WIKIMEDIA
FOUNDATION

January 10, 2020

Docket No. PTO-C-2019-0038: Intellectual Property Protection for Artificial Intelligence Innovation

Comments of the Wikimedia Foundation

Introduction and Background

The Wikimedia Foundation appreciates the opportunity to comment on the US Patent and Trademark Office's Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation. The Wikimedia Foundation is the non-profit organization that hosts and supports a number of free, online, collectively-produced resources, including Wikipedia. The Foundation's objective is to help create a world where every human being can share freely in the sum of all knowledge.

Our mission and projects thus intersect with artificial intelligence and machine learning ("AI/ML") in two major ways. First, we develop AI/ML tools to assist editors in improving our projects, such as finding gaps in Wikipedia's coverage, or identifying poor quality articles or edits.[1] Second, the Foundation's projects are designed to be used, reused, and repurposed, and since they are freely licensed, our volunteer editors' works, organizing, and other contributions are frequently used in training AI/ML systems.

---

[1] *See, e.g.,* Jonathan T. Morgan, *Designing ethically with AI: How Wikimedia can harness machine learning in a responsible and human-centered way*, Wikimedia Foundation (July 2019), https://wikimediafoundation.org/news/2019/07/18/designing-ethically-with-ai-how-wikimedia-can-harness-machine-learning-in-a-responsible-and-human-centered-way/; Miriam Redi et al., *Can machine learning uncover Wikipedia's missing 'citation needed' tags?*, Wikimedia Foundation (April 2019), https://wikimediafoundation.org/news/2019/04/03/can-machine-learning-uncover-wikipedias-missing-citation-needed-tags/; Miriam Redi, *How we're using machine learning to visually enrich Wikidata*, Wikimedia Foundation (March 2018), https://wikimediafoundation.org/news/2018/03/14/machine-learning-visually-enriching-wikidata/.

Our experience with these systems highlights an important consideration when discussing authorship in relation to AI/ML systems: such systems as they exist today are tools developed and deployed by humans. While the term "artificial intelligence" has long been associated with the concept of artificial sentience, as it is used in research and industry today, it most typically means a system that performs tasks that previously required human cognitive processing, such as translating complex and varied audio recordings of speech into text. The most salient aspects of these systems — that they are complex; that their internal decision-making processes can be difficult to predict; or can run autonomously — do not mean that they constitute a cognizable legal entity or an author. Complex algorithms that govern ordinary computer systems do not raise novel questions of authorship; unpredictable processes, such as generating random numbers or simulating chaotic phenomena, are still the tools of those who use them; and autonomous action by machines is at least as old as clockwork.

Rather than beginning from a more theoretical (if not science-fictional) assumption of a created sentience (sometimes called "strong AI"), the Office would likely be better served by addressing the implications of AI/ML tools that are commonly being used and refined today: automated systems designed to perform specific tasks by one party, and deployed in particular applications of those tasks by another. Within that scope, there are a number of questions relevant to IP development, such as how authorship might be attributed between the tool's creator and its user, and the rights implicated by the use of copyrighted works in training AI/ML systems.

There are also several critical considerations not only regarding how AI/ML developments affect IP law, but how IP law may affect the impact of AI/ML systems on individuals' human and legal rights. For instance, copyright policy affects what data sets are used to train AI/ML, frequently leading to skewed and biased systems. Copyright policy should not be so restrictive as to create incentives for AI/ML systems to be trained only on public data, or works from before 1924. Furthermore, when AI/ML systems involve the personal information of individuals, assertions of trade secrets or copyrights should not prevent affected individuals from exercising their privacy rights. Also, when AI/ML systems are used to make determinations that affect people's lives, such as setting credit ratings, bail amounts, or potential for future violations, IP rights should not prevent inspection, auditing, and analysis of systems. Intellectual property laws should not be usable as pretexts to evade accountability.

Response to Questions 1-2:

Question 1 asks whether a work produced by an AI algorithm or process, without the involvement of a natural person contributing expression, could qualify as a work of authorship protectable under US copyright law. While this is a fascinating question, we believe that it has less

salience than it may appear for other aspects of AI/ML and IP law, and may even distort the proper analysis of other copyright questions involving AI/ML.

The current state of machine learning and artificial intelligence encompasses many algorithms and processes that fall far short of a system that could be considered an independent author. The term "artificial intelligence" is commonly used to describe systems that are simply more sophisticated versions of older automated systems that rely upon large sets of data and the repeated application of statistics. They are therefore merely tools that help people work more efficiently. Thus, the concept of a protectable work produced "without the involvement of a natural person" is far more niche — a natural person will necessarily be involved in any works of authorship in the near future.

Moreover, beginning an inquiry into copyright law and AI/ML with this premise runs the risk of skewing or needlessly complicating perspectives on authorship disputes involving the types of AI/ML systems currently in use and that will continue to be used into the future. This is not to say that the questions dealing with the exclusive rights of an artificial being in its creative works should never be addressed; it is merely to say that they likely should be addressed with a lower priority than more fundamental rights for such beings (such as rights to existence, liberty, freedom from involuntary servitude, or freedom of expression), and that even those more basic rights are unlikely to become non-academic questions decades after more prosaic matters need to be litigated or ruled upon.

A more likely model for potential authorship disputes concerning AI/ML would involve one individual using an AI/ML tool that was created by a second individual. This model may provide insight into Question 2 without needing to first make determinations about non-human authorship, and suggest further guidance on how AI/ML systems can be accounted for by current legal doctrines. Presupposing that AI/ML systems are, or could be, non-human authors needlessly inserts an uncertain legal construct into established law.

Currently, non-natural persons cannot be authors. The Copyright Office has consistently taken the stance that works produced through fully randomized machine processes or natural processes may not be registered because protections only extend to work produced by a human (e.g., randomized linoleum flooring designs or naturally "carved" driftwood).[2] This does not prevent works of authorship from containing elements that were created by automated, or even natural, processes. However, in those cases where authorship is found, it is vested in the person who deliberately set the automated process in motion in a creative way, or who contributed human creativity by modifying or combining and arranging the results of natural or automated processes in a sufficiently creative way. A piece of mounted driftwood carved and polished by waves may not be copyrightable, but an artist

---

[2] United States Copyright Office, *Compendium of U.S. Copyright Office Practices* §§906.6-906.7 (2017), https://www.copyright.gov/comp3/chap900/ch900-visual-art.pdf.

might incorporate such a piece into a copyrighted work, or might create a stencil to shape the effects of waves on a beach on a wooden sculpture, and copyright the result.

In this case, a natural person is contributing expression by using a process of nature as a tool. Where, instead of a natural process, an artificial, automated process is involved, we look to the natural person as the potential author, and the separate process, however complicated it may be, as simply a tool that they used. A writer that uses an automated spell-checker, or even an automated thesaurus, is still the author of her written work; we do not, absent some extraordinary circumstance, attribute any authorship to the developer of those automated systems, even if they incorporate AI/ML or otherwise perform their functions without the writer's explicit direction.

In evaluating whether the tool user is the proper author of the creation, existing jurisprudence on copyrightability and traditional standards of copyright, such as the threshold of creativity articulated in *Feist Publications, Inc. v. Rural Telephone Service Co.*,[3] adequately address the matter. In particular, if the overall work created with the assistance of the tool fails to meet the threshold of creativity, *Feist* would still apply and limit the application of copyright protections.

For instance, a sculptor using a 3D printer provides all of the creative input, even if the printer contains operating software that was created or owned by its manufacturer. Even if the sculptor were to make an error allowing the printer to run longer than intended (thus running autonomously, and without the sculptor's immediate creative input) and produce an unexpected result, it would still be the sculptor who would have copyright in the work created as a result of their mistake. If all the sculptor did was turn on the printer and it extruded a blob of plastic, the sculptor's actions would likely fall short of creative expression necessary for authorship, and the resultant blob, no matter how evocative, would simply not be protected by copyright.

As for the contribution of the tool makers — those who designed or contributed to the AI or process, and those who provided or selected inputs to it — they are, in most cases, unlikely to have contributed sufficient expression to the work in order to be considered authors. For instance, even if an AI tool has already been programmed to create images, a person must decide to use the tool to produce a desired output. Therefore there will nearly always be contribution from a natural person present in the result of the image creation tool. The extent to which that person is the author of a protectable work then relies upon *Feist* and other existing doctrines of authorship.[4]

---

[3] 499 U.S. 340 (1991).

[4] Manufacturers or developers of some programs and systems would be considered authors of those programs and systems themselves, insofar as they contain protectable expression. This is distinct from any copyright interest that might exist in outputs made by using those programs. For instance, while Microsoft may have a copyright in its word processing software Word, it has no copyright interest in texts written using Word.

Question 2 also asks whether various types of contributions to an AI-generated work could lead to authorship. As with traditional questions surrounding potential joint works and contributions, the specific nature of the contribution matters a great deal — far more so than the question of whether any system marketed as AI/ML is involved in the process.

AI/ML systems that add substantial amounts of material with very little input from their end users may create situations where assigning authorship may seem counterintuitive, but these questions do not need to be resolved by designating an author, much less deciding that (as per Question 5), the creators of the AI/ML system are the authors. Instead, existing doctrines regarding "thin" copyright may apply: if an end user has made sufficient creative decisions in the use of the AI/ML tool, he might have a copyright in the resulting work, even while a separate end user, making similar decisions with the same tool, may also have a copyright in her independently created work, regardless of its substantial similarity to the first user's.

If the end user's inputs are insufficiently creative to merit authorship, however, there is no particular reason that the AI/ML system's developer should be assigned the copyright. If a simpler algorithmic system were merely choosing from a limited set of hand-coded results created by the developer, those specific results, if protectable expression themselves, could be considered works authored by the developer. For instance, if a 3D printer manufacturer created a model of a horse and included it with the printer as a test pattern to be printed at the push of a button, the manufacturer, and not the button-pusher, might have an interest in the resulting sculpture. However, where the results of the process, such as from an unpredictable AI/ML system, could not be said to be sufficiently determined, selected, or fixed by the developer or other system contributor, the answer may simply be that the work has no legal author and thus is not protected by copyright.

Treating such a work similarly to a product of nature avoids several complications and confusions. For instance, if a default rule held that the developers of a system were the authors of every work made via that system, they could become the holders of copyrights in works whose very existence they were unaware of. AI/ML systems developed in collaborative processes as free / open source software could create countless works with a similar multitude of authors, likely opening the door to punishingly complex disputes over ownership of the works, without a clear reason for any of the parties involved to justify such ownership.

Finally, it is likely that allowing an algorithm or process to create protectable works of authorship absent expressive input from a human would create significant unintended consequences. As one example, consider ubiquitous programs such as word processors or proofreading programs. As they improve, they are beginning to suggest common language, identify points of ambiguity, and provide automated rephrasing of written works. If their contributions rose to a level that might constitute a derivative work, the programmers of a word processor could potentially assert a blocking

ownership right in every work written on their software. Similar AI/ML tools used in other fields could easily create abusive restraints on competition via such rights. Yet if algorithms gain the ability to create protectable works of authorship without proximate human expression, there is a high chance that people far removed from the output of the algorithm will nevertheless try to claim ownership of such outputs.

Response to Question 3:

Question 3 asks whether existing statutory and case law adequately addresses AI algorithms ingesting large volumes of copyrighted material. Even if particular instances of AI training require a fact-specific inquiry, we believe that existing law that deals with the processing of multiple copyrighted works[5] can be applied directly to these issues. As noted above, the relative sophistication or complexity of the system in question should not alter the basic inquiries in copyright law: whether one of an author's exclusive rights under section 106[6] has been implicated, and whether that implication occurs within the bounds of a limitation or exception of sections 107-122.

Response to Question 4:

Question 4 asks whether 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. We believe that they are. Determining infringement in ordinary circumstances frequently requires a court needing to infer whether or not copying has taken place in the absence of an admission, or direct evidence, of copying. The standard test of access to the copyrighted work and substantial similarity between the original work and the AI/ML output should apply just as they would to a defendant who had used technologies other than AI/ML to create a new work.

The same standards should also apply for independent creation: if the system (and the defendant) had no access to the original work, they could not have copied it and thus did not infringe. Similarly, the fair use doctrine would also apply, rendering a sufficiently transformative use a non-infringing one. The fact-specific nature of these doctrines allows them to accommodate whatever relatively minor changes may be introduced by AI/ML systems.

---

[5] *See, e.g., Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015); *Authors Guild v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014); *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630 (4th Cir. 2009); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007).

[6] This would seem to mostly implicate the right of reproduction. The Office should be wary of an over-expansive interpretation of what might constitute the preparation of derivative works in assessing whether an AI trained on copyrighted works is, or contains, derivative works. The Office should also note some of the limitations inherent to the definitions of the Act, such as instances of works so fleeting as to not constitute a "copy." *See Cartoon Network LP v. CSC Holdings, Inc.*, 536 F.3d 121, 130 (2d Cir. 2008).

Response to Question 6:

Question 6 asks whether there are other copyright issues that need to be addressed to promote the goals of copyright law in connection with the use of AI. We wish to emphasize that copyrights and AI/ML may intersect in more critical ways than how AI/ML systems may complicate liability. That more important intersection deals with ongoing questions of how AI/ML and other black-box systems and algorithms are used to make decisions that have drastic effects on individuals' lives.[7]

One substantial problem is bias in AI/ML systems.[8] When selecting training data, developers of AI/ML systems will often select public domain or freely licensed works, both for ease of access and to avoid potential infringement liability exposure. This creates a particular form of selection bias in training data, which can often lead to other forms of more harmful bias in results. For instance, a system ingesting published books in English from before 1923 will likely exclude a disproportionate number of authors of color; systems using biographies and news stories of people from before this cutoff date to correlate them with occupations would fail to associate women with jobs that were denied to them in the early 20th century or before, and carry those systemic biases into the future.[9]

Copyright, along with other types of IP, should also not be misused to create barriers to algorithmic transparency. Already, we are seeing confidentiality being used as an excuse for companies not to make their algorithmic decision making processes available for audit. This is especially concerning in cases where algorithmic decision making is used to make decisions about people's lives, or their rights.[10]

Algorithms today are being developed or are already used to determine prison sentences, loan terms, credit ratings, and bail amounts. Online, algorithms are used to determine what content to surface to a wider audience on social media, and in some cases, what content to remove or block. Individuals should be able to trust that automated systems can and have been audited by someone with their interests in mind, and should be able to understand on some level how a decision was made.

While companies may rely primarily upon trade secret to obscure the inner working of automated decision making, the copyrightability of code has the potential to allow for misuse of

---

[7] *See*, *e.g.*, Frank Pasquale, *The Black Box Society: the Secret Algorithms that Control Money and Information* (2015); Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (2017); Cathy O'Neil, *Weapons of Math Destruction:How Big Data Increases Inequality and Threatens Democracy* (2017).

[8] *See* Aaron Halfaker, *Mitigating biases in artificial intelligences - the Wikipedian way*, Wikimedia Foundation (October 2018), https://wikimediafoundation.org/news/2018/10/10/mitigating-biases-artificial-intelligences-wikipedian-way/.

[9] *See, e.g.,* Amanda Levendowski, *How Copyright Law can Fix Artificial Intelligence's Implicit Bias Problem*, 93 Wash L. Rev. 579, 593 (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3024938.

[10] *Id.* at 597-606.

copyright to deny algorithmic transparency. For instance, public interest organizations or other auditors of such systems will often need to make reproductions of copyrighted code in order to analyze it. Even if they do not explicitly copy source code onto their own devices, these researchers will inevitably make temporary copies in RAM as they interact with or merely use the AI/ML systems. Doing so may expose them to threats of liability for copyright infringement if they are using the software without the permission of the algorithm's developer, or potentially if they use the software contrary to the terms of the software license.[11] Researchers investigating potential discrimination may also be exposed to liability for violating anti-circumvention measures under Section 1201.[12] These are not theoretical threats: recent history contains several examples of companies misusing copyright provisions to prevent transparency and accountability.[13]

The Office should strive to ensure that copyright exceptions and limitations remain robust to allow for beneficial uses like news reporting, research, auditing, and public accountability, and must work with companies and NGOs to build processes by which IP rights may be respected, but not misused to prevent accountability as AI/ML systems affect the full scope of fundamental human rights.

Response to Question 10:

Question 10 asks how AI impacts trade secret law. As mentioned in our response to Q6, we are concerned about the application of IP laws to prevent oversight into algorithmic decision-making, particularly where that decision making will have a substantial effect on people's lives.[14] Just as copyright law should not serve as a pretext to prevent analysis and auditing of AI systems, trade secrets should not be invoked to obscure potentially discriminatory or otherwise actionable results of AI/ML-derived decisions. This would include any rights that individuals have under privacy, data protection, or data access laws to view information held about them, as well as attempts to audit, reverse engineer, or evaluate such systems to ensure they are operating legally and free from invidious discrimination. Citizens' fundamental rights to be free from discrimination, retaliation, intrusion upon privacy, and other harms should take precedence over broad assertions of trade secret.

---

[11] *See, e.g., MDY Indus., LLC v. Blizzard Entm't, Inc.*, 629 F.3d 928, 939-942 (9th Cir. 2010) (discussing possibility that violations of license conditions in use of otherwise legally-obtained software may result in a finding of copyright infringement).

[12] *See MDY*, 629 F.3d at 943-955.

[13] *See, e.g., Online Policy Group v. Diebold, Inc.*, 337 F.Supp.2d 1195 (N.D. Cal. 2004); Edward Felten, *The Chilling Effects of the DMCA*, Slate.com (March 29, 2013), https://slate.com/technology/2013/03/dmca-chilling-effects-how-copyright-law-hurts-security-research.html; Levendowski at 602-610 (noting attempts to use copyright law to prevent reverse engineering).

[14] *See* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 Stanford L. Rev. 1343 (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2920883##.

Response to Question 13:

Question 13 asks whether there are relevant policies or practices that may help inform the USPTO's policies and practices regarding IP rights. With respect to the concerns raised regarding the misuse of copyright and other IP rights to prevent algorithmic accountability, we would refer the Office to the Toronto Declaration,[15] which outlines a human rights framework to be applied to the use of AI/ML systems. Ensuring that AI/ML applications are designed and implemented within a framework of human rights, and not merely according to a *sui generis* set of ethical principles, can more properly ensure that various aspects of the law, including IP law, are coordinated in promoting justice and equality, in addition to science and the useful arts.[16]

Conclusion

The Wikimedia Foundation believes that developments in artificial intelligence and machine learning have comparatively modest effects upon the state of intellectual property law. Many existing doctrines already account for authors using automation to create works without controversy, and robust doctrines already exist to account for the more remote possibilities of non-human creation. Existing case law has demonstrated an ability to account for mass processing of copyrighted works; the particular types of technologies associated with that processing should have less effect on the outcome of such cases than the specific facts of the individual cases themselves. Finally, it is critical that stakeholders in IP law recognize the potential for it to inadvertently perpetuate bias, or be abused by those who would seek to avoid accountability in their algorithmic systems. Stakeholders including policymakers, industry, and civil society must ensure that IP law remain flexible enough not to be misused in this way.

---

[15] The Toronto Declaration, https://www.amnesty.org/download/Documents/POL3084472018ENGLISH.PDF

[16] *See* Anna Bacciarelli, *Ethical AI principles won't solve a human rights crisis*, Amnesty International (June 21, 2019), https://www.amnesty.org/en/latest/research/2019/06/ethical-ai-principles-wont-solve-a-human-rights-crisis/