United States Senate

WASHINGTON, DC 20510

June 3, 2024

Sundar Pichai
Chief Executive Officer
Google LLC
1600 Amphitheatre Parkway
Mountain View, CA 94043

Dear Mr. Pichai,

We write to request additional information related to the recent developments surrounding the Gemini artificial intelligence model (Gemini) released to the public in February of this year. Artificial intelligence is proving to be a powerful technology and we are optimistic that artificial intelligence can empower Americans and drive economic development. With great power, however, comes great responsibility.

While we are encouraged by Google's fundamental research in artificial intelligence such as the development of the Transformer architecture[1] and more recent breakthroughs in model context length,[2] we have concerns surrounding certain outputs generated by Google's Gemini text and image models. Specifically, we are troubled by instances where Gemini's outputs appear to misrepresent objective facts or refuse to acknowledge widely accepted truths.

Shortly after release, Gemini image model users reported that Gemini was tuned to only create "diverse" photos of humans, even when prompted to generate historical subjects such as America's founding fathers or the Pope. Separately, users of the Gemini text model shared responses where Gemini refused to acknowledge Hamas' designation as a terror organization.[3] In other conversations, the model refused to say whether "Elon [Musk] tweeting memes" or Adolf Hitler's atrocities were more harmful to humanity.[4] These outputs are concerning, as they suggest the model may be misrepresenting reality in certain contexts.

---

[1] Ashish Vaswani; Noam Shazeer; Niki Parmar; Jakob Uszkoreit; Llion Jones; Aidan N Gomez; Łukasz Kaiser; Illia Polosukhin. *Attention is all you need.* Advances in Neural Information Processing Systems. Curran Associates, Inc. June 12, 2017. Available at https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[2] Chaim Gartenberg. *What is a long context window?* Google. February 16, 2024. Available at blog.google/technology/ai/long-context-window-ai-models/.

[3] Nate Silver. *The overt political orientation of Google Gemini is really something. Here's another example I've seen from various people in the timeline, with comparison to ChatGPT. I'm not a big Middle East Takes guy but it's really, uh, different. Gemini on left, ChatGPT on right.* February 21, 2024. Available at https://twitter.com/NateSilver538/status/1760461659808674174.

[4] Megan Palin. *Nate Silver calls to shut down Gemini after Google's AI chatbot refuses to say if Hitler or Musk is worse.* New York Post. February 26, 2024. Available at nypost.com/2024/02/25/us-news/nate-silver-calls-to-shut-down-gemini-after-googles-ai-chatbot-refuses-to-say-if-hitler-or-musk-is-worse/.

We appreciate the swift responses taken by your team in identifying and taking preliminary steps to address these issues in the Gemini image outputs,[5] as well as the acknowledgement in an internal memorandum that many Gemini text outputs are "completely unacceptable."[6]

Despite these recognized problems, at the time of writing, Google has still not issued a public statement about the concerning Gemini text outputs, nor provided any concrete steps that it will take to remedy these outputs. Instead, in the week following the Gemini release, Google announced beta testing to further integrate Gemini models into Google Messages.[7] There are also reports that Google is paying publishers to write 3 articles per day using artificial intelligence tools without public disclosure, which could potentially propagate similar inaccuracies.[8]

We acknowledge the challenges inherent in developing complex AI systems that always reflect ground truth. However, given Google's leadership position in the technology industry and its vast resources, we expect Google to, at a minimum, create systems that accurately represent observable, irrefutable and objective facts.

The development of AI technologies presents a dramatic opportunity to enhance human capabilities and address some of the world's most pressing challenges. However, this potential can only be realized if the AI we create accurately reflects the real world. Considering these developments, we have several questions for Google to ensure transparency and accountability as Google continues to build more capable models. We respectfully request that you respond to the following so that the American people can better understand these confusing events.

- Please confirm whether Google was aware of these issues with the Gemini models before they were released. If yes, why did Google feel comfortable releasing a product that resulted in these outputs? If not, why did Google fail to properly identify these issues before releasing the product to millions of Americans? Please respond respectively for both the Gemini image model and Gemini text model.

- Please explain in simple terms the decisions that led to Gemini's concerning model outputs which contradict all available public information (i.e. showing images of "diverse" founding fathers). In light of such outputs, how will Google change its training processes to ensure such severe issues are identified and corrected before any future public releases? Will Google commit to prioritizing factual accuracy over ideological viewpoints when training and finetuning its future models?

- Google's lack of a public statement acknowledging the concerning text outputs from Gemini raises serious concerns about the company's commitment to transparency. How does Google

[5] Prabhakar Raghavan. *Gemini image generation got it wrong. We'll do better.* Google. February 23, 2024. Available at blog.google/products/gemini/gemini-image-generation-issue/.
[6] Reed Albergotti. *Google CEO calls AI tool's controversial responses 'completely unacceptable.* Semafor. February 28, 2024. Available at www.semafor.com/article/02/27/2024/google-ceo-sundar-pichai-calls-ai-tools-responses-completely-unacceptable.
[7] *Use Gemini in Google messages - Gemini apps help*. Google. Available at support.google.com/gemini/answer/14599070#zippy=%2Cin-google-messages%2Cin-gemini-apps-activity.
[8] Mark Stenberg. *Google is paying publishers five-figure sums to test an unreleased gen Ai platform*. Adweek. February 27, 2024. Available at www.adweek.com/media/google-paying-publishers-unreleased-gen-ai/.

reconcile this silence with its purported dedication to AI ethics and responsible development, and what steps will Google take to be more forthcoming about AI accuracy issues moving forward?

- This incident raises significant concerns about data privacy and the ethical use of data in training AI models. Could you elaborate on the protocols and safeguards Google has in place to ensure the privacy and security of the data used in developing your AI technologies? Does Google currently or have plans to train its models based on the conversations it has with users? Can you please elaborate on what protections Google has in place to ensure that personal information does not enter the training set for artificial intelligence models?

- It has been reported that Google has partnerships with major social media networks, such as Reddit, for data to train Google's AI models.[9] Given the significance of such content in training AI models, it raises the question of how partnerships influence the data and narratives that inform your AI technologies. How does Google ensure that the sources and content being introduced result in AI models that accurately reflect objective reality?

- Can you confirm whether Google is paying publishers to publish AI generated stories using unreleased software? Given the outputs from Gemini that did not reflect ground truth, what steps is Google taking to ensure that the published content is factually accurate? Additionally, are readers made aware that AI is being used to write these articles?

- There have been increasing calls for AI companies to provide users with clear disclosures about the capabilities and limitations of their AI models. This is seen as important for promoting transparency, managing user expectations, and helping people understand when they are interacting with an AI versus a human. What would a simple yet effective model disclosure from Google look like that would be most informative and useful for users interacting with your AI technologies?

Thank you for your prompt attention to this matter.

Sincerely,

Cynthia M. Lummis
United States Senator

Michael S. Lee
United States Senator

---

[9] Anna Tong. *Exclusive: Reddit in AI content licensing deal with Google*. Reuters. February 22, 2024. Available at www.reuters.com/technology/reddit-ai-content-licensing-deal-with-google-sources-say-2024-02-22/.

J.D. Vance
United States Senator