



POSITION PAPER

Artificial Intelligence

The Copyright Alliance supports the responsible, respectful, and ethical development and use of artificial intelligence (AI) technologies and a thriving and robust AI economy. The continuing development of AI systems represents a profound achievement of the digital age that brings with it tremendous opportunities. In fact, many in the creative industry are already using or plan to use AI for the creation of a wide range of works that benefit society. But as with many advances in technology, these new opportunities come with challenges.¹

Advancements in AI have led to a new frontier in generative technologies, and thus they are often accompanied by difficult legal questions surrounding both the ingestion of copyrighted works into AI systems and the output generated by those systems. As AI technology continues to evolve and questions arise about how copyright laws apply to the development and use of generative AI models, it's critical that the underlying goals and purposes of our copyright system are upheld and that the rights of creators and copyright owners are respected.

When examining the intersection of AI and copyright, the following general principles must form the basis of a common understanding amongst stakeholders, courts, policymakers, and the public.

1. *When formulating new AI laws and policies, it is essential that the rights of creators and copyright owners be respected.* When making determinations about AI policies, it is vital for policymakers and stakeholders to understand that any new laws and policies relating to AI must be based on a foundation that preserves the integrity of the rights of copyright owners and their licensing markets. The interests of developers who use copyrighted materials for ingestion by AI systems must not be prioritized over the rights and interests of creators and copyright owners.
2. *Longstanding copyright laws and policies must not be cast aside in favor of new laws or policies obligating creators to essentially subsidize the development of AI*

¹ This paper addresses topics specific to the use of copyrighted works for ingestion by AI systems. There are several other questions that will arise as to who, if anyone, is the “author” of a work generated by an AI system, who, if anyone, is responsible for any copyright infringement committed via such system, and the copyrightability of AI-generated works in general. Those subjects are not addressed in this position paper.

technologies. Established copyright laws must not be weakened based on a mistaken belief that doing so is necessary to incentivize the development of AI technologies. This is especially true when there is no evidence of market failure or problems warranting changes to the law. AI-specific statutory exceptions to copyright law that would effectively strip rightsholders of their ability to control and be compensated for the use of their copyrighted works for ingestion purposes are unnecessary and harmful and should be rejected.

3. *The ingestion of copyrighted material by AI systems implicates the right to reproduce copyrighted works.* Section 106(1) of the Copyright Act vests copyright owners with the right to prevent the reproduction of their copyrighted works. When an unauthorized copy is made of a work protected by copyright, there is a violation of the copyright owner's right to reproduce the work, absent a valid defense. It is important to understand that copyright infringement at the input stage is distinguishable from infringement at the output stage because the reproduction right is a "stand-alone" right—it is violated by copying a work (without authority or defense) regardless of whether a specific output of an AI system is infringing.
4. *The ingestion of copyrighted material by AI systems is not categorically fair use.* Determining whether a particular use qualifies for the fair use defense to infringement requires a fact-specific inquiry that is considered on a case-by-case basis. There are no uses that always, categorically qualify as fair use. That is no less true when copyrighted work are used for AI ingestion. In fact, the typical commercial system's ingestion of copyrighted works is particularly unlikely to qualify as fair use when the AI system generates competing works. Courts will need to evaluate fair use defenses involving AI systems the same way they evaluate fair use in all contexts: by applying the four factors set forth in section 107 of the Copyright Act to the specific uses at issue. Under the first factor, ingestion is unlikely to be a transformative use since the output generated by these AI systems will often serve the same exact purpose as the works ingested, especially in the case of music and art. However, even if the use is held to be transformative, as the Supreme Court made clear in *Andy Warhol Foundation v. Goldsmith*, whether a use is transformative is not dispositive of the question of fair use and is merely one of the considerations under the first fair use factor. In addition, under the fourth factor, when courts consider the extent of the "effect of the use upon the potential market for or value of" the works ingested by that system, they may conclude that (i) copyright owners have made licenses available in the market for AI training and therefore allowing a non-licensed user to use the works harms the market for the work; and/or (ii) the market is sufficiently diluted because of the mass generation of substitutional works.
5. *AI companies' mass ingestion of copyrighted works from illicit sources to train their AI systems is unethical, extremely harmful, and should weigh heavily against fair use.* Many leading AI companies have knowingly sourced massive amounts of pirated works from known criminal-level piracy services (and in some cases re-distributed those works). The fact that AI companies intentionally use of these illicit sites and services does presumptively disqualify an AI developer from claiming fair use is antithetical to the foundations of our copyright system and the rights it guarantees. Courts and policymakers must recognize that sourcing mass amounts of pirated copyrighted works is not necessary for the development of

generative AI models and should not be excused as fair use.

6. *AI companies should license works they train on.* AI-copyright policy should not interfere with the free market or the freedom to license works for training. Obtaining a license to use copyrighted works is the best way for AI developers to ensure they avoid infringement liability. Copyrighted works provide immense value to AI developers, and they can and should pay for that value—as many today are already doing.² Many AI licenses exist all over the world and across a wide spectrum of creative disciplines—and the number continues to grow. The existence of this vibrant licensing market weighs against a finding of fair use under the fourth fair use factor. The marketplace should continue to properly value and incentivize creativity, and AI policy should not interfere with the right of copyright owners to choose whether to license, or not to license, their works for AI purposes. In other words, when properly applied, copyright law sets the conditions for the market to prevail.
7. *AI systems must implement safeguards to prevent infringing AI-generated outputs.* Overfitting and allowing prompts that call for copyright protected-material and “in the style of” are more likely to result in AI-generated outputs that infringe one or more copyrighted works. While merely imitating the style of an existing artist does not constitute infringement, it is essential that AI companies implement effective safeguards to prevent the likelihood of output-related infringements. This is yet another reason why AI companies should voluntarily license ingested works because when they do so, the parties can negotiate these safeguards.
8. *Transparency regarding ingestion of copyrighted works by businesses that offer generative AI systems to the public will help ensure that the rights of copyright owners are respected, and that AI development is being implemented in a way that is responsible and ethical.* Adequate and appropriate transparency and record-keeping benefit both copyright owners and AI developers in resolving questions regarding infringement, fair use, and compliance with licensing terms. Transparency has many other benefits unrelated to copyright such as promoting safe, ethical, and unbiased AI systems. Consequently, transparency by businesses that offer generative AI systems to the public is a crucial component of any AI policy. Best practices should include maintaining records of what copyrighted works are being ingested and how those works are being used, except where the AI developer is also the copyright owner of the works being ingested by the AI system. Those records should be publicly accessible and searchable as appropriate and subject to reasonable confidentiality provisions the parties to a license might negotiate as well as the aforementioned exception.

² See the Copyright Alliance’s compilation of licensing deals and partnerships at <https://copyrightalliance.org/artificial-intelligence-copyright/licensing/>.