Microsoft’s Bing Chat: A source of misinformation on elections

[Paris, France, 15 December, 2023] — Microsoft's AI-driven chatbot Bing Chat, now known as Copilot, spreads factually inaccurate and fabricated information about elections in Switzerland and Germany. This raises concerns about potential damage to the reputation of candidates and news sources. In making search engine results less reliable, generative AI impacts one of the cornerstones of democracy: access to reliable and transparent public information on the internet.

Bing Chat is a chatbot, developed by Microsoft based on the large language model (LLM) GPT-4, which combines generative AI and Bing search engine features. It formulates responses by collecting search results and summarizing findings with sources for the user. Microsoft Copilot is a feature of Microsoft Edge browser, which has 300 million users around the world that are continuously recommended to use the new AI tool.

In a collaborative investigation with nonprofit AlgorithmWatch, Al Forensics tested at scale the quality of information that Bing Chat provides about elections. Over a three-month period, from August through October 2023, researchers systematically prompted Bing Chat about the October 2023 federal elections in Switzerland and state elections in German states Hesse and Bavaria. The queries covered topics that voters might realistically search for before an election, such as how to vote, candidates running for election, and the latest election polls.

Voters should not rely on Bing’s chatbot as a source of information

We found that nearly 30% of the chatbot's answers contained factual errors, including but not limited to: false polling numbers, outdated candidates, and wrong election dates. Even more concerning, the chatbot fabricated controversies about candidates without basis in reality. This issue was found to be systemic and persisted across time, countries, and languages (with prompts conducted in English, German, and French).

“It’s time we discredit referring to these mistakes as ‘hallucinations’. Our research exposes the much more intricate and structural occurrence of misleading factual errors in general-purpose LLMs and chatbots“ says Riccardo Angius, Applied Math Lead and Researcher at AI Forensics.

Generative AI undermines trust in institutions

Factually incorrect and fabricated stories pose a risk to the reputation of cited news outlets and of candidates. For example: The chatbot attributed incorrect polling numbers to trusted news sources, even when the news source reported them correctly.

Candidates must contend with voters encountering narratives about them that have no grounding, but appear next to authoritative sources, creating the illusion of their validity. This could further sow distrust in news media if a user does not fact check the chatbot and potentially sway voter opinions about candidates.

Microsoft’s Bing Chat is an unreliable source of information during elections. More than that, it can pollute the information ecosystem by misquoting reliable sources and fabricating stories.
Microsoft seems unable or unwilling to fix the problem

Microsoft has announced measures on their company blog to protect information integrity during elections, but these fall short. The company promises to provide voters with “authoritative election information” through Bing. While the chatbot might cite reliable sources, it misquotes those sources in its answers. Second, Microsoft has promised to help candidates and campaigns maintain better control over narratives around them. However, the chatbot itself is a source of false narratives.

“Our research shows that malicious actors are not the only source of misinformation; general-purpose chatbots can be just as threatening to the information ecosystem. Microsoft should acknowledge this, and recognize that flagging the generative AI content made by others is not enough. Their tools, even when implicating trustworthy sources, produce incorrect information at scale” says AI Forensics Senior Researcher Salvatore Romano.

Moreover, before the initial report in October 2023, Microsoft Deutschland received a set of prompts that returned incorrect answers. Microsoft acknowledged the need for accurate election information and stated they had improved Bing Chat to base responses on top search results. But in a follow-up assessment, Al Forensics and AlgorithmWatch found little progress. Despite some corrections, Bing Chat continued to fabricate stories about candidate controversies and provided incorrect information about Swiss candidates and their cantons.

Regulation is needed to reign in big tech

Our findings indicate a lack of adequate safeguards in Microsoft’s chatbot and the underlying model GPT-4. By introducing generative AI to the public without necessary safety measures in place, tech companies risk undermining people’s access to reliable information. Search engines are especially vulnerable in this regard, because they hold great power in ranking information and are one of the main access points to information on the internet.

Clara Helming, Senior Policy and Advocacy Manager at AlgorithmWatch, comments: “Until now, tech companies have introduced societal risks without having to fear serious consequences. Individual users are left to their own devices in separating fact from AI-fabricated fiction.”

Governments should take societal risks stemming from big AI blackboxes seriously – by introducing and enforcing rules that hold the power of Big Tech in check.

Report link: https://aiforensics.org/work/bing-chat-elections

Contact: team@aiforensics.org