Unlocking Fair Use in the Generative AI Supply Chain: A Systematized Literature Review

Anonymous Authors[1]

Abstract

One of the demands made and accepted during the Writers' Guild of America (WGA) protests[1] included disregarding generative artificial intelligence (GenAI) outputs as source material while determining writing credits and compensation. This demand highlighting the importance of credit and compensation for creative content creators was also voiced by Mr. Ashley Irwin during his congressional testimony[2]. He reasoned that Art 1 section 8(8) of the US Constitution[3] granted rights to authors and inventors to promote Science and arts, thus calling for policies that would regulate AI companies to generate their model using the '3Cs: consent, credit and compensation' for creative workers (creatives).

GenAI companies like OpenAI[4], on the other hand, are pushing for training models as fair use based on basis of the 'transformative'ness of GenAI outputs, relying on Campbell v. Acuff-Rose Music[5] where the Supreme Court averred that "...the goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works."

The challenge for the judiciary and the policymakers lies in how they ideologically interpret these copyright law arguments, given that they both hold ground. On the one hand, the creatives' argument grounds itself in Lockean rhetoric of one's moral right to reap the benefits of one's labor. Conversely, OpenAI employs utilitarian rhetoric that views copyright as an incentivization system to enhance public welfare and advancement (Leaffer, 2019). In addition to this challenge of interpreting cogent arguments, economically, in 2021, the core copyright industries in the US added a value of $1.810 trillion, which was 7.76% of the U.S. GDP. Additionally, employment in the total copyright industries was more than 9.6 million creative workers, representing almost 5% of the total US employment, making balancing this thin line of authorship rights and advancing artistic and scientific progress in society even more difficult(Stoner & Dutra, 2022). GenAI, on the other hand, has transformative potential to generate realistic visual arts, conduct human-like conversations, write programs, and invent new medical treatments(Sun et al., 2024).

Through this systematization of existing user studies (Page et al., 2021), we try to uncover whether advocating for fair use in the GenAI context progresses the objectives of copyright law – to promote science and arts. We began by mapping stakeholders to different stages of the ML supply chain by relying on the GenAI systems supply chain conceptualized by Lee et al. (Lee et al., 2023) as a structuring framework. This systematic literature review seeks to synthesize existing HCI literature available in English for the last 10 years, from 01/01/2014 to 04/30/2024. Based on our search queries, inclusion, and exclusion criteria, the search yielded 5895 papers, forming our analysis unit. Records that were not full papers or duplicate records were removed to result in 855 papers.

Given that this paper is a scoping study for the systematic literature review for these 855 papers, in addition to the review goals and the research questions, this paper seeks to determine the value of conducting a full systematic review. Therefore, this paper restricts its scope and systematically reviews papers from the last 3 years, from 2022 to April 2024, which resulted in 338 papers. A more thorough manual elimination process using title and abstract analysis was conducted to ensure

[1]Art. 72 of Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement: https://www.wga.org/uploadedfiles/contracts/2023_mba_moa.pdf

[2]House holds hearing to examine the intersection of generative AI and copyright law — 05/17/23. Retrieved June 3, 2024 from https://www.youtube.com/watch?v=isTgXmzoaXc

[3]U.S. Constitution. Art. I, § 8(8): https://constitution.congress.gov/browse/essay/artI-S8-C8-1/ALDE_00013060/

[4]OpenAI. 2019. Comment by OpenAI Regarding Intellectual Property Protection for Artificial Intelligence Innovation. Department of Commerce, United States Patent and Trademark Office. Retrieved December 22, 2023 from https://www.uspto.gov/sites/default/files/documents/OpenAI_RFC-84-FR-58141.pdf

[5]Campbell v. Acuff-Rose Music, 510 U.S. 569 (1994)

that only the most relevant papers were reviewed. This round of elimination resulted in 56 papers selected papers. 31 papers were randomly selected for this systematic scoping literature review. After reviewing the selected papers, 10 were eliminated as they did not meet the inclusion/exclusion criteria mentioned in this section, resulting in 20 full papers forming the final dataset.

We then analyzed the final filtered papers using descriptive coding and in-vivo techniques (Saldana, 2021) in MaxQDA 24. We used affective methods such as in-vivo coding and value coding by assigning codes to attributes that interviewed stakeholders, in their interview responses, mention or specify as incentivized or are considered 'good.' The assigned codes were then analyzed using versus coding as an inspiration with the labor theory of copyright and utilitarian theory of copyright as dichotomous groups.

From the analyzed papers, it is evident that while all stakeholders except the end-users add value to the GenAI supply chain through their expertise, the interviewed ML practitioners distinguished between clients' needs and end-users needs, believing that meeting clients' needs and ensuring their involvement determined the efficacy of the model and their workflow and value to the projects. User Experience (UX) practitioners, on the other hand, were found to value their ability to cater to end-users needs by ensuring the accuracy, explainability, and flexibility of the model and its user interface (UI). On the other hand, interviewed creatives believed that their lived experience and vulnerability contributed to their expertise, adding value to their output and the messaging it propagated.

These findings and the plaguing computational concerns of GenAI models call into question OpenAI's claim that GenAI models generate 'transformative work.' Firstly, GenAI, in its present version, has raised significant apprehensions amongst the original creatives and content creators regarding the economic harms they face due to GenAI outputs. Secondly, it is evident from the extant literature that although GenAI has the potential to create transformative work, as of the submission of this paper, GenAI is still plagued with issues of AI mimicry, overfitting, and excess memorization. Therefore, this paper calls for the complementation of copyright law with computational solutions to better implement the letter and spirit of the law. Unless these problems are resolved computationally, GenAI will continue to create pattern-based content that is neither original nor novel, failing to validate its claim of 'transformative work' and hampering its eligibility for the fair use exception.

References

Leaffer, M. A. *Understanding copyright law*. Carolina Academic Press, 2019.

Lee, K., Cooper, A. F., and Grimmelmann, J. Talkin"bout ai generation: Copyright and the generative-ai supply chain. *arXiv preprint arXiv:2309.08133*, 2023.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, 372, 2021.

Saldana, J. *The Coding Manual for Qualitative Researchers*. SAGE Publications Ltd, 2021.

Stoner, R. and Dutra, J. Copyright industries in the us economy: 2022 report. *Available at SSRN 4453588*, 2022.

Sun, Y., Jang, E., Ma, F., and Wang, T. Generative ai in the wild: Prospects, challenges, and strategies. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2024.