

THEIATM

INSTITUTE

CYBERSECURITY & AI ETHICS THINK TANK
WASHINGTON, DC

October 17, 2023

Submitted via: <https://www.regulations.gov/commenton/COLC-2023-0006-0001>

Document Type: Notice

Document Citation: 88 FR 59942

Page: 59942-59949 (8 pages)

Agency/Docket Number: Docket No. 2023-6

Document Number: 2023-18624

U.S. Copyright Office

101 Independence Ave. S.E.

Washington, D.C. 20559-6000

To whom it may concern,

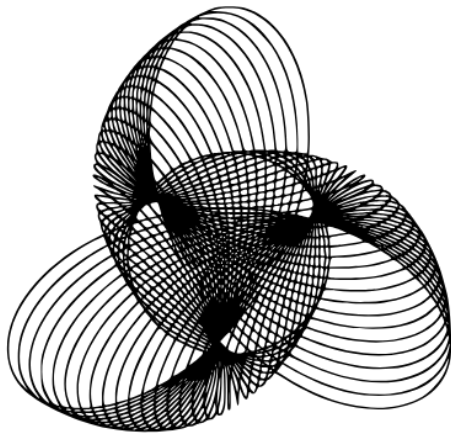
Thank you for the opportunity to comment on Artificial Intelligence and Copyright: a proposed study of the copyright law and policy issues raised by artificial intelligence (“AI”) systems.

From former Government Policy advisors to Silicon Valley innovators, our panel of experts are submitting a response, having analyzed current vulnerabilities, and suggesting robust solutions, in an effort to help guide public policy.

On behalf of the Theia InstituteTM, a Cybersecurity and AI Think Tank, we are submitting public comments in response to the US Patent and Trademark Office’s (USPTO) Notice of Proposed Rulemaking on Artificial Intelligence and Copyright to express our expert opinion and a strategic plan for addressing the unique issues of AI-generated material in the cybersecurity, software development, and risk management sectors—and the of the software industry in particular—that may be distinct from the issues faced by other copyright stakeholders.

Respectfully,

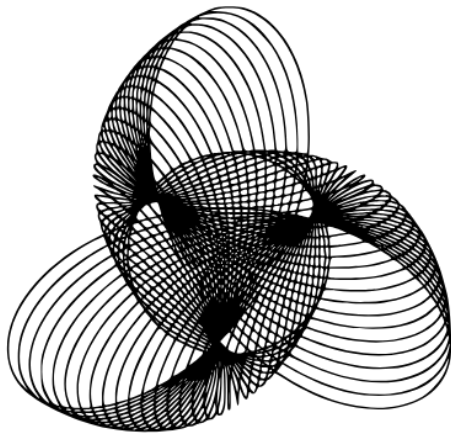
The Theia InstituteTM Editorial Board



Theia InstituteTM Response to USPTO Question

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?

Executive Summary.....	3
USPTO Focus Question.....	3
Industry-Related Software, Risk Management, and Governance Concerns.....	3
Key Recommendation.....	3
Analysis and Recommendations.....	3
Strategic Assessment.....	3
Recommended Action Items.....	4
Additional Context and Informational Materials.....	5
About the Theia InstituteTM	5
Mission Statement.....	5
Consensus-Based Solutions.....	5
Editorial Board.....	6
Press Information.....	6



Executive Summary

Theia InstituteTM Comment on Integrating AI with a Secure Software Development Life Cycle (SDLC) and Addressing Copyright Implications

USPTO Focus Question

USPTO Question: 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?

Industry-Related Software, Risk Management, and Governance Concerns

How can we safely integrate AI into our software development process to make it more secure without introducing quality issues, unauthorized code, and navigate the intricate landscape of copyright, licensing and distribution for AI-generated material?

Key Recommendation

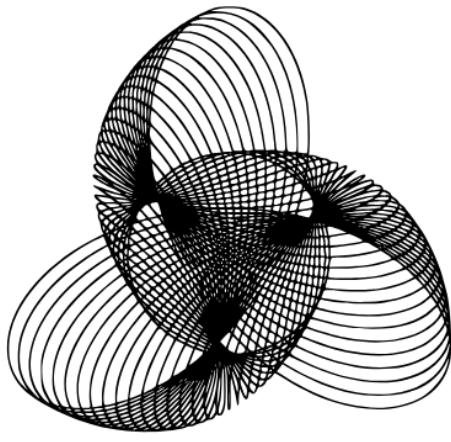
While AI can significantly streamline the software development process, particularly in tasks like unit testing and data generation, it introduces unique challenges in the realm of copyright. As AI-generated material becomes prevalent, distinguishing between human and AI-created content, and determining originality and ownership, becomes crucial. Embracing AI should be coupled with a thorough understanding of copyright implications, especially when distributing AI-generated or AI-assisted code and other informational content.

Analysis and Recommendations

Strategic Assessment

The rapid evolution of AI in software development brings forth not only technical challenges but also significant legal ambiguities, particularly concerning copyrights and distribution. AI-generated content, while not created by humans, may still have elements of originality, leading to questions about its ownership, licensing, and distribution rights. These concerns are particularly acute for software developers who may use AI tools to generate code or other content.

Because our current legal framework for software and content licensing is rooted in copyright law, the training of AI systems on licensed code without respect to that code's copyrights or licensing terms is



THEIA™

INSTITUTE

CYBERSECURITY & AI ETHICS THINK TANK
WASHINGTON, DC

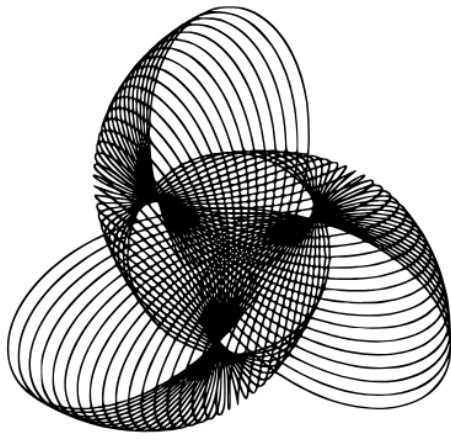
problematic. This is especially true of training data taken from source code or other data made publicly available under licenses but not placed into the public domain when that work is then collectively claimed as proprietary data and/or monetized by machine-learning (ML) or artificial intelligence (AI) vendors without permission from or remuneration to the copyright holders. Other forms of intellectual property (IP) such as patents also suffer from similar issues regarding training data and ownership of the resulting data sets.

For software developers, the introduction of AI-generated code or content can mean navigating a complex web of copyright laws and legal precedents that were not designed with AI in mind. There's a growing need for clear guidelines on how AI-generated material should be copyrighted, licensed, and distributed to ensure fairness, prevent misuse, and encourage innovation.

It is especially important that we ensure that companies building and operating ML and AI systems trained on open-source software (e.g. GPL-licensed works) or other permissively-licensed content (i.e. Creative Commons licenses) cannot then make the results of that training proprietary, place the results behind a paywall, or otherwise appropriate the original contents in ways that violate the original copyright holders' legal ownership and licensing terms. Such activities are not only inequitable on their face, but they also have the potential to create financial barriers to entry that favor the commercial interests of ML and AI companies without regard to the legal rights of copyright holders and licensees, and will be detrimental to the public interests served by making such works available under open-source and open-content licenses.

Recommended Action Items

1. Develop policies that clearly delineate the copyright status of AI-generated material within the software development life cycle (SDLC). This includes understanding the distinctions between human-created and AI-generated content.
2. Clearly articulate the extent to which AI content *currently* constitutes an original work based on general knowledge, fair use of copyrighted works, a derivative work, or other category based on its training and implementation. These definitions and our collective understanding will certainly evolve over time as the sciences of ML and AI evolve, and as our legal and social frameworks adapt to them over time.
3. Engage legal teams to stay updated on evolving copyright laws and their implications on AI-generated content.
4. Establish clear licensing agreements for AI-generated content, especially if such content is to be sold or distributed.
5. Educate developers and stakeholders about the nuances of copyright in the age of AI.
6. Prioritize transparency when using AI to generate content, ensuring that consumers and users are aware of the nature of the content they interact with.
7. Emphasize that while AI can generate content, human oversight remains essential, especially in the realms of quality assurance, ethical considerations, and responsible governance.



THEIATM
INSTITUTE
CYBERSECURITY & AI ETHICS THINK TANK
WASHINGTON, DC

Additional Context and Informational Materials

To further explore the challenges and solutions surrounding the integration of AI into the software development life cycle (SDLC) and the copyright concerns associated with doing so, the institute offers a range of materials. These include podcasts, recorded discussions, and expert opinions. Engaging with these resources will provide stakeholders with a more in-depth understanding of the issue. Please see [our company page on LinkedIn](#) for the full list of content related to each strategic plan, or join the public community discussions in the [Theia InstituteTM Community Discussions](#) group.

About the Theia InstituteTM

Mission Statement

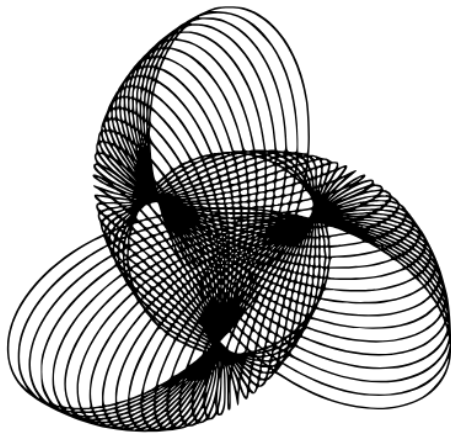
The growing intersection of data, privacy, and technology requires new ways of thinking about security and how it impacts both business and society. The institute's mission is to provide cutting-edge thought leadership that reframes traditional security leadership for the 21st century by addressing traditional cybersecurity, risk management, AI ethics and other related topics. In addition, the institute is also tasked with providing practical security solutions within a modern business context.

Consensus-Based Solutions

The institute provides consensus-based strategic analyses, recommendations, and action items. Each consensus is reached through discussion and debate among the panel's cross-functional members.¹

The members of the institute collectively believe that truly transformative leadership requires transparency and visibility into core processes. To promote those values and enrich materials the institute makes publicly available, panel discussions and deliberations are available in a variety of formats on multiple social media channels.

¹ The views presented by the Theia InstituteTM Editorial Board represent the collective opinion of the board as a body. The board's consensus does *not* represent an endorsement by any institution, employer, or panelist other than that of the board itself.



THEIATM
INSTITUTE
CYBERSECURITY & AI ETHICS THINK TANK
WASHINGTON, DC

Editorial Board

In alphabetical order, the current members of the institute's editorial board are:

- Billings, Q. Wade
- Desmond, Jim
- Engel, Barak
- Jacobs, Todd A.
- Kinon, Dan
- Palmer, Lisa
- Shannon, Doug

Each panelist brings a unique perspective to the topic. Learn more about their views and perspectives by following the panelists' discussions on social media.

Press Information

For further information about the institute's mission, panelists, or recommendations, please contact our public relations team at press@theia.institute. We are always glad to provide quotes, background, speakers, and other content to journalists and media outlets.