67. Schülerwettbewerb des Landtags Baden-Württemberg

2024/2025

Thema 4 Format Erörterung

KI-Fakes im politischen Alltag – Kann man die Gesellschaft vor der Gefahr schützen?

Jonathan Elias Holm

Klassenstufe 11

Friedrich-Schiller-Gymnasium Ludwigsburg

Betreuende Lehrkraft Aileen Straub

Inhalt

1	Einführung	3
1.1	Relevanz und Aktualität	3
1.2	Chancen und Gefahren	3
1.3	Darstellung des Vorhabens.....	3
2.	Begriffsklärung	4
2.1	Deepfakes	4
2.2	(KI-) Fakes.....	4
2.3	Prompt.....	4
2.4	Artefakte	5
2.5	Deliberative Demokratietheorie Jürgen Habermas	5
2.6	Agent-Prinzipal-Theorie	5
3	Hauptteil.....	6
3.1	Darstellung der Vorschläge.....	6
3.2	Bildung und Aufklärung.....	6
3.3	Gütesiegel	7
3.4	Technologische Perspektiven	8
4	Resümee und Evaluation	9
4.1	Probleme zweiter Ordnung	10
5	Fazit.....	11

1 Einführung

1.1 Relevanz und Aktualität

Der Technologiesektor arbeitet schon seit Jahrzehnten an der Ausarbeitung von Realisierungsformen künstlicher Intelligenz. Erst seit Ende 2022 wurde mit der Veröffentlichung des Chatbots „ChatGPT“ eine solche Möglichkeit für die Allgemeinheit zur Verfügung gestellt. Mittlerweile gibt es eine Vielzahl solcher Systeme, die von verschiedenen Herstellern entwickelt worden sind. Inzwischen ist es jedem, der über einen Internetzugang verfügt, möglich, mit diesen zu interagieren. Die Steuerung des Programms erfolgt über Spracheingaben – sogenannte „Prompts“. Je nach Anforderung erstellt die KI darauf basierend Text-, Bild-, Sprach- und sogar Videoformate. Die vom Nutzer angefragten Ergebnisse können den Eindruck erwecken, als wenn sie der Realität entnommen wären. Tatsächlich handelt es sich lediglich um auf Basis von Algorithmen erstellte Datenformate.

1.2 Chancen und Gefahren

Die Programme ermöglichen es, auf eine schier unendliche Weite von Daten zuzugreifen und diese in Sekundenschnelle zielgerichtet zu verarbeiten. Durch die Nutzung können Prozesse der Wissensbeschaffung, Datenverarbeitung und -produktion, bis hin zur Bearbeitung universeller Aufgabenstellungen vereinfacht werden. Die Anwendung dieser Technologien birgt das Potenzial in sich, sozialpolitische Unterschiede zu überwinden, indem neue Zugänge und Partizipationsmöglichkeiten für eine breitere Menge von Menschen, relativ unabhängig vom sozialen Status, geschaffen werden. Auf der anderen Seite ergeben sich damit neue Herausforderungen mit Blick auf fehlende Transparenz in Bezug auf die Unterscheidung von Inhalten aus dem realen Leben und der virtuellen Realität. Hierdurch wird die Gesellschaft vor die Frage gestellt, welche Instrumente nötig sind, um diese Aufgabe zu bewältigen. Gerade im politischen Raum stellt sich diese Frage in einem besonderen Maße, weil die grundlegenden Formen des Zusammenhalts einer Gesellschaft diesem zugrunde liegen. Die Nutzung von KI kann für Desinformationen, Verleumdungen, Falschdarstellungen bis hin zur Verzerrung oder vollständiger Fälschung der politischen Informationsgebung missbraucht werden. Auf nationaler Ebene können politische Prozesse wie Wahlkampf und Berichterstattung beeinträchtigt werden, auf internationaler Ebene weitet sich die Problematik auf die Gefährdung diplomatischer Beziehungen und Vortäuschung globaler Gefährdungslagen aus.

1.3 Darstellung des Vorhabens

Die bisherigen Darstellungen verdeutlichen die Komplexität und besondere Bedeutung der Fragestellung für alle Bereiche des gesellschaftlichen Lebens. Im Folgenden werden grundlegende

Begriffe und Modelle geklärt, um sich auf KI-Fakes im politischen Alltag zu konzentrieren und wie oder ob man die Gesellschaft vor diesen schützen kann. Ziel des Textes ist es, eine Übersicht über verschiedene Umgangsmöglichkeiten mit dieser Herausforderung zu geben, insbesondere wie man demokratischen Gefährdungspotentialen begegnet.

2. Begriffsklärung

2.1 Deepfakes

Der Begriff setzt sich aus „Deep Learning“ und „Fake“ zusammen. „Deep Learning“ beschreibt eine Form des maschinellen Lernens, wobei mehrschichtige neuronale Netzwerke verwendet werden, um die komplexen Entscheidungsmöglichkeiten des menschlichen Gehirns nachzuahmen.^{1,2} „Fake“ wurde als Begriff für die Darstellung von Fälschungen basierend auf Sachverhalten etabliert. Bei einem „Deepfake“ handelt es sich also um datenbasierte Darstellungen, welche durch eine Verzerrung medialer Identitäten hervorgegangen sind und komplexen Algorithmen künstlicher Intelligenz zugrunde liegen. Dem Nutzer stehen hierbei verschiedene Möglichkeiten der Verfälschung zur Verfügung. Besonders gravierend sind personenbezogene Eingriffe, in denen beispielsweise Erscheinungsbild und Sprache von Persönlichkeiten verändert werden, um diesen Aussagen und Vergehen anzuhängen.³

2.2 (KI-) Fakes

Im Umgang mit virtuellen Dateninhalten stellt sich in dieser Thematik die Frage, mit welcher Absicht diese erstellt wurden. Oft kann zum Beispiel keine klare Grenze zwischen Fakes und Satire gezogen werden, was zu falschen Einschätzungen führen kann. Wenn in dieser Arbeit von „Fakes“ gesprochen wird, meint dies die Erstellung von virtuellen Dateninhalten, welche mit der Absicht bestimmte Personengruppen zu manipulieren erstellt wurden.

2.3 Prompt

Ein Prompt ist bei generativer KI eine Eingabe des Benutzers, zu dem das System eine Ausgabe erzeugt. Ein textbasierter Prompt kann Wörter, Buchstaben, Sonderzeichen, Zahlen und Links

¹ <https://www.zdf.de/dokumentation/terra-x/terra-x-deepfakes-100.html>

² <https://www.ibm.com/de-de/topics/deep-learning>

³ https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/Deepfakes/deepfakes_node.html

⁴ <https://wirtschaftslexikon.gabler.de/definition/prompt-125087>

enthalten.⁴ Um das gewünschte Ergebnis zu bekommen, muss der Prompt möglichst eindeutig und umfassend sein.

2.4 Artefakte

Systematische Darstellungsfehler aufgrund fehlender Dateninhalte. Die KI versucht durch Wahrscheinlichkeiten diesem Defizit beizukommen. Deshalb ist es nicht unüblich, dass von KI erstellte Inhalte Logikfehler aufweisen.⁵

2.5 Deliberative Demokratietheorie Jürgen Habermas

Der Begriff *Deliberative Demokratie* bezeichnet sowohl demokratietheoretische Konzepte, in denen die öffentliche Beratung zentral ist, als auch deren praktische Umsetzung. Wesentliches Kennzeichen einer deliberativen Demokratie ist der öffentliche Diskurs über alle politischen Themen, der auch als „Deliberation“ bezeichnet wird.⁶ In der Diskussion wird sich auf dieses Modell bezogen.

In diesem Zuge stellt sich die Frage, welche Anforderungen ein demokratisches Gesellschaftsmodell an die Bevölkerung stellt. Demokratische Gesellschaftssysteme enthalten ein Freiheitsversprechen beziehungsweise den Schutz vor staatlicher Willkür, gleichzeitig wird damit auch die Verantwortung für die politische Repräsentanz an die Bevölkerung übergeben. Diese Verantwortungsübergabe ermöglicht eine Vergrößerung gesellschaftlicher Entwicklungsprozesse durch die potenzielle Partizipation jedes Einzelnen. Angesichts neuer Herausforderungen entwickelt sich die Gesellschaft naturgemäß weiter und stellt damit permanente Anforderungen an die Lern- und Entwicklungsmöglichkeiten der Bevölkerung. Dieses Bewusstsein muss durch den Staat in Form von Bildung gewährleistet werden.

2.6 Agent-Prinzipal-Theorie

Asymmetrische Informationsteilung zwischen Sender und Empfänger in einer Transaktion. Im Anwendungsbeispiel wäre dies das Wissen eine KI benutzt zu haben, dies aber nicht zu kommunizieren. Der Auftragsausführer steht hier vor sogenannten *moral hazards*, also vor der Entscheidung, Informationen gegenüber dem Auftraggeber für seinen eigenen Nutzen zu unterschlagen.⁷

⁵ <https://www.deutschland.de/de/topic/kultur/deepfakes-deutschland-fake-news>

⁶ <https://www.bpb.de/kurz-knapp/lexika/das-europalexikon/176777/deliberation/>

⁷ <https://wirtschaftslexikon.gabler.de/definition/prinzipal-agent-theorie-42910>

3 Hauptteil

3.1 Darstellung der Vorschläge

Kernelemente der Demokratie wie wir sie kennen, wie beispielsweise der öffentliche Diskurs und Persönlichkeitsrechte können durch den Einsatz von KI-Fakes beziehungsweise Deepfakes außer Kraft gesetzt oder zumindest stark eingeschränkt werden. Deshalb sollte es im Interesse aller liegen, die von der Technologie ausgehenden Gefährdungen auf ein Minimum zu reduzieren, ohne auf die Vorteile, welche durch die Nutzung von KI entstehen, zu verzichten. Ein Verbot kommt folglich nicht in Frage und ist auch so gut wie nicht umsetzbar, da dadurch ein großer Teil des Internets national beziehungsweise international zensiert werden müsste. In diesem Fokus werden Ansatzmöglichkeiten zum Vorgehen gegen KI-Fakes diskutiert. Die Lösungsansätze lassen sich in humanistische und technologische Perspektiven, sowie journalistische Standards unterteilen. Humanistische Perspektiven stellen den Menschen in der Gesellschaft in den Mittelpunkt und setzen auf die Potenziale von Bildung und Aufklärung. Setzt man den Schwerpunkt auf technologische Perspektiven, so geht es darum, Technologien zur Überprüfung von Dateninhalten auf deren Richtigkeit beziehungsweise deren Ursprungsquellen zu entwickeln. Journalistische Standards, wie zum Beispiel Gütesiegel, sollen der Gesellschaft dabei helfen, eine Aussage zur Seriosität der Berichterstattung zu treffen.

3.2 Bildung und Aufklärung

Um ein so umfassendes und komplexes Problem zu erfassen und zu lösen, kommt zunächst in den Sinn, die Gesellschaft dahingehend aufzuklären, dass ein Bewusstsein für das Gefährdungspotenzial geschaffen wird. Hierbei soll es darum gehen, auf Dauer angelegte Prozesse anzustoßen, indem alle Schichten der Gesellschaft in die Lage versetzt werden, auf Weiterbildungsangebote zuzugreifen. Um diesen Vorschlag nachhaltig umzusetzen, ist es nötig, die gesamte Gesellschaft zu betrachten und geeignete Angebote zur Verfügung zu stellen. In diesen Angeboten wird darüber aufgeklärt, wie man Deepfakes beispielsweise anhand von Artefakten erkennt. Auf diese Angebote kann sowohl digital als auch in Präsenzveranstaltungen zugegriffen werden. Um eine digitale Erreichbarkeit zu gewährleisten, wird eine interaktive Lernplattform entwickelt, die sich anhand von Nutzerdaten individuell anpasst. Den Nutzern wird die Möglichkeit gegeben, sich auf der Basis von Selbsteinschätzungen spezifisch weiterzubilden. Die Plattform enthält Angebote zur Aneignung von Wissen in Form von Text und Videoformaten, individualisierte Beantwortung von Fragestellungen und sozialen Austausch. Hierdurch entstehen ein Gemeinschaftsgefühl und ein enges Zusammenarbeiten der Gesellschaft in Bezug auf die neuartige Ausgangslage. Über 50% der deutschen Bevölkerung sehen die eigene politische Partizipation als wichtiges Mittel und ermutigen dazu sogar ihre Mitmenschen – so eine Studie der SSOAR zur

politischen Beteiligung in Deutschland 2019.⁸ Ein generelles Interesse zur politischen Eigeninitiative besteht also.

Dies würde die Bedrohungslage auf nationaler Ebene stark eingrenzen und so zum Schutz der Gesellschaft vor KI-Fakes beitragen.

Allerdings ist dieser Vorschlag schwer umsetzbar. In unserer Gesellschaft sind viele verschiedene Gruppen von unterschiedlicher Bildung und Sprache vertreten, was zu Barrieren führt. Um also eine große Menge anzusprechen, müssten jeweils andere Strategien angewandt werden, um diese zu erreichen und nachhaltig aufzuklären. Außerdem muss eine allgemeine Bereitschaft zu einem lösungsorientierten Angehen der Möglichkeit bestehen. Da aber davon auszugehen ist, dass es bedeutende Bevölkerungsanteile gibt, die sich bisher noch nicht tiefgehend mit der Materie auseinandergesetzt haben, und darüber hinaus wenig Informationen über die Nutzung und Möglichkeiten künstlicher Intelligenz erworben haben, wird diese Gruppe tendenziell nicht daran interessiert sein, in diesen Vorschlag Zeit zu investieren. Eine aktuelle Studie der Bitkom e.V. ergab, dass bis jetzt nur rund 40% der Bevölkerung generative KI nutzen.⁹ Um eine erfolgreiche Umsetzung zu gewährleisten, muss dieser Anteil erhöht werden.

Betrachtet man die in der Gesellschaft vorhandenen Ressourcen, wird die Durchführung des Vorschlags in Frage gestellt. Zunächst mangelt es an Spezialisten, die eine solche Plattform zeitnah bereitstellen können, zum anderen verfügt längst nicht jeder Deutsche über die benötigte Hardware, um auf online Angebote zu zuzugreifen. Problematisch wird es bei der Auswahl und Ausbildung von Moderatoren und Fachexperten, aufgrund von mangelnden Kriterien der Überprüfung und Inhaltsstrukturen zur Weiterbildung.

3.3 Gütesiegel

Ein Vorschlag, der auf einen aktiven Einsatz der Bevölkerung abzielt, wird immer auf seine Grenzen stoßen.

Ein etwas pragmatischerer Ansatz ist eine gemeingültige Einführung von Gütesiegeln. Für die Umsetzung des Vorschlags müssen unabhängige Institutionen gegründet werden, um die Daten zu verifizieren. Dies sorgt für eine direkte Bürgerbeteiligung, um so das Gefühl einer Bevormundung zu vermeiden. Aufgabe der Institutionen ist, Dateninhalte von Nachrichtensendern, die ihre Seriosität beweisen wollen, auf KI erstellte Inhalte und deren Wahrheitsgehalt zu prüfen. Die Siegel

⁸ https://www.ssoar.info/ssoar/bitstream/handle/document/68018/ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf?sequence=4&lnkname=ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf

⁹ <https://www.bitkom.org/sites/main/files/2024-10/241016-bitkom-charts-ki.pdf>

werden mithilfe von Wasserzeichen umgesetzt und gehen so gegen Fälschungen vor. Um ein solches Siegel für sich zu gewinnen, müssen die Nachrichtensender ihre Maßstäbe der Recherche offenlegen. Dies schafft die nötige Transparenz, um KI-Fakes zu vermeiden. Die Studie der SSOAR zur politischen Beteiligung ergab ebenfalls, dass 75% Prozent der Bevölkerung bei der Informationsbeschaffung auf Quellen achten und nur auf seriöse zurückgreifen.¹⁰ Somit wird diesem Anteil der Menschen die Entscheidung erleichtert, um Desinformation zu vermeiden.

Problematisch wird hierbei die internationale Durchsetzbarkeit. Die größte Bedrohung geht nicht von nationalen, sondern von ausländischen politischen Feinden aus. Diese würden sich wohl kaum bereit erklären, Transparenz zu schaffen und ihre Maßstäbe offen zu legen, wenn sie die Absicht haben, die nationale Stabilität zu gefährden. Auf der internationalen Ebene der Berichterstattung ist die Gesellschaft möglichen KI-Fakes zur Manipulation folglich schutzlos ausgeliefert, da den Menschen keine eigenen Mittel der Verifizierung zur Verfügung stehen und sie auf die Siegel angewiesen sind.

Außerdem stellt sich die Frage, inwiefern die Institutionen legitimiert werden. Da der Ausruf zur Gründung der Institutionen vom Staat ausgehen würde, ist unklar, wer diese kontrollieren soll. Hinzu kommt, dass das Vertrauen in die Regierung und die Demokratie in den letzten Jahren gesunken ist: Nur 46% der Teilhabenden einer Studie der Körber-Stiftung gaben an, großes oder sehr großes Vertrauen zu haben.¹¹ Darüber hinaus kann man feststellen, dass bei herkömmlichen Gütesiegeln das Ausnutzen von Grauzonen zur scheinbaren Aufwertung des Produkts fast schon alltäglich geworden ist. So machen sich viele Firmen zum Beispiel das sogenannte „Greenwashing“ zunutze, wobei sie ihr Produkt als klimaneutral darstellen wollen. Ob sich ähnliche Sicherheitslücken bei Gütesiegeln in Bezug auf KI-Fakes bilden würden, bleibt abzuwarten.

3.4 Technologische Perspektiven

Aufgrund der *moral hazards*, denen der Mensch bei jedem Vorschlag, bei welchem er eine aktive Rolle spielt, ausgesetzt ist, werden solche Vorschläge immer Lücken aufweisen. Diese können sowohl von den Gesetzgebern oder Organisatoren als auch von der Gesellschaft oder Anwendern ausgenutzt werden. Deshalb ergibt sich die Möglichkeit, den Menschen von der Verantwortung zu entlasten und die Aufgabe der Verifizierung und Überprüfung Technologien zu überlassen. Ein konkreter Vorschlag ist, künstliche Intelligenz zu entwickeln, die auf genau diese Aufgabe spezialisiert ist. Da Inhalte von KI hauptsächlich auf der Basis von Wahrscheinlichkeiten, die einem

¹⁰ https://www.ssoar.info/ssoar/bitstream/handle/document/68018/ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf?sequence=4&lnkname=ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf

¹¹ <https://koerber-stiftung.de/projekte/demokratie-in-der-krise-umfrage-2024/#s41892>

Algorithmus zugrunde liegen, erstellt werden, können andere Algorithmen entwickelt werden, die diese erkennen und kennzeichnen. Erste Projekte gibt es schon: Das vom Bundesministerium für Bildung und Forschung geförderte Projekt „Fake ID“ beschäftigt sich mit der konkreten Ausarbeitung dieser Programme. Der Leiter des Projektes, Uwe Rabeler, äußerte sich in einem Interview optimistisch zur Effektivität des Vorhabens.¹² Bei einer erfolgreichen Realisierung können diese Technologien als einfache Applikation zu bestehenden sozialen Medien hinzugefügt werden, um Inhalte beim Hochladen automatisch zu überprüfen und zu markieren.

Dadurch kann eine hohe Trefferwahrscheinlichkeit gewährleistet werden, da viele mögliche Fehlerfaktoren aus der Gleichung entnommen werden. Die Gesellschaft wird entlastet und kann ohne großen Aufwand und Reformationen effektiv vor der Gefahr geschützt werden. Wirft man einen Blick in die Vergangenheit der technologischen Entwicklungen, so erkennt man, dass innovative technische Verfahren den Menschen meist entlastet haben.

Allerdings muss die prüfende KI für eine zuverlässige Gewährleistung weiterentwickelter sein, als die KI, welche zum Erstellen der Fakes benutzt wurde. Ansonsten können falsche Inhalte, welche zur Manipulation erstellt wurden, als scheinbar richtig gekennzeichnet werden und so die Prüfung umgehen. Da die führenden KI-Konzerne in ständiger Konkurrenz stehen und ihre neusten Entwicklungen nicht zwingend der Öffentlichkeit zur Verfügung stellen, wird immer eine gewisse Unsicherheit bestehen. Da die Firmen meist nach dem Prinzip der Nutzenmaximierung handeln, ist es nicht unwahrscheinlich, dass die Konzerne ihren neuesten Durchbrüchen nicht dem moralisch richtigen, sondern dem Meistbietenden liefern. Da es im Interesse feindlicher Staaten, die über mehr Liquidität als Deutschland verfügen, liegt, das Sicherheitssystem zu unterlaufen und die Gesellschaft zu manipulieren, besteht die Möglichkeit, dass die feindlichen Staaten dies als Möglichkeit sehen. So könnte beispielsweise China KI-Konzerne aufkaufen und ohne das Wissen der deutschen Gesellschaft diesen Fake-News, welche von der scheinbar unfehlbaren KI als ungefährlich eingestuft wurden, als reale Informationslagen vorgaukeln. Im Extremfall zieht dies ein komplettes Versagen der Presse und Informationsgebung mit sich.

4 Resümee und Evaluation

Es wurde festgestellt, dass es sowohl Defizite und Schwierigkeiten auf der Angebots- als auch auf der Nutzerseite gibt.

Humanistische Perspektiven setzen ein aktives Mitwirken der Bevölkerung und ein lösungsorientiertes Bewusstsein voraus. Außerdem liegt es in der Aufgabe des Staates für die

¹² <https://www.bundesregierung.de/breg-de/aktuelles/ki-gegen-deep-fakes-2274090>

nötige Aufklärung und Aufmerksamkeit zu sorgen, damit eine große Menge von Menschen erreicht wird. Hierbei mangelt es an den nötigen Mitteln, diesen Vorschlag schnell umzusetzen.

Allerdings zieht eine erfolgreiche Durchführung eine nachhaltige Aufklärung eines großen Teils der Bevölkerung mit sich, welcher durch eine aktive Bürgerbeteiligung den Zusammenhalt und das Selbstbewusstsein angesichts neuer Herausforderungen stärkt. Die Umsetzung zielt darauf ab, das Volk vor internationalen Gefahren von KI-Fakes zu schützen, da die Gesellschaft dann eigenständig in der Lage ist, diese zu erkennen und nicht auf äußere Faktoren angewiesen ist.

Journalistische Standards setzen Gütesiegel auf nationaler Ebene durch und kennzeichnen vertrauenswürdige Quellen, um den Konsumenten die Ungewissheit abzunehmen. Es liegt im Interesse aller Akteure, die Dateninhalte auf sozialen Medien teilen, diese Siegel für sich zu gewinnen, um ihren Kunden ihre Seriosität zu beweisen.

Jedoch sind hier die Transparenz und Legitimität fragwürdig. Da die Entscheidung, wer ein Siegel erhält, einigen wenigen überlassen ist, können diese Prozesse nicht für die gesamte Öffentlichkeit transparent gemacht werden. Außerdem ist die Gesellschaft auf internationaler Ebene dem Problem weiterhin schutzlos ausgeliefert.

Technologische Perspektiven entnehmen die menschliche Fehlerquote und sind möglicherweise am verlässlichsten, aber weisen ebenfalls eine mangelnde Transparenz auf.

Überschneidende Herausforderungen der Vorschläge beinhalten Fragen der Transparenz, Effizienz und demokratischer Legitimation.

Perspektiven der Bildung und Aufklärung setzen die Kriterien der Transparenz und demokratischer Legitimation auf Kosten der Effizienz in den Vordergrund. Journalistische Gütersiegel erreichen auf nationaler Ebene aufgrund weniger Beteiligter eine hohe Effizienz, weisen aber eine geringe Transparenz und demokratische Legitimität auf. Wohl am effizientesten sind technologische Lösungsansätze, allerdings mangelt es an der Offenlegung der Maßstäbe und öffentlichem Diskurs.

4.1 Probleme zweiter Ordnung

Die Beteiligung des Volkes scheint für die Sicherung der Demokratie zunächst am angemessensten, hat jedoch wie die anderen Vorschläge das Potenzial, Probleme zweiter Ordnung zu schaffen, beziehungsweise Gefahr zu laufen, sich ins Gegenteil zu verkehren. Eine Überbetonung von Seiten der Regierung hinsichtlich einer Problemlage, von welcher die Mehrheit der Bevölkerung noch nicht betroffen ist oder diese nicht erkennt, kann zu Widerstand und Reaktanz führen. Der genannte Sachverhalt ist deshalb so bedeutsam, da die Demokratie als Gesellschaftsform darauf angewiesen ist, dass jeder Einzelne die Aufrechterhaltung der Demokratie als individuelle Lernaufgabe und Entwicklungsziel wahrnimmt.

Auch technologische Perspektiven bergen in sich das Potential, bei einer erfolgreichen Umsetzung Probleme zweiter Ordnung zu generieren. KI-Systeme weisen in diesem Anwendungsbereich an erster Stelle zwei potenzielle Fehlfunktionen auf. Zum einen können Lösungsvorschläge erarbeitet werden, die auf Grund der Dateneingabe exzessiv vorangetrieben werden, ohne dass geeignete Korrekturmaßnahmen vorgenommen werden. Dies ist auf die marginale Transparenz der KI zurückzuführen und die damit verbundene Selbstreferenzialität. Zum anderen kann bei der Nutzung von KI-Systemen nicht unbedingt ausgeschlossen werden, dass eine dynamische Komplexität geschaffen wird, die den eigentlichen Zielen zuwiderläuft. Dies kann zum Beispiel in einer Verselbstständigung erfolgen. Prominent wurde dies am Beispiel der Firma Amazon. Diese nutzte KI-Systeme zur Erleichterung der Personalauswahl, was in einer Diskriminierung von Frauen resultierte.^{13;14} Dies ist darauf zurückzuführen, dass in der Branche die Mehrzahl der Führungspositionen an Männer vergeben ist und die KI dies als vorgesehen erachtete.

Die Einhaltung und Überprüfung journalistischer Standards erscheinen zunächst wünschenswert. Allerdings kann eine solche Standardisierung dazu führen, dass nicht mehr alle Perspektiven von vermeintlich vertrauenswürdigen Medienvertretern repräsentiert werden. Dadurch kann ein Nachrichtenmonopol und infolgedessen eine Vereinseitigung der Informationsgebung entstehen. Aufgrund dessen kann die Meinungsbildung und die Deliberation eingeschränkt werden. Dies kann wiederum in sich eine Bedrohung der Demokratie darstellen.

5 Fazit

Zusammenfassend lässt sich konstatieren, dass jeder angeführte Vorschlag individuell auf Grenzen stößt, weshalb sich die Notwendigkeit ergibt, die Maßnahmen komplementär zu konzipieren. Dadurch wird ein Sicherheitsapparat geschaffen, der auf verschiedenen Ebenen agiert. Auf der einen Seite wird die Bevölkerung in die Lage versetzt, sich eigenständig durch Bildung zu schützen, auf der anderen Seite sorgen unabhängige Institutionen für die nötige Transparenz, um zwischen KI-Fakes und realen Faktenlagen zu unterscheiden.

Zuletzt muss der erweiterte Sicherheitsbegriff betrachtet werden, um das Ausmaß und die Betroffenheit der Bedrohung einzuschätzen. Betrachtet man zunächst die Gefahrendimension, so fällt realistisch gesehen auf, dass nur Wenige aktiv von dem Problem betroffen sind und es sich daher viel mehr um ein diffuses Risiko als eine direkte Gefährdungslage handelt. Dennoch ist es sinnvoll, die Bevölkerung schon früh auf potenzielle Gefahren vorzubereiten, um im Ernstfall ein

¹³ https://www.focus.de/finanzen/boerse/die-tuecken-der-intelligenz-amazon-schaltet-ki-ab-die-bewerbungen-von-frauen-als-minderwertig-erachtete_id_9741890.html

¹⁴ <https://www.boeckler.de/de/boeckler-impuls-vom-algorithmus-diskriminiert-47283.htm>

Chaos zu vermeiden. In Hinsicht auf die Raumdimension wird deutlich, dass sich Herausforderungen bezüglich künstlicher Intelligenz nicht nur auf virtuelle Welten beziehen, sondern auch erheblichen Einfluss auf die reale Welt nehmen können. Man spricht bei der Verknüpfung von virtuellen Technologien mit der physischen Welt vom sogenannten *Internet der Dinge*.¹⁵

Im Zuge der Globalisierung haben sich Internettechnologien auf der ganzen Welt verbreitet, weshalb zum einen ein großer Teil der Weltbevölkerung potenziell gefährdet ist, die Gefährdung aber zum anderen auch von so gut wie jedem ausgehen kann. Der geographische Raum ist daher nicht klar eingrenzbar, die Vorschläge sind aber aufgrund verschiedener Interessen und Staatsformen nur national umsetzbar. Dennoch ist es wünschenswert allgemeingültige Entschlüsse bezüglich geeigneter Präventionsstrategien innerhalb von internationalen Organisationen zu treffen. Betroffene Ebenen der Sachdimension beinhalten vor allem militärische, ökonomische und humanitäre Perspektiven. KI-Fakes können als Provokationen oder Manipulationsstrategien, als Auslöser von kriegerischen Konflikten, als negative Marktbeeinflussung durch Desinformation und zur Spaltung der Gesellschaft verwendet werden. Da viele Sektoren betroffen sind und keine genaue Differenzierung getroffen werden kann, muss nicht ein Bereich, sondern die gesamte Gesellschaft Objekt der Sicherheitspolitik sein. Somit betrifft die Referenzdimension die Bevölkerung als Ganzes.

Um eine effektive Durchsetzung der Vorschläge zu gewährleisten, ist es nötig, dass die angeführten Maßnahmen nicht nur von der Politik, sondern auch von Seiten der Wirtschaft ausgehen. Dies kann allerdings zum Problem werden, da die Weiterentwicklung von KI ohne Einschränkungen im Interesse dieser liegt. Dennoch stehen der Politik Mittel, wie zum Beispiel Subventionen oder andere Anreize zur Verfügung, um eine Kooperation bezüglich der Transparenz und Sicherheit attraktiv zu machen. Auch wenn zuerst die Größe der Herausforderung immens erscheint, sollte man sich vor Augen führen, dass die genannten Vorschläge aussichtsreich sind, um die Risikofaktoren von künstlicher Intelligenz zu beherrschen und eine nachhaltige Gesamtlösung anzustreben. Nachhaltig meint in diesem Sinne, die Einschränkung der Gefährdungspotentiale zu nutzen, um gleichzeitig die Chancen zur Verbesserung gesellschaftlicher Problemlagen und Konflikte auszuschöpfen.

¹⁵ <https://www.bundesnetzagentur.de/DE/Fachthemen/Digitalisierung/Internet/IoT/start.html>

Quellenverzeichnis

- 1) <https://www.zdf.de/dokumentation/terra-x/terra-x-deepfakes-100.html>
(letzter Aufruf:23.11.2024)
- 2) <https://www.ibm.com/de-de/topics/deep-learning>
(letzter Aufruf:23.11.2024)
- 3) https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/Deepfakes/deepfakes_node.html
(letzter Aufruf:24.11.2024)
- 4) <https://wirtschaftslexikon.gabler.de/definition/prompt-125087>
(letzter Aufruf:26.11.2024)
- 5) <https://www.deutschland.de/de/topic/kultur/deepfakes-deutschland-fake-news>
(letzter Aufruf:25.11.2024)
- 6) <https://www.bpb.de/kurz-knapp/lexika/das-europalexikon/176777/deliberation/>
(letzter Aufruf:25.11.2024)
- 7) <https://wirtschaftslexikon.gabler.de/definition/prinzipal-agent-theorie-42910>
(letzter Aufruf:25.11.2024)
- 8) https://www.ssoar.info/ssoar/bitstream/handle/document/68018/ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf?sequence=4&lnkname=ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf
(letzter Aufruf:29.11.2024)
- 9) <https://www.bitkom.org/sites/main/files/2024-10/241016-bitkom-charts-ki.pdf>
(letzter Aufruf:29.11.2024)
- 10) https://www.ssoar.info/ssoar/bitstream/handle/document/68018/ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf?sequence=4&lnkname=ssoar-2020-Schaetz_et_al-Politische_Partizipation_in_Deutschland_2019.pdf
(letzter Aufruf:29.11.2024)

11) <https://koerber-stiftung.de/projekte/demokratie-in-der-krise-umfrage-2024/#s41892>

(letzter Aufruf:29.11.2024)

12) <https://www.bundesregierung.de/breg-de/aktuelles/ki-gegen-deep-fakes-2274090>

(letzter Aufruf:29.11.2024)

13) https://www.focus.de/finanzen/boerse/die-tuecken-der-intelligenz-amazon-schaltet-ki-ab-die-bewerbungen-von-frauen-als-minderwertig-erachtete_id_9741890.html

(letzter Aufruf:24.11.2024)

14) <https://www.boeckler.de/de/boeckler-impuls-vom-algorithmus-diskriminiert-47283.htm>

(letzter Aufruf:27.11.2024)

15) <https://www.bundesnetzagentur.de/DE/Fachthemen/Digitalisierung/Internet/IoT/start.html>

(letzter Aufruf:29.11.2024)