January 10, 2019
Via Electronic MailAIPartnership@uspto.gov

Attention: The Honorable Andrei Iancu
        Under Secretary of Commerce for Intellectual Property
        Director of the United States Patent and Trademark Office

IBM Corporation Comments in Response to "Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation", 84 Fed. Reg. 58141 (October 30, 2019)

IBM thanks the United States Patent and Trademark Office ("Office") for the opportunity to provide comments in response to the Office's Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation. We have organized our comments as direct responses to each of the questions contained in the request.

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?*

Response: No. We believe that a human author is necessary to render a work copyrightable, at least under U.S. law. The copyright clause of the Constitution, the Copyright Act, recent case law (*See, e.g.*, *Naruto v. Slater*, 888 F. 3d 418 (9th Cir. 2018)), and the Compendium of U.S. Copyright Office Practices, § 313.2, consistently contemplate human authorship as required for copyright protection in the U.S. The use of an AI algorithm or process in connection with the creation of a work should not disqualify a work from copyright protection, however, as long as the contribution of the human author contains sufficient originality to meet the "spark of creativity" test.

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"?*

Response: We believe that a work is not entitled to copyright protection unless the human activity considered separately generates a work that is independently copyrightable, even though the human activity need not compose the entirety of the work created with the AI system.

We interpret sub-questions (i)-(v) as inquiring which human actors should have the right to claim to be co-authors of a copyrightable work created substantially by an AI algorithm or process. We note at the outset that we do not believe that an AI system can itself be considered an "author" or a "co-author" of a copyrighted work. Instead, we interpret the question to be directed to which human actors can be considered to be authors or co-authors. Based on that interpretation, we have the following comments about the examples that the Office has posited:

(i)   We do not believe that this example is sufficient to establish co-authorship by itself, because the "work" under consideration is only created and fixed in a tangible medium after the AI algorithm or process is created and trained. The expression of that algorithm or process may well itself be copyrightable, but it does not follow that the author of the algorithm should own new copyrightable works created by implementations of that algorithm or process. Likewise, even if the human creator obtained a patent covering the algorithm or process, that patent would not entitle the patentee to claim ownership of anything created by that algorithm or process, including the copyright on such creation.

(ii)   We believe that this example is analytically the same as (i) above.

(iii)   We believe that this example posits two cases: one is a case where data is used just for training and the other is a case where data is used for purposes beyond simply training. We believe that the selection of inputs for training by itself is too remote from the generation of the ultimate "work" to constitute authorship of that work. However, we do believe that it is possible for a selection of data by a human actor (as embodied, for example, in a copyrightable compilation of previously existing works) to constitute the needed human authorship component to creating a copyrightable work in conjunction with an AI algorithm or process, if that selection directly triggers the creative process whereby the copyrightable work is created.

(iv)   We believe that example (iv) could supply the necessary human authorship -- if that activity involves human activity which separately generates a work that is itself copyrightable – because it is proximate to the creation of the copyrightable work.

(v)   In our view, this example is likely to describe the most common AI scenarios, where a variety of the individual examples listed above are combined in a unified effort to create the ultimate copyrightable work. In practice, AI implementations are typically complex engagements, with unique system training activities and solution customizations, often involving significant professional services in addition to pre-existing AI technologies. In such a case, determining human co-authorship is likely to be a fact-dependent exercise, but could result in the persons performing each of the enumerated activities sharing in authorship of the resulting copyrightable work.

In general, with respect to the question about other contributions a person might make, we would analyze the use of an AI algorithm and process similarly to the use of any other tool. In each case, the use of such a tool may facilitate or augment the creation of copyrightable works rather than rendering such works uncopyrightable. *See, Burrow–Giles Lithographic Co. v. Sarony*, 111 US 53 (1884).

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?*

Response: We believe that the fair use doctrine, as presently interpreted by the courts, allows the use of copyrighted works for system training purposes in the U.S. We also believe that copyrighted works may be copied, searched, and small segments of text or thumbnail images displayed by computer systems. *See, e.g., Authors Guild v. Google, Inc.*, 804 F. 3d 202 (2d Cir. 2015). While the fair use doctrine has enabled the development of AI technologies in the U.S. to date, however, that doctrine is an extremely fact-intensive analysis and thus the scope of its protections is less certain than ideally desirable.

We further note the passage of explicit legislation authorizing and defining text and data mining (TDM) permissions for both commercial and noncommercial purposes in the copyright law of Japan and (at least at the Directive stage) in the European Union. While these legislative developments have been driven by the absence of fair use law in these jurisdictions, they provide the advantage, particularly in the Japan law, of clearly articulating the scope of permitted conduct in TDM-related activities. While we believe that the fair use doctrine has supported the development of AI technologies to date in the United States, we think that adding clarity to the scope of permitted use of copyrighted works for TDM purposes for both commercial and noncommercial entities would further advance the development of AI technologies in the United States as well.

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?*

Response: We do not see any basis under U.S. law for the concept of machine-based liability. Analogizing to other areas of the law imposing liability for the failure of machines to perform as expected (e.g., products liability), the law always seeks to impose liability on human or corporate actors, whether they be owners, operators, or creators. As copyright infringement is fundamentally a statutory tort, legal responsibility for infringements committed by an AI system must lie with the human or corporate actors controlling that system. Since AI systems cannot act without some human control and are certainly created and trained by humans, we expect that the law would impose liability for AI-created infringements on one or more of those actors, depending on the specific facts of the case.

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?*

Response: We believe that an AI algorithm or process cannot be an author of a copyrighted work and therefore cannot be granted ownership of the copyright in any work it has a role in creating. This question asks whether anyone other than a natural person or a legal entity (e.g., an AI algorithm or process) can own a copyright. We believe that the answer is no. Copyrights are property rights and, as such, can only be owned by legal entities or natural persons. Likewise, ownership of copyrights can be exchanged through contracts between and among legal entities and natural persons under current law, but entities other than these (including AI algorithms or processes) have no ability to enter contracts. Therefore, an AI algorithm or process can neither be granted a copyright nor be assigned a copyright under current law. Additionally, if we permitted ownership of property rights to rest with an AI algorithm or process, it is unclear how that entity would be able to enforce or enjoy those rights. For example, how would a machine be able to decide to sue to remedy infringements of its copyright, or to assess and process royalties?

With respect to the cited example, we do not believe that the training of an AI algorithm or process, in the absence of other facts, should give the trainer a copyright interest in the ultimate works generated by the system. Such a right is more likely to inhere to the user or owner of the AI algorithm or process itself.

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

Response: The ultimate goal of copyright law is to "promote the progress of science and the useful arts" (U.S. Constitution, Article 8). While allowing authors to obtain and protect those rights in the form of copyright protection remains critically important, expanding the scope of that protection beyond copyrightable expression itself threatens the promise of AI technology to advance the public good through text and data mining and data analysis. In particular, overly expansive interpretations of derivative works rights, structure / sequence concepts, and prohibitions on copying for analytical (rather than competitive) purposes may result in diminished access to the factual content of those works, which the emergence of AI technologies can help unlock. As observed by the World Intellectual Property Organization ("WIPO") in its 2019 "World Intellectual Property Report," innovation has become increasingly international and collaborative. Optimally, intellectual property law should seek to foster this collaboration while providing appropriate protection for inventors and authors. In the case of AI technologies, we believe that this balance is of great importance. As discussed in our response to Question 3 above, while the fair use doctrine has allowed the development of AI technologies to date, we believe that the ultimate adoption of

explicit TDM rights would further enable future developments in AI by establishing clearer legal parameters for development activity.

7. *Would the use of AI in trademark searching impact the registrability of trademarks? If so, how?*

Response: We believe that AI has the potential to be useful in trademark searching, but that human trademark examiner involvement will still be necessary. In particular, AI may be helpful in assisting with visual / image recognition, as well as semantic analysis. If used in this fashion -- as a tool to assist trademark examiners -- we believe that AI can improve trademark searching results and help reduce improper registrations. We do not believe, however, that AI has the capability to replace the judgment of the trademark examiner, which will still be required to review and assess the evidence assembled by the AI in the examination and registration process.

8. *How, if at all, does AI impact trademark law? Is the existing statutory language in the Lanham Act adequate to address the use of AI in the marketplace?*

Response: We do not believe that AI, when used as a tool to augment the judgment of the examiner, represents a fundamental shift in the trademark registration process and such use would therefore not require a change in statutory language.

9. *How, if at all, does AI impact the need to protect databases and data sets? Are existing laws adequate to protect such data?*

Response: We believe that AI has already had the effect of making data analysis more accessible and increasing the potential for data to be effectively mined for insights. At the same time, data and AI are also very interdependent: Development of effective AI depends on access to a sufficient quantity of unbiased data for its initial development and training and, in the case of machine learning technologies, on its ongoing ability to access and analyze yet more data. Quite literally, reliable and effective AI systems could not exist without access to significant volumes of quality data.

We believe that any discussion of new forms of data protection must be informed by the need to balance the protection of innovation on the one hand with the need to preserve freedom of access to data on the other. We therefore believe that it is important to proceed cautiously with introducing new categories of IP rights. At this point, we see no evidence that supports a need for new IP protections for data.

10. How, if at all, does AI impact trade secret law? Is the Defend Trade Secrets Act (DTSA), 18 U.S.C. 1836 et seq., adequate to address the use of AI in the marketplace?

Response: We believe that a practical effect of using AI technology may be that information (including know-how and insights) may be more readily discoverable without defeating secrecy measures. One of the central premises of trade secret law is that it can be used to protect information that could only be gathered with great difficulty. AI could make it easier to "readily ascertain[] through proper means" such information and thus undermine its status as a trade secret, *see* 18 USC 1839(3)(B). Thus, the practical impact of AI could be to raise the bar on what are considered to be reasonable measures to keep information secret. *See* 18 USC 1839(3)(A). While this potential effect of AI should be acknowledged, we do not believe that it should fundamentally change the legal standards for determining trade secret status, which has always been dependent on a specific, fact-based analysis.

11. Do any laws, policies, or practices need to change in order to ensure an appropriate balance between maintaining trade secrets on the one hand and obtaining patents, copyrights, or other forms of intellectual property protection related to AI on the other?

Response: While we believe the law of trade secrets continues to play a role in intellectual property law generally, we note the findings of WIPO that innovation is increasingly collaborative across organizations and even national boundaries. In our view, U.S. IP policy should encourage disclosure over secrecy. The patent system, which is disclosure based, is better suited to this new collaboration-based innovation model than a trade secret-based system, which favors secrecy and disfavors collaboration. That said, to encourage inventors of new AI technology to file patent applications, rather than rely on trade secret law, more certainty is needed for patentable subject matter eligibility determinations. This will require legislation from Congress.

Under the current vague patent eligibility standard, it is far too uncertain whether a new invention constitutes eligible subject matter. For this reason, inventors are more likely to keep an invention secret instead of filing a patent application, thus depriving the public from learning about the invention.

In addition, patent protection facilitates collaboration because a partner typically relies on pre-existing patents to prevent the other party to a collaboration from taking the innovations disclosed. As today's research initiatives are highly technical and complex and the costs too high for most organizations to bear on their own, companies routinely enter into joint development agreements (JDAs) with research partners. Through these JDAs, partners typically cross-license each other's patents, but only for the purposes of that research effort. If there are background patents on the underlying technology, the partners would likely need to grant a broader cross-license in order to commercialize the resulting technology. If it is difficult or impossible to enforce patents on a particular subject matter, such as AI inventions, there is increased risk that disclosure of innovations to a research partner may result in that partner

commercializing those innovations without just compensation. This discourages collaboration and retards the progress of technology.

Finally, the transparency of AI operations is one of the primary impediments to the widespread acceptance of AI technologies. – If people do not understand enough about how an AI system operates (e.g., how it avoids bias in its analysis), they may be reluctant to accept its insights. Here again, a properly functioning patent system would offer innovators the ability to disclose the workings of AI systems, while preserving protection for their patentable innovations. Reliance on trade secret protection alone would not afford AI innovators with the ability to strike the critical balance between transparency and protection for their innovations.

> *12. Are there any other AI-related issues pertinent to intellectual property rights (other than those related to patent rights) that the USPTO should examine?*

Response: IBM has previously responded to this question in the context of its responses to FRN1. That response is reproduced below:

> "IBM is open to the possibility that new forms of IP protection may be needed as technologies like AI evolve and mature. We believe that many important questions still exist, and we should proceed cautiously and with purpose. The IP protection mechanisms currently in place have served us well for many years even as technology has evolved at an accelerating pace. Therefore, IBM suggests that the focus remain on the basic building blocks of IP currently in place, before creating any new forms of IP protection *ad hoc*.

> Machine learning and trained models are an area of particular importance to IBM. Machine learning is a program or system that is capable of digesting large data sets and utilizes AI algorithms to produce a trained model. In production, the trained model when given new, but similar types of data sets, can provide useful insights. The trained model is often viewed as a black box, with data sets as the input and insights being generated as the output. The trained model can include a learned set of coefficients for use in one or more mathematical functions for that specific trained model. While the output of the trained model and its associated coefficients seems simple, the development of the trained model and its associated coefficients is not at all trivial, and that effort merits some sort of protection.

> IBM believes there may be gaps in IP protection for AI, and specifically gaps in IP protection for protecting the trained model and its associated coefficients. As well, existing IP frameworks such as patents and copyrights may not be available for functional elements of data sets, and in some cases, the useful insights or learnings that are not fixed in a tangible medium of expression. As discussed above with respect to enablement and written description, there are situations where patent protection may not be appropriate. Similarly, copyright cannot protect the mathematical equation or the coefficients due to their functionality. Trade secret protection is also not a sustainable option. Aside from the issues of diminishing the dissemination of

information and slowing technological advancement, trade secret protection may not be adequate for such inventions because of the ease of reverse engineering and the inability to meaningfully obfuscate programming code.

Accordingly, we ask the Office to consider this gap in IP protection for AI and contemplate whether aspects of other IP protection mechanisms may be necessary to adequately protect the trained model, data sets, the coefficients gleaned from the trained model, and the algorithmic equation used to run input data through the AI machine.

Access to data is essential to the development of AI. Therefore, we must be mindful, and temper any alternative IP protection considered with an appropriate fair use doctrine. For example, many companies train machine learning models using text and data mining (TDM), technology-enabled analytics capable of discovering correlation and identifying useful insights undiscovered in data sets. This fair use of data should remain permissible. Stifling access to data could have considerable adverse ramifications for the progress of AI solutions."

The Office has been a consistent supporter of initiatives to increase the availability of federal government data to the public. We encourage the Office to continue its support for this activity, as access to sizeable volumes of unbiased data is essential for the development and training of AI technologies. We particularly urge the Office to support the release of government data under a consistent open data licensing model that adequately and explicitly grants TDM and analytical rights to released datasets, such as the Community Data License Agreements (CDLA), sponsored by The Linux Foundation. Licenses like the CDLA represent a good approach to support collaborative communities built around curating and sharing open data. At a minimum, any such licenses should be free of "no commercial use" restrictions to assure access to such datasets for the broadest potential population of users.

13. Are there any relevant policies or practices from intellectual property agencies or legal systems in other countries that may help inform USPTO's policies and practices regarding intellectual property rights (other than those related to patent rights)?

Response: We believe that the 2018 Amendment to the Japanese copyright statute provides an example of a well-considered legislative definition of a TDM right under the copyright laws.

With respect to other legal doctrines that may impact AI, we note the adoption in Europe of a separate category of database rights, pursuant to the Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases. We also note that some countries, notably the United Kingdom, Ireland, and New Zealand, have adopted statutory provisions recognizing the copyrightability of computer-generated works without a human author. *See*, Copyright, Designs and Patents Act, 1988, c. 48, § 9(3) (U.K.); Copyright Act of 1994, § 5 (N.Z.); Copyright and Related Rights Act 2000, Part I, § 2 (Act. No. 28/2000) (Ire.).

Conclusion

IBM applauds the Office for opening this important dialogue about the intersection of AI and IP rights. We are grateful for the opportunity to comment on and contribute toward meeting the challenges that AI poses for IP policy. We recognize that much uncertainty remains and the stakes are high. The answers to these strategic policy questions could alter the course of AI protection in the U.S., and ultimately our economy. We urge the Office to stimulate discussion with Congress, stakeholders, and the public regarding appropriate and measured IP rights for AI, and we stand ready to continue advising the Office on these issues.

Respectfully submitted,

Manny W. Schecter
Chief Patent Counsel
Intellectual Property Law
IBM Corporation
schecter@us.ibm.com
Voice: 914-765-4260

Christopher W. O'Neill
Associate General Counsel
Intellectual Property Law
IBM Corporation
cwo@us.ibm.com
Voice: 914-765-4484

Valerie Calloway
Senior Trademark Counsel
Intellectual Property Law
IBM Corporation
VALERIE.CALLOWAY@ibm.com
Voice: 914-765-7246

Terrence J. Carroll
Copyright Counsel
Intellectual Property Law
IBM Corporation
tjcarrol@us.ibm.com
Voice: 408-927-3391