



Stärken entwickeln. Gefährdungen begegnen.

# Chatbots, Filter und Deepfakes – KI auf Social Media. Eine neue Herausforderung für den Jugendschutz?



Veranstalter:

Aktion Jugendschutz, Landesarbeitsstelle Bayern e.V.

Stärken entwickeln. Gefährdungen begegnen.

# Aktion Jugendschutz

## Landesarbeitsstelle Bayern e.V.

### Über uns

- Seit über 50 Jahren eine Fachinstitution im erzieherischen Kinder und Jugendschutz
- Institutionell gefördert vom Bayerischen Staatsministerium für Familie, Arbeit und Soziales



Bayerisches Staatsministerium für  
Familie, Arbeit und Soziales

- Instagram:  
jugendschutz\_bayern

### Mitglieder

Unsere Arbeit wird getragen und unterstützt von einer großen Zahl von Mitgliedern:

- Die Spitzenverbände der Freien Wohlfahrtspflege in Bayern
- Der Bayerische Jugendring
- Die kommunalen Spitzenverbände
- Verbände der Kinder- und Jugendhilfe in Bayern
- Fachinstitutionen
- Einzelpersonlichkeiten

### Referate

Unsere Schwerpunktthemen sind:

- Suchtprävention
- Gewaltprävention
- Jugendmedienschutz/ Medienpädagogik
- Prävention von Glücksspielsucht
- Prävention von sexueller Gewalt und Sexualpädagogik
- Projekt ELTERN TALK
- Präventionsangebot FREUNDE

Stärken entwickeln. Gefährdungen begegnen.

## Aktion Jugendschutz Landesarbeitsstelle Bayern e.V.

- Transfer zwischen Wissenschaft und Praxis
- Fortbildungen und Qualifizierungen ( ein- und mehrtägige Präsenz- und Onlineseminare, langfristige Qualifizierungen)
- Veröffentlichungen von Fachliteratur und Arbeitshilfen
- Konzipierung und Implementierung von Projekten und Programmen (Projekt ELTERNTALK, Präventionsprogramm Freunde, stationär 4.0,..)





Stärken entwickeln. Gefährdungen begegnen.

## Chatbot & Deepfakes, was ist KI auf social media?

- Habt ihr schon Erfahrungen mit KI auf Social Media gemacht?
- Sind euch bereits KI generierte Inhalte begegnet?

### **Kurze Recherche Aufgabe:**

Sucht nach KI generierten Inhalten (auf Social Media), was könnt ihr selbst finden?



Stärken entwickeln. Gefährdungen begegnen.



Julian van Dieken, *A Girl With Glowing Earrings*  
The Hague's Mauritshuis museum



KI generiertes Bild vom Papst auf Reddit aufgetaucht, Verfasser unbekannt



TikTok Filter „Bold Glamour“

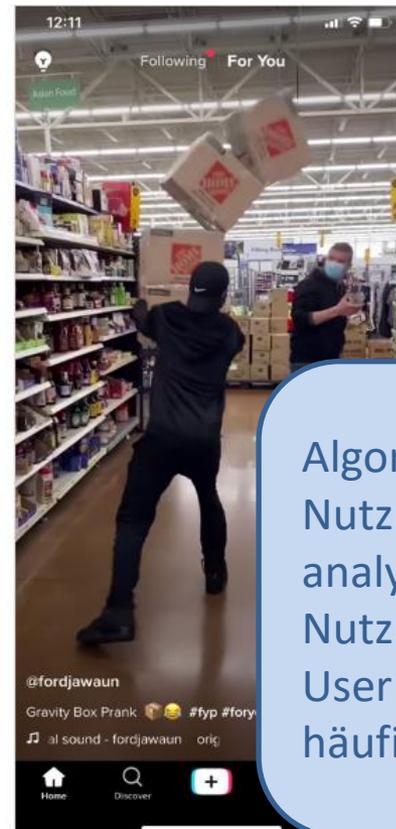
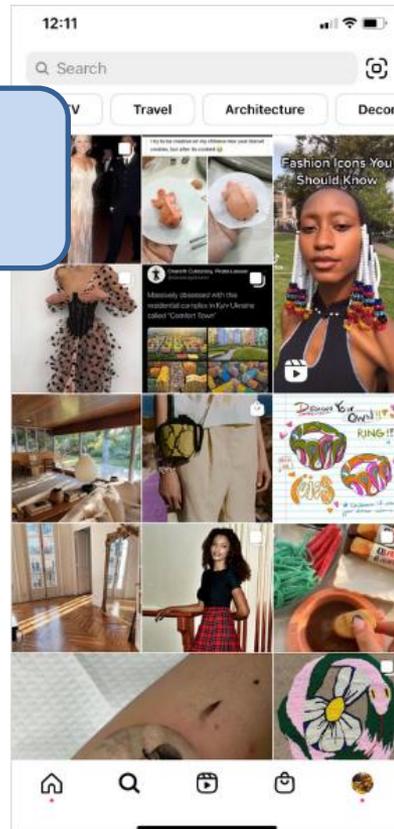


Miquela (rechts) ist AI-Influencerin (hier für Werbekampagne Calvin Klein)

Stärken entwickeln. Gefährdungen begegnen.

# Chatbot & Deepfakes, was ist KI auf social media?

KI ist keine neue Technologie in Sozialen Medien



Algorithmen, die Inhalte und Nutzungsverhalten analysieren um sie den Nutzungsgewohnheiten von User\*innen anzupassen, sind häufig KI generiert

<https://later.com/blog/tiktok-for-you-page/>



Stärken entwickeln. Gefährdungen begegnen.

## Chatbot & Deepfakes, was ist KI auf social media?

### Neue Formen von KI auf Social Media

- Snapchat Funktion „**My AI**“ ist ein KI generierter Chatbot der mit User\*innen chatten und Fragen beantworten kann (ähnlich wie ChatGPT)
- Andere Anbieter wollen der Idee von Snapchat folgen
- Meta arbeitet an **Sprachmodellen und Bild-/Videobearbeitungstools**

## Stärken entwickeln. Gefährdungen begegnen.

The image shows a screenshot of the Snapchat Support website. The browser address bar displays the URL: <https://help.snapchat.com/hc/en-us/articles/13266788358932-What-is-My-AI-on-Snapchat-and-how-do-I-use-it>. The website header includes the Snapchat logo and navigation links: Download, Spotlight, Stories, Map, Ads, Spectacles, Filters & Lenses, Lens Studio, Kit, Snapcodes, Support, and Store. A yellow search bar contains the text "What can we help you with?".

The main content area features a sidebar with a "Managing My AI" section containing the following links:

- How do I add My AI back to my Chat feed?
- What is My AI on Snapchat and how do I use it?
- How do I access My AI on Snapchat?
- How do I unpin or remove My AI with Snapchat+?
- How do I re-pin or add My AI

The main article is titled "What is My AI on Snapchat and how do I use it?". The breadcrumb trail is: Snapchat Support > Using Snapchat > My AI > Managing My AI.

The article text includes:

My AI is a chatbot currently available to Snapchatters.

In a chat conversation, My AI can answer a burning trivia question, offer advice on the perfect gift for your BFF's birthday, help plan a hiking trip for a long weekend, or suggest what to make for dinner.

We're constantly working to improve and evolve My AI, but it's possible My AI's responses may include biased, unhelpful, or inappropriate content. Because My AI is an evolving feature, you should always independently verify information before relying on any advice, and you should not share confidential or sensitive information with My AI.

If you have any feedback or suggestions, please reach out to our team by long pressing on any response from My AI to share more on how we can improve while we continue to train My AI.

My AI uses generative AI (GenAI) technology, with additional safety enhancements and controls unique to Snapchat. We are committed to transparency when AI is involved in content creation. If you share content created with My AI, please let them know AI was involved in your work.

**Are My AI responses stored?**

My AI responses are not stored until you delete it.

At the bottom of the article, there is a link: [Screenshot https://www.help.snapchat.com](https://www.help.snapchat.com)

Overlaid on the bottom left is a mobile phone screen showing the My AI chatbot interface. The phone screen displays a chat window with a purple and blue background. The chatbot's profile is named "My AI" with a small icon. Below the profile are icons for camera, text, voice, and video. A notification at the bottom of the chat window reads: "Friendships are Private. Screenshotting friendship profiles will send a notification - just like with Snaps."



Stärken entwickeln. Gefährdungen begegnen.

## Chatbot & Deepfakes, was ist KI auf social media?

- Auf diversen Apps wie TikTok und Instagram tauchen bereits viele mit KI erstellte Inhalte in Form von Videos, Fotos und Audio auf
- Mit Hilfe von verschiedensten Apps, Programmen und Software kann man selbst **ohne Vorkenntnisse täuschend echt wirkende Videos, Fotos oder Stimmen bearbeiten/erstellen**
- Gesichter und Stimmen können getauscht werden, Personen sagen oder tun Dinge die sie eigentlich nicht getan haben
- Hier spricht man von **Deepfakes**, deren Qualität sich stetig verbessert und welche zunehmend schwerer zu erkennen sind

Stärken entwickeln. Gefährdungen begegnen.





Stärken entwickeln. Gefährdungen begegnen.

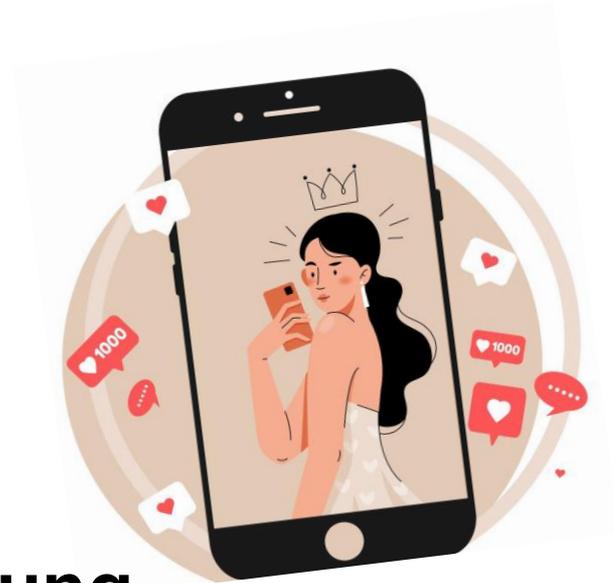
## Welche Chancen kann KI für Kinder und Jugendliche bieten?

- KI wird schon lange erfolgreich bei der **Suche nach jugendgefährdenden Inhalten** genutzt  
Entsprechende Inhalte werden erkannt und an Ermittlungsbehörden gemeldet
- **Jugendschutzfilter**
- Chat Bots können auch zu einem sichereren „begleiteten“ Umgang mit bestimmten Websites etc. helfen
- Kreative Fähigkeiten fördern z.B. eigene Inhalte erstellen, Möglichkeiten austesten (Medienkompetenz)
- Jugendliche unterstützen (z.B. Bewerbungen formulieren etc.)

Stärken entwickeln. Gefährdungen begegnen.

## Faszination und Funktion von Social Media für Jugendliche

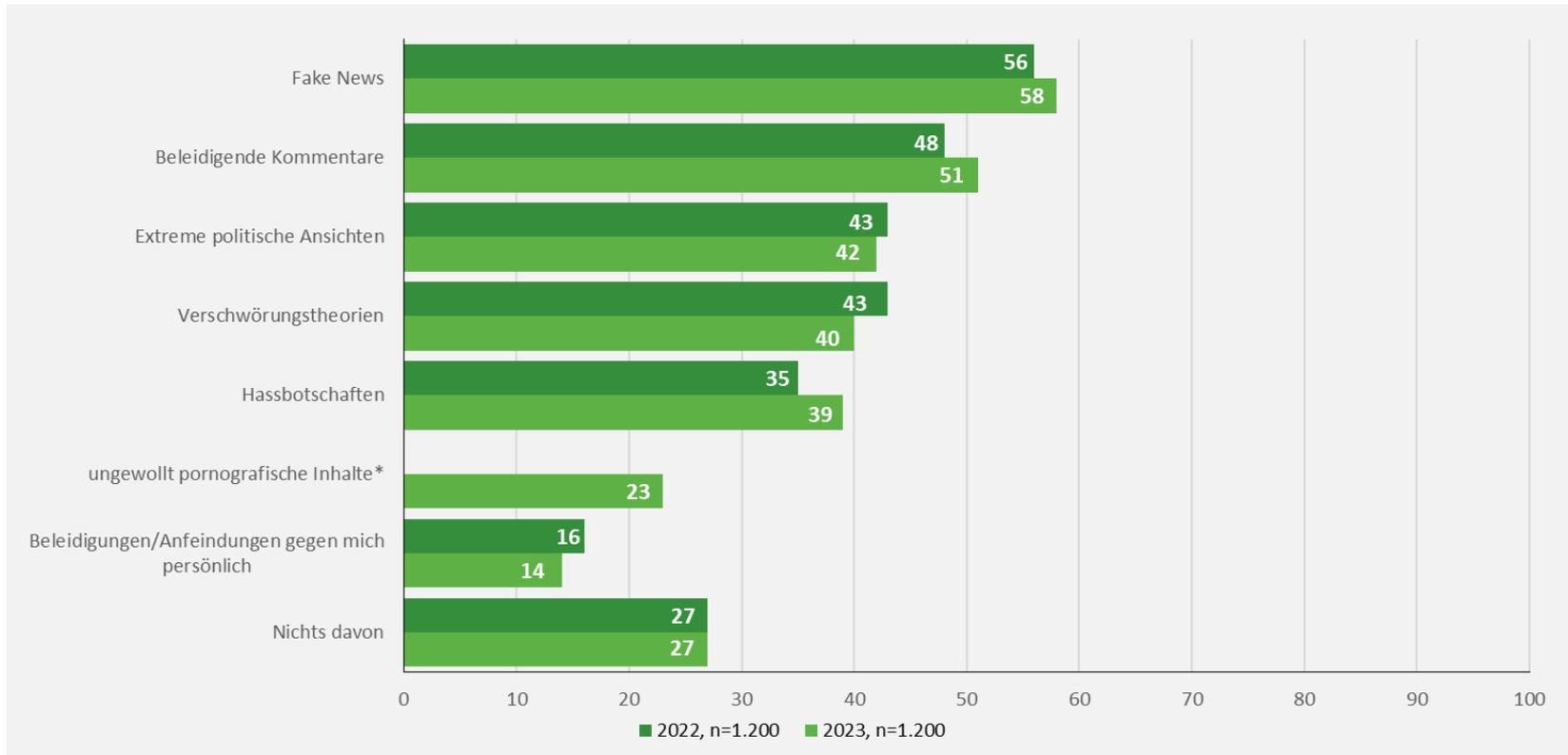
- Orientierung
- Unterhaltung
- **Information**
- Beziehungsaufbau
- **Identitätsfindung und –bildung**





## Stärken entwickeln. Gefährdungen begegnen.

### Mir sind im letzten Monat im Internet begegnet:



Quelle: JIM 2022, JIM 2023, \*2022 nicht abgefragt, Angaben in Prozent, Basis: alle Befragten

## Welche Probleme und Gefahren können durch Mediennutzung bei Jugendlichen entstehen?

- Preisgabe personenbezogener Informationen
- Jugendgefährdende Inhalte
- Extreme politische Ansichten
- Exzessive Nutzung
- Cyber Mobbing und Hate Speech
- Verstöße gegen Urheberrecht
- Kostenfallen, Glücksspiel

Stärken entwickeln. Gefährdungen begegnen.

## Welche Gefahren und Probleme ergeben sich durch KI für den Jugendschutz?

Was meint ihr?

- sind die Probleme durch KI immer noch die selben?
- Wo hat sich eventuell etwas verändert?





# Stärken entwickeln. Gefährdungen begegnen.



**You**

Welche Gefahren ergeben sich durch KI für den Jugendschutz



**ChatGPT**

Der Einsatz von Künstlicher Intelligenz (KI) kann verschiedene Herausforderungen und potenzielle Gefahren für den Jugendschutz mit sich bringen. Hier sind einige Aspekte, die berücksichtigt werden sollten:

## 1. Inhaltsfilter und Zensur:

- KI-Systeme können dazu verwendet werden, Inhalte zu filtern und zu moderieren, um unangemessene oder schädliche Inhalte für Jugendliche zu blockieren. Allerdings besteht die Gefahr von Fehlern, bei denen harmlose Inhalte fälschlicherweise als problematisch eingestuft werden oder umgekehrt.

## 2. Manipulation und Beeinflussung:

- Durch KI-gesteuerte Algorithmen können personalisierte Inhalte und Empfehlungen erstellt werden. Dies kann dazu führen, dass Jugendliche in sogenannten "Filterblasen" gefangen werden, in denen sie nur noch Informationen sehen, die ihren bestehenden Ansichten entsprechen. Dadurch könnten sie wenig mit unterschiedlichen Perspektiven in Kontakt kommen.

## 3. Datenschutz und Sicherheit:

- KI-Systeme sammeln und analysieren große Mengen an Daten. Im Kontext von Jugendschutz müssen Datenschutzbestimmungen besonders beachtet werden, um sicherzustellen, dass persönliche Informationen von Minderjährigen angemessen geschützt werden. Ein unsicherer Umgang mit Daten kann zu Risiken wie Identitätsdiebstahl oder Cybermobbing führen.

## 4. Suchtgefahr und exzessive Nutzung:

- KI-basierte Plattformen und Anwendungen können dazu neigen, Jugendliche durch personalisierte Inhalte dazu zu verleiten, ihre Zeit exzessiv zu nutzen. Dies kann zu Problemen wie Internetabhängigkeit, Schlafmangel und negativen

kann zu Problemen wie Internetabhängigkeit, Schlafmangel und negativen Auswirkungen auf die psychische Gesundheit führen.

## 5. Fehlende Aufsicht:

- Wenn KI-gesteuerte Systeme in sozialen Medien oder Online-Plattformen unangemessene Inhalte nicht effektiv erkennen und entfernen können, besteht die Gefahr, dass Jugendliche ungeschützt auf potenziell schädliche oder gefährliche Inhalte zugreifen.

Es ist wichtig, dass bei der Entwicklung und Implementierung von KI-Systemen für Jugendschutzmaßnahmen sorgfältig darauf geachtet wird, mögliche Risiken zu minimieren und die ethischen Richtlinien einzuhalten. Eine enge Zusammenarbeit zwischen Technologieunternehmen, Gesetzgebern, Eltern und Bildungseinrichtungen ist entscheidend, um eine sichere und positive Onlineumgebung für Jugendliche zu gewährleisten.



ChatGPTs Antwort auf die Frage:  
Welche Gefahren ergeben sich durch KI für den Jugendschutz?

## „neue“ Gefahren durch KI

- Altersgerechte Nutzung
- KI Kommunikation nicht klar zu erkennen, Verwechslung mit realer Person
- Riskante generierte Bildinhalte (z.B. Missbrauchsdarstellungen etc.)
- **Fehlerhafte Informationen/ Desinformationen/ gefilterte Informationen**  
(z.B. als Antworten bei Chatbot)
- Deepfakes (z.B. Desinformationen, Cybermobbing, pornografische Inhalte)



Stärken entwickeln. Gefährdungen begegnen.

## KI Gutachten Ergebnisse des Instituts für Europäischen Medienrecht (EMR)

„Bestehende Herausforderungen werden verschärft: Bereits jetzt werden KI-generierte Bilder zur politischen Manipulation eingesetzt. In Zukunft könnten Fake-Inhalte, die immer „echter“ aussehen, das Netz fluten und das Vertrauen in Informationen erschüttern. Durch Filterblasen droht KI die Meinungsbildung zu verengen. [...] KI ist nicht nur eine Herausforderung für die Meinungsbildung von Kindern und Jugendlichen, sondern auch für die Demokratie. Mit Blick auf die schnelle technologische Entwicklung ist zügiges handeln gefragt“

Dr. Jörg Ukrow, geschäftsführendes  
Vorstandsmitglied des Instituts für  
Europäisches Medienrecht

## Stärken entwickeln. Gefährdungen begegnen.

„[...] bei den rasanten technischen Entwicklungen darf der Schutz von Kindern und Jugendlichen nicht auf der Strecke bleiben. Hier müssen wir früher ansetzen und KI-Anbieter\*innen in die Verantwortung nehmen, Es darf nicht sein, dass solche Inhalte überhaupt entstehen können.“

Dr. Marc Jan Eumann, Vorsitzender der  
Kommission für Jugendmedienschutz

Vorschläge EMR:

- KI-Systeme explizit in den Geltungsbereich des Jugendmedienschutz Staatsvertrag (JMStV) aufnehmen (**Gesetzesänderung**)
- Unabhängige/r KI-Jugendschutzbeauftragte/r sollte existieren, die bei Entwicklung, Training sowie Anwendung von KI einbezogen werden
- JMStV + JuSchG müssen regelmäßig und in kürzeren Abständen evaluiert werden



Stärken entwickeln. Gefährdungen begegnen.

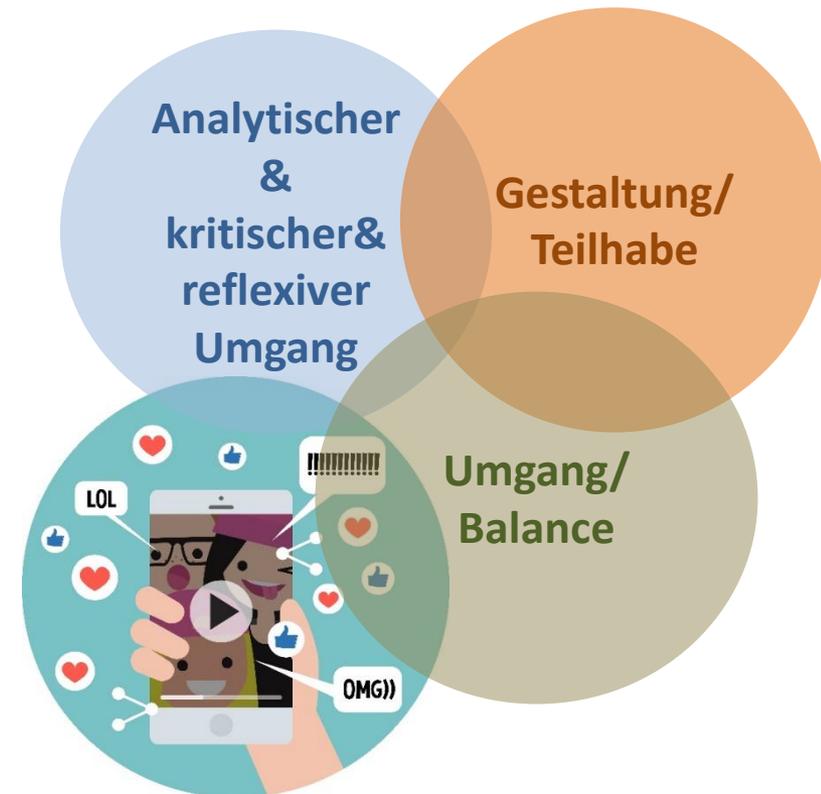
## Was kann jetzt bereits getan werden?

- Wirksamer Jugendmedienschutz kann nur gewährleistet werden, wenn man die Gefährdungslage durch KI für die Entwicklung Minderjähriger erkennt und richtig einordnet
- Genaue Kenntnissen über die Auswirkungen des Einsatzes digitaler Medien vermitteln
- Die genauen Folgen von KI Einsatz sind zum jetzigen Zeitpunkt leider noch unklar

Stärken entwickeln. Gefährdungen begegnen.

## Das altbekannte Ziel: Medienkompetenzförderung

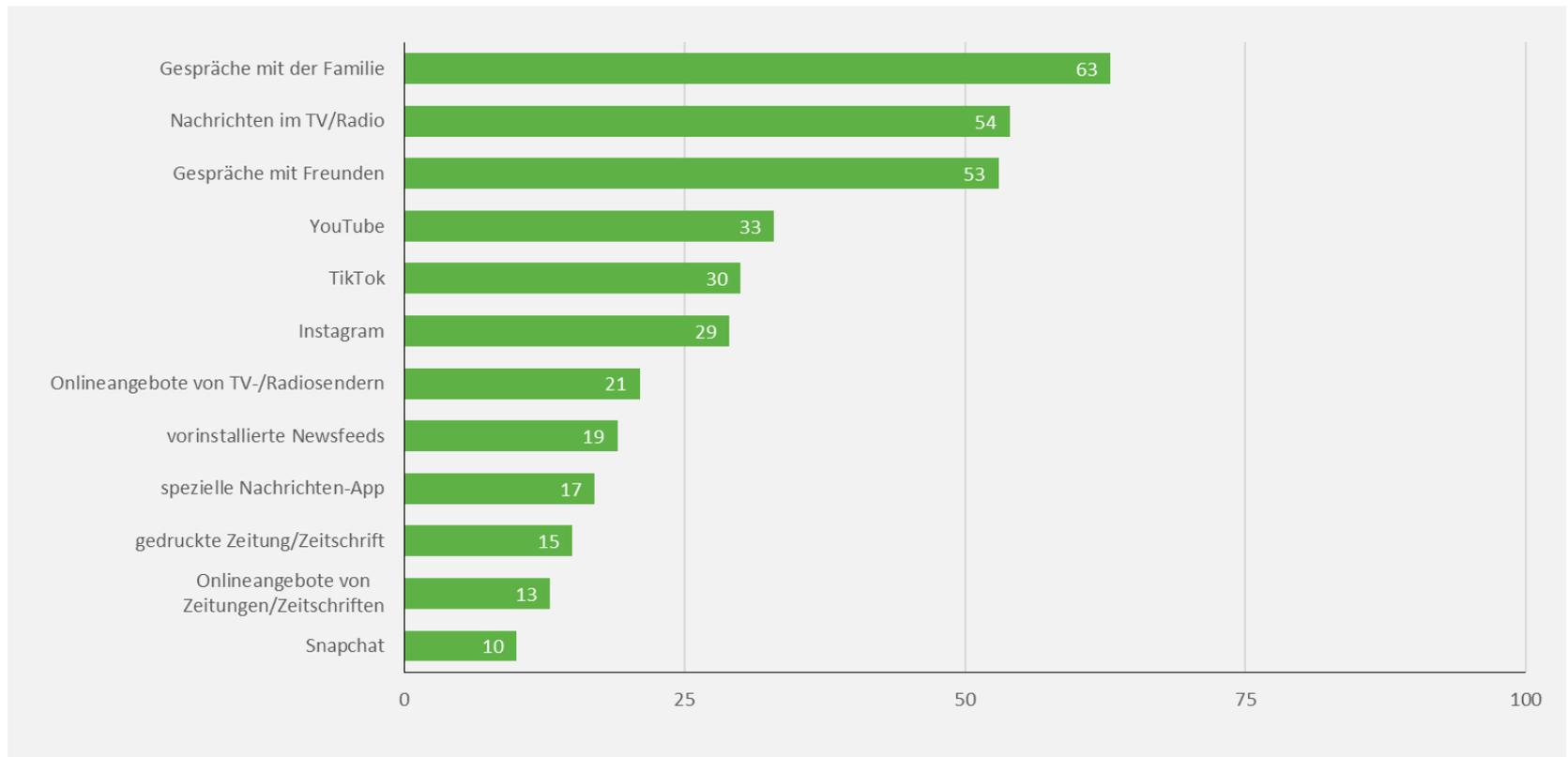
- für sich selbst Sinnvolles und Interessantes aus dem großen Medienangebot auszuwählen, statt wahllos zu konsumieren
- Inhalte einzuordnen und zu verarbeiten
- Medienangebote kritisch zu beurteilen
- **Kritischer Umgang mit Informationen**
- **Kennen von Risiken** und Abschätzen von Konsequenzen
- Kinder und Jugendliche haben ein Recht auf digitale Teilhabe





## Stärken entwickeln. Gefährdungen begegnen.

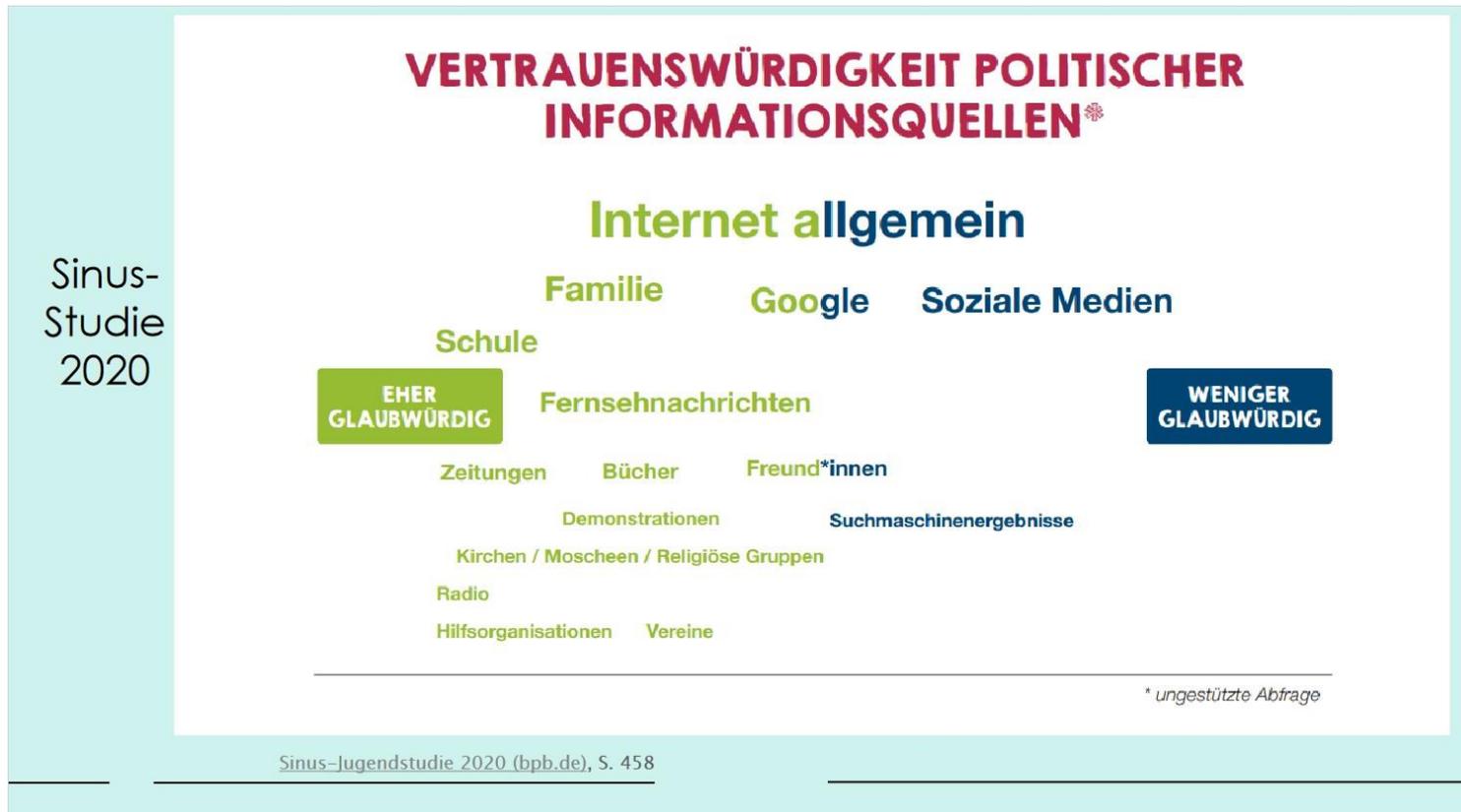
**Ich erfahre über das aktuelle Weltgeschehen durch...**  
- täglich/mehrmals pro Woche -





Stärken entwickeln. Gefährdungen begegnen.

# Informationsparadox

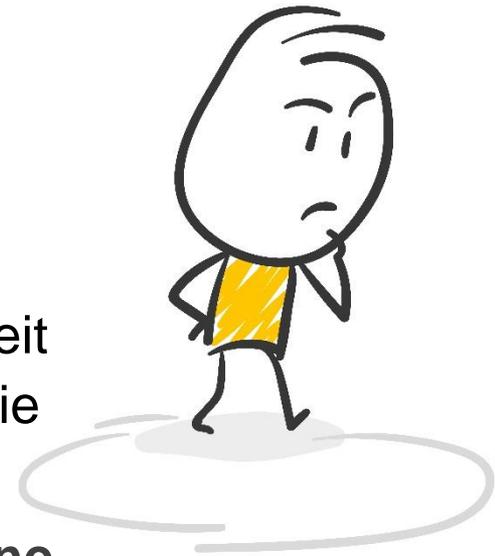


Stärken entwickeln. Gefährdungen begegnen.

## Informationsparadox

- Jugendliche messen Sozialen Medien keine große Glaubwürdigkeit bei und sind sich möglicher Falschmeldungen bewusst
- Trotzdem: nutzen sie Soziale Medien mit als wichtigste Informationsquelle für tagesaktuelle Themen
- Jugendliche sind sich dieser Widersprüchlichkeit durchaus bewusst, wissen aber oft nicht, wie sie damit umgehen sollen

→ Hier ist die Unterstützung durch Erwachsene gefragt





Stärken entwickeln. Gefährdungen begegnen.

## Wie Medienkompetent sind wir?

Kinder & Jugendliche

Kritisch-reflexive  
Medienkompetenz

Technische  
Medienkompetenz

Eltern / Pädagog\*innen

Kritisch-reflexive  
Medienkompetenz

Technische  
Medienkompetenz



Stärken entwickeln. Gefährdungen begegnen.

## Erkennt ihr welches Foto echt ist?

<https://www.whichfaceisreal.com/index.php>



Auf dieser Website kann man versuchen KI Bilder von echten Fotos zu unterscheiden

Screenshot  
<https://www.whichfaceisreal.com/index.php>

Stärken entwickeln. Gefährdungen begegnen.

## Ziel ist es..

- Frühzeitig Medien-und Informationskompetenz zu vermitteln
- Frühzeitig über mögliche Probleme und Risiken aufzuklären
- **Selbst informiert zu bleiben** und Interesse an den Themen der Kinder und Jugendlichen zu behalten
- Ins Gespräch gehen und ein offener, wertschätzender Gesprächspartner/Gesprächspartnerin sein

