WINTERFELDT IP GROUP

1802 Vernon Street NW, Suite 1091
Washington, DC 20009

October 30, 2023

Submitted via regulations.gov

Ms. Suzanne V. Wilson
General Counsel and Associate Register of Copyrights
U.S. Copyright Office
101 Independence Ave. S.E.
Washington, D.C.
20559-6000

Ms. Maria Strong
Associate Register of Copyrights and Director of Policy and International Affairs
U.S. Copyright Office
101 Independence Ave. S.E.
Washington, D.C.
20559-6000

Re: Comments on Artificial Intelligence and Copyright Regulations and Policy Issues under Review by the U.S. Copyright Office

Dear Ms. Wilson and Ms. Strong,

Winterfeldt IP Group (WIPG) submits these comments in response to the U.S. Copyright Office's Notice of Inquiry to address the sudden and meteoric rise of artificial intelligence, especially generative AI technologies, as it relates to copyright regulations and policy issues.

These comments will focus on three primary issues raised by artificial intelligence implicating current copyright law: (I) Human Authorship Requirements as Applied to Generative AI Technologies; (II) Infringement by Generative AI Technologies through Large Language Model Training; and (III) Requirements for Disclosure and Transparency for AI-Generated Works.

As artificial intelligence advances in sophistication and breadth of application to daily tasks and created works, legal guidelines and regulations must appropriately and effectively address challenges to current regimes. Where established provisions and precedents do not sufficiently apply to developing questions surrounding artificial intelligence, the Copyright Office must provide recommendations as to additional regulatory and legislative actions. Such actions must clarify the degree to which artificial intelligence may be utilized throughout the creation of a work before it is ineligible for copyright protection. Industry practices must also be developed to prevent infringement by artificial intelligence large language models trained using copyrighted material. Further, policies must be implemented within the Office as to applicants and works where artificial intelligence was

utilized in some capacity in order to provide transparency with regard to the degree of human authorship. These recommendations and policy implementations will enable the current copyright regime to adapt, evolve, and overcome the regulatory and policy challenges presented by artificial intelligence.

I. Human Authorship Requirements as Applied to Generative AI Technologies

U.S. law has long established, and recent case law has clarified and reinforced, that human authorship is a bedrock principle of copyright eligibility. Protections afforded to creators of expressive works through U.S. copyright provisions purposely aim to enable authors to enrich themselves creatively and in other ways through exclusivity in certain endeavors involving their copyrighted works, including distribution, derivation, and display. The enjoyment of these rights has been understood and defined to lie with human authors.

Although AI presents a challenge to conventional understandings of authorship, the creativity and expression unique to humans as artists cannot be supplanted or subverted by any technology, no matter its sophistication. Where AI may be said in some ways to create, to produce works with merit not entirely dissimilar to the most basic of artistic processes, the intention of human authorship knowingly to create something of higher value and meaning, if only to the author, could never be replaced in terms of artistic and cultural significance and should never be accounted equally with legal weight. As U.S. copyright law has established since its inception, and subsequently explicated through ensuing centuries that have witnessed ever-increasing advances in technologies that have aided in producing works of art, copyright protections should and do apply only to human authors, a qualification that must not be expanded to include creators with artificial intelligence.

II. Infringement by Generative AI Technologies through Large Language Model Training

The challenges of AI faced by current copyright regimes extends beyond qualifications of authorship. Perhaps more concerning is the potential for large-scale, untraceable, and unactionable infringement by Large Language Models such as OpenAI's ChatGPT.

Sarah Silverman has alleged in a complaint filed in the U.S. District Court for the Northern District of California that OpenAI's ChatGPT language models were trained on her copyrighted works without her permission, infringing on her rights. OpenAI had previously disclosed using large internet book databases for their prior language models. For example, OpenAI disclosed that a portion of the training dataset used for ChatGPT-3 came from "internet-based books corpora" that are estimated to have collectively contained hundreds of thousands of titles. Silverman alleged that only a handful of potential "internet-based books corpora" could contain that many titles, and that these were illegal "shadow library" collections. ChatGPT-4 did not disclose the methods or data sets used for language model training. When asked to summarize the works of Silverman, ChatGPT returned largely accurate summaries. Silverman has alleged that this is because ChatGPT retains knowledge or wholesale copies of the works within its data sets and is able to replicate those works in whole or in part, amounting to unauthorized copies in violation of her copyrights.

This case could significantly impact the future of AI machine learning if copyrighted works cannot be used in data sets for language model training. Generative- AI programs do exist currently, such as Adobe's Generative Fill program, that rely solely on stock photos within its own databases to avoid potential infringement of copyrighted works. However, an AI language model is only as good as the data sets that it is fed; utilizing only public domain books for language model training could result in outdated and limited responses. Daniel Gervais, a professor at Vanderbilt Law School, has posited that use of copyrighted material for the training of AI systems may be considered fair use. However, the generation of content using that copyright material may not be considered fair use. This may be a distinction in theory rather than in practice since using copyrighted material to train generative AI programs will necessarily result in content generation. New considerations may arise for significantly transformative content or content that is so distinct from the copyrighted material on which it was based that it is transformative. Limitations on prompts may help promote novel, transformative content. The

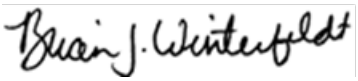
commercial nature of GenAI large language model training will also especially impact fair use analyses of large language model training.

III. Requirements for Disclosure and Transparency for AI-Generated Works

In order to combat the increasing prevalence of generative AI programs and the ease of their accessibility and ability to produce works based on complex human prompts, especially when inviting reproduction or imitation of copyrighted material, the Office must develop and implement guidelines and policies aimed at ensuring transparency through requirements for disclosure of the involvement of artificial intelligence programs in any applications for copyright registration. While works entirely produced through the use of artificial intelligence are ineligible for copyright protections, as the Office and federal courts have clarified, the Office has explained that works in which artificial intelligence was merely a tool partially responsible for the creation of a work will not be discounted for that reason outright. In the recent matter regarding the comic "Zarya of the Dawn," in which the human author employed generative artificial intelligence to produce only the images for the work, the Office granted a limited registration covering the work as a whole, but not the AI-generated images.

Beyond the realm of copyrightability, the use of artificial intelligence in any capacity for producing artistic works has become an issue of contention among artists and consumers alike. Rising sentiments in opposition to the use of AI in particularly rest on the use of copyrighted material for training Large Language Models, as well as the present threat to the livelihoods of artists who already struggle to pursue their craft in an economically viable manner. In order to meet the demands of consumers who wish to support art and artists that do not employ artificial intelligence, the Office should make recommendations for legislation requiring disclosure of the use of artificial intelligence in producing works. Such disclosure requirements will enhance transparency of the degree to which AI is implemented at all stages of production and will allow consumers and the marketplace at-large to dictate and determine whether and how AI should be utilized in the creation of art.

Respectfully submitted,



Brian J. Winterfeldt
Winterfeldt IP Group, PLLC

Joseph Theis
Winterfeldt IP Group, PLLC