6-2022

The Relational Robot: A Normative Lens for AI Legal Neutrality — Commentary on Ryan Abbott, The Reasonable Robot

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

Source Publication:
Jerusalem Review of Legal Studies, Volume 25, Issue 1, June 2022, Pages 24–39

THE RELATIONAL ROBOT: A NORMATIVE LENS FOR AI LEGAL NEUTRALITY

Reviewing *Ryan Abbott, The Reasonable Robot* (Cambridge University Press, 2020)

Carys J. Craig[*]

[Author's original version, forthcoming in *Jerusalem Review of Legal Studies* (2022)]

I. INTRODUCTION

Artificial Intelligence (AI), we are told, is poised to disrupt almost every facet of our lives and society. From industrial labor markets to daily commutes, and from policing tactics to personal assistants, AI brings with it the usual promise and perils of change. How that change will unfold, however, and whether it will ultimately bestow upon us more benefits than harms, remains to be determined. A significant factor in setting the course for AI's inevitable integration into society will be the legal framework within which it is developed and operationalized. Who will AI displace? What will it replace? What improvements will it bring? What damage will it do? The law has the power to shape the answers, but is it up to the task?

This is where Ryan Abbot's book, "The Reasonable Robot," enters the fray. "What is needed," Abbott writes, "is not more or less law but the right law." We need a legal regime "optimized for AI" (p. 3). To this end, Abbott proposes a new guiding tenet for the law's regulation of Artificial Intelligence: *AI legal neutrality*. This would establish, as a principled starting point, the default position that "the law should not discriminate between AI and human behaviour" (p. 3). With a swift and insightful foray through tax, tort, intellectual property, and criminal law, Abbott aims to demonstrate that, "as AI increasingly occupies roles once reserved for people, AI will need to be treated more like people, and sometimes people will need to be treated more like AI" (p. 4).

In what follows, I suggest that AI legal neutrality, as conceived by Abbott, is an interesting proposition but not a satisfactory default principle with which to equip law for the emerging realities of AI. When we scratch beneath the surface, this concept of neutrality or equal treatment is too focused on the individual person/thing, and too reliant on analogical reasoning and false equivalence. It is

[*] Osgoode Hall Law School, York University. My thanks to Omri Rachum-Twaig and Ohad Somech, convenors of "The Reasonable Robot" Symposium at Hebrew University (online) in May 2021, as well as to the symposium participants, and especially to Ryan Abbott for many thought-provoking conversations. Thanks also to Felice Yeung for her excellent research assistance.

therefore too far removed from the normative underpinnings of law, its subjects, and its teleology. Instead, I argue, we need a more substantive notion of law's technological neutrality—one that looks to the law's normative objectives to set the default. The place to start, I suggest, is not with a legal neutrality that, by design, disregards the inherent differences between humans and robots and their respective behaviors; on the contrary, we need to be fully attentive to the dynamics of human-robot relations in social context – and alert to the dangers of overlooking differences. Rights and responsibilities should be allocated with a clear view to the relationships and subjectivities that they shape and the social values they advance.

In Part II, I outline Abbott's argument in respect of legal neutrality and its application to intellectual property law. I then compare its implications in the patent law sphere to its potential significance in the copyright context. Pointing to arguments that I have advanced elsewhere about AI, copyright, and authorship, I suggest that Abbott's legal neutrality principle would likely lead us down the wrong path to bad copyright policymaking. This reveals the problem with the AI legal neutrality default. In Part III, I sketch an alternative approach to understanding technological neutrality and law, concerned with consistency in pursuit of normative objectives rather than formal non-discrimination. This approach can potentially accommodate Abbott's patent policy recommendations without entailing undesirable implications for copyright. I conclude by proposing a relational approach to regulating robots, which would, I believe, bring a much-needed normative lens to legal neutrality.

II. LEGAL NEUTRALITY: PATENT AND COPYRIGHT COMPARED

As Abbott's book convincingly demonstrates, we have arrived at the cusp of a new technological era – a Fourth Industrial Revolution if you will – to which the law must adapt if it is going to "help us realize the benefits of AI while minimizing its risks" (p. 3). When old rules are applied to new technologies, they often produce "unintended consequences" (p. 134) that threaten to undermine the original policy objectives of the law, and so may do more harm than good in the novel social contexts in which they apply. The emergence of increasingly sophisticated AI demands a thoughtful and deliberate response, to be sure – one that is focused on "the functionality of machines and consequentialist benefits" and ultimately directed at "human welfare" (p. 143). It is comforting to imagine that we might still futureproof our laws and regulatory structures to avoid the kinds of human suffering, dislocation, and poverty, on one hand, and massive accumulations of wealth and monopolistic power, on the other, that accompanied prior Industrial Revolutions (p. 5). Abbott is right to insist that

policymakers, when deciding how to legally treat AI, should be asking: "what will result in the greatest social benefit from these technologies…[?]" (p. 143). But is he right that AI legal neutrality will point the way to an answer?

The suggestion is that, at least in many instances, "neutral legal treatment would ultimately benefit human well-being by helping the law better achieve its underlying policy goals" (p. 4); it would therefore provide our policymakers with a sound starting point. From there, however, it is acknowledged that "[d]ifferences between AI and people will occasionally require differential rules" (p. 4). As such, legal neutrality is not dispositive – it "may be departed from when there are good reasons for so doing" – but it is nevertheless "an appropriate default" (p. 4).

It is not always clear, given the various policy permutations and proposed tweaks that follow in each field of law, precisely what larger principle is at work in guiding the many deviations from the default that Abbott ultimately prescribes. Given his preoccupation with optimizing the benefits of AI, as well as the steady attention paid to the dis/incentives put in place by law, the policy prescriptions seem quite uniformly grounded in a law and economics approach, with the attendant drive towards rational choice, efficiency, and big picture "progress" or generalized "welfare" that such an approach typically entails. Thus, where treating AI and human behavior the same would produce sub-optimal outcomes, legal neutrality readily gives way to more efficient alternatives. In this sense, AI legal neutrality is less a matter of principle, perhaps, than it is a pragmatic (if unreliable) shortcut to good (optimally efficient) policy.

Beyond the economic frame, however, Abbott acknowledges the persistent relevance of other principles – transparency, accountability, and beneficence get an express stamp of approval (pp. 4, 142), while mention is also made of privacy, human agency, and non-discrimination (of the equality-of-persons variety) (p. 136) – though the manner of their interaction or relative importance is never really unpacked. Abbott explains the open-ended nature of this thesis: "Intelligent policymaking requires that a decision maker consider how to balance these concepts on a case-by-case basis" (p. 138). It is to Abbott's credit that such principles, as well as the differing cultural and value preferences of different jurisdictions (p. 138), are recognized as significant sources for caveats or corrections to the proposed principle of AI legal neutrality. But is it enough?

The principle of AI legal neutrality seems to me to lack the kind of normative weight that would justify privileging it as the default. Defaults may not be dispositive, but they are important – and where we begin can largely, if not

entirely, determine where we end up. Or, as Ian Kerr would say, "the devil is in the defaults." Kerr explained:

> "Once defaults have settled, opposition to their usage is increasingly difficult: expectations tend to consolidate in favor of their use.… [N]ormative defaults…can lead to what philosophers (in other contexts) have called a hardening of the categories. Once a particular norm is put in place, it becomes more difficult to do otherwise…."[1]

Legal and normative defaults establish baseline presumptions from which any departure or deviation must be justified and defended. Exceptions thus extracted from default norms tend to be hard fought and narrowly drawn. It seems obvious that we should not, in our laudable desire to be proactive, rush to establish a legal default that may favor efficiency but potentially undermine values as fundamental as, say, life and security of the person, equality, or human agency. (There is, for example, little reassuring about Abbott's acknowledgement that "[a]n AI may be more efficient at identifying and eliminating military targets, but there *could be* other reasons not to delegate life and death decisions to an AI" (p. 4, *emphasis added*). There are indeed other reasons.[2])

Abbott explains that "treating AI as if it [morally deserves rights] should only be justified if this would benefit people." He draws the analogy to corporations whose "legal rights exist only to improve the efficiency of human activities such as commerce and entrepreneurship" (p. 4). Seeing the benefits to people characterized in these terms might prompt some of us to ask, "which people?" But even putting that aside, it seems clear that economically valuable rights have a problematic way of compounding to produce more and greater claims to right. The free speech rights now enjoyed by corporations as though they were people (and to the obvious detriment of democracy) are a striking example of where efficiency-rationalized corporate rights can lead us.[3] While Abbott, to his credit, expressly declines to propose that rights in the nature of moral entitlements or legal personhood be bestowed upon AI (p. 4), the principle of AI legal neutrality supports the extension of established legal norms, doctrines, and constructs to AI

[1] Ian Kerr, *The Devil Is in the Defaults*, 4:1 CRITICAL ANALYSIS OF LAW 91, 97-98 (2017).

[2] *See e.g.,* Ian Kerr & Katie Szilagyi, *Asleep at the switch? How killer robots become a force multiplier of military necessity,* in ROBOT LAW 333 (Ryan M. Calo, Michael Froomkin, Ian Kerr, eds., Edward Elgar Pub Ltd, 2016).

[3] *See* Citizens United v. Federal Election Commission, 558 U.S. 310 (2010). *See also* Frank Pasquale, *From Citizens United to Bots United: Reinterpreting 'Robot Rights' as a Corporate Power Grab*, https://www.youtube.com/watch?v=-95lZDqrHPQ (comparing claims made in respect of 'robot rights' to corporate claims-to-right).

when stepping into roles previously occupied by people. In many cases, then, additional legal rights and benefits would be allocated to AI (or its owners) as a matter course – and, once granted, they would be likely to stick and expand (even when subsequent developments reveal more efficient or less harmful alternatives).

Legal rights may be rationalized as means to an end – they may be justified only because of the incentive-based benefits that they promise society as a whole – but this doesn't mean that they are easy to dislodge when they cease to serve social ends or start to do more harm than good. Legal defaults quickly become stubborn dictates while efficiencies predictably morph into perceived entitlements. Indeed, the constantly expanding protection given to intellectual property rights provides the perfect example of this phenomenon. So, let's take Abbott's argument regarding intellectual property as the illustrative example of what AI legal neutrality potentially gets right – and what it risks getting wrong.

A. AI Inventors and Patent Protection

In Chapter 4, Abbott explores the implications of AI for patent law by revisiting the doctrinal requirement that an inventor be a human being. As he explains, almost every jurisdiction in the world currently requires that a patent application list a natural person as the inventor – the human being responsible for the mental act of invention (p. 84-85). As it stands, an AI may be an invent*ion* of a human or a *tool* of a human inventor, but an AI cannot *be* the invent*or*.[4] This is notwithstanding that, as Abbott convincingly argues, AI is capable of creating innovations that would functionally meet or surpass the objective requirements of patentability (p. 77-78).

I find Abbott's argument here quite compelling: "[The] patentability of AI-generated inventions should be based on the inventiveness of an AI's output rather than on a clumsy anthropomorphism because patent law should be interested…in a functionalist solution" (p. 86). If we accept, for now, the common

[4] Abbott's *Artificial Inventor Project*, https://artificialinventor.com/, has since met with some success in its effort to obtain patent protection for inventions generated by the AI tool DABUS. Most notable is the controversial ruling of the Federal Court of Australia in *Thaler v. Commissioner of Patents* [2021] FCA 879 (with Beach J. opining at par. 10: "[I]n my view an artificial intelligence system can be an inventor for the purposes of the Act. First, an inventor is an agent noun; an agent can be a person or thing that invents. Second, so to hold reflects the reality in terms of many otherwise patentable inventions where it cannot sensibly be said that a human is the inventor. Third, nothing in the Act dictates the contrary conclusion.") In July 2021, South Africa's patent office was the first to issue a patent in respect of an invention attributed to DABUS. Applications naming DABUS as inventor have so far been refused in Europe, the US, and the UK.

(though no doubt contestable!)[5] proposition that granting patent protection encourages investment in useful innovation to the benefit of society, then it seems reasonable to assert that patenting AI-generated inventions will therefore encourage more useful AI-generated innovations (p. 83). And if AI is capable of generating novel and useful inventions with a speed and efficiency that outstrips the pace of human inquiry and inventiveness, then it makes sense that we would want to incentivize it for the same reasons that we wish to incentivize innovation in general. It does not matter, Abbott argues, whether AI "thinks" (p. 84): why quibble over romantic tropes like the old "Flash of Genius" (p. 86) if we can enjoy the benefits of this remarkable "progress of science" (p. 79) by extending patent protection to the inventions of AI? And if this is the right result from a functional perspective, then the patenting of AI-generated inventions "ought to be permitted under a dynamic interpretation of the law" (p. 86).

It is not necessary to contest Abbott's position on the patentability of AI-generated inventions, however, to cast doubt on the appropriateness of AI legal neutrality as a guiding principle. But before I get there, I do have to raise a couple of quibbles based not on any commitment to romantic tropes or rigid doctrinal interpretations, but simply holding Abbott's argument to account on its own consequentialist terms.

Arguing that "the term 'inventors' should be afforded the flexibility needed to effectuate [patent law's] constitutional purposes, Abbott writes: "If AI inventorship and AI-generated inventions are to be prohibited, it should only be on the basis of sound public policy" (p. 87). It is worth noting as an aside that the language of "prohibition" here is too strong: no one is proposing prohibiting AI-inventorship or AI-generated inventions, of course, but only prohibiting their protection by the existing patent system. My intention here is not to nitpick. Abbott's meaning is clear enough, but the rhetorical conflation of patents granted with the fact of innovation is a common slip in the patent policy sphere, and it

[5] *See e.g.,* Lisa Larrimore Ouellette, *Patent Experimentalism*, 101 VA. L. REV. 65, 76 (2015) (concluding that none of the current evidence that we have regarding the patent system "resolves whether patents have a net positive effect on innovation, much less their net welfare effect, or whether alternative innovation incentives such as grants, prizes, and tax credits are inferior"). *See also Staff Of S. Subcomm. On Patents, Trademarks, And Copyrights, 85th Cong., An Economic Review Of The Patent System* 80 (Comm. Print 1958) (Patent Study of Fritz Machlup) (observing that "[i]f we did not have a patent system, it would be irresponsible, on the basis of our present knowledge of its economic consequences, to recommend instituting one. But since we have had a patent system for a long time, it would be irresponsible, on the basis of our present knowledge, to recommend abolishing it.") By the same logic, might it also be irresponsible, on the basis of our present knowledge, to recommend extending it to a whole new category of inventions such as those generated by AI?

misstates the connection. Innovation can and does occur in the absence of patent rights, as Abbot expressly notes (p. 84), and there are many reasons why AI-inventorship might continue to flourish undeterred and unabated even in the absence of temporary legal monopolies. To say that AI-inventorship ought to be excluded from patent protection is not to say that it is without value, should be prevented, or shouldn't otherwise be encouraged and rewarded. It could simply be to suggest that patent rights are unnecessary, inefficient, and/or otherwise inappropriate as currently conceived.

As for prohibiting the patenting of AI-generated inventions, I have yet to be convinced, based only on economic theorizing, that there is a real risk of underproduction of AI-innovations in the absence of patent protection that would justify, on demonstrable economic grounds, the need to extend patents to AI-generated inventions. Nor am I satisfied that, given the very different costs and benefits involved in AI innovative processes, the benefits of granting protection are either necessary to offset the costs incurred by the AI owner (given the scale and rapidity at which an AI can operate once programmed and trained) or sufficient to outweigh the costs to society of awarding such monopoly rights over, potentially, vast swathes of new knowledge (especially given that the creator of the inventive AI is already incentivized by intellectual property rights over the AI as invention and/or the copyrighted software code on which it runs.)

This brings me to my second quibble. In the absence of evidence of underproduction[6] or a problem to be solved,[7] why default to granting monopolistic protection rather than leaving inventions free for all to use? Patent protection, like other forms of intellectual property, is a limited and carefully circumscribed exception to the general rule that ideas, knowledge, and innovations can be freely used, shared, and enjoyed. The natural default, in other words, is the *public domain*.[8] We might assume that the logic of incentives applies

[6] *See* Mark Lemley, *The Myth of the Sole Inventor*, 110 MICH. L. REV. 709, 736 (2012) ("We grant patents ... to encourage inventions we wouldn't otherwise get.")

[7] Abbott suggests that the current state of the law is also problematic because patent applicants are incentivized to fail to disclose the use of AI. He further argues that the unacknowledged use of AI will itself undermine human inventorship (p. 83-84). The fact that applicants make misrepresentations in bad faith is not, of course, a good reason to adjust the existing system to accommodate their ambitions. And to the extent that we might seek to protect the category of human inventorship, to simply equate it with AI-inventorship seems like counter-intuitive solution more likely to blur categories than to sanctify them.

[8] *See* James Boyle, THE PUBLIC DOMAIN: ENCLOSING THE COMMONS OF THE MIND (2008); David L. Lange, *Recognizing the Public Domain*, 44 LAW AND CONTEMPORARY PROBLEMS, 147-178 (1981). *See generally* Carys J. Craig, *The Canadian Public Domain: What, Where, and to What End?* 7 CANADIAN JOURNAL OF LAW & TECHNOLOGY 221 (2010).

equally to AI and human-generated inventions, but even then, there are good reasons to think that these inventions might present different problems that demand different, tailored solutions. Perhaps AI-generated inventions, downstream of the AI invention itself, could be freely used without any risk of underproduction, given the scale and rapidity of AI outputs?[9] Perhaps they could benefit from a *sui generis* system of protection that grants limited rights to specified parties (programmers, owners, users) for shorter durations or with more a carefully circumscribed scope of protection, and subject to additional regulatory requirements? We simply don't know yet. Until we do, I would suggest, the public domain should be the preferred legal default – and the normative default should, as a matter of principle, be public access to knowledge and equal enjoyment of the benefits of scientific progress and its applications.[10] To tweak Abbott's assertion, then, "[i]f AI inventorship and AI-generated inventions are to be [*protected*], it should only be on the basis of sound public policy."

But I digress. My purpose here is to suggest that, *even if* Abbott is right about what makes good AI patent policy, this doesn't necessarily support his principle of AI legal neutrality. In fact, it seems to me that his conclusion on AI patent policy (rightly) flows not from the starting point that AI-inventions and human inventions should be accorded neutral treatment as a matter of principle, but rather from the *purpose* of the Patent Act and "the Founders' intent in enacting the Patent and Copyright Clause" (p. 86-87).[11] This is more important than it might seem for reasons that relate to the dynamics of normative defaults.

If we begin with Abbott's principle of AI legal neutrality, we presume that the AI "inventor" should receive the same legal rights as the human inventor, and we depart from this default position only if it can be demonstrated that this would be contrary to human well-being or some other principle that is shown to be "more important" (p. 142). Private rights designed for individuals are thus seamlessly extended to AI entities or their owners as a matter of course in the absence of compelling reasons to withhold protection. And who is to make these arguments against legally neutral protection of AI and human inventors? It seems clear that powerful economic interests will align with more protection for

[9] *See* Clark Asay, *Artificial Stupidity*, 61 WM. & MARY L. REV. 1187, 1207 (2020) (noting that obstacles to patenting in the AI industry "may yield some positive results, such as less frivolous patent litigation and greater freedom to operate generally. Those results, in turn, may help propel the AI industry towards greater levels of AI innovation.")

[10] *Cf.* Article 15 of the *International Covenant on Economic, Social and Cultural Rights.*

[11] I say "rightly" because this starting point would be consistent with the more expansive approach to technological neutrality that I set out below. *See infra*., Part III.B.

the AI owner (given that inventive AI is quite likely to be "owned by large enterprises with sophisticated attorneys" (p. 88)). This presents the risk of conflating already powerful corporate interests with the interests served by AI legal neutrality in the patent sphere – all in the name of advancing public welfare?

If, however, we begin by asking about the purposes of patent law – the public interest objectives at which it is aimed – and conclude (ideally, based on some empirical evidence!) that these interests would be positively advanced by extending patent protection to AI-generated inventions, then so be it. Why not? At that stage, the decision to withhold protection from AI inventions would have to be based on sound public policy because the law would be departing from the steady pursuit of the established social goals that have traditionally justified its prescriptions. Even if Abbott is ultimately correct that AI-generated inventions should receive the same patent protection as their human-created counterparts, then, it matters a great deal how we get there.

The problem with the legal neutrality default becomes even more pronounced when we turn our attention to copyright law.

B. AI "Authors" and Copyright

Abbott's analysis bundles copyright together with patent law under the umbrella term of "intellectual property." The common unifying terminology of intellectual property or "IP" is problematic for a number of reasons,[12] but for our purposes the most troubling is the tendency that it produces to equate entirely unrelated subjects and objects within a single conceptual category and so to disregard fundamental differences between them.

When investigating the possibility of departing from the notion of the human inventor for patent law, Abbott turns to the question of the human author within copyright law. He explains that, while the US Patent Office "has never issued guidance addressing the subject, and there is no case law on the issue of whether an AI could be an inventor, [t]here is…guidance available from the related issue of nonhuman authorship of copyrightable works" (p. 79). Following a review of the status of computer-generated works in the United Kingdom and the human-authorship requirement in the United States, he concludes: "Drawing an analogy from the copyright context, just as the terms 'writings' and 'authors' have been construed flexibly in interpreting the Patent

[12] *See* Carys J. Craig, *Critical Copyright Law and the Politics of 'IP'*, in RESEARCH HANDBOOK ON CRITICAL LEGAL THEORY 301, 306-310 (Emilios Christodoulidis, Ruth Dukes & Marco Goldoni, eds., 2019), https://ssrn.com/abstract=3287377.

and Copyright Clause, so too should the term 'inventors' be afforded the flexibility needed to effectuate constitutional purposes" (p. 87). I applaud the drive towards a purposive interpretation; but I worry about the inventor-author analogy. The result is to blinker the analysis to vital differences that should, taking a purposive approach, point to entirely different conclusions.

Patent law establishes rights in respect of inventions (or, more accurately, the knowledge embodied in novel inventions) while copyright law establishes a different set of rights in respect of literary, dramatic, musical, and artistic works (more accurately, the original expression embodied in those works). Simply put, under each system, different rights are granted to different actors in respect of different subject matters for the purpose of advancing different public policy objectives. Beyond the intangible nature of their respective subject matters, the two systems have very little in common. Abbott does not derive conclusions about patent law from the copyright realm, to be fair, but he does analogize between them when applying the principle of AI legal neutrality to both: according to Abbott, both AI inventorship and AI authorship should by default be treated by law just like human inventorship and human authorship. I disagree.

I have argued elsewhere, together with my coauthor Ian Kerr, that the very concept of AI authorship is oxymoronic.[13] Authorship, we suggested, is not simply the production of works as outputs but is rather a dialogic social practice: a relational act of communication. Understood in this way, authorship is necessarily beyond the reach of AI, which lacks the expressive agency and understanding necessary to engage in a dialogic exchange of meaning.[14] This assertion is not naïve to the current or potential capabilities of AI to create works that are facially "indistinguishable from works of human authorship."[15] Nor does it deny that copyright doctrine and its legal fictions could readily extend to encompass AI-generated works as copyrightable works should policy objectives push us in that direction.[16] Nonetheless, it insists that the very notion of "AI

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

[14] *Id,* at 85.

[15] Annemarie Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author* 2012 STAN TECH L REV 5 (2012).

[16] *See id.* (arguing that the US work-made-for-hire doctrine would offer an appropriate framework for resolving the ownership of AI-authored works by, for example, treating the AI programmer as equivalent to an employer). In the United Kingdom, as Abbott explains (p. 81-82), a legislative amendment already potentially extends copyright to AI-generated works by deeming the author of a computer-generated work to be "the person by whom the arrangements necessary for the creation of the work are undertaken": *Copyright, Designs and Patents Act 1988* (UK), s 9(3). Similar provisions exist in the Republic of Ireland, New Zealand, Hong Kong, and India.

authorship" rests on a category mistake: insofar as it equates AI generative processes with human expressive activities, it makes an error about the ontology of authorship.[17] Even if the AI generative processes and outputs may resemble their human-made counterparts, good copyright policymaking must begin by recognizing that they are fundamentally different in purpose and kind.[18]

Abbott's principle of legal neutrality, if applied, would suggest that these differences should be irrelevant: whether or not AI-generation of works *is* the same as human authorship, our law should treat them *as if* they are equivalent. Elsewhere, Abbott has arrived at this conclusion more explicitly than he does in the book, arguing that works should be protected by copyright if they are "generated by a computer in circumstances such that the computer, if a natural person, would be an author."[19] In so doing, it is to be presumed, the law would further the policy objectives of copyright (*i.e.* the progress of the useful arts) because "this would more accurately take into account contributions by machines, and allow economic incentives to work more efficiently."[20] AI should be regarded simply as stepping into the creative role traditionally reserved for human authors.

The purpose of copyright, however, 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 and reflect the creative agency vital to relational autonomy and human flourishing. Works are not just *things* that circulate; they are *speech* (textual utterances that serve a communicative function). Nor does copyright law create private rights simply for the sake of rewarding "authors," but rather because (and ideally, only to the extent that) these rights serve as "an engine of free expression,"[21] encouraging communication, creativity, and meaning-making.[22] Taking a purposive approach, then, copyright cannot be divorced from the fundamental values reflected in the right of free expression – a right that, as Abbott appears to acknowledge, does not and should not vest in AI. And so, to the extent that

[17] Craig & Kerr, *supra* note 13, at 42.

[18] *See also* Carys Craig, *AI and Copyright,* in ARTIFICIAL INTELLIGENCE AND THE LAW IN CANADA (Florian Martin-Bariteau & Teresa Scassa, eds., 2021) [*AI and Copyright*].

[19] Ryan Abbot, *Artificial Intelligence, Big Data And Intellectual Property: Protecting Computer-Generated Works In The United Kingdom,* in RESEARCH HANDBOOK ON INTELLECTUAL PROPERTY AND DIGITAL TECHNOLOGIES, 322 at 323 (Tanya Aplin, ed., 2020).

[20] *Id,* at 336.

[21] Eldred v. Ashcroft, 537 U.S. 186, 219, 221 (2003) (citing Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 558 (1985)). *See generally* NEIL WEINSTOCK NETANEL, COPYRIGHT'S PARADOX (2008).

[22] *See generally* CARYS J. CRAIG, COPYRIGHT, COMMUNICATION AND CULTURE: TOWARDS A RELATIONAL THEORY OF COPYRIGHT LAW (2011).

Abbott purports to justify giving AI (or its owners) the same copyright as human authors to advance the same constitutional goals and policy objectives, I would suggest that the reasoning is mistaken.

The answer to the conundrum of the so-called 'AI author,' Ian Kerr and I argued, "lies in a proper understanding of what copyright *is for*."[23] Through the instrumentalist frame, the first question we should be asking is not "how should we incentivize the maximum production of original works" but rather, "*why* do we aim to encourage the activity of authorship?" Framed in this way, the goal of incentivizing and so maximizing the production of AI-generated works seems to me to be very far from furthering the normative objectives of the copyright system. Indeed, there are many respects in which it may undermine them.[24] This, in turn, suggests that Abbott's principle of AI legal neutrality does not do what it should to put us on the path to good policymaking in the AI era – instead, it produces what is, in my view, precisely the wrong default position for copyright law. Once again, in the absence of any demonstrated need to specifically incentivize AI-generated works for their own sake, and without any *author* being denied recognition for their intellectual expression, there is no obvious or compelling reason to dislodge what ought to be the principled policy default: such works should freely circulate in the public domain.

I suggested above that the difference between patent inventorship and copyright authorship could provide a revealing window into why legal neutrality is insufficient – indeed, inapposite – as a principled starting point. Applying this principle, the default presumption is that AI inventors should be treated by law like human inventors and, similarly, AI authors should be treated in law like human authors. A look to the normative objectives of copyright and patent law, however, and differences in their respective subject matters and policy implications, suggests that this is not so. There may be good practical reasons to treat AI inventors this way, as Abbott reasonably contends, but there are also good reasons, as I have argued, not to approach the copyright question in the same fashion. If patent law is about maximizing inventions and innovative

[23] Craig & Kerr, *supra* note 13, at 43.

[24] *See* Craig, *AI and Copyright*, *supra* note 18 (arguing that unnecessarily extending the private preserve of copyright over such unauthored works would be contrary to "society's interest in maintaining a robust public domain that could help foster future creative innovation." (Citing CCH Canadian Ltd. v. Law Society of Upper Canada, 2004 SCC 13 at para. 23.)); *See also* Clark D. Asay, *Independent Creation in a World of AI*, 14 FIU L. REV. 201 (2020) (noting that, with the rapid proliferation of copyright-protected AI-generated works, the cultural landscape would become cluttered with "copyright landmines" and ever more difficult for human creators to navigate without legal risk.) But perhaps the greatest harm, in the long run, would flow simply from undermining to the ontological category of "author."

products, copyright is about encouraging creative *processes* and expressive *activities*. There is simply no reason to assume that applying legal neutrality across the board will advance the same social values or preserve a normative equilibrium – so why start there? Why not start by asking what rights, responsibilities, and legal relationships will advance the objectives of the law and social values they reflect? In Part III, I suggest that this is where we must begin if we hope to ensure that AI will help to foster – not subvert – these larger ambitions.

III. THE NORMATIVE LIMITS OF LEGAL NEUTRALITY

Abbott's version of legal neutrality is concerned with identifying functional equivalence in the activities or outputs of AI and humans, and so proposes extending the law's protections and prescriptions "neutrally" where AIs perform actions equivalent in effect to those of their human counterparts. This version of legal neutrality in the face of technological change is something that I have characterized elsewhere as an "intermediate" approach to regulating new technologies: it falls somewhere in between a narrow concept of media neutrality and an expansive principle of substantive technological neutrality.[25] Abbott resists the narrow, formalist approach when he urges lawmakers to look beyond the words of the relevant statutes or existing doctrine, and not to get caught up with applying established rules to the internal operations of AI.[26] Instead, he rightly focuses attention on the functional effects of AI, calling for the consistent application of the law across categories despite the novel technological processes they employ. There are, however, still significant normative limits to this approach.

A. The Risks of Relying on Functional Equivalence

The use of analogy to compare activities across technologies (in Abbott's case, to compare human behaviors to AI operations) can be misleading, causing decisionmakers to overlook differences and fail to perceive the larger paradigm shifts that occur through technological change.[27] The result is a kind of *technology-*

[25] *See* Carys Craig, *Technological Neutrality: Recalibrating Copyright in the Information Age*, 17 THEORETICAL INQUIRIES L. 601 (2016) [*Recalibrating Copyright*]. *See also* 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 (Michael Geist, ed., 2013) [THE COPYRIGHT PENTALOGY]; Gregory R. Hagen, *Technological Neutrality in Canadian Copyright Law*, in THE COPYRIGHT PENTALOGY (*id.*).

[26] *See* Brad A. Greenberg, *Rethinking Technology Neutrality*, 100 MINN. L. REV. 1495 (2016).

[27] *See* Craig, *Recalibrating Copyright*, *supra* note 25, at 611-12.

blindness that can produce, in practice, the very inequities and inconsistencies of result that neutral treatment purports to prevent. If we're honest, the neutrality of both law and technology is always illusive – a myth too often invoked to obscure the realities of substantive inequality, self-interest, or political agenda – and, in this context, it is unlikely to be any different. As such, we should be particularly wary of willful tech-blindness and false claims of functional equivalence when what we need, instead, is clear-sighted recognition of the disruptive force of technological change, the disadvantages it so often compounds, and the practical and political differences it makes.

The neutral treatment of functional equivalents is all well and good in the abstract; but the reality is that "some technologies change the game."[28] In the areas that Abbott explores (and many more besides), AI is a gamechanger. As such, AI legal neutrality may fail to distinguish between contexts where neutral treatment of functional equivalents will produce normative equilibrium and those where it will disrupt the equilibrium, distorting legal constructs, producing undesirable outcomes, and turning guiding principles on their head.

Another normative limit to the vision of legal neutrality is its preference for the regulatory *status quo*. The premise of *The Reasonable Robot* is that, by default, we should place the same standards of behavior and expectations on AI as we do on humans, thereby imposing human-centered norms on AI's operation (and to some extent *vice versa*). But if AI is being interspersed – sometimes invisibly, often insidiously – into the midst of human communities, lives, and relationships, why should we settle for imposing *the same* norms and standards on AI? AI can and should be designed, trained, operationalized, and held to account in a manner that actually improves our human conditions (as opposed to merely to reproducing them more efficiently). Why not impose a higher standard of duty of care, for example, on the programmable AI than we could reasonably expect from our fellow human beings? Why would we not demand more of the machine in the name of advancing social goals, especially when the machine can – without any moral objection – be made the means to our social ends? In this sense, legal neutrality seems to set our normative ambitions too low.

But more than that, as the fiction of "the reasonable robot" suggests, a legal neutrality principle focused on functional equivalence might cause us to overlook, forget, or even fundamentally misunderstand what robots *are* and what they *are not*. The robot is not actually "reasonable," of course, because the robot has no capacity for *reason*. The robot may be programmed to behave in accordance with a standard duty of care; but the robot is incapable of *caring*. It

[28] Kerr & Szilagyi, *supra* note 2, at 349.

may be trained to generate original works, but it is incapable of any act of *expression*. Abbott is not guilty of the kind of techno-romanticism or sentimentalism that would seek to refute these distinctions. But my complaint is a more practical one: there is a risk inherent in this kind of categorical blurring that necessarily accompanies AI legal neutrality and its reliance on false equivalence. It lays the groundwork for misconceptions – mischaracterizing the role of technology, anthropomorphizing the AI, romanticizing the robot – that will surely misdirect our regulatory responses.

Related to this concern is the added risk of obfuscation that reliance on functional equivalence may bring. If we treat AI like humans, do we not hide the actual humans and their corporate structures (their motives, their activities, their power) behind the AI? The robot steps in as the relevant subject thereby permitting the obfuscation of human contributors (the human sources of training data, the programmers, users, owners, and operators) and cloaking their roles and responsibilities in respect of the AI.[29] This is not Abbott's intention, to be sure, as chapters on tort law and criminal punishment make clear. Nonetheless, the fictions of AI-human equivalence can elide the realities of human-robot relations and their unique dynamics in ways that do more to disguise than reveal the human actors, organizations, and power relations behind the AI artifacts.

B. Substantive Tech-Neutrality and the Relational Approach

So, what is the alternative? In the copyright context, I have previously proposed a more expansive principle of *substantive* technological neutrality that looks beyond equal treatment of functional equivalents and focuses instead on the normative objectives of law. As with AI legal neutrality, the principle is by no means dispositive of any policy issue, but it helpfully illuminates the questions that should guide lawmaking. Fundamentally, it reminds us to ask: Given the new realities of the changing technological environment, what rights and responsibilities should the law afford — and subject to what limits — in order to maintain its normative balance and advance its underlying policy objectives? In this sense, the default position is *normative neutrality* (the law's overarching and animating objectives are presumed to remain the same) but adjustments to old

[29] *See* Craig & Kerr, *supra* note 13, at 69-70 ("when we substitute an AI for a human, we are permitting the AI to stand in for significant human expressive activity and relations of communication that occur, invisibly, behind the scenes."). *See also* Astra Taylor, *The Automation Charade* (1 August 2018), logicmag.io (Warning of the "fauxtomation" promulgated by "giving automation more credit than it's actually due. In the process, we fail to see — and to value — the labor of our fellow human beings.")

rules and differential treatment of new technologies are readily endorsed as plausible ways to maintain that equilibrium.

Taking this approach, neutrality is not about the equal treatment of analogous activities but the consistent pursuit of established policy goals when new technologies disrupt our regulatory landscape. If we say, for example, that the goals of copyright are to encourage authorship and foster a vibrant public domain, then we should review the law's treatment of AI-generated works by asking whether their protection under the existing system would advance or undermine these goals.[30] We should not simply analogize between AI and human authorship and extend rights on that basis. In this way, a more substantive vision of normative neutrality contemplates the recalibration of legal doctrine – or even its wholesale reimagination – in response to technological gamechangers like AI.

As we move from a narrow to a more expansive vision of what technological neutrality should entail, so too should the focus shift from individual actors and artifacts to the relationships that the law constructs through legal rights and responsibilities. The law should structure legal relations in ways that reflect and advance shared and basic human values: human agency, autonomy, equality, free expression, and so forth. Taking a relational approach to the regulation of AI, then, we would start by thinking purposively about the relationship between AI and humans: What do we want that relationship to look like? What functions do we want it to fulfil? What human capacities do we want it to foster? Our laws should aim to shape human-robot relations in a manner consistent with these values. It may be that, in some situations, we will build the most beneficial human-AI relations by treating the AI in law as we do humans, as Abbott proposes; but far more often than not, I suspect, the law will serve its public policy purposes best when it recognizes the unique dynamics of human-AI relationships and the social values at stake in their regulation.

Abbott posits the idea of the Reasonable Robot: AI that is subject to the same legal rules and standards of behavior as humans when engaged in the same activities. But my suggestion is that we consider, instead, the Relational Robot. The starting point for our norm-setting should be the AI not *as if* but *as it is* – in its social context, within the human networks in which it is operating, and always with a clear view to the nature of the *thing* (without inviting inapt analogies, the misattribution of human characteristics, or romantic imaginings). If AI legal neutrality is aimed at shaping incentives and finding efficiencies in the regulation of AI, a more substantive approach to normative neutrality is aimed at

[30] See Craig, *AI and Copyright*, *supra* note 18, at Part 1.4.

shaping human-AI relations, and finding out how best to advance our policy goals with a view to the basic human values at stake.

IV. CONCLUSION

'Reasonableness' in tort law does not purport to ascribe characteristics or attributes to particular people but is rather a kind of heuristic device that performs, at least in part, a prescriptive function, shaping our social, physical, and commercial interactions in pursuit of the law's normative objectives. The "reasonable person" does not function as a wholly objective, stable, or predetermined conceptual entity, but as a context-specific standard of behavior with a subjective component, to be applied to complex, situated interactions.[31] To posit the "Reasonable Robot" in any meaningful way is, in a sense, to imagine the *robot-in-relation*, starting with a set of assumptions about the role of the law in shaping our relationships with robots in order to foster or maintain the kind of society we would want to inhabit alongside AI.

Abbott has the ingredients, then, for a more fulsome normative vision of how we should go about regulating AI. But he doesn't ultimately use them in his recipe. Looking at robots in relational context reveals what is lost when we serve up a legally neutral starting point. Our human-machine relationships are not substantively equivalent to human-human relationships – we want to and must understand, value, and regulate them *differently*. Exceptional circumstances or specific contexts may present reasons to treat AIs *as-if* they are humans when performing certain tasks for particular purposes, to be sure;[32] but that is a very different thing from taking neutrality as a principled starting point and reasoning backwards from there, requiring justifications for differential treatment as an exceptional departure from the default norm.

Normative defaults are necessary in times of disruption. As Ian Kerr recognized, "the future is full of question marks, and default rules fill the knowledge-gaps that may arise in future dealings."[33] But the devil is in the defaults – and the equal legal treatment of AI and human actors could quickly become a dangerous default setting.

---

[31] *See generally* Mayo Moran, *The Reasonable Person: A Conceptual Biography in Comparative Perspective*, 14 LEWIS & CLARK L. REV. 1233 (2010).

[32] *See* Jack M Balkin, *The Path of Robotics Law* 6 CAL L REV 45, 55 (2015) (describing the "substitution effect" that occurs when, for particular limited purposes, we treat AIs as special-purpose human beings). *See also* Craig & Kerr, *supra* note 13, at 59-60 ("One must be extremely cautious not to allow the substitution to blur the underlying ontological category that is being substituted. AIs are not persons even if there is practical value, in limited circumstances, in treated them as such.")

[33] Kerr, *The Devil Is in the Defaults*, *supra* note 1, at 100.