HMD Praxis der Wirtschaftsinformatik
https://doi.org/10.1365/s40702-026-01348-y

SCHWERPUNKT

Mit partizipativer Technologieentwicklung Desinformation begegnen: Designprinzipien für ein KI-gestütztes Dashboard im schulischen Kontext

Isabel Bezzaoui · Maximiliane Linde · Jonas Fegert

Eingegangen: 12. Juni 2026 / Angenommen: 14. Oktober 2026
© The Author(s) 2026

Zusammenfassung Die Förderung eines reflektierten Umgangs mit Desinformation gehört zu den zentralen Herausforderungen schulischer Bildung in einer zunehmend digitalisierten Öffentlichkeit. Künstliche Intelligenz (KI) eröffnet dabei neue Möglichkeiten, problematische Inhalte zu analysieren und im Unterricht als Lernanlass zu nutzen. Bislang fehlt es jedoch an praxisnahen und zugleich wissenschaftlich fundierten Gestaltungsansätzen für KI-gestützte Informationssysteme, die im Schulkontext transparent und didaktisch anschlussfähig sind. Vor diesem Hintergrund untersucht der Beitrag, wie sich im Rahmen partizipativer Gestaltungsprozesse Anforderungen und Designprinzipien für ein auf erklärbarer KI gestütztes Dashboard zur Auseinandersetzung mit Desinformation im schulischen Kontext ableiten lassen. Die Studie ist in der gestaltungsorientierten Wirtschaftsinformatik verortet und folgt dem Design Science Research Paradigma. Im Fokus stehen die frühen Phasen der Problem- und Anforderungsanalyse mit Lehrkräften.

Der Beitrag zeigt, dass neben der technischen Erkennungsleistung vor allem Transparenz und Erklärbarkeit, didaktische Anschlussfähigkeit, die aktive Einbindung der Lernenden, eine niedrigschwellige Bedienung sowie begleitende Unterstützungs- und Qualifizierungsangebote für Lehrkräfte über den erfolgreichen Einsatz eines solchen Systems entscheiden. Für die Forschung liefert der Beitrag ein erstes, empirisch fundiertes Set von Gestaltungsprinzipien für KI-gestützte Informations-

Isabel Bezzaoui
Goethe-Universität Frankfurt, Frankfurt, Deutschland
E-Mail: bezzaoui@em.uni-frankfurt.de

✉ Maximiliane Linde · Jonas Fegert
Digital Democracy & Participation, Karlsruher Institut für Technologie, Karlsruhe, Deutschland
E-Mail: maximiliane.linde@kit.edu; linde@fzi.de

Jonas Fegert
E-Mail: jonas.fegert@kit.edu

House of Participation, FZI Forschungszentrum Informatik, Karlsruhe, Deutschland

Published online: 18 November 2026

systeme in einem bislang wenig beforschten Anwendungskontext und verdeutlicht den Wert partizipativer DSR-Ansätze an der Schnittstelle von Wirtschaftsinformatik, Medienkompetenz und schulischer Praxis. Für die Praxis bietet er konkrete Orientierung für die Entwicklung und Einführung KI-gestützter Werkzeuge zur Förderung kritischer Medienkompetenz im Unterricht.

Schlüsselwörter Desinformation · Kritische Medienkompetenz · Explainable AI (XAI) · Design Science Research · Partizipatives Design · Schulische Bildung

Abstract Promoting a reflective approach to disinformation is one of the central challenges of school education in an increasingly digitalized public sphere. Artificial intelligence opens up new opportunities to analyze problematic content and use it as a learning opportunity in the classroom. To date, however, there is a lack of design approaches for AI-supported systems that are both practice-oriented and empirically grounded, and that can be applied in a transparent and didactically adequate manner in schools. This paper examines how requirements and design principles for an explainable AI-supported dashboard to address disinformation in school contexts can be derived through participatory design processes. The study is situated in design-oriented information systems research and follows the Design Science Research paradigm; its focus lies on the early phases of problem identification and requirements analysis with teachers.

The paper shows that, beyond technical detection performance, the successful use of such a system depends above all on transparency and explainability, didactic compatibility, the active involvement of learners, low-threshold usability, and accompanying support and qualification offerings for teachers. The paper provides an initial empirically grounded set of design principles for explainable AI-supported information systems in an application context that has received little scholarly attention to date, and highlights the value of participatory DSR approaches at the intersection of information systems research, media literacy, and educational settings. For practitioners, it offers concrete guidance on developing and implementing AI-supported tools to promote critical media literacy in the classroom.

Keywords Disinformation · Critical media literacy · Explainable AI (XAI) · Design Science Research · Participatory design · School education

1 Einleitung

Die Fähigkeit zur informierten Meinungsbildung und zur Teilhabe ist eine zentrale Voraussetzung demokratischer Gesellschaften. In zunehmend digitalisierten Öffentlichkeiten wird diese Fähigkeit jedoch durch die wachsende Verbreitung von Desinformation herausgefordert. Jugendliche verbringen täglich mehrere Stunden in Online-Social-Networks (OSN) und begegnen dort kontinuierlich potenziell irreführenden oder manipulierten Inhalten. Trotz hoher digitaler Affinität verfügen viele Jugendliche nicht über ausreichende kritische Medien- und Datenkompetenzen, um

Desinformation zuverlässig zu erkennen (Breakstone et al. 2024; Nygren und Guath 2022).

Künstliche Intelligenz trägt einerseits zur Skalierung von Desinformation bei, andererseits eröffnet sie neue Möglichkeiten für deren systematische Analyse. KI-Verfahren können große Mengen plattformspezifischer Daten verarbeiten und darin Muster potenziell manipulativer Kommunikation identifizieren. Methoden der erklärbaren KI (Explainable AI, XAI) können zusätzlich sichtbar machen, auf welchen stilistischen und linguistischen Merkmalen eine solche Einordnung beruht. Indem diese Merkmale im untersuchten Text hervorgehoben werden, wird die Analyse nachvollziehbar und kann als Grundlage für Lern- und Reflexionsprozesse dienen. Für Erkennungssysteme zeigt sich, dass deren Gestaltung sektorspezifische Praxisanforderungen systematisch berücksichtigen muss (Bezzaoui et al. 2025a). Im schulischen Kontext genügt eine hohe Erkennungsgenauigkeit daher nicht, sondern es sind didaktisch anschlussfähige, transparente und für unterschiedliche Zielgruppen verständliche Systeme (Gasevic et al. 2017; Su und Yang 2024) erforderlich. Bislang ist nur begrenzt untersucht worden, wie Informationssysteme, die sich auf Verfahren erklärbarer KI stützen, für den schulischen Einsatz partizipativ gestaltet werden können. Insbesondere fehlt es an belastbaren Gestaltungsprinzipien, die pädagogische Zielsetzungen, technologische Möglichkeiten und die Perspektiven relevanter Akteur*innen systematisch miteinander verbinden.

Hier setzt der vorliegende Beitrag an. Als empirischer und gestalterischer Bezugsrahmen dient das Projekt DisCoBoard, das, gefördert vom Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend, ein KI-gestütztes Dashboard zur Erkennung und Analyse von Desinformation im Kontext der OSN-Nutzung Jugendlicher partizipativ entwickelt. Im vorliegenden Beitrag bedeutet partizipative Entwicklung, relevante Stakeholder in den gesamten Entwicklungsprozess einzubeziehen, mit einem Fokus auf die Anforderungen von Lehrkräften. Im Unterricht dient das Dashboard als didaktisches Instrument, das reale Begegnungen mit Desinformation aufgreift, evidenzbasiert diskutierbar macht und mit der Vermittlung kritischer Medienkompetenz verknüpft. Die anvisierten Nutzer*innen sind Lehrkräfte und Jugendliche ab der Sekundarstufe.

Der Beitrag ist in der gestaltungsorientierten Wirtschaftsinformatik verortet und folgt dem Paradigma des Design Science Research (DSR). DSR zielt darauf ab, für reale Anwendungsprobleme innovative soziotechnische Artefakte zu entwickeln und dabei verallgemeinerbares Gestaltungswissen zu erzeugen (Hevner et al. 2004; Gregor und Jones 2007). Er fokussiert die frühe Problem- und Anforderungsanalyse mit Lehrkräften und leitet daraus Meta-Requirements sowie vorläufige Designprinzipien für ein auf erklärbarer KI gestütztes Dashboard im schulischen Kontext ab. Daraus ergibt sich erstes Gestaltungswissen für die Forschung zu erklärbaren und partizipativ entwickelten Informationssystemen und zugleich praxisnahe Orientierung für die Entwicklung und Einführung solcher Werkzeuge im Schulalltag aus Perspektive der Lehrkräfte.

Daraus ergibt sich die folgende Forschungsfrage: *Wie lassen sich im Rahmen partizipativer Gestaltungsprozesse Anforderungen von Lehrkräften an ein auf erklärbarer KI gestütztes Dashboard zur Förderung kritischer Medienkompetenz im schulischen Kontext erheben und daraus erste Designprinzipien ableiten?*

Der Beitrag verdeutlicht die Relevanz partizipativer Design Science Research für die Gestaltung nachhaltig nutzbarer KI-gestützter Systeme zur Förderung der Medienkompetenz Jugendlicher. Er betont, dass Lehrkräfte Anforderungen nicht nur in Bezug auf Technik, sondern auch hinsichtlich didaktischer Anschlussfähigkeit, Kompetenzorientierung, Transparenz und Unterstützung bei der Nutzung stellen.

2 Forschungshintergrund

Die Auseinandersetzung mit Desinformation ist zunehmend auch Gegenstand der Wirtschaftsinformatik. Als Disziplin, die soziotechnische Systeme gestaltet, ist sie gefordert, Informationssysteme zu entwickeln, die der Verbreitung von Desinformation entgegenwirken und die demokratische Teilhabe stützen (Weinhardt et al. 2024). Der vorliegende Beitrag verbindet diese gestaltungsorientierte Perspektive mit zwei Forschungssträngen, die bislang weitgehend getrennt verlaufen sind: der Integration künstlicher Intelligenz in den Schulunterricht und der Förderung von Medienkompetenz im Umgang mit Desinformation.

2.1 Integration künstlicher Intelligenz im Schulunterricht

Die bildungsbezogene Forschung hebt hervor, dass KI das Potenzial besitzt, Lehr- und Lernprozesse grundlegend zu verändern und qualitativ zu verbessern, etwa durch adaptive Lernumgebungen, automatisiertes Feedback oder datenbasierte Entscheidungsunterstützung (Gill et al. 2024; Mikeladze et al. 2024). Für den vorliegenden Beitrag ist insbesondere relevant, dass KI nicht nur als Automatisierungstechnologie, sondern auch als Bestandteil von Informations- und Reflexionsprozessen im Unterricht verstanden werden kann.

Empirische Forschung zeigt, dass insbesondere KI-gestützte adaptive Lernsysteme positive Effekte auf Lernleistung und Lernmotivation haben können, indem sie Lerninhalte an individuelle Bedürfnisse anpassen und zeitnahes Feedback bereitstellen (Luckin et al. 2016). Qualitative Evidenz zeigt zudem, dass digital gestützte, schülerzentrierte Lernformen zu höherem Engagement beitragen (Timotheou et al. 2023). Trotz dieser positiven Befunde bestehen auch Zweifel an der pädagogischen Eignung und curricularen Anschlussfähigkeit von KI-Lernwerkzeugen (Chiu und Chai 2020) sowie an ihrer systematischen Verknüpfung mit didaktischen Strategien und Lernzielen (Chen et al. 2020). Insbesondere bei KI-basierten Analysesystemen wird betont, dass deren Gestaltung und Evaluation stärker in der pädagogischen Theorie verankert werden müssen (Gasevic et al. 2017) und dass KI-Inhalte altersgerecht und lernwirksam aufbereitet sein sollten (Su und Yang 2024).

Zu den häufig genannten Barrieren neben didaktischen Fragen zählen fehlende KI-Kenntnisse und geringe Selbstwirksamkeitserwartungen bei Lehrkräften, unzureichende curriculare Vorgaben sowie das Fehlen klarer Handlungsleitlinien für den Unterricht (Huang et al. 2026). Darüber hinaus berichtet die empirische Forschung über ausgeprägte Bedenken und teilweise hohe Angst- und Unsicherheitsniveaus bei Lehrkräften (Gouseti et al. 2025). Für die Gestaltung schulnaher KI-Systeme folgt daraus, dass technische Funktionalität allein nicht ausreicht, sondern dass Anforde-

rungen von Lehrkräften hinsichtlich der didaktischen Einbettung und der praktischen Umsetzbarkeit maßgeblich sind. Entsprechend wird ein stärker menschenzentrierter, partizipativer Gestaltungsansatz gefordert, der pädagogische Zielsetzungen, technologische Möglichkeiten und Gerechtigkeitsfragen systematisch miteinander verbindet (Brehmer und Buonassisi 2024; Vieriu und Petrea 2025). Die in diesem Beitrag im Fokus stehende Gruppe der Lehrkräfte wird dabei als wichtiger Stakeholder in der partizipativen Technologieentwicklung im Bildungskontext erachtet (Cober et al. 2015).

Methoden der erklärbaren KI zielen darauf ab, algorithmische Entscheidungen offenzulegen. Im Bereich der schulischen Bildung regen sie den Dialog zwischen Lehrkräften und Jugendlichen an, geben den Jugendlichen Handlungsfähigkeit und Anlass zur Reflexion, und können Vertrauen schaffen sowie die Effektivität von Lernprozessen erhöhen (Khosravi et al. 2022; Chaushi et al. 2023). Erklärungen entfalten jedoch erst Wirkung, wenn sie für die Zielgruppen verständlich und anschlussfähig sind, unter anderem durch partizipative Gestaltung (Schmitt et al. 2025; Khosravi et al. 2022). Wie sich Erklärbarkeit für die Erkennung von Desinformation gestaltungsorientiert umsetzen lässt, zeigt eine vorausgehende DSR-Arbeit, die XAI-Prinzipien systematisch für ein Werkzeug zur Erkennung von Desinformation entwickelt hat (Bezzaoui et al. 2025b). An diese Vorarbeit knüpft der vorliegende Beitrag an, verlagert den Fokus jedoch von der allgemeinen Detektion auf den Schulkontext und dessen pädagogische Anforderungen.

2.2 Desinformation und Medienkompetenz im Schulunterricht

Unter Desinformation wird im Folgenden die absichtliche Erstellung und Verbreitung falscher oder irreführender Informationen verstanden, im Unterschied zur Fehlinformation, die ohne Täuschungsabsicht weitergegeben wird (Wardle 2018). Im schulischen Kontext sollen Lernende daher nicht nur falsche Inhalte erkennen, sondern auch Mechanismen und Verbreitungslogiken einordnen können.

Zahlreiche Studien zeigen, dass junge Menschen Schwierigkeiten haben, die Qualität und Glaubwürdigkeit von Online-Inhalten zu beurteilen (Breakstone et al. 2024; Nygren und Guath 2022), ihre eigenen Evaluationsfähigkeiten teilweise überschätzen und Informationen nicht selten anhand oberflächlicher Signale beurteilen (Coiro et al. 2015; Köbis et al. 2021). Auch Lehrkräfte berichten von mangelnder kritischer Reflexion von Medieninhalten (Herrero-Curiel und La-Rosa 2022). Diese Defizite machen Bildungsinterventionen erforderlich, die gezielt die Bewertungsfähigkeiten von Schüler*innen fördern (Barzilai und Chinn 2020).

Rau und Premo (2025) zeigen, dass gezielte Bildungsinterventionen das Erkennen und das kritische Bewerten irreführender und manipulativer Inhalte effektiv fördern können. Angesichts der Notwendigkeit, Interventionen skalierbar umzusetzen, besteht ein wichtiger Forschungsbedarf darin, effektive Lehrmethoden zu entwickeln, die Lehrkräfte in angemessener Zeit erlernen und anwenden können (Orosz et al. 2024). Vor diesem Hintergrund können XAI-gestützte Systeme angesichts oben besprochener Vorteile Bildungsinterventionen zur kritischen Auseinandersetzung mit Desinformation technologisch unterstützen, indem sie die Analyse konkreter Inhalte strukturieren, Begründungen und Unsicherheiten sichtbar machen und dadurch An-

knüpfungspunkte für angeleitete Reflexion und Diskussion bieten. Wirksame Ansätze gegen Desinformation im Schulkontext sind daher auf reflektierende, didaktisch eingebettete Lernprozesse angewiesen. Bedarfe von Lehrkräften sind dabei systematisch zu ermitteln.

Digitale Medienkompetenz umfasst nicht nur bloßes Faktenwissen, sondern auch die Fähigkeit, mediale Inhalte kritisch zu konsumieren und zwischen verlässlicher und irreführender Information zu unterscheiden (Jeong et al. 2012; Jones-Jang et al. 2021). Gezielte Interventionen können die Fähigkeit zur Differenzierung zwischen wahrheitsgemäßen und falschen Nachrichten signifikant verbessern (Guess et al. 2020; Hameleers 2022). Der reflektierte Umgang mit Desinformation erfordert jedoch auch eine Datenkompetenz, die das Verständnis für die datengetriebene Erzeugung und Verbreitung von Inhalten umfasst (Soßdorf et al. 2024).

Viele Fact-Checking-Tools sind bisher nur bedingt geeignet, um tatsächliche Medienkompetenzen zu vermitteln und lassen sich ohne pädagogische Einbettung nur eingeschränkt im Schulunterricht einsetzen (Nygren et al. 2022). Damit bleibt offen, wie digitale Werkzeuge gestaltet sein müssen, damit sie im Unterricht nicht nur Analyseergebnisse bereitstellen, sondern auch Reflexion, Diskussion und Kompetenzaufbau unterstützen. Genau hier setzt die Entwicklung partizipativ gestalteter und auf erklärbarer KI basierender Informationssysteme an.

3 DisCoBoard: Projektkontext und Forschungsdesign

DisCoBoard ist ein Forschungs- und Entwicklungsprojekt zur partizipativen Gestaltung eines KI-gestützten Dashboards, das Jugendliche beim reflektierten Umgang mit potenzieller Desinformation unterstützen soll. Ziel ist ein browserbasiertes Tool, das Inhalte aus digitalen Informationsumgebungen analysiert, potenziell manipulative Merkmale sichtbar macht und diese im schulischen Kontext als Ausgangspunkt für Reflexion, Diskussion und Kompetenzaufbau nutzbar macht. Das Projekt verbindet damit eine konkrete praktische Problemstellung aus dem Bildungsbereich mit einer gestaltungsorientierten Perspektive der Wirtschaftsinformatik. Methodisch ist DisCoBoard im Paradigma der DSR verortet. Für DisCoBoard ist dieser Zugang besonders geeignet, da nicht nur ein technisches System entwickelt, sondern auch ein didaktisch anschlussfähiges Artefakt für einen komplexen Nutzungskontext gestaltet werden soll.

Das Projekt umfasst vier aufeinander aufbauende Phasen. Der vorliegende Beitrag behandelt lediglich die erste Projektphase, in der die Bedarfsermittlung mit Lehrkräften und pädagogischen Fachkräften im Mittelpunkt steht. Ziel ist es, den Problemraum zu konturieren und erste Lösungsziele zu formulieren sowie daraus Anforderungen und erste Gestaltungsprinzipien für das auf erklärbarer KI gestützte Dashboard abzuleiten. Dieser Schritt steht bewusst am Anfang, denn ohne eine angemessene Analyse der Bedarfe von Lehrkräften lässt sich die effektive Implementierung eines solchen Systems im Unterricht nicht sicherstellen. In Phase 2 werden Schüler*innen als Hauptzielgruppe in die Artefaktgestaltung eingebunden. Ihre Partizipation ist dabei kein nachgelagerter Validierungsschritt, sondern ein zentraler Bestandteil, um ihre Sicht auf Nutzbarkeit, Verständlichkeit und lebensweltliche

Passung des Dashboards systematisch zu erfassen. In den Phasen 3 und 4 wird das Dashboard auf Basis der zuvor erhobenen Anforderungen entwickelt und in Schulen praktisch evaluiert. Rückmeldungen fließen in die iterative Weiterentwicklung ein.

4 Anforderungen für ein interaktives Dashboard zur Desinformationserkennung im Unterricht

4.1 Methodisches Vorgehen

Die Anforderungserhebung folgte einem partizipativen und qualitativ-empirischen Vorgehen. Methodisch orientierte es sich an Prinzipien des nutzerzentrierten und partizipativen Designs, wie sie für die Entwicklung digitaler Bildungswerkzeuge etabliert sind (Sanders und Stappers 2008; Spinuzzi 2005). Die qualitative Ausrichtung ermöglicht es, komplexe pädagogische und technische Anforderungen sowie implizites Praxiswissen sichtbar zu machen.

Die Datenerhebung erfolgte in Form von zwei aufeinander aufbauenden Workshops, die unterschiedliche Funktionen im Erhebungsprozess übernahmen. Die Rekrutierung der Teilnehmenden erfolgte zum einen durch den Praxispartner des Projekts, kosmos b, der über ein bundesweites Netzwerk an Lehrkräften und Bildungsakteuren verfügt, sowie über das Schneeballprinzip. Der erste Workshop fand im Oktober 2025 mit 20 Teilnehmenden in Präsenz statt. Während der Veranstaltung erprobten Lehrkräfte, Praxisvertreter*innen und Forschende gemeinsam Methoden, um den kritischen Umgang mit digitalen Informationen zu fördern. In diesem Rahmen erprobten die Teilnehmenden zwei digitale Demonstratoren, die den praktischen Umgang mit künstlicher Intelligenz und Desinformation exemplarisch veranschaulichen. Über das Interface des ersten Demonstrators erhielten die Akteur*innen die Möglichkeit, Nachrichtenartikel mithilfe eines Sprachmodells zu modifizieren und systematisch die Effekte unterschiedlicher Manipulationsstile auf die Wahrnehmung der Inhalte zu untersuchen. Mittels des Interfaces des zweiten Demonstrators analysierten sie Social-Media-Beiträge sowie Nachrichtenartikel und wendeten eine digitale Taxonomie an, um Indikatoren für Desinformation zu identifizieren. Anschließend reflektierten die Teilnehmenden in Kleingruppen die Einsatzpotenziale KI-gestützter Tools und diskutierten die Möglichkeiten, wie erklärbare künstliche Intelligenz Lernprozesse unterstützen kann. In der abschließenden Plenumsdiskussion wurden die gewonnenen Erkenntnisse konsolidiert und Perspektiven zur Integration der Tools in die schulische Praxis systematisch entwickelt.

Aufbauend auf diesen explorativen Erkenntnissen wurde im November 2025 ein zweiter Workshop als ganztägige Kreativwerkstatt durchgeführt. An dieser Veranstaltung nahmen 11 Personen aus unterschiedlichen institutionellen Kontexten teil, darunter Lehrkräfte verschiedener Schulformen, Bildungsbeauftragte, Medienpädagog*innen sowie Akteur*innen aus der Jugendarbeit. Ziel war die konkrete Erhebung und Ausformulierung von Nutzungsanforderungen für ein interaktives, erklärbares KI-Dashboard zur Erkennung von Desinformation. In moderierten Gruppenphasen entwickelten die Teilnehmenden Nutzungsszenarien, identifizierten didaktische und technische Anforderungen und formulierten erste prototypische Funkti-

onsideen. Zum Einsatz kamen Methoden wie die Erstellung von Personas und User Stories sowie Paper Prototyping und Gruppendiskussionen.

Die Datengrundlage bildeten qualitativ erhobene Diskussions- und Arbeitsergebnisse aus beiden Workshops. Die Beiträge des ersten Workshops wurden durch das Projektteam dokumentiert und konzeptionell ausgewertet. Die Ergebnisse der Kreativwerkstatt erweiterten die zentrale Datengrundlage der Anforderungsanalyse und umfassten schriftlich festgehaltene Gruppenergebnisse, priorisierte Anforderungslisten sowie strukturierte Funktions- und Designvorschläge. Alle Daten wurden offen und nicht standardisiert erfasst, um den Perspektivenreichtum der beteiligten Akteur*innen sowie didaktische, technische und organisatorische Aspekte des schulischen Einsatzes von DisCoBoard angemessen abzubilden.

Die Ergebnisse der beiden Workshops wurden systematisch qualitativ ausgewertet und inhaltlich thematisch geclustert, um zentrale Anforderungsfelder zu identifizieren. Die zusammengefassten Anforderungen wurden in didaktisch-pädagogische, technische, nutzerorientierte und organisatorische Kategorien eingeteilt. Durch iterative Verdichtung des Materials entstanden Meta-Requirements (MRs), die allgemeine Nutzerbedarfe widerspiegeln und als Grundlage für Designentscheidungen dienen. Die Analyse berücksichtigte zudem Barrieren und Herausforderungen in der Schulumgebung, um passfähige Lösungen zu gewährleisten.

4.2 Ergebnisse der Anforderungsanalyse

Die Analyse identifizierte eine Reihe von Meta-Requirements (MR), die grundlegende Anforderungen von Lehrkräften an die Gestaltung eines KI-gestützten Systems zur Auseinandersetzung mit Desinformation im schulischen Kontext beschreiben. Sie lassen sich neun Anforderungsdimensionen zuordnen und bilden die Grundlage

Tab. 1 Meta-Requirements und Designprinzipien

Anforderungsdimension	Meta-Requirements	Designprinzip
Didaktik und Kompetenzförderung	MR-D1, MR-D2, MR-D4, MR-D5	*DP-1* Lernunterstützung und Kompetenzorientierung
Aktive Mitwirkung	MR-D3	*DP-2* Aktive Beteiligung und Co-Kreation der Lernenden
Transparenz und Nachvollziehbarkeit	MR-T1–MR-T6	*DP-3* Erklärbare und nachvollziehbare Analyseprozesse
Usability und Interaktion	MR-U1–MR-U3	*DP-4* Intuitive, niedrigschwellige und zugängliche Bedienung
Technik und Organisation	MR-O1–MR-O3	*DP-5* Integration in schulische Routinen und IT-Infrastruktur
Unterstützung und Qualifizierung	MR-S1–MR-S3	*DP-6* Begleitung durch Materialien, Schulungen und Support
Inhaltlich-konzeptionelle Förderung	MR-C1–MR-C3	*DP-7* Förderung medienkritischen und multiperspektivischen Denkens
Evaluation und Qualität	MR-Q1–MR-Q3	*DP-8* Kontinuierliche Qualitätsprüfung und partizipative Weiterentwicklung
Gestaltung und Kommunikation	MR-V1–MR-V3	*DP-9* Visuelle Klarheit und pädagogisch angemessenes Wording

für die Ableitung gestaltungsleitender Designprinzipien (DP). Diese beziehen sich auf die Gestaltung eines solchen Systems aus Sicht der Lehrkräfte und werden iterativ mit Schüler*innen konkretisiert und weiterentwickelt. Tab. 1 fasst die Anforderungsdimensionen, die zugeordneten MRs sowie die daraus abgeleiteten vorläufigen DPs zusammen.

In der Dimension der Didaktik und Kompetenzförderung soll das System Lernprozesse unterstützen und nicht ersetzen (MR-D1) sowie Reflexion, Analyse und kritische Auseinandersetzung mit Sprache und Manipulation fördern. Dabei ist die Kompetenzorientierung (MR-D2) eng damit verbunden. Hinzu kommen die Anpassung an unterschiedliche Wissens- und Kompetenzniveaus (MR-D4), etwa durch altersgerechte Sprache, sowie die flexible Integration in verschiedene Unterrichtsformate (MR-D5). Aufgrund der Zentralität der Stärkung der Selbstreflexion bildet die aktive Mitwirkung von Schüler*innen eine eigene Anforderungsdimension. Schüler*innen sollen selbstständig Texte oder Inhalte hochladen können (MR-D3).

Eine weitere Dimension betrifft die Transparenz und Nachvollziehbarkeit der XAI. Alle Analyse- und Klassifikationsprozesse müssen erklärbar sein (MR-T1) und mit nachvollziehbaren Begründungen versehen werden (MR-T2). Nutzer*innen sollen verstehen, wie und warum das System zu bestimmten Ergebnissen gelangt (MR-T3). Begriffe, Taxonomien und Vorgehensweisen müssen verständlich sein (MR-T4), und Visualisierungen konsistent und eindeutig vom XAI-System genutzt (MR-T5) werden. Zudem müssen Ergebnisse verlässlich sein, Konfidenzwerte vermittelt, und Unsicherheiten des Modells transparent kommuniziert werden (MR-T6).

In der Anforderungsdimension Usability und Interaktion wurde eine intuitive, niedrigschwellige Bedienbarkeit als Voraussetzung für den Unterrichtseinsatz benannt (MR-U1). Hervorgehoben wurden visuelle Vergleichs- (MR-U2) sowie interaktive Analysefunktionen zur gleichzeitigen Betrachtung mehrerer Texte oder Kategorien (MR-U3).

Technische und organisatorische Rahmenbedingungen sind eine grundlegende Voraussetzung für den Einsatz im Schulalltag. Zentrale Anforderungen betreffen klar definierte Zuständigkeiten für die Wartung von Hardware, Software und der Zugangsverwaltung (MR-O1), eine stabile, störungsarme Funktionsweise (MR-O2) sowie die Kompatibilität mit gängigen schulischen Endgeräten und einen browserbasierten Zugang (MR-O3).

Darüber hinaus wurde ein deutlicher Bedarf an Unterstützungs- und Qualifizierungsangeboten für Lehrkräfte als Anforderung identifiziert: Schulungen und Materialien zur didaktischen und technischen Nutzung (MR-S1), niedrigschwellige Supportangebote wie Tutorials oder FAQ (MR-S2) sowie konkrete didaktische Handreichungen zur Unterrichtsintegration (MR-S3).

Auf inhaltlich-konzeptioneller Ebene steht die Förderung des medienkritischen Bewusstseins im Zentrum (MR-C1). Das System soll Sensibilität für Manipulation, Desinformation und mediale Mechanismen entwickeln und vergleichende Perspektiven, beispielsweise durch generierte Texte mit unterschiedlichen Manipulationsgraden, ermöglichen (MR-C2). Zugleich muss das System Diskussion und Reflexion anregen, statt Ergebnisse als abschließende Bewertung darzustellen (MR-C3).

In der Dimension Evaluation und Qualitätssicherung müssen die Analysen valide und konsistent sein (MR-Q1), Unsicherheiten in der Klassifikation explizit ausgewie-

sen werden (MR-Q2) und eine fortlaufende technische und didaktische Evaluation gewährleistet sein (MR-Q3).

Schließlich spielen Gestaltung und Kommunikation eine zentrale Rolle für die Nutzbarkeit und Verständlichkeit. Farben, Icons und Markierungen müssen nicht nur vom System technisch konsistent genutzt werden (siehe MR-T5), sondern auch eine visuell eindeutige, durch eine Legende erklärte Bedeutung besitzen (MR-V1), ein pädagogisch angemessenes Wording, das Begriffe wie „Fake“, „manipuliert“ oder „wahr“ kontextsensitiv einführt (MR-V2), sowie ein für Jugendliche ansprechendes, modernes Design, das motiviert, ohne inhaltliche Tiefe zu reduzieren (MR-V3).

Auf Basis der Meta-Requirements wurden vorläufige Designprinzipien aus Sicht der Lehrkräfte abgeleitet. Jedes Prinzip operationalisiert eine Anforderungsdimension und wird in den folgenden Iterationen mit Schüler*innen getestet und weiterentwickelt.

DP-1 Lernunterstützung und Kompetenzorientierung. Das System ist an das Vorwissen der Lernenden angepasst und regt sie zum eigenen Denken an, damit Jugendliche unterstützt werden, Manipulationen eigenständig zu erkennen und Medienkompetenz aufzubauen, statt automatisierte Lösungen zu übernehmen.

DP-2 Aktive Beteiligung und Co-Kreation der Lernenden. Das System ermöglicht Mitgestaltung, Experimentieren und Selbstreflexion, denn die Möglichkeit eigene Inhalte in das System zu integrieren, stärkt partizipative Lernformen und intrinsische Motivation.

DP-3 Erklärbare und nachvollziehbare Analyseprozesse. Alle Klassifikationen müssen erklärbar und nachvollziehbar sein. Das System legt seine Kriterien offen, liefert Begründungen und verwendet Visualisierungen, Farbcodes und Beispiele konsistent. Unsicherheiten und Modellgrenzen werden transparent kommuniziert.

DP-4 Intuitive, niedrigschwellige und zugängliche Bedienung. Das System ist intuitiv und niedrigschwellig bedienbar, unterstützt das Vergleichen mehrerer Inhalte und unterbricht den Unterrichtsfluss nicht, sodass die Bedienung für alle Beteiligten einfach bleibt.

DP-5 Integration in schulische Routinen und IT-Infrastruktur. Das System fügt sich in Unterrichtsabläufe und schulische Rahmenbedingungen ein, unterstützt Einzel-, Gruppen- und Plenumsarbeit gleichermaßen, gewährleistet einen stabilen Betrieb und minimiert technische Störungen, damit Lehrkräfte es nachhaltig und effektiv in ihren Unterricht integrieren können. Wartung, Support und Zuständigkeiten sind klar geregelt.

DP-6 Begleitung durch Materialien, Schulungen und Support. Lehrkräfte erhalten Handreichungen, Tutorials und didaktische Materialien, die Orientierung bei der Interpretation der Analysen sowie Vorschläge für Unterrichtseinheiten bieten.

DP-7 Förderung medienkritischen und multiperspektivischen Denkens. Das System macht unterschiedliche Texte und Manipulationsgrade sichtbar, um die Funktionsweise von KI-Modellen und manipulativen sprachlichen Strategien erfahrbar zu machen. Die Systemlogik regt Diskussion und Reflexion an, statt Ergebnisse als abschließend darzustellen.

DP-8 Kontinuierliche Qualitätsprüfung und partizipative Weiterentwicklung. Das System wird iterativ auf Basis von Rückmeldungen von Lernenden, Lehrkräften und pädagogischen Fachkräften weiterentwickelt. Fehler, Unsicherheiten oder inkonsis-

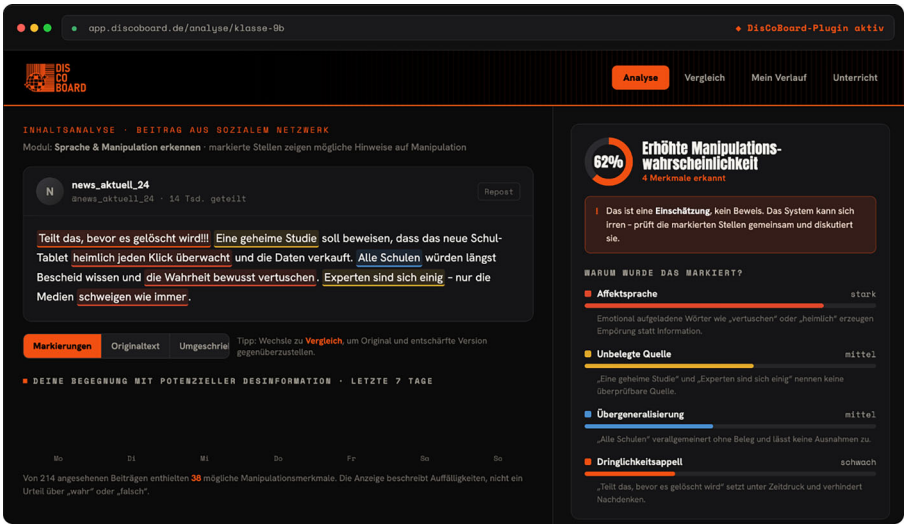


Abb. 1 Erster Klick-Dummy der DisCoBoard-Oberfläche. Markierte Textstellen werden durch die erklärbare KI-Komponente begründet (*rechts*); die Einschätzung wird mit einem expliziten Unsicherheitshinweis versehen, und die individuelle Exposition gegenüber potenzieller Desinformation wird visualisiert (*unten links*)

tente Ergebnisse werden systematisch erfasst und als Lern- und Verbesserungsanlässe genutzt.

DP-9 Visuelle Klarheit und pädagogisch angemessenes Wording. Farben, Icons und Markierungen sind eindeutig und durch Legenden erklärt, das Wording ist kontextsensibel, und das Design ist motivierend, ohne die Komplexität der Analysen zu verschleiern, sodass Jugendliche die Lernerfahrung als sinnvoll wahrnehmen.

Wie sich diese Designprinzipien in eine konkrete Oberfläche übersetzen lassen, veranschaulicht ein erster Klick-Dummy des DisCoBoards (Abb. 1). Sichtbar werden insbesondere die erklärbare Markierung potenzieller Manipulationsmerkmale samt Begründung (DP-3), die transparente Kommunikation von Unsicherheit statt endgültigen Urteilen (DP-3, DP-7), ein pädagogisch angemessenes Wording mitsamt erklärender Legende (DP-9) sowie die Visualisierung der individuellen Begegnung mit potenzieller Desinformation als Ausgangspunkt für Diskussion und Reflexion (DP-1, DP-7). In dieser Form bildet der Klick-Dummy die Grundlage für die folgende Iteration, in der die Oberfläche gemeinsam mit Schüler*innen erprobt und weiterentwickelt wird. Entscheidungen zur Farbgebung, zur pädagogisch angemessenen Wortwahl und zur visuellen Ausgestaltung des Klick-Dummys haben zunächst exemplarischen Charakter. Ob beispielsweise die Gestaltung von Jugendlichen als verständlich, ansprechend und motivierend wahrgenommen wird (DP-9), lässt sich erst unter Einbeziehung ihrer Anforderungen und Perspektiven belastbar beurteilen. Die vorläufigen Designprinzipien werden daher im weiteren iterativen Entwicklungsprozess gemeinsam mit Jugendlichen konkretisiert und überprüft.

5 Diskussion: Beiträge für Forschung und Praxis

Die Anforderungsanalyse zeigt, dass die Gestaltung KI-gestützter Systeme zur Auseinandersetzung mit Desinformation im schulischen Kontext nicht primär eine technische, sondern vor allem eine didaktische und organisatorische Herausforderung darstellt. Im Vordergrund standen aus Sicht der Lehrkräfte Transparenz und Erklärbarkeit, die Integration in schulische Routinen sowie die aktive Einbindung der Lernenden. Die Ergebnisse verdeutlichen damit, dass Akzeptanz und nachhaltiger Einsatz solcher Systeme ebenso von pädagogischer Anschlussfähigkeit wie von technischer Leistung abhängen.

Auf die Forschungsfrage, wie sich im Rahmen partizipativer Gestaltungsprozesse Anforderungen von Lehrkräften und erste Designprinzipien für ein auf erklärbarer KI gestütztes Dashboard ableiten lassen, gibt der Beitrag eine prozessbezogene Antwort. Entscheidend erscheint die gestufte Vorgehensweise: Eine explorative Phase macht den Problemraum und die Potenziale für KI gemeinsam mit Lehrkräften erklärbar, darauf aufbauende Workshops überführen diese Erkenntnisse in konkrete Nutzungsanforderungen, und die qualitative Verdichtung übersetzt sie in Meta-Requirements und Designprinzipien. Die frühe Einbindung von Lehrkräften als Akteur*innen der späteren Implementierung ermöglicht es, didaktische und organisatorische Anforderungen des Schulalltags bereits frühzeitig zu berücksichtigen. Sie eröffnet zugleich die Möglichkeit, nachfolgende Erhebungen mit Jugendlichen fokussierter und zielgerichteter zu gestalten und die vorläufigen Designprinzipien im Rahmen schulischer Handlungsspielräume und Restriktionen zu konkretisieren.

Für die gestaltungsorientierte Wirtschaftsinformatik leistet der Beitrag zweierlei. Erstens stellt er ein empirisch fundiertes Set von Meta-Requirements und Designprinzipien für Informationssysteme, die auf Verfahren erklärbarer KI basieren, aus Sicht von Lehrkräften in einem bislang wenig beforschten Anwendungskontext bereit. Damit knüpft er an Vorarbeiten zur erklärbaren Desinformationserkennung (Bezzaoui et al. 2025b) an und überträgt deren Logik auf einen pädagogischen Nutzungskontext. Zweitens zeigt der Beitrag, wie sich Erklärbarkeit nicht allein als technische Eigenschaft, sondern als zielgruppenbezogene Gestaltungsaufgabe fassen lässt: Die Designprinzipien formulieren Transparenz konsequent aus dem Verständnis der Nutzenden und bestätigen damit, dass Transparenz ohne Adressatenbezug wirkungslos bleibt (Schmitt et al. 2025).

Methodisch verdeutlicht der Beitrag den Wert eines partizipativen DSR-Vorgehens für komplexe soziotechnische Problemräume. Insbesondere die frühzeitige Einbindung von Stakeholdern abseits der Hauptnutzer*innen erwies sich als geeignet, um Gestaltungswissen schrittweise aufzubauen und für weitere Iterationen mit Jugendlichen zielgerichtet und praxisorientiert nutzbar zu machen.

Für die Bildungspraxis zeigt der Anforderungskatalog, dass der erfolgreiche Einsatz KI-gestützter Werkzeuge für Lehrkräfte nicht allein von der Software abhängt, sondern ebenso von didaktischen Materialien, Qualifizierungsangeboten und klar geregelten Supportstrukturen. Die abgeleiteten Designprinzipien formulieren hierfür konkrete Orientierungspunkte, darunter eine niedrigschwellige und geräteunabhängige Bedienung, eine transparente Kommunikation von Unsicherheiten sowie ein pädagogisch sensibles Wording. Für Systementwickler*innen folgt daraus, dass

die pädagogischen und organisatorischen Rahmenbedingungen des Schulunterrichts frühzeitig erhoben und systematisch in die Gestaltung KI-gestützter Systeme einbezogen werden müssen, um die Voraussetzungen für deren praktische Einsetzbarkeit und wirksame Nutzung zu schaffen.

Die Ergebnisse sind vor dem Hintergrund der frühen Projektphase zu lesen. Bislang wurden die Anforderungen ausschließlich mit Lehrkräften und weiteren Bildungsakteur*innen erhoben. Die ausstehende Perspektive der Schüler*innen wird die Meta-Requirements und Designprinzipien erweitern und präzisieren. Zudem beruht die Analyse auf Workshops mit begrenzter Teilnehmer*innenzahl, sodass die abgeleiteten Prinzipien als vorläufig und im weiteren Prozess zu validieren zu verstehen sind. Partizipative Gestaltung lebt von mehrfachen Iterationen und endet nicht mit der Anforderungserhebung. Solange das Artefakt nicht im realen Unterrichtsbetrieb erprobt ist, bleiben Aussagen über seine didaktische Wirksamkeit notwendigerweise begrenzt. Weiterhin müssen in der folgenden Entwicklung mögliche Biases der KI mitbedacht und ihre Erklärungen in einem Rahmen gestaltet werden, der nicht zur Auslagerung kognitiver Prozesse an die KI einlädt. In den folgenden Entwicklungsschritten werden hierfür verschiedene Lerneinheiten und mögliche Erklärungsansätze mit den Jugendlichen evaluiert.

In den nächsten Projektphasen werden die Anforderungen in die Entwicklung des Dashboards überführt und gemeinsam mit Jugendlichen weiterentwickelt und evaluiert. Im nächsten Schritt werden in partizipativen Workshops mit vier Klassen unterschiedlicher Schulformen nach einer inhaltlichen Einführung in Desinformation und KI-gestützte Lösungsansätze gemeinsam Prototypen gebaut und potenzielle Funktionen und Designoptionen des Dashboards evaluiert. Letztere konkretisieren die Anforderungen der Lehrkräfte und illustrieren die Umsetzung allgemein formulierter Designprinzipien in die Lebenswelt der Jugendlichen. Der partizipative Forschungs- und Entwicklungsansatz kann dazu beitragen, das System an pädagogische und organisatorische Rahmenbedingungen anzupassen, gezielt auf die angestrebten Lernziele auszurichten und damit seine langfristige Nutzung im Schulkontext zu unterstützen. Damit verbindet DisCoBoard pädagogische, gesellschaftliche und technische Perspektiven zu einem partizipativen Gestaltungsansatz für auf erklärbarer KI gestützte Lernwerkzeuge zur Erkennung und kritischen Reflexion manipulativer Inhalte und leistet einen Beitrag zur Stärkung demokratischer Kompetenzen im schulischen Kontext.

Autorenrollen Isabel Bezzaoui und Jonas Fegert entwickelten die Konzeption des Beitrags (Conceptualization). Die Datenerhebung (Investigation) sowie die Datenauswertung (Formal analysis) erfolgten durch Isabel Bezzaoui. Den ersten Entwurf des Manuskripts verfassten Isabel Bezzaoui und Maximiliane Linde (Writing – original draft); die kritische Durchsicht und Überarbeitung übernahmen Maximiliane Linde und Jonas Fegert (Writing – review & editing). Die Projektleitung lag bei Isabel Bezzaoui und Jonas Fegert (Supervision), die Projektkoordination bei Isabel Bezzaoui und Maximiliane Linde (Project administration). Alle Autorinnen und Autoren haben die finale Fassung des Manuskripts gelesen und ihrer Veröffentlichung zugestimmt.

Förderung Die dieser Veröffentlichung zugrunde liegende Forschung wurde durch das Programm „Demokratie Leben!“ des Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend im Rahmen des Projekts „DisCoBoard“ gefördert.

Open Access funding enabled and organized by Projekt DEAL.

Datenverfügbarkeit Alle dieser Arbeit zugrunde liegenden Daten sind in diesem Artikel enthalten.

Einhaltung ethischer Richtlinien

Interessenkonflikt Die Autor*innen erklären, dass kein Interessenkonflikt besteht.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf http://creativecommons.org/licenses/by/4.0/deed.de.

Literatur

Barzilai S, Chinn CA (2020) A review of educational responses to the "post-truth" condition: Four lenses on "post-truth" problems. Educational Psychologist 55:107–119. https://doi.org/10.1080/00461520.2020.1786388

Bezzaoui I, Jarvers L, Weinhardt C, Fegert J (2025a) Designing Deepfake Detection Systems: Practitioner Requirements Across Sectors. ICIS 2025 Proceedings.

Bezzaoui I, Stein C, Weinhardt C, Fegert J (2025b) Explainable AI for online disinformation detection: Insights from a design science research project. Electronic Markets 35:66. https://doi.org/10.1007/s12525-025-00799-3

Breakstone J, McGrew S, Smith M (2024) Measuring what matters: Investigating what new types of assessments reveal about students' online source evaluations. Harvard Kennedy School Misinformation Review. https://doi.org/10.37016/mr-2020-133

Brehmer M, Buonassisi V (2024) Educators' Friend—Applying Generative AI to Create Effective Digital Learning Objects for Information Security Education: Toward Initial Design Principles. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2024.003

Chaushi BA, Selimi B, Chaushi A, Apostolova M (2023) Explainable Artificial Intelligence in Education: A Comprehensive Review. In: Longo L (Hrsg) Explainable Artificial Intelligence. Communications in Computer and Information Science, Bd 1902. Springer, Cham, S 48–71 https://doi.org/10.1007/978-3-031-44067-0_3

Chen L, Chen P, Lin Z (2020) Artificial Intelligence in Education: A Review. IEEE Access 8:75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chiu TKF, Chai C (2020) Sustainable Curriculum Planning for Artificial Intelligence Education: A Self-Determination Theory Perspective. Sustainability 12:5568. https://doi.org/10.3390/su12145568

Cober R, Tan E, Slotta J, So HJ, Könings KD (2015) Teachers as participatory designers: two case studies with technology-enhanced learning environments. Instr Sci 43:203–228. https://doi.org/10.1007/s11251-014-9339-0

Coiro J, Coscarelli C, Maykel C, Forzani E (2015) Investigating Criteria That Seventh Graders Use to Evaluate the Quality of Online Information. J Adolesc Adult Lit 59:287–297. https://doi.org/10.1002/jaal.448

Gasevic D, Jovanovic J, Pardo A, Dawson S (2017) Detecting learning strategies with analytics: Links with self-reported measures and academic performance. Learning Analytics 4:113–128. https://doi.org/10.18608/jla.2017.42.10

Gill SS, Xu M, Patros P et al (2024) Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. Internet of Things and Cyber-Physical Systems 4:19–23. https://doi.org/10.1016/j.iotcps.2023.06.002

Gouseti A, James F, Fallin L, Burden K (2025) The ethics of using AI in K-12 education: A systematic literature review. Technology, Pedagogy and Education 34:161–182. https://doi.org/10.1080/1475939X.2024.2428601

Gregor S, Jones D (2007) The Anatomy of a Design Theory. Journal of the Association for Information Systems 8:312–335. https://doi.org/10.17705/1jais.00129

Guess AM, Lerner M, Lyons B, Montgomery JM, Nyhan B, Reifler J, Sircar N (2020) A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. Proceedings of the National Academy of Sciences 117:15536–15545. https://doi.org/10.1073/pnas.1920498117

Hameleers M (2022) Separating truth from lies: Comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands. Information, Communication & Society 25:110–126. https://doi.org/10.1080/1369118X.2020.1764603

Herrero-Curiel E, La-Rosa L (2022) Secondary education students and media literacy in the age of disinformation. Comunicar 30:95–106. https://doi.org/10.3916/C73-2022-08

Hevner AR, March ST, Park J, Ram S (2004) Design Science in Information Systems Research. MIS Quarterly 28:75–105. https://doi.org/10.2307/25148625

Huang R, Yin Y, Zhou N, Lang F (2026) Artificial intelligence in K-12 education: An umbrella review. Computers and Education: Artificial Intelligence 10:100519. https://doi.org/10.1016/j.caeai.2025.100519

Jeong SH, Cho H, Hwang Y (2012) Media Literacy Interventions: A Meta-Analytic Review. Journal of Communication 62:454–472. https://doi.org/10.1111/j.1460-2466.2012.01643.x

Jones-Jang SM, Mortensen T, Liu J (2021) Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. American Behavioral Scientist 65:371–388. https://doi.org/10.1177/0002764219869406

Khosravi H, Shum SB, Chen G, Conati C, Tsai YS, Kay J, Knight S, Martinez-Maldonado R, Sadiq S, Gašević D (2022) Explainable Artificial Intelligence in education. Computers and Education: Artificial Intelligence 3:100074. https://doi.org/10.1016/j.caeai.2022.100074

Köbis NC, Doležalová B, Soraperra I (2021) Fooled twice: People cannot detect deepfakes but think they can. iScience. https://doi.org/10.1016/j.isci.2021.103364

Luckin R, Holmes W, Griffiths M, Forcier L (2016) Intelligence Unleashed: An argument for AI in Education. Pearson, London

Mikeladze T, Meijer PC, Verhoeff RP (2024) A comprehensive exploration of artificial intelligence competence frameworks for educators: A critical review. Eur J Educ 59:e12663. https://doi.org/10.1111/ejed.12663

Nygren T, Guath M (2022) Students Evaluating and Corroborating Digital News. Scandinavian Journal of Educational Research 66:549–565. https://doi.org/10.1080/00313831.2021.1897876

Nygren T, Frau-Meigs D, Corbu N, Santoveña-Casal S (2022) Teachers' views on disinformation and media literacy supported by a tool designed for professional fact-checkers: Perspectives from France, Romania, Spain and Sweden. SN Soc Sci 2:40. https://doi.org/10.1007/s43545-022-00340-9

Orosz G, Faragó L, Paskuj B, Krekó P (2024) Strategies to combat misinformation: Enduring effects of a 15-minute online intervention on critical-thinking adolescents. Comput Human Behav 159:108338. https://doi.org/10.1016/j.chb.2024.108338

Rau MA, Premo AE (2025) Systematic Review of Educational Approaches to Misinformation. Educ Psychol Rev 37:43. https://doi.org/10.1007/s10648-025-10012-8

Sanders EBN, Stappers PJ (2008) Co-creation and the new landscapes of design. CoDesign 4:5–18. https://doi.org/10.1080/15710880701875068

Schmitt V, Bezzaoui I, Jakob C et al (2025) Beyond Transparency: Evaluating Explainability in AI-Supported Fact-Checking. In: Proceedings of the 4th ACM International Workshop on Multimedia AI against Disinformation (MAD'25), co-located with ICMR'25. ACM. https://doi.org/10.1145/3731715.3734512

Soßdorf A, Stein C, Bezzaoui I, Fegert J (2024) Literacies Against Fake News: Examining the Role of Data Literacy and Critical Media Literacy to Counteract Disinformation. MedienPädagogik 59:55–76. https://doi.org/10.21240/mpaed/59/2024.04.10.X

Spinuzzi C (2005) The Methodology of Participatory Design. Technical Communication 52:163–174

Su J, Yang W (2024) AI literacy curriculum and its relation to children's perceptions of robots and attitudes towards engineering and science: An intervention study in early childhood education. Journal of Computer Assisted Learning 40:241–253. https://doi.org/10.1111/jcal.12867

Timotheou S, Miliou O, Dimitriadis Y et al (2023) Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. Educ Inf Technol 28:6695–6726. https://doi.org/10.1007/s10639-022-11431-8

Vieriu AM, Petrea G (2025) The Impact of Artificial Intelligence (AI) on Students' Academic Development. Educ Sci 15:343. https://doi.org/10.3390/educsci15030343

Wardle C (2018) The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder. Digit Journal 6:951–963. https://doi.org/10.1080/21670811.2018.1502047

Weinhardt C, Fegert J, Hinz O, van der Aalst WMP (2024) Digital Democracy: A Wake-Up Call. Business & Information Systems Engineering 66:127–134. https://doi.org/10.1007/s12599-024-00862-x

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