



November 29 , 2023

RESPONSE TO US COPYRIGHT OFFICE NOTICE OF INQUIRY CONCERNING STUDY OF ARTIFICIAL INTELLIGENCE AND COPYRIGHT¹

The Association for Computing Machinery (ACM) is the longest-established and — with more than 50,000 American members — the largest association of individual professionals engaged in all aspects of computing in the United States. A non-profit, non-lobbying, and otherwise wholly apolitical organization, ACM’s mission includes providing unbiased expert technical advice to policymakers on matters of our members’ wide-ranging expertise through the work of ACM’s U.S. Technology Policy Committee (USTPC).

USTPC has reviewed responses submitted to the Copyright Office in this proceeding, particularly by those representing copyright creators and owners. We note that most focus on and express concern about large-language-model generative systems. We share some of those concerns, as noted in item 2 in our June 2023 statement: "Principles for the Development, Deployment, and Use of Generative AI Technologies."² Many other things, however, also might fall under an "AI" rubric. That could have unintended consequences if new regulatory requirements are imposed or the scope of copyright protection is expanded to address these fast-evolving technologies.

We are particularly concerned about the impact of such unintended consequences to academic research in computing, something of particular interest to ACM as many of our members are faculty and students doing that important research and report their peer-reviewed results at conferences and in journals affiliated with ACM and our subject-specific "special interest groups."

For example, a requirement of getting permission to use a copyrighted work to help train an AI model may be an inconvenient cost of doing business for a company providing a “chatbot” to the public. The costs of complying with it, however, can be prohibitive for academic researchers and institutions since it is unlikely that funding agencies, such as the National Science Foundation, would underwrite the time, effort, and expense of contacting every copyright owner.

¹ See <https://www.govinfo.gov/content/pkg/FR-2023-08-30/pdf/2023-18624.pdf> as modified by <https://www.govinfo.gov/content/pkg/FR-2023-09-21/pdf/2023-20480.pdf>

² [Joint Principles for the Development, Deployment, and Use of Generative AI Technologies](#), ACM TPC, Europe/US Technology Policy Committees (June 27, 2023)

This means that researchers who want to try to solve the many problems associated with AI (such as detecting “deep fakes,” preventing “hallucinations,” “unlearning” information, and reducing computing’s energy demands) might have to do their research using limited training material not representative of the real world or be unable to do the research at all. Such a result would be enormously ill-advised and potentially economically disastrous as it is through past academic research that the United States leads the world in employing artificial intelligence productively.

Indeed, researchers in virtually every area of computing research today are exploring how machine learning techniques might be productively applied. This holds true not just in areas that seem closely related to artificial intelligence (such as pattern recognition, message and data understanding, and image generation), but also disciplines such as circuit design and network bandwidth control. That trend is certain to continue. Already, GitHub (now a subsidiary of Microsoft) is offering GitHub Copilot, an AI-based tool to help make software developers more productive.³ Trying to impose the same requirements on those varied uses as in effect for current chatbots risks unintentionally impeding such creative and economically valuable development.

USTPC also notes that, if the scope of a requirement is too broad, it can impose nonsensical burdens. By way of example and analogy, it’s generally agreed that there should be a requirement for informed consent for research on human subjects, such as in clinical trials of new drugs or therapies. But, if defined broadly in this context, “research on human subjects” could mean merely asking a few friends to try a new software system. Regulations should be written to try to avoid this potential overreach, as they have been successfully in the past.

Conversely, too broad an exception can result in a loophole that swallows the statute or regulation that contains it. During the development of the DMCA, some parties sought a blanket exemption from the trafficking restrictions for “encryption research.” Congress realized, however, that such broad provision would allow a circumvention tool to be distributed as an attachment to a self-published “research paper.” Accordingly, Section 1201(g) was enacted to delineate permissible acts of encryption research.

As USTPC member Professor Lee Hollaar cautioned in his initial comments in this proceeding, “Without a precise definition of what is an AI system, it makes little sense to consider special requirements for such systems.” This is particularly true when a provision written with only large-language-model generative systems in mind could be read to also reach the unrelated but critical academic research of ACM members.

USTPC stands ready to assist the Copyright Office to strike the appropriate balance in this regard as it prepares to draft statutory or regulatory language based on the comments received. Please contact the ACM Technology Policy Office for additional information or to arrange a briefing by ACM’s and USTPC’s expert members.

³ The Codex LLM is based on 5 billion lines of open source code. When users of Copilot, a software editing tool, prompt it to help with a programming task, Copilot suggests code that responds to users’ requests. Copilot reportedly now has more than 1 million subscribers and is considered a very useful tool for and by programmers.