Lernende Systeme

GERMANY'S PLATFORM FOR ARTIFICIAL INTELLIGENCE

AI systems and the individual electoral decision

White Paper by Jessica Heesen, Christoph Bieber, Armin Grunwald, Tobias Matzner, Alexander Roßnagel
Working Group IT Security, Privacy, Legal and Ethical Framework

Executive Summary

Fake news, disinformation campaigns, upload filters or even election recommendation apps are no new phenomena. However, the use of AI systems increases their efficiency and significance. In addition, the use of artificial intelligence (AI) does often not meet the transparency requirements of a democratic public and machine action increasingly replaces human action. The discussion about AI is therefore intense and occasionally takes on a „hype character", oscillating between two extremes: far-reaching hopes for better systems on the one hand and diverse fears such as loss of control, surveillance, dependency and discrimination on the other. The debate's „hype character" can be observed in the relationship between democracy and AI – especially with regard to individual opinion-forming in the context of elections.

The aim of this white paper is to examine how AI systems can support the opinion formation process in the context of democratic elections and determine the conditions under which potential problems could be remedied. By adopting a demand-driven approach, the experts of the working group IT Security, Privacy, Legal and Ethical Framework of the Plattform Lernende Systeme pursue a factual analysis.

Potentials: Simplification of information and mobilization processes

In the context of elections, AI systems can theoretically be used in a variety of ways (see Figure 1). They hold considerable potential for simplifying information and mobilisation processes and enhancing their efficiency. At present, however, these potentials are hardly being realized, as AI systems have so far only been used in a very selective manner:

Figure 1: Potential of AI systems in elections

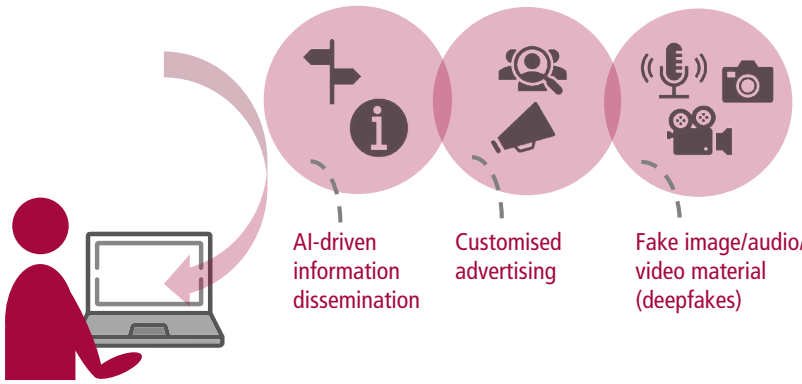
- **Election recommendation apps:** These apps – such as the German „Wahl-O-Mat" – are examples of how AI systems could be used in the run-up to future elections. Currently, these systems contain only a few automated processes and hardly any machine-based learning processes. In the future, further developments toward learning systems are conceivable (for example, as „digital assistance systems" or „digital twins").
- **Organisation of the election campaign:** At present, AI systems are still barely used by election campaigners in Germany. The objective is to improve communication with members in the context of specific election campaigns: party members can use an election campaign app to coordinate the election campaign more efficiently by identifying particularly promising regions and target groups through automated evaluation processes.
- **Election forecasts:** AI systems can also be used to develop and improve election forecasts.

Challenges: Use of AI systems to influence voting decisions

One possible danger lies in the unconscious manipulation of individual voting decisions (the formation of opinion before the election, the election campaign or the motivation to vote) with the help of AI systems. In this context, the platforms themselves have faced criticism for using their direct access to data as a

means of manipulating individual opinion. When it comes to indirectly influencing an election with the help of AI systems, three aspects in particular – but not only regarding social media platforms – need to be considered (see Figure 2):

Figure 2: Effects of AI systems on voting



- **AI-driven information dissemination:** Information dissemination, increasingly AI-driven by search engines and social networks, functions according to different criteria than it is the case with the press or broadcasting. These modes of operation can easily be influenced by automation. The automated processes of AI give the impression that the behaviour originates from a human and therefore pose a high risk of manipulating eligible voters, as they can give more validity to misinformation.
- **Creation of personality profiles for personalized advertising:** Another risk lies in the creation and use of personalized content („personalization"/"micro-targeting"). When personal profiles are used in combination with results from behavioural research, there is an increased risk of manipulation. At the present time, however, it is unsettled to what extent these principles from advertising can be transferred to the public before elections.
- **Creation of fake image, audio or video material:** With the help of AI systems deepfakes can be developed and used for very different objectives. Many deepfakes show politically active persons performing actions and/or making statements that they never made (96% in the pornographic area and especially in relation to women). Conversely, the technology also provides a backdoor for people who are criticized for actions and/or statements they have made: They increasingly claim that the recordings incriminating them are a deepfake.

Risk management strategy based on AI systems

The above mentioned distorting and manipulating impacts can be countered using tools based on AI systems as part of a risk management strategy.

- **Electoral content moderation:** With the help of this strategy, content on social networks can be moderated by the social media platforms as well as by the regulatory authorities. The aim is to remove misinformation, hate speech, violent content or even deepfakes. However, platforms in particular are criticised for often not being transparent or comprehensible as to why certain content was removed (AI-based upload filters) and other content not.
- **Detection of disinformation:** Intentionally deployed disinformation campaigns on the Internet are particularly booming in times of crisis and in the run-up to elections. Social bots are often used to distribute a political agenda across different accounts. AI systems can make a valuable first contribution to the detection of fake news. So far, platforms use AI to detect suspicious patterns in content or to declare content as election advertising.
- **Balanced reporting – balancing media bias:** Tendentially biased reporting („media bias") arises from a cleverly placed choice of words and topics that present the information provided in a certain light. AI can help identify biased information offers and provide alternative information.

It should be noted, however, that this can lead to an interplay, as the use of AI systems in this context can again lead to new problems (intentional misuse as well as unintentional errors).

Social significance

In view of the increasing availability of data and the associated advancing spread of AI technology, the question must be asked as to how AI can contribute to the preservation and strengthening of democracy and the rule of law. In order to gauge the societal significance of the potential uses of AI systems in this context, a legal assessment is necessary.

Legal assessment: European and national law have so far attempted to describe the use of AI technologies in the context of elections by means of general guidelines aimed at ensuring compatible coexistence. European regulations include the Datenschutz-Grundverordnung (DSGVO), the Platforms for Business Regulation (P2B-VO) and the European Commission's regulatory proposal on AI; national regulations include the Grundgesetz (GG), the new Bundesdatenschutzgesetz (BDSG-neu), the Telemediengesetz (TMG) and the Medienstaatsvertrag (MStV). For platform operators, there are also regulations such as the Netzwerkdurchsetzungsgesetz (NetzDG), which obliges the deletion of illegal information.

Relationship between the state and democracy: The increasing availability of data and its processing will increasingly raise questions in the future about how we want to shape „the state" and „democracy". This involves a complex interplay of civil liberties, data management and norms of democratic coexistence, which itself must be the subject of participatory procedures and political deliberation processes. Simply because of the importance of democratic elections, yet also because the relationship between AI and democracy is often viewed critically in public, it is reasonable to keep observing and reflecting these developments and to look for ways to use AI to strengthen democracy.

Legal framework and options for platform operators

Normative regulatory approaches primarily address the control and management of processes as well as the dissemination and application of AI in connection with elections as effective measures. They thus also provide the framework for measures that platform operators can take:

- **Electoral content moderation:** Transparency of selection criteria as well as the right to justification – these basic principles – are of great importance when dealing with AI and algorithmic systems for the selection of information. Especially in the context of democratic elections, the validity of the selected information is crucial.
- **Combating disinformation:** Although there are currently no specific legal regulations for the area of „deepfakes" in Germany, general abstract regulations have been applied to date. In the international context, however, there are already case studies for specific legal regulations of AI-related Internet content in connection with elections.
- **Labelling:** Another instrument against false and manipulated information is the labelling of content with warnings. These warnings indicate to users that fact checkers doubt the claims of the article and refer to further verified sources.
- **Platform policies:** Binding standards based on the model of regulated self-regulation would be a starting point, as is common in Germany and many other EU countries for the media sector.

Possible design options

To strengthen the potentials of AI systems in the context of elections for individual opinion formation and to mitigate risks, possible design options are outlined that address different actors:

Social media platform operators should...

- Create transparency of moderation processes
- Implement effective grievance mechanisms in case of suspected wrong decisions
- Develop and implement general standards for social media platforms
- Promote community policing

Policy makers should...

- Promote digital sovereignty through plural content and infrastructure
- Restrict micro-targeting
- Grant research and research-based non-governmental organizations access to relevant social media platform data
- Expand research funding to combat disinformation using AI systems
- Support data journalism and net-policy formats
- Promote consistent prosecution of crimes (including deepfakes created with the help of AI) on social networks

Research and research-related nongovernmental organizations should...

- Develop procedures for mobilizing ethical guidance
- Improve existing AI applications for detecting, labelling, and deleting disinformation and its dissemination channels, and develop additional AI applications
- Analyse what organizational structure is target-appropriate to curb the development and spread of disinformation, including disinformation created with the help of AI

The (critical) public should...

- Build up (digital) competencies for evaluating (dis)information
- Use public communication platforms as neutral sources of information

AI developers should...

- Use and explore opportunities for explainable AI (XAI) and fair AI
- Take a responsible approach to product development
- Expand risk analysis

Imprint
Editor: Lernende Systeme – Germany's Platform for Artificial Intelligence | Managing Office | c/o acatech | Karolinenplatz 4 | D-80333 München | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Follow us on Twitter: @LernendeSysteme | Status: September 2021 | Photo credit: scyther5/iStock/Title

This executive summary is based on the white paper *AI systems and the individual electoral decision – opportunities and challenges for democracy*, Munich, 2021. The authors are members of the working group IT Security, Privacy, Legal and Ethical Framework of Plattform Lernende Systeme. The original version of this publication is available at: https://www.plattform-lernende-systeme.de/publikationen.html

SPONSORED BY THE

Federal Ministry of Education and Research

acatech
NATIONAL ACADEMY OF SCIENCE AND ENGINEERING