

PERSPECTIVES AND IMPLICATIONS OF USING AI IN ELECTIONS / CEC OF GEORGIA
GIORGI KALANDARISHVILI, CHAIRPERSON OF THE CEC

It is a special honor for us to participate in today's event, that gives us opportunity to enrich our knowledge on the use of artificial intelligence in elections and with the joint efforts ensure the series of decisions that will support the process in the future.

So, I would like to share with you our vision, which is mainly based on our practice and on some of our findings from various document analysis that have been done on this matter.

As we all are aware technologies create pool of amazing opportunities to easy our work in any field, however, usually these innovations are followed by risks and challenges that needs to be promptly addressed. Otherwise, the harm can be as great as the benefit, and the same applies to the use of AI not only in elections, but in any aspect of our life, as it might be followed by the harmful content.

As stated in the most recent researches about the impact of AI it "greatly influences and impacts the way people seek, receive, impart and access information and how they exercise their right to freedom of expression in political campaigns, especially through online platforms and social media networks."¹ Stating this we can conclude that AI may have a tremendous impact on political life including the field of elections.

According to several resolutions by UN bodies have reaffirmed that "the same rights people have offline must be protected online".

In general, the public debate on artificial intelligence fluctuates between two positions: Far-reaching hopes for the optimization of socially-relevant processes, and on the other hand, manifold fears such as loss of control, surveillance, dependency and discrimination.

A possible risk of AI is the manipulation of individual voting decisions prior to an election through targeted disinformation campaigns. Particularly in times of crisis and in the run-up to elections, the danger of interested parties disseminating misleading information via the internet increases. **Social bots based on AI processes** are often used to amplify disinformation campaigns that are also carried out simultaneously in the press and on television. For example, automated social bots distribute their posts across different accounts simultaneously or disguise themselves by communicating in a human-like manner.

AI systems can also be used to create deep-fakes (images, videos, or audio files manipulated by AI) that blur the lines between reality and fiction.

Speaking about the Election processes in Georgia and the examples of using AI in this process we can state that we already experienced negative influence of AI in political processes including elections. A very large-scale disinformation campaign that we encountered during the post- election period in 2020 after

¹ SPOTLIGHT ON ARTIFICIAL INTELLIGENCE AND FREEDOM OF EXPRESSION
https://www.osce.org/files/f/documents/9/f/456319_0.pdf page 13

the Parliamentary Elections was significantly generated by AI driven Bots and trolls that created a tangible influence over the public opinion.

Social bots were actively used in 2020 during the negative campaign rolled after the parliamentary elections.

In response to this case after 2020 election, we established a responsive mechanism aiming to diminish the impact of widespread disinformation. We have created an internal unit – Information Analysis Division within the Strategic Communication Department of the CEC, with the trained team who 24 hours monitor the media and social networks, promptly detects the disinformation cases and develops the ways of response in order to stop the process spreading of disinformation and provide public with accurate election-related information. The unit produces reports and uses the most recent methods of media monitoring and analysis. The unit also uses AI, which searches for an article in online publications by words and information is provided promptly by using a special platform. This newly developed tool utilized through outsourcing proved to be affective so far and it is still on its way of advancing.

Despite the aforementioned negative implications of AI on public discourse, AI can also contribute to **better balancing media content. AI can help to identify biased information and provide alternative coverage.** In general, but especially in the context of elections, there are tendencies toward one-sided reporting. For this reason, platform operators are relying on the use of automated analyses that enable rapid detection and assessment of media bias and, if necessary, can serve as a basis for counter-offers. Social bots can also help in countering false and tendentious reporting. They can automatically disseminate verified information and, if necessary, interactively respond to questions in the context of elections and overall political opinion-forming.

Social media platforms are also increasingly using AI systems to detect suspicious patterns in content before elections, or to detect content as election advertising. **Special self-regulatory platform rules** now apply to such content on all known platforms. In this way, AI systems can make a valuable first contribution to the detection of fake news and support citizens in reaching informed opinions.

But still platform operators could foster the establishment of **independent boards** consisting of experts and representatives of civil society. Together with the operators, these boards would discuss common standards and codes of conduct on questions of algorithm-based electoral content management, as well as on all other topics relating to a pluralistic and non-discriminatory democratic public sphere.

It is to be expected that the systems will improve and that progress will be made, especially in the area of detecting false or even tendentious reporting. It can therefore be assumed that more and more tools for detecting fake news or deep-fakes will be offered in the future. Until now, though, most browser plug-ins and like applications need to be actively installed by the users and therefore already presuppose a critical awareness of disinformation, limiting their efficacy.

The algorithmic content moderation systems already in existence have been criticized for often being unaccountable and poorly understood. For example, the decision as to why some content is removed, and other content is not, is not transparent and comprehensible. One of the key tools of electoral content moderation is AI-based upload filters.

In this context it is required to determine the attribution of rights and responsibilities:

- Who should decide which content should be removed, for which reasons, when and how?
- How do we know what gets deleted, and whether what gets deleted violates laws or not (e.g., on hate speech, harmful content (slander), deep fakes, ...)? In other words, how do we know that an AI generated content monitor does not mistakenly remove legitimate content, and by removing it thus violates the freedom of expression?

The distinction between illegal and/or harmful content, in addition to concerns of algorithmic governance and data protection and for the potential of AI to be fully realized, are necessary to define any **regulatory framework**. A regulatory framework needs to define the level of accountability, transparency and accessibility requirements for AI and other algorithmic decision systems to uphold and, indeed promote, the freedom of expression in political competition.

It should be mentioned that according to exploration results first steps are being taken to establish regulative framework over the use of AI, thus the progress is still in the stage researching the issue. Not only the EU, but also the whole world is in the position of seeking the ways to “Establish a clear and strong legal framework on AI that ensures respect for freedom of expression online and offline, privacy and the right to participate in public affairs, in line with international standards agreed upon in international and regional treaties.”²

It should be mentioned that there are a number of regional standards developed on AI and human rights, which are applicable in parts of the OSCE region, in particular by the Council of Europe and the EU. However, it is necessary that awareness and a better understanding of the impact of AI related issues are promoted across the entire OSCE region. In particular, it is important to have a more systematic overview of the regional approaches and methodologies. To this extent, more regional and country-specific studies on positive and negative practices, as well as the exchange of knowledge and expertise at all levels, should be encouraged.

The **principle of transparency** in AI and algorithmic systems is of great importance. The planned Digital Services Act requires service providers to explain in a transparent and comprehensible way which measures their in-house regulations comprise and how they work.

A further contribution to transparency and the creation of trustworthy public communication is the **obligation to label content** generated by AI. The European Commission's regulatory proposal on AI already envisages a corresponding labeling obligation.³ Another instrument (that doesn't rely on AI) against false and manipulated information is the labeling of content with warnings that indicate to users that fact checkers doubt the claims of the article and provide references to other verified sources. In this way, users are prevented from assuming that false and manipulated information is true and from spreading it.

² <https://www.osce.org/files/f/documents/a/3/483638.pdf> page 9

³ <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>

However, as I already stated in the beginning, all of the new technologies also give us a plenty of positive opportunities to prevent voters from harmful content and AI can be used to create positive affect.

For example, bots can be programmed to step in when people share articles that contain known misinformation. We can deploy micro-targeting campaigns that help to educate voters on a variety of electoral issues and enable them to make up their own minds.

At the end, considering the complexity of the impact that these technologies can have on freedom of expression and access to information, there is a genuine risk that such technologies could have a detrimental impact on fundamental freedoms, especially when driven by commercial or political interests.⁴ However, we are at the stage of studying of approaches which at some stage already proved their effectiveness in many of the countries and we stay cautious to any new risks that new technologies may bring with them, in order to take preventive measures instructed by the adopted guidelines.

Thank you for your attention.

For additional remarks:

Studies show that bias in AI design can disproportionately affect the removal of minority groups' content. There is an additional risk of unwanted consequences when AI, having been trained in a certain cultural setting, is being used in societies with different cultural communication rules. A recent study of Twitter content, for instance, written in standard American English and African-American English, has demonstrated evidence of systematic racial bias of tweets written in African-American English. The study concluded that "these systems may discriminate against the groups who are often the targets of the abuse we are trying to detect". Besides the problem of over-removals, AI can also fail to remove genuinely illegal content, leading to serious harm for groups at risk of discrimination.

⁴ SPOTLIGHT ON ARTIFICIAL INTELLIGENCE AND FREEDOM OF EXPRESSION

https://www.osce.org/files/f/documents/9/f/456319_0.pdf page 7