ECPA
EVANGELICAL CHRISTIAN PUBLISHERS ASSOCIATION

5801 S. McClintock Dr. Suite 104
Tempe AZ 85283
t 480.966.3998
www.ecpa.org

Electronic submission via *www.regulations.gov*

October 30, 2023

The Hon. Suzanne V. Wilson
General Counsel and Associate Register of Copyrights
The Hon. Maria Strong
Associate Register of Copyrights and Director of Policy and International Affairs
U.S. Copyright Office
101 Independence Avenue SE
Washington, DC 20559-6000

Re: Notice of Inquiry and Request for Comments on Copyright Law and Artificial Intelligence Systems (Docket No. 2023—6)

Dear Associate Register Wilson and Director Strong:

The Evangelical Christian Publishers Association (ECPA) appreciates the opportunity to provide its views to U.S. Copyright Office on matters related to copyrighted works and the evolving field of artificial intelligence (AI).

ECPA is the national trade association for evangelical Christian book publishers. ECPA's mission is to equip its member-publishers so they can more effectively make the Christian message more widely known. Religious publishers comprise a portion of the book publishing industry accounting for $1.27 billion in annual revenue and thousands of jobs at both commercial and non-profit publishers.[1] Further, Christian publishers serve as a resource to Americans of diverse religious and political persuasions on matters of spirituality, morality, philosophy, and yes, advancing forms of technology like AI.

The year 2023 saw the rapid adoption and advancement of AI technologies in the U.S. copyright industries as well as everyday American life. Newly developed AI tools afford the book publishing industry tremendous opportunities in streamlining tasks in editorial, marketing, and almost every corner of a publishing organization. But ECPA approaches this new horizon with cautious optimism, seeking to promote a legal and regulatory framework that encourages the practical evolution of AI technologies for the betterment of numerous areas of life while guarding against risks and adverse consequences that may be unforeseen or ignored by individual actors in the space. In approaching these comments, ECPA encourages policy-makers to consider decisions that include: (i) the respect of copyright and authors in the training of AI models and the corresponding AI-generated outputs; (ii) a system of transparency, disclosure, and multi-stakeholder involvement in the development and training of AI, not only for the consideration of those stakeholders' interests but to create AI technologies whose outputs are reliable and fair; and (iii) the legal recognition within the copyright framework of human creativity adapting to new technological tools and the legal protection of the creative expressions birthed by human authors with the use of powerful assistive machines.

[1] AAP 2022 StatShot Annual Report. Publishers.org

Below ECPA responds to the questions that inquire into the U.S. copyright structure's suitability and adaptability to issues raised by generative AI:

1. **As described above, generative AI systems have the ability to produce material that would be copyrightable if it were created by a human author. What are your views on the potential benefits and risks of this technology? How is the use of this technology currently affecting or likely to affect creators, copyright owners, technology developers, researchers, and the public?**

Like the book publishing industry as a whole, the advent of generative AI presents opportunities and risks to the Christian publishing community. Much has been made of the potential efficiencies created by AI tools in industries from healthcare to national defense, and publishing is no exception. These efficiencies may be even more valuable to independent publishers and church-owned presses which often employ fewer staff members than larger, corporate-owned publishers such as the "Big Five" publishers (Penguin/Random House, Hachette Book Group, Harper Collins, Simon and Schuster, and Macmillan). The expanded capacity created by generative AI tools may significantly increase the productivity potential of marketing, editorial, and other areas within most traditional publishing houses. Existing AI tools already assist editors in discovering plot and language inconsistencies. Marketing departments may use AI to crawl the web for mentions of previously published-titles and develop defined marketing plans more likely to reach interested customers. However, the treatment of works utilizing AI tools within the current policy and regulatory structure diminishes the value that these tools, when properly trained, may bring to the Christian publishing community. A system in which publishers may utilize these quickly-evolving tools without jeopardizing the exclusive rights the publishing industry relies on for economic viability would further the interests of the publishing community and copyright itself.[2]

While ECPA and its members approach the technological advancements brought by AI tools with cautious optimism, they also have concerns. The rapid evolution of generative AI technology risks outpacing the legal and moral framework that could place protective parameters around its implementation in many sectors. While early decisions from the Copyright Office address the copyrightability of AI outputs[3], the mass infringement of copyrighted works for the purpose of training AI large language models is of equal concern to publishers. The large-scale reproduction of works to which publishers control exclusive rights for the purpose of training models capable of producing competing works would be detrimental to the publishing industry and the very purposes of copyright. The trampling of not only copyright but also rights of publicity by "deep fakes" and the proliferation of misleading information are concerning to not only the creative industries but society as a whole. The ease with which generative AI creates content that may be brought to market risks a flood of content and information, much of which will be of a low standard in both artistic expression and information value, that drowns out talented voices honed by professional-quality editing and publishing.

---

[2] Article I, Section 8, Clause 8 of the U.S. Constitution grants to Congress "Power . . . To promote the Progress of Science and useful Arts, by securing for limited Times to Authors . . . the exclusive Right to their respective Writings . . . ".

[3] See United States Copyright Office, *"Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence"* (March 16, 2023) at 16190-16194.

2. Does the increasing use or distribution of AI-generated material raise any unique issues for your sector or industry as compared to other copyright stakeholders?

Generative AI raises new challenges for the publishing industry as a whole but the Christian publishing community faces a unique concern of its own: the creative genesis of work intended to convey spiritual truths or the application of those truths to the human condition. Publishers of all stripes will maneuver through the changing landscape of AI-generated material distribution with various priorities. AI-generated works may mimic the style or genre of popular authors, damaging the market for an author's books. The sheer amount of content capable of being developed in such a short time will test marketing and promotion teams' abilities to cut through increased clutter (although generative AI may assist in such a task). This proliferation of content exacerbates one of publishers' greatest challenges even in the pre-generative AI marketplace: discoverability. While AI technologies offer innovative marketing opportunities, the sheer quantity of published material available to consumers could proliferate at a rate never previously experienced. That Amazon instituted a policy limiting the number of AI-generated titles uploaded per day by Kindle Direct Publishing users highlights this issue.[4]

But the Christian publishing industry specifically faces challenges of its own in both its publication of the Bible and texts interpreting the Bible and commenting on other matters of Christian theology and living. Many Christians hold that the Bible itself is of divine origin or is at least inspired by the divine. Some also believe that other Christian authors receive divine prompting in authoring works other than the Bible. ECPA holds that while copyright ownership remains within the realm of human authorship[5], there are spiritual implications for the use of AI in the creation of content. Not that technology cannot be used as a tool to assist authors in creating original works or derivatives works, but that Christian publishers possess a responsibility to both allow for the spiritual influence of works in human authors and display transparency when certain materials originate by other means. Policies and industry standards requiring transparency and identification of AI-generated content allows the Christian publishing community to better steward the use of AI tools in an open and honest manner.

AI also presents Christian publishers with additional layer of risk to AI-generated works created in the style of a particular author. This risk applies across the publishing industry when considering the potential for AI-generated works to substitute for an author's past or future works. But for Christian publishers, the risk is not limited to this economic effect typically caused by copyright infringement. For a theologian-author, AI creates a serious reputational and theological risk. An AI-generated work containing non-sensical and even heretical misapplications of the Bible that appears to be written by a known human theologian may confuse readers. In this sense, generative AI risks creating works which proliferate confusion in the marketplace, damaging the reputation of authors and faith of readers. The negative consequences of such a scenario are not dissimilar from those which trademark law seeks to guard against by protecting a recognized source of particular goods or services.

8.1. In light of the Supreme Court's recent decisions in *Google v. Oracle America* and *Andy Warhol Foundation v. Goldsmith*, how should the "purpose and character" of the use of copyrighted works to train an AI model be evaluated? What is the relevant use to be

---

[4] See *Update on KDP Title Creation Limits*, Kindle Direct Publishing Forum, Sept. 18, 2023, https://www.kdpcommunity.com/s/article/Update-on-KDP-Title-Creation-Limits?language=en_US&forum=KDP%20Forum#top

[5] See *Penguin Books U.S.A. v. New Christian Church Full Endeavor*, 288 F. Supp. 2d 544 (S.D.N.Y. 2003)

analyzed? Do different stages of training, such as pre-training and fine-tuning, raise different considerations under the first fair use factor?

ECPA finds troubling the line of argumentation that the "ingestion" of copyrighted works for the purpose of training an AI model should be considered a "transformative use" under the first fair use factor.[6] AI proponents' assertion that electronically reproducing copyrighted works to train AI models is transformative from the purpose of the original use conveniently ignores the very reason for which the AI model is trained: to "learn" the use of language from its training datasets to a sufficient degree to allow it to predict and reproduce language in the same manner. Its very purpose is to be able to reproduce language like the data in its training datasets.

The U.S. Supreme Court in *Goldsmith*[7] focused on the specific use of the secondary work at issue (as opposed to the content's expression). Supporters of the AI industry may argue that the ingestion of copyrighted works for AI-training purposes seeks to understand the unprotectable elements of the work—a fundamentally different purpose than the purpose of the original, ingested work. But the purpose of AI-ingestion is not merely to train the AI model in use of language generally (in a text-based example) but to understand how *a particular author uses language in a particular work*. Once the AI model ingests sufficient data, it will be capable of mimicking the original work. It is not only the "unprotectable" elements of the original works that are then capable of being copied, but the protectable language choices of the authors of ingested data. By training an AI system to recognize the "style" of a writer in a particular work, the AI model is then able to respond to prompts to create works using the same turns of phrase and language choices, creating a work which would substitute for that of the original author. As Justice Sotomayor noted in *Goldsmith*, the issue with the first factor "relates to the problem of substitution."[8] Ingesting copyrighted works so that a machine is capable of copying, reproducing, or even creating derivative works of the original transgresses this prohibition on substitutions for the same purpose. To borrow the facts at issue in *Goldsmith*, an AI model trained on the works of photographer Goldsmith would be capable of responding to a prompt to create an image of Prince in Goldsmith's style in such a way as to copy the protectable elements of Goldsmith's photograph, just as Warhol did.

This capability distinguishes the "ingestion" of copyrighted works in AI models from the scanning performed by Google in *Authors Guild v. Google, Inc.*[9] While Google reproduced entire copyrighted works, its purpose was to create a searchable database that would only display snippets of the reproduced works. Google never sought to make entire works available or create works derived from those it scanned. The purpose of Google's secondary work could not substitute for the purpose of the original work, namely reading the text as a literary work. The purpose of large language model AI, however, is substitutionary in nature: to "learn" the ingested data so well that it can manipulate language to mimic a human author, making it capable of replicating copyrighted works already in existence or creating new derivative works.

---

[6] See Defendant's Motion to Dismiss, p. 2 at ¶¶ 8 – 11, *Kadrey v. Meta Platform's Inc.*, No. 3:23-cv-03417-VC (N.D. Cal. Sept. 18, 2023) ("Use of texts to train LLaMA to statistically model language and generate original expression is transformative by nature and quintessential fair use—much like Google's wholesale copying of books to create an internet search tool was found to be fair use in *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015)").

[7] *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 143 S. Ct. 1258 (2023)

[8] *Id.* at 1274.

[9] *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015)

8.5. Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured? Should the inquiry be whether the outputs of the AI system incorporating the model compete with a particular copyrighted work, the body of works of the same author, or the market for that general class of works?

The effect on the market for copyrighted works used to train AI models and the infringement risks of AI outputs may be best addressed by contractual and licensing agreements between the relevant parties. As described above in ECPA's response to no. 8.1, the risk of infringement in AI-generated outputs directly results from infringement in the unauthorized use of copyrighted works to create datasets and train AI models.

The economic effect on copyright works by any potential infringing AI-generated outputs may then be bargained by the parties best suited to determine such value. The parties may take into account AI-generated competing works relative to a particular copyrighted work, an author's body of works, or a general class of works.

No inquiry into the fourth factor would be complete without considering the established infrastructure and market for licensing that already exists within the publishing industry. Virtually every commercial publisher responds to permissions requests for various uses of copyrighted works under their exclusive license. AI proponents should not be permitted to hide behind the fourth fair use factor on grounds that no market yet exists for works used in training AI models or included in datasets. But allowing this narrow market to grow within the established broad licensing market for literary works serves all parties in fairly compensating rights holders and ensuring AI developers have access to quality data to produce reliable AI.

9. Should copyright owners have to affirmatively consent (opt in) to the use of their works for training materials, or should they be provided with the means to object (opt out)?

ECPA holds that an opt-in approach as part of a broader licensing scheme would be consistent with the rights of copyright and the existing licensing framework of the publishing industry. In few instances are the exclusive rights of 17 U.S.C. § 106 made available for use by third parties without the consent of the copyright owner. And even in those instances, such as compulsory mechanical licenses under 17 U.S.C. §115, the law dictates that the rights holder be compensated for the use. An opt-out approach for AI models is inappropriate because of the capabilities of generative AI. Once a copyrighted work is ingested, the AI will be capable of producing infringing or derivative works of the original work. For a rights holder unaware of the use of his or her work, opting-out of further "ingestion" may be a remedy too late for effect.

9.4. If an objection is not honored, what remedies should be available? Are existing remedies for infringement appropriate or should there be a separate cause of action?

U.S. copyright law's history of adapting to technological change supports a conclusion that current remedies are adequate to incentivize the respect of rights as AI technology develops, provided those remedies are made available to rights holders in the AI context. The fair use doctrine provides sufficient guardrails to ensure rights owners do not overstep in protecting their works. As long as the fair use doctrine is not expanded to permit new exceptions such as text-and-data mining or ingestion of

copyrighted works into generative AI models capable of creating competing AI-generated works, significant changes to the existing copyright framework would be premature. The publishing industry's established licensing marketplace and practices provide a foundation for a licensing scheme to allow their exclusive rights to be utilized in AI technologies.

15. In order to allow copyright owners to determine whether their works have been used, should developers of AI models be required to collect, retain, and disclose records regarding the materials used to train their models? Should creators of training datasets have a similar obligation?

Policy should require that both AI developers and training dataset creators maintain accurate records regarding: (i) the content within the training datasets ingested by an AI model; (ii) the relevant copyright information of each piece of data, including the identity of the copyright owner or exclusive licensee and whether the content was properly licensed; and (iii) the assessment used to interrogate the AI algorithm and determine whether it is sufficiently trained for public deployment.

As the adage of "garbage in, garbage out" meets a new technological application, training datasets without source records or disclosure requirements would make AI accountability and reasonable regulation difficult. Respecting the rights of rights holders, detecting underlying biases, and uncovering risks associated with use of an AI are all dependent on the provenance AI's training data. The European Union's AI Act already includes a disclosure requirement to make publicly available a summary of copyrighted works used to train an AI.[10] Record retention and disclosure not only protects the interests of copyright industries such as publishing but creates greater public trust in the reliability of AI after deployment.

16. What obligations, if any, should there be to notify copyright owners that their works have been used to train an AI model?

Disclosure is paramount to a legal and regulatory framework that respects the rights of copyright owners and offers needed transparency in the emerging generative AI field. A system of licensing in which AI developers obtain the necessary rights to include copyright works in training data would satisfy such a notification requirement.

18. Under copyright law, are there circumstances when a human using a generative AI system should be considered the "author" of material produced by the system? If so, what factors are relevant to that determination? For example, is selecting what material an AI model is trained on and/or providing an iterative series of text commands or prompts sufficient to claim authorship of the resulting output?

Based on ECPA's knowledge of existing AI technology, ECPA contends that a work produced by an AI algorithm or process absent sufficient involvement by a natural person is not protectable as a work of authorship under U.S. copyright law. However, ECPA also contends that for AI to serve as a meaningful and transformative tool for the publishing industry, the reality of AI involvement in content creation

---

[10] *Report on the Proposal for a Regulation of the European Parliament and of the Council on Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, at 211 (May 22, 2023), https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_EN.pdf (Plenary Sitting on May 22, 2023); *see also* Artificial Intelligence Act, Eur. Parl. Doc. P9 TA(2023)0236 (2023) (Texts Adopted on June 14, 2023)

must be recognized in a legally meaningful manner. The publishing industry would benefit from a policy recognizing the human ingenuity and creativity found in utilizing AI tools to produce a certain output.

Until now, the Copyright Office has rejected registering copyright for portions of works originating with a generative AI.[11] The basis of these decisions is that AI-generated work lacks sufficient predictability and control by a natural person to support a finding of human authorship. However, ECPA supports the position that copyright eligibility due to human authorship is a matter of intentionality and expression even through the use of new technologies. A photograph taken by camera was once also considered to be "merely 'a reproduction on paper of the exact features of some natural object or of some person' made by a machine."[12]

Is the artist who repeatedly enters prompts until the output matches his or her desired expression significantly different from the artist who continues to dab paint on the canvas until the image matches the painter's vision? Does the graphic artist utilizing an AI exhibit less control and predictability than the abstract expressionist artist throwing paint at a canvas? The artist selects the color and approximate target at which to aim but where each speck and spot lands could hardly be described as intentional. Does the digital artist choosing a color and providing specific inputs not at least exert the same level of control over the outcome as the artist seeking to be the next Jackson Pollock? Both have an approximation of what the output will be but neither controls the exact placement of each fleck of color. But each may take a brush or palette knife or an additional prompt to further change the work into what the artist envisioned.

To be sure, the recognition of sufficient human input into an AI to warrant copyright protection of the output will require development as a body of law. It will not be as clean as the status quo rejecting each portion of a work which is AI-generated. But copyright law is already equipped to manage determinations with factors whose importance varies with each case and whose decisions fall more on a spectrum than in one category or another. The fair use doctrine has developed in this manner for decades. Recognition of AI as a legitimate tool for protectable creation will allow publishers to utilize these tools without fear of jeopardizing exclusive rights in the resulting work.

Copyright supports the dissemination of ideas necessary to a democratic society. To allow authors and publishers to both further this purpose and make use of new, developing technologies, policies must promote clarity in considering human creativity and intentionality to create outputs of expression qualifying for copyright protection. A hardline policy excluding legal protection for any work inclusive of AI contributions obtained through intentional and specific human efforts will not benefit the publishing industry in the long-term.

21. Does the Copyright Clause in the U.S. Constitution permit copyright protection for AI-generated material? Would such protection "promote the progress of science and useful arts"? If so, how?

As described in response to no. 18, copyright protection is warranted for works evidencing sufficient human intentionality and creativity in the input such that output matches the author's intended expression in tangible form. In this instance, the natural person is the author more than the machine,

---

[11] See United States Copyright Office, *"Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence"* (March 16, 2023) at 16190-16194.

[12] *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884)

thus satisfying the Constitutional requirement. If the issue is one of control and predictability, fine-tuning repeatedly until the final expression is satisfactory demonstrates the author's ultimate control of the final work, even if each iteration leading up to the final expression may be subject to unpredictability. It is the return to the tool until the expression is correct that shows authorship, regardless of the tool.

Such a recognition of human authorship by AI-generated tools expands the purpose of copyright's promotion of "sciences and the useful arts" to an even wider array of Americans. Does a disabled artist who cannot hold a paintbrush lose the ability to own copyright in her expression because she must input hundreds of prompts instead of making micro brush strokes? Does the tool with which she creates her output—making repeated adjustments until it is just right—disqualify her work from legal protection? Where AI-generated work is the product of insignificant human intent and creativity, the status quo is correct. But where AI acts as a tool in the hands of an author, wielding it to create intended expression, AI follows a litany of prior technological advances in expanding the way in which humans express themselves.

22. Can AI-generated outputs implicate the exclusive rights of preexisting copyrighted works, such as the right of reproduction or the derivative work right? If so, in what circumstances?

AI-generated outputs may infringe upon the rights of copyright, including the exclusive rights of reproduction and to create derivative works. An AI-developer or creator of a training dataset which, without authorization or license of the copyright owner, includes copyright works within the data ingested by its model, necessarily does so with the intent that the resulting AI will be capable of producing outputs similar to the copyrighted work. This is yet another justification for a licensing framework that allows the affected parties to address the risks of infringement and corresponding economic damage.

Additionally, users of the AI model who intentionally create a substantially similar work via generative AI also violate the copyright owner's exclusive rights. By entering prompts inducing an AI to generate output like an existing copyrighted work or particular author, the user intentionally creates a substantially similar work. Intending an AI-created output to appear or sound like a copyrighted work is no less infringing because the tool used to create such a similar work was an AI and not a pen or brush. Once again, a pre-ingestion licensing framework would allow AI developers and rights holders to apportion risk and cost, and allow users to select generative Ais in which their own risk is known in advance.

Sincerely,

Jeff Crosby
President and CEO