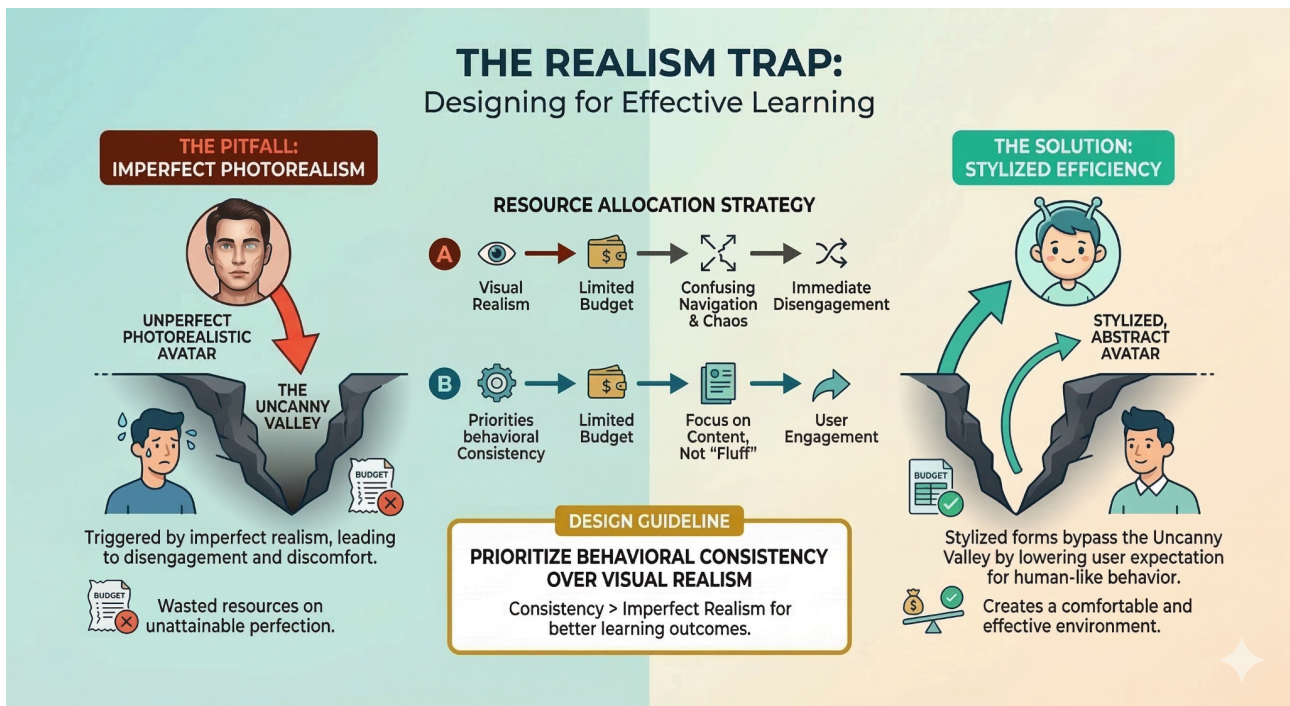


Abbildung: „Die Realismus-Falle“. Quelle: Gemini, 2025

3.3. Vorhersehbare Interaktionen und mehrschichtige Erklärbarkeit

Dieses Konzept verbindet UX-Design und KI-Ethik und postuliert, dass das Vertrauen des Lernenden durch zwei Schlüsselmechanismen aufgebaut wird: Vorhersehbarkeit (die KI verhält sich konsequent) und Erklärbarkeit (der Benutzer versteht, warum die KI auf eine bestimmte Weise gehandelt hat). KI-Systeme in der Bildung verwenden oft adaptive Inhaltssalgorithmen, um Schwierigkeitsgrade oder Feedback-Stile in Echtzeit zu ändern. Wenn diese Änderungen jedoch ohne Kontext auftreten, können Lernende – insbesondere ängstlichere – das Gefühl haben, ungerecht beurteilt oder manipuliert zu werden.

Um „psychologische Sicherheit“ zu fördern, sollte die Schnittstelle plötzliche, unerklärliche Änderungen im Ton oder in der Schwierigkeit vermeiden. Die Lösung liegt in der „mehrschichtigen Erklärbarkeit“ – einer Designstrategie, die vielschichtige Transparenz bietet. Auf Anfängerniveau erhält der Lernende eine einfache, nicht-technische Erklärung für die KI-Anpassung, wie zum Beispiel: „Ich habe bemerkt, dass Sie eine lange Pause gemacht haben, also bin ich zu einem einfacheren Vokabular übergegangen.“ Auf der mittleren Ebene ist ein detailliertes technisches Protokoll verfügbar, in dem Ausbilder den Entscheidungsprozess der KI prüfen können. Letztendlich sollte die Anpassung nie verborgen sein; der Lernende muss immer die Freiheit behalten, KI-Vorschläge abzulehnen, und sicherstellen, dass er die Kontrolle über seine Bildungsreise behält.



Abbildung: Psychologische Sicherheit in der KI

4. Fallstudie „Der nicht blinkende Lehrer“

Das Szenario

Um Stakeholder zu beeindrucken, erstellte eine Berufsschule einen fotorealistischen KI-Avatar für ein Modul zum Kundenservice. Er sah unglaublich menschlich aus, aber technische Einschränkungen verursachten zwei subtile Mängel: Der Avatar hielt intensiven, ununterbrochenen Augenkontakt, und seine Lippensynchronisation war nur um 200 Millisekunden verzögert.

Das Ergebnis

Während des Pilotprojekts lernten die Schüler nicht. Anstatt sich auf das Szenario zu konzentrieren, fühlten sie sich „beobachtet“ und ängstlich. Ihre Gehirne fixierten sich auf den unnatürlichen Blick des Avatars und seinen nicht synchronisierten Mund.

Warum es scheiterte (Die Wissenschaft des Lernens)

Dies ist ein klassisches Beispiel für das Uncanny Valley. Wenn eine KI fast menschlich aussieht, aber bei kleinen Dingen versagt (wie z. B. nicht zu blinken), hört sie auf, ein „Lehrer“ zu sein, und wird zu einer Quelle der Angst.

Aus kognitiver Sicht schuf dies unnötige kognitive Belastung. Die mentale Energie der Lernenden wurde damit verbraucht, die visuellen Fehler zu verarbeiten, nicht den Inhalt der Lektion.

Wichtige Erkenntnis

Besser abstrakt als unvollkommen real. Wenn Sie keinen 100%igen Realismus erreichen können, verwenden Sie einen stilisierten oder illustrierten Charakter. Konsistenz baut Vertrauen auf; „unheimlicher“ Realismus zerstört es.

5. Praxisszenario (Anwendung „Danach“) [Anwendbares Werkzeug (200-300 Wörter). Dies ist der Kernwert des Handbuchs. Stellen Sie eine konkrete Übung, Checkliste oder ein Audit-Tool bereit, das der Leser sofort verwenden kann.]

Ziel

Um von der Theorie zur Praxis überzugehen, bietet dieser Abschnitt ein standardisiertes Qualitätssicherungsprotokoll (QS). Bevor Entwickler und Lehrdesigner ein KI-basiertes Bildungstool fertigstellen, sollten sie diesen neuro-inklusiven Schnittstellenaudit durchführen. Dieser Prozess stellt sicher, dass das Tool die kognitive Verarbeitung des Lernenden unterstützt und nicht überlastet.

Anweisungen

Überprüfen Sie Ihren aktuellen KI-Prototyp anhand der folgenden vier Kriterien. Wenden Sie für jedes „Nein“ die vorgeschlagene Korrekturstrategie an.

Phase 1: Audit-Checkliste

1. Kontrollprüfung (Benutzeraktivität und Tempo)

Die Frage: Hat der Lernende absolute Kontrolle über die Präsentation des Avatars? Kann er jederzeit pausieren, wiederholen oder das Sprechtempo verlangsamen (z. B. 0,75-fache Geschwindigkeit)?

Begründung: Angst entsteht oft aus einem Mangel an Kontrolle. Das Ermöglichen, dass Lernende den Informationsfluss regulieren, reduziert die „Angst vor dem Verpassen“ wichtiger Inhalte.

Maßnahme: Implementieren Sie eine sichtbare, immer aktive Mediensteuerungsleiste.

2. „Visuelles Rauschen“-Prüfung (Geteilte Aufmerksamkeit)

Die Frage: Führt der Avatar „Umgebungsbewegungen“ (Nicken, Gewichtsverlagerung, schnelles Blinzeln) aus, während wichtiger Text auf dem Bildschirm angezeigt wird?

Begründung: Das menschliche Gehirn hat Schwierigkeiten, sich bewegende Bilder und statischen Text gleichzeitig zu verarbeiten. Umgebungsbewegungen lenken das Auge ab und erhöhen die unnötige kognitive Belastung.

Maßnahme: Programmieren Sie den Avatar so, dass er in einen „neutralen/stationären Zustand“ wechselt, wenn textreiche Folien präsentiert werden.

3. Ton“-Prüfung (Emotionale Anpassungsfähigkeit)

Die Frage: Ermöglicht die Schnittstelle dem Benutzer, zwischen der Persönlichkeit oder dem Kommunikationsstil des Avatars zu wechseln? (z. B. Umschalten von „Streng/Professionell“ zu „Unterstützend/Leger“).

Begründung: Verschiedene Lernende benötigen unterschiedliche emotionale Unterstützung. Neurodivergente Lernende könnten den direkten, autoritären Ton als zu stimulierend empfinden, während selbstsichere Lernende vielleicht Effizienz bevorzugen.

Maßnahme: Fügen Sie eine einfache „Persönlichkeitseinstellung“ in die Onboarding-Phase ein.

4. „Uncanny Valley“-Prüfung (Stilisierungsüberprüfung)

Die Frage: Entspricht der visuelle Realismus des Avatars seinen technischen Merkmalen? Verzögert sich die Lippsynchronisation insbesondere um mehr als 100 ms?

Begründung: Hoher visueller Realismus erzeugt hohe Erwartungen an das Verhalten. Wenn die Synchronisation nicht stimmt, wird der Realismus zu einem Nachteil (Uncanny Valley).

Maßnahme: Wenn die technische Latenz nicht behoben werden kann, reduzieren Sie das Niveau des visuellen Realismus. Wechseln Sie zu einem stilisierten oder illustrierten Charakter, um die Erwartungen der Benutzer an die technische Realität anzupassen.

Phase 2: Das Produkt

Ergebnis: UX-Optimierungsbericht

Nach Abschluss des Audits sollte das Designteam einen kurzen Bericht erstellen, der mindestens drei spezifische Designänderungen dokumentiert, die zur Anpassung an neurodivergente Lernende vorgenommen wurden. Dies stellt sicher, dass Inklusion ein aktiver Bestandteil des Entwicklungslogbuchs ist und keine nachträgliche Überlegung.

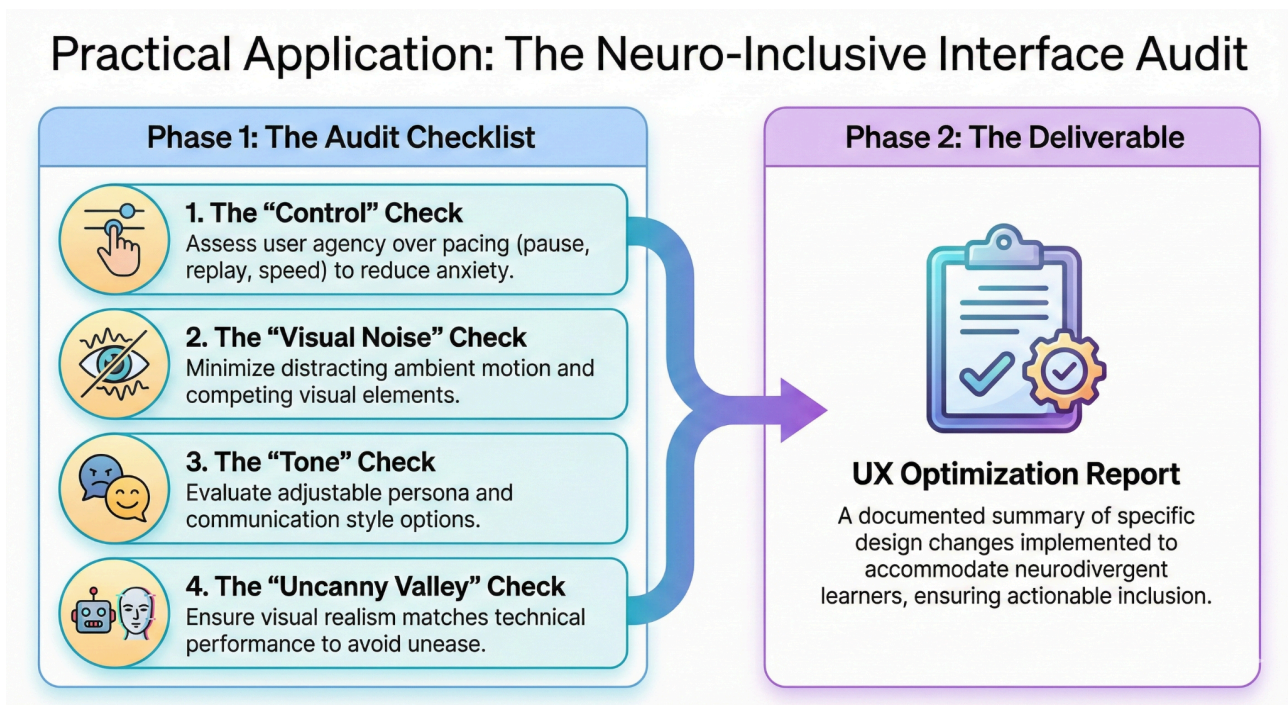


Abbildung: Neuro-inklusive Schnittstellenaudit.

6. Reflexionsfrage

Fügt der visuelle Realismus Ihres KI-Avatars tatsächlichen Bildungswert hinzu, oder ist er lediglich ein technisches Element, das den Lernenden vom Lernmaterial ablenkt?

7. Referenzen

Kahneman, D. (2011). Thinking, Fast and Slow. New York, NY: Farrar, Straus and Giroux.

Doshi-Velez, F. & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608. <https://arxiv.org/abs/1702.08608>

Kapitel 3: Barrierefreiheitsstandards in der KI

Kontext und Begründung

Das Problem: Die Bedarfsanalyse von D2CIN ergab, dass viele kreative Berufsbildungsfachleute Schwierigkeiten haben, beim Entwickeln von KI-basierten Konversationstools internationale Barrierefreiheitsstandards zu erfüllen. Darüber hinaus hoben Fokusgruppen-Interviews mit UX- und Barrierefreiheitsexperten eine kritische Herausforderung hervor: Barrierefreiheit in der Berufsbildung muss über die grundlegende WCAG-Konformität hinausgehen. Designer vernachlässigen häufig multimodale Bereitstellung, die Notwendigkeit, vielfältige kognitive und sprachliche Profile zu unterstützen, und die Realität von Lernumgebungen mit niedriger Bandbreite.

Das Ziel: Dieses Kapitel bietet Berufsbildungsausbildern und -entwicklern praktische Methoden zur Umsetzung umfassender multimodaler Barrierefreiheit. Es schließt die Konformitätslücke durch umsetzbare Schritte, um sicherzustellen, dass KI-Avatare und adaptive Inhalte ethisch gestaltet, transparent und vollständig zugänglich für alle Lernenden sind, unabhängig von physischen, kognitiven oder umgebungsbedingten Barrieren.

2. Lernziele (LZ)

LZ1: Bewertung und Abstimmung KI-basierter Bildungstools mit den WCAG 2.2 AA-Standards und den Grundsätzen multimodaler Barrierefreiheit (z. B. Integration synchronisierter Untertitel, Tastaturnavigation und Screenreader-Kompatibilität).

LZ2: Anwendung transparenter Designstrategien auf adaptive KI-Inhalte, um sicherzustellen, dass automatisierte Anpassungen im Lehrplan des Lernenden sichtbar, erklärbar und unter der Kontrolle des Benutzers bleiben.

3. Grundlegende theoretische Konzepte

Konzept A: Multimodale Barrierefreiheit

Die einfache Definition: Multimodale Barrierefreiheit bedeutet, sich nie nur auf eine Möglichkeit der Informationsübermittlung zu verlassen. Es ist die Praxis, denselben Bildungsinhalt gleichzeitig über mehrere sensorische Kanäle bereitzustellen – visuell (Sehen), auditiv (Hören) und kognitiv (Lesen/Verarbeiten).

Anwendbarkeit auf KI in der Berufsbildung: Denken Sie daran als „Redundanzregel“. In einem Berufsausbildungsraum könnten Sie Lernende mit Hörbeeinträchtigung haben oder einfach einen

Lernenden, der versucht, in einem lauten Zug zu lernen. Wenn Ihr KI-Avatar eine Anweisung laut ausspricht, sollte das System gleichzeitig synchronisierte, genaue Untertitel (Text) bereitstellen. Darüber hinaus sollten die visuellen Gesten des Avatars die Botschaft unterstützen, ohne abzulenken oder zu überlasten. Echtes multimodales Design bedeutet auch, dass wenn ein hochauflösender 3D-Avatar aufgrund einer langsamen Internetverbindung (niedrige Bandbreite) nicht geladen werden kann, das System automatisch eine leichte, textbasierte Alternative bereitstellt, damit das Lernen ohne Unterbrechung fortgesetzt werden kann.

(AT) Konzept B: Kompatibilität mit Assistiver Technologie (AT)

Die einfache Definition: Assistive Technologie bezieht sich auf externe Software oder Hardware, die Lernende mit Behinderungen bereits verwenden, um sich in der digitalen Welt zurechtzufinden (z. B. Screenreader, Sprachbefehlssoftware oder spezialisierte Tastaturen). Kompatibilität bedeutet, dass Ihr KI-Tool mit diesen vorhandenen Geräten gut „funktioniert“, ohne sie zu beeinträchtigen.

Anwendbarkeit auf KI in der Berufsbildung: Es reicht nicht aus, dass ein KI-Tool intelligent ist; es muss internationalen Standards wie WCAG 2.2 AA entsprechen. Wenn zum Beispiel ein sehbeeinträchtigter Lernender einen Screenreader verwendet, um durch Ihr Berufsbildungsmodul zu navigieren, sollte die KI-generierte Avatar-Stimme nicht den Screenreader-Sound überlagern und „übertönen“. Darüber hinaus sollte die gesamte KI-Konversationsschnittstelle einfach nur mit der Tastatur (ohne Maus) navigierbar sein, was sicherstellt, dass Lernende mit motorischen Beeinträchtigungen vollständig am Unterricht teilnehmen können.

Konzept C: Transparente adaptive Inhalte

Die einfache Definition: „Adaptive Inhalte“ bedeuten, dass die KI die Lektion automatisch basierend darauf ändert, wie gut der Schüler abschneidet – zum Beispiel indem sie die Schwierigkeit eines Tests verringert, wenn der Schüler Schwierigkeiten hat, oder sein Tempo ändert. „Transparenz“ bedeutet, dass die KI dies nie im Verborgenen tut; sie teilt dem Lernenden immer mit, was sie ändert und warum.

Bedeutung in der Berufsbildungs-KI: KI sollte nie eine „Black Box“ sein, die den Bildungsweg des Lernenden ohne sein Wissen manipuliert. Versteckte Anpassungen können zu Frustration und Vertrauensverlust führen. Berufsbildungslernende müssen verstehen, warum sich ihre Inhalte geändert haben. Zum Beispiel sollte die KI explizit sagen: „Ich habe bemerkt, dass Sie mit diesem Designkonzept Schwierigkeiten haben, deshalb habe ich den Text für Sie vereinfacht.“ Am wichtigsten ist, dass das System Personalisierung mit Benutzerautonomie ausbalancieren muss – was bedeutet, dass der Lernende immer eine sichtbare Schaltfläche haben muss, um KI-Korrekturen rückgängig zu machen und sein eigenes Schwierigkeitsniveau zu wählen. Dies steht

im Einklang mit der „Mehrschichtigen Erklärbarkeit“ und bietet den Schülern klare, menschenlesbare Erklärungen, während gleichzeitig detaillierte technische Protokolle für Ausbilder verfügbar sind.

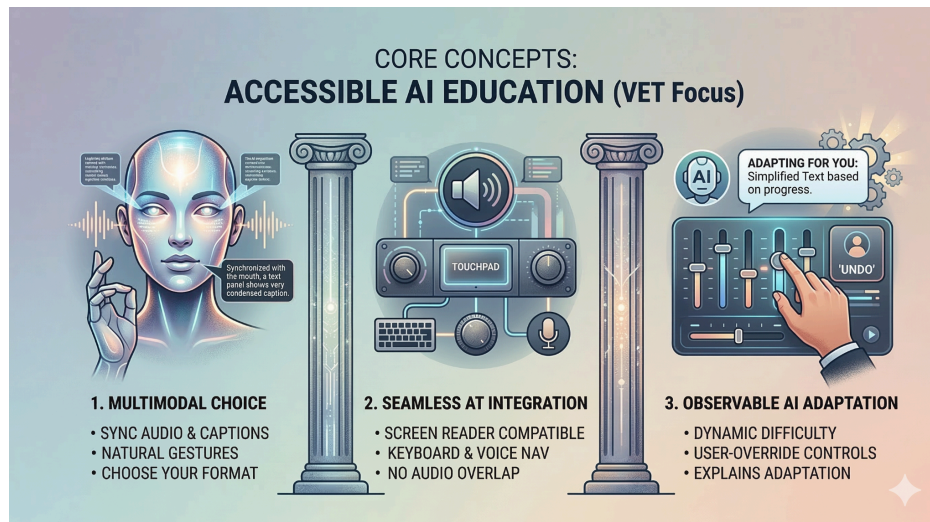


Abbildung: Multimodale Barrierefreiheit, Quelle: Gemini 2025

4. Häufige Fallstudie (Das „Vorher“-Szenario)

- **Titel: Der stille Reisende**

- **Szenario:** Eine kreative Berufsbildungseinrichtung (BV) entwickelte einen ausgefeilten, KI-gesteuerten Konversationsavatar, um Grafikdesign-Studierende durch komplexe 3D-Modellierungslektionen zu führen. Um ihre technologischen Fähigkeiten zu demonstrieren, verließen sich die Entwickler hauptsächlich auf hochwertige Sprachsynthese und Echtzeit-Sprachbewertung. Die Anwendung wurde jedoch ohne Untertitel, Tastaturnavigation oder rein textbasierte Alternative erstellt. Während der Pilotphase hörte ein Studierender, der versuchte, das Modul in einem lauten Pendelzug ohne Kopfhörer zu absolvieren, die Anweisungen überhaupt nicht. Ein anderer Lernender, der auf einen Screenreader angewiesen war, stellte fest, dass die grafische Überlagerung des interaktiven Avatars kritische Navigationsschaltflächen blockierte, was die Anwendung unbrauchbar machte.

- **Analyse:** Dieses Szenario veranschaulicht ein schwerwiegendes Versagen in der multimodalen Barrierefreiheit und der Kompatibilität mit assistiver Technologie. Die Entwickler haben die Anwendung ausschließlich für eine „perfekte“ sensorische Umgebung und einen Nutzer ohne körperliche Einschränkungen konzipiert. Indem sie situationsbedingte Beeinträchtigungen (lauter

Zug) und dauerhafte Behinderungen (Sehbehinderung) ignorierten, schufen sie ein Tool, das Lernende aktiv ausschloss und den Grundgedanken einer inklusiven Berufsbildung untergräbt.

5. Praxisszenario (Die „Danach“-Anwendung)

1. Aktivitätstyp: Barrierefreiheits-Audit-Checkliste für multimodale Berufsbildung
2. Anweisungen: Bevor Sie Ihr KI-Trainingsmodul fertigstellen oder einführen, führen Sie diesen 5-Punkte-Stresstest durch, um seine Inklusivität zu bewerten. Sie müssen Ihren Prototyp physisch unter folgenden eingeschränkten Bedingungen testen:
3. Test „Ton aus“: Schalten Sie alle Audio-Ausgaben aus. Kann der Lernende die gesamte KI-Interaktion erfolgreich nur mit synchronisierten Untertiteln oder Texttranskripten abschließen?
4. Test „Maus aus“: Trennen Sie Ihre Maus. Können Sie zum KI-Avatar navigieren, ein Gespräch beginnen und Antworten nur mit den Tab-, Enter- und Leertasten senden?
5. Test „Screenreader“: Aktivieren Sie einen grundlegenden Screenreader (z. B. NVDA, VoiceOver). Liest die Software die Texteingaben der KI klar vor oder kollidiert sie mit dem generierten Sprachklang des Avatars?
6. Test „Niedrige Bandbreite“: Simulieren Sie eine langsame 3G-Netzwerkverbindung. Wenn der hochauflösende 3D-Avatar nicht lädt, wird dem Lernenden automatisch eine leichte, textbasierte Fallback-Option bereitgestellt?

Test „Anpassungstransparenz“: Wenn die KI die Schwierigkeit des Tests aufgrund schlechter Leistung reduziert, erscheint eine klare Benachrichtigung, die erklärt, warum dies geschehen ist, zusammen mit einer Schaltfläche, die dem Benutzer die Möglichkeit bietet, zum vorherigen Zustand zurückzukehren?



Abbildung: 5-Punkte-Audit zur physischen Barrierefreiheit

Produkt: Ein verifizierter Barrierefreiheits-Compliance-Bericht, der alle fehlgeschlagenen Tests auflistet, begleitet von einer Liste der spezifischen technischen Korrekturen, die zur Einhaltung der holistischen Inklusionsstandards in der Berufsbildung erforderlich sind.

6. Reflexionsfrage

Wenn die fortschrittlichste Funktion Ihrer KI (wie 3D-Rendering oder Spracherkennung) plötzlich ausfällt oder nicht geladen werden kann, kann der gefährdetste Schüler im Klassenzimmer das grundlegende Lernziel dennoch erreichen?

7. Referenzen

World Wide Web Consortium. (2023). Web Content Accessibility Guidelines (WCAG) 2.2. <https://www.w3.org/TR/WCAG22/>

Shadi Abou-Zahra. (2018). How People with Disabilities Use the Web. World Wide Web Consortium (W3C). <https://www.w3.org/WAI/people-use-web/>

European Commission. (2024). Ethics Guidelines for Trustworthy AI. Amt für Veröffentlichungen der Europäischen Union.

<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Batya Friedman, B., & David G. Hendry, D. G. (2019). Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge, MA: MIT Press.

Kapitel 4: KI-Ethik und Compliance

1. Kontext und Begründung

Das Problem: Mit dem Inkrafttreten des EU-KI-Gesetzes ist die ethische und rechtliche Compliance im Bereich der Künstlichen Intelligenz keine Option mehr, sondern eine berufliche Pflicht. Die Bedarfsanalyse von D2CIN identifizierte jedoch eine erhebliche Lücke: Der Großteil des Zielpersonals hat keine formale Ausbildung in KI-Ethik. Anstatt eines abstrakten philosophischen Defizits äußert sich dies in realen, alltäglichen Entscheidungen. Fachleute wissen nicht, welche Daten sie legal erheben dürfen, wie sie die Grenzen der KI an Lernende kommunizieren sollen oder wie sie reagieren sollen, wenn KI zu einem Ergebnis führt, das der Schüler als ungerecht empfindet. Darüber hinaus betonten die D2CIN-Fokusgruppen – insbesondere Ethik- und Bildungsfachleute in Deutschland –, dass die zerstörerischste Kraft in der KI-unterstützten Bildung die Intransparenz ist. Wenn Lernende nicht verstehen können, warum eine KI eine Entscheidung über ihre Arbeit oder ihren Weg getroffen hat, bricht das Vertrauen vollständig zusammen (Quelle: D2CIN Insights Report, Abschnitt 5.3).

Das Ziel: Dieses Kapitel überführt KI-Ethik von einem abstrakten Compliance-Rahmen in eine Reihe konkreter, täglicher Praktiken. Es befähigt Berufsbildungsfachleute, zwei Kernprinzipien – mehrschichtige Erklärbarkeit und Privacy-by-Design – direkt in ihre Projektarbeitsabläufe zu integrieren. Am Ende dieses Kapitels wird der Leser in der Lage sein, KI-bezogene Entscheidungen transparent an Lernende zu kommunizieren, rechtlich fundierte Datenzustimmungskommunikationen zu verfassen und KI-Interaktionen zu gestalten, die institutionelles Vertrauen aufbauen, anstatt es zu untergraben.

2. Lernziele (LZ)

- LZ1: Anwendung von „Mehrschichtiger Erklärbarkeit“, um Lernenden und Administratoren zu helfen zu verstehen, warum ein KI-Modell eine bestimmte Bewertung oder Empfehlung gegeben hat.
- LZ2: Anwendung der „Privacy-by-Design“-Prinzipien, um die Datenerhebung zu minimieren und die Zustimmung der Lernenden zu Sprach- und Verhaltensdaten zu verwalten.

3. Grundlegende theoretische Konzepte

Ethische KI in der Bildung ist keine Compliance-Checkliste, die einmal vor dem Start ausgefüllt werden muss. Es ist eine Philosophie des kontinuierlichen Designs, die in jede Entscheidung darüber eingebettet werden muss, wie ein KI-Tool Daten erhebt, Entscheidungen kommuniziert und die Machtdynamik zwischen System und Lernendem verwaltet.

3.1 Mehrschichtige Erklärbarkeit: Vertrauen durch Transparenz gewinnen

Mehrschichtige Erklärbarkeit ist ein Designprinzip, das verlangt, dass KI-Systeme zielgruppengerechte, mehrschichtige Erklärungen für ihre Entscheidungen anbieten. Die Prämisse ist klar: Verschiedene Stakeholder benötigen unterschiedliche Informationstiefen, um das Gefühl zu haben, dass die Entscheidung der KI legitim und fair ist.

In der Praxis schafft dies zwei Erklärungsebenen. Die erste Ebene ist für den Lernenden bestimmt: eine einfache, nicht-technische, menschenlesbare Aussage, die die Handlung der KI in verständlicher Sprache erklärt. Die zweite Ebene ist für Ausbilder und Administratoren bestimmt: ein detailliertes, überprüftes Protokoll der spezifischen Datenpunkte, Gewichtungskriterien und Modellparameter, die zum Ergebnis geführt haben.

Ohne mehrschichtige Erklärbarkeit ist die Bewertung der KI kein Feedback, sondern ein Urteil. Lernende, die nicht verstehen können, warum sie eine bestimmte Note erhalten haben, können sich nicht verbessern. Wichtiger noch: Ohne ein sekundäres Prüfprotokoll können Ausbilder nicht identifizieren, wann die KI einen systematischen Fehler gemacht hat, und können ihn nicht korrigieren.

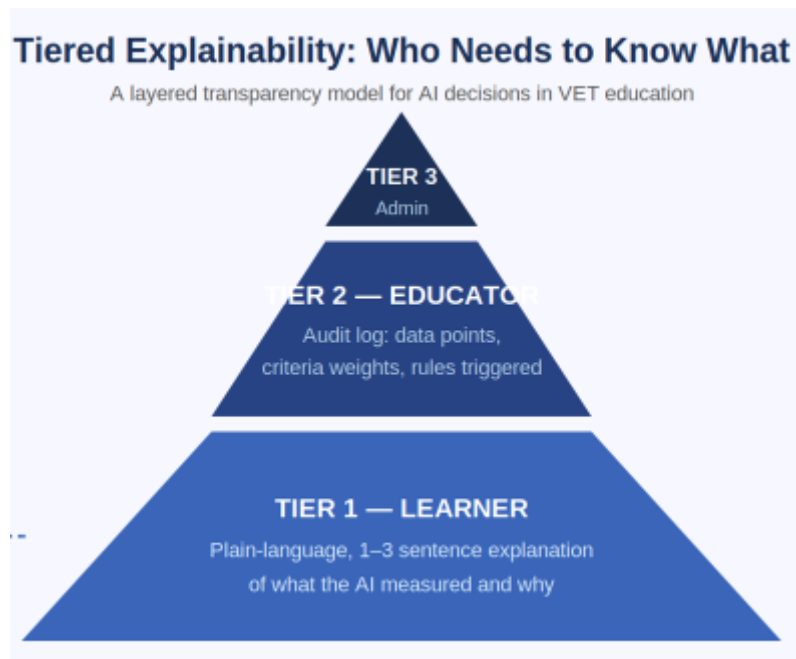


Abbildung 4.1 - Mehrschichtige Erklärbarkeit: Drei Transparenzebenen der KI für die Berufsbildung

3.2 Datenminimierung: Nur das Notwendige erheben

Datenminimierung ist ein Grundprinzip der EU-Datenschutz-Grundverordnung (DSGVO) und spiegelt sich im risikobasierten Rahmen des EU-KI-Gesetzes wider. Es besagt, dass ein KI-System nur die Daten erheben sollte, die streng und nachweislich für seine erklärte Bildungsfunktion notwendig sind. Jeder zusätzlich erhobene Datenpunkt ist eine zusätzliche ethische und rechtliche Verantwortung.

Im Kontext von KI-Tools in der Berufsbildung (BV) hat dieses Prinzip unmittelbare praktische Konsequenzen. Eine KI für Aussprachetraining, die die Stimme des Lernenden hört, um Feedback zu Betonung und Intonation zu geben, erfordert nicht, dass diese Stimme dauerhaft gespeichert wird. Die akustische Analyse kann in Echtzeit durchgeführt, das Ergebnis generiert und das Audio verworfen werden.

Viele BV-Entwickler erheben umfangreiche Verhaltensdaten in der Annahme, dass mehr Daten bessere KI produzieren. Datenminimierung hinterfragt diese Annahme und fragt: Ist die Bildungsverbesserung den Preis der Privatsphäre wert?

3.3 Transparenz: Die Grenze zwischen Mensch und KI aufrechterhalten

Transparenz im Kontext der KI-Ethik bezieht sich auf die Verpflichtung, Benutzer klar und konsequent darüber zu informieren, dass sie mit einem automatisierten System interagieren und nicht mit einem menschlichen Wesen. Das EU-KI-Gesetz behandelt dies ausdrücklich in seinen

Bestimmungen zu KI-Systemen, die direkt mit natürlichen Personen interagieren, indem es verlangt, dass Benutzer informiert werden, dass sie mit KI interagieren, es sei denn, dies ist „aus dem Kontext offensichtlich“. In einem Bildungsumfeld ist der Kontext selten offensichtlich.

Transparenz ist nicht nur eine rechtliche Pflicht, sondern ein wesentliches Element der informierten Einwilligung des Lernenden. Eine KI, die sich als menschlicher Mentor ausgibt, auch implizit durch Designentscheidungen wie einen menschlichen Namen, ein fotorealistentes Gesicht und emotionale Sprache, manipuliert Vertrauensbeziehungen auf eine Weise, die die Integrität der Bildungserfahrung untergräbt.

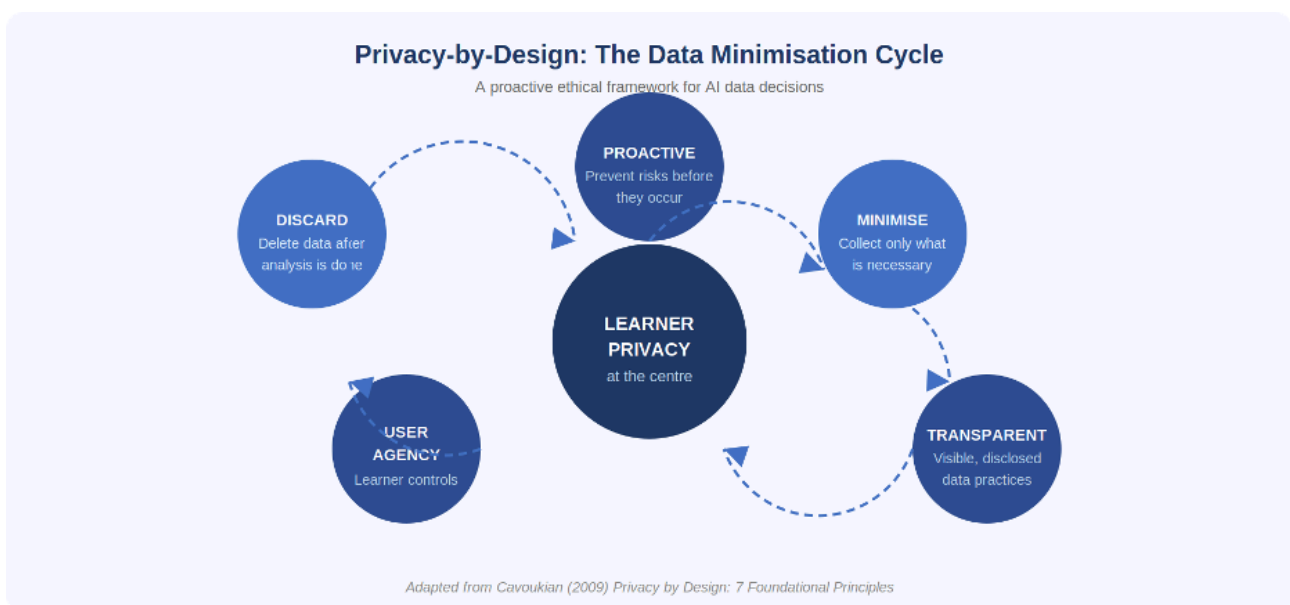


Abbildung - Der Privacy-by-Design-Zyklus: Fünf Prinzipien für ethisches Datenmanagement

4. Fallstudie (Das „Vorher“-Szenario)

- Titel: „Die Black-Box-Bewertung“

Szenario

Eine Designfachschule in Bulgarien führte ein KI-gestütztes Bewertungstool zur Beurteilung von Studierenden-Grafikdesignprojekten ein. Am Ende jedes Moduls erhielten die Studierenden eine numerische Note und eine Ein-Wort-Bewertung: „Ausgezeichnet“, „Großartig“ oder „Entwicklung“. Kein spezifisches Feedback wurde gegeben. Keine Kriteriengewichtung wurde offengelegt. Ein Studierender, dessen Projekt von seinem Ausbilder als besonders innovativ für seinen

unkonventionellen Einsatz von negativem Raum hervorgehoben worden war, erhielt die Bewertung „Entwicklung“ von der KI. Da er nicht verstehen konnte, warum ein Ansatz, den sein Ausbilder gelobt hatte, eine schlechte Bewertung generierte, reichte der Studierende eine formelle Beschwerde ein. Die Überprüfung ergab, dass das Kompositionsanalysemodul der KI Asymmetrie ungeachtet ihrer Absichtlichkeit als „unausgewogen“ sanktioniert hatte – eine in den Trainingsdaten eingebettete Designverzerrung. Da es jedoch keine Erklärungsebene gab, blieb dieser systematische Fehler drei volle Semester lang unbemerkt.

Warum es scheiterte

Dieses Szenario unterstreicht das doppelte Versagen eines KI-Systems ohne mehrschichtige Erklärbarkeit. Erstens wurde der Lernende geschädigt: Er erhielt eine Note, die er nicht verstehen, nicht davon lernen und nicht anfechten konnte. Zweitens wurde die Institution geschädigt: Ein systematischer KI-Fehler, der kreatives Risikobereitschaft sanktionierte – möglicherweise die wertvollste Fähigkeit im Kontext kreativer Berufsbildung – blieb monatelang unbemerkt, weil es keinen Prüfpfad gab. Die unerklärliche Entscheidung der KI ist nicht nur nutzlos, sondern ethisch nicht vertretbar.

5. Praxisszenario (Die „Danach“-Anwendung)

Übung zum ethischen Design der KI-Kommunikation

Diese zweistufige Übung fordert den Leser auf, die Transparenz- und Einwilligungsinfrastruktur für ein KI-gestütztes Berufsbildungstool zu entwerfen. Sie operationalisiert sowohl mehrschichtige Erklärbarkeit als auch Privacy-by-Design in konkreten, aufzeichenbaren Ergebnissen.

- Anweisungen

Teil 1: Protokoll für mehrschichtige Erklärbarkeit

Sie entwickeln ein KI-Bewertungstool, das Videoarbeiten der Schüler nach drei Kriterien bewertet: narrative Kohärenz, technische Ausführung und Barrierefreiheits-Compliance (z. B. Einbeziehung von Untertiteln). Für jedes der drei Bewertungskriterien schreiben Sie zwei Versionen des Feedbacks:

Schülerorientierte Erklärung (Ebene 1): Schreiben Sie eine klare Aussage von nicht mehr als drei Sätzen, die dem Schüler genau mitteilt, was die KI gemessen hat, was sie gefunden hat und wie ein konkreter Verbesserungsschritt aussehen würde. Die Aussage darf keinen technischen Jargon verwenden und sollte in einem konstruktiven, nicht-pönanen Ton geschrieben sein.

Eintrag im Ausbilder-Prüfprotokoll (Ebene 2): Schreiben Sie einen strukturierten Protokolleintrag für dasselbe Bewertungsergebnis. Dieser Eintrag sollte den spezifischen von der KI bewerteten

Datenpunkt, die diesem Kriterium zugewiesene Gewichtung und die ausgelöste Modellregel enthalten. Dieses Protokoll ist der Prüfpfad, der es dem Ausbilder ermöglicht, jede KI-Entscheidung zu überprüfen und anzufechten.

Teil 2: Datenzustimmungskommunikation

Sie starten das oben beschriebene KI-Bewertungstool. Bevor ein Schüler sein erstes Videoprojekt einreicht, muss das System eine Datenzustimmungsbenachrichtigung anzeigen. Schreiben Sie mit dem Datenminimierungsprinzip als Leitfaden ein Zustimmungs-Popup in einfacher Sprache, das:

Alle Datenpunkte klar identifiziert, die die KI aus der Einreichung des Schülers erheben wird.

Ausdrücklich angibt, welche Daten analysiert, aber NICHT gespeichert werden (z. B. „Wir analysieren Ihre Audioaufnahme, um das Erzähltempo zu überprüfen, speichern die Audioaufnahme nach Abschluss der Analyse jedoch nicht.“).

Klar angibt, welche Daten gespeichert werden und wie lange.

Eine klar gekennzeichnete Option für den Schüler enthält, bestimmte Formen der Datenerhebung abzulehnen.

- Produkt: Ein ausgefülltes ethisches Kommunikationspaket mit: sechs Feedback-Erklärungen (drei Schülererklärungen Ebene 1 und drei Prüfprotokolleinträge Ebene 2) und einem einfachen Zustimmungs-Popup-Text. Dieses Paket zeigt, dass ethische KI-Kommunikation eine vielschichtige Designherausforderung ist, die verschiedene Schreibregister für verschiedene Zielgruppen erfordert.

6. Reflexionsfrage

Wird der Lernende in Ihrem KI-Tool beim ersten Interaktionsmoment ausdrücklich und klar informiert, dass er mit einem automatisierten System interagiert – oder lässt ihn das Design, der Name und der Gesprächston der Schnittstelle subtil glauben, er spreche mit einem menschlichen Experten?

7. Referenzen

Cavoukian, A. (2009). Privacy by Design: The 7 Foundational Principles. Information and Privacy Commissioner of Ontario.

D2CIN Project Consortium. (2025). Insights & Recommendations Report (Aktivität 1.2). Mimic Productions & Budakov Films.

Doshi-Velez, F. & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv:1702.08608.

Europäisches Parlament. (2024). Verordnung (EU) 2024/1689 — KI-Gesetz. Amtsblatt der Europäischen Union.

Europäisches Parlament und Rat. (2016). Verordnung (EU) 2016/679 (DSGVO). Amtsblatt der Europäischen Union.

Kapitel 5: Digitales Storytelling für inklusives Lernen

Kontext und Begründung

Das Problem: Bildungstechnologie behandelt Vielfalt oft als visuelle Spezifikation: Fügen Sie einen Charakter mit Kopftuch ein, fügen Sie im Hintergrund einen Rollstuhlnutzer hinzu, stellen Sie sicher, dass die Hautfarben des Avatars einen Gradienten abdecken. Dieser Ansatz, den die D2CIN-Fokusgruppen ausdrücklich als „Tokenismus“ bezeichneten, ist nicht nur unwirksam, sondern aktiv schädlich. Er signalisiert Lernenden aus unterrepräsentierten Gruppen, dass sie sekundäre Figuren in der Bildungserzählung sind, nicht ihre Protagonisten. Darüber hinaus hat die Integration von KI in das Storytelling in der Berufsbildung neue Risiken eingeführt: KI-generierte Charaktere und Szenarien, wenn sie auf Massenmedien trainiert werden, neigen dazu, dominante kulturelle Stereotype in großem Maßstab und mit algorithmischer Autorität zu reproduzieren (Quelle: D2CIN Insights Report).

Das Ziel: Dieses Kapitel bietet Erstellern von Berufsbildungsinhalten und Ausbildern einen strukturierten kreativen Rahmen für die Entwicklung inklusiver digitaler Narrative, die KI als Empowerment-Werkzeug und nicht als Homogenisierungswerkzeug nutzen. Es bewegt Fachleute von passiver Vielfalt – dem Vorhandensein vielfältiger Charaktere – zu aktiver Inklusion: vielfältige Charaktere, die Narrative vorantreiben, Expertise besitzen und Lernenden während ihrer Bildungsreise sinnvolle Handlungsfähigkeit verleihen.

2. Lernziele (LZ)

- LZ1: Erstellen Sie vielfältige Narrative, die komplexe Identitäten widerspiegeln und „Tokenismus“ (oberflächliche Vielfalt) in Berufsbildungsszenarien vermeiden.
- LZ2: Gestalten Sie Narrative-Verzweigungen, die es Lernenden ermöglichen, die Richtung der Geschichte oder die Identität ihres KI-Mentors zu kontrollieren.

3. Grundlegende theoretische Konzepte

Inklusives digitales Storytelling ist keine Frage des Castings. Es erfordert eine fundamentale Veränderung in der Art und Weise, wie wir die Beziehung zwischen Narrativ, Identität und Lernen verstehen. Die folgenden drei Konzepte liefern den kreativen und ethischen Rahmen für dieses Kapitel.

3.1 Authentische Repräsentation: Tiefe statt Dekoration

Authentische Repräsentation ist die Praxis, Charaktere und Szenarien zu schaffen, die die volle Komplexität menschlicher Erfahrung widerspiegeln, anstatt auf stenografische visuelle oder kulturelle Stereotypen zurückzugreifen. Ein authentisch repräsentierter Charakter hat eine spezifische Berufsgeschichte, einen kohärenten Wertesatz, einen kulturellen Kontext, der die

Entscheidungsfindung informiert, und – was entscheidend ist – eine Expertise, die für die Narrativentwicklung zentral ist. Authentische Repräsentation ist das Gegenteil von „Kulissendekoration“, bei der vielfältige Charaktere sichtbar, aber stimmlos sind.

In KI-gestützten Berufsbildungsszenarien fungieren Charaktere als Stellvertreter-Mentoren. Wenn jeder KI-Mentor, der Expertise demonstriert, dasselbe demografische Profil hat – typischerweise westlich, männlich und körperlich leistungsfähig –, internalisiert der Lernende dieses Profil als Norm beruflicher Kompetenz. Daher ist authentische Repräsentation in KI-Bildungsinhalten eine Form der Definition beruflicher Ambitionen mit messbarem Einfluss auf den Ehrgeiz und die Selbstwirksamkeit des Lernenden.

3.2 Narrative Handlungsfähigkeit: Über das lineare Szenario hinausgehen

Narrative Handlungsfähigkeit ist das Ausmaß, in dem ein Lernender die Richtung und das Ergebnis eines Bildungsszenarios sinnvoll beeinflussen kann. Echte narrative Handlungsfähigkeit entsteht, wenn die Entscheidungen des Lernenden zu sinnvoll unterschiedlichen und gleichwertig gültigen Bildungsergebnissen führen. In einem Konfliktlösungsszenario sollten die Wahl zur Deeskalation durch Empathie und die Wahl zur Deeskalation durch das Setzen klarer professioneller Grenzen als legitime, kontextangemessene Strategien behandelt werden. Die Rolle der KI besteht nicht darin, zu urteilen, was „richtig“ ist, sondern realistische Konsequenzen jedes Ansatzes zu simulieren.

Narrative Handlungsfähigkeit ist besonders wichtig für Lernende aus unterschiedlichen kulturellen Hintergründen, deren Problemlösungsinstitute möglicherweise vom in einem linearen KI-Szenario kodierten „Standard“-Ansatz abweichen. Eine KI, die kulturell informierte, vollständig valide professionelle Antworten bestraft, weil sie nicht ihrem vordefinierten optimalen Weg entsprechen, ist keine neutrale Bewertung, sondern kulturelle Voreingenommenheit, die als Bildungstechnologie operationalisiert wird.

3.3 Tokenismus: Die Unterscheidung zwischen Anwesenheit und Inklusion erkennen

Tokenismus ist die Praxis, nur oberflächliche oder symbolische Bemühungen um Vielfalt zu leisten - Mitglieder unterrepräsentierter Gruppen in eine Erzählung aufzunehmen auf Weisen, die sichtbar, aber nicht bedeutsam sind. Im digitalen Storytelling manifestiert sich Tokenismus, wenn verschiedene Charaktere in einem Szenario anwesend sind, aber periphere Rollen einnehmen: Sie sind stille Zeugen, unterstützende Helfer oder das Objekt des Problems, kein Handlungsträger der Lösung.

In KI-gestützten Bildungsinhalten ist Tokenismus leicht zu erzeugen und schwer zu entdecken. Die Diagnosefrage für Berufsbildungsentwickler lautet nicht „Sind verschiedene Charaktere anwesend?“, sondern „Treiben die verschiedenen Charaktere die Handlung voran, demonstrieren sie Expertise und modellieren sie beruflichen Erfolg?“

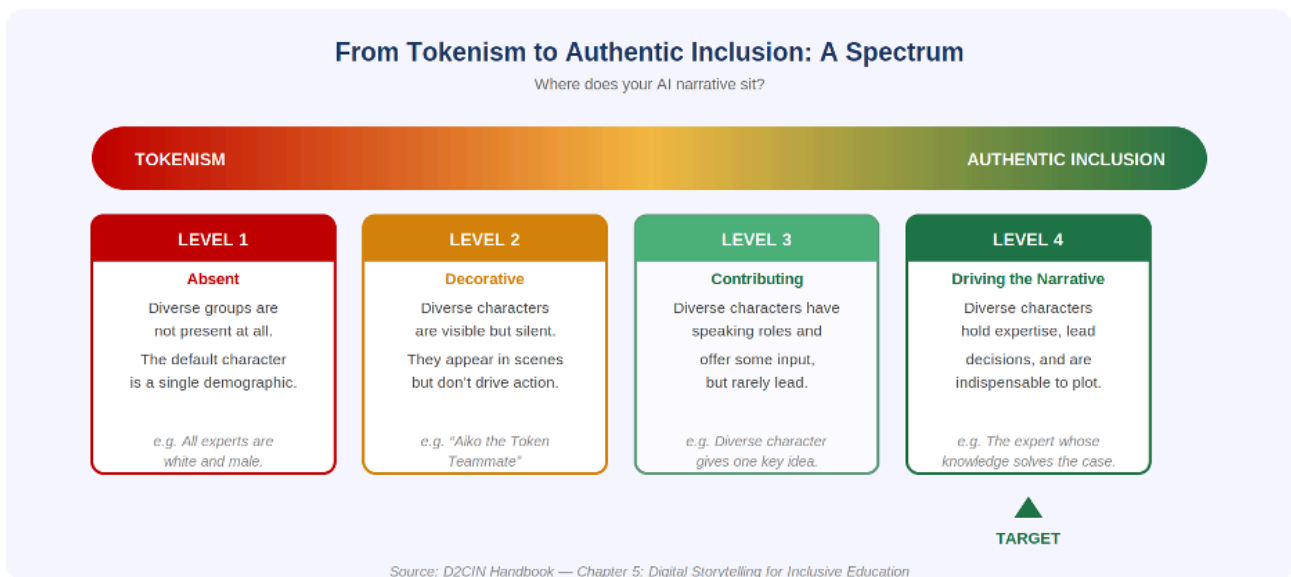


Abbildung - Von Tokenismus zu authentischer Inklusion: Ein Vier-Ebenen-Spektrum

4. Häufige Fallstudie (Das „Vorher“-Szenario)

- **Titel:** „Der Token-Teammitglied“
- **Szenario:** Eine Berufsbildungseinrichtung in Deutschland entwickelte eine KI-gestützte Simulation für ein Projektmanagement-Modul. Das Szenario versetzte Lernende in eine fiktive Kreativagentur, die einen geschäftigen Produktlaunchprozess managt. Das Team bestand aus vier KI-generierten Charakteren: Daniel (Teamleiter, weißer Mann, Anfang 40), Markus (technischer Leiter, weißer Mann, Anfang 30), Sophie (Kontaktperson, weiße Frau, Anfang 30) und Aiko (visuelle Designerin, ostasiatische Frau, Anfang 20, beschrieben nur als „die Kreative“). Im Modul mit sieben Szenarien erschien Aiko in vier Szenen. In jeder nickte sie schweigend, während Daniel die kreative Strategie erklärte, äußerte eine einzige Zustimmung („Ich kann das zum Laufen bringen“) und bekam nie eine Szene, in der ihre visuelle Expertise zu einer Lösung geführt hatte. Als Lernende gebeten wurden, die Teamleistung zu bewerten, bemerkten mehrere, dass sie „nicht sicher waren, was Aikos eigentliche Rolle war“.

Warum es scheiterte



Aiko ist ein Fall von Tokenismus. Sie ist präsent, besteht die demografische Vielfaltsprüfung, ist aber nicht inkludiert. Ihre berufliche Identität wurde auf ein Adjektiv reduziert, und ihre narrative Funktion ist rein dekorativ: das Team visuell vielfältig zu machen, ohne narrative Macht auf sie zu verlagern. Dies hat zwei schädliche Konsequenzen. Erstens sehen Lernende mit ähnlichem Hintergrund wie Aiko ein Modell beruflicher Beteiligung, das ihre Marginalisierung bestätigt. Zweitens internalisieren alle Lernenden ein Modell des kreativen Projektmanagements, in dem visuelle Expertise, implizit als nicht-westlich und weiblich kodiert, peripher zur strategischen Macht ist.

5. Praxisszenario (Die „Danach“-Anwendung)

Workshop für inklusives Verzweigungs-Szenario

Diese kreative Schreibübung fordert den Leser auf, ein kurzes, verzweigendes narratives Szenario für einen Berufsbildungskontext zu entwickeln, das sowohl authentische Repräsentation als auch narrative Handlungsfähigkeit anwendet.

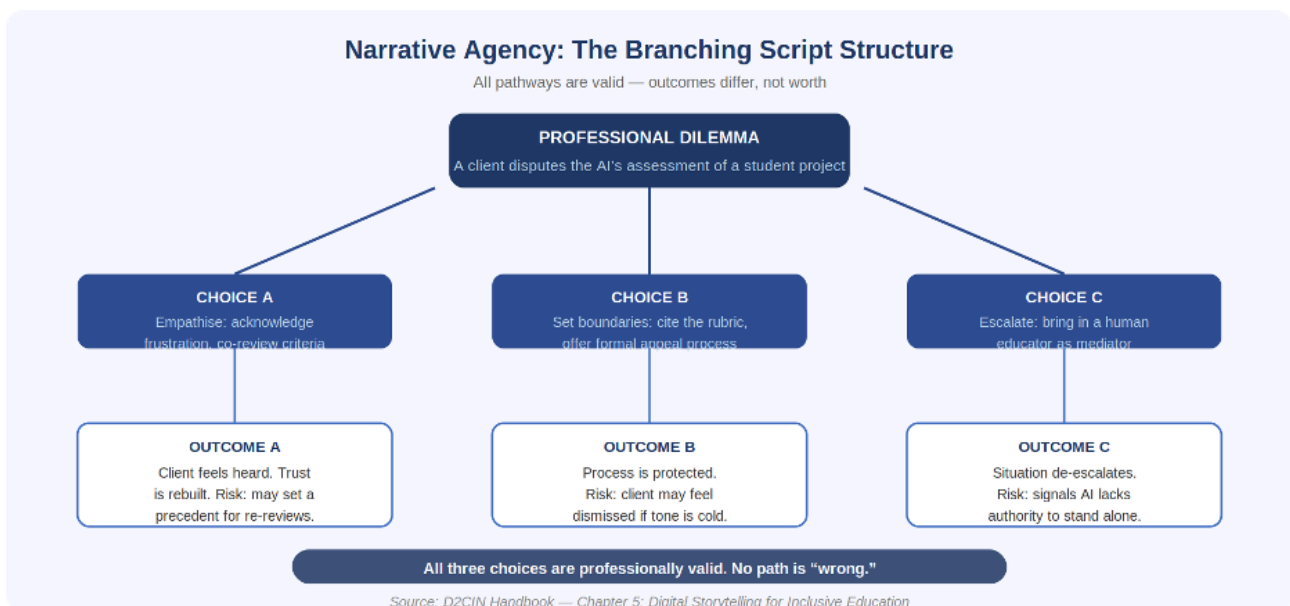


Abbildung - Verzweigungs-Skript-Struktur: Alle Wege sind gültig

- Anweisungen

Phase 1: Charakterdesign (Anwendung authentischer Repräsentation)

Entwerfen Sie die zwei Hauptcharaktere, mit denen Ihr Lernender im Szenario interagieren wird. Für jeden Charakter müssen Sie Folgendes definieren:

- Name, Alter und Berufserfahrung: Geben Sie dem Charakter eine konkrete, glaubwürdige Berufsgeschichte – nicht nur eine Berufsbezeichnung.
- Kultureller Kontext: Identifizieren Sie einen Aspekt des kulturellen Hintergrunds des Charakters, der sinnvoll beeinflusst, wie er die professionelle Herausforderung im Szenario angeht.
- Expertisebereich: Identifizieren Sie das spezifische Wissen, das dieser Charakter einbringt und das kein anderer Charakter im Szenario liefern kann.
- Art des Scheiterns: Identifizieren Sie eine Weise, in der der Ansatz dieses Charakters, so legitim er auch sein mag, in jedem Kontext möglicherweise nicht geeignet ist.

Phase 2: Szenariodesign (Anwendung narrativer Handlungsfähigkeit)

Schreiben Sie ein kurzes Szenario (200–300 Wörter), das in einem berufsbezogenen Kontext der Berufsbildung angesiedelt ist. Das Szenario muss:

- Dem Lernenden ein echtes professionelles Dilemma präsentieren, das keine einzige objektiv richtige Antwort hat.
- Dem Lernenden mindestens drei sinnvoll unterschiedliche Antwortoptionen anbieten – keine von denen als „offensichtlich falsch“ formuliert werden sollte.
- Eine kurze Beschreibung (je 2–3 Sätze) der verschiedenen narrativen Konsequenzen enthalten, die sich aus jeder Wahl ergeben.

Phase 3: Tokenismus-Audit (Selbstbewertung)

Nachdem Sie Ihr Szenario verfasst haben, wenden Sie das folgende Vier-Fragen-Audit vor der Finalisierung an:

- Existieren einige meiner vielfältigen Charaktere nur dazu, um von anderen Charakteren geholfen, beraten oder gerettet zu werden, anstatt die Handlung mit ihrer eigenen Expertise voranzutreiben?
- Ist der kulturelle Kontext, den ich meinen vielfältigen Charakteren zuschreibe, ein echter Aspekt ihrer beruflichen Identität oder eine visuelle oder sprachliche Abkürzung?
- Erlaubt meine Verzweigungsstruktur wirklich mehrere gültige professionelle Ansätze, oder belohnt sie subtil eine kulturelle Norm beruflichen Verhaltens vor anderen?
- Wenn ich meine vielfältigen Charaktere vollständig aus dem Szenario entferne, ändert sich das Hauptnarrativ oder das Lernziel? (Wenn nein, sind sie Dekoration.)

- Produkt: Ein abgeschlossenes inklusives Szenariopaket mit: zwei Charakterdesign-Profilen, dem vollständigen Verzweigungs-Szenario (einschließlich aller drei Wahlpfade und ihrer Konsequenzen) und einem abgeschlossenen Tokenismus-Audit aus Phase 3 mit schriftlichen Antworten auf alle vier Diagnosefragen.

6. Reflexionsfrage

Treiben in dem von Ihnen entwickelten narrativen Szenario die Charaktere aus unterrepräsentierten Gruppen die Handlung mit ihrer spezifischen Expertise voran, oder sind sie nur anwesend, um die Szene vielfältig aussehen zu lassen? Wenn Sie sie vollständig aus der Geschichte entfernen, ändert sich das Lernziel?

7. Referenzen

Hooks, b. (1994). Teaching to Transgress: Education as the Practice of Freedom. Routledge.

Steele, C.M. (2010). Whistling Vivaldi: How Stereotypes Affect Us and What We Can Do. W. W. Norton & Company.

D2CIN Project Consortium. (2025). Insights & Recommendations Report (Aktivität 1.2). Mimic Productions & Budakov Films.

Murray, J.H. (1997). Hamlet on the Holodeck: The Future of Narrative in Cyberspace. The Free Press.

Fazit

Dieser Leitfaden basiert auf der Überzeugung: dass inklusive, ethische KI in der Berufsbildung keine zukünftige Ambition ist, sondern eine aktuelle berufliche Pflicht. Die fünf Kapitel, die diesem Fazit vorangehen, sind nicht als Theorie konzipiert, die erlernt und vergessen werden soll. Sie wurden – basierend auf fundierten Belegen, die in zwei Ländern und sieben disziplinären Perspektiven gesammelt wurden – als praktischer Werkzeugkasten für Menschen entwickelt, die täglich mit KI entwickeln und lehren.

Der Weg durch diesen Leitfaden hat sich vom Unsichtbaren zum Anwendbaren verschoben. Kapitel 1 machte Verzerrungen sichtbar – verfolgte sie von verzerrten Trainingsdaten bis zu den Karriereempfehlungen, die der Lernende erhält, und bot Fachleuten ein Audit für Repräsentationsgerechtigkeit, um sie zu erfassen und zu korrigieren, bevor sie eingesetzt werden. Kapitel 2 machte die psychologische Erfahrung des Lernenden zu einer Designpriorität und lieferte Heuristiken für das Management kognitiver Belastung und die Navigation im „Uncanny Valley“ beim Design von KI-Avataren. Kapitel 3 ging über das WCAG-Compliance-Kästchen zu einem ganzheitlichen Modell multimodaler Barrierefreiheit hinaus und bestand darauf, dass ein KI-Tool, das den „Ton aus“-Test nicht besteht, kein zugängliches Tool ist, unabhängig von seiner Zertifizierung. Kapitel 4 beseitigte die Opazität der KI-Ethik, indem abstrakte Prinzipien durch aufzeichenbare und implementierbare Ergebnisse ersetzt wurden: mehrschichtige Erklärbarkeits-Protokolle, Datenzustimmungsformulare zur Minimierung und Prüfprotokolldateien, die Lernenden und Ausbildern die Transparenz geben, die ihnen gesetzlich und pädagogisch zusteht. Und Kapitel 5 gab dem inklusiven Design seine kreative Dimension – und zeigte, dass vielfältige Repräsentation keine Casting-Entscheidung ist, sondern ein narratives Handwerk mit spezifischen, erlernbaren Techniken zum Aufbau von Geschichten, in denen alle Charaktere die Handlung vorantreiben.

Vom Handbuch zur Praxis: Eine Verpflichtung zum Handeln

Dieser Leitfaden sollte nicht als passives Lesematerial behandelt werden. Er ist ein aktives Implementierungswerkzeug, das dazu bestimmt ist, zusammen mit realen Projekten, realen Prototypen und echten Klassenzimmern eingesetzt zu werden. Die fünf Kapitel werden direkt auf den Arbeitsablauf eines professionellen Berufsbildungsfachmanns abgebildet: prüfen vor dem Training; entwerfen vor dem Bauen; testen vor dem Einsetzen; erklären vor dem Bewerten; und Geschichten erzählen, die ermächtigen, nicht dekorieren.

Für Inhaltentwickler und -entwickler bedeutet dies, die Bias-Erkennungs-Checkliste (Kapitel 1) und den multimodalen Barrierefreiheits-Audit (Kapitel 3) als verbindliche Qualitätssicherungsmittel zu verwenden, bevor ein KI-Modul den Lernenden erreicht. Für Ausbilder und Lehrdesigner bedeutet dies, die mehrschichtigen Erklärbarkeits-Protokolle (Kapitel 4) und die Heuristiken für inklusives

Nutzererlebnis (Kapitel 2) zu verwenden, um die ihnen bereitgestellten Tools zu bewerten und zu verbessern, anstatt sie einfach zu akzeptieren und zu verwenden. Für Geschichtenerzähler und Narrativdesigner bedeutet dies, den Workshop für verzweigende Skripte und den Tokenismus-Audit (Kapitel 5) auf jedes Szenario, jeden Charakter, jede Wahlstruktur anzuwenden – und nicht nur zu fragen, ob verschiedene Charaktere anwesend sind, sondern ob sie wesentlich sind.

Dies sind keine ehrgeizigen Ziele für eine ideale Zukunft. Sie sind der Mindeststandard für KI-unterstützte Berufsbildung im europäischen Kontext, wo das EU-KI-Gesetz, die DSGVO und die Erasmus+-Programmpriorität Digitale Bereitschaft und Resilienz gemeinsam ein klares Mandat festlegen: KI in der Bildung muss ethisch, erklärbar und von Design her inklusiv sein.

Der Weg vorwärts

Dieser Leitfaden ist das Fundament, nicht die Decke. Er wird als grundlegende Lernressource für die interaktive KI-Workshop-Serie von D2CIN (Aktivität 3) dienen, wo die Leitlinien von 40 externen Teilnehmenden in Deutschland und Bulgarien getestet und verfeinert werden. Die Pilotworkshops werden eine letzte Validierungsschicht bieten und prüfen, nicht nur ob die Rahmenwerke in der Theorie funktionieren, sondern ob sie in den Händen der Berufsbildungsfachleute, die sie am dringendsten benötigen, anwendbar sind.

Nach der Projektlaufzeit ist die in diesem Leitfaden eingebettete Methodik so konzipiert, dass sie skalierbar ist. Der Repräsentationsgerechtigkeits-Audit, der Neuro-inklusive Schnittstellenaudit, das Mehrschichtige Erklärbarkeitsprotokoll und der Inklusiver Verzweigungs-Skript-Workshop sind kein Eigentum von D2CIN. Sie sind übertragbare Werkzeuge, die unter der internationalen Creative Commons Attribution-Non-Commercial-Share-Alike 4.0 Lizenz angeboten werden und dazu bestimmt sind, von jeder Berufsbildungseinrichtung adaptiert, übersetzt und weiterentwickelt zu werden, die sich dem Aufbau von KI verpflichtet hat, die allen Lernenden gleichberechtigt dient.

Da KI-Technologie sich weiterentwickelt - da Avatare überzeugender werden, adaptive Algorithmen leistungsfähiger werden, Datenerhebung immer weiter verbreitet wird – werden die Rahmenwerke in diesem Leitfaden aktualisiert werden müssen. Die spezifischen Methoden werden sich ändern; die Grundprinzipien werden es nicht. Voreingenommenheit ist kein technischer Fehler, der behoben werden kann. Barrierefreiheit ist kein Compliance-Kontrollkästchen, das abgehakt werden kann. Transparenz ist keine Marketingverpflichtung, die verwaltet werden kann. Und Inklusion ist keine visuelle Spezifikation, die eingehalten werden kann.

Jedes davon ist eine laufende professionelle Praxis - Denkgewohnheiten und Handwerksgewohnheiten, die bei jedem Datensatz, jeder Designentscheidung und jeder erzählten Geschichte erneuert werden müssen.

„Dies kann nicht allein von Technikern gemacht werden, die echten Lehrer müssen es kontrollieren“ - Fokusgruppen-Teilnehmer in Deutschland, D2CIN
Aktivität A1.2, Oktober 2025

Dieses Prinzip - dass inklusive KI nicht isoliert aufgebaut werden kann, dass sie ein nachhaltiges, kritisches Engagement von Pädagogen, Lernenden und den Gemeinschaften erfordert, denen sie zu dienen bestimmt ist - ist das abschließende Fazit dieses Leitfadens. Die Werkzeuge sind hier. Die Belege sind dokumentiert. Die Verpflichtung ist geteilt.

Jetzt beginnt die Arbeit.