Lernende Systeme
DIE PLATTFORM FÜR KÜNSTLICHE INTELLIGENZ

KI im Superwahljahr 2024

Whitepaper

Bieber, C., Heesen, J., Grunwald, A.
& Rostalski, F.
AG IT-Sicherheit, Privacy, Recht und Ethik

Kurzfassung

Die zunehmende Verbreitung generativer Künstlicher Intelligenz (KI) wie ChatGPT stellt Demokratien vor neue Herausforderungen. Demokratien brauchen den öffentlichen Diskurs, der wiederum Vertrauen und Verlässlichkeit voraussetzt. Falsche oder ungenaue, aber gleichzeitig überzeugende, Informationen in sozialen Medien untergraben die Wissensbasis für diese Diskurse. Fake News und Desinformation sind zwar auch im Netz nichts Neues, nehmen mit dem Aufkommen generativer KI aber zu. Damit sind mögliche Auswirkungen von generativer KI auf Demokratien zu einem wichtigen Thema geworden – gerade im Superwahljahr 2024. Schließlich ist mehr als die Hälfte der Weltbevölkerung in über 40 Ländern zur Stimmabgabe aufgerufen.

Welchen (direkten) Einfluss generative KI dabei tatsächlich auf Wahlen und die politische Meinungsbildung hat, lässt sich zum jetzigen Zeitpunkt aufgrund der unzureichenden Datenlage noch nicht belegen. Jedoch sind unabhängig davon demokratische Gesellschaften gefordert, sich mit dieser in unserem Alltag bereits fest etablierten Technologie mit all ihren Vor- wie Nachteilen auseinanderzusetzen und geeignete Maßnahmen zur Eindämmung von Manipulationen und Einflussnahme – auch mithilfe generativer KI – zu ergreifen, wie Herkunftspflicht oder Stärkung der gesellschaftlichen KI-Kompetenzen.

Auswirkungen generativer KI auf Meinungsbildung und Wahlen

Zahlreiche reale Beispiele lassen bereits erkennen, welche (potenziellen) Auswirkungen generative KI auf Wahlen und die politische Meinungsbildung haben kann und welches Potenzial sich daraus für zukünftige Wahlkämpfe ergeben könnte: So wurden bei den US-Vorwahlen mithilfe einer täuschend echten KI-Stimme von Präsident Joe Biden falsche Informationen rund um die Vorwahlen der Demokratischen Partei für die US-Präsidentschaftswahl verbreitet. In der Slowakei tauchte vor den Nationalratswahlen eine gefälschte Audiodatei auf, die den Parteichef Michal Šimečka und eine Journalistin in einen Stimmenkauf verwickeln sollte. In Deutschland war Bundeskanzler Olaf Scholz Ziel eines Deepfakes, und es kursierte ein gefälschtes Video der Tagesschau. Das zeigt: Der Einsatz von generativer KI im politischen Kontext ist bereits Realität – in Form von täuschend echt gefälschten Bildern, Videos oder Audioaufnahmen als sogenannte Deepfakes im Vorfeld von Wahlkämpfen.

Beispiele für KI-unterstützte Einflüsse auf Wahlen

KI-Tools können an verschiedenen Stellen im Wahlprozess eingesetzt werden und sowohl Bedrohungsszenarien auslösen als auch unterstützend oder absichernd wirken. Auf vielfältige Weise kann KI insbesondere im **Wahlkampf unterstützen**. Ein Beispiel ist der Einsatz von KI-Chatbots wie dem amerikanischen Chatbot „Ashley", der im US-Wahlkampf 2020 tausende von Menschen anrief, um für die Kandidatin Shamaine Daniels zu werben. Diese KI-Bots ermöglichen es Kandidatinnen und Kandidaten, massenhaft personalisierte Anrufe in Wahlkämpfen zu tätigen und damit eine Reichweite zu erzielen, die bisher nicht möglich war.

Die Herausforderung besteht jedoch darin, sicherzustellen, dass solche Systeme keinen ungewollten Bias entwickeln: Auch wenn die KI-Stimme wie im Fall von Ashely (erkennbar als nicht menschliche Stimme) zu Beginn des Gesprächs selbst darauf hinweist, dass sie eine KI ist. Der Einsatz von KI-Chatbots seitens Politikerinnen und Politiker, Parteien oder Institutionen der öffentlichen Meinungsbildung wie Verbände oder Kampagnennetzwerke stellt daher besondere Anforderungen an Transparenz oder Datenqualität. Vor allem in Deutschland mit Blick auf KI-Regelungen wie den AI Act der EU gelten Transparenzpflichten für KI-Systeme, die direkt mit Menschen interagieren: KI-Chatbots müssen sich demnach als solche zu erkennen geben. Im Wahljahr 2024 werden diese Regelungen allerdings noch nicht greifen. Solche „Robo-Calls" werden in Deutschland aufgrund von Datenschutzgesetzen jedoch keine Rolle spielen, da Werbeanrufe ohne Zustimmung verboten sind.

Der zunehmende Einsatz von Sprachmodellen und Chatbots liefert zugleich die Basis für **(falsche) Wahlinformationen**. Ein qualitatives Problem solcher Modelle wie ChatGPT sind die sogenannten „Halluzinationen", bei denen Sprachmodelle fehlerhafte oder erfundene Informationen ausgeben. Dies kann zu irreführenden Wahlinformationen führen, wie etwa falsche Zuordnungen von Kandidatinnen und Kandidaten oder Parteien. Außerdem besteht die Gefahr, dass Sprachmodelle menschliche Vorurteile verstärken. Dies ist vor allem dann problematisch, wenn Menschen den Modellen zu sehr vertrauen und die Informationen nicht weiter überprüfen. Dennoch können gut regulierte Chatbots auch dazu beitragen, Wählinnen und Wähler besser und individueller zu informieren, um die politische Beteiligung zu fördern.

Generative KI-Systeme, insbesondere große Sprachmodelle, bieten vor allem politischen Akteuren neue Möglichkeiten, um **zielgruppenorientierte Inhalte zu erstellen**. Parteien können damit Wahlprogramme an aktuelle Stimmungen der Wählerschaft anpassen und zielgruppenorientierte Inhalte generieren, etwa durch sprachliche Anpassung an verschiedene Wählergruppen oder die Übersetzung in Leichte Sprache. Populistische Parteien hingegen könnten diese Technologien jedoch nutzen, um gezielt polarisierende und auf unterschiedliche Zielgruppen zugeschnittene Inhalte zu verbreiten, die politische Stimmungen ausnutzen und verstärken.

Neben unabsichtlichen Fehlinformationen durch KI-Sprachmodelle ist die **gezielte böswillige Manipulation** von Wählerinnen und Wählern durch **sogenannte Deepfakes** – also täuschend echt wirkende Bild-, Audio- oder auch Videoaufnahmen – ein Problem. Manipulation an sich ist kein neues Phänomen, aber mithilfe von generativer KI erreichen die Möglichkeiten zur Einflussnahme damit eine neue Dimension. Im Kontext politischer Einflussnahme werden sie unter anderem genutzt, um Politiker falsch zu zitieren oder in ein negatives Licht zu rücken und so das Vertrauen in seriöse Informationen im Wahlkampf zu untergraben. In diese Richtung wirkt auch die Deepfake-Pornographie, etwa durch die Erpressung von in der politischen Öffentlichkeit stehenden Persönlichkeiten, überwiegend Frauen.

Deepfakes erleichtern die Manipulation der öffentlichen Meinung erheblich, insbesondere unterstützt durch die schnelle Verbreitung in sozialen Medien, was eine erhebliche Bedrohung für die Demokratie darstellen kann. KI-gestützte Desinformation könnte das Vertrauen in politische Institutionen untergraben und gesellschaftliche Spaltungen vertiefen. Allerdings ist noch unklar, ob KI-generierte Fake News einen tatsächlich messbaren Einfluss auf Wahlergebnisse haben oder sich von herkömmlichen digitalen Falschinformationen unterscheiden. Dennoch: Demokratische Systeme müssen ihre Prozesse schützen, auch wenn direkte Auswirkungen auf Wahlergebnisse schwer nachzuweisen sind.

ANWENDUNGSBEISPIEL

Deepfake-Bildmaterial zur Beeinflussung afroamerikanischer Wählerinnen und Wähler

Quelle: Spring, 2024

Maßnahmen zur Sicherung der Integrität demokratischer Wahlen

Unabhängig von der aktuellen Relevanz der KI-beeinflussten Wahlmanipulation müssen Demokratien wachsam bleiben und geeignete Maßnahmen hinsichtlich rechtlicher, gesellschaftlicher und technischer Art ergreifen.

Durch **rechtliche Maßnahmen** lassen sich Anforderungen wie Transparenz, Nachvollziehbarkeit, Qualität der Trainingsdaten oder Reichweitenbegrenzung beim Einsatz von KI im politischen Kontext umsetzen. So stuft der AI Act der EU solche KI-Systeme, die Wahlen beeinflussen können, als Hochrisiko-KI ein. Damit müssen diese Systeme Qualitätsstandards erfüllen und transparent gekennzeichnet werden. Auch der Digital Services Act (DSA) fordert von großen Plattformen Risikoprüfungen und Maßnahmen gegen Desinformation. Nach dem deutschen Netzwerkdurchsetzungsgesetz (NetzDG) ist es zudem möglich, Accounts oder Bots, die Falschinformationen verbreiten, zu sperren oder deren Reichweite zu begrenzen. Bei derartigen Überlegungen und Handlungen ist jedoch der Eingriff und damit der Schutz der Meinungs- und Redefreiheit durch staatliche – aber auch private Akteure wie soziale Netzwerke – stets zu beachten und zu prüfen.

Darüber hinaus ist es wichtig, die Kompetenzen der Bürgerinnen und Bürger als **gesellschaftliche Maßnahme** zu fördern, um die Widerstandsfähigkeit gegenüber KI-Manipulationen zu stärken. Eine KI-kompetente Öffentlichkeit ist weniger anfällig für Desinformation. Bildungsangebote, die den kritischen Umgang mit Medien und KI fördern, sollten daher gezielt ausgebaut werden. Auch Journalistinnen und Journalisten sowie Medienschaffende müssen im Umgang mit KI geschult werden, sowohl bei der Nutzung als auch in der Bewertung von Quellen. Gleichzeitig bleibt die Sicherung demokratischer Prozesse eine gesamtgesellschaftliche Aufgabe – die Verantwortung darf nicht allein den Bürgerinnen und Bürgern überlassen werden.

Darüber hinaus gibt es verschiedene **technische Maßnahmen**, um KI-generierte Desinformation zu verhindern oder zumindest zu vermindern. Dazu gehören beispielsweise kryptographische Herkunftsnachweise und Wasserzeichen. Herkunftsnachweise nutzen digitale Signaturen zur Sicherstellung der Authentizität von Inhalten, können aber nur verlässliche Inhalte bestätigen. Wasserzeichen kennzeichnen synthetische Inhalte und ermöglichen deren Erkennung und Entfernung. Zudem können KI-Methoden auch zur Erkennung von Deepfakes eingesetzt werden, indem sie Unregelmäßigkeiten in Gesichtern, Stimmen und Videos identifizieren. Allerdings ist der Einsatz solcher Technologien begrenzt, da Angreifer ihre Methoden anpassen können, was zu einem ständigen Wettlauf um technologische Überlegenheit führt. Große Tech-Unternehmen haben jedoch bereits erkannt, dass ihre Produkte weniger anfällig für Missbrauch gestaltet werden müssen.

Ausblick

Generative KI-Systeme sind mittlerweile ein fester Bestandteil unserer Lebenswirklichkeit und werden sowohl kurz- als auch langfristig unverzichtbar sein. Wählerinnen und Wähler sowie Akteure in Parteien, Politik und Journalismus müssen sich dieser Realität stellen. Dazu gehört auch, dass generative KI für bösartige Zwecke mit dem Ziel der Einflussnahme auf demokratische Prozesse und Meinungsbildung missbraucht werden kann.

Ob und inwieweit Wahlen durch generative KI manipuliert werden können, bleibt jedoch noch unklar; das Superwahljahr 2024 könnte erste wissenschaftliche Studien liefern. Alle Beteiligten sind unabhängig davon aufgerufen, sich der Chancen und Risiken von generativer KI bewusst zu werden und Missbrauch mit geeigneten Maßnahmen vorzubeugen. Dies ist entscheidend, um die Integrität demokratischer Prozesse und das Vertrauen in KI-Systeme zu schützen – ein Prozess, der kontinuierlich angepasst werden muss.

Impressum
Herausgeber: Lernende Systeme – Die Plattform für Künstliche Intelligenz | Geschäftsstelle | c/o acatech | Karolinenplatz 4 | D-80333 München | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Folgen Sie uns auf X: @Lernende Systeme | LinkedIn: de.linkedin.com/company/plattform-lernende-systeme | Mastodon: social.bund.de/@LernendeSysteme | Stand: Juli 2024 | Bildnachweis Rawpixel/Shutterstock/Titel

Diese Kurzfassung entstand auf Grundlage des Whitepapers *KI im Superwahljahr 2024. Generative KI im Umfeld demokratischer Prozesse*. Whitepaper aus der Plattform Lernende Systeme, München, 2024. Es wurde erstellt von Mitgliedern der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik. *https://doi.org/10.48669/pls_2024-5*

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

acatech
DEUTSCHE AKADEMIE DER
TECHNIKWISSENSCHAFTEN