MARK R. WARNER
VIRGINIA

United States Senate
WASHINGTON, DC 20510-4606

COMMITTEES:
FINANCE
BANKING, HOUSING, AND URBAN AFFAIRS
BUDGET
INTELLIGENCE
RULES AND ADMINISTRATION

May 24, 2024

Shira Perlmutter
Register of Copyrights and Director, U.S. Copyrights Office
Library of Congress - Copyright Office
101 Independence Avenue, SE
Washington, DC 20559-6000

Dear Ms. Perlmutter,

I write today in response to the petition submitted to your office that proposes a new exemption for "Security Research Pertaining to Generative AI Bias" as part of the Copyright Office's ninth triennial rulemaking proceeding under the Digital Millennium Copyright Act (DMCA). I understand a number of stakeholders have submitted public comments to weigh in on this petition, including a letter from the Department of Justice. Ultimately, I urge the Copyright Office to consider expanding the existing good-faith security research exemption to cover both security and safety flaws or vulnerabilities, where safety includes bias and other harmful outputs.

As the leader of bipartisan legislation to improve the security of AI systems and the Co-Chair of the Senate Cybersecurity Caucus, I recognize the importance of independent security research. The existing DMCA exemption for good-faith security researchers plays a critical role in empowering a robust security research ecosystem that identifies vulnerabilities and risks to systems around the world, facilitating their remediation, and preventing future exploitation by threat actors that could lead to incidents. We must continue to promote this important work and understand that, although AI is software at its core, the non-deterministic nature of AI systems means that security vulnerabilities are no longer the only type of flaw that can be introduced and enable misuse. As the *AI Risk Management Framework*, developed by the National Institute of Standards and Technology (NIST), emphasizes, AI risks differ from traditional software risks in key ways - including increased opacity and barriers to reproducibility, complex and non-deterministic system dependencies, more nascent testing and evaluation frameworks and controls, and a "higher degree of difficulty in predicting failure modes" for so-called "emergent properties" of AI systems.[1]

Due to the difficulty in understanding the full range of behaviors in AI systems - particularly as models are introduced in contexts that diverge from their intended use - the scope of good-faith research has expanded to the identification of safety flaws caused by misaligned AI systems, as

[1] National Institute of Standards and Technology. "Artificial Intelligence Risk Management Framework (AI RMF)." NIST Special Publication 800-223. Gaithersburg, MD: National Institute of Standards and Technology, 2023.

well as research into how AI systems can reflect and reproduce socially and economically harmful biases. This research into bias and other harmful outputs is essential to ensuring public safety and equity while enabling continued innovation, public trust, and adoption of AI. Therefore, it is crucial that we allow researchers to test systems in ways that demonstrate how malfunctions, misuse, and misoperation may lead to an increased risk of physical or psychological harm.

At the same time, as the Department of Justice letter emphasized, a hallmark of the research exemption has been the good faith of security researchers. In the absence of regulation, many AI firms have voluntarily adopted measures to address abuse, security, and deception risks posed by their products. Given the growing use of generative AI systems for fraud, non-consensual intimate image generation, and other harmful and deceptive activity, measures such as watermarks and content credentials represent especially important consumer protection safeguards. While independent research can meaningfully improve the robustness of these kinds of authenticity and provenance measures, it is vital that the Copyright Office ensure that expansion of the exemption does not immunize research that intends to undermine these vital measures; absent very clear indicia of good faith, efforts that undermine provenance technology should not be entitled to the expanded exemption.

The existing exemption has been an important contributor to the multistakeholder effort to improve information security by enabling the "good-faith testing, investigation, and/or correction of a security flaw or vulnerability" in computer programs.[2] As you review the public comments on this new petition, I urge you to consider expanding the good-faith security research definition to include both security and safety flaws or vulnerabilities, where safety includes bias and other harmful outputs. In considering this expansion, I urge the Copyright Office to continue to bind the exemption to research that is conducted in a safe environment, primarily to enhance the security or safety of computer programs, without facilitating copyright infringement. Further, I encourage careful consideration of the exemption's application to any research on technical measures that protect the authenticity or provenance of content from generative AI models.

Sincerely,

Mark R Warner

Mark R. Warner
United States Senator

[2] Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, 86 Fed. Reg. 59,640 (codified at 37 C.F.R. pt. 201), 2021.