AUTHORSHIP IN WORKS CREATED BY AI

Arth Nagpal

INTRODUCTION

In recent times, Artificial Intelligence (AI) has entered the realm of content creation, a development that could either be hugely beneficial, or irreparably detrimental, to the art industry. Copyright law protects works of authorship, with the U.S. Copyright Office adding its own embellishments to preclude non-human authors. It is imperative for an in-depth analysis of AI-generated content in light of copyright law with an eye on the merits of this human-centric approach adopted in copyright practice and whether it can be countered using modern theories of authorship.

The issue that this paper grapples with is that of copyrightability of the content created by an AI machine and related points of discussion, such as authorship and liability upon infringement while exploring means to reconcile AI-created content with the existing legal framework, suggesting modifications where necessary.

NATURAL LANGUAGE PROCESSING: INSIDE THE AI

Qualcomm defines machine learning as 'a subset of AI that refers to a machine's ability to think without being externally programmed.'[1] Conventional devices are programmed with a set of rules for how to act, and then this takes the form of if-then-else statements. But machine learning enables devices to continuously think about how to act based on data they intake.[2] This allows the machine to build on the user's input, as well as its own artificially-created content that arose as a result of its comprehension of, and its being constrained by, the user's input.

AI provides analytic tools to construct stories deriving from previously published content, particularly their semantics and syntax. Natural language processing (NLP) views the process of language analysis as being 'decomposable into a number of stages, mirroring the theoretical linguistic distinction drawn between syntax, semantics and pragmatics.'[3] The sequence of events in NLP is as follows: First,

---

[1] *Everything you need to know about AI*, ONQ BLOG (Mar. 27, 2019), https://www.qualcomm.com/news/onq/2019/03/27/everything-you-need-know-about-ai (last visited Nov. 3, 2019).

[2] *Id.*

[3] Robert Dale, *Classical Approaches to Natural Language Processing*, *in* HANDBOOK OF NATURAL LANGUAGE PROCESSING 4 (Nitin Indurkhya & Fred Damerau eds., 2d. ed. 2010).

a particular sentence is analyzed on the basis of its syntax, arranging it in a manner such that semantic analysis is more efficient and accurate. Second, the sentence is analyzed semantically, to extrapolate its literal meaning. Third, a pragmatic analysis is made whereby the meaning of the text is determined in light of its context.[4] Semantic analysis is especially significant, since the ultimate goal for humans and NLP is to understand the utterance by incorporating information supplied by the utterance into one's knowledge base, or performing some action in response to it.[5]

Natural language generation (NLG) is the process through which thought is rendered into language, and from a computational perspective, the equivalent of a person with something to say is a computer program.[6] In the process of NLG, the machine moves from intention to text as opposed to the movement from text to its intention through comprehension in NLP. Hence, in an NLG, what is 'known' is the generator's awareness of its speaker's intentions and mood, its plans and the 'content and structure of any text the generator has already produced'.[7] It has already been noted above that an AI machine relies on the input it receives, as well as the embellishments it has made over-and-above that input, to determine subsequent content creation. The AI's problem is basically that of choosing how to 'signal its intended inferences from an oversupply of possibilities along with that what information should be omitted and what must be included.'[8]

AI: CONTRIBUTION TO LITERATURE

AI has pervaded much of digital run-of-the-mill writing and research in addition to futuristic endeavors such as driverless cars and the like. Tasks like crunching data, eliminating fake news – and more relevant to this paper – generating output, have overtime come to be outsourced to machines.[9] News agencies have begun extensively using AI to construct simplistic stories (e.g. Washington Post's Heliograf) through NLG.

Structured data is fed into the software and is 'processed through conditional logic', so that the output sounds like a human-

---

[4] *Id.*

[5] David D. Palmer, *Text Processing*, *in* HANDBOOK OF NATURAL LANGUAGE PROCESSING 9 (Nitin Indurkhya & Fred Damerau eds., 2d. ed. 2010).

[6] David D. McDonald, *Natural Language Generation*, *in* HANDBOOK OF NATURAL LANGUAGE PROCESSING 121 (Nitin Indurkhya & Fred Damerau eds., 2d. ed. 2010).

[7] *Id.* at 122.

[8] *Id.*

[9] Corinna Underwood, *Automated Journalism – AI Applications at the New York Times, Reuters, and Other Media Giants*, EMERJ (Nov. 17, 2019), https://emerj.com/ai-sector-overviews/automated-journalism-applications/.

generated piece of content.[10] There have been instances of AI writing books of fiction as well, for instance a Jack Kerouac-inspired narration of a road trip, the utterances being described as being both 'profound' and 'nonsense'.[11] Other commonly used applications of AI are suggestive text, intelligent software to improve writing, and AI-written poetry – poetry based on conceptualization and abstraction through various figures of speech and semantic devices in a moving way. Alternatively, AI is also used in architecture design, using a concept called 'generative design', characterized by the head of an Alphabet (Google's parent company) lab as 'working with an all-powerful, really painfully stupid genie.'[12]

These applications, although scintillating, raise questions about the copyrightability of AI content, especially in cases of authorship claims and alleged infringement. The distinctiveness of human writing vis-a-vis machine-generated writing has in recent years converged, so much so that it has become increasingly difficult to tell which is which.[13]

PRE-REQUISITES FOR AUTHORSHIP IN COPYRIGHT LAW

U.S. copyright law, rather than expressly protecting copyrightable works, grants copyright to 'original works of authorship'.[14] Hence, it becomes imperative to understand the contours of authorship – in addition to the newly developed requirements for authorship created by modern scholars on causation[15] and mental effects[16] – since it subsumes the prerequisites for copyrightability in itself. .[17] The basic principles of copyright law – as is commonly known – are originality and independent creation.[18]

---

[10] Bernard Marr, *Artificial Intelligence Can Now Write Amazing Content – What Does That Mean For Humans?*, FORBES (Mar. 29, 2019, 1:23 A.M.), https://www.forbes.com/sites/bernardmarr/2019/03/29/artificial-intelligence-can-now-write-amazing-content-what-does-that-mean-for-humans/#daae83a50ab0.

[11] Brian Merchant, *When AI Goes Full Jack Kerouac*, ATLANTIC (Oct. 1, 2018) https://www.theatlantic.com/technology/archive/2018/10/automated-on-the-road/571345/.

[12] Derek Thompson, *The Spooky Genius of AI*, ATLANTIC (Sept. 28, 2018) https://www.theatlantic.com/ideas/archive/2018/09/can-artificial-intelligence-be-smarter-than-a-human-being/571498/.

[13] *See Bot or not*, BOTPOET.COM, http://botpoet.com/; *Did a Human or a Computer Write This?*, N.Y. TIMES (Mar. 7, 2015), https://www.nytimes.com/interactive/2015/03/08/opinion/sunday/algorithm-human-quiz.html.

[14] 17 U.S.C. § 102(a) (2019).

[15] Shyamkrishna Balganesh, *Causing Copyright*, 117 COLUM. L. REV. 1, 4 (2017).

[16] Christopher Buccafusco, *A Theory of Copyright Authorship*, VIRG. L. REV. 1229, 1232 (2016).

[17] Russ VerSteeg, *Defining "Author" For Purposes of Copyright*, 45 AM. U. L. REV. 1323, 1329 (1996).

[18] Feist Publications Inc. v. Rural Telephone Service Co., 111 S. Ct. 1281 (1991).

The law is clear on entitlement of independently created works to copyright protection even if they bear some similarity with previously published content. However, a copyright will not vest on these conditions alone; fixation of the work 'in a tangible medium' is also a prerequisite for copyright protection.[19]

Through judicial pronouncements, the definition of an 'author' for the purposes of copyright law has been clearly established. Though the Copyright Act itself does not attempt to define it, courts in the U.S. have examined issues of conflicts in ownership rights from the lens of authorship. In *Community for Creative Violence v. Reid*,[20] Justice Marshall defines an author as "the party who actually creates the work, that is, the person who translates an idea into a fixed, tangible expression entitled to copyright protection," Professor Nimmer also mulls over the impending reality of non-human authors but leaves the question of the interpretation of 'person' unanswered.[21] Elaborating on the general principle that the creator of the work is its author, barring cases of 'works for hire', a district court in *Lindsay* v. *R.M.S. Titanic*[22] accorded probative significance to the presence of control in the determination of the true author of the work in question.

Professor Buccafusco propounds an alternative theory for copyright authorship on the basis of the intention to 'produce some mental effect' in an audience, whereupon copyright will subsist not in the mental effect produced but the 'manner or form by which it is produced if the manner is original, minimally creative and fixed in a tangible medium'[23] Authorship here has been equated with the intention to create a mental effect. As a corollary, no component of a work that does not entail authorship can be copyrighted.[24] This stems from a conspicuous distrust for the idea/expression dichotomy used to ascertain copyrightability of components of a work, something that Professor Nimmer also subtly recognizes through his analysis of protectable ideas through contract and confidential relationships.[25]

Professor Balganesh prescribes the theory of causation, using the example of the 'monkey selfie' case,[26] relying on the clarification issued by the U.S. Copyright Office vesting copyright protection only in human authors. Balganesh infers this to mean that while Slater – the photographer responsible for leaving the camera in the vicinity of

[19] PAUL GOLDSTEIN, COPYRIGHT: PRINCIPLES, LAW AND PRACTICE 379 (1989).
[20] 490 U.S. 730 (1989).
[21] 1 DAVID NIMMER, NIMMER ON COPYRIGHT 5-5 (1997).
[22] No. 97-civ-9248 (HB), 1999 U.S. Dist. LEXIS 15837 (S.D.N.Y. 1999).
[23] *See* Buccafusco, *supra* note 16, at 1232.
[24] *Id.* at 1233.
[25] *See* NIMMER, *supra* note 21, at 19D-41.
[26] Naruto v. Slater, 888 F.3d 418 (9th Cir. 2018).

the star monkey – played some role in the creation of the photograph, that role was "insufficient to make him its author, since the real creator of the work was the monkey...Slater's failure to press the shutter button himself rendered him ineligible to be characterized as the author of the photograph. In other words, this failure was treated as having broken the causal connection to the work."[27] Moreover, he denigrates the idea of originality as a gauge for copyrightability, since it focuses almost entirely on the work itself rather than the process of creation.[28] This makes a strong case in favor of artificial authorship, since the overarching question under this theory ultimately is, in Balganesh's words, 'whether the conceiver who designed the project and was responsible for it could be characterized as its author under copyright law, despite the fact that others had actually created the project's individual components.'[29]

LOCATING ARTIFICIAL AUTHORS IN THE CONTEMPORARY LEGAL FRAMEWORK

Copyright law is rife with unique concepts that make the ownership and transfer of copyright in works possible through and between entities. In the context of a work created by a machine, these provisions and concepts in the law can be examined and applied in conjunction with one another to bring artificially-created work under the purview of the copyright framework as it currently stands.

A.   Works made for hire

The phenomenon of a 'work made for hire' prevails in an employer-employee arrangement, wherein the employer usurps the copyright in any work created by the employee in the scope of the employment. This has been embodied under Section 201 of the U.S. Copyright Act:

> "In the case of a work made for hire, the employer or other person for whom the work was prepared is considered the author for the purposes of this title, and, unless the parties have expressly agreed otherwise in a written agreement by them, owns all of the rights comprised in the copyright." [30]

The 'work for hire' doctrine can just as readily be adapted to allow works created by AI to acquire protection under copyright law, by bringing a minor modification in the law to allow the developer or the licensee of the AI to be considered the 'deemed owner' of the

---

[27] *See* Balganesh, *supra* note 15, at 4.
[28] *Id*. at 7.
[29] *Id*. at 14.
[30] 17 U.S.C § 201(b).

copyright. This has been extensively considered by Professor Hristov in his paper on the copyright dilemma in AI.[31] He states, 'a relative interpretation of terms 'employee' and 'employer' within the made for hire doctrine, as opposed to rigidly defining them in accordance with agency law, is one of the most effective ways to allow transfer of AI generated works to human authors.'[32] This method appears to be useful in vesting copyright ownership with a 'legal/natural person instead of a non-human with no legal protection'.[33]

However, the U.S. Supreme Court's decision in *Community for Creative Violence v. Reid* poses an impediment in the prospects of the 'work for hire' classification of AI-generated works. importing the law of agency, the Supreme Court opined thus:

> "In determining whether a hired party is an employee under the general common law of agency, we consider the **hiring party's right to control the manner and means by which the product is accomplished**. Among the other factors relevant to this inquiry are the skill required; the source of the instrumentalities and tools; the location of the work; the duration of the relationship between the parties; whether the hiring party has the right to assign additional projects to the hired party; the extent of the hired party's discretion over when and how long to work; the method of payment; the hired party's role in hiring and paying assistants; whether the hiring party is in business; the provision of employee benefits; and the tax treatment of the hired party."[34] (emphasis added)

The 'Reid factors', as they have come to be called, lay down stringent conditions that need to be fulfilled to establish an employer-employee relationship. Although there have been instances courts bending these rules to accommodate the rapidly-changing technological environment[35] – such as giving a leeway to startup employees and bringing their work under the doctrine irrespective of their unconventional contractual arrangements with the hiring

---

[31] Kalin Hristov, *AI and the Copyright Dilemma*, 57 J. F. PIERCE CTR. INTELL. PROP. 431(2017).
[32] *Id*. at 442.
[33] *Id*. at 447.
[34] *See* Cmty. for Creative Violence v. Reid, 490 U.S. 730 (1989).
[35] Elaine D. Ziff, *Keeping Current: Intellectual Property: The 'Work for Hire' Doctrine and Start-up Technology Companies*, BUS. L. TODAY BLOG (Apr. 30, 2011),
https://www.americanbar.org/groups/business_law/publications/blt/2011/04/keeping_current_ziff/.

company[36] – they do not stray much from the 'work for hire' doctrine explained in *Reid*, in that the control still largely remains with the employer. This regime has, however, found academic support[37] and remains the most ideal proposal for modification in the law to protect AI-created works.

B.   Joint authorship arrangement

In copyright law, joint authorship is defined as a collaborative work between two or more authors, with each creating something independently copyrightable and intending to merge their contributions into a single work.[38] The question that we grapple with here is whether an AI is an author for the purposes of vesting of copyright. In a theoretical sense, an AI-created work qualifies the *de minimis* threshold of originality prescribed under Feist.[39] Coupled with independent creation, depending upon the sources used by the AI, this work is eligible for copyright protection.

As regards the theories explored above about authorship presupposing causation and mental effects, causation between a machine and its content is an easy link. The only impediment to authorship is the non-personhood of the machine, which has been dealt with to some extent below. Nonetheless, the law could be tweaked to allow a special joint authorship setup, wherein the copyright of works created by an AI will vest in the AI or the user responsible for the output generated. This not only resolves the concern of protection of AI-created works, but also distinguishes between works created by the same AI, albeit utilization by persons other than the original programmer or user. The ideal arrangement for creation of Ai-driven content has been conjectured to be a collaboration with humans, since most works of this nature involve a human 'guide' of sorts.[40]

C.   Personhood of AI

Legal personhood has been granted to human beings and corporations largely for their capacity to make decisions as distinct entities with a modicum of rationality and deliberation. Humans i.e. adults are categorized as 'persons' by dint of being alive and overtime

---

[36] JustMed Inc v. Byce, 600 F.3d 1118 (9th Cir. 2010); Woods v. Resnick, 725 F. Supp. 2d 809 (W.D. Wis. 2010).
[37] Annemari Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author*, STAN. TECH. L. REV. 1, 26 (2012).
[38] 17 U.S.C. § 101.
[39] Feist Publications Inc. v. Rural Telephone Service Co., 111 S. Ct. 1281 (1991).
[40] Natalie Shoemaker, *Japanese AI Writes a Novel, Nearly Wins Literary Award*, BIGTHINK (Mar. 24, 2016), https://bigthink.com/natalie-shoemaker/a-japanese-ai-wrote-a-novel-almost-wins-literary-award.

having developed the cognitive ability to rationalize thoughts and express them in a coherent and comprehensible manner. Sentience has been considered to be an essential factor in according personhood to adult humans.[41] Kurki's theory of legal personhood for AI is pertinent here. According to him, only AIs that are of 'ultimate value' could be purely 'passive legal persons' for the purposes granting certain fundamental protections,[42] vesting copyright for instance.

Legal personhood for AI can have dystopian consequences, but if a regime of differential personhoods is adopted, it would eliminate technical questions of propriety and validity of copyright. Identifying AI as either active – on the basis of their capacity to act – or passive –based on the presence of a human principal – the complexity of liability considerations would stand reduced [43] In support of passive personhood for AI machines, it is known in NLG circles that computers are ultimately dumb.[44] The reason why human authorship is starkly distinct from artificial authorship is the deficiency of emotion and rhetoric in computer programs towards the people who are using them. In the absence of the richness of information, perspective, and intention that the user or programmer brings to their expression, computers have no basis for making the decisions that go into natural utterances.[45]

D.   Fair use

Fair use constitutes the most significant defense to an infringement allegation and hence, can be utilized in favor of copyrightability of AI-created content as well. Professor Nimmer dealt with the idea of computer-generated content in the style of a certain author and its infringement[46] well before its extensive utilization as seen today. He inferred that the only way for the computer to learn the stylistic devices of the author Jacqueline Susann was to 'read' her works the same way a human would. He then proceeds to draw a parallel with instances of reverse engineering, and how this act of 'reading' merely entails analyzing and thereby extracting out unprotectable elements.[47]

---

[41] VISA AJ KURKI, A THEORY OF LEGAL PERSONHOOD 9 (2019).
[42] *Id*. at 177.
[43] *Id*.
[44] *See* David D. McDonald, *Natural Language Generation*, *in* HANDBOOK OF NATURAL LANGUAGE PROCESSING 123 (Nitin Indurkhya & Fred Damerau eds., 2d. ed. 2010).
[45] *Id*.
[46] David Nimmer, *Brains and Other Paraphernalia of the Digital Age*, HARV. J. L. & TECH. 1, 44 (1996).
[47] *Id*.

It has been established that AI machines derive considerably from inputs from humans, who in the quest for creativity may utilize existing works to build on. Problems arise when the "datasets used to train AIs include copyrighted works without the permission of the rightsholder", whereupon the only defense available is fair use.[48] Views for and against fair use in artificial authorship have been examined by researchers in this field.[49] Nonetheless, The four factor test for fair use enshrined under Section 107 of the U.S. Copyright Act 1976 can be arguably read in favor of AI-generated content, due to their transformative value,[50] lack of commerciality,[51] extrapolation of the idea rather than the expression,[52] and the impact on the potential market of the original author.[53]

CONCLUSION

Artificial authorship is both a boon and a bane for society, since it brings with it a myriad possibilities of creativity while at the same time lugging along a host of legal conundrums that are not easy to resolve, irrespective of the theories or principles one chooses to adopt. Like fan fiction, AI-created content is replete with copyrightable elements which ultimately lead to the cultural enrichment of the people at large through their randomness in perspective. Primarily for this reason, it needs to be protected to further the basic tenets of copyright law. A fairly robust regime for recognizing artificially-created works simplifies liability considerations as well; one would certainly rather have someone to blame than let the work remain unprotected and fall into the public domain like the 'monkey selfie'.

In light of the suggestions made above, certain safeguards can be introduced into AI systems as well. Stringent guidelines against infringement must be formulated so as to preclude AIs from copying the expressive elements of authors it emulates. Alterations can be made to the code to ensure that the machine does not replicate a copyrighted work extensively, with a round of screening before

---

[48] Andrew Tarantola, *Modern copyright law can't keep pace with thinking machines*, ENGADGET (Dec. 13, 2017) https://www.engadget.com/2017/12/13/copyright-law-ai-robot-thinking-machines/?guccounter=1.

[49] Elise De Geyter, *Inside Views: The Dilemma of Fair Use and Expressive Machine Learning: An Interview With Ben Sobel*, INTELL. PROP. WATCH (Aug. 23, 2017), http://www.ip-watch.org/2017/08/23/dilemma-fair-use-expressive-machine-learning-interview-ben-sobel/.

[50] Campbell v. Acuff-Rose Music, 510 U.S. 569 (1994).

[51] 17 U.S.C. § 107(1).

[52] Vanderhye v. iParadigms, LLC, 562 F.3d 630 (4th Cir. 2009).

[53] *See Campbell*, 510 U.S. at 590.

publication to confirm the non-infringing nature of the AI-generated work.

Copyright law, as it currently stands, is grossly inadequate to account for such technological advancements. Drawing from the theories of authorship explored above with respect to AI-generated works, a new test for copyrightability can be envisioned, doing away with the requirement of independent creation since it is increasingly difficult to dissociate authors, irrespective of their sentience, from their 'inspirations'. A policy decision needs to be made between the convenience of vesting copyright in an AI machine and its programmer or user, depending upon factors such as efficiency, fairness and the overall contribution to the society at large.