Written evidence submitted by The Alan Turing Institute

The Alan Turing Institute (hereafter 'the Institute') is the UK's national institute for data science and artificial intelligence. Our teams working on Defence and Security challenges conduct policy and technical research on emerging technologies. Within this context, understanding and mitigating the threat of AI-enabled information operations to UK election security is a high priority. We therefore welcome the opportunity to provide evidence to the Joint Committee's investigation.

The role of generative AI in upcoming elections (opportunities)

1.1 Since the release of OpenAI's ChatGPT in November 2022, advancements in generative AI have significantly improved the accessibility of AI systems to non-technical audiences, as well as the realism, personalisation and scale of AI-generated content.[1]

1.2 Given the operational complexity and resources involved in running election campaigns, new generative AI tools offer a number of opportunities to both **'level the playing field' for under-resourced political campaigns** and **assist election management processes**.[2]

1.3 Firstly, generative AI could enhance political advertisements by generating speeches and promotional content which can be **tailored to different voter demographics or campaign issues**. Coupled with the speed at which these outputs can be created, this may reduce the logistical burden of microtargeting activities.[3] Such efforts have already been evidenced in the recent Argentinean presidential elections, where candidates utilised AI-generated posters to appeal to different voter blocs.[4]

1.4 Secondly, chatbots in particular could help to **improve the connection between political candidates and voters,** owing to the real-time response capabilities of these tools and the ability for chatbots to translate outputs into several languages.[5] Some US Congressional candidates have already deployed chatbots which can engage with voters in 20 different languages.[6] Similarly, chatbots could **simplify voting information,** such as summarising registration processes and polling times or locations.[7]

1.5 Finally, generative AI could **add an extra layer of election security and resilience**. For instance, in future elections it may assist in proofreading election materials alongside

---

[1] Ardi Janjeva et al., *The Rapid Rise of Generative AI: Assessing risks to safety and security*, (December 2023: CETaS), 14-16, https://cetas.turing.ac.uk/sites/default/files/2023-12/cetas_research_report_-_the_rapid_rise_of_generative_ai_-_2023.pdf.

[2] Ethan Bueno de Mesquita et al., *Preparing for Generative AI in the 2024 Election: Recommendations and Best Practices Based on Academic Research* (University of Chicago Harris School of Public Policy and the Stanford Graduate School of Business), 11, https://harris.uchicago.edu/files/ai_and_elections_best_practices_no_embargo.pdf.

[3] Christina LaChapelle & Catherine Tucker, 'Generative AI in Political Advertising', Brennan Center for Justice, 28 November 2023, https://www.brennancenter.org/our-work/research-reports/generative-ai-political-advertising.

[4] Jack Nicas & Lucia Cholokian Herrera, 'Is Argentina the First A.I. Election?', The New York Times, 15 November 2023, https://www.nytimes.com/2023/11/15/world/americas/argentina-election-ai-milei-massa.html

[5] Mekela Panditharante et al.,, 'Artificial Intelligence, Participatory Democracy, and Responsive Government', Brennan Center for Justice, 3 November 2023, https://www.brennancenter.org/our-work/research-reports/artificial-intelligence-participatory-democracy-and-responsive-government.

[6] Rashi Shrivastava, 'This Congressional Candidate Is Using AI To Have Conversations With Thousands Of Voters', Forbes, 12 December 2023, https://www.forbes.com/sites/rashishrivastava/2023/12/12/this-congressional-candidate-is-using-ai-to-have-conversations-with-thousands-of-voters/?sh=4d33615c10f7.

[7] Edgardo Cortes et al.,, 'Safeguards for Using Artificial Intelligence in Election Administration', Brennan Center for Justice, 12 December 2023, https://www.brennancenter.org/our-work/research-reports/safeguards-using-artificial-intelligence-election-administration.

human reviewers to ensure legal compliance (e.g. with mail-in ballots), as well as providing helpful analysis on polling processes. This includes optimal public transport routes for those voting in different areas, travel time estimations for voters assigned to certain polling stations and future polling location recommendations based on traffic patterns.[8]

The role of generative AI in upcoming elections (threats)

2.1 Despite these potential opportunities, generative AI could also create a number of risks for upcoming elections. '**Deepfakes'** are considered one of the most significant threats in this respect, which relate to AI-generated content that blur the boundaries between real and fake individuals or events.[9] However, it is important to consider the nuances between how different modalities are used rather than merely talking about deepfake content in general.

2.1.1 For instance, **voice cloning and audio deepfakes** are likely to pose greater problems for countering disinformation than other formats. This is due to both the ease of cloning voices with a small amount of audio, alongside the challenges in detecting when the content has been manipulated compared to more obvious signs in AI-generated imagery or videos (e.g. atypical movements).[10]

2.1.2 Deepfake content could also be used for different objectives. On the one hand, political candidates may publish AI-generated videos, imagery and audio which seek to **discredit or undermine their opponents**, in ways that may not always be reflective of real-life events or remarks. This was demonstrated by a recent deepfake of London Mayor Sadiq Khan falsely making inflammatory remarks about Armistice Day protests, where he raised concerns over how it could have led to "serious disorder".[11]

2.1.3 Domestic and foreign malign actors may further contribute to these incidents in order **to undermine public confidence in the wider information environment**. CETaS researchers have created a framework which breaks down the role of generative AI in the political information ecosystem and may help in making sense of different threats (see below).

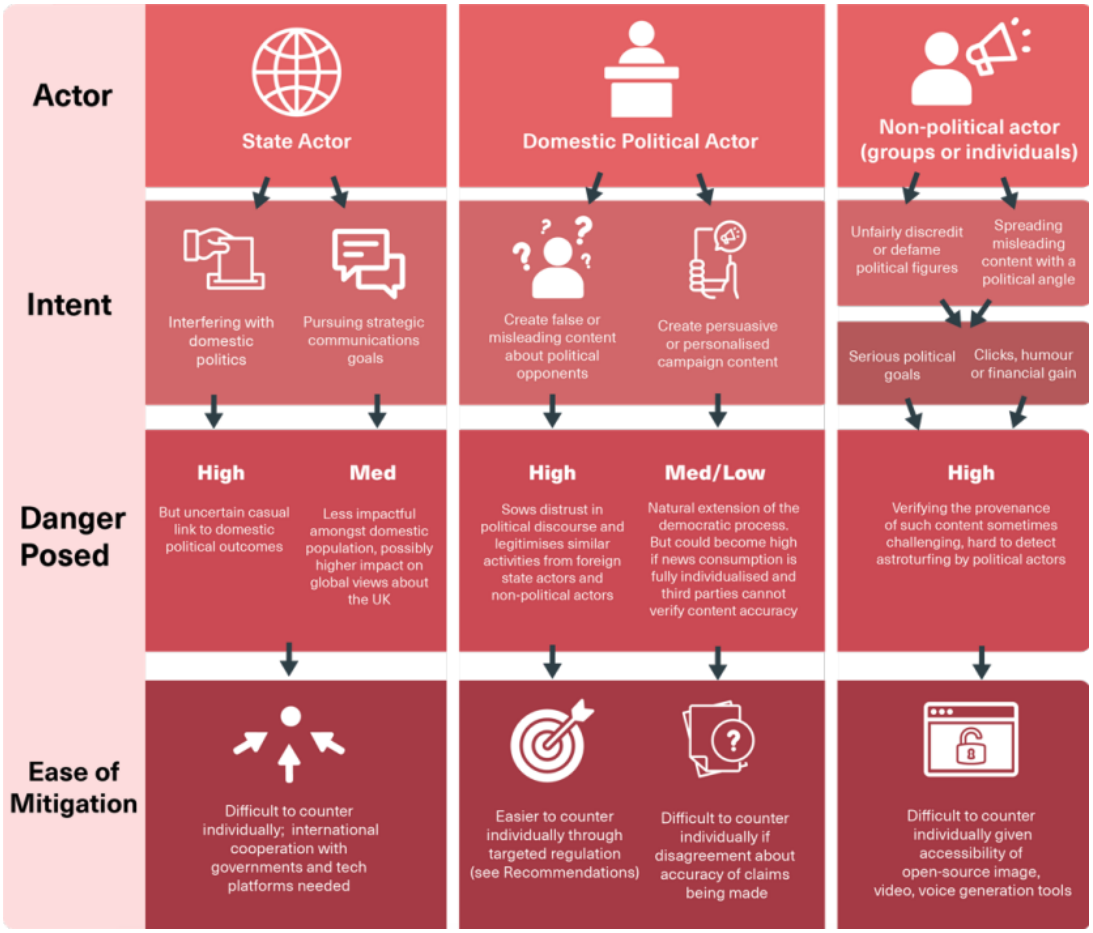
---

[8] Ibid.

[9] Dave Johnson & Alexander Johnson, 'What are deepfakes? How fake AI-powered audio and video warps our perception of reality', Business Insider, 15 June 2023, https://www.businessinsider.com/guides/tech/what-is-deepfake?r=US&IR=T

[10] Hannah Murphy, 'Audio deepfakes emerge as weapon of choice in election disinformation', Financial Times, 23 January 2024, https://www.ft.com/content/bd75b678-044f-409e-b987-8704d6a704ea.

[11] Marianna Spring, 'Sadiq Khan says fake AI audio of him nearly led to serious disorder', 14 February 2023, https://www.bbc.co.uk/news/uk-68146053.



Source: https://cetas.turing.ac.uk/sites/default/files/2023-12/cetas_research_report_-_the_rapid_rise_of_generative_ai_-_2023.pdf

2.1.4 Known as the '**liar's dividend**', certain political candidates have already sought to avoid accountability for actions by claiming that credible allegations are deepfakes. In India, audio clips released which purported to reveal evidence of corruption that an opposition leader denied as AI-generated later turned out to be authentic by experts.[12]

2.1.5 Finally, deepfakes could be used in **targeted voter suppression efforts.** In particular, 'robocalls' – or fake recordings purporting to be political candidates or election officials – may seek to discourage voters from going to polling stations, by referring to fake station closures, extreme weather incidents or public transport issues. Such strategies could try to sway the outcome of key marginal seats in certain directions.[13] A recent robocall circulated in the US state of New Hampshire which sought to portray US President Joe Biden dissuading Democrat voters from voting in the primary elections is one example of how these activities may materialise.[14]

---

[12] Nilesh Christopher, 'An Indian politician says scandalous audio clips are AI deepfakes. We had them tested', Rest of World, 5 July 2023, https://restofworld.org/2023/indian-politician-leaked-audio-ai-deepfake/.

[13] R. Michael Alvarez et al., *Generative AI and the Future of Elections*, (July 2023: Center for Science, Society and Public Policy), https://lindeinstitute.caltech.edu/documents/25475/CSSPP_white_paper.pdf.

[14] Alex Seitz-Wald & Mike Memoli, Fake Joe Biden robocall tells New Hampshire Democrats not to vote Tuesday, NBC News,, 22 January 2024, https://www.nbcnews.com/politics/2024-election/fake-joe-biden-robocall-tells-new-hampshire-democrats-not-vote-tuesday-rcna134984?utm_source=Schaake+Newsletter&utm_campaign=134a4735ce-

2.2 However, it is important to note that **deepfakes form one part of the disinformation 'toolkit' which may be enhanced by generative AI.** As such, there is a need to consider other threats that could arise in upcoming elections. One of these includes **'hack and leak' operations**. Versions of large language models (LLMs) such as 'WormGPT' have emerged which can create highly-sophisticated malware code for use in cyberattacks and digital intrusions.[15] Such tools could be deployed during an election cycle to cause various problems, from seeking to access and edit voter databases for undermining the integrity of ballot counts, to 'doxing' (e.g. leaking) sensitive personal information on election staff in an attempt to intimidate them.[16]

2.3 Alongside 'hack and leak' operations, a further threat may involve **voter mimicking efforts**. Social media 'bots' are a well-known tactic utilised within election contexts by malicious actors to try and manipulate public discourse on key political issues, helping to shape voter attitudes and behaviour towards favoured candidates. However, they typically suffer from clear signs of fakery, such as repetitive content being spread and numerous grammatical errors.[17] Yet with new generative AI developments, so-called 'agents' represent a far more powerful form of automated disinformation. These artificial entities can sense their environment, make decisions and take actions with little human control or intervention.[18] Through such capabilities, malign actors could release large volumes of autonomous agents on social media platforms to spam tailored narratives and realistic disinformation to select voter groups, whilst also dynamically responding to discussions. In doing so, this could create an illusion of political consensus on certain topics or amplify divisive content from extremist ends of the political spectrum.[19]

Prepared by:

Samuel Stockwell, Research Associate, Centre for Emerging Technology and Security and Ardi Janjeva, Research Associate, Centre for Emerging Technology and Security

23 February 2024

EMAIL_CAMPAIGN_2022_04_12_06_47_COPY_01&utm_medium=email&utm_term=0_3beb31fac3-134a4735ce-603047422

[15] Janjeva et al., 2024, 28.

[16] Cybersecurity and Infrastructure Security Agency (CISA), *Risk in Focus: Generative AI and the 2024 Election Cycle*, 18 January 2024, https://www.cisa.gov/sites/default/files/2024-01/Consolidated_Risk_in_Focus_Gen_AI_Elections_508c.pdf.

[17] Kris Shaffer, 'Spot a Bot: Identifying Automation and Disinformation on Social Media', Medium, 5 June 2017, https://medium.com/data-for-democracy/spot-a-bot-identifying-automation-and-disinformation-on-social-media-2966ad93a203.

[18] Janjeva et al., 2024,49.

[19] Neil Sahota, 'AI And The Shadow Over Democracy: The Rising Threat To Global Elections', Forbes, 2 February 2024, https://www.forbes.com/sites/neilsahota/2024/02/02/ai-and-the-shadow-over-democracy-the-rising-threat-to-global-elections/?sh=104aa22b657b.