Lernende Systeme

GERMANY'S PLATFORM FOR ARTIFICIAL INTELLIGENCE

AI in the super election year of 2024

White Paper

Bieber, C., Heesen, J., Grunwald, A. & Rostalski, F.
Working Group IT Security, Privacy, Legal and Ethical Framework

Executive Summary

The increasing spread of generative artificial intelligence (AI) such as ChatGPT poses new challenges for democracy. Democratic societies need public discourses, which in turn require trust and reliability. False or inaccurate, but at the same time convincing, information in social media undermines the knowledge base for these discourses. Although fake news and disinformation are nothing new on the internet, they are on the rise with the advent of generative AI. The potential impact of generative AI on democratic societies has therefore become an important topic – especially in the super election year of 2024. Finally, more than half of the world's population in over 40 countries will be called to vote.

The (direct) influence that generative AI actually has on elections and the formation of political opinion cannot yet be proven due to the insufficient data available. However, regardless of this, democratic societies are called upon to deal with this technology, which is already firmly established in our everyday lives, with all its advantages and disadvantages, and to take appropriate measures such as a duty of origin or strengthening social AI skills to curb manipulation and influence – including with the help of generative AI.

Effects of generative AI on opinion-forming and elections

Numerous real-life examples already show the (potential) impact generative AI can have on elections and political opinion-forming and the potential this could have for future election campaigns: In the US primaries, for example, a deceptively real AI voice of President Joe Biden was used to spread false information surrounding the Democratic Party primaries for the US presidential election. In Slovakia, a fake audio file emerged ahead of the national elections that was intended to implicate party leader Michal Šimečka and a journalist in a vote-buying scheme. In Germany, Chancellor Olaf Scholz was the target of a deepfake and a fake video of the Tagesschau news program was circulated. This shows: The use of generative AI in a political context is already a reality – in the form of deceptively real fake images, videos or audio recordings as so-called deepfakes in the run-up to election campaigns.

Use-Cases of AI-supported influences on elections

AI tools can be used at various points in the electoral process and can both trigger threat scenarios as well as having a supporting or safeguarding effect. AI can provide **support** in a variety of ways, particularly in **election campaigns**. One example is the use of AI chatbots such as the American chatbot "Ashley", which called thousands of people to campaign for candidate Shamaine Daniels during the 2020 US election campaign. These AI bots enable candidates to make mass personalized calls in election campaigns and thus achieve a reach that was previously not possible. The challenge, however, is to ensure that such systems do not develop an unwanted bias: Even if, as in the case of Ashely (recognizable as a non-human voice), the AI voice itself indicates at the beginning of the conversation that it is an AI.

The use of AI chatbots by politicians, political parties or public opinion-forming institutions such as associations or campaign networks therefore poses particular challenges in terms of transparency and data quality. Particularly in Germany, with regard to AI regulations such as the EU's AI Act, transparency obligations apply to AI systems that interact directly with humans: AI chatbots must therefore identify themselves as such. However, these regulations will not yet take effect in the 2024 election year. However, such "robo-calls" will not play a role in Germany due to data protection laws, as advertising calls without consent are prohibited.

The increasing use of language models and chatbots also provides the basis for **(incorrect) voting information**. A qualitative problem of models such as ChatGPT are the so-called "hallucinations", in which language models output incorrect or invented information. This can lead to misleading information about elections, such as incorrect assignments of candidates or parties. There is also a risk that language models reinforce human prejudices. This is particularly problematic if people have too much trust in the models and do not check the information further. Nevertheless, well-regulated chatbots can also help to provide voters with better and more personalized information to encourage political participation.

Generative AI systems, especially large language models, offer political actors in particular new opportunities to **create target group-oriented content**. Parties can use them to adapt election programs to the current mood of the electorate and generate target group-oriented content, for example by adapting the language to different voter groups or translating it into plain language. Populist parties, on the other hand, could use these technologies to disseminate polarizing content tailored to different target groups in order to exploit and reinforce political sentiment.

In addition to unintentional misinformation through AI language models, **the targeted malicious manipulation** of voters **through so-called deepfakes** – i.e. deceptively real-looking images, audio or video recordings – is a problem. Manipulation in itself is not a new phenomenon, but with the help of generative AI, the possibilities for exerting influence reach a new dimension. In the context of political influence, they are used, among other things, to misquote politicians or cast them in a negative light, thus undermining trust in serious information during election campaigns. Deepfake pornography also works in this direction, for example by blackmailing public political figures, predominantly women.

Deepfakes greatly facilitate the manipulation of public opinion, especially aided by rapid dissemination on social media, which can pose a significant threat to democracy. AI-powered disinformation could undermine trust in political institutions and deepen social divisions. However, it is still unclear whether AI-generated fake news actually has any measurable impact on election results or whether it differs from conventional digital misinformation. Nevertheless, democratic systems must protect their processes, even if direct effects on election results are difficult to prove.

EXAMPLE

Deepfake-Image to influence voting behaviour of afroamerican voters

Source: Spring, 2024

Measures to safeguard the integrity of democratic elections

Irrespective of the current relevance of AI-influenced election manipulation, democracies must remain vigilant and take appropriate legal, social and technical measures.

Legal measures can be used to implement requirements such as transparency, traceability, quality of training data or range limitation when using AI in a political context. For example, the EU's AI Act classifies AI systems that can influence elections as high-risk AI. These systems must therefore meet quality standards and be labeled transparently. The Digital Services Act (DSA) also requires large platforms to carry out risk assessments and take measures against disinformation. According to the German Network Enforcement Act (NetzDG), it is also possible to block accounts or bots that spread false information or limit their reach. However, when considering and acting in this way, the interference and thus the protection of freedom of opinion and speech by state – but also private actors such as social networks – must always be taken into account and examined.

In addition, it is important to promote the skills of citizens as a **social measure** to strengthen resilience to AI manipulation. An AI-competent public is less susceptible to disinformation. Educational programs that promote a critical approach to media and AI should therefore be expanded in a targeted manner. Journalists and media professionals must also be trained in dealing with AI, both in the use and evaluation of sources. At the same time, safeguarding democratic processes remains a task for society as a whole – responsibility must not be left to citizens alone.

In addition, there are various **technical measures** to prevent or at least reduce AI-generated disinformation. These include, for example, cryptographic proof of origin and watermarks. Proofs of origin use digital signatures to ensure the authenticity of content but can only confirm reliable content.

Watermarks identify synthetic content and enable it to be recognized and removed. AI methods can also be used to detect deepfakes by identifying irregularities in faces, voices and videos. However, the use of such technologies is limited as attackers can adapt their methods, leading to a constant race for technological superiority. However, major tech companies have already recognized the need to make their products less vulnerable to abuse.

Outlook

Generative AI systems are now an integral part of our lives and will be indispensable in both the short and long term. Voters and players in political parties, politics and journalism must face up to this reality. This also includes the fact that generative AI can be misused for malicious purposes with the aim of influencing democratic processes and opinion-forming.

Whether and to what extent elections can be manipulated by generative AI remains unclear, however; the super election year 2024 could provide the first scientific studies. Regardless of this, all parties involved are called upon to be aware of the opportunities and risks of generative AI and to prevent misuse. This is crucial to protect the integrity of democratic processes and trust in AI systems – a process that needs to be continuously adapted.

Imprint
Editor: Plattform Lernende Systeme – Germany's Platform for Artificial Intelligence | Managing Office | c/o acatech | Karolinenplatz 4 | D-80333 Munich | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Follow us on X: @Lernende Systeme | LinkedIn: de.linkedin.com/company/plattform-lernende-systeme | Mastodon: social.bund.de/@LernendeSysteme | Status: June 2024 | Photo credit: Rawpixel/Shutterstock/Title

This executive summary is based on the white paper *KI im Superwahljahr 2024. Generative KI im Umfeld demokratischer Prozesse*, Munich, 2024. The authors are members of the working group IT Security, Privacy, Legal and Ethical Framework. https://doi.org/10.48669/pls_2024-5

SPONSORED BY THE

Federal Ministry of Education and Research

acatech
NATIONAL ACADEMY OF SCIENCE AND ENGINEERING