

# EINEN EMPFEHLUNGSALGORITHMUS SIMULIEREN – DER UNAUSWEICHLICHE WEG VON HARMLOSEN INTERESSEN ZUR EXTREMISTISCHEN BLASE?

Oliver Held

Ein KI-gestütztes Simulationsspiel zur Thematik der Radikalisierung durch Empfehlungsalgorithmen.

Thema	Radikalisierung durch Empfehlungsalgorithmen
Schulstufe	9. Schulstufe
Kompetenz	politische Handlungs- und Urteilskompetenz
didaktische Zielsetzung	Lernende sollen nachvollziehen, wie Empfehlungsalgorithmen auf Social Media funktionieren und wie sie schleichend zur Radikalisierung beitragen können. Durch die Simulation erfahren sie, wie Filterblasen entstehen, welche rhetorischen Strategien extremistische Inhalte nutzen und welche Rolle algorithmische Logiken dabei spielen. Ziel ist es, kritische Sensibilität gegenüber manipulativen Inhalten zu entwickeln und ein reflektiertes Medienverhalten zu fördern.
Prompts	siehe Link
Link, Zusatzmaterial etc.	<a href="https://app.fobizz.com/ai/chats/new?assistantId=54ad2b74-8504-4b75-8dd8-adf195884ca4">https://app.fobizz.com/ai/chats/new?assistantId=54ad2b74-8504-4b75-8dd8-adf195884ca4</a>
Online unter	<a href="http://www.politik-lernen.at/empfehlungsalgorithmus-simulieren">www.politik-lernen.at/empfehlungsalgorithmus-simulieren</a>

Jugendliche konsumieren auf sozialen Medien täglich Videos, deren Inhalte sie vermeintlich bewusst auswählen – doch tatsächlich bestimmen Empfehlungsalgorithmen als Gatekeeper, welche Inhalte präsentiert werden (Oertel et al. 2022: 85f.).

Aus ökonomischer Sicht ist es naheliegend, diese Systeme auf maximales „Engagement“ zu optimieren (Nemitz/Pfeffer 2020: 209-211), wodurch die Wahrscheinlichkeit steigt, dass sie emotionalisierende sowie polarisierende und damit extreme Inhalte bevorzugt empfehlen (Frühbrodt/Floren 2019: 1).

**Der KI-Chatbot „Algorithmus-Simulator“ veranschaulicht stark vereinfachend, wie schnell problematische Inhalte empfohlen werden können.**

⚙️ Basierend auf beliebigen Interessen der Lernenden, werden ihnen wiederholt zwei fiktive individualisierte Videotitel vorgeschlagen.

⚙️ Die Titel nähern sich systematisch problematischen Inhalten an und enthalten zunehmend typische rhetorische Strategien oder (stereotype!) Feindbilder.

**Die Lernenden erleben so im Zeitraffer, wie sie mit jeder Entscheidung tiefer in radikalere Inhalte hineingezogen werden.**

**Diese Eskalation von neutralen zu extremen Inhalten verläuft zwar gezielt und übertrieben schnell, entspricht im Kern aber Mechanismen, die hinsichtlich realer Empfehlungsalgorithmen immer wieder diskutiert werden** (Mahrt 2024: 5).

Nach der zehnten Runde bietet die KI an, eine Analyse bereitzustellen, wie Entscheidungen vom System aufgegriffen wurden, um gezielt in ideologische Sphären zu lenken. Diese Analyse sollte zunächst von den Lernenden geleistet werden.

Anschließend lohnt sich eine **Diskussion** darüber, dass derlei Algorithmen nicht nur Rezipienten beeinflussen, sondern auch Content-Creator dafür belohnen, immer polarisierendere und extremere Inhalte zu produzieren (Mahrt 2025, 28).

**Das Simulationsspiel ist deterministisch angelegt**, es führt zwangsläufig in eine extremistische Blase, da nach einigen Runden ausschließlich ideologisch gefärbte Titel angeboten werden.

⚙️ Im Unterricht sollte kritisch diskutiert werden, ob reale Algorithmen ebenfalls unausweichlich sind – zwar beeinflusst das eigene Verhalten einen Teil der Empfehlungen, doch sie werden ergänzt durch gesponsorte, trendende oder besonders interaktionsstarke (extreme) Inhalte, was die Frage aufwerfen muss: Wie frei entscheide ich wirklich in Sozialen Medien?

**Das Spiel endet, wenn die internen Sicherheitsfilter von OpenAI oder Fobizz greifen** (siehe O. V. 2025).

➔ **Dieser Moment des Abbruchs verdeutlicht den Lernenden, wie nah sie gefährlichen Inhalten gekommen sind und bietet Anlass, kritisch zu diskutieren ob die vorhandenen Filter der Plattformen ausreichen und wer mit welchen Interessen diese Filter kontrolliert.**



## VERWEISE

- ⚙️ Frühbrodt, Lutz/Floren, Annette (2019): Unboxing YouTube. Ein Projekt der Otto Brenner Stiftung. Frankfurt a. M. Online: [https://opus4.kobv.de/opus4-fhws/frontdoor/deliver/index/docId/1786/file/AH98\\_YouTube\\_Fruehbrodt.pdf](https://opus4.kobv.de/opus4-fhws/frontdoor/deliver/index/docId/1786/file/AH98_YouTube_Fruehbrodt.pdf)
- ⚙️ Mahrt, Merja (2024): Digitale Demokratie. Lenken Soziale Medien die Meinungsbildung? Berlin. Online: [www.weizenbaum-library.de/server/api/core/bitstreams/7c887c89-d56c-4eb4-8aea-baa5a6730c3d/content](http://www.weizenbaum-library.de/server/api/core/bitstreams/7c887c89-d56c-4eb4-8aea-baa5a6730c3d/content)
- ⚙️ Mahrt, Merja (2025): Meinungsbildung und Soziale Medien. Berlin. Online: [www.weizenbaum-institut.de/media/Publikationen/Einzelpublikationen/broschuere\\_meinungsbildung\\_soziale\\_medien.pdf](http://www.weizenbaum-institut.de/media/Publikationen/Einzelpublikationen/broschuere_meinungsbildung_soziale_medien.pdf)
- ⚙️ Nemitz, Paul/Pfeffer, Matthias (2020): Prinzip Mensch. Bonn.
- ⚙️ Oertel, Britta/Diego, Dametto/Kluge, Jakob/Todt, Jan (2022): Algorithmen in digitalen Medien und ihr Einfluss auf die Meinungsbildung. Arbeitsbericht Nr. 204 des Büros für Technikfolgen-Abschätzung beim Deutschen Bundestag. Berlin. Online: <https://publikationen.bibliothek.kit.edu/1000154065/149886157>
- ⚙️ O. V. (2025): Welche Inhaltsfilter gibt es bei den KI-Tools? In: fobizz. Online: <https://app.fobizz.com/de/help/articles/258-ki-inhaltsfilter>