

Echt oder gemacht?

Künstliche Intelligenz, Manipulation und Vertrauen in Social Media

Kurzbeschreibung

Mit Künstlicher Intelligenz können Bilder, Videos, Texte, Stimmen und ganze Geschichten erzeugt werden, die täuschend echt wirken. In sozialen Medien begegnen Jugendliche zunehmend Beiträgen, die KI-generiert, manipuliert oder automatisiert verbreitet sind – oft ohne, dass dies klar erkennbar ist. Auch extremistische Akteure nutzen diese technischen Möglichkeiten, um Emotionen zu verstärken, Glaubwürdigkeit zu simulieren oder bestimmte Deutungen plausibel erscheinen zu lassen.

Unsere Methode sensibilisiert Jugendliche für KI-basierte Inhalte im Netz, zeigt Möglichkeiten der Manipulation auf und stärkt die Fähigkeit, Inhalte kritisch zu hinterfragen. Ziel ist nicht, KI-generierte Beiträge eindeutig von „realen“ Inhalten unterscheiden zu können, sondern einen reflektierten, kritischen und kompetenten Umgang mit Unsicherheit, Täuschungspotenzialen und Vertrauen im digitalen Raum zu entwickeln.

Ziele der Methode

Die Methode unterstützt die Teilnehmenden dabei:

- KI-generierte und manipulierte Beiträge in sozialen Medien bewusster wahrzunehmen
- zu verstehen, wie Glaubwürdigkeit und Authentizität technisch erzeugt werden können
- zu erkennen, wie leicht Inhalte heute verändert oder neu erstellt werden können
- Strategien für einen reflektierten Umgang mit Unsicherheit und Zweifel zu entwickeln

Zielgruppe

- Jugendliche, unabhängig von Vorwissen oder Intensität der Social-Media-Nutzung

Rahmenbedingungen

- **Dauer:** 60–90 Minuten
- **Gruppengröße:** 6–20 Teilnehmende
- **Setting:** Gruppenraum mit WLAN
- **Technik:** Smartphones / Tablets der Einrichtungen
- **Material:** Moderationskarten, Stifte
- **Empfohlenes KI-Tool:** [duck.ai](#) (keine Speicherung privater Daten)

Darüber hinaus sollten sich Fachkräfte mit **Datenschutzbestimmungen** der jeweilig verwendeten KI-Tools befassen. Eine Einführung in das Thema KI und Informationen zum Datenschutz finden sich in der interaktiv plus Ausgabe Künstliche Intelligenz – Einstieg in die kreative Arbeit mit KI-Modellen. Zum Nachlesen den nebenstehenden QR-Code scannen.



INFO

In der Praxis haben sich **Einrichtungsaaccounts** bewährt, um die Daten und die Persönlichkeitsrechte der Jugendlichen zu schützen. Es kann auch sinnvoll sein, technische Geräte (Smartphones/Tablets) der Einrichtung zu nutzen und einen Beamer zu verwenden. Natürlich können auch die Endgeräte der Teilnehmenden genutzt werden, wenn alle einverstanden sind und es sinnvoll erscheint.



Echt oder gemacht?

Künstliche Intelligenz, Manipulation und Vertrauen in Social Media

Inklusiv gedacht

Damit auch Jugendliche mit einer Behinderung teilhaben können, haben wir hier einige Tipps zusammengestellt:

- Informationen in verschiedenen Formaten anbieten, z. B. Verwendung von Einfacher Sprache, visuelle Symbole / Piktogramme und Hilfsmittel, Gebärdensprache.
- Medieninhalte barrierefrei gestalten, z. B. durch Untertitel und Audiodeskription. Dabei darauf achten, dass Untertitel und Hintergrund genug Kontrast haben, damit die Untertitel deutlich zu lesen sind.
- Weitere Informationen und Hinweise für die inklusive Gestaltung von Medienprojekten:
 - Netzwerk Inklusion mit Medien (nimm!): www.inklusive-medienarbeit.de
 - nimm!-Akademie: www.nimm-akademie.nrw
 - Tool-Tipps bei Instagram unter [@deinnimm](https://www.instagram.com/deinnimm)

Vorbereitung

Die Fachkraft macht sich grundlegend mit folgenden Begriffen vertraut (ohne technische Vertiefung):

- künstliche Intelligenz
- KI-generierte Inhalte
- Manipulation
- Deepfakes

Zur fachlichen Vorbereitung und Orientierung eignen sich zusätzlich folgende Angebote:

- klicksafe.de – Erklärungen und Materialien zu KI, Deepfakes und digitaler Manipulation
- jugendschutz.net – Analysen zu Manipulationsstrategien und extremistischen Online-Praktiken
- bpb.de – Einordnungen zu KI, Desinformation und Medienkompetenz

Die Methode selbst arbeitet bewusst **ohne umfangreiche Beispielmaterien**, um Überforderung zu vermeiden und den Fokus auf Wahrnehmung und Reflexion zu legen.

Ablauf der Methode

Schritt 1 Einstieg – Kann ich dem trauen?

Die Teilnehmenden werden eingeladen, über ihre bisherigen Erfahrungen mit KI im Netz zu sprechen.

Mögliche Impulsfragen:

- Wo begegne ich KI in meinem Alltag?
- Habe ich schon Beiträge gesehen, bei denen ich mir unsicher war, ob sie echt sind?
- Woran merke ich, dass ich einem Inhalt vertraue?

Der Austausch erfolgt offen und ohne Bewertung.

Schritt 2 Erkennen und hinterfragen – Was wirkt glaubwürdig?

Die Gruppe setzt sich mit der Frage auseinander, wodurch Inhalte glaubwürdig wirken können – unabhängig davon, ob sie echt oder KI-generiert sind.

Gemeinsam wird reflektiert:

- Welche Details erzeugen Vertrauen?
- Was lässt Beiträge authentisch erscheinen?
- Wo entstehen erste Zweifel?

Dabei wird deutlich gemacht:

Nicht jede Manipulation ist erkennbar – und genau darin liegt eine zentrale Herausforderung im Umgang mit KI-basierten Inhalten.

Echt oder gemacht?

Künstliche Intelligenz, Manipulation und Vertrauen in Social Media

Schritt 3 Selbst ausprobieren – Wie leicht geht Manipulation?

In Kleingruppen probieren die Teilnehmenden selbst einfache KI-Anwendungen aus, z. B. Text- oder Bildgeneratoren mit neutralen Aufgabenstellungen.

Mögliche Aufgaben:

- ➔ Einen kurzen Text verfassen, in dem „Fakten“ frei erfunden sind.
- ➔ Ein Bild generieren und es mit einem emotionalisierenden Titel versehen.

Anschließend wird gemeinsam reflektiert:

- ➔ Wie schnell entsteht ein glaubwürdiger Inhalt?
- ➔ Was überrascht?
- ➔ Wo liegt das Risiko, wenn solche Beiträge gezielt eingesetzt werden?

Abhängig von Gruppensituation, Zeit und technischer Ausstattung kann optional auch ein einfacher KI-basierter Videogenerator genutzt werden. Dies setzt jedoch einen bereits reflektierten Umgang mit der Technik voraus und ist kein notwendiger Bestandteil der Methode. Durchführende sollten sich hierbei unbedingt mit Fragen zum Datenschutz auseinandersetzen (Infos dazu finden sich u.a. in der o.g. interaktiv plus).

Für den Arbeitsschritt kann duck.ai – anonyme Nutzung möglich – verwendet werden.

Schritt 4 Einordnung und Handlungsfähigkeit – Wie gehe ich künftig damit um?

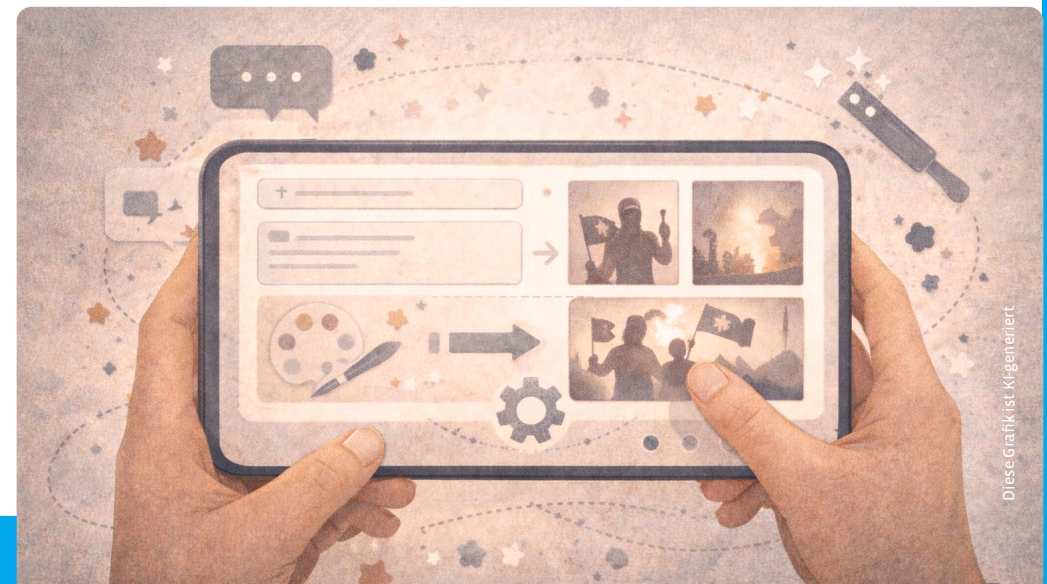
Im Plenum wird gemeinsam überlegt:

- ➔ Was kann ich tun, wenn ich mir bei einem Beitrag unsicher bin?
- ➔ Wann hilft es, innezuhalten und einen zweiten Blick auf Inhalte zu werfen?
- ➔ Wie kann ich mit Zweifeln umgehen, ohne allem misstrauisch zu begegnen?

Ziel ist es, individuelle Strategien für einen bewussten und reflektierten Umgang mit KI-Inhalten zu entwickeln.

Reflexionsfragen (Beispiele)

- ➔ Muss ich immer wissen, ob ein Inhalt echt ist?
- ➔ Wann kann ich mir sicher sein?
- ➔ Was hilft mir, mich nicht emotional manipulieren zu lassen?



Echt oder gemacht?

Künstliche Intelligenz, Manipulation und Vertrauen in Social Media

Pädagogische Hinweise

- ➔ Unsicherheit ist kein Zeichen von Naivität, sondern ein wichtiger Schutzmechanismus.
- ➔ Nicht jede Täuschung ist sofort erkennbar.
- ➔ Die Rezeption einzelner KI-Inhalte führt nicht automatisch zu Radikalisierung.
- ➔ Offenheit, Austausch und Nachfragen stärken die Orientierung.

Fazit

Unsere Methode stärkt Jugendliche darin, KI-basierte Beiträge und Inhalte in sozialen Medien bewusster wahrzunehmen und einzuordnen. Sie vermittelt keine Scheinsicherheit, sondern fördert einen reflektierten Umgang mit Unsicherheit, Vertrauen und Manipulation – zentrale Kompetenzen in einer zunehmend KI-geprägten Mediensozialisation.



Diese Grafik ist KI-generiert



Unser Autor

Mehmet Koç B.A. Soziale Arbeit und Islamische Studien, sowie M.A. Forschung in der Sozialen Arbeit. Er entwickelt und erprobt innovative Ansätze digitaler Präventionsarbeit auf Social Media, u.a. zur Dekonstruktion von ‚islamistischen Narrativen‘ – und bringt seine Perspektiven in den fachlichen und gesellschaftlichen Diskurs ein.

Aktuell ist er unabhängiges Mitglied im Expertenbeirat von KN:IX connect | Verbund Islamismusprävention und Demokratieförderung.



Praxistipps für die Medienarbeit
mit jungen Geflüchteten



Lizenz: CC BY-NC-ND 2.0

Ein Angebot der



Gefördert vom

Ministerium für Kinder, Jugend, Familie,
Gleichstellung, Flucht und Integration
des Landes Nordrhein-Westfalen

