Grüne Leitlinien für den Einsatz von Künstlicher Intelligenz (KI) im Wahlkampf

Viele KI-Anwendungen bieten eine Reihe von Chancen: Sie können den Zugang zu Informationen erleichtern und damit die politische Willensbildung unterstützen, Datenanalysen verbessern, dabei helfen Wähler*innenpotenziale besser zu verstehen oder die Effektivität von Kampagnen stärken. Viele KI-Anwendungen können aber auch erhebliche Risiken bergen: Mit einem Klick können ganz leicht überzeugende Falschinformationen und -inhalte generiert werden. Durch Bot-Angriffe mit gefälschten Social-Media-Accounts wird das Netz massenhaft gezielt mit Inhalten überflutet, um den Diskurs zu verzerren. Mit Blick auf die Wahlen können diese Manipulationstechniken das gegenseitige Vertrauen innerhalb unserer Gesellschaft weiter beschädigen, Debatten beeinflussen und eine Bedrohung für den demokratischen Prozess darstellen.

In Übereinstimmung mit unseren grünen Werten sind wir überzeugt, dass KI eigentlich dazu beitragen sollte, unsere Demokratien, globale Menschenrechte und individuelle Freiheiten zu stärken, ohne dass sie einen Schaden verursachen.

Im Vorfeld der EU-Wahlen 2024 wollen wir uns für eine umweltfreundliche und ethisch vertretbare Nutzung von Künstlicher Intelligenz einsetzen. KI hat nicht nur die politische Landschaft geprägt, sondern auch eine neue Form von Wahlkampf etabliert. Dieser Wandel wird durch den kontinuierlichen technologischen Fortschritt beschleunigt und stellt unsere demokratische Gesellschaft - und uns als grüne Partei - vor große Herausforderungen.

Mit diesen Leitlinien wollen wir sicherstellen, dass die sichere und transparente KI-Nutzung innerhalb unserer Wahlkämpfe gewährleistet ist und Risiken minimiert werden.

1. Kontinuierliche, menschliche Überwachung bei KI-generierten Inhalten

Die Ergebnisse von KI-Tools hängen stark von der Qualität, der Quantität und Gewichtung der einzelnen Datensätze ab, mit denen sie trainiert werden. KI kann möglicherweise ungenaue, irreführende oder nicht aktuelle Aussagen generieren. Die Zuverlässigkeit und Objektivität der Ergebnisse zum aktuellen Zeitpunkt sollten daher stets hinterfragt werden.

Daten, mit denen die KI trainiert wurde, können urheberrechtlich geschützt sein – zum Beispiel Textbausteine, Begriffe oder Bilder. Dadurch können die KI-generierten Inhalte unter Umständen eine Urheberrechtsverletzung darstellen.

Wir verpflichten uns zur Überwachung und Prüfung von KI-generierten Inhalten, einschließlich einer sorgfältigen Faktenprüfung. Wir übernehmen und veröffentlichen nicht ungeprüft Entscheidungen und Ergebnisse von KI-Systemen. Wir stellen sicher, dass jegliches von uns möglicherweise produzierte KI-generierte Material vor der Veröffentlichung einer menschlichen Aufsicht unterzogen wird - das gilt insbesondere für Situationen wie Wahlkämpfe, in denen ein fehlerhafter Inhalt schwerwiegende Folgen haben könnte.

2. Schutz von persönlichen und sensiblen Daten

Da sich viele Anbieter von KI-Tools das Recht vorenthalten, die Eingaben bzw. Daten zur Optimierung ihrer Dienste (Training von Algorithmen) weiter zu nutzen, dürfen personenbezogene und sensible Daten an keiner Stelle verwendet werden.

Wir setzen auf das Prinzip der Datensparsamkeit. Wir nutzen und verwenden so viele Daten wie nötig und so wenig Daten wie möglich. Daten werden nur in dem Umfang erhoben, gespeichert und verarbeitet, wie es für die Erfüllung der Aufgaben nötig ist.

Wir beachten die geltenden Datenschutz-, Datensicherheits- und Urheberrechtsbestimmungen und verfolgen die Debatten zu diesen Themenfeldern und leiten Erkenntnisse daraus ab.

Wir verpflichten uns, keine personenbezogenen und andere sensiblen vertraulichen Daten bei der Nutzung von KI-Tools zu verwenden (bspw. mit Prompts). Das gilt auch für die Daten von Dritten, die in anderen Zusammenhängen bezogen und/oder verarbeitet wurden.

3. Transparenz bei KI-generierten Inhalten

Wir verpflichten uns KI-generierte Inhalte transparent zu kennzeichnen und auf die Verwendung von KI zur Nachahmung von Stimmen oder physischem Erscheinungsbild einer Person zu verzichten - außer, wenn die Person ihre ausdrückliche Zustimmung dazu gegeben hat.

4. Berücksichtigung von algorithmischer Voreingenommenheit

Die Funktionsweise einer KI basiert darauf, dass sie ausschließlich das lernt und wiedergeben kann, was in ihren Trainingsdaten vorhanden ist. Somit fungiert sie als "Echo des Internets", das erhebliche Nachteile mit sich bringt. Dazu gehören die Diskriminierung von marginalisierten Gruppen sowie die Verbreitung von Vorurteilen und Rassismus. Ein weiteres Problem besteht in dem sogenannten “Drift” der Trainingsdaten, das bedeutet, dass das Modell auf veraltete Informationen zurückgreift und keine Berücksichtigung aktueller Ereignisse findet. Ein konkretes Beispiel ist, dass z.B. ChatGPT keine Antworten zu Themen wie Zeitenwende und Energiekrise geben kann.

Wir verpflichten uns, uns bei Veröffentlichung von KI generiertem Material vor dem Einfluss von algorithmischer Voreingenommenheit zu schützen, insbesondere was die von uns veröffentlichten Bilder und Texte betrifft. Das muss die menschliche Aufsicht stets berücksichtigen.

5. Einhaltung der nationalen Gesetze und EU-Vorschriften

Wir verpflichten uns, uns strikt an unsere nationale Gesetzeslage und EU-Gesetze zu halten, die sich mit der Frage der KI-Nutzung für politische Organisationen befassen. Wir müssen uns im Rahmen unserer eigenen politischen Forderungen messen lassen. Unser Vorgehen wird im Einklang mit dem Digital Services Act und dem AI Act stehen.

Hinweis: Diese Richtlinien werden regelmäßig aktualisiert, um neue technologische Entwicklungen im Bereich KI zu berücksichtigen.(Stand: Q1/2024)

CHECKLISTE FÜR DEN EINSATZ VON KI

1 • Klare Ziele definieren

- Identifiziere die spezifischen Ziele und Anwendungsfälle und prüfe, ob Du dafür tatsächlich den Einsatz von KI benötigst
- Stelle sicher, dass der KI-Einsatz die übergeordneten politischen Ziele von BÜNDNIS 90/DIE GRÜNEN unterstützt.

2 • Datenverfügbarkeit und Datenqualität sicherstellen

- Überprüfe welche Daten Du für den KI-Einsatz wirklich brauchst und dann die Qualität der verfügbaren Daten, die für Verwendung von KI benötigt werden
- Stelle sicher, dass Du Maßnahmen ergreifst, die den Schutz von sensiblen und persönlichen Daten sicherstellen (s. Guidelines 2.)

3 • Sicherheit

- Implementiere Sicherheitsmaßnahmen, um eigene KI-Anwendungen vor Cyberangriffen und Datenlecks zu schützen.

4 • Ressourcenplanung:

- Ermittle die benötigten Ressourcen, einschließlich finanzieller Mittel, Fachkenntnisse, Infrastruktur und Zeitrahmen für die Anschaffung und Implementierung der KI-Anwendungen (z.B. ChatGPT Pro Account).

5 • Auswahl der richtigen KI-Anwendung

- Schaue Dir verschiedene KI-Anwendungen an, um diejenigen auszuwählen, die die Erfordernisse unserer Guidelines und zu Deinen verfügbaren Ressourcen passen.

6 • Transparenz und Haftung

- Kennzeichne mit KI-Anwendungen erstellte Inhalte so präsent wie möglich
- Überprüfe die rechtlichen und ethischen Aspekte des KI-Einsatzes, einschließlich Datenschutz, Sicherheit, Fairness, Transparenz und Haftung.

7 • Integration mit bestehenden Systemen

- Berücksichtige die Integration von KI-Anwendungen mit den bestehenden Systemen und Prozessen in Deiner Organisation, um einen reibungslosen Betrieb sicherzustellen.

8 • Schulung und Kompetenzaufbau

- Stelle sicher, dass die Mitarbeiter*innen, die die KI-Anwendung verwenden sollen, über die erforderlichen Fähigkeiten und Kenntnisse verfügen, um mit den neuen KI-Technologien umzugehen.
- Biete angesichts des raschen Entwicklungstempos bei KI-Technologien so oft wie möglich Schulungen an, um das Verständnis und die Kompetenz im Umgang mit der zu implementierenden KI-Anwendung zu fördern.

9 • Messung und Evaluation

- Definiere realistische Messgrößen und Leistungsindikatoren (KPIs), um den Erfolg der KI-Anwendung zu messen und zu bewerten.

- Führe regelmäßige Bewertungen durch, um sicherzustellen, dass die KI-Anwendung die erwarteten Ergebnisse liefert und die übergeordneten politischen Ziele von BÜNDNIS 90/DIE GRÜNEN unterstützt.

10 • Skalierbarkeit und Flexibilität

- Berücksichtige die Skalierbarkeit und Flexibilität der KI-Anwendung, um zukünftige Anforderungen und Veränderungen in der Organisation zu berücksichtigen.

11 • Kontinuierliche Verbesserung

- Implementiere Mechanismen zur kontinuierlichen Verbesserung der KI-Anwendungen und -Prozesse basierend auf Rückmeldungen und neuen Erkenntnissen.

4 Tipps, wie du mit KI-generierten Fake News im Netz umgehen kannst

Im Internet verbreiten sich zunehmend Bilder und Videos, die mithilfe künstlicher Intelligenz erstellt oder bearbeitet wurden. Diese Medien zeigen Szenen oder Personen, die nie in der Realität stattgefunden haben. Insbesondere im politischen Umfeld werden sie für gezielte Desinformation genutzt, um den öffentlichen Diskurs zu beeinflussen. Aufgrund der schnellen Verbreitung, insbesondere auf Social Media, ist es wichtig, manipulierte Inhalte erkennen zu können.

Neben einem Blick in die Kommentare, ob nicht vielleicht dort schon auf KI generierten Inhalt aufmerksam gemacht wird, kommen hier 4 Tipps für das Erkennen von KI generierten Inhalten.

1. Auf Details und optische Mängel achten.

Auf den ersten Blick wirken die Bilder oder Videos meistens täuschend echt, doch bei genauem Hinschauen erkennt man oft kleine Unstimmigkeiten oder Fehler. Bei hochauflösenden Versionen lohnt es sich, ins Bild hinein zu zoomen, um optische Ungereimtheiten zu erkennen. Vielleicht sind Teile des Bildes verschwommen, die eigentlich scharf sein sollten. Oder es treten Schatten an Stellen auf, wo eigentlich keine sein dürften. Spiegelnde Flächen können ebenfalls Probleme verursachen.
Achte auf häufige Fehler wie zusätzliche Finger, unnatürliche Zahn- oder Brillengestelle. KI-Generatoren machen oft Fehler bei Details wie Ohren, Augen und Haaren. Schau nach, ob die Körperproportionen der abgebildeten Personen stimmig sind. Unnatürliche Verhältnisse, wie zu kleine Hände oder unpassende Kopf-Füße-Verhältnisse, können auf KI-Generierung hinweisen. Zudem wirken Bilder von KI-Programmen, wie Midjourney oder Dall-E, oft idealisiert und ästhetisch perfekt. Misstraue zu perfekten Darstellungen und prüfe nach, ob die Inhalte möglicherweise künstlich bearbeitet wurden. Manche KI Generatoren setzen ein Wasserzeichen auf das Bild (bspw. Dall-E mit den 5 Farbkästchen unten rechts).

Hinweis: OpenAI hat vor Kurzem mit "Sora" ein Programm vorgestellt, mit dem ki-generiertes Video-Material von höchster Qualität erstellt werden kann. Daher sollte man insbesondere bei "zu perfekten" Videos skeptisch sein.

2. Kritisch bleiben und selbst recherchieren.

Verlasse dich nicht auf den ersten Eindruck, selbst wenn das Material authentisch erscheint. Teile keine verdächtigen Inhalte im Netz, ohne sie vorher selbst zu prüfen. Es empfiehlt sich, eine kurze Recherche durchzuführen, um Informationen zu sammeln. Oft werden Inhalte aus dem Zusammenhang gerissen und sind vor langer Zeit in einem völlig anderen Kontext entstanden. In solchen Fällen kann eine Bilder-Rückwärtssuche mit Suchmaschinen wie Google schnell Klarheit schaffen. Zudem kann man überprüfen, ob die schriftlichen Angaben plausibel erscheinen: Wann und wo wurde das Material zuerst veröffentlicht? Decken sich die Angaben mit anderen Informationen? Was haben Medien an dem Tag berichtet?

3. Keine Zeit für eigene Recherche? Verdächtige Bilder und Videos melden.

Wenn dir die Zeit oder das Hintergrundwissen fehlt, um selbst zu recherchieren, kannst du verdächtiges Material kostenlos an Online Dienste wie bspw. https://wasitai.com/,

https://www.aiornot.com/ oder Meldestellen wie den Correctiv-Faktencheck, den Faktencheck der dpa oder die Faktenchecker der Deutschen Welle schicken. Alternativ empfiehlt es sich, Inhalte direkt auf den Social Media Plattformen zu melden.

4. Expert*innen-Tipp: Mit Geolokalisierung den Ort bestimmen.

Geolokalisierung ist eine etablierte Methode, um Fälschungen aufzudecken: Dabei identifiziert man Objekte oder Merkmale im Hintergrund des Bildes - beispielsweise spezifische Gebäude, Straßenschilder, Gegenstände oder Landschaftsmerkmale wie Bäume oder Pflanzen. Hier kannst du Karten- oder Bilderdienste nutzen, um den Ort zu identifizieren. Vergleiche die erkannten Objekte und Merkmale mit realen Gegebenheiten. Passen sie zu dem, was man in dem Video vermeintlich sehen soll? Unstimmigkeiten können hier Aufschluss über eine mögliche Manipulation geben.