AI & Elections: How can we ensure we know what we're voting for?

A St George's House Consultation

Report

Monday, 30th September – Tuesday, 1st October 2024

Contents

Summary

The consultation on "AI and Democracy" at St George's House calls for immediate action across all sectors to safeguard British elections against existing risks which, in the near future, could be rapidly worsened by AI. The group concluded that the primary solution for protecting free and fair elections is an impartial "Social Observatory" to provide the public with regular updates on the quality of information available on social media platforms and advise government bodies on policy.

While there is justifiable global concern about the existential threats of AI, there should be an emphasis on short-term, practical solutions with measurable results such as public education, verification of information online and use of existing regulation to ensure safety for all. The main danger of AI in elections is its ability to accelerate vulnerabilities which are already known to us, in particular targeted misinformation on social media platforms and the potential for bad actors to threaten participants in the electoral process including candidates, journalists and administrators.

The consultation focused on British elections since all societies and cultures require their own approaches to developing technology. International conversations, and a diversity of ideas, are vital for nuanced and rapid responses to AI that help it act as a tool for the good of humanity.

Background

In October 2024, the second consultation on AI at St George's House met to discuss the impact of AI on democratic elections. In preparation for the meeting, members of the group read several research papers. The report by the Alan Turning Institute's Centre for Emerging Technology on "AI-Enabled Influence Operations: The Threat to the UK General Election" was a key reference. It acknowledged that while there was no current indication of interference by AI in British elections, research is still limited and there is potential for interference in the future which requires immediate action. This conclusion was unanimously shared by those at the consultation who emphasised the potential threats from "bad actors", including targeted misinformation and intimidation to those involved in electoral processes.

Members of the consultation were divided into 4 discussion groups based on their expertise and interests:

1. **Players**: Identified key actors related to AI and the electoral process, including the British government, foreign governments, leading AI development companies, the public, media, regulatory bodies and social media companies.

2. **Experience**: Shared evidence of existing risks to the electoral system such as key channels of disinformation and erosion of trust in the democratic system, as well as how these vulnerabilities could develop as AI use becomes more common.

3. **Regulation**: Focused on potential legislation to guard against AI interference, although it also emphasised that codes of conduct between candidates and parties could have a greater impact for certain risks like intimidation and erosion of public trust.

4. **Solutions**: Examined ways to safeguard democracy beyond legislation, such as educational resources for the public, and information gathering.

Once the groups had shared information in plenary, participants agreed that an impartial "observatory" should be established to inform the British public about AI and democracy.

The groups retired for a second time to discuss how such a body might function. The consultation subsequently combined these ideas in the call to action included in this report.

Group Discussions

Players: Who is involved in the electoral process and potential risks to democracy posed by AI?

The "Players" group focused on positive and negative actors most likely to affect the relationship between AI and democracy, particularly those influencing public trust and information available to the voting public. The categories of players mentioned were those in industry (tech firms), research, campaigners, extremists, domestic government (electoral officials, the Civil Service and regulators), foreign government, media and the public.

The main risk posed is AI "turbo-charging" existing threats to democracy that might ultimately lead to a total lack of trust in an election result or create an "overload" of information which would make democratic decision making by voters impossible. Initial discussions of a solution by this group called for a "societal level change" considering the amount of time required to bring about regulation which might not be possible in the face of development in AI and the speed at which the technology might impact existing vulnerabilities. Such a change would largely involve a greater common understanding of AI and its risks, both within institutions and throughout the general public. The observatory would be an efficient body for education, advice and regularly updated reports on threats.

Experience: Is there evidence of AI interference in British elections, and what are the existing vulnerabilities in British democracy that could be worsened by AI?

The "Experience" group focused on AI's ability to falsify information and its impact on the electoral process. It was emphasised that women are major targets of abuse and intimidation, in particular through deepfake sexual abuse, which could have a devastating impact on the safety of current representatives and the number of female candidates willing to stand for elected roles in the future. Two other key points were made about AI interference in the electoral process. Firstly, it is currently difficult to measure impact since there are few recorded metrics. Secondly, it is difficult to disrupt a paper voting system. This method is currently used in the UK and may be important for future safeguarding against interference. Beyond this, it is hard to assure protections, especially considering both direct and indirect forms of influence. Public perception of the impact of AI on an election, or electoral processes in general, was noted as capable of undermining voting outcomes. This risk is part of a wider issue of online manipulation of public perception, often revolving around political "trigger points".

Content creation and microtargeting (specific targeting of swing voters) were addressed separately to account for long-term and short-term risks to electoral processes. Short-term risks to electoral processes include the indirect creation of echo-chambers and attacks on infrastructure, such as false claims that polls are closed as well as falsifying voters and candidates. Long-term risks from content creation include hyperbole (projection of fear and overpromises) and superspreaders of misinformation who exploit "faultlines" and induce an erosion of trust in democratic processes. However, AI could become a tool both for protecting the public against false information online and sharing accurate information. The group highlighted the positive impact of verification tools and the generation of quick results on search engines like Google. While there are no clear indicators of AI negatively impacting electoral processes so far, it is important to establish methods to measure it. Other initial solutions included verification systems for elections, focusing on perceptions of AI through advancing public understanding, watermarking, establishing individual accountability and improving public trust in power structures especially amongst young people. Since online identification could be detrimental to democratically vital work such as whistle-blowing and campaigning, various mechanisms could be used according to context. These steps would develop a social resilience to developing technologies.

Regulation: What legislation, if any, should be applied, developed or implemented to safeguard elections against AI threats?

The "Regulation" group shared the consultation's wider opinion that there should be greater regulation of social media platforms. This should be procedural regulation of moderation and data use: for example, opt-in models or explicit consent regarding information for political use. There should also be clear justification for content filtration to avoid invisible censorship.

The group emphasised the importance of maintaining public confidence in institutions, and warned against overregulation. Instead, existing legislation surrounding misinformation, harassment and disinformation could be referred to in future legal cases. Online anonymity was also highlighted as an issue, since bots exploiting such anonymity could misrepresent public concerns and create smear campaigns. The group believed the biggest impact would be on local government due to limited resources and the comparative lack of public participation. This presents a risk to the safety of local representatives and administrators as well as to the public since local government closely impacts our daily lives. Like the "Experience" group, this group called for protected online identities that could be verified by platforms and regulators when required for the public interest.

Law, however, is a long-term not a short-term solution, and disinformation is difficult to regulate due to its subjectivity. Rather than introducing major legal reforms, the group suggested a code of conduct agreed upon by those standing for office (there is an existing code for elected members). An emphasis on decency as a cultural rather than legal expectation would address existing issues like *ad hominem* attacks and falsehoods that are not caused by AI but could be worsened with the technology. However, since it is difficult to get different political parties to talk to each other at the local level, the Electoral Commission could be empowered to enforce measures where the code is ignored. The Representation of the People Act might be an appropriate vehicle for such instances.

It was also proposed that there should be meeting places, particularly online, for a tolerant electorate where ideas could be expressed freely and safely to minimise opportunity for bad actors to exploit tensions by AI interference or other means. Another suggestion for improving social resilience was a series similar to the Reith Lectures aimed at helping young people to understand electoral processes and engage in democratic debate.

Solutions: What other practical and technical solutions are there for potential threats posed by AI to elections?

The "Solutions" group put forward actions for education and verification in order to rebuild trust and react quickly to bad actors. A key aspect of this would be a non-governmental social observatory that could provide long-term measurements of AI interference as well as immediate reports (which were compared to meteorological information) for the general public. A clear understanding of how companies develop their models would improve AI's use as a tool for good since it is influenced by real-world biases, largely from specific cultures. This would require a legal responsibility for social media platforms and tech companies to provide the social observatory with information, including pre-trained models and data sets. AI could be a service to democracy through utilising its potential for rapidly sharing verified information and educating the public about information dynamics. To improve self-regulation, there could also be a British Kitemark equivalent and a ranking of veracity to inform users which social media platforms would provide them with the most verifiable information. Additionally, it could be beneficial to make it illegal to advertise politically through social media.

Education is vital for tackling disinformation. People are active, not passive, recipients of information. The EU AI Act requires companies to train their employees in AI principles and application. The same should be required for governments to inform citizens about the technology. A Raspberry Pi equivalent for teaching in schools would provide greater understanding about AI. Games that simulate aspects of the political system and the use of AI would also be engaging. Considering the impact AI will continue to have on our information gathering, the group suggested schools should also feature more skill-sets involved in rhetoric and debate. It is worth noting that subjects like English (and similar subjects like Classics, Film and Drama), Geography, History, RE, Philosophy, Journalism, Law and Politics already exist for this purpose in the British education system. This would indicate that the humanities and social sciences are invaluable for the development of social resilience and should be invested in at all levels of learning with greater emphasis on their capacity for autonomy, analysis and communication.

Groups Conclusions: The Social Observatory

The social observatory was expanded upon by all groups. The "Players" group presented it as "CERN for AI and Social Media" since the research centre has a positive public image. It was agreed by the consultation that the observatory should avoid drawing political attention and becoming a topic of debate. For this reason there should be an emphasis on "accountability" and "tech-enabled threats" rather than potentially inflammatory or attention-grabbing phrases like "AI" or "whistleblowing". In an apolitical manner likened to food standards agencies, the observatory would provide media literacy for the public and include a lighthouse approach of shining a light on good practice. The focus would first be on the UK with the opportunity to develop internationally once consolidated.

Call to Action

In response to the growth of AI and the impact it is expected to have on democratic elections around the world, the consultation group at St George's house calls for an independent social observatory to protect and inform the public, and safeguard institutions which enable free and safe voting.

AI has the capacity to 'turbocharge' existing risks such as misinformation or harassment campaigns. Women, in particular, are already experiencing online violence and threats through the use of AI generated images at an increasing rate which could impact their involvement as political representatives and journalists in future elections. For this reason, the observatory should monitor and report on wider online and offline threats to the British public such as failure to fact check by social media companies or interference by foreign actors.

Immediate action ahead of potential crises is vital. However, there is great potential for AI to improve democratic processes and empower voters at a faster rate, especially through targeted information which can encourage more people to participate in elections and increase accessibility to democratic action. AI is neither malign nor benign. The success of AI as a tool for good is based on accountability and collaboration at all levels of society at home and across the globe. This consultation focused on elections in the UK with an awareness that international cooperation and discussion would have great benefit for humanity and develop diverse solutions which could be implemented according to different cultural practices and regulation.

The greatest danger posed by AI is the erosion of trust held by the public for institutions and individuals involved in electoral processes. It was noted that young people have been particularly affected by this loss of trust. This requires reliable information and agreed codes of conduct between parties, especially during election periods. Reliable information can only be assured through increased expectation on social media and AI companies to be transparent about their data sets and algorithms. The social observatory would monitor and provide regular ratings for the public about the accuracy of information on each social media platform. While the tech sector is encouraged to develop its self-regulation, it is unrealistic to expect individual companies to prioritise public safety over profit. This is particularly important as leading social media companies like Meta and X, formerly known as Twitter, decrease their fact-checking tools and their own codes of conduct for users. It is, however, recognised that many tech companies and leaders in the industry have done positive work, particularly during the so-called "Year of Democracy", to ensure advertising standards and verified information during elections around the globe.

While regulation is necessary in this process of safeguarding, the social observatory must be independent to counter misuse by politicians, political parties or other actors. To protect social media and tech companies, information from their platforms should be provided to the observatory rather than government itself. It is noted that the concept of fact-checking has taken on partisan significance which would pose a challenge to the observatory. The consultation is emphatic in its belief that reliable information is vital for the safety of all people across the political spectrum. Rhetoric about truth in the past decade has had a detrimental effect on public trust. While there will always be differing opinions in a healthy democracy, the majority of information in an election can be verified and provided to the public with assurance. As a result of potential opposition to the observatory, it is vital that there is collaboration across political parties, institutions, media outlets and tech companies.

Considering that there are numerous actors currently threatening elections in the UK and around the globe, the social observatory would prevent risks that have not presented themselves yet. The worst-case scenarios discussed by the group all focused on hypothetical totalitarian regimes. It should be recognised that these discussions were largely based on works of literature that have become universal points of reference like Huxley's *Brave New World* and Orwell's *1984*. The artistic community should be included in discussions of AI and its potential in order to expand simulation activities and communication about AI with the general public. Similarly, it would benefit all those involved in AI discussion to consume a broad range of historic and contemporary art for an understanding of public perception and anxieties.

Another risk posed by AI to democracy is a lack of public education about electoral processes and the technology itself. The observatory could collaborate with schools and universities to teach young people about AI; one example of this was a Raspberry Pi equivalent for AI. Education requires reliable information across the media. Traditional media outlets and social media platforms, including creators, should be included in this process. There is also potential for in-house artists in the social observatory and collaboration with creative bodies to engage the public and empower them to discuss AI themselves. Therefore, education about AI should aim to focus on autonomy and action rather than fostering fear. As technologies develop, so should humanity's understanding of them.

ST GEORGE'S HOUSE

WINDSOR CASTLE
HONI·SOIT·QUI·MAL·Y·PENSE
COLLEGE OF ST GEORGE

For more information about
Consultations at St George's House
visit www.stgeorgeshouse.org

St George's House, Windsor Castle, Windsor SL4 1NJ

T +44 (0)1753 848848 E house@stgeorgeshouse.org F +44 (0)1753 848849

@StGeorgesHouse @StGeorgesHouseWindsor