AI Derivatives: The Application to the Derivative Work Right to Literary and Artistic Productions of AI Machines

Daniel J. Gervais, PhD*

This Article predicts that there will be attempts to use courts to try to broaden the derivative work right in litigation either to prevent the use of, or claim protection for, literary and artistic productions made by Artificial Intelligence (AI) machines. This Article considers the normative valence of, and the (significant) doctrinal pitfalls associated with, such attempts. It also considers a possible legislative alternative, namely attempts to introduce a new sui generis right in AI productions. Finally, this Article explains how, whether such attempts succeed or not, the debate on rights (if any) in productions made by AI machines is distinct from the debate on text and data mining exceptions.

I. INTRODUCTION

Artificial Intelligence (AI) machines are increasingly used to assist authors in creating copyrightable works, but they are also at the point where they can "create" literary and artistic productions autonomously, in the sense that the "cause" of the work is not human.[1] To do so, AI machines use machine-learning algorithms and process data corpora often consisting of existing copyrighted works.[2]

Enter the derivative work right contained in 17 U.S.C. § 106(2) (hereinafter "The Copyright Act"). The Copyright Act provides an exclusive right "to prepare derivative works based upon the copyrighted work"[3] and defines "derivative work" in part as any work "*based upon one or more preexisting works.*"[4] This definition of the right could loosely be used as a *definition* of machine-learning when applied to the creation of literary and artistic productions because AI machines can produce literary and artistic content (output) that is *almost necessarily "based upon"* a dataset consisting of preexisting works.[5] For example, an AI machine can be "fed" existing works composed by J.S. Bach and produce a new musical composition "in the style of Bach."[6] Or

---

[1] I discuss elsewhere how "cause" (a version of proximate, not simple, cause) is an appropriate notion to determine whether a literary or artistic work has a human author. *See generally* Daniel Gervais, *The Machine as Author*, 105 IOWA L. REV. 2053 (2020); Daniel Gervais, *The Human Cause,* RESEARCH HANDBOOK ON INTELLECTUAL PROPERTY AND ARTIFICIAL INTELLIGENCE (R. Abbott, ed., forthcoming 2022). I give multiple examples of those types of productions in both sources. This Article uses "machine" as a generic term that may apply to a computer using AI software but could also cover machines capable of movement, such as a robot painting on canvas.

[2] Machine learning is the dominant form of AI. *See* Roberto Iriondo, *Machine Learning (ML) vs. Artificial Intelligence (AI) – Crucial Differences,* TOWARDS AI (Oct. 15, 2018), https://medium.com/datadriveninvestor/differences-between-ai-and-machine -learning-and-why-it-matters-1255b182fc6 ("'Machine learning [ML] is the study of computer algorithms that allow computer programs to automatically improve through experience.' — ML is one of the ways we expect to achieve AI. Machine learning relies on working with large datasets by examining and comparing the data to find common patterns and explore nuances[,]" quoting Prof. Tom M. Mitchell, former Chair of the Machine Learning Department at Carnegie Mellon University).

[3] 17 U.S.C. § 106(2).

[4] 17 U.S.C. § 101 (emphasis added).

[5] A well-known example is the machine named e-David, which produces paintings using a complex visual optimization algorithm that "takes pictures with its camera and draws original paintings from these photographs." Shlomit Yanisky-Ravid, *Generating Rembrandt: Artificial Intelligence, Copyright, and Accountability in the 3A Era—The Human-Like Authors Are Already Here—A New Model,* 2017 MICH. STATE L. REV. 659, 662 (2017); *see also* Section III.A *infra.*

[6] *See* Gaëtan Hadjeres & François Pachet, *DeepBach: A Steerable Model for Bach Chorales Generation*, ARXIV 1 (Dec. 3, 2016), https://arxiv.org/pdf/1612.01010v1.pdf

it can scan works by Rembrandt and produce a new painting in the style of the Dutch master.[7] One can readily see how a broad interpretation of the derivative work right resulting from a wide-ranging reading of the statutory definition could provide a very far-reaching tool to copyright owners.

A quick terminological point should be made before proceeding further. This Article uses the term "production" to refer to the perceptible output of an AI machine.[8] If this input belongs in the literary and artistic category, then it is a literary and artistic production. This terminological distinction is necessary because for that production to be a "work" (of authorship) protected by the Copyright Act, two conditions must be met. First, that production must be original and, second, it must be fixed.[9] Originality is not defined in the statute but the Supreme Court defined it as meaning that the work must be independently created by its author and must embody expression that is at least minimally creative.[10] The test is satisfied when the work is the result of creative choices made by the author.[11] I have argued elsewhere that originality implies that the work must be the result of *human* creative choices.[12] To be as thorough as possible, this Article will treat this question as unsettled in the caselaw, though it is worth noting that the United States

---

(last visited Feb. 15, 2022); *see also* William T. Ralston, *Copyright in Computer-Composed Music: HAL Meets Handel*, 52 J. Copyright Soc'y U.S.A. 281, 283–91 (2004).

[7] *See* Yanisky-Ravid, *supra* note 5, at 663.

[8] 17 U.S.C. § 101 refers to objects protected by the Copyright Act as "original works of authorship." In adopting "production" as its terminological stance, this Article borrows from language used in Article 2 of the most important international copyright treaty, the Berne Convention for the Protection of Literary and Artistic Works of, Sept. 9, 1886, as revised at Paris, July 24, 1971, 1161 U.N.T.S. 30 [hereinafter Berne Convention]. The Berne Convention had 180 member States as of January 2022. The United States became a party to the Convention on Mar. 1, 1989. *See Contracting Parties: Berne Convention*, World Intell. Prop. Org., https://www.wipo.int/treaties/en/ShowResults.jsp?lang=en&treaty_id=15 (last visited Dec. 12, 2021). The term "output" can also be used generically and would include literary and artistic productions made by a machine.

[9] *See* Restatement of the Law, Copyright, § 1(a) (Am. Law Inst., Tentative Draft No. 2, 2021).

[10] *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 346 (1991) (citing Trade-Mark Cases, 100 U.S. 82, 94 (1879); Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53 (1884); *see also* Restatement of the Law, Copyright, *supra* note 9, at §§ 5–7.

[11] *See* Restatement of the Law, Copyright, *supra* note 9, at § 7(b).

[12] *See* Daniel Gervais, *The Machine as Author*, 105 Iowa L. Rev. 2053, 2098–99 (2020); *see also* Daniel Gervais, *Self-Driving Culture*, Kluwer Copyright Blog (Nov. 25, 2021), http://copyrightblog.kluweriplaw.com/2021/11/25/self-driving-culture/.

Copyright Office has taken the same position defended by this Author, as have a number of senior scholars.[13]

This Article proceeds as follows. In Part II, it briefly explains how AI technology works and how AI machines produce literary and artistic outputs. In Part III, it recalls the ins and outs of the derivative work right and then its potential new role in the AI context. In Part IV, to situate and properly cabin its prescriptive findings, it considers possible new legislation to introduce a sui generis right in AI productions and assesses the impact of rights in AI productions, if any, on the ongoing discussions about text and data mining exceptions.

II. THE AI "CREATION" PROCESS & ITS COPYRIGHT RAMIFICATIONS

The purpose of this Part is to offer a brief glimpse into AI technology only to the extent necessary to discuss the derivative work right in detail in the next Part. At the most basic level, AI is, first and foremost, computer code running on one or more computers, often but not necessarily connected to the Internet, that can learn from a dataset—sometimes a huge dataset—hence the label "Big Data."[14] This process is known as machine-learning, which today is "the dominant AI technology."[15] Machine-learning can be *supervised* (by humans), or not.

---

[13] *See* U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 306 (3d ed. 2017); Pamela Samuelson, *Allocating Ownership Rights in Computer-Generated Works*, 47 PITT. L. REV. 1186, 1208 (noting that protecting machine productions by copyright would "over-reward[] the programmer, particularly in light of the fact that the programmer is no more able to anticipate the output than anyone else"); Jane C. Ginsburg, *The Concept of Authorship in Comparative Copyright Law*, 52 DEPAUL L. REV 1063, 1066 (2003) (arguing that an author is the "*human being* who exercises subjective judgment in composing the work and who controls its execution") (emphasis added); Sam Ricketson*, People or Machines: The Berne Convention and the Changing Concept of Authorship*, 16 COLUM. J.L. & ARTS 1, 8 (1991) (noting that at the outset of the Berne Convention, "[d]espite the omission of a definition of 'author,' . . . there was nonetheless a basic agreement between the contracting states as to the meaning of the term, and, because of this, it was thought unnecessary to define it"). *See generally* Jane C. Ginsburg & Luke Ali Budiardjo, *Authors and Machines*, 34 BERK. TECH. L.J. 343 (2019) (discussing the way in which AI can produce literary and artistic content and explaining why machine productions are almost always at least in some way the product of human authorship). Naturally, the Internet as "global meme factory" probably makes it easier than ever to produce derivative content. *See* Daniel Gervais, *Authors, Online*, 38 COLUM. J.L. & ARTS 385, 385 (2014).

[14] Big Data can be defined according to "three essential features, a fourth that, though not essential, is increasingly typical, and a fifth that is derived from the other three (or four). Those features are volume, veracity, velocity, variety, and value." Daniel Gervais, *TRIPS Meets Big Data, in* BIG DATA AND GLOBAL TRADE LAW 160–61 (Mira Burri ed., 2021).

[15] U.K. INFORMATION COMMISSIONER'S OFFICE & THE ALAN TURING INSTITUTE, EXPLAINING DECISIONS MADE WITH AI 7 (2020), https://iapp.org/media/pdf/resource_center/ico_

"*Unsupervised*" in this context means that the system is "trained on a dataset without explicit instructions or labelled data."[16] Situated between supervised and unsupervised learning, *reinforcement learning* is a third mode of machine-learning, in which humans verify what the machine learned on its own and hopefully correct mistakes, often using sampling techniques.[17]

Machine-learning in all three modes is used both "to discern and operationalise patterns in data."[18] It uses a set of "computational methods using experience to improve [its] performance or to make accurate predictions."[19] Using machine-learning, an AI system can "automatically generate heuristics" and make autonomous determinations of various kinds.[20] It can adjust its "behavior to enhance [its] performance on some task through experience."[21] A machine can, for example, be shown pictures of cats and dogs and then learn the features of each so that it can recognize cats and dogs it has never "seen" before.[22] The quality of the learning process is obviously dependent on the quality of the training data, as some well-documented disastrous examples have brought to light.[23]

The machine-learning function can take the form of "deep learning," a subset of machine-learning using a layered structure of

---

explaining_decisions_with_ai.pdf [hereinafter EXPLAINING DECISIONS MADE WITH AI]; *see also* Iriondo, *supra* note 2.

[16] *See* EXPLAINING DECISIONS MADE WITH AI, *supra* note 15, at 7.

[17] *See Leslie* Pack *Kaelbling*, Michael L. Littman & Andrew W. Moore, *Reinforcement Learning: A Survey*, 4 J. A.I. RSCH. 237, 237 (1996).

[18] Michael Veale, Governing Machine Learning that Matters 33 (2019) (Ph.D. dissertation, University College London), https://discovery.ucl.ac.uk/id/eprint/10078626/1/thesis_final_corrected_mveale.pdf.

[19] MEHRYAR MOHRI, AFSHIN ROSTAMIZADEH & AMEET TALWALKAR, FOUNDATIONS OF MACHINE LEARNING 1 (MIT Press 2d ed. 2018).

[20] WOLFGANG ERTEL, INTRODUCTION TO ARTIFICIAL INTELLIGENCE 102 (Ian Mackie ed., Nathaniel Black trans., Springer-Verlag London Ltd. 1st ed. 2011). AI programmers use several different algorithmic techniques, depending (usually) on the task at hand. For a detailed overview, see EXPLAINING DECISIONS MADE WITH AI, *supra* note 15, at 115–19.

[21] Harry Surden, *Machine Learning and the Law*, 89 WASH. L. REV. 87, 89 (2014).

[22] *See* Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 WASH. L. REV. 579, 592 (2018).

[23] For example, when Google's AI created a link between images of Black people and gorillas. *See* James Vincent, *Google "Fixed" Its Racist Algorithm By Removing Gorillas From Its Image-Labeling Tech*, THE VERGE (Jan. 12, 2018, 10:35 AM), https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai. Another example is when a new Microsoft AI chatbot quickly turned racist by "learning" on social media. *See* James Vincent, *Twitter Taught Microsoft's AI Chatbot to Be a Racist Asshole in Less Than a Day*, THE VERGE (Mar. 24, 2016, 6:43 AM), https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist.

algorithms allowing the machine to learn and make predictions and *decisions on its own*.[24] Deep learning has been called "the true challenge to artificial intelligence," namely "solving the tasks that are easy for people to perform but hard for people to describe formally—problems that we solve intuitively, that feel automatic, like recognizing spoken words or faces in images."[25] With deep learning, one could say— acknowledging that metaphors are intellectual shortcuts—that the computer has its own, autonomous brain.[26] Importantly, deep learning is *automated* and often (if not almost always) removed from direct human input or control.[27]

There are various ways to make AI systems learn and produce better. One of them is the development of General Adversarial Networks ("GANs"), a technological path likely to grow the affordances of AI systems both qualitatively and quantitatively.[28] GANs' potential is huge, because they can learn to mimic any distribution of data. That is, "GANs can be taught to create worlds eerily similar to our own in any domain: images, music, speech, prose," sometimes using smaller datasets.[29]

---

[24] *See* Robert D. Hof, *Deep Learning: With Massive Amounts of Computational Power, Machines Can Now Recognize Objects and Translate Speech in Real Time. Artificial Intelligence is Finally Getting Smar*t, MIT Tᴇᴄʜ. Rᴇᴠ. (Apr. 23, 2013), https:// www.technologyreview.com/technology/deep-learning/.

[25] Iᴀɴ Gᴏᴏᴅғᴇʟʟᴏᴡ, Yᴏsʜᴜᴀ Bᴇɴɢɪᴏ & Aᴀʀᴏɴ Cᴏᴜʀᴠɪʟʟᴇ, Dᴇᴇᴘ Lᴇᴀʀɴɪɴɢ 1 (MIT Press 2016).

[26] *See* Patrick Grieve, *Deep Learning vs. Machine Learning: What's the Difference?*, Zᴇɴᴅᴇsᴋ (Jan. 23, 2020), https://www.zendesk.com/blog/machine-learning-and-deep-learning/; *see also* Claudio Masolo, *Supervised, Unsupervised and Deep Learning*, Tᴏᴡᴀʀᴅs Dᴀᴛᴀ Sᴄɪ. (May 7, 2017), https://bit.ly/2BydnE8.

[27] This has now gone mainstream. *See* William Vorhies, *Automated Deep Learning – So Simple Anyone Can Do It*, Dᴀᴛᴀ Sᴄɪ. Cᴇɴᴛ. (Oct. 4, 2018), https://www.datascience central.com/profiles/blogs/automated-deep-learning-so-simple-anyone-can-do-it.

[28] Yann LeCun, Facebook's AI Research Director and a professor at NYU, described GANs as "the most interesting idea in the last 10 years in [machine learning]." Yann LeCun, *What Are Some Recent and Potentially Upcoming Breakthroughs in Deep Learning*?, Qᴜᴏʀᴀ (July 28, 2016), https://www.quora.com/What-are-some-recent-and-potentially-upcoming-breakthroughs-in-deep-learning. GANs are "adversarial" because two machines work one against the other, creating a constant feedback loop that increases the quality of outputs. *See* AI Wɪᴋɪ, A Bᴇɢɪɴɴᴇʀ's Gᴜɪᴅᴇ ᴛᴏ Gᴇɴᴇʀᴀᴛɪᴠᴇ Aᴅᴠᴇʀsᴀʀɪᴀʟ Nᴇᴛᴡᴏʀᴋs (GANs), https://wiki.pathmind.com/generative-adversarial-network-gan [hereinafter Bᴇɢɪɴɴᴇʀ's Gᴜɪᴅᴇ] (last visited Feb 15. 2022).

[29] Bᴇɢɪɴɴᴇʀ's Gᴜɪᴅᴇ, *supra* note 28. More specifically, GANs use an actor-critic model, as one machine, "called the generator, generates new data instances, while the other, the discriminator, evaluates them for authenticity; i.e., the discriminator decides whether each instance of data it reviews belongs to the actual training dataset or not.". *Id.*; *see also* Jᴀᴋᴜʙ Lᴀɴɢʀ & Vʟᴀᴅɪᴍɪʀ Bᴏᴋ, GANs ɪɴ Aᴄᴛɪᴏɴ: Dᴇᴇᴘ ʟᴇᴀʀɴɪɴɢ ᴡɪᴛʜ Gᴇɴᴇʀᴀᴛɪᴠᴇ Aᴅᴠᴇʀsᴀʀɪᴀʟ Nᴇᴛᴡᴏʀᴋs 3 (Manning Publications Co., 2019).

GANs have been used to create literary and artistic content. Indeed, a painting produced by a GAN was sold at auction in 2018 for $432,500.[30]

What the technology described above boils down to is that a large amount of the inputs needed for an AI machine to produce *contemporary* literary and artistic content is inevitably preexisting copyrighted works that are still protected, unlike the Bach and Rembrandt examples mentioned above for which the term of protection (assuming it even existed in some form at the time) has long expired.[31] In deep learning scenarios, a large dataset is typically used. In the case of GANs, a smaller set of preexisting works may well be sufficient.[32]

III. A NEW ROLE FOR THE DERIVATIVE WORK RIGHT?

A. *Overview of the Derivative Work Right*

If there is one definition in the United States Copyright Act that raises eyebrows and uses language that cannot *literally* mean what it says, it is the definition of "derivative work." That definition in relevant part reads as follows: "a work *based upon one or more preexisting works*, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted."[33] What this Article means when it says that the definition cannot literally mean what it says is simply that human creations are often, if not almost always, based upon some other work that the author has read, seen, consulted, experienced, or been influenced by in some other way.[34] There is a so-called general osmosis of creativity through both space and time that

---

[30] James Vincent, *Christie's Sells Its First AI Portrait for $432,500, Beating Estimates of $10,000,* THE VERGE (Oct. 25, 2018, 1:03 PM), https://www.theverge.com/2018/10/25/18023266/ai-art-portrait-christies-obvious-sold.

[31] The current term of protection for most works under U.S. law is the life of the author plus seventy years. 17 U.S.C. § 302(a).

[32] Hence the frequent use of the term Big Data. For an explanation and definition, see Daniel J. Gervais, *TRIPS Meets Big Data*, *supra* note 14.

[33] 17 U.S.C. § 101 (emphasis added).

[34] For example, it is well-known that to learn creative writing or art humans learn from existing masterpieces and other works. *See* Daniel Gervais, *The Derivative Right, or Why Copyright Protects Foxes Better than Hedgehogs,* 15 VAND. J. ENT. & TECH. L. 785, 851 (2013) [hereinafter Gervais, *Foxes*] ("By copying a master's work, the 'pupil' might at least get a glimpse of the great author's mind, which would seem like a normatively desirable process. '*L'art naît d'un regard sur l'art*,' as the French would say: art is born from a view on existing art.").

the famous saying often attributed to Isaac Newton that we "stand on the shoulder of giants" captures in part.[35]

Given the "or in any other form" language in the definition of "derivative work," which signals that the definition is not exhaustive, an attempt to read the broad language of the first part of the statutory definition (the "based upon" clause) can be restrained by the enumeration that follows in application of the *ejusdem generis* rule.[36] One can argue that the list captures the major forms of derivation that come under the derivative work umbrella and that the opening clause may then just capture what I have elsewhere labelled "penumbral derivatives."[37]   Other arguments to limit the reach of the right exist. Professor Paul Goldstein for example, has argued that, considering the enumeration, the statutory text is intended primarily to protect certain licensing markets, though he acknowledges that even following that approach sometimes courts have incorrectly overextended the right's scope.[38]

The risk of an overextension by interpretation is real.[39]   As this Article adumbrated in the Introduction and explains further in the next Part, with AI the debate takes on a potentially more complex hue.[40]   As we apply the derivative work right to AI productions, we face obstacles stemming from controversial and uncontroversial aspects of the derivative work right. Let us begin with the latter.

---

[35]   *See generally* ROBERT K. MERTON, ON THE SHOULDERS OF GIANTS 8–12 (1993). Professor Bridy, for example, has argued along those lines "all cultural production is inherently derivative." Annemarie Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author*, 2012 STAN. TECH. L. REV. 5, 12 (2012).

[36]   On the *ejusdem generis* rule, see *Garcia v. United States*, 469 U.S. 70, 74 (1984): When general terms follow an enumeration of persons or things, such general words should not be construed in their widest extent, but should apply only to persons or things of the same general kind or class as those specifically mentioned. In 17 U.S.C. § 101, of course, the general words of the "based upon" clause precede instead of follow, but the canon could still be invoked. The canon, however, "cannot be used to 'obscure and defeat the intent and purpose of Congress' or 'render general words meaningless.'" United States v. Kaluza, 780 F.3d 647, 661 (5th Cir. 2015).

[37]   *See* Gervais, *Foxes*, *supra* note 34, at 808.

[38]   PAUL GOLDSTEIN, GOLDSTEIN ON COPYRIGHT § 7.3 (3d ed. 2012); Paul Goldstein, *Derivative Rights and Derivative Works in Copyright*, 30 J. COPYRIGHT SOC'Y U.S.A. 209, 221 (1983) (noting that "[i]t is no coincidence that the principal cases establishing broad rights against infringement by derivative works characteristically involve situations in which the alleged infringer had at some earlier point sought a license.").

[39]   I am not the first scholar to point to the interpretive risks associated with the "based upon" clause. *See, e.g.*, Naomi Abe Voegtli, *Rethinking Derivative Rights*, 63 BROOK. L. REV. 1213, 1263–67 (1997) (proposing a narrow formulation of the derivative right, in addition to broader fair use and a possible compulsory license).

[40]   *See infra* Section III.A.

B. *(Relatively) Uncontroversial Aspects of the Derivative Work Right*

Some aspects of the derivative work right, in this Article's submission, are mostly uncontroversial. The first is that a derivative work for which one claims protection under federal law must be a *work*. This requirement implies that a derivative work must itself be original to be eligible for copyright protection.[41] Second, unless one is prepared to violate the basic canon of statutory construction that guards against superfluity, the derivative work right cannot have the exact same scope as the right of reproduction.[42] This logically implies that, although the Venn diagram of the two rights would show considerable overlap, some derivative works are not reproductions. Third, because the definition clearly refers to preexisting works, what is derived from must be one or more "works," i.e., works protected by copyright.[43] Fourth, the notion of originality applied to the protection of derivative works requires that the person claiming to have authored a derivative work must have added or transformed one or more preexisting works in some way, shape, or form.[44]

By combining the four above points, one can posit fairly safely that the derivative work right, properly applied and understood, is situated in a zone between (and occasionally "beyond"[45]) reproduction, on the one hand, and uses that are inspired by, but not infringing (because they

---

[41] *See* RESTATEMENT OF THE LAW, COPYRIGHT, *supra* note 9.

[42] *See* GEORGE COSTELLO, CONG. RSCH. SERV., STATUTORY INTERPRETATION: GENERAL PRINCIPLES AND RECENT TRENDS 12 (2006) ("[S]tatutes should be construed 'so as to avoid rendering superfluous' any statutory language.").

[43] *See* 17 U.S.C. §§ 101, 106, *supra* note 4 and accompanying text. Arguably, this also includes works that were protected but whose term of protection expired, although in such a case the derivative work right would be irrelevant because it could no longer be enforced.

[44] The Copyright Restatement draft refers to this as "original expression contributed by the author of the derivative work." *See* RESTATEMENT OF THE LAW, COPYRIGHT, *supra* note 9, § 3(c).

[45] This term is used in that very context in *Lone Ranger Television, Inc. v. Program Radio Corp.*, 740 F.2d 718, 722 (9th Cir. 1984) ("[T]he protection of derivative rights extends *beyond mere protection against unauthorized copying* to include the right to 'make other versions of, perform, or exhibit the work.'") (emphasis added) (quoting Russell v. Price, 612 F.2d 1123, 1128 n.16 (9th Cir. 1979). Professor Ochoa has argued in a similar vein that the "reason the right to prepare derivative works was drafted without any reference to fixation was to make sure that public performances of derivative works would be covered." *See* Tyler T. Ochoa, *Copyright, Derivative Works and Fixation: Is Galoob a Mirage, or Does the Form(GEN) of the Alleged Derivative Work Matter?*, 20 SANTA CLARA HIGH TECH. L.J. 991, 1020 (2004).

are not "based upon"), an earlier work, on the other hand.[46] A final point that belongs in this Section on less controversial aspects is that a derivative work is not protected if it is infringing, as the statute itself makes clear.[47]

Beyond the points made in the previous paragraph, the analysis quickly enters more troubled waters.

C. *The Originality Controversy*

As the previous Section just explained, *to be protected* as a derivative work, a literary or artistic production must meet the originality condition applicable to other works of authorship. Does that mean that, *to infringe* the derivative work right (belonging to a third party), the derivative work must also be original? This is a controversial area of law. Recall that, to be protected as a work, a work must normally be both fixed and original.[48]

Professor Goldstein opined that the derivative work right may be infringed even if the derivative production would *not* qualify for protection as a work.[49] The Ninth Circuit agrees, but as we move forward, one must ask whether this means that the derivative work may be infringed by a production that need not be original, need not be fixed, or both.

In two cases with similar fact patterns, a defendant took images from artwork prints or books and placed them on tiles that he then offered for sale.[50] The defendants' actions unquestionably did not

---

[46] This Article agrees with the characterization suggested in Timothy Everett Nielander, *The Mighty Morphin Ninja Mallard: The Standard for Analysis of Derivative Work Infringement in the Digital Age*, 4 TEX. WESLEYAN L. REV. 1, 2 (1997), as follows:

> On the continuum between an exact reproduction of protected property, and the creation of an original work, lies a gray zone. This zone is a mixture of protected works—printed art, art on digital media, digital and analog music, and other works recognized as deserving intellectual property protection—that can be mixed and matched with other works to create new works. American law recognizes protection of this form of copying as *derivative rights*.

[47] 17 U.S.C. § 103(a) ("[P]rotection for a work employing preexisting material in which copyright subsists does not extend to any part of the work in which such material has been used unlawfully."). *See, e.g.*, Gracen v. Bradford Exch., 698 F.2d 300, 302 (7th Cir. 1983).

[48] *See supra* notes 9–10 and accompanying text.

[49] Paul Goldstein, *Derivative Rights and Derivative Works in Copyright*, *supra* note 38, at 231 n.75 (1983) ("[T]he Act does not require that the derivative work be protectable for its preparation to infringe.").

[50] Mirage Editions, Inc. v. Albuquerque A.R.T. Co., 856 F.2d 1341, 1342 (9th Cir. 1988); Munoz v. Albuquerque A.R.T. Co., 38 F.3d 1218 (9th Cir. 1994).

involve copying (reproduction) of a protected work. These cases are interesting because precedents in which the derivative work right is infringed generally also involve a reproduction. In contrast, these cases clearly delineate an area for the derivative work right that the reproduction right does not cover.[51] In *Mirage Editions*, referring to the statutory language that a derivative work must "recast, transform[] or adapt[]" one or more preexisting works in some form, the Ninth Circuit found that the tiles amounted to derivative works.[52] In drawing a distinction between this type of mounting on tiles and traditional art reframing (which the court acknowledged was a non-infringing activity), the court found that the use of a resin to glue the pictures on the tiles made a difference.[53] In that case, fixation was obviously not an issue and originality was not discussed.

In another case, the Ninth Circuit found that it made "no difference that the derivation may not satisfy certain requirements for statutory copyright registration itself."[54] The works at issue were both voice recordings of scripts from the "adventures" of the Lone Ranger and the duplication, remixing, and distribution of those recordings "just as if he had hired the actors, sound effects crew, and producers originally used for the tapes to do a second interpretation of the scripts for an audience."[55]

In another case, the plaintiff argued that the derivative work right contained in § 106(2) was "intended to expand the definition of derivative works to include any work *based on* a copyrighted work."[56] Strangely perhaps, the court states that the plaintiff cited "no authority to support this novel proposition."[57] There was, however, one authority in plain view: the text of the statute adopted in 1976.[58] This seems to suggest that the "based upon" clause is not likely to create a huge new zone of exclusivity for copyright owners compared to the previous (1909) Act.[59]

---

[51] *Mirage Editions*, 856 F.2d at 1343 ("The protection of derivative rights extends beyond mere protection against unauthorized copying.")

[52] *Mirage Editions*, 856 F.2d at 1343–44.

[53] *Id.*

[54] Lone Ranger Television, Inc. v. Program Radio Corp., 740 F.2d 718, 722 (9th Cir. 1984). The current (1976) Copyright Act entered into force on January 1, 1978. The facts of the case (*Lone Ranger*) predate this entry into force.

[55] *Lone Ranger*, 740 F.2d at 722.

[56] Litchfield v. Spielberg, 736 F.2d 1352, 1357 (9th Cir. 1984) (emphasis added).

[57] *Id.*

[58] *See supra* note 4.

[59] The "based upon" language was added in 1976. *See supra* note 4.

Then in *Galoob v. Nintendo*, a case involving a technology that allowed Nintendo video game players to alter features of a game (e.g., by increasing the number of lives of the player's character), the Ninth Circuit held that the derivative work right could be infringed even without the production by the defendant of a "fixed" derivative work, thus eliminating a fixation requirement from the infringement equation.[60] To reach this conclusion, the court found that the statute, which equates "creation" with "fixation," was not to be read as containing an indirect definition of the notion of "work," only as providing a definition of the *moment* of creation/fixation of a work.[61] The court also noted that the statutory definition of "derivative work" did not explicitly require fixation.[62] Interestingly, the court considered a Seventh Circuit case concerning a chip used to speed up video games where the court found that the defendant's chip contained a derivative work.[63] The Ninth Circuit drew a distinction because the technology in the Seventh Circuit case copied the computer chip, whereas the technology at play in *Galoob* did not.[64] Yet, as the Seventh Circuit itself acknowledged, the Ninth Circuit noted that it was a "stretch" to find the chip contained a derivative work, a stretch that the Ninth Circuit considered a bridge too far.[65] As this Article sees it, the Ninth Circuit cases focus largely on the *protection of the market* for the original work and potential derivatives made or authorized by the copyright owner. Indeed, this Article agrees with Professor Goldstein's view that a derivative work "effectively creates a new work for a different market."[66]

There are potentially dissonant notes in Ninth Circuit jurisprudence, however. For example, in two cases, that Circuit agreed with the proposition that a "work will be considered a derivative work only if it would be considered an infringing work if the material which it has derived from a prior work had been taken without the consent of a

---

[60] Lewis Galoob Toys, Inc. v. Nintendo of Am., Inc., 964 F.2d 965, 967–68 (9th Cir. 1992), as amended (Aug. 5, 1992) ("A derivative work must be fixed to be protected under the Act, see 17 U.S.C. § 102(a), but not to infringe.").

[61] *Galoob*, 964 F.2d at 968, 970.

[62] *Id.*; *Lone Ranger*, 740 F.2d at 722

[63] *Galoob*, 964 F.2d at 969 (citing Midway Mfg. Co. v. Artic Int'l, Inc., 704 F.2d 1009, 1013–14 (7th Cir. 1983)).

[64] *See id.*

[65] *Id.* (citing *Midway Mfg.*, 704 F.2d at 1014).

[66] *Lone Ranger,* 740 F. 2d, at 721 (quoting Paul Goldstein, *Derivative Rights and Derivative Works in Copyright*, 30 J. COPYRIGHT SOC'Y U.S.A. 209, 217 (1983)).

copyright proprietor of such prior work."[67] This would seem to imply that the derivative work must meet all the requirements for protection as a work. The Nimmer treatise and several appellate circuits prefer that approach.[68] According to Nimmer, a "work will be considered a derivative work only if it would be considered an infringing work if the material which it has derived from a preexisting work had been taken without the consent of a copyright proprietor of such preexisting work."[69]

Other circuits have taken what looks like a different approach. In a case involving a fact pattern fairly similar to *Muñoz* and *Mirage*, the Seventh Circuit considered the Ninth Circuit precedents, but found that the tiles self-evidently neither "recast" nor "adapted" the preexisting works.[70] The court then took a deeper look at the "transformed" prong but concluded that the preexisting works were not transformed in the slightest, a view with which many copyright scholars agree.[71] As the Court noted, the "art was bonded to a slab of ceramic, but it was not changed in the process. It still depicts exactly what it depicted when it left Lee's studio."[72] In deciding that the works lacked the required originality, the Seventh Circuit ostensibly relied in part on a case predating the entry into force of the 1976 Copyright Act in which the Second Circuit en banc found that plastic reproductions of a public domain cast iron "banks" were not original.[73] The Second Circuit decided in that case to follow "the school of cases in this circuit and elsewhere supporting the proposition that to support a copyright there must be at least some substantial variation, not merely a trivial variation such as might occur in the translation to a different medium."[74]

The Seventh Circuit went a step further than *Lee v. A.R.T.* in *Gracen*.[75] It held that paintings of the character Dorothy in *The Wizard of Oz* (based on stills from the movie) were *not* original enough to be

---

[67] U.S. v. Taxe, 540 F.2d 961, 965 n.2 (9th Cir.1976); *see also* Litchfield v. Spielberg, 736 F.2d 1352, 1357 (9th Cir. 1984).

[68] MELVILLE B. NIMMER & DAVID NIMMER, 1 NIMMER ON COPYRIGHTS § 3.03 (1997).

[69] *Id.* § 3.01.

[70] Lee v. A.R.T. Co., 125 F.3d 580, 582 (7th Cir. 1997).

[71] *See* Pamela Samuelson, *The Quest for a Sound Conception of Copyright's Derivative Work Right,* 101 GEO. L.J. 1505, 1551 (2013) ("Scholars agree that Lee is more persuasive than Mirage, and the most recent case to have confronted the choice between Mirage and Lee followed the latter.").

[72] *Lee*, 125 F. 3d, at 582.

[73] L. Batlin & Son, Inc. v. Snyder, 536 F.2d 486 (2d Cir. 1976) (en banc).

[74] *Id.* at 491.

[75] Gracen v. Bradford Exch., 698 F.2d 300 (7th Cir. 1983).

protected by copyright.[76] The painting even changed the background using a different scene from the movie. All to no avail.[77] Yet the court also noted that a more or less exact replica of a scene captured from nature would be original enough.[78] *Gracen* would seem to support the idea of a heightened originality standard for derivative works, which most courts have rejected.[79] Finally, the Eleventh Circuit also discussed *Mirage*, which it described as "much-criticized," and found the epoxy glue argument (as compared to traditional framing) "highly questionable."[80]

This tour d'horizon allows us to answer the two above-mentioned questions. First, if a derivative work is legally produced, then must it meet the originality and fixation requirements *to be protected* as a work? The answer is yes, and the applicable standard is the same as for other types of works.[81] The legal nature of the derivative work can stem from an authorization from the copyright owner (directly, as a valid sub-license, etc.), from an exception such as fair use, or because the underlying work is no longer protected.[82] Second, *to infringe* the right to prepare derivative works in 17 U.S.C. § 106(2), must the alleged derivative work meet the originality standard, even if the work need not be fixed? Here, we may face a circuit split, with the Seventh and Eleventh Circuits on one side, and the Ninth Circuit on the other—though not all of its precedents are necessarily neatly lined up in single file. Moving forward, either we accept that to infringe the derivative work right the defendant must have contributed some originality, or the

---

[76] *See id.*

[77] "We do not consider a picture created by superimposing one copyrighted photographic image on another to be 'original'—always bearing in mind that the purpose of the term in copyright law is not to guide aesthetic judgments but to assure a sufficiently gross difference between the underlying and the derivative work to avoid entangling subsequent artists depicting the underlying work in copyright problems." *Id.* at 305.

[78] *See id.*

[79] *See* RESTATEMENT, *supra* note 9, § 3, comment e ("From time to time, some courts have suggested that there is a special, heightened originality standard for derivative works. But most have rejected that view. . . . The better view, and the one adopted by this Restatement, is one that applies a consistent originality standard to derivative and non-derivative works.").

[80] Peter Letterese & Assocs., Inc. v. World Inst. of Scientology Enters., 533 F.3d 1287, 1299 (11th Cir. 2008).

[81] *Id.*

[82] "Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work [...] is not an infringement of copyright," 17 U.S.C. § 107. This means that fair use covers copyright rights contained in those two sections, including the right to prepare derivative works.

Ninth Circuit correctly stated the law in not looking for originality in the defendant's production, even though its application of that law to the facts of *Mirage* and *Muñoz* may well have been incorrect.[83]   To summarize where we are:

> (a) something must be added or done to a preexisting work to make a derivative work;[84]
> (b) to be protected as a derivative work, a production must be original and fixed, and at least partially non-infringing;[85]
> (c) it seems well-established that a production can infringe the derivative work right even if it is unfixed; and
> (d) is the derivative work infringed if a production, whether fixed or unfixed, recasts, adapts or transforms one or more preexisting works even if said production is not original?

Before moving to a suggested answer, however, a brief look at the difference between derivation and reproduction is in order.

D. *The Relationship between Reproduction and Derivation*

As already noted, the preparation of a derivative work (what one could label derivation) and reproduction are two of the exclusive rights of a copyright owner.[86] There is little doubt that the two rights, if seen as separate sets on a Venn diagram, would overlap considerably. Take the adaptation of a novel to the stage. Much of the novel's expression would likely be copied in the play, in particular, of course, the

---

[83]   A question embedded in the first alternative is whether the originality added to preexisting works must be self-standing, a matter on which there is some degree of disagreement among scholars. *See* WILLIAM F. PATRY, PATRY ON COPYRIGHT, § 12:14.50 (2019) ("There is no necessary correlation between originality and infringement. The statute is to contrary."). *But see* 2 PAUL GOLDSTEIN, COPYRIGHT § 7.3 (3d ed. 2005) (suggesting that a derivative work must contain material that is "capable of standing on its own as a copyrightable work."). The latter statement is certainly correct as to whether a protectable derivative work was produced.

[84]   Courts have used several tests to decide whether "enough" was added to make a production a derivative work, but generally in the context of deciding whether it was protectable, not whether it was infringing. For a discussion, see NIMMER § 3.54 (2020). There is no disagreement that this Author can see about the need for a protectable derivative work to be original. The question is whether originality (but not fixation) is required to infringe.

[85]   The text of the statute suggests rather clearly that it is only parts of a derivative work that contain infringing material that are unprotected. 17 U.S.C. § 103(a) ("[P]rotection for a work employing preexisting material in which copyright subsists does not extend to any part of the work in which such material has been used unlawfully.").

[86]   17 U.S.C. § 106(1)–(2).

dialogues.[87] Which exclusive right would be infringed? In a nutshell, the infringement standard for the right of reproduction is one of substantial similarity between protectable elements of the plaintiff's work and the defendant's work, which essentially means protected expression.[88] In such a case, a reproduction right analysis may be justified yet this is principally, as this Article sees it, a matter for the derivative work right.

That right can apply even if or when the right of reproduction reaches its limits. This is not surprising, as the domain of both rights must be different to avoid superfluity of the statutory text.

A full discussion of the relationship between 17 U.S.C. §§ 106(1) and (2) is beyond the scope of this Article. In a previously published work, I argued that, though the outcome of the analysis (that is, determining whether the defendant's work is prima facie infringing or non-infringing) is often the same, the analytical path is different.[89] In the right of reproduction context, the infringement stems, as just noted, from the copying of protected expression.[90] In the case of derivation, the test is whether creative choices that gave the plaintiff's work its originality were copied.[91] The latter analysis could be said to operate at a higher level of abstraction, as it searches for the actual cause of the originality instead of the original expressive elements in themselves.[92] Adding a teleological layer to the analysis, recall, in addition, that the

---

[87] Sheldon v. Metro-Goldwyn Pictures Corp., 81 F.2d 49, 55–56 (2d Cir. 1936) (noting that one can also infringe by taking elements from a novel without any of the dialogues when making an adaptation as a motion picture because "some of it is plainly drawn from the novel; but that is entirely immaterial; it is enough that substantial parts were lifted.").

[88] *See, e.g.*, Concrete Mach. Co., Inc. v. Classic Lawn Ornaments, Inc., 843 F.2d 600, 606 (1st Cir. 1988) ("Copying therefore is generally established by showing that the defendant had access to the copyrighted work and that the offending and copyrighted articles are 'substantially similar.'"). On the limitation that this test is applied to protectable elements, see Compulife Software Inc. v. Newman, 959 F.3d 1288, 1306 (11th Cir. 2020) ("If the defendant carries this burden as to any portion of the copied material, that material should be filtered out of the analysis before comparing the two works."). Expression is used here to distinguish ideas, which are not protected (under 17 U.S.C. § 102(b)), and "protected expression" is used to indicate that some expression is not protected; Bucklew v. Hawkins, Ash, Baptie & Co., LLP., 329 F.3d 923, 929 (7th Cir. 2003) (For example, if in the case of a merger between idea and expression (in which case the § 102(b) exclusion applies) or the expression constitutes a scène à faire, that is elements that are "so rudimentary, commonplace, standard, or unavoidable that they do not serve to distinguish one work within a class of works from another.").

[89] *See generally* Gervais, *Foxes*, *supra* note 34.

[90] *See supra* note 86.

[91] *See* Gervais, *Foxes*, *supra* note 34 at 839–47.

[92] *See id.*

purpose of derivation is to adapt or recast the preexisting works for a different context, such as those listed as illustrations in the statute.[93]

E. *Application to AI*

The application of the previous Section's findings to the first question above to the AI context means that to be protected under federal law, the production of an AI machine must have originality. As this Author has argued elsewhere, this requires human authorship.[94] Courts would make what this Author considers a grave mistake if they protected literary and artistic productions without (human) originality.[95]

The answer to the second question is the one that this Article will now answer more fully. It can be framed as follows: if one considers that, to infringe the derivative work right, a production must itself be a work and, therefore, original (but not protected, even if original, if it is infringing or unfixed, as noted above), then *machines that by their nature cannot produce originality* (due to a lack of human cause) *cannot infringe this right*.

The solution to this quandary rests in part on a crucial doctrinal point. If the above proposition is true, it would mean that a machine translation, adaptation, or other derivation from one or more preexisting works would be scot-free under 17 U.S.C. § 106(2) if it did not result demonstrably and sufficiently from human creative choices, because a non-original production cannot be a work, and, therefore, it cannot be a derivative work.[96] Naturally, the machine production could still be—and in many cases, as we will see shortly, would likely be considered—a reproduction, one must recall that 17 U.S.C. §§ 106(1) and 106(2) must have different real estate to cover.[97] The key takeaway to use going forward in the analysis is that, under the Ninth Circuit test, a machine could infringe 17 U.S.C. § 106(2) but not under the Seventh (or Eleventh's) Circuit's.

Let us imagine seven different scenarios. For this analysis, let us assume that the works used are (still) protected by copyright, that is, the

---

[93] 17 U.S.C. § 101 ("[T]ranslation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation," although Congress clearly intended this list to be non-exhaustive when it added "or any other form in which a work may be recast, transformed, or adapted.").

[94] *See supra* note 13 and accompanying text. *See generally* Gervais, *The Machine as Author*, *supra* note 1.

[95] *See* Gervais, *The Machine as Author*, *supra* note 1, at 2102–03.

[96] Gervais, *The Machine as Author*, *supra* note 1, at 2100–01.

[97] *See supra* Sections III.C, D.

expiration of the term of protection will be set aside for illustrative purposes.[98]

Scenario 1. AI machine produces a painting "based upon" a dataset comprising 100 Impressionist and post-Impressionist paintings showing landscapes from the South of France. The corpus includes works by Bonnard, Cezanne, Matisse, Morisot, Renoir, Signac, and others. The machine's output gives (if the reader will pardon the pun) the same impression as a "traditional" Impressionist painting but it does not *include any distinguishable element from any of the paintings in the dataset.*[99]

In this first scenario, as this Article sees it, the computer does not in fact *derive*; instead, it finds correlations and patterns to use as a matrix for its own production. As there is no substantial similarity between the painting used in the machine-learning context, the issue of substantial similarity does not apply. When it comes to derivation, the aggregation of choices into a "blend" where preexisting works are no longer individually identifiable means that we are not in the presence of an infringing derivative work—even assuming that what the machine has produced is a "work." True, the correlations probably reflect in some distant way the creative choices that authors of preexisting works incorporated into their expression, but that does not recast, transform or adapt identifiable preexisting works.

Scenario 2. AI machine produces a painting of the famous Japanese bridge in Monet's garden at Giverny "based upon" a dataset comprising thirty paintings of the bridge by various artists. The machine's output resembles existing paintings, including some by Monet, but none are a replica.

While the analysis in this second case is fairly similar to the previous scenario, it is possible that the output, though not a replica, will bear some substantial similarity to one or more of the works used in the machine-learning context, which should trigger a reproduction right analysis, work-by-work. It is much less likely that the machine's output would infringe the right to prepare derivative works for three reasons. First, if there are identifiable, substantially similar elements from individual works in the data corpus in the machine's output, then, as just noted, this should be considered under 17 U.S.C. § 106(1). Second, it

---

[98] 17 U.S.C. §§ 302(a), 305 (establishing that for a typical work of art under current U.S. rules, the copyright lasts for the life of the author plus seventy years, extended until December 31st of the seventieth year after the author's death.).

[99] OXFORD ILLUSTR. ENCYCL. OF THE ARTS 218 (J.J. Norwich ed., Oxford Univ. Press 1990) (explaining that the term "Impressionist" was actually used derisively to refer to the impression that works that tried "to capture the effects of light on various surfaces, particularly in open-air settings" tried to create).

strikes the Author as functionally impossible to isolate creative choices of one author of one preexisting work in a way that would not be in fact a copy of the actual expression. Third, teleologically, there is no adaptation or recasting here of the type that the statute suggests is the purpose of 17 U.S.C. § 106(2). It is a painting similar to those in the dataset.

Scenario 3. AI machine produces a painting of the same bridge in the same garden as in the previous scenario, but "based upon" five paintings all by Claude Monet.

In this third scenario, there is a higher risk that the reproduction right would be engaged because there is a smaller, finite number (five) of well-identified works. The derivative work right could also be relevant under a broad understanding of "based upon" since all the creative choices are by the same author. One could argue that, if creative choices were "transferred" from the five paintings to the machine-produced painting, then that could trigger the application of right. But for the same reasons as in the previous scenario, this Article suggests that the analysis should be carried out under § 106(1). It is not the purpose of § 106(2) to be the glazing on the § 106(1) donut, as it were. Each right has its own target and purpose.

Scenario 4. AI machine produces a painting "based upon" a dataset of five abstract paintings that contain geometric shapes such as circles, squares, and triangles.

The main difference between the third and fourth scenarios in terms of the infringement analysis is that common symbols are not protected expression.[100]

Scenario 5. AI machine, using a 3D printer, creates sculptures using any image or set of no more than five images as its basis. The user of the AI machine has obtained a license to reproduce the images. When using more than one image, the machine selects parts of each. The shape of the sculpture is based on a large database of modern sculptures and images and is designed to fit what the machine "recognizes" in the image or images it is asked to "turn into" a sculpture.

This fifth scenario stands directly in the analytical path of this Article. The images that were reproduced would normally trigger a § 106(1) analysis, but the reproduction is licensed. Then, as images were transformed into a sculpture, a stronger case of derivation can be made than in previous scenarios. This is not quite as straightforward as in the

---

[100] Copyright protects original expression, not common symbols, letters. *See* RESTATEMENT OF THE LAW, COPYRIGHT, *supra* note 9, § 5(c).

so-called tile cases, where images were simply cut and pasted.[101] Moreover, the market for sculptures is possibly quite different than the market for the images themselves. A real *recasting* of the image(s) takes place. Absent fair use or some other defense, the application of § 106(2) may thus be warranted.

Scenario 6. AI machine produces music "based upon" the dataset in scenarios 1, 2, and 3 and a vast database of musical recordings containing classical, jazz, and rock pieces and information about human emotional reactions to colors and shapes, harmonies, and melodies. In creating the music, the machine tries to correlate images and sounds that produce similar emotional reactions.

In this sixth scenario, there is no reproduction of any of the paintings. Is there a derivation? Again, a broad reading of the statute could lead to this conclusion and music does appear to be a different market and a type of transformation that could lead one to argue for the application of § 106(2). Yet, what is "taken" from the painting does not appear in the output. The paintings become pure data, as it were. The distance between the paintings and the musical output is such that neither § 106(1) nor § 106(2) is triggered.

Scenario 7. AI machine produces a translation into English of all the novels that won the top literary prizes in fiction in French, Italian, Japanese and Spanish. No licensing arrangement is in place.[102]

This scenario demonstrates the problem with the "only original derivative works can infringe" approach combined with a human authorship requirement (and recall that abandoning that requirement could have catastrophic effects[103]). The machine would be able to produce these free translations without a license unless one were also able to make a case under the reproduction right, which takes us back to the circularity problem: if every violation of 17 U.S.C. § 106(2) is also a violation of 17 U.S.C. § 106(1), then Congress wasted legislative ink.

Does it make sense to achieve an outcome where machine productions can produce scot-free (under 17 U.S.C. § 106(2), at least) but not humans? Asked that way, the answer emerges. Machines should

---

[101]   *See supra* Section III.C for a description of the tile cases.

[102]   This isn't that far-fetched. For examples of sites that offer this service as of early 2022, *see* Dragneel, *Top 5 Machine Translation Sites for Novels,* DRAGNEELCLUB (Jan. 1, 2022), https://dragneelclub.com/top-5-machine-translation-sites-for-novels/.

[103]   *See supra* notes 12 and Gervais, *The Machine as Author*, *supra* note 1. In this case, this would likely take the floor from under the market for human translators, and that expertise would be lost.

not be placed in a better position than humans.[104] Then a teleological layer can be added to the analysis. If the purpose of the statute is to allow new creations to be inspired by existing ones, "hopefully, ad infinitum," in a cycle that "makes copyright 'the engine of free expression,'" then *whose expression* are we talking about?[105]

The Supreme Court also wrote that "[b]y establishing a marketable right to the use of one's expression, copyright supplies the economic incentive to create and disseminate ideas."[106] This followed in the wake of another pronouncement by the Court, three decades earlier, that "[t]he economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement *of individual effort* by personal gain is the best way *to advance public welfare through the talents of authors* and inventors in 'Science and useful Arts.'"[107]

As I read these encapsulations of copyright law's policy underpinnings, the reward to an author is for what that author has created, and what is then "derived" (as the term is defined in copyright law) from her work. As explained above, the key distinction between reproduction and derivation is that the former copies the expression of a protected work, while the latter means reusing the creative choices that make a work original.[108] Applied to the AI context, the fact that the human programmer of the AI machine can get copyright protection for

---

[104] If one wants to claim that machines should have equal rights, then one can endeavor to prove that point. This Article most certainly will not.

[105] Golan v. Gonzales, 501 F.3d 1179, 1183 (10th Cir. 2007) (citing Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 558 (1985)) (emphasis added). As the Supreme Court noted, "the Framers intended copyright itself to be the engine of free expression." *Harper & Row*, 471 U.S. at 558.

[106] *Harper & Row*, 471 U.S. at 558.

[107] Mazer v. Stein, 347 U.S. 201, 219 (1954) (emphasis added).

[108] *See* Gervais, *Foxes*, *supra* note 34, at 807 ("The qualitative part of the reproduction inquiry focuses chiefly on the *form* of what was taken while the derivation inquiry looks at a deeper level of appropriation, namely at whether the *creative choices* that made the primary work worthy of copyright protection were taken."). There is then a separate question as to whether the person using expression and creative choices contained in one or more previous works has transformed the same, which would then trigger a fair use analysis under the now common text of Transformativeness. *See* R. Anthony Reese, *Transformativeness and the Derivative Work Right*, 31 COLUM. J.L. & ARTS 467, 476 (2008) ("[I]n cases in which the court found that an alleged infringer had violated—or could be found to have violated—the derivative work right, courts showed no inclination to treat the transformation involved in the preparation of the derivative work as "transformativeness" in analyzing the first fair use factor."). For a comparative law analysis, see Mary W. S. Wong, *"Transformative" User-Generated Content in Copyright Law: Infringing Derivative Works or Fair Use?*, 11 VAND. J. ENT. & TECH. L. 1075, 1112–1114 (2009).

the code she produced is not controversial, but then to stretch the beam of protection to outputs produced by the machine using that code is at least one step removed—a crucial step. The test will be whether the creative choices *made by the program's author* (or arguably by the user, if applicable) are present in the machine's output.[109] If not, protecting that output *as the work of the programmer* (or user) is incompatible with both fundamental doctrinal tenets of copyright and its policy purpose, and it over-rewards the programmer (or user).[110]

This Article thus concludes that productions need to be original to infringe 17 U.S.C § 106(2).[111] Under an originality test, the question would be whether a production, if fixed and to the extent that it is non-infringing, would be protectable. In this Article's view, that is the incorrect inquiry. Originality, as defined in *Feist*, is a protection threshold, not an infringement standard.[112]

IV. DIFFERENT DEBATES ABOUT OTHER AI-RELATED RIGHTS AND EXCEPTIONS

In this short Part, this Article explores two related areas of law to ensure that the arguments contained up to this point are neither misunderstood nor applied out of context. Those two areas are the calls for a new, sui generis right in AI productions, and the relationship between the protection of AI outputs and the need for text and data mining exceptions.

A. *Sui generis rights*

If courts follow the doctrinal and/or normative views expounded in this Article and espoused in whole or in part by several other scholars, then there would be no copyright protection for AI productions that do not have a human cause.[113] As the previous Section explicated, AI machines that process datasets consisting of protected works will, however, be able to produce outputs that infringe both the reproduction and, in some cases, the derivative work right. The first of those conclusions will quickly find in its way a normal reflex in our market-

---

[109] *See* Gervais*, The Machine as Author, supra* note 1, at 2095–96.

[110] *See* Samuelson, *supra* note 13, at 1208 (It would "over-reward[] the programmer, particularly in light of the fact that the programmer is no more able to anticipate the output than anyone else."). Anticipation is of course directly related to causality.

[111] In saying that, however, this Article would still take the view that, as applied to their facts (tiles on which images were glued), *Mirage* and *Munoz* were wrongly decided. *See supra* note 50.

[112] *See supra* note 10 and accompanying text.

[113] *See* Gervais, *The Human Cause, supra* note 1.

2022]   *AI DERIVATIVES*   1133

based economy to appropriate any value that can be traded.[114] Hence, if copyright is not available, calls for a new, so-called sui generis right will likely emerge. Indeed, they already have.[115]

Sui generis protection of databases, as it exists in the European Union (EU), is predicated not on progress of science and useful arts or human progress more generally but rather on investment protection.[116] The temptation to create such a right is simple, yet ill-founded. The argument goes as follows: there is investment in AI and any investment should benefit from some form of legal protection. This claim is either false or at least a major overstatement.[117] Then, the temptation is also based on the fact that sui generis rights do not require originality, which obviates the doctrinal difficulty explained in the previous Part.[118]

There are several solid arguments that cast serious doubt on the usefulness of a sui generis right in AI machine productions. First, this Author has been unable to find convincing evidence of underinvestment in AI research and development.[119] Second, AI machines may in short

---

[114] *See* Gregory Alan Bonadies, Property Rights and International Trade: An Institutional Determinant of Export Structure, at 14 (Dec. 2016) (Ph.D. dissertation, The University of Southern Mississippi) https://aquila.usm.edu/cgi/viewcontent.cgi?article=1904&context=dissertations ("Governments exist to develop and implement systems of property right laws to control the use and exchange of property for public and private purposes including production.").

[115] *See* Ana Ramalho, *Will Robots Rule the (Artistic) World?: A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems*, 21 J. INTERNET L. 1, 20 (2017) ("It also is possible to consider the grant of a *sui generis* right in AI-created works, much as the EU legislature decided to do for makers of databases. The common rationale here is protection of investment."); Ricketson, *supra* note 1, at 36–37 (suggesting that humans who own or use machines might "obtain strong and effective protection under a neighboring rights or sui generis regime").

[116] *See* Ramalho, *supra* note 115; Katleen Janssen & Jos Dumortier, *The Protection of Maps and Spatial Databases in Europe and the United States by Copyright and the Sui Generis Right*, 24 J. MARSHALL J. COMPUT. & INFO. L. 195, 214 (2006) (noting that the EU's sui generis database right "protects the investment in a database, and not creativity").

[117] There is some form of protection, for example against expropriation or, in the case of foreign investment by multinationals, under investor-state-dispute-settlement ("ISDS"), but there is no rule that any and all investment must be protected by law. *See* Daniel Gervais, *Investor-State Dispute Settlement: Human Rights and Regulatory Lessons from* Lilly v. Canada, 8 U.C. IRVINE L. REV. 459, 466 (2018) ("ISDS provides multinational corporations a right to sue states that are parties to an investment treaty (such as a bilateral investment treaty or BIT) or a trade agreement containing an investment protection chapter for direct or indirect expropriation.").

[118] *See* C.D. Freedman, *Should Canada Enact a New Sui Generis Database Right?,* 13 FORDHAM INTELL. PROP., MEDIA & ENT. L.J. 35, 37–38 (2002) (explaining that EU law protects "both original and unoriginal databases").

[119] Quite the opposite. For example, the International Data Corporation forecasts accelerating growth over the next few years. *See IDC Forecasts Improved Growth for*

order be able to produce massive amounts of literary and artistic output, which will quickly risk crowding out the field if each production was protected by exclusive copyright rights, or even a sui generis right against some form of reuse.[120] Third, while the United States does not have sui generis database protection (despite several failed attempts in Congress to enact it), it does not seem to have provided the EU with a noticeable comparative advantage in database investment.[121]   If anything, the database directive is viewed as a policy failure.[122]

Against this backdrop, it seems reasonable to suggest that proponents of a new sui generis right in literary and artistic productions of AI machines bear the burden of proof. They should not be allowed to rely simply on hollow-ringing policy clichés. At the very least, one should be able to identify a market failure, and that failure cannot be the insufficiency of the rate of replacement of human authors in the marketplace, as this is the exact opposite of what policy endeavors should aim for.[123] Lastly, if and when the proponents of a sui generis right are able to discharge their burden of proof, there will remain "immense challenges" in defining the exact contours of the right.[124]

Global AI Market in 2021, INT'L DATA CORP. (Feb 23, 2021), https://www.idc.com/getdoc.jsp?containerId=prUS47482321.

[120] *See* Patrick Zurth, *Artificial Creativity? A Case Against Copyright Protection for AI-Generated Works*, 25 UCLA J.L. & TECH. 1, 15 (Spring 2021). EU law, namely the Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases, provides a right against certain forms of "extraction" of data, rather than traditional copyright exclusive rights.

[121] On attempts to introduce EU-style law in the United States, see Janssen & Dumortier, *supra* note 116, at 223 ("Since the introduction of the European database directive in 1996, a number of bills have been introduced in the U.S. Congress, some of which were based on the database directive . . . ."). *See also* Samuel E. Trosow, Sui Generis *Database Legislation: A Critical Analysis*, 7 YALE J.L. & TECH. 534, 627 (2004–05); Daniel Gervais, *The Protection of Databases*, 82 CHI.-KENT L. REV. 1109–1169 (2007).

[122] *See* European Commission, *Commission Staff Working Document: Executive Summary of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases*, SWD (2018) 147 final (Apr. 25, 2018), https://digital-strategy.ec.europa.eu/en/library/staff-working-document-and-executive-summary-evaluation-directive-969ec-legal-protection-databases ("[T]he *sui generis* right continues to have *no proven impact* on the overall production of databases in Europe, nor on the competitiveness of the EU database industry.").

[123] For a more complete explanation on this point, see Gervais, *The Machine as Author*, *supra* note 1.

[124] Peter K. Yu, *Data Producer's Right and the Protection of Machine-Generated Data*, 93 TUL. L. REV. 859, 929 (2019).

B. *Text and Data Mining*

Up to this point, this Article has essentially focused on the output of AI machines. Text and data mining ("TDM") is the flip side of the process—namely, the input. It refers to "computational processes for applying structure to unstructured electronic texts and employing statistical methods to discover new information and reveal patterns in the processed data."[125] The terms "text" and "data" are "broad enough to include fixed images, sound recordings, and audio-visual works."[126] The ultimate example of TDM might be Google scanning entire libraries of books, a process that the Second Circuit determined was a fair use.[127] The EU, whose legislative experiment about databases was discussed in the previous Section, has experimented in this area as well, with the adoption in 2019 of a Directive on Copyright in the Digital Single Market that provides two separate provisions on TDM.[128] The first exception allows TDM "for the purposes of scientific research" by "research organisations and cultural heritage institutions."[129] The second exception is not limited to any particular category of user, but rightsholders are able to opt out.[130] The Directive also explains that "there is widespread acknowledgment that text and data mining can, in particular, benefit the research community and, in so doing, support innovation."[131] Other countries have also adopted similar TDM exceptions and limitations, including Japan and Singapore.[132]

---

[125] Eleanor Dickson et al., *Synthesis of Cross-Stakeholder Perspectives on Text Data Mining with Use-Limited Data: Setting the Stage for an IMLS National Forum* 5, https://www.ideals.illinois.edu/bitstream/handle/2142/100055/IMLSNationalForum_TDM_DiscussionPaper.pdf?sequence=2&isAllowed=y (last visited Jan. 30, 2022); *see also* Veale, *supra* note 18 and accompanying text.

[126] Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y. U.S.A. 291, 294–95 (2019).

[127] Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97 (2d Cir. 2014); Authors Guild v. Google, Inc., 804 F.3d 202, 207 (2d Cir. 2015). There is little doubt in this Author's mind that the main value of this corpus is not in making snippets available (though that is what gave the use its fair nature), but the possibilities afforded by the mining of the data. At this point, only Google and those it partners with can reap that benefit.

[128] Directive 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, 2019 O.J. (L 130) 92.

[129] *Id.* art. 3(1).

[130] *Id.* art. 4(3).

[131] *Id.* recital 8.

[132] On Japan, see generally Tatsuhiro Ueno, *The Flexible Copyright Exception for 'Non-Enjoyment' Purposes – Recent Amendment in Japan and Its Implication*, 70 GRUR INT'L 145 (2021). For a summary of the Singapore provisions, see Alban Kang & Pin-Ping Oh, *Coming Up in Singapore: New Copyright Exception for Text and Data Mining*, BIRD & BIRD:

In the United States, the solution is more likely to come from courts interpreting the fair use doctrine than from Congress.[133] Be that as it may, the debate about allowing TDM inputs, which can have a positive valence in many areas, for example, in assisting scientific research, should be dissociated from the separate matter of the protection that applies, or could apply, to the outputs of AI machines.

V. Conclusion

Lawmaking by courts and legislators in the coming years will include the unprecedented challenge of adapting "our" legal system to a new, artificially intelligent "species" that can mimic and indeed surpass humans at many tasks involving the one feature that had distinguished humans from other species, and indeed allowed the dominion of humans over those species.[134] One of the specific challenges will be the adaptation of copyright law to literary and artistic productions of AI machines. This Article makes a contribution to the exploration of this challenge by explicating how the derivative work right, sometimes seen as copyright's poor cousin in a house occupied mostly by the rights of production and public performance, may be called upon to step forward in litigation to either protect machine productions or find them infringing, in part because the very process of AI-based machine learning, which leads to those productions, is arguably a form of derivation. This Article explained why the scope of the derivative work right should be cabined in that context and considered the caselaw concerning the requirement that a derivative work be original and suggests how it should be applied in the AI context. Finally, this Article also briefly discussed proposals to create a non-copyright, sui generis right in machine outputs, and clarified the linkage between the debates about copyright in such outputs and exceptions designed to allow text and data mining.

News Ctr. (Sept. 19, 2021), https://www.twobirds.com/en/news/articles/2021/ singapore/coming-up-in-singapore-new-copyright-exception-for-text-and-data-mining.

[133]   *See* Sag, *supra* note 126, at 366 (explaining that U.S. courts do not see TDM as interfering with the protection of copyrighted works).

[134]   *See generally* Daniel Gervais, *Towards an Effective Transnational Regulation of AI*, AI & Soc'y (2021).