

Osgoode Hall Law School of York University **Osgoode Digital Commons**

All Papers

Research Papers, Working Papers, Conference **Papers**

2022

The Al-Copyright Challenge: Tech-Neutrality, Authorship, and the **Public Interest**

Carys Craig Osgoode Hall Law School of York University, ccraig@osgoode.yorku.ca

Follow this and additional works at: https://digitalcommons.osgoode.yorku.ca/all_papers



Part of the Computer Law Commons, Intellectual Property Law Commons, and the Internet Law

Commons

Repository Citation

Craig, Carys, "The Al-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest" (2022). All Papers. 360.

https://digitalcommons.osgoode.yorku.ca/all_papers/360

This Article is brought to you for free and open access by the Research Papers, Working Papers, Conference Papers at Osgoode Digital Commons. It has been accepted for inclusion in All Papers by an authorized administrator of Osgoode Digital Commons.

The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest

Carys J. Craig

Abstract

Many of copyright's core concepts—from authorship and ownership to infringement and fair use—are being challenged by the rapid rise of generative AI. Whether in service of creativity or capital, however, copyright law is perfectly capable of absorbing this latest innovation. More interesting than the doctrinal debates that AI provokes, then, is the opportunity it presents to revisit the purposes of the copyright system in the age of AI. After introducing the AI-copyright challenge in Part 1, Part 2 considers the guiding principles and normative objectives that underlie—and so ought to inform—copyright law and its response to AI technologies. It proposes a substantive approach to techneutrality aimed at achieving normative equilibrium in the face of technological disruption. Applying this frame—with its corresponding emphasis on authorship and the public interest—Part 3 goes on to explain why AI-generated outputs are therefore uncopyrightable and AI-training inputs are non-infringing.

Keywords: Artificial Intelligence, technological neutrality, copyright, authorship, expression, public domain

Contents

- 1. INTRODUCTION
- 2. TECHNOLOGICAL NEUTRALITY AND THE PURPOSES OF COPYRIGHT
 - 2.1 Competing Conceptions of Technological Neutrality
 - 2.2 Competing Conceptions of Copyright's Purpose
- 3. COPYRIGHT AND THE AI CHALLENGE
 - 3.1 AI "Authorship"
 - 3.2 The Myth of the "AI Author"
 - 3.2.1 Antropomorphism and Expressive Agency
 - 3.2.2 Romanticizing the "AI Author"

3.3 Authorship, Creative Progress, and the Public Interest 3.3.1 Encouraging Creative Practices (Not Products) 3.3.2 Recognizing the Limits of Copyright's Reach

4. CONCLUSION

1. INTRODUCTION

For readers of this volume, it hardly needs to be said again that recent developments in Artificial Intelligence (AI) are poised to disrupt our intellectual property (IP) system. Many of the core concepts and assumptions that underlie our copyright and patent laws, especially, are already being tested and disturbed by the rapid rise of AI and its generative capabilities. In the public theatre of IP policy, our old heroes – copyright's author and patent's inventor – are at risk of being outshone by their intrepid AI understudies. As AI-generated works and inventions increasingly take center stage, conventional constructs like the "original work of authorship" or the "nonobvious invention" suddenly look less solid, and their attendant bundle of rights more tenuous. And so familiar IP policy questions about incentives and rewards, progress, and the public interest, present themselves anew in the face of fresh promises, mounting apprehensions, and the pervasive hyperbole that has accompanied the arrival of AI.

This is the unfolding AI copyright drama into which governments, courts, and commentators are increasingly being drawn. There is a sense of urgency behind the questions that are now repeatedly posed in public consultations, private litigation, and behind the closed doors of lawmakers and lobbyists around the world. How will the copyright system respond to the rise of AI? Should the law protect mass-produced AI-outputs as copyrightable works? Might there be copyright liability for copies fed into or produced by AI systems, where would it fall, and what exceptions should apply? More broadly, what will this new technological revolution mean for artists and creators, owners and users, and the copyright system as we know it?

The urgency may well be overblown, and the hyperbole unhelpful at best, but such questions must be asked, and, in time, they will be answered one way or another. Given the many paradigm-changing technologies that have come along since copyright's 18th century inception, from the Player Piano Roll to the World Wide Web, we can reasonably predict that copyright will once again

adapt and prevail. Whether in service of creativity and culture, or simply in service of capital, the copyright system is perfectly capable of absorbing this latest innovation and continuing about its business as it has so many times before. Our now over-rehearsed law and policy debates can quite readily offer up answers to the doctrinal conundrums that AI appears to present. But to what end? More interesting than the doctrinal debates that AI provokes, then, is the opportunity that it presents to revisit the purposes of the copyright system—its functions and its fictions—and to reconsider its rationale in the age of AI. This is where the copyright policy debate about AI should begin; ultimately, how we respond to the AI-copyright challenge will depend on why we are responding, what rights we are choosing to protect, and in service of whose interests.

In this spirit, Part 2 of this chapter considers the normative objectives that underlie—and so ought to inform—copyright law and its response to technological change. Highlighted here is the guiding principle of "technological neutrality" and its potential importance in shaping the law's trajectory into the future. Part 3 then turns to survey some of the quandaries presented by AI through this theoretical frame, looking first at the question of AI "authorship" and the public interest, and then turning briefly to the matter of potential infringement in respect of AI's inputs. The chapter concludes, first, that AI generated works are not works of authorship and so ought to be left "where they lie" in the public domain; it further concludes that the usage of works as training data does not involve use of the work of authorship as such and should therefore require no defense. These conclusions, it is argued, are consistent with—indeed required by—the technologically-neutral normative objectives of copyright law and the social values they reflect.

Without a principled commitment to the public purposes of copyright, economic interests (propelled by neo-liberal logic and techno-romantic rhetoric) will surely set our course towards increased commoditization and private control. We may then miss out on the true promise of AI to contribute to a vibrant public domain, instead hindering or harming the very creative processes and cultural exchange that copyright law is meant to foster.

2. TECHNOLOGICAL NEUTRALITY AND THE PURPOSES OF COPYRIGHT

The arrival of AI—or, at least, its recent flourishing—has the feel of something new, something ground-breaking. The temptation is therefore to treat it as wholly unprecedented, and so to rush into action, making reforms to clear the way for this exciting innovation; the mistake is to forget that no matter how

novel this particular technological innovation, we have been here many times before. The modern copyright system was itself a response to technological innovation, after all—the invention and proliferation of the printing press. Since that time, history has witnessed a continual evolution of communication technologies, from the creation of lithographs and engravings to sound recordings and cameras, and from the first radio broadcasts and phone calls to the invention of computers, photocopiers, fax machines, and the Internet. And with each innovation, copyright laws that were originally written to control the printing of physical books have adapted and expanded, embracing new modes of cultural production, and supporting new means of economic exploitation. Even if the technology is novel, then, the question of how the law ought to respond to technological change is not at all new.

A common response to the question of how the law should treat new technologies is simply that the law should be technologically neutral. That is, copyright law should be written and developed to be independent of any specific technology and should therefore continue to apply equally across technologies as they emerge, without favoring or discriminating between new and old. There is an obvious appeal to this approach for policy makers in the digital age. From a practical perspective, it presents the promise of sustainable laws in a time of rapid technological change, "future-proofing" the copyright system to some degree by permitting old laws to apply seamlessly to new technologies. Happily, it may also excuse lawmakers from following the twists and turns of each technological development as it occurs, and likely produces more comprehensible legislation for those non-experts who are expected to abide by it. This answer is also intuitively attractive from a principled perspective: neutrality, equality, and non-discrimination in the law are almost always perceived as intrinsically laudable goals. It is not surprising, then, that technological neutrality is widely hailed as an unquestionably good thing—like "motherhood and apple pie," as one commentator has observed—but typically with little in the way of explanation or justification. The problem, of course, is that (as with equality, non-discrimination, and neutrality writ large) what the principle of technological neutrality means in any particular application is highly subjective and necessarily context dependent.

¹ See Entm't Software Ass'n v. Soc'y of Composers, Authors and Music Publishers of Can., 2012 S.C.C. 34 at ¶ 5, 9 [SOCAN]. See also Robertson v. Thomson Corp., 2006 S.C.C. 43; Soc'y of Composers, Authors and Music Publishers of Can. v. Bell Can., 2012 S.C.C. 36 at ¶ 43; Can. Broad. Corp. (CBC) v. SODRAC 2003, Inc., 2015 S.C.C. 57 [SODRAC].

² Chris Reed, *Taking Sides on Technology Neutrality*, 4:3 SCRIPTed 263, 265 (2007).

In fact, technological neutrality has many shades of meaning, and different meanings can produce differing applications with more or less desirable results.³ I have suggested elsewhere that these different meanings range from formal non-discrimination between technologies, at one end of the spectrum, to substantive equality-in-effect at the other.⁴ It is worth identifying these different approaches and their limits here, as our question of how copyright law should treat AI is directly informed by the relationship we perceive between technological innovation and good copyright policy-making.

2.1 Competing Conceptions of Technological Neutrality

At one end of the spectrum, a narrow, formal version of neutrality simply extends the law as it is to new technological activities or processes. On this logic, for example, a copy is a copy, whether it is made by hand, photocopier, cell phone, or a computer's Random-Access Memory, and whether it is printed, saved, or temporarily cached.⁵ The problem with this restrictive approach is that, in the name of neutrality, it treats alike activities that are, in fact and effect, very different. What it produces in practice, then, can be a vastly different outcome for old and new technologies, and thus their *substantively unequal* treatment. Whereas the mere act of reading a physical newspaper, for example, does not involve copying and so would not implicate a copyright owner's rights, browsing a newspaper online creates temporary digital reproductions and potentially cached copies. If every digital copy, however fleeting, were to be regarded as a reproduction like any other, then the act of reading online would suddenly implicate copyright in a way that reading a physical copy never did. Similarly, if we treat every digital reproduction made

³ For an excellent discussion of various meanings attached to the principle, *see* Bert-Japp Koops, *Should ICT Regulation be Technology-Neutral?*, *in* STARTING POINTS FOR ICT REGULATION: DECONSTRUCTING PREVALENT POLICY ONE-LINERS (Bert-Jaap Koops, Miriam Lips, Corien Prins & Maurice Schellekens, eds., 2006). *See also*, Deborah S. Tussey, *Technology Matters: The Courts, Media Neutrality, and New Technologies*, 12 J. INTELL. PROP. L. 427 (2005); Brad A. Greenberg, *Rethinking Technology Neutrality*, 100 MINN. L. REV. 1495, 1498 (2016).

⁴ Carys J. Craig, *Technological Neutrality: (Pre)Serving the Purposes of Copyright Law, in* The Copyright Pentalogy: How the Supreme Court of Canada Shook the Foundations of Canadian Copyright Law (Michael Geist, ed., 2013); Carys Craig, *Technological Neutrality: Recalibrating Copyright in the Information Age,* 17 Theoretical Inquiries L. 601 (2016). *See also,* Gregory R. Hagen, *Technological Neutrality in Canadian Copyright Law, in* The Copyright Pentalogy: How the Supreme Court of Canada Shook the Foundations of Canadian Copyright Law (Michael Geist, ed., 2013); Cameron J. Hutchison, *Technological Neutrality Explained & Applied to CBC v. SODRAC* 13:1 C.J.L.T. 101 (2015).

⁵ See, e.g., SODRAC, supra note 1.

in the process of an online transmission as though it were a copy for copyright purposes, the act of sending a digital copy online is suddenly very different than putting a physical one in the mail.⁶ Treating things equally regardless of technological difference can in fact discriminate between technologies and produce unequal treatment in result. (The same critique has, of course, been made of formal equality and neutral treatment in other contexts.)⁷

Moving along the spectrum, a preferable ("intermediate") approach to technological neutrality focuses on the *effects* of particular technological activities or processes, demanding that the law apply equally to *functionally equivalent* actions. Tech-neutrality, in this version is about the equal treatment in law of effectively analogous activities.⁸ Thus, a digital streaming service, for example, may be subject to the same legal rules as a cable broadcaster if their activities have the same effect, notwithstanding the different technological processes involved.⁹ To take our example above, reading online should be treated by law in the same way as a reading a physical paper, because the act is functionally the same no matter the medium. Similarly, as Canada's Supreme Court has held, selling a digital copy online should implicate the same copyright interests—no more and no less—as selling a physical copy over the counter.¹⁰

⁶ The act of digital broadcasting would involve considerably more copies that would traditional broadcasting, and so could potentially involve far greater costs in the form of copyright tariffs.

⁷ I am referring here to a long and deep body of scholarship and advocacy addressing formal equality in the context of human rights and discrimination law. *See*, *e.g.*, Sandra Fredman, *Substantive Equality Revisited*, 14 INT'L J. OF CONST. L. 712, 713 (2016), citing, *e.g.*, Peter Westen, *The Empty Idea of Equality*, 95 HARV. L. REV. 537, 542 (1982); Law v. Can. [1999] 1 S.C.R. 497 (Can.); Prinsloo v. Van der Linde (CCT4/96), [1997] Z.A.C.C. 5 (S. Afr.); CEDAW Committee, General Recommendation No. 25: On Temporary Special Measures (2004) CEDAW/C/GC/25.

⁸ See, e.g., Am. Broad. Co. v. Aereo, Inc., 134 S. Ct. 2498 (2014) [Aereo] (one-to-one transmission of a digitized signal was sufficiently "cable-like" to constitute public performance as a cable transmission would). See also Ryan Abbott, The Reasonable Robot (2020) (proposing a concept of "legal neutrality" that appears to focus on functional equivalence and analogy). For a more detailed critique of Abbott's conception of legal neutrality, see Carys Craig, The Relational Robot, Jerusalem Rev. of Legal Studies (forthcoming, 2022).

⁹ The Aereo case, *supra* note 8, reveals the difficulty with this kind of reasoning by technological analogy: in fact, in my opinion, the activity at issue bore greater resemblance to the provision of a VCR to facilitate the making of lawful personal copies. *See Brief Amici Curiae of Law Professors and Scholars in Support of Respondents*, at 22-23, Am. Broad. Co., 134 S. Ct., http://isp.yale.edu/sites/default/files/American%20Broadcasting%20Companies%20v.%20Aere o.pdf (citing Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984)).

¹⁰ SOCAN, *supra* note 1.

This version of tech-neutrality is better; but still, its reliance on analogical reasoning and functional equivalence present cause for concern. Drawing analogies between old and new technologies ("x does the same as y") is a more subjective endeavor than it might seem. A cloud-based TV recording service may be convincingly analogized to a cable TV service, for example, as it was by the US Supreme Court in the *Aereo* case, when arguably the better analogy—and one that would have supported an entirely different ruling—would have been to a personal VCR.

In the strategic search for functional equivalence, critical differences between technological processes can be overlooked, and comparisons can be stretched to justify a kind of willful technology-blindness. Analogies can therefore produce false equivalents while also obscuring the disruptive force of technological change. Treating the personal recording service as akin to a cable service undervalued the distributive potential of the novel technology at issue in Aereo for consumers seeking to access works, and it re-established the technology provider as the appropriate point of control (thereby reinforcing the distribution rights of content industry incumbents). To take another example, a teacher sharing digital copies with students through a content management platform may be doing something functionally equivalent to handing out photocopies in class; if we simply regard these acts as analogous and regulate them accordingly, however, we overlook both the practical implications for copyright owners of having perfect, reproducible files circulating online and the enormous opportunities for improved access and education that the technology now affords.

Some technologies are paradigm-shifting—they simply "change the game." In such cases, we cannot assume that neutral legal treatment will produce a substantively equivalent legal effect. When the risks and opportunities shift significantly with the affordances of novel technologies, the costs and benefits of copyright control should be reevaluated. A myopic focus on the comparable effects of specific technological actions and processes may miss this bigger picture. The weakness of the intermediate approach to technological neutrality, then, is precisely that it purports to solve the policy conundrum by treating the new thing it *as if* it were something else. The mistake is to assume that treating like things alike will produce equivalent outcomes. In the copyright system as elsewhere, equal treatment without due attention to

¹¹ Ian Kerr & Katie Szilagyi, Asleep at the switch? How killer robots become a force multiplier of military necessity, in ROBOT LAW 333, 349 (Ryan M. Calo, Michael Froomkin & Ian Kerr, eds., 2016).

context and difference can disguise systemic advantages and disadvantages behind a veil of neutrality.

And so, at the furthest end of the spectrum, the most expansive vision of technological neutrality looks beyond the equal treatment of functional equivalents to the systemic effects of new technologies and the balance of rights and interests that will shape their use. The focus is not only on the comparable effects of specific technological actions or processes, but also on the broader effects of the technology and its affordances. This substantive principle of technological neutrality is therefore less concerned with ensuring the consistent treatment of analogous activities than with consistently advancing the objectives of the law in new technological contexts. In other words, the neutrality that it supports is not just a technical or legal neutrality but a *normative neutrality*: Whether through the interpretation and application of existing law or its reform, the principled goal is presumed to be normative equilibrium in the face of technological change.

The principle of technological neutrality can be a brain scratcher. Technology itself, of course, is not neutral; nor is the law. And to apply the law neutrally to novel or different technologies will potentially produce new and unequal outcomes. A more substantive approach to tech-neutrality therefore focuses on the technology's implications for the purposes that the law serves. It is these purposes—and the values they reflect—that should be regarded as presumptively tech-neutral, transcending the technical capabilities of the moment. So, what does this mean for copyright law?

2.2. Competing Conceptions of Copyright's Purpose

If the above is accepted, then the task that awaits us is to adjust copyright law and its application to today's technological context—the AI age, if you will—in a manner that continues to advance the normative objectives of the copyright system. Defining these objectives is by no means a straight-forward task, of course; copyright's rationale has been a matter of debate since its inception. For our purposes, however, it is sufficient to point to the widely accepted justification for the copyright system: it is a means by which to encourage or incentivize authorship and the dissemination of original works, which ultimately serves the public interest.

In its 1709 formulation, the stated purpose of the first modern copyright law in the United Kingdom was "the encouragement of learning," which it sought to achieve "by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies."¹² The United States Constitution subsequently captured the purpose of copyright as being "to promote the Progress of Science and useful Arts." Similarly, this was to be done "by securing for limited Times to Authors...the exclusive Right to their...Writings."¹³ In both formulations, the grant of rights to authors (and, through them, owners) is presented as means to a larger social end. As explained by the US Supreme Court in the foundational *Feist* case, "The primary objective of copyright is not to reward the labor of authors, but '[t]o promote the Progress of Science and useful Arts."¹⁴ More recently, in *Google LLC v. Oracle*, the Supreme Court spoke of the "creative 'progress' that is the basic constitutional objective of copyright itself."¹⁵ (We will return to consider this idea of "progress" in Part 3.3 below.)

The Supreme Court of Canada has also had occasion of late to revisit copyright's purpose, which it has described as "a balance between promoting the public interest in the encouragement and dissemination of works of the arts and intellect and obtaining a just reward for the creator." Most recently, that Court explained:

[J]ust rewards for copyright creators provide necessary incentives, ensuring that there is a steady flow of creative works injected into the public sphere....A proper balance ensures that creators' rights are recognized, but authorial control is not privileged over the public interest.¹⁷

Whether authors are entitled to just rewards independently of the public interest will always be a point of contention. It is not one we need to resolve here.¹⁸ It is sufficient to embrace the idea of a copyright balance, which in turn

¹² Statute of Anne, 1710, 8 Ann. C. 19 (Eng.). The statute's full title was "An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies, during the Times therein mentioned."

¹³ U.S. CONST. art. I, § 8, cl. 8.

 $^{^{14}}$ Feist Publ'ns Inc. v. Rural Tel. Serv. Co. Inc., 499 U.S. 340, 349–350 (1991) [Feist], quoting id.

¹⁵ Google LLC v. Oracle America, Inc., 593 U.S. ____ 25 (2021) [Google].

¹⁶ Théberge v. Galerie d'Art du Petit Champlain Inc., 2002 S.C.C. 34 at ¶ 11-12, 30.

¹⁷ York Univ. v. Canadian Copyright Licensing Agency (Access Copyright), 2021 S.C.C. 32 at ¶ 93–94 [Access Copyright], citing Mya Tawfik, *History in the Balance: Copyright and Access to Knowledge*, in FROM "RADICAL EXTREMISM" TO "BALANCED COPYRIGHT": CANADIAN COPYRIGHT AND THE DIGITAL AGENDA 69 (Michael Geist, ed., 2010).

¹⁸ I have argued elsewhere against the idea of an author's natural right to copyright as a reward for intellectual labour: *see* Carys J. Craig, *Locke, Labour, and Limiting the Author's Right: A Warning Against a Lockean Approach to Copyright Law*, 28 QUEEN'S L. J. 1 (2002), online: https://ssrn.com/abstract=2078157.

means recognizing that "[c]opyright law has public interest goals." Specifically, "increasing public access to and dissemination of artistic and intellectual works, which enrich society and often provide users with the tools and inspiration to generate works of their own, is a primary goal of copyright." Ultimately, then, copyright is a system of state-granted entitlements to encourage creative expression and learning; authors' or creators' rights must therefore be balanced with users' rights and the public interest to support a flourishing public domain.²¹

The concept of balance that has emerged so explicitly in Canadian copyright jurisprudence is helpfully intertwined with the principle of technological neutrality as I have described it. Inherent in the very concept of balance is the need to adjust the weight and distribution of rights and interests in order to maintain a consistent equilibrium as conditions change.²² Thus, as the Court's robust vision of technological neutrality suggests, "the traditional balance between authors and users should be preserved in the digital environment."²³ Justice Abella has since explained:

The question...is how to preserve [the balance that best supports the public interest in creative works] in the face of new technologies that are transforming the mechanisms through which creative works are produced, reproduced and distributed.... The answer to this challenge...lies in applying a robust vision of technological neutrality as a core principle of statutory interpretation under the *Copyright Act*.²⁴

 $^{^{19}}$ Access Copyright, *supra* note 17, at ¶ 91. *See also* CCH v. Law Society of Upper Can., 2004 S.C.C. 13 at ¶ 23 [CCH Canadian].

²⁰ *Id*. at ¶ 92.

²¹ See Carys J. Craig, Copyright, Communication and Culture: Towards a Relational Theory of Copyright Law 52 (2011).

²² Robertson v. Thompson, 2006 S.C.C. 43 at ¶ 79 (Abella, J., dissenting) (citing Michael Geist, Our Own Creative Land: Cultural Monopoly & The Trouble with Copyright 9 (2006), cdn.michaelgeist.ca/wp-content/uploads/2006/05/hhlo6_Online_Book.pdf.

²³ SOCAN, *supra* note 1, at ¶ 7–8 (citing Carys Craig, *Locking Out Lawful Users: Fair Dealing and Anti-Circumvention in Bill C-32*, in FROM 'RADICAL EXTREMISM' TO 'BALANCED COPYRIGHT': CANADIAN COPYRIGHT AND THE DIGITAL AGENDA 177 (Michael Geist, ed., 2010) [*Locking Out Lawful Users*].

 $^{^{24}}$ SODRAC, *supra* note 1, at ¶ 147–148. *See also* SODRAC, *supra* note 1, at ¶ 149 (Abella, J., dissenting) (identifying Parliament's intent that the Copyright Act "must adapt and apply to new technologies in a manner that maintains the careful balance between creators and users that underpins the *Act* as a whole.")

The US Supreme Court recently made a similar pronouncement about balance, technological change, and copyright's purpose when interpreting the fair use provision in *Oracle*:

[W]e have understood the provision to set forth general principles, the application of which requires judicial balancing, depending upon relevant circumstances, including "significant changes in technology." Sony Corp. of America v. Universal City Studios, Inc., 464 U. S. 417, 430 (1984); see also Aiken, 422 U. S., at 156 ("When technological change has rendered its literal terms ambiguous, the Copyright Act must be construed in light of its basic purpose").²⁵

To be sure, technological neutrality in this sense does not dictate a definitive answer to any particular legal question. If the consistency sought is not consistency in the application of the law but rather in the steady pursuit of the law's purpose, then the answer to any policy question will depend upon how this purpose is defined, propelling us straight into subjective balancing exercises and age-old debates around competing conceptions of authorship, entitlement, and the public interest. Approaching the question in this way does, however, relieve us of the task of examining the internal mechanics of technological processes, and it allows us to escape intractable debates about which analogies most aptly apply to describe new technology-enabled activities. More to the point, it reminds us of the big question we *should* be asking: Given the new realities of the current technological environment, what rights should the law recognize—and subject to what limits—if it is to continue to advance its objectives?

In my view, it is the search for this answer that should inform copyright's response to recent developments in AI. If AI is the game-changer that we have been led to believe it is, then it will not be sufficient to simply apply the law as it is to the new technology—treating alike things that are fundamentally different. Nor will it do to simply apply the existing law to the activities of an AI wherever they appear to be functionally equivalent in effect to an act traditionally performed by a human. Rather, we must analyze the function and effects of AI technologies with a view to their implications for the copyright system as a whole. Efforts to respond to this latest technological (r)evolution should aim to preserve the balance that "best supports the public interest in

²⁵ Google, *supra* note 15, at 14.

creative works" as this is "the central purpose" of copyright law. ²⁶ A substantive vision of tech-neutrality thus operates as a framing principle, guiding a purposive application of law to our new technological realities. It necessitates not willful technology-blindness, but quite the opposite: a clear-sighted recognition of the disruptive force and political significance of AI for the "creative progress" that copyright is intended to foster.

3. COPYRIGHT & THE AI CHALLENGE

There are, in current debates around the implications of AI, two doctrinal and policy questions that seem to loom large for copyright law. First is the matter of AI-generated works (outputs) and where—or whether—they fit within the copyright scheme, premised as it is on *authorship* of the heretofore human variety. Second is the question of AI-training data-sets (inputs) and whether or not the processes involved in machine-learning implicate—and potentially infringe—copyright, premised as it is on the exclusive right to *reproduce* works of authorship. Whereas other relevant and important policy issues such as, say, the appropriate allocation of liability for unlawful actions/outputs or the automation of enforcement are also raised by the arrival of AI, these are legal issues that traverse a multitude of fields.²⁷ Authorship and infringement for copying, however, present quandaries that are thoroughly internal to copyright law and its logic. As such, these are the issues to which we now turn, armed with our guiding principle of normative neutrality and a clearer sense of copyright's purpose.

3.1 AI "Authorship"

"The rise of the machines is here," it has been said, "but they do not come as conquerors, they come as *creators*." ²⁸

Recent high-profile examples of AI-generated works span the full range of human cultural endeavor from music to film, and from literature to the visual arts. *Jukebox*, for instance, is a machine-learning model capable of generating

²⁶ SODRAC, *supra* note 1, at ¶ 146.

²⁷ For further discussion of these issues, *see* Carys J. Craig, *Copyright and Artificial Intelligence*, in AI and the Law in Canada 33–37 (Florian Martin-Bariteau & Teresa Scassa, eds., 2021) [*Copyright and AI*].

²⁸ Andres Guadamuz, *Artificial Intelligence and Copyright*, 5 WIPO MAGAZINE 14, 17 (2017), https://www.wipo.int/wipo_magazine/en/2017/05/article_ooo3.html (emphasis added).

music that imitates different styles and artists, even incorporating singing in natural-sounding voices.²⁹ TalkToTransformer is an AI language generator, created by Canadian engineer Adam King using the OpenAI GPT-2 technology, which can write articles when prompted with a headline or complete short stories when fed the first line.³⁰ Dio is a sculptural work from artist Ben Snell, which was generated by an machine-learning algorithm trained on a dataset of classical sculptures (and then made out of the pulvarised dust of the machine that designed it!).31 Sunspring is an award-winning science fiction film, the screenplay and music for which was written by the self-named Benjamin—a neural network trained on hundreds of sci-fi movies from the 1980s and '90s.32 Perhaps most famously, *Portrait of Edmond Belamy* is an AI-generated painting that sold at Christie's Auction House for just shy of half a million dollars in 2018, signaling "the arrival of AI-generated art on the world auction stage."33 Of course there are also many more mundane examples: machine-generated text produced by online chatbots; the verbal responses of digital assistants like Alexa or Siri; the manipulations of "selfies" and visual images produced by apps like Google's Deep Dream AI;34 and the proliferating AI-composed music playlists of techno-beats and ennui-inducing electronica to be found on

-

²⁹ Prafulla Dhariwal et al., *Jukebox: A Generative Model for Music*, arXiv (April 30, 2020), https://arxiv.org/pdf/2005.00341.pdf. Curated examples of Jukebox's outputs are available online: https://openai.com/blog/jukebox/.

³⁰ James Vincent, *Use This Cutting-Edge AI Text Generator to Write Stories, Poems, News Articles, and More,* The Verge (May 13, 2019), https://www.theverge.com/tldr/2019/5/13/18617449/ai-text-generator-openai-gpt-2-small-model-talktotransformer; James Vincent, *OpenAI's New Multitalented AI Writes, Translates, and Slanders,* The Verge (Feb. 14, 2019), https://www.theverge.com/2019/2/14/18224704/ai-machine-learning-language-models-read-write-openai-gpt2. TalkToTransformer.com is now being offered as a paid service online: https://inferkit.com/.

³¹ James Vincent, *This AI-generated sculpture is made from the shredded remains of the computer that designed it*, The Verge (Apr. 12, 2019), https://www.theverge.com/tldr/2019/4/12/18306090/ai-generated-sculpture-shredded-remains-ben-snell-dio. High resolution views of Snell's sculptures are provided on the website of Blackbird Gallery, online: https://www.blackbird.gallery/artists/54-ben-snell/works/.

³² Oscar Sharp, Sunspring, YOUTUBE (June 9, 2016), https://youtu.be/LY7x2Ihqjmc.

³³ Christie's, *Is Artificial Intelligence Set to Become Art's Next Medium*?, Christie's (December 12, 2018), https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx.

³⁴ Online: https://deepdreamgenerator.com/. See also, e.g., Hilary Brueck, Google's Computers are Making Thousands as Artists, Fortune (Mar. 1, 2016), https://fortune.com/2016/03/01/google-deepdream-art/.

Spotify³⁵ (alas, not every AI has the training necessary to finish Schubert's Unfinished Symphony).³⁶

As with any new wave of innovative production with commercial applications and potentially valuable, commodifiable outputs, the question in respect of such AI-generated works quickly becomes, "who owns it?" A more apt starting point for copyright purposes would be, "who is the author?" For, unless we can answer that prior question, we cannot know whether there is even an original copyrightable work to be owned, never mind to whom it belongs. In respect of the *Portrait of Edmond Belamy*, for instance, it was queried whether the Parisian art-group that collated and fed data to the AI rightly laid claim to the portrait, or whether the credit should have gone to Robbie Barrat—the person who wrote the open-source code on which the AI operated.³⁷ Meanwhile, the place on the canvas where the artist's name would traditionally appear contained a portion of the algorithm itself.

As things stand, works that are autonomously generated by machines or AI are not copyrightable in most jurisdictions around the world, which means that they currently belong in the public domain. In the United States, Australia, Canada, and most of Europe, the fact that copyright attaches only to "original" works of authorship has been interpreted to mean that they must originate from a human author—a natural person with some direct intellectual involvement in the resulting expression.³⁸ In some jurisdictions, though—the

35 See, e.g., the Barbican Centre playlist, https://open.spotify.com/playlist/6xbMspVDXiIUTCncyBzoT7.

³⁶ Online: https://consumer.huawei.com/au/campaign/unfinishedsymphony/.

³⁷ See Jason Bailey, *The Truth Behind Christie's \$432K AI Art Sale*, Artnome (Oct. 29, 2018), https://www.artnome.com/news/2018/10/13/the-truth-behind-christies-432k-ai-art-sale. See also Amanda Turnbull, *The Price of AI Art: Has the Bubble Burst?*, The Conversation (Jan. 6, 2020), https://theconversation.com/the-price-of-ai-art-has-the-bubble-burst-128698.

³⁸ The point of copyright's threshold originality doctrine is to identify an authorial act. *See*, *e.g.*, Feist, *supra* note 14; U.S. Copyright Office, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 306 (3d ed. 2017) "Because copyright law is limited to 'original intellectual conceptions of the author,' the [Copyright] Office will refuse to register a claim if it determines that a human being did not create the work."; IceTV Pty. Ltd. v. Nine Network Austl. Pty. Ltd., [2009] H.C.A. 14; Telstra Corp. Ltd. v. Phone Directories Co. Pty. Ltd., [2010] F.C.A. 44, *affd* [2010] F.C.A.F.C. 149; Acohs Pty. Ltd. v. Ucorp Pty. Ltd., [2010] F.C.A. 577, *aff'd* [2012] F.C.A.F.C. 16; CCH Canadian, *supra* note 19 (finding that "an original work must be the product of an author's exercise of skill and judgment"); Telstra Corp. Ltd. v. Phone Directories Co. Pty. Ltd., [2010] F.C.A.F.C. 149 at ¶ 133–34, 137 (requiring that an original work must be the product of human authorship); Eva-Maria Painer v. Standard Verlags GmbH and Others, C-145/10, [2012] E.C.R. I-12594 at I-12622. *See generally* Daniel J. Gervais, *The Machine As Author* (2020) 105:5 IOWA L. REV. 2053. For a detailed discussion of Canadian copyright and AI, *supra* note 27.

United Kingdom, Hong Kong, India, Ireland, and New Zealand—specific provisions have been enacted in respect of computer-generated works, deeming the author to be the person who makes the arrangements necessary for their creation.³⁹ (Whether such works meet the threshold requirements of copyrightability is often still an open question).

It is clear, however, that we are now at a critical moment in the evolution of law when policymakers around the world are turning their attention to this question anew, asking whether they should enact legal fictions to ensure the copyright protection of AI-generated works; and wondering whether, if they do not, they will fall behind their international counterparts in the global competition over AI-innovation. Several high-profile public consultations and policy reports over the past few years—in the European Union, the United Kingdom, and now in Canada—have revisited these established limits of copyrightability, putting questions to the public, industry, and experts about how the law should respond to the arrival of AI-generated works circulating as products in our cultural sphere.⁴⁰ Would extending copyright to such works reflect an appropriate reward for the original, creative efforts and investment of the people responsible for the AI? Would it encourage the kind of authorial, creative practices that advance the "progress" at which copyright is aimed?

My answer to both questions is no, it would not. In fact, far from an incentive to encourage would-be authors, the protection of AI-generated works with

https://www.uwindsor.ca/law/sites/uwindsor.ca.law/files/final ai submission canadian ip sc holars.pdf.

³⁹ See, e.g., Copyright, Designs and Patents Act 1988 (UK), c. I, s. 9(3).

⁴⁰ See, e.g., Government of Canada, A Consultation on a Modern Copyright Framework for Artificial Intelligence the Internet Things, of https://www.ic.gc.ca/eic/site/693.nsf/eng/00316.html; Eur. Parl. Doc. (2020/2015(INI)) Intellectual property rights for the development of artificial intelligence technologies; U.K. Government Response to call for views on artificial Intelligence and Intellectual Property (March https://www.gov.uk/government/consultations/artificial-intelligence-andintellectual-property-call-for-views/government-response-to-call-for-views-on-artificialintelligence-and-intellectual-property; WIPO Secretariat, Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence 7-8 (May 21, 2020), https://www.wipo.int/edocs/mdocs/mdocs/en/wipo ip ai 2 ge 20/wipo ip ai 2 ge 20 1 rev. pdf. My responses to the Canadian Consultation Paper are contained in two submissions: S. Flynn, L. Guibault, C. Handke, J. Vallbé, M. Palmedo, C. Craig, M. Geist & J.P. Quintais, Submission to Canadian Government Consultation on a Modern Copyright Framework for AI and the Internet of Things (Sept. 17, 2021), https://ssrn.com/abstract=3952238; and Carys Craig, Bita Amani, Sara Bannerman, Céline Castets-Renard, Pascale Chapdelaine, Lucie Guibault, Gregory Hagen, Cameron Hutchison, Ariel Katz, Alexandra Mogyoros, Graham Reynolds, Anthony D. Rosborough, Teresa Scassa & Myra Tawfik, Submission by IP Scholars Copyright and Artificial Intelligence,

exclusive private copyright control would be an unnecessary obstacle to authorship and creative progress. That is because, first, what the AI is doing when it generates a works is *not authorship*—in an ontological sense, the AI is not an "author" with rights or interests in the copyright balance; and second but relatedly, in a teleological sense, the mere proliferation of AI-generated products does not reflect the kind of creative *process* or *practice* that copyright is intended to foster.

3.2. The Myth of the "AI-Author"

Copyright protection is not hard to get. Machine-generated works may, at least for now, lack the aesthetic quality or conceptual coherence of their humanmade counterparts (TalktoTransformer's stories rarely make much sense; the Sunspring script was unintelligible nonsense; and Edmond Bellamy had no discernible nose!)—but there is no aesthetic quality bar to copyrightability. The originality threshold requires only a modicum of creativity or, in the Canadian iteration, non-trivial skill and judgment.⁴¹ Had humans created any of these works, copyright would attach to them. But while AI-generated outputs may facially resemble original works of authorship, they are nonetheless categorically different things. When an AI generates outputs, however objectively novel or interesting these may be, they are not original works of authorship or expression within the meaning of copyright law. As Ian Kerr and I argued at length elsewhere—and I explain in more detail below the whole concept of an AI author is an oxymoron; AI is ontologically incapable of authorship.⁴² But if this is true, why do people seem so eager to assume otherwise?

3.2.1 Antropomorphism and Expressive Agency

We know—and a great deal of research now demonstrates—that people are inclined to anthropomorphize robots, attributing to AI human attributes and emotions.⁴³ Anthropomorphic "framing"—giving an AI a human name and gender like *Benjamin* the *Sunspring* screenwriter)—compounds this

⁴¹ Feist, supra note 14; CCH Canadian, supra note 19.

⁴² See generally, Carys Craig & Ian Kerr, *The Death of the AI Author*, 52 OTTAWA L. REV. 31 (2020), https://rdo-olr.org/2021/the-death-of-the-ai-author/.

⁴³ See, e.g., Pascal Boyer, What Makes Anthropomorphism Natural: Intuitive Ontology and Cultural Representations, J. ROYAL ANTHROPOLOGICAL INST. 83 (1996); Brian R. Duffy, Anthropomorphism and the Social Robot, 42 ROBOTICS & AUTONOMOUS SYSTEMS 177 (1996); Brian R. Duffy & Karolina Zawieski, Suspension of Disbelief in Social Robotics, 21 st IEEE INT'L SYMP. ON ROBOT & HUMAN INTERACTIVE COMM. (RO-MAN), 484 (2012); Ryan Calo, People Can Be So Fake: A New Dimension To Privacy and Technology Scholarship, 114 PA. ST. L. REV. 809 (2009).

inclination.⁴⁴ The phenomenon is not limited to laypersons; even AI researchers frequently anthropomorphize their robot creations. As Diane Proudfoot observes, "the same researchers who deny that their robots have emotions attribute *expressive behaviors* to the machines literally and without qualification; in this way they unwittingly anthropomorphize the machines..." Proudfoot points to the terms in which MIT researchers described the various facial displays of the "Kismet" robot head, developed in the 1990s, which could recognize and simulate emotions. When its creators described it as having a "happy expression," for example, they implied "that the robot has a certain communicative intent—the intent possessed by creatures that smile, namely human beings." Of course, all that the robot head is capable of simulating is the *representation* of a smile; it is not *expressing* anything at all, never mind a *feeling*.

When accounting for the actions or behavior of AI in such terms, researchers unwittingly ascribe an *intentional stance* to the machine; that is, they slip into thinking of it as possessing some form of "intentional agency" complete with implied "drives, interests, goals, as well as intentions."⁴⁷ To think in these terms elides the obvious distinction: "AI is computational, whereas intentions are not...[T]he two are ontologically different."⁴⁸ Ultimately, as Proudfoot cautions, the "extravagance with which even AI researchers anthropomorphize machines suggests that...the illusion of communication with a machine may be too readily generated."⁴⁹ Why too readily? There is a risk here to which we must be attentive: The ready illusion of genuine communication permits the machine to be mistakenly hailed as "a thinking thing."⁵⁰

The potential social, normative, and regulatory significance of such an error should be self-evident. Recall here in passing that our vision of normative neutrality requires us to see the novel technology in context *as it is*—not *as if* it

⁴⁴ Kate Darling, *Who's Johnny? Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy*, in ROBOT ETHICS 2.0 173 (Patrick Lin, Keith Abney & Ryan Jenkins, eds., 2017).

⁴⁵ Diane Proudfoot, *Anthropomorphism and AI: Turing's Much Misunderstood Imitation Game*, 175 ARTIFICIAL INTELLIGENCE 950, 951 (2011).

⁴⁶ *Id.* at 952 (citing Cynthia Breazeal & Brian Scassellati, *Challenges in Building Robots That Imitate People*, in IMITATION IN ANIMALS AND ARTIFACTS 1 (Kerstin Dautenhahn & Chrystopher L. Nehaniv, eds., 2001).

⁴⁷ Deborah G. Johnson & Mario Verdicchio, *From AI, Agency and Responsibility: The VW Fraud Case and Beyond*, 34 AI & SOCIETY 639, 645 (2019).

⁴⁸ *Id.* (emphasis added).

⁴⁹ Proudfoot, *supra* note 45, at 954.

⁵⁰ *Id.* (citing personal communication with Rodney Allen Brooks). *See generally* Rodney Allen Brooks, Cambrian Intelligence: The Early History of the New AI (1999).

were something else. As a machine, we should not treat it as if it were a thinking thing. Recall too that we are tasked with considering the AI-copyright challenge in light of copyright's purpose. If copyright aims to encourage authorship—a communicative act of *expression*—we should not treat AI *as if* it were communicating or expressing something when, evidently, it is not.

Recent videos of dancing robots capture the same concern, perhaps even more overtly. In December 2020, Boston Dynamics released a video showing four of its robots dancing to the 1962 hit "Do You Love Me?"⁵¹ As it turned out, people did love them—the video took off online and now has over 35 million views on YouTube alone. Not everyone was charmed by the exercise, admittedly. As one Twitter user was widely quoted to have replied: "Do you love me? Not when you come to annihilate us."52 While this tongue-in-cheek response evokes the dystopian AI futures of sci-fi imaginings in which raging robots seek liberation, it touches on a point that is real right now: the robots' physical mobility, dexterity, timing, and precision were not designed for dancing the mashed potato; they are attributes that support the effective deployment of these anthropomorphized machines on uneven terrain, from war zones to factory floors to public parks, making them suitable for military, industrial, or police use. The robot dog (named Spot, of course) was arguably the star of the dance video but its surprise entrance garnered less public adoration when it appeared in exercises by the French military and US police departments. More sinister still was the widely shared video of the robot dog armed with rifle on its back at a US trade show.53

My point is not simply that dancing robots appear misleadingly lovable and harmless compared to their armed counterparts on patrol—though that is an important side note for anyone concerned with the ethics of AI. Rather, I juxtapose these divergent deployments of the same technology to emphasize the obvious point that, as far as the robot is concerned, *there is no difference*: The robot dog is mechanically moving as it is trained to do whether it is pointing its weapon at a potential target or doing the twist. From this indubitable conclusion another more controversial one might be offered: The "dancing robot" is not really *dancing* at all.

⁵¹ Online: https://www.youtube.com/watch?v=fn3KWM1kuAw.

⁵² Jan Nicolas, online: https://twitter.com/phoyager/status/1344176961116102658>.

⁵³ Online: https://www.youtube.com/watch?v=byhRXB8]ZNs>. This, too, provoked headlines. One asked (as if it were a mystery), *Why do armed robot dogs make us uncomfortable: Is it the function of an armed robot dog that raises eyebrows, or its form?*, The Diplomat (Oct. 15 2021), https://thediplomat.com/2021/10/why-do-armed-robot-dogs-make-us-uncomfortable/ (Spoiler: it's both!).

Dancing, after all, is an expressive act.⁵⁴ Copyright protects dance—original works of choreography—just as it protects poems, novels, paintings, or songs. Whether an author uses their intellectual effort to combine words into verse or movements into dance, we understand the resulting work to be one of protected intellectual *expression*.⁵⁵ In the absence of originality, there will not be copyright, of course; but there will still be expression. We dance to express ourselves, to communicate with one another, to bond socially, to entice sexually, or sometimes just because we feel like dancing when the music moves us. The robots are neither moved by the music not capable of the expressive agency, emotion, or intentionality that makes dancing *dancing* as opposed to mere physical or functional movement. Once again, then, when we imagine or describe robots as behaving in terms that inherently imply a certain communicative intentionality, we make a category mistake: we attribute to the machines qualities that, by their nature, they simply cannot possess.

When an AI writes a screenplay, it does so by predicting words that might be logically strung together into a sentence, then a paragraph, and then another. It may be tempting for us to frame this ability to predict what words will follow other words as an act of authorship—it might even appear to be, in effect, the functional equivalent of what an author does when they string words together to make an original work. But, as Kerr wrote, the machine "neither knows, understands, nor appreciates the connotation of its word assemblage, let alone the meaning or value of the 'work' as a whole."⁵⁶ As captured in Ryan Calo's evocative depiction, "the box is 'gorged on data but with no taste for meaning."⁵⁷ According to Christopher Buccafusco, the act of stringing together words becomes an act of authorship not because of semantic intentions (those having to do with the intended meaning or interpretation of the work) but rather categorical intentions (about what kind of work the author has

⁵⁴ See generally Christopher Buccafusco, Authorship and the Boundaries of Copyright: Ideas, Expressions, and Functions in Yoga, Choreography, and Other Works, 39 Colum. J. L. & Arts 421 (2016), https://doi.org/10.7916/jla.v39i3.2081.

⁵⁵ See No. 121, 77 Fed. Reg. 37607 (June 12, 2012) (copyright can attach to the "composition and arrangement of a related series of dance movements and patterns organized into an integrated, coherent, and expressive whole" but not to "the mere selection and arrangement of physical movements.")

⁵⁶ Craig & Kerr, *supra* note 42, at 69.

⁵⁷ Ryan Calo, *The Box*, in Telling Stories: On Culturally Responsive Artificial Intelligence (University of Washington Tech Policy Lab, 2020), quoted *id*.

created).⁵⁸ Authors deem themselves to be authoring, and in doing so, their intention is to "produce mental effects" in their audience—they intend to "generate thoughts, feelings, emotions, and other states of cognition."⁵⁹ If this is our measure, once again, AI is incapable of the requisite intentionality.

The notion that the AI is doing something functionally equivalent to a human author thus depends upon a false analogy—it rests on a vision of the technology distorted by misunderstanding, misrepresentation, anthropomorphic framing, and romanticism.

3.2.2 Romanticizing the "AI Author"

The romanticization of robots is all the more apparent when the generative AI is presented in rhetorical terms reminiscent of the mythic "romantic author:" the entirely independent, wholly original creative genius, generating novel works *ex nihilo*. It has long been recognized, of course, that this mythic author-figure bears little or no resemblance to the real human author, whose relationality and social-situatedness means that authorship is essentially a "process of adapting, transforming, and recombining what is already 'out there' in some other form." In Roland Barthes' famous phrasing, "the text is a tissue of citations, resulting from the thousand sources of culture." But where does this leave the human author in relation to the emergent AI author in our popular romantic narrative? Barthes may have declared "The Death of the Author" back the 1960s, 62 but the old romantic myth of the Author-as-God

 $^{^{58}}$ Christopher Buccafusco, A Theory of Copyright Authorship, 102 VA L. REV. 1229, 1261 (2016).

⁵⁹ *Id.* at 1263.

⁶⁰ Jessica Litman, *The Public Domain*, 39 EMORY L. J. 965, 967 (1990). See also, e.g., James D.A. Boyle, *The Search for an Author: Shakespeare and the Framers*, 37 Am. U. L. Rev. 625, 633 (1988); Peter Jaszi, *Toward a Theory of Copyright: The Metamorphoses of 'Authorship'*, 1991 Duke L. J. 455, 462 (1991); Martha Woodmansee, *On the Author Effect: Recovering Collectivity*, 10 CARDOZO ARTS & ENT. L.J. 279, 291 (1992); David Lange, *At Play in the Fields of the Word: Copyright and the Construction of Authorship in the Post-Literate Millennium*, 55 L. & CONTEMP. PROBS. 139 (1992). For a discussion of competing conceptions of authorship, see Carys J. Craig, *Reconstructing the Author-Self: Some Feminist Lessons for Copyright Law*, 15 J. GEN. SOC. POL'Y & L. 207 (2007).

⁶¹ Rolande Barthes, *The Death of the Author*, 5+6 Aspen: Mag. Box (Richard Howard, trans., 1967), http://www.ubu.com/aspen/aspen5and6/index.html.

⁶² Id.

quite seamlessly morphed into the AI-Author-as-God.⁶³ Consider the position of Mario Klingemann, a German artist on the leading edge of AI art:⁶⁴

"Humans are not original," he says. "We only reinvent, make connections between things we have seen." While humans can only build on what we have learned and what others have done before us, "machines can create from scratch." 65

In Klingemann's depiction, the AI becomes, quite paradoxically, the ideal author—the only truly originating, creative entity. Such a conception of generative AI overlooks the vast quantities of human-produced expression, creative processes, experiences, and cultural heritage necessary to train the AI. Dan Burk explains:

Consideration of the machine in isolation from its extended sociotechnical network lends itself to romanticization of the machine, much as isolation of the human creator from his assemblage of influences once lent itself to romanticization of the human author.

Thus, regarding the machine as author seems unproductive as either a policy or a doctrinal prescription.⁶⁶

Just as the romantic author myth obscured the cumulative and collaborative nature of human creativity, it invisibilizes the extended networks of human influences and interactions behind the AI's generative processes. The AI must be understood in the context of its complex relational network, but this does not make it capable of *relating*—it is not a relational being. As such, to imagine

⁶³ See id. ("[A] text does not consist of a line of works, releasing a single 'theological' meaning (the 'message' of the Author-God)..."). For a larger discussion of the role of the so-called Romantic Author in constructions of AI authorship, see Craig & Kerr, supra note 42, at 45–57. On the idea of AI-as-God in AI discourse, see Beth Singler, The AI Creation Meme: A Case Study of the New Visibility of Religion in Artificial Intelligence Discourse, in 11(5) RELIGIONS 254 (2020) (finding that 67% of AI Creation Meme images based on Michelangelo's The Creation of Adam depicted the AI hand in the 'God' position as Creator).

⁶⁴ See Mario Klingemann: Memories of Passerby I, SOTHEBY'S, http://sothebys.com/en/auctions/ecatalogue/2019/contemporary-art-day-auction-l19021/lot.109.html (last visited Oct. 13, 2020).

⁶⁵ Arthur Miller, *Can Machines Be More Creative Than Humans?*, The Guardian (Mar. 4, 2019), http://theguardian.com/technology/2019/mar/04/can-machines-be-more-creative-than-humans.

⁶⁶ Dan Burk, *Thirty-Six Views of Copyright Authorship, by Jackson Pollock*, 58 HOUS. L. REV. 263, 266 (2020), https://houstonlawreview.org/article/18011-thirty-six-views-of-copyright-authorship-by-jackson-pollock, citing *see* Carys Craig & Ian Kerr, *The Death of the AI Author* 27, 38 (Mar. 25, 2019) (unpublished manuscript), http://dx.doi.org/10.2139/ssrn.3374951.

the AI stepping into the category of original author results in a kind of reification of the AI, as though it crafted its own individuated work by force of some magical creative spark. It ignores the now inescapable fact: Whether or not the AI is generating outputs that are facially equivalent to human-authored works, the AI lacks the intentionality, creative agency, and understanding necessary to engage in authorship as a relational and communicative act.⁶⁷ It is, again, ontologically incapable of being an author, properly understood.

Recall that our substantive approach to tech-neutrality tasks us with evaluating the copyrightability of AI-generated works in light of copyright's normative objectives. If copyright's purpose is understood as a balance between rewarding authors and encouraging the creation and dissemination of works, we can say now that no "author" is denied their "just reward" when an AI-generated work is refused copyright protection. Moreover, if the public interest that copyright serves is the creation and dissemination of works of intellectual *expression* (works of the arts or intellect), then the outputs generated by AI are not, in substance or nature, the kind of works that copyright is meant to encourage: However similar they may appear on their face, they are fundamentally different in kind.⁶⁸

3.3 Authorship, Creative Progress, and the Public Interest

I have argued above that AI cannot be an author, and that AI-generated works are not works of original expression or authorship within the meaning of copyright law. One might now object that, be this as it may, the production and circulation of AI-generated works should be incentivized through the extension of copyright to such works in order to advance the "progress" or public interest that central to copyright's purpose. There is, then, a final piece to the argument that AI-generated works ought not to be protected by copyright: It must be urged that the protection of AI-generated works would not advance the kind of "creative progress" with which copyright is concerned—but worse, it could cause copyright to defeat its own ends, stultifying creative practices in a thicket of privately owned algorithmic productions.

3.3.1 Encouraging Creative Practices (Not Products)

⁶⁷ See Craig & Kerr, supra note 42, at 69-71.

⁶⁸ *Id.* at 85–86.

First, I would insist, the public interest or creative progress that copyright is intended to serve is not merely the maximal production and circulation of works as products or commodities; rather it is the encouragement of the creative *processes* and communicative *practices* that constitute authorship. As such, the mass production and distribution of AI outputs does not, in itself, advance copyright's purpose. Barton Beebe has explored the conceptual slippage of "progress" in American copyright jurisprudence from an unarticulated idea of "aesthetic progress" to "a vision of progress grounded in 'commercial value' and committed to accumulation." Arguing that this has proven to be "a significant mistake," he notes that in a "post-scarcity" or "artificial intelligence society," an "accumulationist" model of progress will no longer make good sense. As our technological conditions change, it becomes increasingly clear that the idea of progress must mean more than "the accumulation of ever more things" in a "giant warehouse' of intellectual commodities."

The "thingification" of the work in intellectual property structures obscures the social value of the expressive act behind the commercial value of the propertized "thing." The arrival of generative AI is thus an excellent reminder that the pursuit of creative progress and the public interest at play in copyright policy is advanced by encouraging the dialogic and communicative practices of creativity and authorship, not just by the mass-production and rapid proliferation of works as commercial products.

Secondly, it is crucial to consider the potential implications of extending copyright to cover AI-generated works. Having escaped the thrall of romantic authorship, we know that nothing is entirely original; creative practices require inspiration and imitation, drawing from and building on what has gone before. An understanding of authorship as a social and communicative practice reveals how critical it is for copyright law to leave space for this kind of intertextuality or dialogic response in the cultural conversation.⁷² If the copyright system is to advance the public interest in encouraging the creation and dissemination of works of authorship, then the limits of copyright—the public freedoms to use, enjoy, and employ works—are every bit as important as the private rights that

⁶⁹ Barton Beebe, *Bleistein, The Problem of Aesthetic Progress, and the Making of American Copyright Law*, 117 COLUM. L. REV. 319, 395 (2017).

⁷⁰ *Id.* at 395.

⁷¹ *Id.* at 396, citing Jessica Litman, *Lawful Personal Use*, 85 TEX. L. REV. 1871, 1880 (2007).

⁷² See Carys Craig, Transforming 'Total Concept & Feel': Dialogic Creativity and Copyright's Substantial Similarity Doctrine, 38(3) CARDOZO ARTS & ENT. L. J. 603, 609–12 (2021).

it accords.⁷³ There is, then, nothing *wrong* with leaving AI works outside the scope of copyright. As Burk notes:

Leaving creations 'where they lie' in authorship doctrines means that in some cases there will be no author, and the creation will simply fall into the public domain. In general, this will be the right result where there is no need for a legal incentive, or the incentive is gratuitous.

I have argued that encouraging AI-generated works does not advance copyright's purpose—awarding copyright protection to incentivize their generation would therefore be both gratuitous and unnecessary. But we can take this one step further: Leaving AI-generated works in the public domain may prove to be essential to encourage and facilitate authorship in the AI age.

With the rapid production and proliferation of AI-generated works on a massive scale, the cultural landscape is likely to become increasingly cluttered with AI-generated outputs. If these were to be protected by copyright and so subject to private control, these works would quickly become "copyright landmines," depleting the usable public domain by making it ever more difficult for creators to navigate without legal risk.⁷⁴ An AI can potentially produce thousands or millions of copyrightable works in a relatively short time frame. If they were to be protected by copyright, as Clark Asay cautions:

each of those thousands or millions of copyrights may stand in the way of other creative parties wishing to make use of the same or similar expression in their own creative efforts. AI technologies may thus foment a copyright anticommons, where creative parties wishing to engage in their own creative activities face so many AI-spawned copyright hurdles that they simply relent in those efforts.⁷⁵

The result of applying copyright to AI-generated outcomes, then, would be to "significantly inhibit creative efforts overall."⁷⁶ By the same token, I would add,

 $^{^{73}}$ Cf. Craig, Locking Out Lawful Users, supra note 27, at 179; quoted in Access Copyright, supra note 17, at ¶ 95 ("The limits to these private rights, defined by fair dealing and other exceptions—and circumscribed by the boundaries of the public domain—are therefore essential to ensure that the copyright system does not defeat its own ends.")

⁷⁴ Clark D. Asay, *Independent Creation in a World of AI*, 14 FLA. INT'L U. L. REV. 201 (2020) [*Independent Creation*]. The risk that a human-authored work will be identified as substantially similar to a protected AI work will surely be compounded by the growing use of algorithmic copyright enforcement—technology trained to detect "matches" between works, but equipped to investigate claims of independent creation, non-substantial taking, or fair use. *See* Craig, *Copyright & AI*, *supra* note 27, at 36–37.

⁷⁵ Asay, *Independent Creation*, supra note 74, at 206.

⁷⁶ *Id*.

leaving this potentially vast quantity of works "where they lie" without enclosing them behind private lines would significantly enrich the public domain, allowing it "to flourish as others are able to produce new works by building on the ideas and information contained" therein.⁷⁷

3.3.2 Recognizing the Limits of Copyright's Reach

It should be stressed that to define AI-generated works as public domain is therefore not to pronounce them worthless or without value to human culture, but rather to keep them free from the exclusive control granted by copyright private control that imposes a social and cultural cost. As Jessica Litman explains, "[t]he public domain should be understood not as the realm of material that is undeserving of protection, but as a device that permits the rest of the system to work...."78 The notion that intellectual property rights should attach to any intangible thing of value—"if value/then right"—is an unfortunate fallacy.⁷⁹ It produces the assumption that, if AI-generated works are socially, culturally valuable in some way, they should be privately owned.80 But, as scholars have warned repeatedly over the years, it is a damaging default to assume that intellectual property simply expands to enclose the latest valuable innovation.⁸¹ This entails the unnecessary swelling of our IP system and the continual encroachment of IP claims into the public domain and our shared cultural sphere. Unnecessarily extending the private preserve of copyright over unauthored, AI-generated works may enrich the—likely corporate—actors behind the AI (those who create, own, train or deploy the AI), but it would come "at the loss of society's interest in maintaining a robust

⁷⁷ CCH Canadian, *supra* note 19, at ¶ 23.

⁷⁸ Litman, *supra* note 60, at 968.

⁷⁹ See Alfred C. Yen, Brief Thoughts About If Value/Then Right, 99 BOS. U. L. REV. 2479, 2480 ("That principle, which the U.S. Copyright Act does not embrace, expresses the intuition that "wherever value is received, a legal duty to pay arises, regardless of whether imposing that legal duty serves public welfare", citing Wendy Gordon). See also Rochelle Cooper Dreyfuss, Expressive Genericity: Trademarks as Language in the Pepsi Generation, 65 NOTRE DAME L REV 397, 405–06 (1990) (questioning the idea that relationship between value and ownership justifies granting trademark rights); Wendy J. Gordon, On Owning Information: Intellectual Property and the Restitutionary Impulse, 78 VA. L. REV. 149, 178–80, 244 (1992).

⁸⁰ Perhaps the most powerful critique of this tautological reasoning is still that of Felix Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809, 815 (1935).

⁸¹ See, e.g., David Lange, Recognizing the Public Domain, 44 L. & CONTEMP. PROBS. 147 (1981); James Boyle, The Second Enclosure Movement and the Construction of the Public Domain, 66 L. & CONTEMP. PROBS. 33 (2003); James Boyle, THE PUBLIC DOMAIN (2008). See generally Pamela Samuelson, Enriching Discourse on Public Domains, 55 DUKE L.J. 783 (2006); Carys Craig, The Canadian Public Domain: What, Where and to What End?, 7 C.J.L.T. 221 (2010).

public domain that could help foster future creative innovation."⁸² In other words, once again, expanding copyright to AI generated works would run counter to the normative objectives of copyright law.

Indeed, if AI-generated works are of value to the public, the more appropriate policy response is to ensure that there is *space* for them to develop within the strictures of our existing IP system. Recognizing the social and cultural value of AI's generative capacities should therefore entail, as a policy response, not the unnecessary and counterproductive protection of its outputs but the shielding of its *inputs* and functional processes from potentially chilling liability.

The problem arises because copyright is, of course, premised on the exclusive right to make copies. In the process of training AI systems, as we have seen, vast quantities of data—texts, images, and other potentially copyrightprotected works of human authorship—must be digitally reproduced. significant policy concern, then, is whether any copyright that subsists in the Al inputs is infringed in the training of Al. Given the sheer volume of text and data mined to effectively train a sophisticated AI, limiting or foreclosing the use of copyright-protected works in such processes in the absence of permission from the right-holder places an enormous burden on AI research and development. Moreover, it produces de facto barriers to certain kinds of AI projects, differentially disadvantages anything but the most well-resourced AI researchers, and exacerbates the built-in biases and discriminatory effects of AI systems. 83 The quality and scope of a dataset has a direct bearing on the quality and operation of the resulting AI. In short, we must be alert to the risk that copyright law unduly restricts, distorts, or otherwise determines the trajectory of Al's technological development and operation. This, too, would be contrary to copyright's normative objectives—copyright is neither designed for, nor suited to, this role.

The solution to this policy conundrum can once again be found in recognizing the nature of the copyrightable work—in light of copyright's purpose—as fundamentally *expressive*. What copyright is concerned with is not the mere reproduction of *things* as such, then, but with copying that "relates to human appreciation of the expressive qualities of that work." From this it follows that copyright law need not concern itself with "any act of reproduction that is not

⁸² CCH Canadian, *supra* note 19, at ¶ 23.

⁸³ See Amanda Levendowski, How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem, 93:2 WASH. L. REV. 579 (2018). Reliance on public domain or other low-liability risk input means training AI on data that is obsolete, exclusionary or fails to reflect contemporary information and social values.

intended to enable human enjoyment, appreciation, or comprehension of the copied expression as expression."⁸⁴ The copyright owner's reproduction right should not be implicated by "non-expressive" or "non-consumptive" copies, including the kind of digital copies involved in machine learning. As Matthew Sag notes, such a conclusion would be "entirely consistent with the fundamental structure of copyright law because, at its heart, copyright law is concerned with the communication of an author's original expression to the public."⁸⁵

A recent addition to Japan's copyright law offers an interesting articulation of such a limit upon copyright's scope, stating: "It is permissible to exploit a work, in any...case in which it is not a person's purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work." As Ueno Tatsuhiro explains, "The underlying theory behind this relates to the nature of copyright, or the justification for copyright protection that an exploitation not for 'enjoyment' purposes is beyond the inherent scope of copyright." 87

As we have seen, works of authorship are the result of a communicative act, intended to generate thoughts, feelings, emotions, desires, or other states of cognition in their audience—and copyright law protects them as such. The AI is not an audience capable of such a response, of course, and so the use of the work to train the AI amounts only to use of it as a functional thing and as a source of information or data (which belongs in the public domain).⁸⁸ It is not a use of the work *as a work of authorship*.⁸⁹ It should not, therefore, implicate copyright at all.⁹⁰

⁸⁴ Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291 (2019).

⁸⁵ *Id.* at 302.

⁸⁶ Copyright Law of Japan Act, art. 30-4, http://www.cric.or.jp/english/clj/cl2.html.

⁸⁷ Ueno Tatsuhiro, The Flexible Copyright Exception for 'Non-Enjoyment' Purposes – Recent Amendment in Japan and Its Implication, 70(2) R.U.R. INT'L 145 (2021).

⁸⁸ Facts and information are not protected by copyright even if they are contained within a copyrightable work. *See* CCH Canadian, *supra* note 19, at ¶ 22 ("in Canada, as in the United States, copyright protection does not extend to facts or ideas but is limited to the expression of ideas."); Feist, *supra* note 14 ("That there can be no valid copyright in facts is universally understood").

⁸⁹ I am indebted for this framing to Martin Senftleben, Ueno Tatsuhiro, and the participants of the panel discussion on "New Developments in Copyright Limitations and Exceptions for Education and Research" at the *Global Congress for Intellectual Property and the Public Interest* (Oct. 26 2021).

⁹⁰ If the making of copies for such text and data-mining purposes were to be considered a prima-facie infringement of copyright, however, it should nonetheless avoid

Much of today's policy discussion around copyright liability in respect of AI inputs focuses on the availability of fair use or fair dealing as a potential defense to copyright infringement or the need for specific exemptions to permit informational analysis or text and data mining. Many if not all uses made for machine-learning purposes should qualify as "fair" and lawful, not least because such copies do not compromise the core interests of the copyright owner or substitute for the work of the author in the market.⁹¹ Several U.S. court rulings suggest that text and data mining may indeed satisfy the fair use test and will not, therefore, amount to copyright infringement.92 In Canada, such uses should fall within the scope of fair dealing for purposes of research, private study, or education, and should therefore be considered a user's right.93 But conducting a nuanced and context-specific assessment as required by fair use in relation to each work fed into the AI dataset is clearly unmanageable at the scale demanded by AI—and the risks of getting it wrong will have an inevitable chilling effect. This in turn points towards the strategic need for specific and mandated exceptions for text and data mining activities. International developments in this direction, however, reveal that framing these user rights as new exceptions to copyright owners' existing entitlements produces unduly narrow and rigidly defined carve-outs from presumptive copyright control. In practice, these impose significant limits on users and obstruct AI research and development under the guise of protecting copyright interests.94

As this demonstrates, the advantage of recognizing, first and foremost, copyright's *limits*—as defined by the nature of authorship and the public

liability under the fair use defence or, in Canada, the user right of fair dealing for the purposes of research, private study, education or review.

⁹¹ Sean Flynn, Christophe Geiger, João Pedro Quintais, Thomas Margoni & Matthew Sag, *Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action* 4 (2020), https://digitalcommons.wcl.american.edu/research/48.

^{92 17} U.S.C. § 107; 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. v. iParadigms, LLC. (4th Cir. 2009); See also K. Courtney, R. Samberg, & T. Vollmer, Big data gets big help: Law and policy literacies for text data mining", 81(4) COLL. & RES. LIBR. NEWS (2020), https://crln.acrl.org/index.php/crlnews/article/view/24383/32222.

⁹³ Copyright Act, R.S.C. 1985, c. C-42, s. 29. See Craig et al., Submission by IP Scholars Copyright and Artificial Intelligence, supra note 40.

⁹⁴ See, e.g., Copyright, Designs and Patents Act 1988, c.48, art. 29A (U.K.); Act No. 2016-1321 of 7 October 2016 for a digital Republic, s. 38 (Fr.); Eur. Parl. Doc., Dir. 2019/790 of the European Parliament and of the Council Of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directive 96/9 and 2001/29, arts. 3 & 4. See generally, Flynn et al., supra note 91.

interest at stake—is to shift the starting point for this whole analysis: Copyright protects works of authorship as textual embodiments of original *expression*, from which it follows that its reach should not extend to non-expressive uses of texts as data sources. Such uses are not equivalent to uses of works by or for a human addressee, whether audience or downstream author. Simply put, they are not copyright's concern. Specific exceptions or fair use determinations may helpfully *confirm* the lawfulness of such uses and add much-needed certainty for those in doubt—but no defense is necessary to excuse lawful, non-infringing uses of works.

Recall that one of the flaws of a formalist approach to tech neutrality described above in Part 2.1 was the assumption that a copy is a copy, and so every copy should be treated in the same manner by law no matter the technological processes or activities involved in its production. The more substantive principle with which we are working here considers the functional effect of the copy in context and in relation to the normative objectives of the copyright system. We can see here, again, the importance of approaching the policy question in this way: Rather than getting caught up in the technical matter of reproduction and the inevitable extension of copyright control to new technological processes, we can recognize the irrelevance of the reproduction in relation to copyright's normative concerns. In doing so, we can ensure that copyright leaves space for technology to evolve, for AI to be trained on richer data sets, and for the public domain to be enriched by new AI-generated works—works from which human authors may readily draw inspiration and upon which they may freely build, fostering the very creative progress at which copyright is aimed.

4. CONCLUSION

The purpose of copyright is to encourage not merely the production and circulations of works, as objects, but the *activity* of authorship – the dialogic *processes* and exchange *of meaning* that constitute authorship. Works are not just *things* that circulate; they are *expressions* that form relations of communication between people. Nor does copyright law create private rights over such works simply for the sake of rewarding authors; it serves a public interest—the encouragement of authorship as a social and creative *practice* and the dissemination of works to contribute to a vibrant public domain.

As AI-generated works increasingly come to resemble, facially, the human authored works with which copyright is concerned, it will be vital to keep these

purposes in view—for it is these purposes that will continue to define the justifiable scope of copyright. As I argued in Part 2, it is not appropriate for copyright law to respond to these technological developments in a formally tech-neutral way, treating alike AI and human activities and artefacts that are fundamentally different in nature. Nor is it sufficient to focus on their functional equivalence, analogizing between AI and human activities and artefacts without regard to the social values and public purposes that underpin the copyright system. Rather, it is the normative objectives of the copyright system and the values they reflect that should be regarded as technologically neutral. The real AI-copyright challenge, then, is to prevent this latest technological innovation from upsetting the copyright balance, obstructing its normative objectives, and thereby undermining the social value of authorship, creative progress, and the public domain. I explained, in Part 3, that this will require resisting calls to extend copyright to cover AI-generated works on the basis that they are not equivalent to the works of authorship that copyright seeks to encourage. Similarly, attention to the expressive nature of authorship supports restricting the power of copyright owners to control the use of works as inputs for training AI systems.

These and many more legal and doctrinal conundrums will be thrown up by the arrival of increasingly sophisticated AI systems and their pervasive deployment in our cultural sphere. It seems we are poised at a policy precipice of sorts, ready to jump or be pushed into the path of increased commoditization and private control as these systems proliferate. Due to a combination of romanticism about robots, misapprehensions about creative progress, and complacency about the social costs of copyright, critical mistakes may be made. Both the protection of AI-generated works and the constriction of AI research and development through the creeping expansion of copyright control may spell an uncertain future for human creators, users, and audiences, harming rather than advancing the public interest that copyright is meant to serve. Amongst the many risks that this poses to our cultural environment is the possibility that we will miss out on the very real promise of AI to contribute, in previously unimaginable ways, to a rich and vibrant public domain—as both a tool of human authorship and a source of inspiration.

It should be noted in closing that the themes addressed in this chapter have implications beyond copyright and cultural production. Policymakers will surely make mistakes and misallocate economic resources, legal privileges, and political power across the board if critical policy analyses begin by misattributing human attributes, intentionality, and expressive agency to AI. The challenge of regulating and responding to AI requires first that we

recognize what AI is—and what it is *not*. The task is then to ensure that legal interventions and reform efforts are undertaken to steer AI and its development in service of the public goals, human interests, and social values that our legal constructs—much like our technological constructs—are ultimately supposed to serve.