## Before the United States Patent and Trademark Office

## Department of Commerce

| | | |
|---|---|---|
| In the Matter of | **)** | |
| | **)** | |
| **Impact of Artificial Intelligence** | **)** | Docket No. PTO-C-2019-0038 |
| **("AI") Technologies on Copyright** | **)** | |
| | **)** | |

Document Number: 2019-23638

## Comments of the Authors Guild, Inc.

## Submitted by Mary Rasenberger

The Authors Guild submits this statement on behalf of its almost 10,000 members in response to the USPTO's Request for Public Comment concerning the impact of artificial intelligence ("AI") on copyright law and policy. Our members represent every cross-section of the writing community, including novelists, poets, journalists, academic and textbook authors, genre writers, traditionally and self-published authors. Copyright is the lifeblood of our members' art and trade, and as such the questions posed by the PTO are highly salient to present and future challenges facing us.

The purpose of copyright is to incentivize the creation of new works—including literary works, which contribute so greatly to our nation's store of knowledge and culture. Advancements in artificial intelligence and machine learning raise daunting concerns that cut to the core of copyright law and have important consequences for the future of the creative industries. These issues are difficult, but must be proactively addressed to ensure that the incentives for human creativity continue to function without interfering with the growth of AI. The Authors Guild would like to thank the USPTO for opening this timely conversation.

### Defining Artificial Intelligence

The umbrella term "artificial intelligence" broadly refers to "a set of techniques aimed at approximating some aspect of human or animal cognition using machines."[1] This term is often used together with "machine learning," which encompasses particular

---

[1] Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 404–05 (2017).

methods whereby machines are trained to develop cognitive semblance [2] Machine learning typically involves ingesting and "analyzing vast amounts of data and deriving general principles [to enable machines to] improve their ability to accomplish tasks."[3] AI is not a singular algorithm, process, or device, but a combination of various algorithms, processes, and devices configured to perform some tasks or produce certain results. The human actors behind an AI system are often similarly various, consisting of developers, programmers, and users.

Given the wide range of artificial intelligence models and their applications, it is important at the outset to understand the term accurately with respect to copyright law. AI systems that are competent at human expressive activity have been called generative machines;[4] these generative machines, more so than other forms of AI, directly concern core copyright law issues. One example of such a generative use of AI is ING and Microsoft's *The Next Rembrandt*, which used machine learning to create a 3-D printed Rembrandt painting based on 168,263 Rembrandt painting fragments.[5] The AI-generated "Portrait of Edmond de Belamy," which fetched an astounding $432,500 at a 2018, is another example. Computer scientists and machine learning enthusiasts have been experimenting with generating literary writing for decades,[6] and a sci-fi novel written by a Japanese AI has even come close to winning a literary prize.[7] Yet another example of an AI generative machine is OpenAI's "Musenet," an algorithm that can generate 4-minute musical compositions.[8] Much of the recent advancements in AI generated creative works can be attributed to improvements in neural networks and the algorithms designed to recognize patterns in data that are loosely modeled on the human brain.[9]

Generative AI can be classified based on the degree of autonomy they possess. For instance, in Professor Jane Ginsburg's offers the following typology, which we will refer to in our comments to clarify the issues:

Ordinary tools: These machines "rely solely on the creative contributions of their

---

[2] Bernard Marr, *What is the Difference Between Artificial Intelligence and Machine Learning?* FORBES (Dec. 6, 2016), https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#3a39f2972742.

[3] Jane C. Ginsburg & Luke Ali Budiardjo, *Authors and Machines,* 34 BERKELEY TECH. L.J. 343, 401

[4] *See generally id.*

[5] Mary Brown, "New Rembrandt to be unveiled in Amsterdam, THE GUARDIAN (Apr. 5, 2017), https://www.theguardian.com/artanddesign/2016/apr/05/new-rembrandt-to-be-unveiled-in-amsterdam.

[6] *See generally* David Streitfeld, *Computer Stories: A.I. is Beginning to Assist Novelists*, N.Y. TIMES (Oct. 18, 2018), https://www.nytimes.com/2018/10/18/technology/ai-is-beginning-to-assist-novelists.html; John Seabrook, *The Next Word: Where will Predictive Text Take Us*, NEW YORKER (Oct. 14, 2019), https://www.newyorker.com/magazine/2019/10/14/can-a-machine-learn-to-write-for-the-new-yorker.

[7] Ben Schiller, *This Japanese Novel Authored by a Computer is Scarily Well-Written*, FAST COMPANY (Mar. 28, 2016), https://www.fastcompany.com/3058300/this-japanese-novel-authored-by-a-computer-is-scarily-well-written.

[8] MUSENET-AI, https://openai.com/blog/musenet/#fn2 (last visited Jan. 10, 2020).

[9] Larry Hardesty, *Explained: Neural Networks,* MIT NEWS (Apr. 14, 2017), http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414 ("Neural nets are a means of doing machine learning, in which a computer learns to perform some task by analyzing training examples. Usually, the examples have been hand-labeled in advance. An object recognition system, for instance, might be fed thousands of labeled images of cars, houses, coffee cups, and so on, and it would find visual patterns in the images that consistently correlate with particular labels.").

users, and for which the creative contributions of the machines' designers are minimal, nonexistent, or not apparent in the resulting work."[10] The most common example of an ordinary tool is MS Word. Copyright in works resulting from use of these tools belongs to the user.

Partially-generative machines: These systems "combine the creative contributions of both the user and the designer of the tools, those creative contributions being inseparably fused in the resulting work—form the center of the spectrum."[11] Works generated by these machines create complicated issues of copyright ownership due to the combination of user and developer contributions, and absent a licensing arrangement, are likely to be works of joint authorship.

Fully-generative machines: These machines "rely entirely on the creative contributions of their designers and do not require any creative choices made by the users (who simply turn the machine on or tell it to "create")—form the other end of the spectrum.[12] Depending on various factors, especially the degree of involvement by the designers, works resulting from use of these machines may either be attributed to the authorship of the designers or be authorless works.

**1. Should a work produced by an AI algorithm or process, without the involvement of a natural person contributing expression to the resulting work, qualify as a work of authorship protectable under U.S. copyright law? Why or why not?**

There is no basis in existing U.S. copyright law for providing nonhuman creation with copyright incentives, and we do not believe that there is any reason to change the law to allow for nonhuman expression to be treated as copyrightable authorship. While there is a natural impulse here to find an author in whom to ascribe authorship, not all works need be protected by copyright. Uncopyrightable works are not a novel issue for copyright law, and indeed works of conceptual art such as Chapman Kelley's Wildflower Works have been excluded from copyright protection for lacking human authorship.[13] Any fixed, original human expression in the AI machines that generate new works, however, might be copyrightable works, and where the AI is functioning as a human author's worker or amanuensis[14], then the resulting work may also be copyrightable human authorship.

**Human Authorship Requirement**

The Compendium III of U.S. Copyright Office Practice explained that "The U.S. Copyright Office will register an original work of authorship, *provided that the work was created by a human being* (emphasis added)." The Compendium goes on to explain that copyright law

---

[10] Ginsburg & Budiardjo, *supra* note 3, at 405.

[11] *Id.*

[12] *Id.*

[13] *See* Kelley v. Chicago Park District, 635 F.3d 290, 304 (7th Cir. 2011) ("A living garden like Wildflower Works is neither 'authored' nor 'fixed' in the senses required for copyright."); *see also* Agnieszka Kurant, *Phantom Capital, Hybrid Authorship, and Collective Intelligence*, 39 COLUM. J.L. & ARTS 371 (2016).

[14] *See* discussion *infra* pp. 5-7.

[O]nly protects "the fruits of intellectual labor" that "are founded in the creative powers of the mind." Because copyright law is limited to "original intellectual conceptions of the author," the Office will refuse to register a claim if it determines that a human being did not create the work.[15]

We believe this to be sound policy consistent with the Constitution and supported by case law going back more than a hundred years. The intellectual property clause of the U.S. Constitution empowers Congress to protect the exclusive rights of *authors and inventors* in their writings and discoveries; and although the Constitution does not define authors and inventors as human beings or natural persons, it is clear from the context and the usage of these words at the time that the framers meant humans—not animals or paranormal or artificial beings. A century of Supreme Court jurisprudence further validates the view that authorship is necessarily a product of a *human* intellect. In *Burrow-Giles Lithographic Co. v Sarony*, the Supreme Court defined an author as someone "to whom anything owes its origin; originator; maker; one who completes a work of science or literature," and writing as the process "by which the ideas in the mind of the author are given visible expression."[16]

The Court further expounded on the concept of authorship in the 1903 case *Bleistein v. Donaldson Lithographing Co.*, where it infused authorship with the uniqueness of the human creator's individual personality.[17] Even a "very modest grade of art," Justice Oliver Wendell Holmes, Jr. wrote in the opinion, "has in it something irreducible, which is one man's alone. That something he may copyright."[18] *Bleistein*— and *Sarony* to a lesser extent—stand for the proposition that a copyrightable work is the product of human intellect, with an ascertainable human point of origin. This fundamental premise is not altered by copyright law's accommodation of corporate ownership under the work-made-for-hire doctrine, where works are created by human employees or contractors, or by joint-ownership, where the copyright law ascribes copyright ownership of a joint work to each human author who contributed original authorship.[19] Over the decades, the concept of human authorship has been tested under a range of novel conditions, from authorship by spiritual beings to forces of nature and animals,[20] and each test has only further reified the premise of human authorship.

The reasons the copyright law protects only human author are rooted in the constitutional purpose of incentivizing the progress of arts and sciences.[21] Forces of

---

[15] U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 306 (3d ed. 2017) (citations omitted).

[16] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 58 (1884).

[17] Bleistein v. Donaldson Lithographing Co.188 U.S. 239, 250 (1903).

[18] *Id.*

[19] COMPENDIUM (THIRD) §§ 505.1, 505.3, and 503.1

[20] *See* Penguin Books U.S.A., Inc. v. New Christian Church of Full Endeavor, Ltd. No. 96 Civ. 4126 (RWS), 2000 WL 1028634 (S.D.N.Y. July 25, 2000); *Kelly*, 635 F.3d 290; Naruto v. Slater, 888 F.3d 418, 426 (9th Cir. 2018) (holding that "animals. . . lack statutory standing to sue under the Copyright Act*"); see also* Annemarie Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author*, 2012 STAN. TECH. L. REV. 1, 18–20 (2012) (discussing non-human putative authors).

[21] U.S. CONST. art. 1, § 8, cl. 8.

nature, machines and works made by animals are not incentivized by copyright; nor are there policy reasons to allow people to be able to claim ownership in them—say, for instance, in a beautiful rock or a bird's song. [22] AI machines, in spite of their anthropomorphic portrayals in works of fiction, also do not need incentives to create. It is the humans and human-run corporations who we want to incentivize with intellectual property. Humans create the algorithms, processes and tools that result in copyrightable authorship. Copyright protects those works of human intellect, and conjoined with current trademark, trade secret, and patent laws, provides ample incentives for human and human-run corporations to develop and use AI systems.

We are not aware of any works or authorship currently created solely by AI, without human involvement, although it is conceivable that might be the case in the future. To the extent that AI systems do create original works of expression on their own, especially with respect to expressive works in the various arts—works that express aspects of the human condition, such as books, poetry, music and visual arts—we do *not* believe there is a need to provide copyright protection. (If the software industry recognizes a need to protect AI-written code, as some have argued, then we suggest a separate *sui generis* form of protection, rather than revising copyright law generally and interfering with other copyright industries.) Granting copyright protection works that are generated by an AI machine without significant human authorship—as a result of ingesting mass amounts of existing human works of authorship, identifying patterns, and mimicking them in some sense through the rules it is fed—does not advance the goals of copyright in any way. It neither advances the particular field of expression since it merely mashes up and regurgitates existing expression, nor does it add to human understanding. AI will never add that spark of human creativity–the elements that makes art true art. Moreover, allowing wholly AI generated works to be copyrightable will put human authored works at a huge competitive disadvantage since AI can produce works faster and cheaper.

The result will not just be one more disruption in the work force; we will be a much poorer society because those AI-created works, no matter their superficially similarity to works of human provenance, will lack the experience and emotion of the human artist. Great books and other works of art help us understand ourselves and each other; they make us empathetic and help us understand our own lives and time and place in the world; the arts take us out of the mundane of our day-to-day lives and help us feel and see more truths about human existence. That experience makes the arts so fundamentally important to every human society. AI generated works rehash, mash-up, and rework what they are fed; they do not tell or emote.

---

[22] Ginsburg & Budiardjo, *supra* note 3, at 448 ("[W]e should not assume that we need copyright-like protection to stimulate the production of authorless outputs. Absent an author, the premise underlying incentive justifications requires substantiation. One must inquire whether these outputs in fact need the impetus of exclusive rights, or if sufficient incentives already exist, for example higher up the chain, through copyright or patent protection of the software programs, patent protection of the specialized machinery to produce different kinds of outputs, and copyright protection of the database the software consults. Trade secrets and contracts may also play a role in securing the outputs.").

As the Supreme Court has noted, copyright is "an engine of free expression."[23] Bestowing copyright on AI-generated works will ensure that the marketplace can no longer afford anything but; and copyright law will no longer function as our founders intended—to build free marketplaces where works of human intellect can be freely traded and human creators of expressive works can be paid. The human arts won't die altogether—because we still will be human after all, but the type of quality works that advance us as a society – those created by trained, paid professional, often with corporate investment, will disappear.

**Human Copyrightable Authorship in Works Created for or Using AI**

It is possible for AI-enabled works to be created as the human-intended, predictable output of a fully-generative AI machine. In such cases, the resulting works of original authorship could be the copyrightable authorship of the human designers who created the AI machine. Acts of authorship can be attributed to an agent or amanuensis worker in the manner that large sculptures or art installations, for example, often are actually "made" by workers under creator's direction. As Professor Ginsberg explains, "[t]he law [following agency rules] attributes authorship to the "mastermind," whose detailed conception so controls [the work's] subsequent execution that the individuals carrying out the embodiment exercise no creative autonomy."[24] An AI system following the detailed guidance of its users and/or programmers and under their authority would be an amanuensis, and authorship in the work generated by the AI would be attributed to the human masterminds. *The Next Rembrandt* might be found to be an example of such a work, since the computer scientists and art historians behind the project gave the AI system specifically detailed instructions on *what* to create and *how* to create it, feeding it a limited set of data from which the results could be predicted within a range. What they conceived was so specific—a Rembrandt-like portrait of a middle age man in certain costume in a particular pose—that one could find that they jointly were the "mastermind" and the AI system was their worker. Consequently, authorship in the work (as long as it is found to be original) could be collectively attributed to them.

If a human user, instead of programmers and developers of the AI system, directs the *what* and *how* (i.e., provide the detailed conception for execution by the AI) of an original AI-generated work, authorship would be attributed to the user. In other words, human authorship can be found in outputs using AI technologies closer to the spectrum to ordinary tools than generative machines. For instance, AI-based image processors that allow users to restore, stylize, and enhance images can fall into this category: the end user might create copyrightable expression (in the form of a derivative work).[25] On the other hand, if a user only commands the production of the output, without providing original expressive determinative choices to that output, the user's command is no more than an idea for the resulting work and excluded from copyright protection by Section102(b)[26]

---

[23] Harper & Row v. Nation Enterprises, 471 U.S. 539, 558 (1985)

[24] *Id.* at 358.

[25] *See e.g.* Christopher Thomas, *Deep learning based super resolution, without using a GAN*, Towards Data Science (Feb. 24, 2019), https://towardsdatascience.com/deep-learning-based-super-resolution-without-using-a-gan-11c9bb5b6cd5.

[26] 17 U.S.C. § 102(b)

6

of the Copyright Act. It is the detailed instructions for the *what* and *how* of the output that can establish copyrightable authorship in the output.

**2. Assuming involvement by a natural person is or should be required, what kind of involvement would or should be sufficient so that the work qualifies for copyright protection? For example, should it be sufficient if a person (i) designed the AI algorithm or process that created the work; (ii) contributed to the design of the algorithm or process; (iii) chose data used by the algorithm for training or otherwise; (iv) caused the AI algorithm or process to be used to yield the work; or (v) engaged in some specific combination of the foregoing activities? Are there other contributions a person could make in a potentially copyrightable AI-generated work in order to be considered an ''author''?**

From the outset, we will note that it is impossible to answer this question in the abstract without knowing the exact specifications of the AI technology. Nevertheless, we have found Professor Ginsburg agency/amanuensis theory useful in thinking about copyrightable authorship in AI systems and their output. We respond to the scenarios in the questions below:

(i) and (ii): As we discussed in our answer to the previous question, it is possible for an entity or persons who design an AI system to have a copyright claim under the amanuensis theory in AI-generated output that comprises an original work of authorship– if the person(s), in designing the system provided sufficiently detailed parameters for the execution of a work that the person intended or foresaw, and that work meets copyright law's originality requirements.[27] If more than one person or entity (on a work-made-for-hire basis) designs the system, then they may have a joint copyright in the output as joint authors.

There may be multiple copyrightable works created or incorporated in the development of an AI system. AI systems may incorporate pre-existing, copyrightable works of authorship, with separate copyright ownership. Programmers and developers may contribute separable works of copyrightable authorship to an AI system; programmers who assemble the code together to construct an AI system may have copyright in the arrangement as a "collective work."[28]

(iii) A person who "chose[s] data used by the algorithm for training or otherwise" may have copyright in the training dataset as a compilation work if contributing sufficient selection and arrangement original authorship.[29] However, it seems implausible that the construction of a dataset would amount to sufficiently detailed direction to the system to avail a trainer's copyright claim in the resulting work.

(iv) A person who caused AI to yield a new work may have a copyright claim in the resulting work if their contribution goes beyond a mere command and embodies a degree of creative planning, and direction of the work under the amanuensis theory. Also, if the AI is a mere tool and the user is providing most of the creative expression, then the

---

[27] COMPENDIUM (THIRD) § 308.
[28] COMPENDIUM (THIRD) § 509.
[29] COMPENDIUM (THIRD) § 308; *see also* 17 U.S.C. § 103.

user would be the copyright owner in a result. For example, the user of an AI-based music generation tool such as MuseNet, which allows users to combine instrumentation and musical styles (such as Baroque and bluegrass) to create an original musical composition, would be the author of the copyrightable expression in the output, if any.[30]

(v) Any person who contributes substantial copyrightable expression with the mutual intent of creating a joint work may be considered a co-author of the output along with others who meet the requirements for joint authorship.[31] In all cases, whether contributors are co-authors or not, the copyright interests of multiple parties will have to carefully delineated through transfers, licenses, co-authorship agreements, and terms of use for downstream users. Down the road, AI machines may incorporate the contributions of so many different authors and pre-existing work[32] that an automated licensing/tracking system will become necessary.

**3. To the extent an AI algorithm or process learns its function(s) by ingesting large volumes of copyrighted material, does the existing statutory language (e.g., the fair use doctrine) and related case law adequately address the legality of making such use? Should authors be recognized for this type of use of their works? If so, how?**

Existing statutory language and case law do not adequately address the potential legal issues created by AI's ingestion of large volumes of copyrighted material. Machine learning methods involve the copying of mass volume of either noncopyrightable data or of copyrightable works, the latter which particularly concerns us. Courts have addressed the ingestion of mass copying of copyrighted works for AI and search engine purposes under the doctrine of fair use; and they have found fair use in each case.[33] As AI grows more sophisticated, continued reliance on fair use has the potential to cause enormous harm to creators and the expressive arts, as described in response to Question 1.

When thinking about the application of fair use to the ingestion on works for AI training, it is helpful to distinguish between 1) ingestion for training non-generative AI algorithms and process; i.e., AI that perform functions that do not result in expressive activity; and 2) ingestion for training generative machines that create expressive works. Needless to say, the unauthorized (unlicensed) ingestion of copyrighted works to generate new competitive creative works will ultimately cause market harm to the value of human-created copyrighted works that the AI machines essentially mimic in style and essence. Those types of uses should *not* be permitted without authorization. In some cases,

---

[30] MUSENET, *supra* note 8.

[31] *See, e.g*., Thomson v. Larson, 147 F.3d 195 (2d Cir. 1998).

[32] Katherine Forrest, *Copyright Law and Artificial Intelligence: Emerging Issues,* J. COPYRIGHT SOC'Y U.S.A. 65:4, 355 (2018).

[33] *See e.g.* Authors Guild v. Haithitrust, 755 F.3d 87 (2d Cir. 2014) (finding full-text searches of scanned books to be fair use); Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007) (finding reproducing thumbnails to be fair use); Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015) (finding search function and limited display to be fair use); Fox News Network, LLC v. TVEyes, Inc. Nos. 15-3885, 15-3886 (2d Cir. Feb. 27, 2018) (suggesting that the recording and database functions would qualify as fair use, but rejecting fair use on the grounds that service allowed users to view substantial segments of copyrighted transmissions in a way that could deprive Fox of licensing revenue).

ingesting of copyrighted works for functional purposes might also interfere with the value of copyrighted works.

Application of fair use by the courts is not a rational or practical way to make decisions on when uses need to be authorized or compensated. Rules on when works may be ingested with or without authorization or renumeration should be created though coherent policymaking and not in an ad-hoc manner by courts who only have the facts of a specific cases in front of them. Too much is at risk. Many courts in the last decade have exhibited a tendency to over-emphasize transformative use in their analyses of the fair use defense, often sidelining the important fourth fair use factor that instructs considering potential market harm to the copyrighted works. As Mary Rasenberger and June Besek observed, this "ascendency of transformative use, and in particular, "functional transformation," gives rise to concern that the fair use pendulum has now swung too far away from its roots and purpose, now enabling new business models rather than new works of authorship."[34]

Where works are ingested to train AI to perform a non-expressive function, that output is of course less likely to compete with the ingested works: and in some cases, a use might be one that we want to allow without compensation or permission because there is no potential for market harm and the use benefits society. But not all cases of functional use should be deemed fair use. There may be a downstream adverse impact on the value of copyrighted works that a court could easily overlook based on the limited set of facts presented to it in the particular case. For instance, in the case of Google Books Search, Google ingested millions of books to train its computers to read and write in human language.[35] It copied expressive language for expressive purposes and at the same time created a book search engine that could in some cases serve as a substitute for buying a book. The harm to a particular book from Google Book Search arguably is not great, but the collective harm caused by the entities that have ingested books on the basis of that decision, such as Internet Archives' Open Library, and the many others who have or will followed suit, is growing and is very definitely chipping away at authors incomes. Authors' incomes have declined by a mean of over 40% in the last decade. The expansion of the fair use doctrine plays a significant role in that decline in that it has eradicated much of authors' U.S. licensing income for library, educational and other uses that are arguably "fair" under the recent broad interpretations of transformative use. [36] (We understand that other types of creators' incomes have also had precipitous drops in the digital age.) A coherent policy review will take into account the overall effect of allowing fair use to swallow potential licensing markets – something that courts cannot do when looking at the facts of the particular case in front of them.

---

[34] Mary Rasenberger & June Besek, *The Authors Guild v. Google: The Future of Fair Use?*
https://www.authorsguild.org/wp-content/uploads/2015/12/AG-v-Google-The-Future-of-Fair-Use.pdf
[35] Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).
[36] *Authors Guild Survey Shows Drastic 42 Percent Decline in Authors Earnings in Last Decade* (Jan. 5, 2019), https://www.authorsguild.org/industry-advocacy/authors-guild-survey-shows-drastic-42-percent-decline-in-authors-earnings-in-last-decade/.

**Collective Licensing for Commercial Machine-Learning**

If we want to preserve a sustainable economy for human authorship in the U.S., we need to create a way for human authors whose works are ingested by AI algorithms and processes to receive remuneration for such use. As noted above, this will eventually require a sophisticated, automated licensing system. Courts have sympathized with technology companies who conducted mass copying for functional purposes for the obvious reason that obtaining permissions on a work-by-work basis is highly impractical. The only reasonable means of enabling licenses for mass use is collective licensing. For this reason, the Authors Guild proposed an extended collective license for out of print books to the Copyright Office.[37] Some copyright owners object to mandatory or the opt-out nature of extended collective licensing because they want to control the use of their works, but compensated use would greatly alleviate pressures on the creative industries and is far preferable to allowing mass ingestion of works for functional purposes on a fair use basis – what in practice amounts to a free compulsive license.

Mass ingestion of copyrighted works to create competing creative works – as we have already seen for romance novels – should never be subject to extended or compulsory licensing, or permitted under fair use. The risk of grave harm to the markets for the ingested works is too great.

The EU has taken a policy driven approach. Articles 3 and 4 of the recent EU DSM Directive establishes clear exceptions for non-commercial, research text and data mining, and reserves authorization with the rightsholders when the use is commercial. We are not advocating for adoption of the Directive's provisions on text and data mining, but they provide a framework for how to think about the issues.[38]

> "In certain instances, text and data mining can involve acts protected by copyright, by the sui generis database right or by both, in particular, the reproduction of works or other subject matter, the extraction of contents from a database or both which occur for example when the data are normalised in the process of text and data mining. Where no exception or limitation applies, an authorisation to undertake such acts is required from rightholders."[39]

The Directive evinces a widespread circumspection about the potential of text and data mining functionalities hurting the creative economy. The recital cited above is also consonant with the EU's adoption of Article 17, which explicitly directs online content-sharing providers to gain authorization from the rights-holders. Restricting the copyright immunities of tech-businesses encourages licensing—something that can only be managed globally through a robust collective licensing scheme.

---

[37] Comments of The Authors Guild to the Copyright in the Matter of Mass Digitization Pilot Program, *available at* https://www.copyright.gov/policy/massdigitization/comments/Authors%20Guild.pdf

[38] Directive 2019/790/EU 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 96/9/EC and 2001/29/EC, OJ (L 130), 113-114, *available at* https://eur-lex.europa.eu/eli/dir/2019/790/oj.

[39] *Id.* at 93.

The Authors Guild has long been a proponent for a collective licensing scheme that would allow authors to authorize their work for text and data mining uses. Considering that data mining and patter recognition lies at the core of machine learning and AI systems, and recognizing their power to create disruptions in the creative industries, we urge the USPTO and sister-agencies to revisit collective licensing as a possible solution to rights management and enforcement in the digital realm.

**4. Are current laws for assigning liability for copyright infringement adequate to address a situation in which an AI process creates a work that infringes a copyrighted work?**

No. The technological tilt towards increasing automation and diminishing role of legal actors in directing AI actions raises serious concerns about the adequacy of current legal rules to address the types of infringement that AI will be capable of creating and causing.

Particularly troubling is the application of the volitional conduct line of cases to infringement where AI is the proximate cause. Courts may find that the AI caused the infringement and that neither the creator of the AI system or any human user had the requisite volition.[40] *Cartoon Network, LP v. CSC Holdings, Inc.* and its progeny could easily be applied to find a lack of necessary volition by companies that produce automated technologies that infringe, including AI systems that infringe on a mass basis. Furthermore, case law on secondary liability offers scant bright-line rules for deciding whether an AI developer would be deemed liable as a contributory infringer for inducing or otherwise causing the infringement, or as a vicarious infringer for reaping the financial benefits from infringement that they can control. Developers of an AI system would be held liable for indirect infringement under the *Sony, Napster* and *Grokster* line of cases only if the AI technology's primary purpose is infringement or if the technology does not have has "substantial non-infringing uses."[41]

This presents a serious challenge for copyright law, as it allows beneficiaries of copyright infringement—the owners of the technology aiding the infringement—to evade liability and leave rightsholders without recourse. Down the road one can imagine AI systems that infringe for their own purposes. The copyright law will have to find a way to ascribe infringement to legal entities or persons.

AI systems dramatically increase the potential for rapid and large-scale infringement. And as they become more widespread and capable of autonomously creating infringing works, courts will need clear legal rules to attribute liability. We cannot emphasize the importance of proactively amending the law to account for AI generated infringement.

---

[40] *See, e.g.*, Cartoon Network, LP v. CSC Holdings, Inc., 536 F.3d 121 (2d Cir. 2008).
[41] Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417 (1984); A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (2001); MGM Studios, Inc. v. Grokster, Ltd., 545 U.S. 913 (2005).

**5. Should an entity or entities other than a natural person, or company to which a natural person assigns a copyrighted work, be able to own the copyright on the AI work? For example: Should a company who trains the artificial intelligence process that creates the work be able to be an owner?**

There is nothing under U.S. law to prevent an entity from owning copyright of copyrightable works of authorship that are part of AI systems or are works that are created by persons or entities using AI as a tool or generating works using AI as an amanuensis. Copyrights in AI related works are assignable to entities just as any other copyrights are under Section 201(d) of the Copyright Act.[42] As explained in our answers to Question 1 and 2, existing rules of copyright authorship, ownership and transfer are sufficient for addressing these questions.

**6. Are there other copyright issues that that need to be addressed to promote the goals of copyright law in connection with the use of AI?**

Returning to the human authorship requirement, some argue that the focus on the human intellect as the source of authorship evinces an inability to imagine the possible existence of non-human cognition. But the framers included copyright in the first article of the Constitution because they believed that the mind and labor of authors to be critical for the functioning of the democratic polity they imagined. In the debates around the commercial uses of copyright, we tend to forget copyright's role in shaping our democracy as the engine of self-determination, ideas, and expression. The Constitution created copyright as an incentive to continue the tradition of free expression, inquiry and discovery. We need to keep this purpose in mind going forward, and ensure that authors and inventors have those incentives.

Developments in artificial intelligence are both exciting and unsettling. They force us to interrogate our beliefs about human nature and art. On the other hand, AI-based technologies like "deep fakes" can also distort facts and muddle the truth to dangerous consequences. But fundamentally AI is a tool, albeit with remarkably generative potential that could have grave consequences for human society and culture. Pointing out the fact that we need responsible policy and clear legal rules to manage our inevitable transition to a world where human consciousness will interact and interface with machine-based forms of intelligence is not neo-Luddism or "crying wolf." Neither is resisting the notion that machine-based forms of intelligence are somehow superior and will render the work of human artists, authors, composers, and poets—whose speculation and expression allows to make sense of our past, present, and future—meaningless. A world where machines in the form of AI dominate our arts is not a world most of us will want to live in. If human creative industries are left without protection against the inevitable competition of AI-created works, they will surely flounder, and we will all be worse off.

---

[42] 17 U.S.C. § 201(d)