

1. Thema: Echt jetzt?

Das ist Leonie. Sie ist für den Klimaschutz aktiv, ernährt sich vegan und findet, Frauen sollen stark und unabhängig sein.

Doch neulich haben wir sie im Urlaub von einer ganz anderen Seite erwischt.

Verändere den Hintergrund in eine Karibik-Strandlandschaft. Ersetze die Hose durch Shorts. Mache aus dem Herzzeichen zwei Hände, die einen Hamburger halten. Ergänze einen Jetski, einen Liegestuhl mit Sonnenschirm und ein Flugzeug im Hintergrund. Ändere den T-Shirt-Aufdruck und setze ihr eine Sonnenbrille auf.

Dieses Bild geht jetzt im Netz viral. Eine gefälschte Datei – auch Deepfake genannt –, die von einem Programm der Künstlichen Intelligenz – kurz KI – erstellt wurde.

Durch KI können Bilder kinderleicht und täuschend echt verändert werden. Du kannst also sprichwörtlich deinen eigenen Augen nicht mehr trauen.

2. Thema: Was ist KI?

Wie kommt es nun dazu, dass Programme ein Foto nach deinen Vorstellungen verändern und manipulieren können?

Grundlage dafür ist die sogenannte Künstliche Intelligenz, kurz KI. Bei der KI wird all das, was ein Mensch imstande ist, zu erlernen und sich auszudenken, auf einen Computer übertragen. Dieser führt dann entsprechend logische Handlungen oft viel schneller und ausdauernder aus.

Du kennst KI schon aus deinem Alltag. Beispielsweise bei Sprachassistenten oder Navigationsgeräten handelt es sich um KI-Systeme.

Technologien, die neue Inhalte künstlich erzeugen, nennen wir generative KI. Sie generieren, also erzeugen, beispielsweise Texte, Bilder, Video- oder Audioinhalte.

Sie können Referate, Sach- oder Wissenschaftstexte verfassen, Bild-, Video- oder Audiodateien erstellen und Musik komponieren oder Kunststile imitieren und Gemälde entwerfen.

Der Chatbot ChatGPT zählt zu den bislang meistverwendeten Text generierenden KIs. Bots – der Begriff ist von Roboter abgeleitet – sind Computerprogramme, die nach ihrer Aktivierung ohne menschliches Zutun automatisiert im Internet agieren. Ein Chatbot beispielsweise imitiert menschenähnliche Kommunikation und beantwortet Fragen in verständlicher Sprache.

Die KI ist also eine Musterschülerin, die alles abschreibt. Kann die KI wissen, dass ihr Ergebnis richtig ist? Wenn du alle Abschlussprüfungen in deinem Bundesland kennst, aber vom Thema sonst keine Ahnung hast, schreibst du dann die beste Prüfung?

3. Thema: Algorithmen: Fluch und Segen

Programme zur KI machen sich sogenannte Algorithmen zunutze, um komplexe Aufgaben zu lösen.

Grundlage für Algorithmen sind eine Vielzahl an Daten im Netz. Aus diesen Daten versucht die KI, wiederkehrende Muster zu erkennen und Zusammenhänge abzuleiten. Auf dieser Basis kommt die KI zu ihren Entscheidungen. Die KI begründet ihre Ergebnisse nicht.

Die KI versucht, intelligentes Verhalten nachzuahmen. Dies kann sie nur in dem Umfang, in dem ihr Daten zur Verfügung stehen. Die Qualität dieser Daten kann sie nur über einen Vergleich von Daten filtern.

Die generative KI trainiert sich selbst. Sie greift dabei auf ihre Trainingsdaten zurück und entwickelt sich weiter. Wir nennen das maschinelle Lernen. In der Folge lernt die KI so auch aus von KIs generierten Inhalten. Der menschliche Einfluss verringert sich damit.

Trainingsprogramme von textbasierten generativen KI-Systemen sind:

Das KI-System sucht nach Häufigkeiten aufeinanderfolgender Wörter.

Auch Menschen trainieren das System, beispielsweise, indem sie für dieses musterhafte Texte erstellen oder von der KI erstellte Texte bewerten.

Die generative KI passt ihr Verhalten diesem Lernprozess an.

Welche Risiken gibt es bei Entscheidungen nach Algorithmen? Wäre es möglich, die KI zu täuschen und mit falschen Inhalten zu manipulieren?

4. Thema: Manipulation durch KI

Eine KI hat kein Gewissen oder ethische Kriterien. Sie arbeitet nach Aufgabenstellungen. Sie unterscheidet nicht zwischen Wertvollem oder einer bösen Absicht. Das System fragt auch nicht, was mit dem erstellten Material passiert.

So sind KI-Systeme beispielsweise eine nützliche Hilfe bei Manipulationen und für Menschen, die die Gesellschaft spalten und die Demokratie abschaffen wollen.

So gingen 2023 von einer KI-generierte Deepfake-Fotos von angeblichen Müllbergen nach einer Pizzabestellung beim Bundesparteitag der Grünen durchs Netz. Diese sollten der Partei schaden.

Bei den Vorwahlen der US-Wahl 2024 hat ein KI-generierter Telefonroboter Menschen mit einer täuschend echten Stimme von Präsident Joe Biden angerufen und sie aufgefordert, nicht zur Wahl zu gehen.

Die Verunsicherung um KI wird auch genutzt, um Verschwörungen zu erzählen. So hat die AfD nach den massenhaften Demonstrationen für Demokratie 2024 gezielt verbreitet, die Bilder von den Demonstrationen wären Deepfakes, und in Wahrheit hätten sehr viel weniger Menschen daran teilgenommen.

Überlege dir selbst ein Deepfake zur Diskriminierung von Menschen. Erkläre die Wirkung deiner erdachten Manipulation.

5. Thema: Wer und was steckt hinter Manipulationen?

Mit KI können Bilder und Videos kinderleicht manipuliert werden. Es können Menschen sogar Aussagen, die sie nie getätigt haben, in den Mund gelegt werden.

So können Fake-News verbreitet und Verschwörungen erzählt werden, ohne dass wir deren wahre Urheberschaft und die genauen Absichten dahinter kennen.

Schon jetzt wird mit gezielter Desinformation unsere Gesellschaft gespalten mit dem Ziel, die Demokratie abzuschaffen.

Das funktioniert nur, wenn wir Falschmeldungen und Verschwörungen glauben und Informationen nicht überprüfen und so auf Deepfakes hereinfliegen.

Gezielte Desinformationen und Manipulationen kommen dabei von politisch Handelnden im In- und Ausland.

Radikale Kräfte im Inneren versuchen, die öffentliche Meinung zu beeinflussen, um Menschen oder Gruppen zu diskriminieren und zu isolieren.

Letztlich geht es um einen gesellschaftlichen Einfluss und den Zuwachs an Macht durch Wahlergebnisse oder die Abschaffung des politischen Systems.

Auch aus dem Ausland gibt es politische Kräfte oder Geheimdienste, die beispielsweise versuchen, Wahlen zu beeinflussen oder mit Verschwörungserzählungen den sozialen Frieden zu stören, um so das politische System zu schwächen.

Recherchiere Fälle von Wahlmanipulationen oder von Verschwörungserzählungen durch Deepfake.

6. Thema: Chancen und Risiken für die Demokratie

Eines ist sicher: Künftig wird die KI all unsere Lebensbereiche prägen.

Der Algorithmus könnte einseitig beeinflusst werden. Oft richten sich Internetseiten nicht nach pluralistischen, allgemein zugänglichen und integrativen öffentlichen Debatten aus.

Es werden oftmals „neue“ Informationen zu Inhalten gezeigt, die häufig angeklickt werden und mit dem bisherigen Online-Verhalten der Nutzenden übereinstimmen.

Das Internet als Datenbasis für KI bildet im Moment also nicht den breiten gesellschaftlichen Diskurs ab, sondern gibt Diskussionen und Informationen – häufig in vielen Filterblasen – wieder.

Generative KI kann die Demokratie aber auch stärken. Entscheidend für eine demokratiefreundliche KI sind die Transparenz der Algorithmen und überprüfbare Trainingsdaten.

Dann könnte KI Desinformation und Cyberangriffe auch verhindern und einen vertrauenswürdigen Zugang zu Informationen ermöglichen.

Chancengleichheit und Vielfalt könnten durch KI-Tools gefördert und vorurteilsfrei Daten erhoben werden. Dies ist beispielsweise bei Bewerbungen anwendbar und verhindert Diskriminierung.

Befrage selbst eine KI, beispielsweise ChatGPT, wo sie die Chancen und Risiken von KI für die Demokratie sieht. Ergänzt, erörtert und diskutiert die Ergebnisse in der Klasse.

7. Thema: Fakten- und Quellencheck

KI-generierte Bilder lassen sich aktuell meistens noch durch genaues Hinsehen als Fälschungen entlarven. Aber die Technologien entwickeln sich rasant, und Auffälligkeiten werden in Zukunft seltener.

Suche hochauflösende Versionen der verdächtigen Bilder und zoome in Details. Überprüfe, ob alle Körperteile vom Größenverhältnis zueinander passen. Details wie beispielsweise Hände, Ohren oder Brillen werden oft nicht vollständig oder verschwommen dargestellt.

Achte auch auf realistische Hintergründe und Übergänge von einzelnen Elementen zum Hintergrund. Mache eine Rückwärtssuche nach Ursprungsbildern oder Elementen daraus, beispielsweise bei Google Images oder TinEye.

Stammen die Inhalte von seriösen Quellen wie bekannten Medien, Instituten, Behörden oder Verbänden? Sieh dir das Impressum der Internetseite und die Web-Adresse an.

Indizien für die Echtheit eines Videos können sein, wenn es auf seriösen Nachrichtenseiten und nicht auf dubiosen Webseiten oder Plattformen zu finden ist. Gibt es bereits einen Faktencheck im Netz, oder finde ich über Suchmaschinen Hinweise auf Urheberschaft oder Entstehung?

Deckt sich das Video mit üblichen Aussagen und Verhalten der Person oder steht es im Widerspruch dazu?

Wer genauer hinsieht, geht Fake nicht so schnell auf den Leim.