



Osgoode Hall Law School of York University

Osgoode Digital Commons

All Papers

Research Papers, Working Papers, Conference
Papers

2021

AI and Copyright

Carys Craig

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](#)

Chapter 1

AI and Copyright

Carys J. Craig

Overview

This chapter examines the most pertinent issues facing copyright law as it encounters increasingly sophisticated artificial intelligence (AI). It begins with a few introductory examples to illuminate the potential interactions of AI and copyright law. Section 1 tackles the question of whether AI-generated works are copyrightable in Canada and who, if anyone, might own that copyright. This involves a doctrinal discussion of “originality” (the threshold for copyrightability) as well as reflections on the meaning of “authorship,” and concludes with the suggestion that autonomously generated AI outputs presently (and rightly) belong in the public domain. Section 2 turns to consider issues of copyright infringement. First, it addresses the law in respect of AI inputs (the texts and data used to train AI systems, which may themselves be copyrightable works) and highlights the need for greater limits and exceptions to ensure that copyright law does not obstruct best practices in the development and implementation of AI technologies. It then examines the matter of potentially infringing AI outputs (which may, of course, resemble copyright-protected, human-created works), identifying current uncertainties around independent creation, agency, and the allocation of liability. Section 3 addresses the deployment of AI in automated copyright-enforcement, emphasizing its increasingly critical role in shaping our online environment and citizens’ everyday encounters with copyright enclosures. The chapter concludes with reflections on the risks and opportunities presented by AI in the copyright context, and identifies key gaps and questions that remain to be answered as copyright law and policy adjust to evolving AI technologies.

Key Challenges and Issues

The key challenges and issues in this area are as follows :

- With AI now producing outputs that are facially indistinguishable from works of human authorship, it must be established whether copyright subsists or ought to subsist in such AI-generated works, or whether they more appropriately belong in the public domain.
- If it is determined that copyright does or should attach to the outputs of AI systems, it remains to be established who—or what—should be the first owner in the absence of an identifiable human author.
- There is a pressing need for adequate and explicit exceptions for text and data mining to ensure that copyright law does not obstruct or distort the research, development, and operation of AI systems.
- There is a lack of clarity around the potential primary and secondary copyright liability of programmers, providers, and users of AI systems in relation to infringing AI processes and outputs.

RECOMMENDED CITATION

Carys Craig, “AI and Copyright” in Florian Martin-Bariteau & Teresa Scassa, eds., *Artificial Intelligence and the Law in Canada* (Toronto: LexisNexis Canada, 2021), ch. 1.

Discover the full collection at <http://aisociety.ca/ailawbook>

- AI and machine learning are playing an increasingly pervasive role in the automated enforcement of copyright online in a manner that threatens to undermine due process and upset the traditional balance of copyright owners' and users' rights.
- There is an increasingly urgent need to identify and assess the potential impact of sophisticated AI technologies on the pursuit of copyright policy objectives, and to develop legislative and regulatory responses that ensure copyright's substantive technological neutrality as AI continues to evolve.

Synopsis

| | |
|---|----|
| Introduction | 3 |
| 1. The Copyrightability of AI-Generated Works: Authorship and Ownership | 5 |
| 1.1. Conditions for the Subsistence of Copyright | 5 |
| 1.2. Ownership of Copyright | 6 |
| 1.3. The Obstacle of Originality | 7 |
| 1.4. A Purposive Approach to the Copyrightability of AI outputs | 9 |
| 2. AI and Copyright Infringement..... | 11 |
| 2.1. AI Inputs: Training Algorithms | 11 |
| 2.2. AI Outputs: Liability and for Whom/What?..... | 13 |
| 2.3. AI and Automated Copyright Enforcement..... | 15 |
| 3. Risks and Opportunities | 16 |
| 4. Key Gaps in the Law..... | 17 |

Introduction

“The rise of the machines is here,” it has been said, “but they do not come as conquerors, they come as *creators*.”¹ Once the distant imaginings of philosophers and science fiction writers, artificially intelligent machines have, in recent years, become capable of generating artifacts—literary, dramatic, musical, and artistic works—that possess all of the external attributes of human creativity. If we have arrived in an age of “digital authorship”—an age in which digital works (themselves artifacts) can, “relatively autonomously, produce other works that are indistinguishable from works of human authorship”²—then we have surely entered a newly tumultuous time for copyright law and policy.

The copyright system, which dates back to the early 18th century and the days of the printing press, is the body of law that grants exclusive rights over works of authorship: the sole right to copy or reproduce a work, to publish it, and to perform it or communicate it to the public.³ It also grants moral rights of attribution and integrity to authors,⁴ and establishes a system of neighbouring rights for performers, broadcasters, and the makers of sound recordings.⁵ With each major development in recording and information technologies over the past three centuries, from the Player Piano Roll to the cassette tape, and from the invention of radio to the World Wide Web, copyright law has responded, adapted, and (rightly or wrongly) expanded to embrace new forms of communication and economic exploitation. But with each paradigm-changing shift comes new policy quandaries, and new opportunities to rethink old assumptions and established systems of control.

Just as the arrival of the Internet and digital technologies disrupted long-held assumptions about the copying and dissemination of expressive works that had underpinned the copyright system from its inception, so too is artificial intelligence (AI) now disrupting long-held assumptions about creative production and consumption. Perhaps the law will simply adapt and expand to incorporate this new technology, as it has done so many times before—or perhaps this latest technological revolution will be the one that fundamentally alters or unseats the copyright system as we know it. Time will tell, but for now the task at hand is to identify some of the most pertinent issues that copyright law will surely have to address as it encounters the rising challenge of AI.

Recent high-profile examples of AI-generated works span the full range of human cultural endeavour from music to film, and from literature to the visual arts. By way of example, *Jukebox* is a machine-learning model capable of generating music that imitates different styles and artists and even incorporates singing in natural-sounding voices,⁶ while the self-named *Benjamin* is the neural network that generated the screenplay for *Sunspring*, an award-winning science fiction film.⁷ *TalkToTransformer* is an AI language generator, created by Canadian engineer Adam King using the OpenAI GPT-2 technology, which can write articles when

¹ Andres Guadamuz, “Artificial Intelligence and Copyright” (2017) 5 WIPO Magazine 14 at 17 [emphasis added], online: *WIPO Magazine* https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html.

² Annemarie Bridy, “Coding Creativity: Copyright and the Artificially Intelligent Author” (2012) 5 *Stan. Tech. L. Rev.* 1 at 3.

³ *Copyright Act*, R.S.C. 1985, c. C-42, s. 3. See also *Entertainment Software Association v. Society of Composers, Authors and Music Publishers of Canada*, 2012 SCC 34 at para. 42.

⁴ *Copyright Act*, R.S.C. 1985, c. C-42, s. 14.1. Performers are also granted certain moral rights under s. 17.1.

⁵ *Copyright Act*, R.S.C. 1985, c. C-42, ss. 15, 18, 21.

⁶ Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford & Ilya Sutskever, “Jukebox: A Generative Model for Music” (30 April 2020), online: *arXiv* <https://arxiv.org/pdf/2005.00341.pdf>. Curated examples of Jukebox’s outputs are available online: <https://openai.com/blog/jukebox/>.

⁷ Oscar Sharp, “Sunspring” (9 June 2016), online: *YouTube* <https://www.youtube.com/watch?v=LY7x2Ihqjmc>.

prompted with a headline or complete short stories when fed the first line.⁸ *Dio* is a sculptural work from artist Ben Snell, which was generated by a machine-learning algorithm trained on a dataset of classical sculptures (and then made out of the pulverized dust of the machine that designed it).⁹ Perhaps most famously, *Portrait of Edmond Belamy* is an AI-generated painting that went under the hammer at Christie's Auction House in 2018, selling for almost half a million dollars and thereby signalling, according to Christie's, "the arrival of AI-generated art on the world auction stage."¹⁰

A final example from the visual arts might helpfully ground the discussion that follows. Canadian artist Adam Basanta recently attracted attention—and a copyright infringement claim—for his project *All We'd Ever Need Is One Another*. This mixed-media installation autonomously generates over a thousand images each day using two flatbed scanners, tipped over and facing each other, continuously running scanning cycles on randomized software-controlled settings. The images are then subjected to an automated "validation process" conducted by a machine-learning algorithm trained to identify patterns that resemble existing images in a database of contemporary abstract art.¹¹ When a randomly generated image bears at least an 83% likeness to a known artwork, it is automatically uploaded to a dedicated website where it is displayed as an "art-factory" output, and assigned a title that cross-references the similar human-made art.¹²

In what some had hoped would become Canada's first AI-meets-copyright lawsuit, Montreal artist Amel Chamandy alleged that Basanta had infringed copyright in her photographic work, *Your World Without Paper*, in the production of an image identified as an 85.81% match. Chamandy's statement of claim alleged that "the process used by the Defendant to compare his computer-generated images to Amel Chamandy's work necessarily required an unauthorized copy of such a work to be made," and sought statutory damages of up to \$20,000.¹³ While the parties have now settled, the scenario provides a window into thinking through the copyright issues potentially provoked by AI-generated works, from copyright's subsistence to its infringement. First up is the question of whether the products or outputs of AI systems are the kind of original works of authorship in which copyright might vest.

⁸ James Vincent, "Use This Cutting-Edge AI Text Generator to Write Stories, Poems, News Articles, and More" (13 May 2019), online: *The Verge* <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" (14 February 2019), online: *The Verge* <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/>.

⁹ James Vincent, "This AI-generated sculpture is made from the shredded remains of the computer that designed it" (12 April 2019), online: *The Verge* <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/>.

¹⁰ Christie's, "Is Artificial Intelligence Set to Become Art's Next Medium?" (12 December 2018), online: *Christie's* <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>.

¹¹ Adam Basanta explains the process in a video, "All We'd Ever Need is One Another" (19 May 2018), online: *Vimeo* <https://vimeo.com/268772915>.

¹² Adam Basanta, "All We'd Ever Need is One Another" (Last accessed 18 October 2020), online: *All we'd ever need is one another* <http://allwedeveryneed.com/>.

¹³ Statement of Claim at para. 30, quoted by Teresa Scassa, "Artist Sued in Canada for Copyright Infringement for AI-Related Art Project" (4 October 2018), online: *Teresa Scassa* http://www.teresascassa.ca/index.php?option=com_k2&view=item&id=286.

1. The Copyrightability of AI-Generated Works: Authorship and Ownership

To be clear, copyright protection is not hard to get. Machine-generated works may, for now, typically lack the aesthetic quality or conceptual coherence of their human-made counterparts (the art-factory output was no match for Chamandy’s evocative imagery or message; *TalktoTransformer*’s stories rarely made much sense; parts of the *Sunspring* script were unintelligible nonsense; and the mythical *Edmond Bellamy* was missing a nose)—but there is no aesthetic quality bar to copyrightability. If humans had directly created any of these works, copyright would attach to them. In Canada, as elsewhere, “the general irrelevance of aesthetics has become a cornerstone of copyright jurisprudence.”¹⁴ Far from requiring human ingenuity, copyright vests without much quibble in the most mundane and utilitarian of works.¹⁵ Nor is there a bureaucratic or administrative barrier to obtaining protection: copyright vests automatically in “every original literary, dramatic, musical and artistic work”¹⁶ as soon as the work is fixed in a more or less permanent tangible form¹⁷—a subject matter that is defined to “include[] every original production in the literary, scientific or artistic domain.”¹⁸ The standards for copyright protection are not qualitatively high; there are, nevertheless, significant *substantive* barriers to copyrightability that could (and, in my view, should) exclude machine-generated works *ab initio* from the current copyright system.

1.1. Conditions for the Subsistence of Copyright

Copyright’s subsistence in a work is conditional on meeting a geographical requirement that ties the work to Canada or its international treaty partners. Specifically—and tellingly—copyright shall subsist if “the author was, at the date of the making of the work, a citizen or subject of, or a person ordinarily resident in, a treaty country.”¹⁹ The geographic condition can also be met by first publication or, in the case of cinematographic works, the location of the maker’s corporate headquarters.²⁰ However, the significance of this provision lies in the obvious implication that an “author” of a “work” is a *natural person*—a citizen or, in any event, someone who *resides* somewhere.

The conclusion is further reinforced by the duration of copyright’s subsistence in a work, which is determined by the natural life span of the author. As required by the *Berne Convention* (*Berne*), copyright in Canada subsists for “the life of the author, the remainder of the calendar year in which the author dies, and a period of fifty years following the end of that calendar year.”²¹ Sam Ricketson noted almost 30 years ago that

¹⁴ Alfred C. Yen, “Copyright Opinions and Aesthetic Theory” (1998) 71:2 S. Cal. L. Rev. 247 at 249. See *e.g.* *Hay v. Sloan*, [1957] 12 D.L.R. (2d) 397 at 398–400 (Ont. H.C.); *DRG Inc. v. Datafile Ltd.*, [1988] 2 F.C. 243 at para. 12 (F.C.T.D.).

¹⁵ See *e.g.* *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13; *Delrina Corp. v. Triolet Systems Inc.*, [1993] O.J. No. 319 (Ont. Ct. J.G.D.), *aff’d* (2002), 58 O.R. (3d) 339 (Ont. C.A.); *U & R Tax Services Ltd. v. H & R Block Canada Inc.*, [1995] F.C.J. No. 962 (F.C.T.D.); *Keatley Surveying Ltd. v. Teranet Inc.*, 2016 ONSC 1717; *Geophysical Service Inc. v. Encana Corp.*, 2016 ABQB 230.

¹⁶ *Copyright Act*, R.S.C. 1985, c. C-42, s. 5(1). This is subject to conditions in paragraphs (a)–(c), discussed in Section 1.1.

¹⁷ *Canadian Admiral Corp. Ltd. v. Rediffusion Inc.*, [1954] Ex. C.R. 382 (Ex. Ct.).

¹⁸ *Copyright Act*, R.S.C. 1985, c. C-42, s. 2.

¹⁹ *Copyright Act*, R.S.C. 1985, c. C-42, s. 5(1)(a). A treaty country is defined in s. 2 to mean “a Berne Convention country, UCC country, WCT country or WTO Member”.

²⁰ *Copyright Act*, R.S.C. 1985, c. C-42, s. 5(1)(b)–(c).

²¹ *Copyright Act*, R.S.C. 1985, c. C-42, s. 6. See also *Berne Convention for the Protection of Literary and Artistic Works* (1886, as amended by *Paris Act* of 1971), Art. 7(1).

“such a provision would be inappropriate in the case of non-human entities, which may have an infinite existence.”²² There are some copyright-protected subject matter for which term is not pegged to a natural life—works falling under Crown copyright, non-dramatic cinematographic works, performer’s performances, broadcasts, and sound-recordings. With the exception of Crown copyright, however, these are categories of *unauthored* subject-matter. For works of authorship, the clear implication is, again, that the author is a natural person—with heirs and an estate—who lives and dies. Machines do not die (because, of course, machines do not live).

The author is also, under *Berne*’s mandate, vested with moral rights that protect his or her honour and reputation from prejudice, and entitle the author to be associated with the work or to remain anonymous.²³ These unassignable moral rights, which can be bequeathed or passed with the estate and similarly last 50 years after death, assume an intimate personal connection between the author and his or her original work. In Ricketson’s assessment, such protections “make no sense other than in relation to human authors.”²⁴ Once again, the law’s prescriptions appear to assume a human author—one with a finite life, a sense of honour, and a personal reputation to uphold.

While there is no explicit statutory requirement that an author be human, then, there is textual support within Canada’s *Copyright Act* from which such a requirement can readily be inferred, and judicial efforts to define the nature of original works of authorship lend further credence to this conclusion.²⁵

1.2. Ownership of Copyright

It might be objected, at this point, that there are many instances in which copyright is held by a non-human entity. Indeed, copyright functions in the modern marketplace much as any other alienable commodity that can be freely transferred or licensed for value or otherwise,²⁶ with the result that it is frequently owned by corporate persons as opposed to those of the human variety. More to the point, any copyright in works made in the course of employment is owned by the employer—typically a corporate entity—from the moment that the right vests, never belonging to the human author at all.²⁷ And we have already seen that works falling under Crown copyright belong not to a human author, but to the Crown. It is important, then, to parse the matters of authorship and ownership, and to emphasize that, in Canada, all copyrightable “works” (with the arguable exception of non-dramatic cinematographic works) have *authors*—and more importantly, that these authors are, invariably, human.

Here, a distinction should be drawn with the United States (US) copyright doctrine applicable to works for hire. Where a work falls within the work-for-hire category, the US law deems the employer to be the author of the work and pegs its duration to the work’s creation or first publication rather than the author’s date of death.²⁸ Some US commentators have suggested that it would therefore not be a particularly significant stretch

²² Sam Ricketson, “People or Machines: The Berne Convention and the Changing Concept of Authorship” (1991) 16:1 Colum. V.L.A. J. L. & Arts 1 at 11.

²³ *Copyright Act*, R.S.C. 1985, c. C-42, ss. 14.1, 28.1. See also *Berne Convention for the Protection of Literary and Artistic Works* (1886, as amended by *Paris Act* of 1971), Art. 6bis.

²⁴ Sam Ricketson, “People or Machines: The Berne Convention and the Changing Concept of Authorship” (1991) 16:1 Colum. V.L.A. J. L. & Arts 1 at 11.

²⁵ See Section 1.3.

²⁶ *Copyright Act*, R.S.C. 1985, c. C-42, s. 13(4).

²⁷ *Copyright Act*, R.S.C. 1985, c. C-42, s. 13(3).

²⁸ 17 U.S.C. § 101 (defining “work made for hire”), § 302(c) (specifying a terms of 95 years from first publication or 120 years creation, whichever expires first).

of this legal fiction to similarly deem authorship to lie with a non-human author in the case of AI-generated works.²⁹ But in Canada, the person who creates a work in the course of employment, though never the copyright owner, is and remains the “author”; his or her lifetime determines the copyright’s duration; and moral rights vest in the author unless waived. Of course, deeming provisions and legal fictions are available tools for lawmakers in Canada too—distinctions can always be statutorily drawn between authors-in-fact and authors-in-law³⁰—but the departure from current copyright doctrine would be considerably more profound. Any such distinction would require legislative amendment and careful consideration of who (or what) should be deemed to be the author-in-law of AI-generated works—whether the AI programmer, its user, or the AI itself (in which case, first ownership would have to be allocated elsewhere).

The underlying problem, however, is that unlike conventional works-made-for-hire, in the case of AI-generated works, there is no human author-in-fact. Such a legal fiction would therefore risk undermining the concept of authorship as central to copyright law—and also as a particular sort of human expressive endeavour that copyright is intended to encourage.³¹

1.3. The Obstacle of Originality

In copyright doctrine, the corollary of authorship is originality. The main substantive gate-keeper of copyrightability—the “foundation stone of copyright”³²—is the requirement that a work be “original.”³³ It is not sufficient that a work *looks like* something a person could have done: the originality threshold is not an objective assessment but a subjective one, which asks us to examine not the output of a production process, but the *process* that led to the output. In *CCH Canadian Ltd. v. Law Society of Upper Canada*³⁴ (*CCH*), the Supreme Court of Canada has defined this to mean that a work “must be more than a mere copy of another work,” but it “need not be creative, in the sense of being novel or unique.”³⁵ Rather, in order to be protected, an author’s expression must involve “skill and judgment.” “Skill” is defined as “the use of one’s knowledge, developed aptitude or practised ability in producing the work,” while “judgment” involves “the use of one’s capacity for discernment or ability to form an opinion or evaluation by comparing different possible options in producing the work.”³⁶ The amount of skill and judgment involved “must not be so trivial that it could be characterized as a purely mechanical exercise.”³⁷

The threshold is low and is, of course, likely to be met by the programmer of the code on which an AI system runs. The computer program (however functional it may be) is treated as a protected literary work in its own right.³⁸ Moreover, the originality threshold may be met by the software user who employs the program

²⁹ See Annemarie Bridy, “The Evolution of Authorship: Work Made by Code” (2016) 39:3 Colum. J. L. & Arts 395 at 400.

³⁰ Annemarie Bridy, “Coding Creativity: Copyright and the Artificially Intelligent Author” (2012) 5 Stan. Tech. L. Rev. 1 at 24–26. Such a distinction used to be drawn for the authorship of photographs under *Copyright Act*, R.S.C. 1985, c. C-42, s. 10(1) (repealed by the *Copyright Modernization Act*, R.S.C. 2012, c. 20, s. 6).

³¹ Carys Craig & Ian Kerr, “The Death of the AI Author” (2020) 52:1 Ottawa L. Rev. 31 at 61–62..

³² *Robertson v. Thomson Corp.*, 2006 SCC 43 at para. 35.

³³ *Copyright Act*, R.S.C. 1985, c. C-42, s. 5(1).

³⁴ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13.

³⁵ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 16.

³⁶ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 16.

³⁷ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 16.

³⁸ *Copyright Act*, R.S.C. 1985, c. C-42, s. 2, defined “literary work” as including “computer programs,” which are in turn defined as “a set of instructions or statements, expressed, fixed, embodied or stored in any manner, that is to be used directly or indirectly in a computer in order to bring about a specific result.” See also *Delrina Corp. v. Triolet Systems Inc.*, [1993] O.J. No. 319 (Ont. Ct. J.G.D.), aff’d (2002), 58 O.R. (3d) 339 (Ont. C.A.).

as a technical tool to assist in the production of works that nonetheless ultimately reflect his or her own authorial skill and judgment.³⁹ Where such technical instruments require human input or intervention in the production of a work, the skilled human user is likely to be hailed an original author (tenuous though the claim may be).⁴⁰ What remains to be determined, however, is whether the originality threshold can be met by the outputs of an autonomous AI over which both the human programmer and user have limited influence and little or no authorial control.

The *CCH* test has thus far been assumed to mean that “a human author is required to create an original work for copyright purposes.”⁴¹ But does “developed aptitude or practised ability” necessitate a human capacity for learning and improvement that is fundamentally distinct from the machine-learning processes accomplished by AI? Do the “deep learning methods” of a Generative Adversarial Network, for example, involve or merely mimic the kind of “developed aptitude” that legal doctrine demands? Are machines, though surely incapable of “forming an opinion,” nonetheless exercising a “capacity for discernment” when they select between different available options (as when Basanta’s algorithms select the best match)? Does an automated process, though entirely dependant on the copying of a vast quantity of data, involve more than “mere copying”? And finally, is the AI-generative process, though carried out by a machine, more than purely “mechanical”?

One can imagine reasonably compelling legal arguments that point to technical AI processes, deliberately designed to emulate human creativity, and analogize them to the mental mechanics, if you will, of human authorship. Yet, when one takes a step back, brushes off the “doctrinal mud,”⁴² and simply asks whether an AI system is *itself* exercising “intellectual effort”⁴³ in the “expression of ideas,”⁴⁴ it seems abundantly clear that the answer must be *no*. Ultimately the originality inquiry is an attempt to establish whether the intellectual process involved was an *authorial* one. The real question underlying this inquiry, then, is whether machines, by definition, can *be* authors. As argued elsewhere, the answer to this larger ontological question is a categorical *no*.⁴⁵ Authorship entails expressive agency—an intention to communicate, to engage in dialogic relations with others—that AI simply cannot possess. The illusion of communication with a machine is readily generated by today’s AI and is predictably exacerbated by our natural tendency to anthropomorphize robots; but it is simply a category mistake to hail the AI as “a thinking thing” capable of engaging in the activity of *authorship* that the copyright system is intended to encourage.⁴⁶

³⁹ See *Geophysical Service Incorporated v. Encana Corporation*, 2016 ABQB 230 at para. 83 (finding sufficient skill and judgment in processed seismic data because “[t]he raw data is not simply pumped into a computer and a useful product comes out. The evidence is clear that the processed product can be quite different depending on the skill of the processor”).

⁴⁰ See *Geophysical Service Incorporated v. Encana Corporation*, 2016 ABQB 230 at para. 91 (finding sufficient originality in raw seismic data because “[e]ven though many technical instruments are used in the production of seismic data, they require human intervention, in the form of expert scientific skill and judgment to make them work”). However, the author disagrees with the Court’s conclusion in this case that the skill and judgment involved was authorial in nature.

⁴¹ *Geophysical Service Incorporated v. Encana Corporation*, 2016 ABQB 230 at para. 88.

⁴² See Pamela Samuelson, “Allocating Ownership Rights in Computer-Generated Works” (1985) 47:4 U. Pitt. L. Rev. 1185 at 1200: “Only those stuck in the doctrinal mud could even think that computers could be ‘authors’.” See also Carys Craig & Ian Kerr, “The Death of the AI Author” (2020) 52:1 Ottawa L. Rev. 31.

⁴³ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 16.

⁴⁴ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 22.

⁴⁵ Carys Craig & Ian Kerr, “The Death of the AI Author” (2020) 52:1 Ottawa L. Rev. 31.

⁴⁶ Diane Proudfoot, “Anthropomorphism and AI: Turing’s Much Misunderstood Imitation Game” (2011) 175:5-6 *Artif. Intell.* 950 at 954. See generally Carys Craig & Ian Kerr, “The Death of the AI Author” (2020) Ottawa L. Rev. [forthcoming], online: SSRN <https://ssrn.com/abstract=3374951>.

Even putting aside ontological inquiries in favour of pragmatic policy-making, however, a purposive approach to copyrightability leads to a similar conclusion.

1.4. A Purposive Approach to the Copyrightability of AI outputs

Notably, the Canadian originality standard is said to lie between the “industriousness” standard for originality traditionally found in the United Kingdom (UK) and common law jurisdictions, and the “creativity” approach adopted in the US and traditionally associated with civil law jurisdictions.⁴⁷ Divergent approaches have thus far been taken towards machine-generated works in jurisdictions with different originality thresholds (which in turn reflect different underlying philosophies of copyright and its purpose).

In the UK, for example, where the traditional originality threshold was low and copyright was typically regarded as a reward for labour, a legislative amendment ostensibly swept computer-generated works into copyright’s protective fold by deeming the author to be “the person by whom the arrangements necessary for the creation of the work are undertaken”⁴⁸ (but limiting protection to a flat 50 years and denying any moral rights).⁴⁹ Meanwhile, in Europe, where copyright is philosophically regarded as a matter of *droit d’auteur*, it is widely agreed that machine-generated works fail to satisfy the threshold requirement that the work expresses “the author’s own intellectual creation”⁵⁰—by which it is meant that “it reflects the author’s personality” in the sense that “the author was able to express his creative abilities in the production of the work by making free and creative choices.”⁵¹ In the US, with its more utilitarian conception of copyright’s purpose, originality also requires at least a minimal degree of creativity and entails a human authorship requirement.⁵² Even in Australia, which employs a “skill and labour” standard closer to the traditional UK approach, courts have repeatedly insisted that an original work must demonstrably be “the product of human authorship.”⁵³

In Canada, the Supreme Court explicitly defined originality to reflect copyright’s balance between “the public interest in promoting the encouragement and dissemination of works of the arts and intellect and obtaining a just reward for the creator.”⁵⁴ It is with this purpose in mind, then, that we should consider the question at hand. In the case of truly machine-generated works, there is no “creator” to claim the reward of copyright. The AI programmer already obtains any copyright reward to which he or she can lay claim in the software program itself. The user of the AI can claim copyright in any outputs that sufficiently involve his or

⁴⁷ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 15. See Carys J. Craig, “Resisting ‘Sweat’ and Refusing Feist: Rethinking Originality After CCH” (2007) 40:1 U.B.C. L. Rev. 69.

⁴⁸ *Copyright, Designs and Patents Act 1988* (U.K.), c. 8, s. 9(3).

⁴⁹ *Copyright, Designs and Patents Act 1988* (U.K.), c. 8, s. 12(7), 79(2)(c), 81(2). It remains unclear what standard of originality (if any) is to determine the *subsistence* of copyright in such works, see Tim W. Dornis, “Artificial Creativity: Emergent Works and the Void in Current Copyright Doctrine” (2020) 22:1 Yale J. L. & Tech. 1 at 17–18.

⁵⁰ *Infopaq International A/S v. Danske Dagblades Forening*, C-5/08, [2012] E.C.R. I-6624 at I-6644.

⁵¹ *Eva-Maria Painer v. Standard Verlags GmbH and Others*, C-145/10, [2012] E.C.R. I-12594 at I-12622. See Daniel J. Gervais, “The Machine As Author” (2020) 105:5 Iowa L. Rev. 2053.

⁵² *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991); U.S. Copyright Office, *Compendium of U.S. Copyright Office Practices*, 3rd ed. (Washington, D.C.: United States Copyright Office, 2017) at § 306: “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”.

⁵³ *IceTV Pty. Ltd. v. Nine Network Australia Pty. Ltd.*, [2009] H.C.A. 14; *Telstra Corporation Ltd. v. Phone Directories Co. Pty Ltd.*, [2010] F.C.A. 44, aff’d [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.

⁵⁴ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 23.

her skill and judgment in the expression of ideas.⁵⁵ It remains to be asked whether the public interest in the encouragement and dissemination of works of the arts and intellect nonetheless demands the protection of autonomous AI-generated work. One might query whether such works are “works of the arts and intellect” at all. But even accepting that this is so—perhaps simply because they circulate in our culture and are received as such by their audience—one wonders whether their production and dissemination require encouragement in service of the public interest. Self-evidently, the machine requires no incentive to create.⁵⁶

If there is evidence to support an instrumentalist argument that further encouragement of AI-generated artifacts is indeed desirable—that there is, in the absence of copyright protection, a risk of underproduction—then policy-makers may wish to turn their minds to whether a *sui generis* system of protection or a new neighbouring right is warranted. Such a system would presumably offer a more appropriately tailored and limited monopoly over machine-produced works, with shorter terms, no moral rights, and clearly designated rightholders (specifically, the parties whose behaviour or investment the incentive is intended to encourage).⁵⁷ In the absence of any demonstrated need to specifically incentivize AI-generated works, however, and without any *author* being denied protection, there is no obvious or compelling reason to dislodge what ought to be a policy default: that the works belong in the public domain.

It should be emphasized that copyright is intended to function in service of a vibrant public domain.⁵⁸ To define AI-generated works as public domain is not to pronounce them worthless or without value to society, but to keep them free from the exclusive control granted by copyright—private control that imposes a social cost, and should therefore be the exception and not the rule. Unnecessarily extending the private preserve of copyright over such unauthored works would come “at the loss of society’s interest in maintaining a robust public domain that could help foster future creative innovation.”⁵⁹ With the rapid proliferation of 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.⁶⁰ By the same token, limiting copyright to original works of *authorship* can help “the public domain to flourish as others are able to produce new works by building on the ideas and information contained”⁶¹ in AI-generated works. As the Supreme Court reminds us:

The need to strike an appropriate balance between giving protection to the skill and judgment exercised by authors in the expression of their ideas, on the one hand, and leaving ideas and elements from the public domain free for all to draw upon, on the other, forms the background against which the arguments ... must be considered.⁶²

Against this background, it is suggested, arguments in favour of protecting AI outputs as copyrightable works are currently unconvincing.

⁵⁵ See Jane C. Ginsburg, Jane C. & Luke Ali Budiardjo, “Authors and Machines” (2019) 34:2 Berk. Tech. L.J. 343.

⁵⁶ Pamela Samuelson, “Allocating Ownership Rights in Computer-Generated Works” (1986) 47:4 U. Pitt. L. Rev. 1185 at 1199.

⁵⁷ See Pamela Samuelson, “Allocating Ownership Rights in Computer-Generated Works” (1986) 47:4 U. Pitt. L. Rev. 1185 at 1226.

⁵⁸ *Théberge v. Galerie d’Art du Petit Champlain inc.*, 2002 SCC 34 at para. 32; *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 23.

⁵⁹ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 23.

⁶⁰ Clark D. Asay, “Independent Creation in a World of AI” (2020) 14 Fla. Intl. U.L. Rev. 201 at 215.

⁶¹ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para 23.

⁶² *Cinar Corp. v. Robinson*, 2013 SCC 73 at para. 28.

2. AI and Copyright Infringement

The core principle of balance informs not only copyright's subsistence but also its scope. In the Supreme Court's words, "[t]he proper balance ... lies not only in recognizing the creator's rights but in giving due weight to their limited nature."⁶³ When establishing the extent of copyright owners' exclusive control, then, it is important to recall that "excessive control" can upset the copyright balance by "unduly limit[ing] the ability of the public domain to incorporate and embellish creative innovation in the long-term interests of society as a whole, or creat[ing] practical obstacles to proper utilization."⁶⁴ With the Court's cautionary note in mind, this section considers the matter of copyright infringement by AI systems, the risks of overprotecting AI inputs, and the vital role of copyright limits and exceptions in supporting the development and operation of AI technologies.

2.1. AI Inputs: Training Algorithms

Machine-learning processes require vast quantities of input data on which to "train." This is what enables an AI system to identify and dynamically learn from regularities or patterns and thereby make predictions and generate new outputs. A significant concern, then, is whether any copyright that subsists in the AI *inputs* is infringed in the training of AI. 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 rightholder 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.⁶⁵ 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 AI's technological development and operation. Copyright is neither designed for, nor suited to, this role.

The problem arises because copyright includes the sole right to reproduce the work in any material form⁶⁶—an exclusive right to control copies that has been capaciously defined, not least by Canada's Supreme Court, which recently held that it is infringed even by "broadcast incidental copies" of works—digital copies that never reach an audience but are made solely to facilitate transmissions.⁶⁷ There are compelling arguments to be made that the reproduction right should not be implicated by "non-expressive" or "non-consumptive" copies, including the kind of digital copies involved in machine learning. If copyright is ordinarily concerned with copying that "relates to human appreciation of the expressive qualities of that work," then it need not concern itself with "any act of reproduction that is not intended to enable human enjoyment, appreciation, or comprehension of the copied expression as expression."⁶⁸ In other words, copies made purely for machine-learning processes, text, or data mining are not "material" to the copyright scheme, and should be beyond the scope of the copyright owner's control. This conclusion would be "entirely consistent with the fundamental

⁶³ *Théberge v. Galerie d'Art du Petit Champlain inc.*, 2002 SCC 34 at para. 31.

⁶⁴ *Théberge v. Galerie d'Art du Petit Champlain inc.*, 2002 SCC 34 at para. 32.

⁶⁵ See Amanda Levendowski, "How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem" (2018) 93:2 Wash. L. Rev. 579. 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.

⁶⁶ *Copyright Act*, R.S.C. 1985, c. C-42, s. 3(1).

⁶⁷ *Canadian Broadcasting Corp. v. SODRAC 2003 Inc.*, 2015 SCC 57 at para. 55.

⁶⁸ Matthew Sag, "The New Legal Landscape for Text Mining and Machine Learning" (2019) 66 J. Copyr. Soc. U.S.A. 291 at 301.

structure of copyright law because, at its heart, copyright law is concerned with the communication of an author's original expression to the public."⁶⁹

Even where copies made for machine-learning processes *prima facie* implicate the reproduction right, however, there are limits and exceptions in the *Copyright Act* that could render many such copies lawful. Temporary reproductions made for technological processes do not infringe copyright, for example, but only if the reproduction is an essential part of the technological process, made only to facilitate a non-infringing use, and exists only for the duration of the technological process.⁷⁰ While some reproductions made in the research, training, and operation of AI systems will be transitory, many more will be more permanent, including the copies that construct the dataset mined by the AI. Additional reproductions and communications of the dataset may also be necessary in order for other researchers to use and test it for accuracy, replicability, and transparency.⁷¹

It is therefore necessary to look to the broader *fair dealing* defence in respect of these uses. Fair dealing in Canada is a *user right*, which, the Supreme Court has repeatedly explained, "must not be interpreted restrictively."⁷² It is understood to be available, however, only for uses made for statutorily enumerated purposes. These include research, private study, education, criticism, and review, and so could potentially embrace certain AI inputs made for such purposes, which are liberally construed.⁷³ From here follows a contextual examination of the fairness of particular uses, taking into account the purpose and character of the use, the nature of the work, the amount used, any alternatives available, and the effect of the use on the market for the work.⁷⁴ Again, many uses made for machine-learning purposes are likely to be "fair" when subjected to such analysis, 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.⁷⁵ But the need for such a context-specific assessment in relation to each work fed into the AI dataset in order to determine the lawfulness of its use is clearly unmanageable on the scale with which we are concerned.

It should also be noted, given the unpredictability of this analysis, that the risks in this regard are prohibitively high: statutory damages in Canada range from \$500 to \$20,000 per infringement for commercial purposes, and from \$100 to \$5000 for non-commercial infringements.⁷⁶ In Basanta's case, for example, Chamandy claimed \$20,000 in statutory damages for commercial use of a single protected work. Now imagine the potential impact of statutory damages for typical text and data mining activities that can involve the reproduction of millions of individual works in a training dataset.⁷⁷ The statute provides room for judicial discretion where the formula for damages would produce a result "grossly out of proportion with the

⁶⁹ Matthew Sag, "The New Legal Landscape for Text Mining and Machine Learning" (2019) 66 J. Copyr. Soc. U.S.A. 291 at 302.

⁷⁰ *Copyright Act*, R.S.C. 1985, c. C-42, s. 30.71.

⁷¹ See 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" (2020) at 4, online: *American University Washington College of Law Digital Commons* <https://digitalcommons.wcl.american.edu/research/48>.

⁷² *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 48.

⁷³ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 51.

⁷⁴ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 53.

⁷⁵ 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" (2020) at 4, online: <https://digitalcommons.wcl.american.edu/research/48>.

⁷⁶ *Copyright Act*, R.S.C. 1985, c. C-42, s. 38.1(1)(a).

⁷⁷ See Teresa Scassa, "Artist Sued in Canada for Copyright Infringement for AI-Related Art Project" (4 October 2018), online: *Teresa Scassa* https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=286.

infringement,”⁷⁸ but the staggering copyright liability risks produced by any such aggregate calculation are an obvious obstacle to AI research and development, making the importance of limiting copyright’s reach in this regard all the more apparent.

Internationally, there is currently an evolving patchwork of limits and exceptions that can accommodate text and data mining for machine learning, to a greater or lesser degree.⁷⁹ In Canada, a recent review of the *Copyright Act* recommended a legislative amendment “to facilitate the use of a work or other subject-matter for the purpose of informational analysis.”⁸⁰ There is clearly an appetite and a need for a broad and clear exception—one that should extend beyond research and non-commercial use—to prevent copyright from unnecessarily obstructing the development and compromising the capabilities of AI systems.

2.2. AI Outputs: Liability and for Whom/What?

As previously discussed, AI systems generate works that objectively resemble human creations. It follows that such works may, by design, be substantially similar to existing copyright-protected works (as demonstrated by Basanta’s art-factory “matches”), and so a final question to consider is whether such outputs can infringe copyright—and if so, who, if anyone, should be liable for the infringement.

The copyright doctrine of independent creation means that a work produced independently, without copying, will not infringe copyright in another’s pre-existing work even if it is identical. Where a substantial similarity between an AI output and a protected work is simply a matter of coincidence (as where Basanta’s scanners independently captured images without copying) the similarity is no basis for liability. However, if an AI is trained on a dataset that includes a particular protected work and subsequently produces a substantially similar output, then it is harder to chalk the similarity up to coincidence: there is access to the protected work and so the necessary causal connection. Is the protected input the *causa sine qua non* of the output, but for which the particular output would not have been made?⁸¹ The inscrutability of the algorithm’s operation may make it impossible to say. One could argue that, by virtue of the automated generative system, all of its outputs are independently created. Alternatively, it might be argued, the AI’s outputs are effectively derived from its inputs, and the combination of access plus substantial similarity is sufficient to establish *prima facie* infringement.⁸² But if that is the case, who should be liable?

Copyright infringement is a matter of strict liability and does not require knowledge; but it does require causation⁸³ and, in line with common law principle of tortious liability, some degree of responsibility for, or control over, the unlawful act.⁸⁴ In the US jurisprudence, this has been articulated as a requirement of some

⁷⁸ *Copyright Act*, R.S.C. 1985, c. C-42, s. 38.1(3)(a)–(b).

⁷⁹ See e.g. E.U., *Directive (EU) 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 Directives 96/9/EC and 2001/29/EC*, [2009] O.J., L. 130/92, arts. 3–4; *Copyright, Designs and Patents Act 1988* (U.K.), c. III, s. 29A.

⁸⁰ House of Commons, Standing Committee on Industry, Science and Technology, *Statutory Review of the Copyright Act* (June 2019) at 87–89, online: *House of Commons* www.ourcommons.ca/Content/Committee/421/INDU/Reports/RP10537003/indurp16/indurp16-e.pdf.

⁸¹ See *Francis, Day & Hunter Ltd. et al. v. Bron et al.*, [1963] 2 All E.R. 16 at 27, Diplock L.J., cited in *Gondos v. Hardy et al.*; *Gondos v. Toth et al.* (1982), 38 O.R. (2d) 555 at para. 32 (Ont. H.C.J.).

⁸² See Clark D. Asay, “Independent Creation in a World of AI” (2020) 14 Fla. Intl. U. L. Rev. 201 at 214-215.

⁸³ *Copyright Act*, R.S.C. 1985, c. C-42, s. 27(1).

⁸⁴ See e.g. *CBS Songs Ltd. v. Amstrad Consumer Electronics plc.*, [1988] 2 All E.R. 484.

“element of volition” with respect to the infringing conduct.⁸⁵ The volitional act requirement, while usually unproblematic in direct infringement cases, “comes right to the fore” when assessing the liability of someone “who does nothing more than operate an automated, user-controlled system.”⁸⁶

A volitional act requirement coheres with the intuition that individuals should be held responsible only for that which was under their control.⁸⁷ In the case of AI outputs, the allocation of liability should depend on who, if anyone, has the necessary degree of control over the production of an infringing copy. In some cases, this may be the AI programmer, in others it may be the AI user, or indeed there may be joint liability. Even if these individuals are not responsible for *doing* the infringing act, it could be argued that they “authorize” it—a cause of action that similarly amounts to primary infringement in Canada where the authorizer can be said to “sanction, approve [or] countenance”⁸⁸ the infringing act. Importantly, “a person does not authorize infringement by authorizing the mere use of equipment.”⁸⁹ Moreover, “[c]ourts should presume that a person who authorizes an activity does so only so far as it is in accordance with the law.”⁹⁰ On this basis, for example, the unpredictable production of an infringing output by the AI system may not implicate the programmer as the mere provider of the AI technology. However, the presumption is rebutted where a “certain relationship or degree of control”⁹¹ exists between the authorizer and the infringer. This could implicate the user of an AI system, for example, but only where the user has a sufficient degree of control over its outputs. One might query whether, normatively or ontologically, the authorization of a machine is equivalent to the authorization of a person under one’s control, but the notion of a machine as the tool or agent of a tortfeasor is by no means a stretch of the legal imagination.

Questions remain regarding liability for the truly autonomous production of copies by a machine over which no human actor has *de facto* control. In such a case, can the AI itself be said to infringe? The *Copyright Act* states that “it is an infringement of copyright for *any person* to do ... anything that ... only the owner of the copyright has the right to do.”⁹² It might seem intuitive that AI systems are incapable of exercising the kind of volition or autonomous agency necessary to attract liability. Taking a more functional approach, however, some have argued that machines should be regarded as possessing the requisite volition or agency, as a matter of law, to be capable of infringement.⁹³ Such a conclusion need not be at odds with the assertion that AIs lack the *expressive* agency necessary for authorship: a copier need not be an author, after all, and indeed the less a copier resembles an author, the more likely that their activity is infringing reproduction as opposed to lawful

⁸⁵ See *e.g. Religious Technology Center v. Netcom On-Line Communication Services, Inc.*, 907 F. Supp. 1361 (N.D. Cal. 1995) at 1370; *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*, 536 F. (3d) 121 (2d Cir. 2008) at 131. See generally Mala Chatterjee & Jeanne C. Fromer, “Minds, Machines, and the Law: The Case of Volition in Copyright Law” (2019) 119:7 Colum. L. Rev. 1887.

⁸⁶ *American Broadcasting Companies, Inc. v. Aereo, Inc.*, 134 S. Ct. 2498 (2014) at 2513–2514, Scalia J. dissenting.

⁸⁷ See Mala Chatterjee & Jeanne C. Fromer, “Minds, Machines, and the Law: The Case of Volition in Copyright Law” (2019) 119:7 Colum. L. Rev. 1887 at 1901.

⁸⁸ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 38, citing *Muzak Corp. v. Composers, Authors, and Publishers Association of Canada, Ltd.*, [1953] 2 S.C.R. 182 at 193.

⁸⁹ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 38, citing *Muzak Corp. v. Composers, Authors, and Publishers Association of Canada, Ltd.*, [1953] 2 S.C.R. 182 at 193.

⁹⁰ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 38, citing *Muzak Corp. v. Composers, Authors, and Publishers Association of Canada, Ltd.*, [1953] 2 S.C.R. 182 at 193.

⁹¹ *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13 at para. 38, citing *Muzak Corp. v. Composers, Authors, and Publishers Association of Canada, Ltd.*, [1953] 2 S.C.R. 182 at 193 and *De Tervagne v. Belœil (Town)*, [1993] 3 F.C. 227.

⁹² *Copyright Act*, R.S.C. 1985, c. C-42, s. 27(1) [emphasis added].

⁹³ Mala Chatterjee & Jeanne C. Fromer, “Minds, Machines, and the Law: The Case of Volition in Copyright Law” (2019) 119:7 Colum. L. Rev. 1887 at 1909–1911.

production or transformation.⁹⁴ But even so, who is to be sued? Applying a principal-agent analysis, the machine's volitional act *qua* "agent" could potentially produce liability for whichever party notionally fits the role of "principal"—most likely the person that creates or deploys the AI.⁹⁵

The matter may appear more academic than practical, as the exploitation of infringing copies produces its own form of secondary liability that could readily extend to those seeking to exploit infringing AI outputs. It can be secondary infringement for a person to sell, rent, or distribute copies of works and other subject matter where the person knows or should have known that the copies infringe copyright.⁹⁶ This could capture the sale or distribution of both infringing copies within a training dataset, for example, and, potentially, infringing copies produced by a generative AI—but, strictly speaking, the latter will be "infringing copies" only if an autonomous AI is deemed capable of infringing copyright in the first place.

2.3. AI and Automated Copyright Enforcement

A final issue at the intersection of AI and copyright is the increasing instrumentalization of AI and machine-learning technologies for the purpose of monitoring and controlling the use of copyright-protected works in digital environments. Anyone who uses the Internet knows that copyright norms are now encoded and enforced through automated copyright bots tasked with identifying protected content online and issuing takedown notices, blocking content, or monetizing it. Indeed, this is likely the most common context in which people now encounter copyright restrictions in their daily lives.

The problems associated with algorithmic copyright enforcement are many.⁹⁷ Structurally, the algorithms deployed for this task are frequently not fit for purpose. Studies have indicated that up to 30% of automated takedown requests are problematic in the sense that there are issues with the accuracy of the "matching" between the library of protected content and the new content.⁹⁸ Also, because algorithms are designed to perform a binary infringing/non-infringing analysis, they typically fail to recognize the complex layering of rights that subsist in respect of any particular content. Worse, they are ostensibly unable to take into account copyright limitations or exceptions that rely on discretion, context, or nuance (such as the substantial similarity analysis or the identification of lawful fair dealing). Such concerns are exacerbated by the absence of transparency and predictability of "black-box" decision-making, as well as a paucity of due process: the removal of content prior to any determination of its lawfulness, and the subsequent absence of a meaningful right of appeal or redress for errors, or penalty for unwarranted removal. From a Canadian perspective, it should be noted that the algorithms tend to encode US copyright doctrines (with the notable

⁹⁴ See *Cinar Corp. v. Robinson*, 2013 SCC 73 at paras. 39–40. See also discussion in Carys J. Craig, "Transforming 'Total Concept & Feel': Dialogic Creativity and Copyright's Substantial Similarity Doctrine" (2020) *Cardozo A. & Ent. L. J.* [forthcoming], online: SSRN <https://ssrn.com/abstract=3691280>.

⁹⁵ Mala Chatterjee & Jeanne C. Fromer, "Minds, Machines, and the Law: The Case of Volition in Copyright Law" (2019) 119:7 *Colum. L. Rev.* 1887 at 1913.

⁹⁶ *Copyright Act*, R.S.C. 1985, c. C-42, s. 27(2).

⁹⁷ For further discussion, see Carys Craig & Bob Tarantino, "'An Hundred Stories In Ten Days': COVID-19 Lessons for Culture, Learning, and Copyright Law" (2020) 57:3 *Osgoode Hall L.J.* 567.

⁹⁸ Jennifer M. Urban, Joe Karaganis & Brianna L. Schofield, "Notice and Takedown in Everyday Practice" (2017), online: SSRN <https://ssrn.com/abstract=2755628>.

exception of fair use),⁹⁹ disregarding jurisdictional differences and effectively rendering redundant domestic provisions such as Canada's non-commercial user-generated content exception.¹⁰⁰

Perhaps the most pernicious outcome of automated copyright enforcement is that it skews the copyright ecology itself: the inaccurate mobilization of algorithmic infringement analyses becomes the “social, legal and creative default,” and the choices of creators and audiences become “informed, manufactured, and ultimately distorted by the architecture of regulation.”¹⁰¹ The burden shifts to users and the public to challenge (when the process permits it) enforcement actions by copyright owners and platforms, and so to prove the lawfulness of their own uses. Automated copyright enforcement technologies therefore upset the balance between users' rights and owners' rights and so threaten to obstruct the purposes of Canada's copyright system.

3. Risks and Opportunities

It should by now be clear that the rapid evolution of AI technologies presents various risks and opportunities within the copyright field. One risk seems to weigh particularly heavily in our collective imagination: AI systems will soon displace human creators as the source of original cultural works. The sheer volume and rapidity with which AI can produce its outputs, and the increasingly elusive distinction between AI-generated and human-authored works, suggests that this risk is real. Moreover, the protection of these AI-generated works would exacerbate the risk that human creators are crowded out of the cultural landscape by encouraging the production and exploitation of such works, and by cluttering the field with “copyright landmines”—AI-generated works that human creators must avoid substantially replicating for fear of liability.

On the opportunity flip-side, however, we might note that a primary purpose of copyright is to encourage the creation of literary, dramatic, musical, and artistic works, and so the capacity to produce such works without the costly ongoing investment of human intellectual effort—and without the limits consequently imposed by copyright law—could produce a welcome windfall for the public domain.

Similarly, it should be noted that machine learning and text and data-mining technologies offer enormous promise when it comes to creative problem-solving and the generation of new knowledge. This is why it is particularly important that we recognize the risk that copyright's overprotection currently poses to the research, development, training, and transparency of AI technologies. Without adequate limits and exceptions to shield text and data mining from costly copyright liability, there is a risk that copyright law will stymie technological development and exacerbate the well-documented problems caused by incomplete, superficial, or biased training datasets.

At this moment in time, when the technology is developing and norms are unsettled, we have the opportunity to ensure that a balanced copyright system offers the necessary incentives and provides the vital space for the ongoing development of AI technologies—in furtherance, it is to be hoped, of the public interest. The greater risk is that, instead, copyright law simply will simply expand to capture the economic value of emerging AI technologies and their outputs, protecting corporate interests instead of authors, restricting user rights and impoverishing the public domain.

⁹⁹ See *Lenz v. Universal Music Corp.*, 801 F (3d) 1126 (9th Cir. 2015) (requiring consideration of fair use when filing a takedown notice).

¹⁰⁰ *Copyright Act*, R.S.C. 1985, c. C-42, s. 29.21. See Henning Grosse Ruse-Khan, “Automated Copyright Enforcement Online: From Blocking to Monetization of User-Generated Content” (2020), online: *SSRN* <https://ssrn.com/abstract=3565071>.

¹⁰¹ Dan L. Burk, “Algorithmic Fair Use” (2019) 86:2 *U. Chicago L. Rev.* 283 at 303–306.

4. Key Gaps in the Law

The discussion above has revealed several key gaps in copyright law regarding AI that remain to be filled. In respect of copyright subsistence, the issue that has attracted the most attention is the apparent gap in copyright protection when it comes to AI-generated works over which no human can claim authorial control. Currently, in Canada, such unauthored outputs likely fail to satisfy copyright's originality threshold and are therefore not protected by copyright. Whether the public domain status of such works is a "gap" in the law or, rather, an appropriate limit established by the law remains a matter of debate. If it is thought that copyright should extend to AI-generated works, another gap quickly reveals itself: it is unclear who, if anyone, should be deemed to be the author, and in whom first ownership of copyright should vest. If new protections are to be afforded to AI-generated works, it further remains to be determined, as a matter of evidence-based policy-making, what the appropriate scope and duration of such rights would be.

Gaps are also evident in the law regarding AI and copyright infringement. Most pressing is a gap in the existing copyright limits and user rights, which may fail to provide adequate protection for the text and data-mining processes that are vital to training AI systems. Until this gap is addressed by the enactment of a broad and clear exception specifically for informational analysis, and/or an expansion of fair dealing, copyright law threatens to chill the research and development, transparent operation, and commercialization of AI.

There is also a gap in the law regarding potential copyright liability for apparently infringing acts carried out by AI systems and beyond the control of the AI developer or user. It is currently unclear what degree of involvement or control by the programmers, providers or users of AI could produce direct or indirect liability in such cases. It also remains to be determined whether autonomous AI systems can have sufficient agency to infringe copyright law, and if so, where the resulting liability should lie.

Finally, with respect to AI and automated copyright enforcement, a critical gap exists between Canada's carefully calibrated copyright law and the increasingly pervasive architecture of algorithmic control.

Such gaps in the law cannot be satisfactorily filled, however, until we tackle a larger gap in our normative and conceptual thinking about the appropriate interaction of AI and copyright law: filling this normative gap requires much more careful consideration of the extent to which the copyright system can and should play a role in encouraging, facilitating, restricting, and/or regulating the ongoing evolution of AI. A robust, substantive principle of technological neutrality¹⁰² should guide any efforts to "update" copyright law in response to AI technologies: changes to the allocation of rights and responsibilities within the copyright system must be made with a view to maintaining the appropriate balance between protection and access, in furtherance of copyright's public policy goals and the social values it seeks to foster.

Further Reading

Asay, Clark D., "Independent Creation in a World of AI" (2020) 14 Fla. Intl. U. L. Rev. 201.

Bridy, Annemarie, "The Evolution of Authorship: Work Made by Code" (2016) 39:3 Colum. J. L. & Arts 395.

¹⁰² *Entertainment Software Association v. Society of Composers, Authors and Music Publishers of Canada*, 2012 SCC 34 at para. 8. See generally, Carys J. Craig, "Technological Neutrality: Recalibrating Copyright in the Information Age" (2016) 17 Theor. Inq. L. 601.

Burk, Dan, “Algorithmic Fair Use” (2019) 86:2 U. Chicago L. Rev. 283.

Chatterjee, Mala & Jeanne C. Fromer, “Minds, Machines and the Law: The Case of Volition in Copyright Law” (2019) 119:7 Colum. L. Rev. 1887.

Craig, Carys & Ian Kerr, “The Death of the AI Author” (2020) 52:1 Ottawa L. Rev. 31.

Elkin-Koren, Niva, “Fair Use by Design” (2017) 64:5 U.C.L.A. L. Rev. 1082.

Flynn, Sean, 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” (2020), online: <https://digitalcommons.wcl.american.edu/research/48>.

Gervais, Daniel, “The Machine As Author” (2020) 105:5 Iowa L. Rev. 2053.

Levendowski, Amanda, “How Copyright Law Can Fix Artificial Intelligence’s Implicit Bias Problem” (2018) 93:2 Wash. L. Rev. 579.

Samuelson, Pamela, “Allocating Ownership Rights in Computer-Generated Works” (1986) 47:4 U. Pitt. L. Rev. 1185.

Acknowledgements

The author thanks Felice Yeung for her excellent research assistance. The author is also forever grateful for many inspiring and thought-provoking conversations about AI and authorship with the wonderfully insightful—and sorely missed—Ian R. Kerr.