

# Licensing of Text for Generative AI: Learnings from Non-AI Licensing Practices

Regan Smith\*

## INTRODUCTION

Licensing developments for generative Artificial Intelligence (“AI”) Large Language Models (“LLMs”) are moving fast, fueled by developer demand for access to, and the ability to use, increasing amounts of high-quality textual data.<sup>1</sup> With this constant demand for quality literary works come questions around how licensing practices can enable technological developments while preserving the contours of copyright law and sufficiently incentivizing human authorship of books, journalism, and other literary content for human readers and AI uses alike.

While some venture that licensing for generative AI purposes is “impossible,” many companies have negotiated partnerships with media publishers or publishing houses for generative AI uses.<sup>2</sup> Meanwhile, others query whether there is a need to build or adjust licensing systems to better facilitate licensing of textual content, whether through regulatory updates, increased use of collecting societies, or augmenting data management infrastructure.

Before declaring the status quo of marketplace licensing insufficient, it makes sense to take stock of where we have been, where we are, and where we might be going. This short piece hypothesizes that some current bumps in generative AI licensing stem from uncertainty in an emerging market, not inherent difficulties in licensing at scale for

---

\* SVP and General Counsel, News/Media Alliance. The views expressed are her own and not necessarily those of her employer. Thank you to Johannes Munter for his research assistance and review.

1. See Paul Sweeting, *Generative AI & Licensing: A Special Report*, VARIETY (Oct. 1, 2024), <https://variety.com/vip-special-reports/generative-ai-content-licensing-special-report-1236157051/> [<https://perma.cc/Z5B8-8VM9>] [<https://web.archive.org/web/20250223224314/https://variety.com/vip-special-reports/generative-ai-content-licensing-special-report-1236157051/>].

2. Dan Milmo, *‘Impossible’ To Create AI Tools Like ChatGPT Without Copyrighted Material*, *OpenAI Says*, GUARDIAN (Jan. 8, 2024), <https://www.theguardian.com/technology/2024/jan/08/ai-tools-chatgpt-copyrighted-material-openai> [<https://perma.cc/3T9F-XUEY>] [<https://web.archive.org/web/20250122113600/https://www.theguardian.com/technology/2024/jan/08/ai-tools-chatgpt-copyrighted-material-openai>].

© 2025 Regan Smith. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction, provided the original author and source are credited.

professionally published content. Given that generative AI is still in its nascency, content licensing is not close to a one-size-fits-all standard. The time is ripe for marketplace developments, and experimentation in private arrangements between rightsholders and users. The Article also provides a brief primer in copyright principles of licensing regulation and overviews guideposts for collective management of content, based on experiences outside of AI. While voluntary collective licensing can play a valuable role in the AI licensing market, these guideposts may assist authors and other licensees as they consider whether, with whom, and on what terms to affiliate with a licensing intermediary.

## I. LICENSING FOR LLMs IS A RAPIDLY EMERGING SPACE

Despite over three dozen copyright litigations currently testing the boundaries of fair use for AI uses, significant licensing activity is already taking place. In the past two years, there have been reports of hundreds of deals struck between AI developers and rightsholders.<sup>3</sup> In addition, several startups have launched services to support the

---

3. See, e.g., Trishla Ostwal, *Dow Jones Wins AI Licensing Deals with More Than 4,000 News Outlets Like the AP*, ADWEEK (Nov. 12, 2024), <https://www.adweek.com/media/dow-jones-lands-more-than-4000-ai-licensing-deals-with-news-outlets-like-the-ap/> [https://perma.cc/DRG3-48XU] [https://web.archive.org/web/20250223223554/https://www.adweek.com/media/dow-jones-lands-more-than-4000-ai-licensing-deals-with-news-outlets-like-the-ap/]; Ella Creamer, *HarperCollins To Allow Tech Firms To Use Its Books To Train AI Models*, GUARDIAN (Nov. 19, 2024), <https://www.theguardian.com/books/2024/nov/19/harpercollins-tech-firms-books-train-ai-models-nonfiction-artificial-intelligence> [https://perma.cc/F8DY-9ZXL] [https://web.archive.org/web/20250223223801/https://www.theguardian.com/books/2024/nov/19/harpercollins-tech-firms-books-train-ai-models-nonfiction-artificial-intelligence]; Matilda Battersby, *Wiley Confirms It Is Seeking Further Generative AI Partnerships as Its Quarterly Revenue Rises To £296m*, BOOKSELLER (Sept. 6, 2024), <https://www.thebookseller.com/news/wiley-confirms-it-is-seeking-further-generative-ai-partnerships-as-its-quarterly-revenue-rises-to-296m> [https://perma.cc/87LJ-G7CH] [https://web.archive.org/web/20250121172941/https://www.thebookseller.com/news/wiley-confirms-it-is-seeking-further-generative-ai-partnerships-as-its-quarterly-revenue-rises-to-296m]; Kyle Wiggers, *Perplexity Expands Its Publisher Program*, TECHCRUNCH (Dec. 5, 2024), <https://techcrunch.com/2024/12/05/perplexity-expands-its-publisher-program/> [https://perma.cc/8LTZ-S9SQ] [https://web.archive.org/web/20250121173549/https://techcrunch.com/2024/12/05/perplexity-expands-its-publisher-program/]; Emma Roth, *Microsoft Is the Mystery AI Company Licensing HarperCollins Books, Says Bloomberg*, VERGE (Nov. 19, 2024), <https://www.theverge.com/2024/11/19/24300893/microsoft-ai-training-deal-harpercollins-report> [https://perma.cc/W5ZT-8END] [https://web.archive.org/web/20250121174437/https://www.theverge.com/2024/11/19/24300893/microsoft-ai-training-deal-harpercollins-report]; Sara Fischer, *Scoop: Meta Strikes Multi-Year AI Deal with Reuters*, AXIOS (Oct. 25, 2024), <https://www.axios.com/2024/10/25/meta-reuters-ai-news-facebook-instagram> [https://web.archive.org/web/20250218041601/https://www.axios.com/2024/10/25/meta-reuters-ai-news-facebook-instagram]; Guardian Staff and Agencies, *OpenAI Signs Multi-Year Content Partnership with Condé Nast*, GUARDIAN (Aug. 20, 2024), <https://www.theguardian.com/technology/article/2024/aug/20/conde-nast-open-ai-deal> [https://perma.cc/8XUN-CGG4] [https://web.archive.org/web/20250121175126/https://www.theguardian.com/technology/article/2024/aug/20/conde-nast-open-ai-deal]; Sara Fischer, *Exclusive: The Atlantic, Vox Media Ink Licensing, Product Deals with OpenAI*, AXIOS (May 29, 2024), <https://www.axios.com/2024/05/29/atlantic-vox-media-openai-licensing-deal> [https://web.archive.org/web/20250129002733/https://www.axios.com/2024/05/29/atlantic-vox-media-

infrastructure needed for different aspects of these transactions.<sup>4</sup> Considering these developments, one would be hard-pressed to say that marketplace licensing is impossible.

Indeed, marketplace arrangements are taking off even amidst rapid experimentation around how generative AI offerings may use publisher content. Generally, content licensing requires reaching a shared understanding regarding what content is desired and how it will be used so that the parties can agree on fair terms and conditions.<sup>5</sup> For generative AI, licensing discussions raise product questions regarding how the developer intends to use its model or the content it is interested in accessing and licensing. Many publishers operate robust licensing and partnership arms, regularly handling requests for reprints, excerpts, derivative works, etc., and identifying the expected uses of their content by an AI developer—including outputs or deployments of LLMs—can help compare those uses with other, existing revenue streams. Meanwhile, facing an emerging and competitive field, AI developers may be reluctant to commit to strict limits on the use of copyrighted materials, and some may attempt to hedge around legal questions by purporting to license only the “access to” works. A

openai-licensing-deal]; Kyle Orland, *Reddit Cashes In on AI Gold Rush with \$203M in LLM Training License Fees*, ARS TECHNICA (Feb. 23, 2024), <https://arstechnica.com/ai/2024/02/reddit-has-already-booked-203m-in-revenue-licensing-data-for-ai-training/> [<https://web.archive.org/web/20250121175555/https://arstechnica.com/ai/2024/02/reddit-has-already-booked-203m-in-revenue-licensing-data-for-ai-training/>]; Angela Cullen & Jackie Davalos, *OpenAI To Pay Axel Springer Tens of Millions To Use News Content*, BLOOMBERG (Dec. 13, 2023), <https://www.bloomberg.com/news/articles/2023-12-13/openai-axel-springer-ink-deal-to-use-news-content-in-chatgpt> [<https://perma.cc/CT43-FE8D>] [<https://web.archive.org/web/20250121175820/https://www.bloomberg.com/news/articles/2023-12-13/openai-axel-springer-ink-deal-to-use-news-content-in-chatgpt>].

4. See, e.g., CREATED BY HUMANS, <https://www.createdbyhumans.ai/> [<https://perma.cc/JHU4-PHNM>] [<https://web.archive.org/web/20250130030651/https://www.createdbyhumans.ai/>] (last visited Jan. 21, 2025) (“AI Rights licensing platform for books”); TOLLBIT, <https://tollbit.com/> [<https://perma.cc/H4EM-R64J>] [<https://web.archive.org/web/20250122001551/https://tollbit.com/>] (last visited Jan. 21, 2025) (showing a service that offers analytics, bot payroll, and content licensing solutions); DAPPIER, <https://dappier.com/> [<https://perma.cc/C2WK-A8N4>] [<https://web.archive.org/web/20250130143904/https://dappier.com/>] (last visited Jan. 21, 2025) (allowing publishers to create and monetize AI agents through a marketplace); PRORATA.AI, <https://www.prorata.ai/> [<https://web.archive.org/web/20250125095155/https://www.prorata.ai/>] (last visited Jan. 21, 2025) (providing an AI search engine based exclusively on licensed data, using an algorithmic solution to distribute payments based on usage); HUMAN NATIVE AI, <https://www.humannative.ai/> [<https://perma.cc/7ESE-BJ3Y>] [<https://web.archive.org/web/20250122062925/https://www.humannative.ai/>] (last visited Jan. 21, 2025) (providing a two-sided marketplace for publishers and AI developers); SCALEPOST, <https://www.scalepost.ai/> [<https://perma.cc/3SWT-2R64>] (last visited Jan. 21, 2025) (providing a marketplace for developers and publishers); CALLIOPE NETWORKS, <https://calliopenetworks.ai/> [<https://perma.cc/VUU9-JD6C>] [<https://web.archive.org/web/20250130030315/https://calliopenetworks.ai/>] (last visited Jan. 21, 2025) (aggregating and licensing textual and audiovisual content to AI developers on a revenue-share basis).

5. See *What Is Content Licensing?: The Ultimate Guide*, N.Y. TIMES, <https://nytlicensing.com/latest/our-brand/what-is-content-licensing/> [<https://perma.cc/7HYW-236F>] [<https://web.archive.org/web/20250121183053/https://nytlicensing.com/latest/our-brand/what-is-content-licensing/>] (last visited Jan. 6, 2025) (describing discussions in licensing agreements, “such as which pieces of content, type of usage, platforms, length of time, etc.”).

recent skim of [openai.com](https://openai.com) advertises multiple uses for ChatGPT, ranging from “reducing health insurance costs and improving patient care” by insurers to ChatGPT search, which “blends the benefits of a natural language interface with the value of up-to-date sports scores, news, stock quotes, and more.”<sup>6</sup> It is therefore perhaps not a coincidence that early partnership announcements spoke in general terms, such as using publisher content to “develop . . . generative AI models,”<sup>7</sup> or to “share access to select news content and technology.”<sup>8</sup>

## II. REGULATORY PRINCIPLES URGE CAUTION IN USURPING PRIVATE ARRANGEMENTS BETWEEN RIGHTSHOLDERS AND USERS

Discussions on licensing content at “scale” have been ongoing since well before the passage of the 1976 Copyright Act, which was the same time as the Copyright Clearance Center opened its doors to facilitate large scale licensing for reprographic works. Almost fifty years later, there are several models for aggregated licensing of copyrighted works across different industries for various uses, including recorded music, music publishing, fine and graphic arts, photography, news media, journals, books, and motion pictures. Aggregated licensing takes a variety of forms, including through licensing or administration agents, collective management organizations (e.g., performance rights organizations and other member-based licensing organizations), syndicates, co-operatives, special purpose joint ventures, click-wrap licenses (both inbound and outbound), extended collective licenses, and, in the U.S., a discrete set of statutory licenses for limited audio and video uses.

Most are voluntary participation models that naturally emerged in response to market dynamics. The United States presumption is for marketplace licensing, with the government stepping in to override private contracting only in very rare conditions. As former Register of Copyrights Marybeth Peters explained in 2004, a compulsory license is a “last resort” mechanism,<sup>9</sup> because by its nature, it “limits an author’s bargaining power. It deprives the author of determining with whom and on what terms

---

6. *Oscar*, OPENAI, <https://openai.com/index/oscar/> [<https://web.archive.org/web/20250213133338/https://openai.com/index/oscar/>] (last visited Jan. 6, 2025); *Introducing ChatGPT Search*, OPENAI (Oct. 31, 2024), <https://openai.com/index/introducing-chatgpt-search/> [<https://web.archive.org/web/20250128224727/https://openai.com/index/introducing-chatgpt-search/>] (last visited Jan. 6, 2025).

7. Rick Merritt, *Moving Pictures: NVIDIA, Getty Images Collaborate on Generative AI*, NVIDIA BLOG (Mar. 21, 2023), <https://blogs.nvidia.com/blog/generative-ai-getty-images/> [<https://perma.cc/N4VC-KMG6>] [<https://web.archive.org/web/20250121235140/https://blogs.nvidia.com/blog/generative-ai-getty-images/>].

8. Sara Fischer, *Exclusive: AP Strikes News-Sharing and Tech Deal with OpenAI*, AXIOS (July 13, 2023), <https://www.axios.com/2023/07/13/ap-openai-news-sharing-tech-deal> [<https://web.archive.org/web/20250118003213/https://www.axios.com/2023/07/13/ap-openai-news-sharing-tech-deal>].

9. Section 115 of the Copyright Act: *In Need of an Update?: Hearing Before the Subcomm. on Courts, the Internet, and Intellectual Property of the H. Comm. on the Judiciary*, 108th Cong. 27 (2004) (statement of Marybeth Peters, Register of Copyrights, U.S. Copyright Office).

he wishes to do business.”<sup>10</sup> For that reason, she explained that (at that time) the United States was one of only two countries with a compulsory license regulating musical works.<sup>11</sup> In 2005, Register Peters elaborated further:

The Copyright Office has long taken the position that statutory licenses should be enacted only in exceptional cases, when the marketplace is incapable of working. One could argue that it is difficult to say that the marketplace is incapable of working . . . when the marketplace has never been given a chance to succeed.<sup>12</sup>

And, as the Copyright Office more recently explained when considering the now-discontinued cable satellite license, the presence of a regulated license can depress the development of business models that would otherwise emerge in the marketplace and add costs to administer a compulsory license upon the licensor and licensees.<sup>13</sup>

For similar reasons, from a copyright legal perspective, the marketplace appears fairly well-suited to address the different copying and uses for generative AI development and offerings. While generative AI’s mass taking of copyrighted content has prompted some questions regarding whether licensing regulation could be helpful,<sup>14</sup> it remains early days. Imposing a government-mandated structure could ossify licensing parties, terms, and conditions (ranging from permitted services to remuneration) before the market can mature, meanwhile imposing intermediary administrative costs. In addition, as others noted,<sup>15</sup> many purported concerns regarding licensing solutions specifically relate to user-generated content, such as social media threads, user-generated uploads, consumer reviews, or blog posts, rather than professionally produced content by publishers. While it is important to be able to address those copyright interests, this short piece is focused primarily on publisher-produced, first-party content. Without diving into separate considerations for other content, it is safe to say that it would be unprecedented to require AI licensing to be a “one stop shop” or to allow concerns for licensing other types of content to override the rights of professional publishers of journalism and other literary works, who rely upon copyright to invest in the creation of new works.

10. *Id.* at 20.

11. *Id.*

12. *Music Licensing Reform: Hearing Before the Subcomm. on Intellectual Property of the Comm. on the Judiciary*, 109th Cong. 13 (2005) (statement of Marybeth Peters, Register of Copyrights, U.S. Copyright Office).

13. See U.S. COPYRIGHT OFF., U.S. COPYRIGHT OFFICE ANALYSIS AND RECOMMENDATIONS REGARDING THE SECTION 119 COMPULSORY LICENSE 4–5 (June 3, 2019), <https://copyright.gov/laws/hearings/views-concerning-section-119-compulsory-license.pdf> [<https://perma.cc/3JWC-FGLR>] [<https://web.archive.org/web/20250124222807/https://copyright.gov/laws/hearings/views-concerning-section-119-compulsory-license.pdf>].

14. Notice of Inquiry and Request for Comments on Artificial Intelligence and Copyright, 88 Fed. Reg. 59,942, 59,946 (Aug. 30, 2023) (“Are there any statutory or regulatory approaches that have been adopted or are under consideration in other countries that relate to copyright and AI that should be considered or avoided in the United States?”).

15. See Matthew Stratton, *Market-Based Licensing for Publishers’ Works Is Feasible. Big Tech Agrees*, 48 COLUM. J.L. & ARTS (forthcoming 2025).

### III. COLLECTIVE MANAGEMENT ORGANIZATIONS (CMOS) OFFER A WAY TO AGGREGATE LICENSES, AND BOTH LICENSORS AND LICENSEES CAN CONSIDER ESTABLISHED CRITERIA WHEN EVALUATING CMO ROLES IN AI LICENSING

While not the only option to facilitate efficient, aggregated rights clearance, collective management is one solution for licensing mass quantities of copyrighted works for generative AI uses. With collective management, rightsholders are allowed—sometimes required—to administer their rights through a collective management organization, or CMO.<sup>16</sup> As the World Intellectual Property Organization (“WIPO”) explains it, CMOs can assist the interest of both parties in cases where it is not practical to seek specific permission from each rightsholder for a use.<sup>17</sup> This section briefly outlines different types of CMOs and benchmarks for parties to consider when evaluating use of a given intermediary to facilitate licensing for generative AI uses.

Complementary to direct licensing, voluntary collective management can be a valuable option, where authors and other rightsholders choose whether or not to affiliate with a CMO, or even choose among CMOs, such as when choosing to affiliate with a performing rights organizations in the U.S.<sup>18</sup> As the Copyright Office explained, “[v]oluntary collective licensing does not require legislation.”<sup>19</sup> By contrast, mandatory or compulsory collective management can operate as a limitation on rights interests. Rightsholders objecting to license terms, or how a CMO administers the license (such as its operations, royalty distribution, or governance issues) may find recourse avenues limited to filing a complaint with the collective or a supervising government agency.

Extended collective licensing (“ECL”) is sometimes suggested as a hybrid approach, where a set of copyrighted works (including out-of-commerce and so-called orphan works) are presumed to be included in a license unless the rightsholders controlling those works opt out of the license. Because most CMOs operate with rightsholders as members, and act to represent the interests of their members, an ECL mandate allows the CMO to “extend” its representation to include nonmember works. U.S. copyright law does not include ECL, and when the Copyright Office previously studied the

---

16. See *Collective Management of Copyright and Related Rights*, WORLD INTEL. PROP. ORG., <https://www.wipo.int/en/web/copyright/collective-management> [https://perma.cc/99CR-8WU9] [https://web.archive.org/web/20250124225619/https://www.wipo.int/en/web/copyright/collective-management] (last visited Jan. 24, 2025).

17. *Id.*

18. The Copyright Clearance Center (“CCC”) is an example of voluntary collective licensing in the U.S. See *About CCC*, COPYRIGHT CLEARANCE CTR., <https://www.copyright.com/company-about/> [https://perma.cc/2PFS-6YUS] [https://web.archive.org/web/20250124225844/https://www.copyright.com/company-about/] (last visited Jan. 24, 2025); U.S. COPYRIGHT OFF., LEGAL ISSUES IN MASS DIGITIZATION: A PRELIMINARY ANALYSIS AND DISCUSSION DOCUMENT 32–34 (Oct. 31, 2011), [https://www.copyright.gov/docs/massdigitization/USCOMassDigitization\\_October2011.pdf](https://www.copyright.gov/docs/massdigitization/USCOMassDigitization_October2011.pdf) [https://perma.cc/NX5C-9ZG5] [https://web.archive.org/web/20250124230020/https://www.copyright.gov/docs/massdigitization/USCO MassDigitization\_October2011.pdf].

19. U.S. COPYRIGHT OFF., *supra* note 18, at 33.

feasibility of a pilot program to facilitate certain mass digitization projects, it found a notable lack of stakeholder support and consensus.<sup>20</sup> Meanwhile, some European nations have limited ECL systems. The European Union Copyright Directive Article 12 provides that, where a CMO is operating within its mandates,<sup>21</sup> and it is “typically onerous and impractical” to obtain individual authorizations from rightsholders,” a state may choose to extend that license to non-member rightsholders under certain conditions, including by requiring the CMO to be “sufficiently representative” of the relevant rightsholders, by guaranteeing “equal treatment” to all rightsholders, allowing rightsholders to “easily and effectively” opt out and implementing “appropriate publicity measures” to inform rightsholders of the extended nature of the CMO.<sup>22</sup>

While collective licensing can provide efficiency to the licensing parties, the addition of a CMO structure inherently adds a layer of operational complexity. CMOs bring a separate set of management issues to be considered to ensure the licensing agent is performing its desired role.<sup>23</sup> Further, by consolidating negotiations into a single entity, which must adequately represent a variety of members, there is less flexibility than with individual negotiations, which may be able to be more experimental. Over decades, these issues have been considered by various governments and nongovernmental organizations—such as the International Confederation of Societies of Authors and Composers (“CISAC”), a global association of 227 authors societies—and have been incorporated into various CMO governance documents. A few core themes may be useful to consider in connection with generative AI licensing, to evaluate a CMO against direct licenses or other aggregated licensing structure alternatives.

### A. DISTRIBUTION

Any organization collecting and distributing royalties must have a strong ability to distribute those royalties fairly, transparently, and accurately.<sup>24</sup> CISAC’s guidance provides that “[a CMO] shall . . . base its distributions on actual usage of Works or, if

---

20. See Letter from Karyn A. Temple, Acting Register of Copyrights and Director, U.S. Copyright Off., to Charles E. Grassley, Chairman, S. Judiciary Comm., and Dianne Feinstein, Sen., S. Judiciary Comm. (Sept. 29, 2017), <https://www.copyright.gov/policy/massdigitization/senate-letter.pdf> [<https://perma.cc/NV3X-CYY8>] [<https://web.archive.org/web/20250124230500/https://www.copyright.gov/policy/massdigitization/senate-letter.pdf>].

21. Although uncommon in the U.S., many nations may designate a CMO with a mandate to license a particular set of rights.

22. Directive 2019/790, of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/6/EC and 2001/29/EC, art. 12, 2019 O.J. (L 130).

23. Cf. WORLD INTELL. PROP. ORG., *supra* note 16.

24. See Directive 2014/26, of the European Parliament and of the Council of 26 February 2014 on Collective Management of Copyright and Related Rights and Multi-Territorial Licensing of Rights in Musical Works for Online Use in the Internal Market, art. 13(1), 2014 O.J. (L 84) (“[A] collective management organisation [shall] regularly, diligently and accurately distribute[] and pay[] amounts due to rightsholders in accordance with the general policy on distribution referred to in point (a) of Article 8(5).”).

not practicable, on the basis of a statistically valid sample of actual usage of Works or, in the case of reprography, on potential usage,”<sup>25</sup> and the International Federation of Reproduction Rights Organisations (“IFRRO”)’s Code of Conduct echoes these principles by asking RROs to “distribute remuneration received to rightsholders; efficiently and expeditiously; approximating actual use as far as possible . . .”<sup>26</sup> In some sectors, CMOs emerged to administer licenses where it is impractical to count the actual uses of the copyrighted work (such as live music performances). In other cases, including the U.S. compulsory license for digital music streaming, usage reported by licensees allows the collective to make distributions based on the actual use of copyrighted works.<sup>27</sup> For born-digital uses—including generative AI where training sets already copiously document which works were ingested for training, and technical tools are developing to track model calls to content used to “augment” LLM responses<sup>28</sup>—the starting presumption should be that distribution policies are based on, or closely approximate, actual usage.

---

25. See CONFÉDÉRATION INTERNATIONALE DES SOCIÉTÉS D’AUTEURS ET COMPOSITEURS, PROFESSIONAL RULES FOR DRAMATIC, LITERARY AND AUDIOVISUAL SOCIETIES ¶ 18 (2019) <https://members.cisac.org/CisacPortal/cisacDownloadFileSearch.do?docId=34448>

[<https://perma.cc/8M34-92WX>]

[<https://web.archive.org/web/20250124232932/https://members.cisac.org/CisacPortal/cisacDownloadFileSearch.do?docId=34448>].

26. *Code of Conduct*, INT’L FED’N REPROD. RTS. ORGS., § 5.1, <https://ifrr.org/page/code-of-conduct/> [<https://ifrr.org/page/code-of-conduct/>] [<https://perma.cc/UU2U-EPNG>] [<https://web.archive.org/web/20250124193251/https://ifrr.org/page/code-of-conduct/>] (last visited Jan. 24, 2025) [hereinafter, IFRRO CODE OF CONDUCT].

27. See, e.g., 37 C.F.R. § 210.29 (2025); INT’L FED’N OF THE PHONOGRAPHIC INDUS., CODE OF CONDUCT FOR MUSIC INDUSTRY MUSIC LICENSING COMPANIES, § 5(2) (May 2021) [hereinafter, IFPI], <https://www.ifpi.org/wp-content/uploads/2021/06/MLCs-Code-Of-Conduct-May-2021.pdf> [<https://perma.cc/CC3T-XELF>] [<https://web.archive.org/web/20240315155444/http://www.ifpi.org/wp-content/uploads/2021/06/MLCs-Code-Of-Conduct-May-2021.pdf>].

28. See, e.g., NEWS/MEDIA ALL., WHITE PAPER: HOW THE PERVASIVE COPYING OF EXPRESSIVE WORKS TO TRAIN AND FUEL GENERATIVE ARTIFICIAL INTELLIGENCE SYSTEMS IS COPYRIGHT INFRINGEMENT AND NOT A FAIR USE (2023), <https://www.newsmediaalliance.org/wp-content/uploads/2023/10/AI-White-Paper-with-Technical-Analysis.pdf> [<https://perma.cc/BX7K-F46G>] [<https://web.archive.org/web/20250309195128/https://www.newsmediaalliance.org/wp-content/uploads/2023/10/AI-White-Paper-with-Technical-Analysis.pdf>] (analyzing datasets used for AI training); GEORGE WUKOSON & JOEY FORTUNA, THE PREDOMINANT USE OF HIGH-AUTHORITY COMMERCIAL WEB PUBLISHER CONTENT TO TRAIN LEADING LLMs (2024), <https://www.ziffdavis.com/wp-content/uploads/2024/11/The-Predominant-Use-of-High-Authority-Commercial-Web-Publisher-Content-to-Train-Leading-LLMs.pdf> [<https://perma.cc/KC4Y-DRV7>] (reviewing publicly available datasets); *TollBit for RAG*, TOLLBIT, <https://tollbit.com/ai/rag/> [<https://perma.cc/QS36-MM49>] [<https://web.archive.org/web/20250124200834/https://tollbit.com/ai/rag/>] (last visited Jan. 24, 2025) (“TollBit provides your business operations teams ledgers to track incoming ad revenue, and automate distributions to content and data partners.”); *ProRata Announces Gist.ai, New AI Search Engine Based Entirely on High-Quality Licensed Content*, BUS. WIRE (Dec. 9, 2024), <https://www.businesswire.com/news/home/20241209071998/en/ProRata-Announces-Gist.ai-New-AI-Search-Engine-Based-Entirely-on-High-Quality-Licensed-Content> [<https://web.archive.org/web/20250124201216/https://www.businesswire.com/news/home/20241209071998/en/ProRata-Announces-Gist.ai-New-AI-Search-Engine-Based-Entirely-on-High-Quality-Licensed-Content>] (“ProRata employs a proprietary algorithmic process to score and determine multi-source content



## B. REPRESENTATION DUTY

Most CMOs operate on a membership basis, representing authors, publishers, or other rightsholders who have affiliated with them. It is not uncommon for national or societal codes of conduct to require CMOs to act in the interest of their rightsholder members, and on a nondiscriminatory basis.<sup>29</sup>

## C. ADMINISTRATIVE COSTS.

As new businesses emerge to help administer generative AI licensing at scale, it is important that entities operate efficiently and avoid layering unnecessary costs between the AI developers and creative rightsholders. As WIPO's *Good Practice Toolkit for CMOs* explains, "it should be a key objective for a CMO to provide high quality rights management services at the lowest possible cost, thus maximizing the Distributions to Rightsholders."<sup>30</sup> This principle is echoed by others, including the Societies' Council for the Collective Management of Performers' Rights ("SCAPR"),<sup>31</sup> the association for performers' CMOs, and the IFRRO Code of Conduct.<sup>32</sup> Even CISAC, which navigates a broad membership across many nations and uses, has set a maximum floor for

---

attribution, enabling content owners to receive credit and compensation for their material on a per-use basis.").

29. See, e.g., WORLD INTEL. PROP. ORG., WIPO GOOD PRACTICE TOOLKIT FOR COLLECTIVE MANAGEMENT ORGANIZATIONS (THE TOOLKIT): A BRIDGE BETWEEN RIGHTHOLDERS AND USERS 44 (2025) [hereinafter, WIPO TOOLKIT] ("CMOs shall accept rightsholders as members if they fulfil the membership requirements which shall be based on objective, transparent and non-discriminatory criteria.") (quoting Code of Economic Law, Book XI, Title 5) (Belg.)); CODE OF CONDUCT FOR COPYRIGHT COLLECTING SOCIETIES § 2.2(a) at 4 (2022) (Austl.), <https://www.copyrightcodeofconduct.org.au/code> [<https://perma.cc/TR6R-PZ8W>] [[https://web.archive.org/web/20250309200733/https://static1.squarespace.com/static/5c454daba2772c47f0dcd6ea/t/6285be6522bb0b2d9bfc7b29/1652932200304/Code\\_of\\_Conduct\\_May\\_2022.pdf](https://web.archive.org/web/20250309200733/https://static1.squarespace.com/static/5c454daba2772c47f0dcd6ea/t/6285be6522bb0b2d9bfc7b29/1652932200304/Code_of_Conduct_May_2022.pdf)] ("The membership of a Collecting Society will be open to all eligible creators of copyright material, and to anyone who owns or controls copyright material . . . in accordance with the Constitution of the Collecting Society."); IFPI, *supra* note 27, at 1 ("MLCs shall act in the best interests of all the right holders they represent, whether directly or via agreements with other MLCs. MLCs shall offer their services and conduct their operations in a fair, effective and non-discriminatory manner and in compliance with the applicable legislation.").

30. WIPO TOOLKIT, *supra* note 29, § 6.4.1. See also Directive 2014/26, *supra* note 24, art. 12(2) ("Deductions shall be reasonable in relation to the services provided by the collective management organisation to rightsholders . . ."); *id.* art. 12(3) ("Management fees shall not exceed the justified and documented costs incurred by the collective management organisation in managing copyright and related rights.").

31. THE SOCIETIES' COUNCIL FOR THE COLLECTIVE MGMT. OF PERFORMERS' RTS., CODE OF CONDUCT § 5 (May 2024) <https://www.scapr.org/file/2024/09/SCAPR-Code-of-Conduct-approved-May-2024.pdf> [<https://perma.cc/9HHA-F4S8>] [<https://web.archive.org/web/20250309201838/https://www.scapr.org/file/2024/09/SCAPR-Code-of-Conduct-approved-May-2024.pdf>] ("The performers shall pay only the costs which are objectively necessary for the protection of their rights and interests and for the effective management of their rights.").

32. RROs must inform rightsholders of administrative deductions, which must be "proper and reasonable and in proportion to their actual efforts and services delivered" and administration must be designed "so as to minimize total administrative costs that are being deducted." IFFRO CODE OF CONDUCT, *supra* note 26, §§ 1.1.7, 5.3.

administrative costs of 30%; according to a questionnaire of the Asia-Pacific Economic Cooperation (“APEC”), CMOs reported “significant variation . . . in management fee[s], which ranged from 5% to 30%.”<sup>33</sup>

In the relatively well developed U.S. music market, administrative rates are much lower than 30%, particularly for digital uses where licensees transfer data to help the CMO administer the license and distribute royalties.<sup>34</sup> For example, in 2023 the Mechanical Licensing Collective (“MLC”), which administers streaming mechanical rights, reported an administrative fee of 3.07%;<sup>35</sup> SoundExchange, which administers certain sound recording performance licenses, reported a range of 4–6% in administrative rates;<sup>36</sup> and the American Society of Composers, Authors and Publishers (“ASCAP”) takes “10% for operating expenses and pay[s] the rest[] as royalties.”<sup>37</sup> Finally, while most CMOs deduct an administration fee from royalties otherwise payable to rightsholders, the MLC is unique in that the *licensees* are obligated by statute to bear its administrative costs. Regardless of who assumes the costs, it remains critical that a CMO have sufficient incentives and tools of governance to ensure that it operates with robust efficiency.

In the emergent market for generative AI licenses, some startups are rumored to request take rates ranging from 25–50%. Those initiatives may—or may not—perform distribution delivery or data normalization and processing roles (such as attribution technologies) beyond those of traditional CMOs to support a higher administrative fee. Continued evaluation and scrutiny in potential licensing partners is appropriate to

33. *How Do Authors' Societies Work?*, CONFÉDÉRATION INTERNATIONALE DES SOCIÉTÉS D'AUTEURS ET COMPOSITEURS, <https://members.cisac.org/CisacPortal/page.do?id=50> [<https://perma.cc/3W7P-LBHL>] [<https://web.archive.org/web/20250124202324/https://members.cisac.org/CisacPortal/page.do?id=50>] (last visited Jan. 24, 2025); ASIA-PAC. ECON. COOP., GUIDELINES ON THE BEST LICENSING PRACTICES OF COLLECTIVE MANAGEMENT ORGANIZATIONS (CMOs) TO MSMEs 36 (2019).

34. Note, the U.S. is the largest geographical market for all CMO uses, and music is the largest type of use for collective management globally. *See CISAC Collections Grew by 7.6% To a Record High of €131 Billion in 2023, Boosted by a 9.6% Rise in Digital Revenues To €46 Billion*, CREATIVE INDUS. NEWSL. (Oct. 24, 2024), <https://creativeindustriesnews.com/2024/10/cisac-collections-grew-by-7-6-to-a-record-high-of-e131-1-billion-in-2023-boosted-by-a-9-6-rise-in-digital-revenues-to-e4-6-billion/> [<https://perma.cc/RCA3-F73S>].

35. THE MECHANICAL LICENSING COLLECTIVE, INITIAL SUBMISSION OF THE MECHANICAL LICENSING COLLECTIVE (THE MLC) 35 (2024), <https://www.copyright.gov/rulemaking/mma-designations/2024/initial-submissions/mlc-initial-submission-2024.pdf> [<https://perma.cc/6D4S-DGYH>] [<https://web.archive.org/web/20250309233626/https://www.copyright.gov/rulemaking/mma-designations/2024/initial-submissions/mlc-initial-submission-2024.pdf>].

36. *Frequently Asked Questions*, SOUNDEXCHANGE, <https://www.soundexchange.com/frequently-asked-questions> [<https://perma.cc/GM7C-QGJS>] [<https://web.archive.org/web/20240922235237/https://www.soundexchange.com/frequently-asked-questions/>] (last visited Jan. 24, 2025) (“SoundExchange pays royalties directly to music creators, offering monthly payments at one of the lowest administrative rates in the music industry (between 4–6%) and paying out most royalties within 45 days of receipt.”).

37. Todd Longwell & Geoff Mayfield, *Variety's 2024 Music Legal Elite Report Covers Top Attorneys Who Rock Showbiz*, VARIETY (Oct. 9, 2024), <https://variety.com/lists/music-legal-elite-report/clara-kim-4/> [<https://perma.cc/P8A5-9YV7>] [<https://web.archive.org/web/20250222075605/https://variety.com/lists/music-legal-elite-report/clara-kim-4/>] (profiling ASCAP executive, who also noted “the others take a minimum of 15%”).

ensure that intermediaries develop in a supportive fashion that reduces, rather than compounds, licensing inefficiencies and costs.

#### D. TRANSPARENCY.

Collecting societies have a duty of transparency to their members, to ensure the responsible discharge of representation, distribution, efficiency, and operational obligations. Transparency is especially important where participation is compelled, but it remains important even where members can choose whether or not to affiliate with a given CMO. As the head of the U.S. music publishers association recently said regarding music PROs, “I believe that you have a fundamental right to know what it costs you to use a particular collection society,” and the U.S. Copyright Office was recently asked to study, *inter alia*, “the level of information currently provided by PROs to the public.”<sup>38</sup> Along with periodic statements and educational material, a CMO’s annual report is intended to function as an important tool of transparency, and various governments and bodies have looked to specify the content of these reports and other disclosure requirements.<sup>39</sup> For the same reason, audit rights can be important tools, and may be negotiated for or required by regulation.<sup>40</sup>

#### E. OTHER OPERATIONAL CONSIDERATIONS.

At a basic level, a licensing agent or collecting society needs the operational acumen to perform its function accurately and efficiently, and data management is a paramount obligation. While the manifold operational considerations involved in collective management for LLM licensing are beyond the scope of this article, it is worth noting

---

38. Ashley King, *House Judiciary Committee Demands Greater Transparency from ASCAP, BMI, Other PROs*, DIGIT. MUSIC NEWS (Sept. 13, 2024), <https://www.digitalmusicnews.com/2024/09/13/house-judiciary-committee-pros/> [https://perma.cc/W4E4-E5WM] [https://web.archive.org/web/20250309234003/https://www.digitalmusicnews.com/2024/09/13/house-judiciary-committee-pros/]; Letter from Reps. Jim Jordan, Chairman, H. Judiciary Comm., et al., to Shira Perlmutter, Register of Copyrights (Sept. 11, 2024), <https://thetrichordist.com/wp-content/uploads/2024/09/mic-coalition-letter-to-co-re-gmr.pdf> [https://perma.cc/4C2D-668X] [https://web.archive.org/web/20250126174009/https://thetrichordist.com/wp-content/uploads/2024/09/mic-coalition-letter-to-co-re-gmr.pdf] (asking the Copyright Office to study the operational aspects of various U.S. PROs).

39. See, e.g., Directive 2014/26, *supra* note 26, art. 21 (requiring membership terms, licensing contracts, tariffs, discounts, distribution policies, management fees, revenue deductions, representation agreements, and dispute procedures to be made public); CONFÉDÉRATION INTERNATIONALE DES SOCIÉTÉS D’AUTEURS ET COMPOSITEURS, PROFESSIONAL RULES FOR MUSICAL SOCIETIES §11 (2018) (outlining annual report requirements including requirement to detail revenue, expenses, and distribution criterion); WIPO TOOLKIT, *supra* note 29, at 21 (“A collective society . . . must answer within a reasonable time all reasonable requests from the public for information about its repertoire of works, performer’s performances, sound recordings or communication signals.”) (quoting Copyright Act, R.S.C. 1985, c C-42, art. 70(11) (as amended up to June 22, 2016) (Can.)); BRITISH COPYRIGHT COUNCIL, PRINCIPLES OF GOOD PRACTICE FOR COLLECTIVE MANAGEMENT ORGANISATIONS—POLICY FRAMEWORK 34 (2011) (outlining multiple transparency practices).

40. See 17 U.S.C. § 115 (d)(3)(L).

the importance of well-functioning operations when considering whether to encourage or even obligate the use of a CMO. The EU requires that CMOs have “sufficient capacity to process electronically, in an efficient and transparent manner, data needed for the administration of such licences, including for the purposes of identifying the repertoire and monitoring its use, invoicing users, collecting rights revenue and distributing amounts due to rightholders.”<sup>41</sup> The speed at which a CMO can process usage and issue payments is important for members depending on distribution income and can vary widely, ranging from seventy-five days to nine months or even years.<sup>42</sup> In many sectors, the use of standard identifiers and metadata requirements accompanies licenses for digital works, and already efforts are underway to shore up metadata relevant to uses of copyrighted works for generative AI purposes.<sup>43</sup>

#### IV. CONCLUSION

The explosion of generative AI technologies provides challenges and opportunities for publishers and developers to partner and experiment with licensed uses of literary content. As we peer past the cusp of early licensing announcements, it is helpful to look to copyright principles, and current licensing practices for non-AI uses, and to encourage the private marketplace to evolve and iterate in beneficial ways. History suggests we are unlikely to find, or need, a one-size-fits-all solution, given the immense amount of content and variegated uses in the mix. Instead, AI developers, like other businesses that depend on copyrighted content, may require scaling up business and compliance functions to effectuate transactions and mitigate risk.

Following established copyright law, a principle of conservatism should be applied to allow market solutions to blossom before considering whether persistent licensing gaps remain or any adjustment is needed. While collective licensing, on a voluntary basis, may be helpful to facilitate generative AI licensing, mandatory collective

---

41. Directive 2014/26, *supra* note 24, art. 24.

42. See *What Is the Payment Timeline?*, MECH. LICENSING COLLECTIVE, <https://help.themlc.com/en/support/what-is-the-payment-timeline> [https://perma.cc/78W4-PBFV] [https://web.archive.org/web/20250225064437/https://help.themlc.com/en/support/what-is-the-payment-timeline] (last visited Jan. 26, 2025) (“The MLC will issue monthly royalty statements approximately 75 days following the end of each calendar month.”); European Commission Memorandum 14/79, Directive on Collective Management of Copyright and Related Rights and Multi-Territorial Licensing—Frequently Asked Questions, at 8 (Feb. 4, 2014), [https://ec.europa.eu/commission/presscorner/api/files/document/print/en/memo\\_14\\_79/MEMO\\_14\\_79\\_EN.pdf](https://ec.europa.eu/commission/presscorner/api/files/document/print/en/memo_14_79/MEMO_14_79_EN.pdf) [https://perma.cc/QSG6-GARJ] [https://web.archive.org/web/20250126183028/https://ec.europa.eu/commission/presscorner/api/files/document/print/en/memo\_14\_79/MEMO\_14\_79\_EN.pdf] (“[C]ollective management organisations should regularly and diligently pay royalties to rightholders—no later than nine months . . .”).

43. See, e.g., COAL. FOR CONTENT PROVENANCE & AUTHENTICITY, <https://c2pa.org/> [https://perma.cc/98CG-8SRS] [https://web.archive.org/web/20250305102122/https://c2pa.org/] (last visited Jan. 26, 2025); *The Invention of Trust for Digital Media*, LICCIUM, <https://liccium.com/> [https://perma.cc/JVQ8-RAC8] [https://web.archive.org/web/20250307031637/https://liccium.com/] (last visited Jan. 26, 2025).

management of rights is far from the only option. When considering options and evaluating potential intermediaries, potential members and licensees can look to established CMO guideposts relating to membership and representation obligations, distribution policies, administrative fees, transparency, and operational acumen.