IBM.

May 22, 2024

Senator Mark Warner
United States Senate
Washington, DC 20510

Dear Senator Warner,

IBM appreciates your longstanding commitment to enacting smart artificial intelligence policies and your continued leadership in Congress on this important issue. You have been an engaged partner with us on a myriad of technology policies, and we share your concerns on election integrity.

Deepfakes are one of the most pressing challenges posed by generative AI, given the potential for bad actors to use them to undermine democracy as well as exploit artists and performers, and harass and harm everyday people.

That is why we signed the *Tech Accord to Combat Deceptive Use of AI in 2024 Elections* ("Tech Accord") as part of our long-standing commitment to the development and deployment of trustworthy AI. We appreciate this opportunity to share more about how IBM is implementing commitments in response to your May 14, 2024 letter.

As detailed below, we have technical solutions that can help clients govern generative AI models throughout the AI lifecycle, including enabling organizations to provide transparency and to proactively detect and mitigate risk. We also are collaborating externally to raise awareness, increase technical know-how, and help lay the foundation for AI safety across the world.

We look forward to continuing the dialogue with you, and advancing legislation, to address deepfake harms.

Sincerely,

Christina Montgomery
Chief Privacy & Trust Officer
IBM Corporation

IBM Response to Senator Warner's Questions

Question 1: What steps is your company taking to attach content credentials, and other relevant provenance signals, to any media created using your products? To the extent that your product is incorporated in a downstream product offered by a third-party, do license terms or other terms of use stipulate the adoption of such measures? To the extent you distribute content generated by others, does your company attach labels when you assess – based on either internal classifiers or credible third-party reports – to be machine-generated or machine-manipulated?

Question 5: Has your company worked to develop widely-available detection tools and methods to identify, catalogue, and/or continuously track the distribution of machine-generated or machine-manipulated content?

IBM Response to Questions 1 and 5

As a leading enterprise AI company, IBM provides AI-enabled technologies and services to government and corporate entities in critical infrastructure areas such as financial services, telecommunications, and healthcare. These entities rely on IBM as the trusted partner helping to advance responsible AI with a multidisciplinary, multidimensional approach. IBM believes that companies have a responsibility to develop technical and governance solutions to advance trustworthy AI. Thus, we offer technology and consulting services that place trust at the core of products and offerings and enable our clients to do likewise.

First, our integrated watsonx platform solution enables our clients to develop and deploy trustworthy AI. As part of that platform, watsonx.governance helps organizations govern generative AI models throughout the AI lifecycle with a suite of tools that enable organizations to proactively detect and mitigate risks and deploy models that are trustworthy, fair and transparent. Overall, watsonx.governance serves as a crucial differentiator for organizations looking to harness the power of AI while upholding ethical standards and minimizing risks.

Second, IBM made its open-source AI 360 toolkits available to everyone to help make AI more explainable, fair, robust, private, and transparent. These toolkits help enable organizations to develop and deploy trustworthy AI with free and open source tools.

Third, as a leading AI consulting firm, we work across our own IBM watsonx technology and an open ecosystem of partners to deliver any AI model, on any cloud, guided by ethics and trust.

Finally, in addition to developing our own solutions, IBM also collaborates with other expert organizations on issues of data provenance, transparency and trustworthy AI. As a member of the Data & Trust Alliance, IBM is actively contributing to the development of new practices and tools that can help advance the responsible use of data and AI across industries and disciplines. IBM is one of 19 cross-industry companies co-creating data provenance standards to help organizations determine if data is suitable and trusted for use. These standards currently are being tested and will offer a practical approach for organizations to gain insights into the provenance and suitability of data being used to train AI.

Question 2: What specific public engagement and education initiatives have you initiated in countries holding elections this year? What has the engagement rate been thus far and what proactive steps are you undertaking to raise user awareness on the availability of new tools hosted by your platform?

Question 3: What specific resources has your company provided for independent media and civil society organizations to assist in their efforts to verify media, generate authenticated media, and educate the public?

IBM Response to Questions 2 and 3

IBM fully agrees with the proposition noted in the letter that this requires a "multi-stakeholder approach." As such IBM leads numerous collaboration and innovation initiatives to identify and mitigate risks posed by generative AI.

First, IBM co-founded the Notre Dame-IBM Technology Ethics Lab to promote interdisciplinary research, thought, and policy leadership in artificial intelligence and other technology ethics by engaging with relevant stakeholders to drive solutions to real-world challenges. Through the Tech Ethics Lab, IBM is supporting the Pulitzer Center's "AI Spotlight Series" to enable 1,000 reporters and editors from around the world to cover and shape media coverage of AI. The program emphasizes training journalists from the Global South and from communities that are underrepresented in media, giving them the knowledge and tools to report to the public on AI and its potential impacts on society, the economy, and political systems.

Second, IBM is a founding member of the AI Alliance, bringing together more than 70 leading organizations to advance open, safe, and responsible AI. Alliance members focus on education and creating benchmarks, tools, and methodologies to enable evaluation of trustworthy generative AI.

Third, IBM is a member of the NIST AI Safety Institute Consortium, which brings together more than 200 organizations to develop science-based and empirically backed guidelines and standards for AI measurement and policy, laying the foundation for AI safety across the world. Efforts include numerous initiatives focused on mitigating risk associated with generative AI that will support addressing risks associated with mis/disinformation.

Lastly, we recognize that promoting equitable access to AI skills is also crucial to tomorrow's workforce. We announced a new goal to train 2 million learners in AI skills by the end of 2026 through expanded education collaborations and new generative AI coursework, including education on AI ethics, on IBM SkillsBuild.

In conclusion, IBM is committed to the advancement of trustworthy and responsible AI, including efforts in furtherance of the Tech Accord. We look forward to continuing to collaborate with your office and other stakeholders to develop comprehensive solutions that harness the benefits of responsible AI while protecting against its misuse.