Written Statement for AI Insight Forum: Elections & Democracy

Protecting Elections and Safeguarding Democracy in the Age of AI

Matt Masterson, Director of Information Integrity
Democracy Forward Program
Microsoft Corporation

November 8, 2023

Introduction

Thank you for the opportunity to participate in the Senate AI Insight Forum on Elections and Democracy. I am Matt Masterson, Director of Information Integrity for Microsoft's Democracy Forward Program.

One year from now millions of Americans will head to the polls to cast their vote for President, Congress, and many critical state and local races. They will be casting their ballots amidst war and conflict, concerns of nation-state electoral interference, and the rise of new and powerful technological tools, specifically generative AI. It is not unusual for elections to occur amid major societal and technological change. From Lincoln's use of the telegraph to Roosevelt's use of the radio, to the impact TV had on the Kenndy v. Nixon debates and to the use of social media in the 2008 Presidential campaign, emerging technology has been front and center in American Democracy.

AI is a powerful and transformative technology that can bring many benefits to the political process but also poses significant challenges and risks, especially for the integrity and security of democratic processes. As we enter a historic election season, with over two billion people across the globe having the opportunity to vote in significant nationwide contests, we face unprecedented threats from nation-state actors, cybercriminals, and malicious actors who seek to undermine public trust, manipulate information, and disrupt election operations. These could be amplified by the rapid development and proliferation of generative AI, which can create realistic and deceptive content that can mislead voters, impersonate candidates, and sow doubt.

Microsoft believes technology companies have a responsibility to help protect democratic processes and institutions globally from cyber-enabled threats. Microsoft's Democracy Forward team works to safeguard open and secure democratic processes, promote a healthy information ecosystem, and advance corporate civic responsibility. We partner with governments, non-governmental organizations, academics, political campaigns, and industry to protect democratic processes around the world. This is why it is vital that we, and other technology providers like us, build our products centered around the principles, which is why we developed our Responsible AI framework as we support democratic institutions around the world. We have created and iterated on corporate standards that embody these principles and defines more precise practices and policies for our teams to follow. We've implemented the standards through training, tooling, and testing systems that continue to mature rapidly. This is supported by additional governance processes that include monitoring, auditing, and compliance measures.

This framework enables us to identify risks and threats, such as those we see in the coming year as the 2024 elections approach. We view the primary risks and opportunities to elections through the lens of three key areas: Video & Image Provenance, Voting Information Integrity, and AI Enabled Cybersecurity.

Video and Image Provenance

Campaigns and election offices have historically been targets for sophisticated cyberattacks and foreign influence operations and we should anticipate that this election cycle will be no different. As we consider how adversaries may use this new technology to attack candidates in particular, the manipulation of their likeness to create a malicious depiction of the candidate is high on the list. To be sure, this is not a new threat entirely. The ability to alter a photo or edit a video deceptively has existed as long as photos and videos themselves have. While this threat is not new, the proliferation of generative AI image creation tools and face swapping software has become widely accessible. No longer does one need to have a background in photoshop or a computer science degree to deceptively alter images or videos of political candidates.

To address this threat, we suggest three key mitigations:

1- **Content Provenance Standards:** Candidates should be able to assert when media, specifically videos and photos, originates from their campaign. Industry and government should align on a common standard to watermark images and videos so that the campaigns (and other at-risk organizations) can claim what is authentic, thereby enabling consumers to reject that which is not. Microsoft recommends the Coalition for Content Provenance and Authenticity's (C2PA) content credentials digital watermark standard developed by a coalition including Adobe, TruePic, ARM, Microsoft, and others.

2- **Transparency & Labeling:** Providers of generative AI image creation tools have an obligation to provide a means of transparency regarding the origin of the images their technology creates. At Microsoft we attach the C2PA Content Credentials to every image created by our Bing Image Creator product. This is a baseline responsibility for generative AI technology providers. We are proud to have contributed to and support the Partnership on AI's (PAI) Responsible Practices for Synthetic Media framework, a helpful guide for how technology providers can responsibly develop synthetic media.

3- **Candidate Recourse:** Candidates should have recourse when their likeness is maliciously manipulated or generated by AI for the purpose of deceiving the public during a campaign.

Voting Information Integrity

Voters are the foundation of democracy and have a right to transparent and authoritative information regarding when, where, and how to vote. Elections are fast-moving and constantly changing events and nation-state adversaries continue to attempt to interfere with the systems and information environment. These operations could be further aided by generative AI, which can create false or misleading information that can confuse or mislead voters creating doubt and distrust about the process or those who run it. Intentional deception is only part of the risk regarding integrity in voting information online. Not all generative AI-enabled tools are grounded in the most recent information and may inadvertently provide voters with inaccurate information.

At Microsoft, we are continuing to improve our AI-powered Bing Chat, so answers are grounded in search results with proper sourcing, trustworthiness ratings, and reporting and feedback mechanisms. Resources for search results are displayed clearly so users can understand the third-party sites the answer is derived from. Bing Chat also integrates third-party trustworthiness ratings to help users assess content sources and access relevant, reliable, and high-authority content in Bing Chat.

We implemented safeguards to help minimize the risk that bad actors can use Bing generative AI experiences to create ungrounded or inaccurate information. These efforts, many of which are described in The New Bing: our Approach to Responsible AI, include the use of classifiers and metaprompts, provenance tools, enhanced reporting functionalities, and robust operations and incident response. Bing's generative AI features also prohibit using these services to attempt to create or share information that is fraudulent, false, or misleading, including with respect to elections.

When an AI chatbot is grounded in a search index, one key advantage is the voter can see the source of the information provided and even navigate to that page for additional information. Bing is committed to ensuring the first site a voter engages with regarding voting information is the most authoritative source of that information- the state's election site. Bing will also feature a rich election experience starting with the U.S. presidential primaries. Users will have the opportunity to discover who is running for office, understand their positions and issue stances - all from authoritative sources - and quickly find the info they need to make informed voting decisions.

While it is imperative for voters to have access to authoritative election information it is also critical that we can identify, track and expose nation station driven information operations targeting them and ultimately, our democracy. There must be a commitment to transparently communicate about these operations with voters to lessen their impact. We saw this approach work successfully with relation to Russia's invasion of Ukraine. In the days leading up to and after the invasion governments, the private sector and non-profits worked quickly to downgrade intelligence and share information to expose the Russian plans and corresponding narratives being pushed in Ukraine and around the world. This pro-active broad sharing of intelligence and information limited the scope and impact of the Russian information operations and empowered allies to respond quickly. As the 2024 election approaches Microsoft's Threat Analysis Center (MTAC) is utilizing AI to assist its analysts to identify nation state supported media and messengers to expose information operations around the world.

AI and Cybersecurity

As we continue to face a complex and evolving cyber threat landscape, from nation-state attacks to ransomware-as-a-service, AI technology will provide an important and powerful cybersecurity tool for defenders to detect, block and mitigate at speed and scale. As Microsoft noted in its most recent Digital Defense Report, a Gartner study estimates that the global cost of cybercrime is expected to hit $10.5 trillion dollars by the end of 2025. Nation-state adversaries continue to deploy sophisticated techniques to conduct cyber-attacks, with US critical infrastructure as a top target. Just one example of this is "Volt Typhoon" a China based, state-sponsored actor that

Microsoft's Threat Intelligence discovered had compromised various US critical infrastructure organizations in the US and Guam—likely for prepositioning purposes to disrupt communications in case of a US-China conflict-- through highly sophisticated techniques that makes detection very difficult.

Given these threats, embedding and enhancing technology with AI will uplevel the cybersecurity posture of critical infrastructure—including election infrastructure-- and organizations large and small, through detecting anomalies, blocking malware and deploying mitigations quickly and managing a cyber incident more effectively. Using AI to enhance cybersecurity is not just aspirational, it's happening. Microsoft successfully used AI technology to detect and then block malware that Russia deployed against a Ukrainian shipping company, and we recently released Security Copilot in preview, which will help organizations more effectively manage their security environment and incident response capabilities. Other technology companies as well are pursuing the same path when it comes to using AI to fuel cybersecurity protections. While we recognize there are important conversations to be had about how our adversaries may use AI to, for example, generate more realistic phishing emails or potentially develop malware, we should not lose sight of its important benefits.

In the election and campaign space, to help campaigns and election authorities navigate these threats and protect themselves from cyberattacks, Microsoft offers a no-cost security service, AccountGuard. AccountGuard is built for highly targeted, often low-resourced organizations and includes free access to YubiKeys (hardware authentication devices), additional security support for consumer accounts, and streamlined notifications to our customers if attacked by a nation-state cyber actor. We also offer M365 for Campaigns, which is tailored specifically for political campaigns to provide premier security offerings at our lowest nonprofit price point so it is accessible to campaigns of all sizes.

With modern AI advancements analyzing trillions of security signals daily, we have the potential to build a safer, more resilient cyber ecosystem. AI can help by automating and augmenting many aspects of cybersecurity, such as threat detection, response, analysis, and prediction. And it can also help address the cybersecurity talent gap by bridging knowledge and technical gaps as well as workflows, threat actor profiles, and incident reports across teams. The benefits that AI brings to cybersecurity are critical to addressing the complex global cyber threat environment we now face and at Microsoft, we are committed to developing an "AI-based cyber shield" to keep our customers safe and help secure the cyber ecosystem.

Conclusion

As we navigate this critical election season, it is more important than ever to remain engaged and vigilant, taking proactive steps to protect democratic institutions. Microsoft is taking these steps to help safeguard democracy during this time of rapid technological change and an ever-evolving range of threats. No one person, institution, or company can guarantee our elections are free and fair, but by working together we can make meaningful progress. The strength of our democracy depends on our actions, and Microsoft remains committed to doing our part to support campaigns, elections, and the democratic process during this transformative time.