THE COLUMBIA
SCIENCE & TECHNOLOGY
LAW REVIEW

VOLUME 22 STLR.ORG NUMBER 2

ARTICLE

GROUND-LEVEL PRESSING ISSUES AT THE INTERSECTION OF AI AND IP

David Kappos* and Asa Kling†

Humankind has always sought to solve problems. This impetus has transformed hunters and gatherers into a society beginning to enjoy the fruits of the fourth industrial revolution. As part of the fourth industrial revolution, and the increased computing power accompanying it, the long-theorized concept of artificial intelligence ("AI") is finally becoming a reality. This raises new issues in myriad fields—from the moral and ethical implications of replacing human activity with machines to who will own inventions created by AI. While these questions are worth exploring, they have already received a fair amount of coverage in popular and theoretical writing. This paper will take a different direction, focusing on the current and near-future issues arising on the ground at the intersection of AI and intellectual property ("IP"). After providing a brief overview of AI, we will analyze legal issues unique to AI, including access to data, patent requirements, open source licenses and trade secrecy. We will then suggest best practices for obtaining and preserving IP protection for AI-related innovations through the United States and European Union IP systems.[1] By addressing these issues, the intellectual property system will be better positioned to do its part in unlocking AI's immense potential.

* David J. Kappos is a partner at Cravath, Swaine & Moore LLP. Previously, he served as the Under Secretary of Commerce and Director of the United States Patent and Trademark Office from August 2009-January 2013.

† Asa Kling is a partner at Naschitz Brandes Amir & Co. Previously he served as the Director of the Israel Patents Office, Commissioner of Patents, Designs and Trademarks from May 2011-June 2017.

[1] This work represents an integrated United States/European Union approach to the issues on the basis of the authors' respective experience in the United States and Israel (which closely follows European doctrines in this field).

I. DEFINING AI

No definitive definition of AI exists. There are numerous definitions centering on a shared concept: machines that can perform, and become better than humans at performing, a task traditionally performed by humans with little or no human oversight.[2] Even in antiquity, the Greek god of metalworking, Hephaestus, was said to have created golden machines to serve him. The concept behind this myth repeats in various forms and various fields throughout history, but it was not referred to as "artificial intelligence" until 1955 when researchers from Dartmouth, Harvard, IBM and Bell Telephone Laboratories coined the term to describe research they sought to perform.[3] In the sixty-five years since they coined the term, many researchers in the field have advanced artificial intelligence from a hypothetical concept to the powerful tool it is today.

Various genres of AI have come into existence in recent decades. There is "narrow AI," also referred to as "weak AI."[4] Narrow AI is AI that performs a specific individual task—e.g., playing chess[5] or reading CT scans—but cannot perform other tasks.[6] Narrow AI's one-dimensional nature (which borders with mere sophisticated automation) stems from its reliance on a singular dataset. If a dataset does not contain information about how to perform a task, narrow AI cannot

---

[2] *See, e.g.*, WIPO, Technology Trends 2019: Artificial Intelligence 19 (2019) [hereinafter WIPO AI Report].

[3] *See* Gil Press, *A Very Short History of Artificial Intelligence (AI)*, FORBES (Dec. 30, 2016, 9:09 AM), https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/#5ab2ac176fba.

[4] WIPO AI REPORT, *supra* note 2, at 19.

[5] Although chess is discussed in the WIPO AI Report as an example of early-stage AI, some consider playing chess as mere high-level automation. Rather than solving unforeseeable occurrences, certain "AIs" play chess by computing all potential moves and outcomes before selecting upcoming moves. *See* WIPO AI REPORT, *supra* note 2, at 19.

[6] Tannya D. Jajal, *Distinguishing Between Narrow AI, General AI and Super AI*, MEDIUM (May 21, 2018), https://medium.com/@tjajal/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22.

perform that task.[7] In other words, narrow AI can only perform (and improve its performance of) the task it was designed to perform; it does not have the ability to adapt and perform additional tasks. Narrow AI is in use extensively today.

Alternatively, there is "general AI," also referred to as "strong AI."[8] General AI involves machines that can "perform any intellectual task that a human being can."[9] Unlike narrow AI, general AI may not be dependent on the dataset that feeds it, but instead may create its own abstract "thought process" to evolve upon its initial dataset to solve problems not necessarily identifiable at the outset of the system's design. General AI and superintelligent AI, which is AI that surpasses human intelligence, do not exist today.[10] However, the investment of extensive resources in some fields (e.g., autonomous vehicles) is quickly bringing about machines capable of matching human ability.

Within narrow AI, the focus of current development is machine learning, which is an iterative process through which AI systems are trained to recognize and react to certain situations based upon repeated exposure to similar situations and their outcomes.[11] Machine learning encompasses state-of-the-art technologies called neural networks and deep learning. Neural networks "are computing systems with interconnected nodes that work much like neurons in the human brain. Using algorithms, they can recognize hidden patterns and correlations in raw data, cluster and classify it, and – over time – continuously learn and improve."[12] Deep learning uses neural networks and massive computing power to sift through large amounts of data and identify complicated patterns such as those underlying real-time language translation.[13] These are just a few of the many terms of art in the AI field. As AI technologies continue to develop, new AI nomenclature will continue to proliferate. For the present, an understanding of narrow AI and its component parts will enable engagement with the state of AI today and the urgent issues facing AI from an intellectual property perspective. Further, tackling the "simple" case of narrow AI now may provide tools to cope with the broader implications of general AI in the future.

---

[7] *Id.*

[8] *Id.*

[9] *Id.*

[10] *See* WIPO AI REPORT, *supra* note 2, at 19.

[11] *See Machine Learning: What It Is and Why It Matters*, SAS, https://www.sas.com/en_us/insights/analytics/machine-learning.html (last visited Apr. 5, 2021).

[12] *Neural Networks: What They Are & Why They Matter*, SAS, https://www.sas.com/en_us/insights/analytics/neural-networks.html (last visited Apr. 5, 2021).

[13] *See* SAS, *supra* note 11.

II. THE STATE OF AI AND INVENTIONS TODAY

To some, "AI" is a magic word akin to abracadabra. Pick a problem, apply AI, and the problem is solved. To these folks, AI is already making inventions and substituting for human judgment.[14]

In reality, AI represents a paradigm shift, but it is not magic. While AI can currently analyze human-provided data to create a food container shaped like a snowflake and a flickering lamp,[15] it has not yet created meaningful inventions. State-of-the-art AI is a useful tool that innovators are using to make inventions, and AI-related inventions are increasingly the subject of patent coverage.[16] Such AI-related inventions are not monolithic and can be roughly divided between (1) AI techniques, (2) functional AI applications and (3) broader fields of AI application.[17] The most important AI techniques include methods such as machine learning, logic programming and fuzzy logic.[1819] Prominent AI functional applications include computer vision, natural language processing, speech processing and robotics.[20] And the broader fields of such applications include areas like transportation, telecommunication, life and medical sciences, personal devices and security.[21] Currently, the majority of AI-related inventions pertain to functional applications of specific narrow AI.[22] In the near future, advancements to AI techniques will increase the number of AI patent applications.

With these advancements, the questions surrounding the patentability of AI will only grow in significance. Some guidance has been put forth by patent offices, including the United States Patent and Trademark Office ("USPTO") and the European Patent Office ("EPO"),[23] that has provided helpful direction in evaluating the patentability of inventions relating to abstract ideas and artificial intelligence.

---

[14] *See* Erik Sherman, *Robots Aren't Coming for Jobs: AI Is Already Taking Them*, FORBES (Oct. 25, 2018, 3:54 PM), https://www.forbes.com/sites/eriksherman/2018/10/25/robots-arent-coming-for-jobs-ai-is-already-taking-them/#7469086c374e.

[15] Leo Kelion, *AI System 'Should Be Recognized as Inventor'*, BBC TECHNOLOGY (Aug. 1, 2019), https://www.bbc.com/news/technology-49191645.

[16] *See* WIPO AI REPORT, *supra* note 2, at 13 (noting that over 50% of all patents in the AI field have been published since 2013).

[17] *Id.* at 13, 95–101.

[18] *Id.* at 42 fig.3.4.

[19] The WIPO AI Report defines (i) machine learning as an "AI process that uses algorithms and statistical models to allow computers to make decisions without having to explicitly program it to perform the task"; (ii) logic programming as the "use [of] facts and rules to make decisions, without specifying additional intermediary steps, in order to achieve a particular goal"; and (iii) fuzzy logic as a "decision-making approach which is... [based] on 'degrees of truth.'" *Id.* at 146.

[20] *Id.* at 46 fig.3.11.

[21] *Id.* at 50 fig.3.17.

[22] *Id.* at 40.

[23] *E.g.*, 2019 Revised Patent Subject Matter Eligibility Guidance, 84 Fed. Reg. 50 (Jan. 7, 2019); October 2019 Patent Eligibility Guidance Update, 84 Fed. Reg. 55942 (Oct. 17, 2019); EUR. PAT. OFF., GUIDELINES FOR EXAMINATION IN THE EUROPEAN PATENT OFFICE, pt. G, ch. II, § 3.3.1 (2021) [hereinafter EPO GUIDELINES], http://documents.epo.org/projects/babylon/eponet.nsf/0/C4B20952A0A7EF6BC125868B002A5C61/$File/epo_guidelines_for_examination_2021_hyperlinked_en.pdf. As discussed in greater detail in Section III.A of this Comment, this is not to say that all AI is patentable subject matter.

This guidance is a good first step, but it is too general to answer all of the questions raised by the multi-faceted character of AI-related inventions. Therefore, important questions about how AI can, and should, be protected by intellectual property laws still need to be answered to unlock AI's many opportunities. Some of these questions are discussed below.

III. THE LESS DISCUSSED, BUT MORE URGENT, ISSUES FOR AI AND IP

Beyond the mostly theoretical issues dominating public discussion of AI and IP, there are less discussed, but more urgent (or near-future), issues requiring attention today.

A. *AI and Patent-Eligible Subject Matter*

Despite the famous saying that patentable subject matter includes "anything under the sun that is made by man,"[24] there are judicially created exceptions to this maxim that prevent applicants from receiving U.S. patents for laws of nature, physical phenomena, and abstract ideas.[25] Similarly, the European Union explicitly precludes certain types of inventions from receiving patents.[26] Over the past decade, the United States Supreme Court has tightened subject matter eligibility requirements through a series of decisions interpreting Section 101 of the Patent Act.[27] Two of these cases, *Alice* and *Mayo*, set out the current test used to determine whether computer-implemented inventions (including AI-related inventions) are patentable subject matter.[28] Under the *Alice*/*Mayo* test, patent-eligible subject matter is separated from patent-ineligible subject matter by first asking whether the claims covering the subject matter are "directed to" patent-ineligible concepts (i.e., laws of nature, physical phenomena or abstract ideas).[29] If the answer to this first question is "yes," then the claims are examined to determine whether they contain an inventive concept sufficient to transform the patent-ineligible subject matter into a patent-eligible application.[30] This test has created tremendous uncertainty in the United States about the patent eligibility of computer-implemented inventions.[31] For instance, whether a method of remotely monitoring the transfer of data between

---

[24] *See* Diamond v. Chakrabarty, 447 U.S. 303, 309 (1980) ("Congress intended statutory subject matter to include anything under the sun that is made by man. This is not to suggest that § 101 has no limits or that it embraces every discovery. The laws of nature, physical phenomena, and abstract ideas have been held not patentable.") (citation omitted).

[25] *Id.*

[26] Convention on the Grant of European Patents art. 52(2), Oct. 5, 1973, 1065 U.N.T.S. 199 (excepting "discoveries, scientific theories and mathematical methods" as well as "schemes, rules and methods for performing mental acts, playing games or doing business, and programs for computers" from the definition of patentable inventions).

[27] *See generally* Alice Corp. Pty. v. CLS Bank Int'l, 573 U.S. 208 (2014); Ass'n for Molecular Pathology v. Myriad Genetics, Inc., 569 U.S. 576 (2013); Mayo Collaborative Servs. v. Prometheus Labs., Inc., 566 U.S. 66 (2012).

[28] *Alice*, 573 U.S. at 217–18.

[29] *Id.* at 217.

[30] *Id.* at 217–18.

[31] *See* Manny Schecter, *Patent Subject Matter Eligibility 101*, IPWATCHDOG (May 8, 2018), https://www.ipwatchdog.com/2018/05/08/patent-subject-matter-eligibility-101/id=96928/.

a local and remote computer[32] or a method for monitoring a casino poker game to assign players to tables based on availability[33] should be considered patent-eligible subject matter is not immediately clear. Since *Alice*, both courts and patent examiners applying this test have treated similar technologies unevenly, and there is no sign that AI inventions have been spared.[34]

To provide improved certainty about whether an invention is patent eligible, the USPTO issued guidance in January 2019.[35] As part of this guidance, the USPTO clarified its view that "a claim is not 'directed to' a judicial exception if [it] is integrated into a practical application of that exception."[36] The guidance included a hypothetical claim covering "a computer-implemented method of training a neural network for facial detection"—i.e., an AI system designed to recognize faces:[37]

> A computer-implemented method of training a neural network for facial detection comprising: collecting a set of digital facial images from a database; applying one or more transformations to each digital facial image including mirroring, rotating, smoothing, or contrast reduction to create a modified set of digital facial images; creating a first training set comprising the collected set of digital facial images, the modified set of digital facial images, and a set of digital non-facial images; training the neural network in a first stage using the first training set; creating a second training set for a second stage of training comprising the first training set and digital non-facial images that are incorrectly detected as facial images after the first stage of training; and training the neural network in a second stage using the second training set.[38]

According to its guidance, the USPTO does not read this claim to be "directed to" patent-ineligible subject matter.[39] Instead, it notes that "[w]hile some of the limitations may be based on mathematical concepts, the mathematical concepts are not recited in the claims. Further, the claim does not recite a mental process because the steps are not practically performed in the human mind."[40] Based on this interpretation of the claim, the USPTO deemed the claim eligible subject matter.[41]

---

[32] *See* Helios Software, LLC v. SpectorSoft Corp., No. CV 12-081-LPS, 2014 WL 4796111 (D. Del. Sept. 18, 2014).

[33] *See* Ameranth, Inc. v. Genesis Gaming Sols., Inc., No. SACV 11-00189 AG (RNBx), 2014 WL 7012391 (C.D. Cal. Nov. 12, 2014).

[34] USPTO, OFFICE OF THE CHIEF ECONOMIST, IP DATA HIGHLIGHTS NO. 3, ADJUSTING TO *ALICE*: USPTO PATENT EXAMINATION OUTCOMES AFTER *ALICE CORP V. CLS BANK INTERNATIONAL* (2020), https://www.uspto.gov/sites/default/files/documents/OCE-DH_AdjustingtoAlice.pdf.

[35] 2019 Revised Patent Subject Matter Eligibility Guidance, 84 Fed. Reg. 50 (Jan. 7, 2019).

[36] *Id.* at 50.

[37] USPTO, SUBJECT MATTER ELIGIBILITY EXAMPLES: ABSTRACT IDEAS 8–9 (2019), https://www.uspto.gov/sites/default/files/documents/101_examples_37to42_20190107.pdf.

[38] *Id.*

[39] *Id.* at 9.

[40] *Id.*

[41] *Id.*

Similarly, the EPO has released recent guidance addressing the eligibility of AI subject matter for patent protection.[42] According to this guidance, the EPO believes AI is "*per se* of an abstract mathematical nature."[43] At the first stage, if a claim is directed to an AI algorithm and the claim does not include subject matter exhibiting any technical character, the claim will be excluded from patentability.[44] But at the second stage, if a claim is directed to an AI algorithm contributing to the technical character of the invention, then the claim will not be excluded from patentability.[45] Put another way, the EPO will not grant a patent merely covering an algorithm, but it will consider granting a patent covering the technical implementation of an algorithm while taking into account all features, technical as well as non-technical, of the implementation (provided the other patent requirements are met).[46] This guidance suggests that the EPO considers AI-related inventions to be a subtype of software-implemented inventions that will be subject to the general eligibility questions raised by all software-implemented inventions.

Thus far, patent offices have provided more guidance than courts about whether AI technology is patent eligible. This lack of guidance from the courts stems from the current lack of litigation involving AI-related patents. While the WIPO AI Report found there were 339,828 patent families related to AI, only 1,264 of these families have been involved in litigation.[47] That is less than half a percent. In comparison, the U.S. patent litigation rate between 2007 and 2016, which compares the number of patents litigated to the number of patents issued in that span, was about 1.9%.[48] Furthermore, a yearly average of only 1.2% of AI-related patent families have been opposed.[49]

Based on the cases that have been litigated, courts appear to ground their patent eligibility decisions on the specificity of the claims; the more specific the claims, the more likely courts will find they are patent eligible. For instance, the District Court for the Northern District of California found that a patent covering a method of automatically generating an ensemble of machine learning modules was ineligible subject matter because the claims covered the general concept of predictive analytics rather than any specific application of predictive analytics.[50] However, the Federal Circuit recently clarified that claims directed to a "specific

---

[42] EPO GUIDELINES, *supra* note 23, pt. G, ch. II, § 3.3.1.

[43] *Id.*

[44] *Id.* pt. G, ch. II, § 3.3 (invoking Articles 52(2) and 52(3) of the European Patent Convention).

[45] *Id.*

[46] This guidance leaves open what constitutes a technical purpose or technical implementation. The EPO's Boards of Appeal have previously held that merely stating that an algorithm is computer-implemented without providing any specific implementation details does not meet the EPO's patent eligibility requirements. Eur. Pat. Off., Tech. Bd. App., Case T 0306/10 § 4.6 (Feb. 4, 2015), https://www.epo.org/law-practice/case-law-appeals/pdf/t100306eu1.pdf.

[47] WIPO AI REPORT, *supra* note 2, at 23, 35, 111.

[48] David Pridham, *The Patent Litigation Lie*, FORBES (Apr. 13, 2017, 12:24 PM), https://www.forbes.com/sites/davidpridham/2017/04/13/the-patent-litigation-lie/?sh=7da192ee7ea9.

[49] WIPO AI REPORT, *supra* note 2, at 115.

[50] PurePredictive, Inc. v. H2O.AI, Inc., No. 17-cv-03049-WHO, 2017 WL 3721480 (N.D. Cal. Aug. 29, 2017).

asserted improvement in computer capabilities" were directed to patentable subject matter.[51] Based upon these cases and guidance from patent offices, AI innovators should highlight the specific technical applications of their inventions in their patent applications. Examples of specific technical applications of AI include "the use of a neural network in a heart-monitoring apparatus for the purpose of identifying irregular heartbeats" and the "classification of digital images, videos, audio or speech signals based on low-level features (e.g. edges or pixel attributes for images)."[52] On the other hand, "[c]lassifying abstract data records or even 'telecommunication network data records' without any indication of a technical use being made of the resulting classification"[53] is not a *per se* technical purpose. Moreover, an applicant who merely states a "generic purpose such as 'controlling a technical system'"[54] has not done enough to claim a specific technical application. Therefore, patent offices and courts currently seem to support the eligibility of AI-related functional applications using specific narrow AI and narrowly tailored implementations of infrastructural AI methods, but suggest that broadly claimed infrastructural AI methods will be *per se* ineligible for lack of specific technical character.

B. *AI and Novelty*

One hurdle inventors seeking to obtain patent protection for AI face is that, at a certain level, many algorithms are not new. Therefore, patent applications seeking to obtain coverage for the algorithms may not meet the novelty requirements found in patent laws around the world. In the United States, Section 102 prevents the patenting of "inventions" that have already been invented and made available to the public.[55] The concept of AI is not new. Modern AI emerged in the 1950s and will soon be a septuagenarian.[56] After a series of AI winters and summers, the current "newness" of AI largely stems from improvements in hardware that enable previously unimaginable amounts of data to be processed.[57] These datasets produce the new results, not the widely commoditized algorithms that process the datasets.[58]

Although deep machine learning may provide for dynamically evolving algorithms, many current algorithms are stable and already commoditized. Section 102 will preclude "inventors" from obtaining patents over such anticipated algorithms. This may compel inventors to claim the interaction between datasets

---

[51] Finjan, Inc. v. Blue Coat Systems, Inc., 879 F.3d 1299, 1303, 1305–06 (Fed. Cir. 2018).

[52] EPO GUIDELINES, *supra* note 23, pt. G, ch. II, § 3.3.1.

[53] *Id.*

[54] *Id.* pt. G, ch. II, § 3.3.

[55] 35 U.S.C. § 102.

[56] WIPO AI REPORT, *supra* note 2, at 13.

[57] Eleanor Cummins, *Another AI Winter Could Usher in a Dark Period for Artificial Intelligence*, POPULAR SCIENCE (Aug. 29, 2018), https://www.popsci.com/ai-winter-artificial-intelligence/.

[58] *See* Boris Wertz, *Data, Not Algorithms, Is Key to Machine Learning Success*, VERSIONONE (Jan. 6, 2016), https://versionone.vc/data-not-algorithms-is-key-to-machine-learning-success/ (noting that as early as 2016, "[a]lgorithms have largely been commoditized by now, so a machine learning company built around publicly accessible data isn't defensible").

and algorithms in their patent applications. No matter how much effort is expended compiling datasets, intangible collections of data are not patentable subject matter.[59] It will be difficult to craft claims reciting unpatentable datasets interacting with anticipated algorithms and to receive patent protection for, or ultimately enforce, such claims.[60] The difficulty will increase if the datasets are already available to the public, as there will be a stronger likelihood that the claimed "invention" was anticipated by applying widely commoditized algorithms to publicly available data. Barring a small number of exceptions,[61] such prior public use will preclude "inventors" from obtaining patent protection for their "inventions" in light of Section 102's novelty requirement.[62] Instead, as discussed above, inventors must focus their patent applications on the technical applications of their AI-related inventions, instead of on the datasets or the algorithms, to maximize their chances of obtaining patent protection.

C. *AI and 35 U.S.C. § 112 Disclosure Requirements*

AI is nothing without data. It relies on big, clean datasets[63] to generate meaningful results.[64] AI implementers are focusing on techniques that can provide results with smaller datasets,[65] but there is still a long way to go even to accumulate the data needed for AI to recognize everyday objects with the efficiency of a toddler.[66] This extreme need for, and reliance on, big data raises serious enablement and written description challenges for inventors seeking to protect their AI inventions with patents.

In the United States, Section 112 of the Patent Act imposes two important requirements on applicants seeking to obtain patents. First, the patent's

---

[59] Digitech Image Techs. v. Elecs. for Imaging, 758 F.3d 1344, 1350 (Fed. Cir. 2014).

[60] *See* WIPO AI REPORT, *supra* note 2, at 17, 111–17 (analyzing patent litigation involving AI-related patents and hypothesizing that the low enforcement rate of AI-related patents may be due to the difficulty of proving infringement of AI patents).

[61] 35 U.S.C. § 102(b).

[62] 35 U.S.C. § 102(a).

[63] By "big data," we mean large amounts of data—i.e., "data sets [that] are so voluminous that traditional data processing software just can't manage them." *What Is Big Data?*, ORACLE, https://www.oracle.com/big-data/guide/what-is-big-data.html (last visited Apr. 5, 2021). By "clean data," we mean data that does not contain incomplete, inaccurate or inconsistently formatted data. *See* Awan-Ur-Rahman, *What Is Data Cleaning? How to Process Data for Analytics and Machine Learning Modeling?*, TOWARDS DATA SCIENCE (Nov. 19, 2019), https://towardsdatascience.com/what-is-data-cleaning-how-to-process-data-for-analytics-and-machine-learning-modeling-c2afcf4fbf45.

[64] WIPO AI REPORT, *supra* note 2, at 34 ("A key advantage to AI lies in its ability to analyze huge datasets and identify patterns and correlations that may pass unnoticed in smaller, or more piecemeal, data."). *See also* Bernard Marr, *Why AI Would Be Nothing Without Big Data*, FORBES (June 9, 2017, 12:29 AM), https://www.forbes.com/sites/bernardmarr/2017/06/09/why-ai-would-be-nothing-without-big-data/#312d9c854f6d.

[65] WIPO AI REPORT, *supra* note 2, at 141.

[66] *Id.* at 8 ("For example, a toddler can usually recognize a cat after just one encounter, but a computer still needs more than one example to learn. We need to find ways to train computers on training datasets as small as 100, or even 10.").

specification must contain a written description of the invention.[67] This written description must "describe the claimed invention in sufficient detail that one skilled in the art can reasonably conclude that the inventor had possession of the claimed invention."[68] Second, the patent's specification must enable the patent's claims.[69] In other words, the specification must contain sufficient information about the claimed invention "so that any person skilled in the art can make and use the invention without undue experimentation."[70] In recent years, the Supreme Court has tightened its interpretation of Section 112, especially in the software field.[71] Now, a patent's specification must provide those having ordinary skill in the patent's art "reasonable certainty" about the scope of the claims.[72]

It may be difficult for patent applications covering AI to both meet the increasingly rigorous disclosure requirements set forth in Section 112 and provide adequate patent protection to the covered AI. Unlike a process for making potash,[73] truly innovative AI is not static. Instead, innovative AI changes over time to increase its efficiency and enhance its performance.[74] In some instances, AI can change in ways not understood by the AI's inventors.[75] If the inventors do not understand aspects of the AI they invented, they will struggle to meet the written description and enablement requirements discussed above. Inventors might be tempted to broadly claim a method of solving the problem the AI addresses. However, such broad claims may well face questions about whether they cover patentable subject matter or mere abstract ideas.[76] Further, these claims may *per se* fall into the trap of claiming the problem rather than any solution. While it is important to identify problems, problems themselves cannot be patented.[77] Moving

---

[67] 35 U.S.C. § 112(a).

[68] MPEP § 2163 (9th ed. Rev. 10, June 2019).

[69] 35 U.S.C. § 112(a).

[70] MPEP § 2164.01 (8th ed. Rev. 8, Aug. 2012) (citing *In re* Wands, 858 F.2d 731, 737 (Fed. Cir. 1988)).

[71] Nautilus, Inc. v. Biosig Instruments, Inc., 572 U.S. 898, 901 (2014) (replacing the Federal Circuit's indefiniteness test with a narrower test that invalidates a patent for indefiniteness "if its claims, read in light of the specification delineating the patent, and the prosecution history, fail to inform, with reasonable certainty, those skilled in the art about the scope of the invention"). *See also* Maliha Khan, *The Supreme Court's Tightening of Patent Definiteness & the Impact of* Nautilus v. Biosig *on the Software Patent Industry*, 4 AM. U. BUS. L. REV. 353 (2015).

[72] *Id.*

[73] Potash is a compound used in fertilizer. The first patent issued by the United States government covered a process for making potash. Grace Bradford, *First U.S. Patent Issued Today in 1790*, SUITER|SWANTZ (July 31, 2020), https://suiter.com/first-u-s-patent-issued-today-in-1790/.

[74] *See* Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 89 (2014) ("[Algorithms] are capable of changing their behavior to enhance their performance on some task through experience.").

[75] Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J. L. & TECH. 889, 891 (2018).

[76] *See* Alice Corp. Pty. v. CLS Bank Int'l, 573 U.S. 208, 212 (2014) (holding claims drawn to the abstract idea of intermediated settlement that merely require generic computer implementation are patent-ineligible subject matter).

[77] 35 U.S.C. § 101 ("Whoever invents or discovers any new *and useful* process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.") (emphasis added).

forward, patent offices may seek to solve this Gordian knot by lowering enablement requirements for certain applications of "black box" AIs that support specific functional applications.

Even where inventors do understand how the AI they invented works, questions remain about what such inventors must disclose to both prove to others that they know how their invention works and to enable others to practice their invention—i.e., to meet Section 112's written description and enablement requirements (or for that matter Article 83 of the European Patent Convention's requirements regarding disclosure of an invention). After all, "[c]ompliance with the written description requirement is essentially a fact-based inquiry that will necessarily vary depending on the nature of the invention claimed."[78]

Since AI derives its value from data, inventors could potentially be required to disclose the datasets upon which their AI runs to enable any person skilled in the art to make and use the invention without undue experimentation. Historically, Section 112 has required patentees to disclose underlying products that are critical unique elements in the process of making their inventions. The patentee must make such disclosures when it cannot sufficiently describe how to obtain those underlying products in another way.[79] Further, in analogous data-dependent life science technologies, "[p]atent applications which contain disclosures of nucleotide and/or amino acid sequences must contain, as a separate part of the disclosure, a paper copy disclosing the nucleotide and/or amino acid sequences and associated information" using the USPTO's prescribed format.[80] This requirement has at times resulted in massive amounts of paper being delivered to the USPTO to satisfy applicants' Section 112 requirements.[81] If the USPTO finds that claims related to AI cannot meet Section 112's requirements without access to the AI's datasets, then history tells us the USPTO may well require the disclosure of such datasets.[82]

Beyond the logistical challenge of organizing and disclosing terabytes of data to the USPTO, this path to Section 112 compliance also raises important questions about the treatment of proprietary, confidential and personal data. Most data used by AI are proprietary data.[83] When properly organized, collections of proprietary

---

[78] Enzo Biochem, Inc. v. Gen-Probe, Inc., 323 F.3d 956, 963 (Fed. Cir. 2002) (internal quotations omitted).

[79] *See In re* Argoudelis, 434 F.2d 1390 (C.C.P.A. 1970) (finding the deposit of a microorganism necessary to practice claims drawn to a fermentative method of producing two novel antibiotics into a public depository when the patentee could not otherwise provide adequate description of how to obtain the microorganism sufficient to satisfy Section 112).

[80] 37 C.F.R. 1.821(c) (2019).

[81] *See* Don Levin, *Electronic Filing of Patent Documents at the U.S.P.T.O.*, 88 J. PAT. & TRADEMARK OFF. SOC'Y 727 (2006).

[82] Similarly, the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure creates a broad international framework that allows inventors to sufficiently enable inventions involving microorganisms by depositing a sample of the microorganism with any international depositary authority. *See generally* Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure, Apr. 28, 1977, 32 U.S.T. 1241, T.I.A.S. 9768. Such an international depositary framework could be developed to sufficiently enable AI-related inventions.

[83] *See* WIPO AI REPORT, *supra* note 2, at 141.

data can be extremely valuable.[84] Such collections can include data obtained from third-party data brokers under a duty to keep the data confidential.[85] They also include personal data subject to increasingly rigorous data privacy regimes.[86] Even collections of "anonymized" data may contain personal data, as the same increasing computing power and increasing volume of data empowering the rise in AI also enables the de-anonymization of once-anonymous data into personal data.[87] Due to these concerns, AI owners may face a Hobson's choice: disclose data underlying their AI to obtain a patent and face contractual and regulatory liability, or fail to disclose data underlying their AI and fail to obtain patent coverage.

Finally, one line of Section 112 cases could require an AI owner to disclose a proprietary algorithm even if the owner is not seeking to patent the algorithm. If the owner is seeking to patent either a method that is practiced on the algorithm or a product created by the algorithm, this doctrine may apply. Prior to being subsumed by the Federal Circuit, the United States Court of Customs and Patent Appeals held in *In re Ghiron* that if the practice of a claimed method requires a "particular apparatus" to perform the method, then "it is axiomatic that the application must therefore provide a sufficient disclosure of that apparatus if such is not already available."[88] Similarly, in analyzing the patentability of certain derivatives of clavulanic acid, the same United States Court of Customs and Patent Appeals in *In re Howarth* posited that if certain intermediary products, in this case, the clavulanic acid itself, are required to make an end product or practice a process, those intermediary products must be disclosed to obtain patent protection for the end product or process.[89] While these two cases are old, they are still included in the USPTO's Manual of Patent Examining Procedure and could be used to support an argument that an AI inventor must disclose its algorithm to obtain patent

---

[84] Wolfgang Kerber, *Governance of Data: Exclusive Property vs. Access*, 47 INT'L REV. INTELL. PROP. & COMPETITION L. 759, 759 (2016) ("[D]ata is indeed a new and critical input resource for the digital economy, and the striving of businesses for collecting and analyzing these potentially very valuable data is a rational strategy.").

[85] *See generally* FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 11–13 (2014), https://www.ftc.gov/system/files/documents/reports/data-brokerscall-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (discussing how data brokers obtain consumer information); STEPHEN P. MULLIGAN ET AL., CONG. RSCH. SERV., DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY (2010).

[86] *E.g.*, Regulation 2016/679, of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1 [hereinafter GDPR]. Regulation 2016/679 is more commonly referred to as the GDPR and provides data subjects increased rights to control how their data is used and does not recognize decisions of administrative authorities in non-EU countries requiring the disclosure of personal data unless the decision is based on an international agreement. *Id.* art. 48. *See infra* Section III.G for further discussion on privacy regulation and IP.

[87] *See* ARVIND NARAYANAN & VITALY SHMATIKOV, ROBUST DE-ANONYMIZATION OF LARGE DATASETS (HOW TO BREAK ANONYMITY OF THE NETFLIX PRIZE DATASET) (2008), https://arxiv.org/pdf/cs/0610105.pdf.

[88] *In re* Ghiron, 442 F.2d 985, 991 (C.C.P.A. 1971).

[89] *In re* Howarth, 654 F.2d 103, 105 (C.C.P.A. 1981).

protection for a method practiced on the algorithm or for an end product created using the algorithm.[90]

D. *AI, Trade Secrecy and Access to Data*

The importance of data to AI cannot be overstated. Headlines such as *The Economist*'s "*The World's Most Valuable Resource Is No Longer Oil, But Data*" underscore this fact.[91] Companies with big, clean datasets reap their benefits,[92] while startups scrap and scrape to obtain sufficient data to train their algorithms. The increasing importance of data has prompted new questions about what legal and technical measures can, and should, protect data.

The Federal Circuit has stated that "[d]ata in its ethereal, non-physical form is simply information that does not fall under any of the categories of eligible subject matter under section 101."[93] Put another way, data is not patent-eligible subject matter. Similarly, data itself is not copyrightable subject matter.[94] While courts have found that certain compilations of data can receive copyright protection,[95] such protection "is limited to the particular selection or arrangement. In no event may copyright extend to the facts themselves."[96] Further, compilations must involve a modicum of creativity in order to obtain copyright protection; the more the compilations appear to be mere collections of data instead of organized selections, the more likely courts will find the compilations ineligible for copyright protection.[97] Barring the creation of a new *sui generis* property right for data, the only IP protection available is trade secrecy.

The term "trade secret" has several (largely consistent) definitions under state and federal law in the United States.[98] Taking the federal definition of a "trade

---

[90] MPEP § 2164.01(b) (8th ed. Rev. 8, Aug. 2012); *see also* EPO GUIDELINES, *supra* note 23, pt F, ch. III, § 1.

[91] *The World's Most Valuable Resource Is No Longer Oil, but Data*, THE ECONOMIST (May 6, 2017), https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

[92] *See generally Facebook Artificial Intelligence Blog*, FACEBOOK, https://ai.facebook.com/blog/ (last visited Apr. 5, 2021). *See also* John S., *Benefits and Advantages of Data Cleansing Techniques*, Big Data Made Simple (Feb. 11, 2021), https://bigdata-madesimple.com/benefits-and-advantages-of-data-cleansing-techniques.

[93] Digitech Image Techs. v. Elecs. for Imaging, 758 F.3d 1344, 1350 (Fed. Cir. 2014).

[94] Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 348 (1991) ("The same is true of all facts—scientific, historical, biographical, and news of the day. They may not be copyrighted and are part of the public domain available to every person.") (internal quotations omitted). *See also* 17 U.S.C. § 102(b) (clarifying that concepts, principles and discoveries—i.e., factual data—may never receive copyright protection).

[95] *See Feist Publ'ns, Inc.,* 499 U.S. at 350 ("A factual compilation is eligible for copyright if it features an original selection or arrangement of facts."); CCC Info. Servs., Inc. v. Maclean Hunter Mkt. Reps., Inc., 44 F.3d 61, 65–66 (2d Cir. 1994).

[96] *See Feist Publ'ns, Inc.,* 499 U.S. at 350–51.

[97] *See, e.g.*, Warren Publ'g, Inc. v. Microdos Data Corp., 115 F.3d 1509 (11th Cir. 1997) (finding a comprehensive collection of data about a cable system to be copyright-ineligible subject matter).

[98] Many other countries have similar, country-specific trade secrecy laws. *E.g.*, Directive 2016/943, of the European Parliament and of the Council of 8 June 2016 on the Protection of

secret" as an example, a trade secret can cover "all forms and types of financial, business, scientific, technical, economic, or engineering information . . . whether tangible or intangible, and whether or how stored, compiled, or memorialized physically, electronically, graphically, photographically."[99] However, in order for data to qualify as a trade secret, the owner of the data must take reasonable measures to keep the data secret and the data must derive independent economic value from not being generally known or readily ascertainable by other persons.[100]

Because of these two secrecy-related requirements, one of the biggest challenges facing AI implementers is "[t]he lack of access to data, particularly for public research organizations or smaller players."[101] Since data loses its trade secret protection upon disclosure, companies collecting proprietary data are incentivized to keep it secret and are discouraged from disclosing it to others who can use it for public or private benefit.[102] Further, these two secrecy-related requirements have the effect of denying trade secret protection for certain compilations of data, as any publicly available data would not "derive independent economic value from not being generally known or readily ascertainable."[103] Therefore, even trade secrecy, which explicitly covers certain types of data, cannot consistently provide AI implementers adequate IP protection.

Further, technical measures to prevent data scraping[104] may be equally ineffective at protecting certain types of data. In a recent interlocutory appeal before the United States Court of Appeals for the Ninth Circuit, the court affirmed a district court's grant of a preliminary injunction in favor of a startup data analytics company against LinkedIn. The injunction barred LinkedIn from (i) invoking claims based on various federal and state statutes against the startup's scraping of data posted on LinkedIn's platform and (ii) "putting in place any legal or technical measures with the effect of blocking [the startup's] access to public profiles."[105] While the Ninth Circuit's ruling merely upholds a preliminary injunction and neither resolves the parties' underlying legal dispute nor purports to allow the startup to access LinkedIn's non-public data, it does indicate that courts will allow data aggregators to scrape and use publicly available data. Companies seeking to enter the AI field without large collections of data will likely cheer this decision and any other decisions that open more data to the public. Conversely, companies with large collections of data, such as LinkedIn, may move to make less of the data on their platforms public-facing (thereby creating more data silos). Since data is the lifeblood of AI, fights about access to data are likely to continue and even escalate

---

Undisclosed Know-How and Business Information (Trade Secrets) Against Their Unlawful Acquisition, Use and Disclosure, 2016 O.J. (L 157) 1.

[99] 18 U.S.C. § 1839(3).

[100] *Id.*

[101] WIPO AI REPORT, *supra* note 2, at 141.

[102] *Id.*

[103] 18 U.S.C. § 1839(3).

[104] Data scraping, which is also referred to as web scraping, screen scraping and web data extraction, is a technique to collect large amounts of data from websites and other unstructured data sources. At its most basic form, it includes copying the contents of a website and saving them locally in some structured format—e.g., an excel file.

[105] HiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 992 (9th Cir. 2019).

as the AI field develops. At their extreme, such fights over data—and the commensurate market power it gives rise to—could lead to anti-competitive practices.[106]

E. *AI and Open Source*

An emerging issue at the intersection of AI and IP is the advent of open source data. While the term "open source" is generally associated with software,[107] it can also apply to patents[108] and, more recently, data.[109] In October 2017, the Linux Foundation introduced a family of open source data agreements called the Community Data License Agreements (each, a "CDLA").[110] The CDLAs are modeled on open source software agreements.[111] According to the Linux Foundation, the motivation behind creating and introducing the agreements was "to assemble the critical mass of data for [big data analytics, machine learning and AI technologies] to analyze."[112] Currently, there are two CDLA versions. One version, called the CDLA–Permissive, provides data to recipients but does not place any additional sharing requirements on the recipients.[113] The other version, called the CDLA–Sharing, is not permissive, and requires recipients of data under the license to make any subsequent redistributions of the data under the CDLA.[114] The IBM Data Asset eXchange, which provides access to open source data sets under the CDLA, has launched a variety of projects, including, for instance, the Weather Project, which uses local climate data collected by John F. Kennedy Airport to allow for weather forecasting models.[115] Open source data can promise easier access to large amounts of data, which, for certain technologies reliant on AI, like self-driving cars, may provide the incentive necessary to promote industry buy-in. However, as with open source software licenses, it is likely that even more restrictive open source data licenses, imposing more pervasive and aggressive (including "viral") sharing requirements on proprietary data, will be created and

---

[106] *See infra* Section III.G for further discussion on antitrust concerns raised by data silos.

[107] *E.g.*, *The Open Source Definition*, OPEN SOURCE INITIATIVE (Mar. 22, 2007), https://opensource.org/osd (defining "open source" in the context of software).

[108] *E.g.*, Elon Musk, *All Our Patent Are Belong to You*, TESLA (June 12, 2014), https://www.tesla.com/blog/all-our-patent-are-belong-you.

[109] *Linux Foundation Debuts Community Data License Agreement*, THE LINUX FOUNDATION (Oct. 22, 2017), https://www.linuxfoundation.org/press-release/2017/10/linux-foundation-debuts-community-data-license-agreement.

[110] *See id.*

[111] *See id.*

[112] *Id.*

[113] *Community Data License Agreement – Permissive, Version 1.0*, THE LINUX FOUNDATION, https://cdla.io/permissive-1-0/ (last visited Apr. 5, 2021). While this license does not restrict how recipients may use data, it does impose certain restrictions on recipients including providing proper credit and attribution for any data they ultimately distribute. *Id.* § 3.1(c).

[114] *Community Data License Agreement – Sharing, Version 1.0*, THE LINUX FOUNDATION, https://cdla.io/sharing-1-0/ (last visited Apr. 5, 2021).

[115] *IBM Developer Model Asset Exchange: Weather Forecaster*, GITHUB: IBM, https://github.com/IBM/MAX-Weather-Forecaster (last visited May 3, 2021).

adopted.[116] Moving forward, AI implementers will need to track where they source their data in the same way software engineers must track where they source their base code to avoid "opening"—attaching distribution or sharing obligations to—their valuable, proprietary data.[117]

F. *AI and Persons Having Ordinary Skill in the Art*

As discussed above, AI is a powerful tool. It is already increasing mankind's ability to invent, and it will continue to increase the level of invention as progress is made towards realizing currently theorized superintelligent AI.[118] This increased ability may impact a legal concept found throughout patent law: the hypothetical "Person Having Ordinary Skill in the Art" ("PHOSITA").

The PHOSITA standard is used in determining what constitutes an obvious invention for purposes of Section 103 and, as mentioned above, what constitutes adequate enablement and written description for purposes of Section 112. For Section 103, an invention is obvious "if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a [PHOSITA] to which the claimed invention pertains."[119] This is an objective analysis, and the determination of who constitutes a PHOSITA depends on the field of the invention.

Based on this portion of the patent statute, courts have classified inventions created by choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success—i.e., inventions that are obvious to try—as obvious (and therefore unpatentable).[120] While courts have recognized that an invention can be non-obvious even though it would have been obvious to try (i) "to vary all parameters or try each of numerous possible choices until one possibly arrived at a successful result, where the prior art gave either no indication of which parameters were critical or no direction as to which of many possible choices is likely to be successful" or (ii) "to explore a new technology or general approach that seemed to be a promising field of experimentation, where the prior art gave only general guidance as to the particular form of the claimed invention or how to

---

[116] *See, e.g.*, *GNU General Public License Version 2*, OPEN SOURCE INITIATIVE, https://opensource.org/licenses/GPL-2.0 (last visited Apr. 5, 2021). While "viral" open source licenses may come about, open source can have a positive impact on AI development if its underlying rationales encourage the creation of open developer platforms on which innovators can create new, easily incorporable inventions.

[117] On the topic of open source, it is also worth mentioning that AI algorithms, as software, can be subject to open source software licenses. Therefore, AI implementers should track where they, and their algorithms, source their code.

[118] Superintelligent AI is theorized AI that is able to successfully perform any intellectual task that could be undertaken by the human brain or the hypothetical ability of a machine to far surpass the human brain. WIPO AI REPORT, *supra* note 2, at 19; *see also* NICK BOSTROM, SUPERINTELLIGENCE, PATHS, DANGERS, STRATEGIES (Keith Mansfield ed., 2014).

[119] 35 U.S.C. § 103.

[120] KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 421 (2007).

achieve it,"[121] the continued development of AI's ability to analyze vast quantities of data points and outcomes may lead to a state where the non-obviousness of many inventions is called into question. As AI technology has developed, AIs have gained an increasing ability to process numerous combinations of parameters. Eventually, there may come a time when an AI engine can try so many parameters that one could task an AI with solving many problems with a reasonable expectation of success, resulting in what may be considered an invention made by AI. If so, many inventions could become unpatentable, and any issued patents rendered invalid, as obvious to try.[122]

Moving forward, patent practitioners will likely need to highlight in their patent specifications why the claimed inventions would not be obvious to individuals using AI to solve the relevant problem. By focusing on how the inherent capabilities of AI do not address the problem a claimed invention solves, or do not readily facilitate the creation of the claimed invention, patent practitioners will be better positioned to show their claimed inventions are non-obvious.

G. *AI and Emerging Regulatory and Liability Risks*

While not matters of IP law *per se*, AI innovators must also account for the growing regulatory and liability risks that have the potential to disrupt their business models. The incoming head of the European Commission vowed that "in [her] first 100 days in office, [she] will put forward legislation for a coordinated European approach on the human and ethical implications of Artificial Intelligence."[123] This declaration has been interpreted to mean there will be increased regulation and antitrust scrutiny of AI and AI implementers in the European Union.[124] Adding to this concern is the recent election of Margrethe Vestager for a second term as the European Union's competition commissioner with added powers.[125] Under Commissioner Vestager, the antitrust branch of the European Commission has already fined big data companies such as Alphabet billions of dollars[126] and is currently reviewing the data collection and use practices of other large companies such as Amazon.[127] Along with European regulators, "[s]everal agencies, including the Justice Department and the Federal Trade Commission, have taken aim at tech companies, as have attorneys general" in the

---

[121] MPEP § 2143(I)(E) (9th ed. Rev. 10, June 2019) (citing *In re* Kubin, 561 F.3d 1351 (Fed. Cir. 2009)).

[122] *See* Ryan Abbott, *Everything Is Obvious*, 66 UCLA L. REV. 2 (2018).

[123] Ursula von der Leyen (candidate for President of the European Commission), *Political Guidelines for the Next European Commission 2019-2024* (July 16, 2019), https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf.

[124] *See* Valentina Pop, *No Relief for Big Tech Under New EU Leadership*, WALL ST. J. (Sept. 2, 2019, 1:26 PM), https://www.wsj.com/articles/no-relief-for-big-tech-under-new-eu-leadership-11567428651?mod=article_inline.

[125] *See* Valentina Pop & Sam Schechner, *Amazon Faces New EU Antitrust Charges*, WALL ST. J. (Nov. 10, 2020 8:34 AM), https://www.wsj.com/articles/amazon-faces-new-eu-antitrust-charges-11605003489.

[126] Rochelle Toplensky, *EU Fines Alphabet's Google €1.5bn for Antitrust Violations*, FIN. TIMES (Mar. 20, 2019), https://www.ft.com/content/e26b4ae0-4b00-11e9-8b7f-d49067e0f50d.

[127] Pop & Schechner, *supra* note 125.

United States.[128] Some have always thought that there is tension at the intersection of IP and antitrust enforcement.[129] While "[t]rue conflicts between antitrust and intellectual property rights are relatively rare,"[130] the specter of this tension is beginning to impact the nascent AI field. However, the main goal of antitrust law and IP law is the same: promotion of economic welfare through competition and investment. AI implementers must champion the pro- and dynamic-competition advantages AI offers to help ensure antitrust enforcement does not stymie the promising AI field.

Along with antitrust scrutiny, increasing adoption of AI technologies will also likely lead to increased data privacy concerns. There has been a global shift towards increased data privacy in the past decade.[131] The European Union's General Data Protection Regulation ("GDPR") went into force on May 25, 2018,[132] the California Consumer Privacy Act went into effect on January 1, 2020,[133] and various countries around the world are exploring other data privacy regimes.[134] These regulations impose numerous requirements on data controllers, including data minimization requirements (which include implementing procedures for data subjects to practice their rights to erasure) and strict consent requirements (which obligate data processors to only process data for consented-to purposes).[135] As discussed above, AI derives its value from the datasets it analyzes. In the European Union, whenever an AI implementer processes personal data, the implementer becomes either a data controller or data processor.[136] As such, the AI implementer is subject to the GDPR's harsh penalties for noncompliance.[137] These penalties include fines of up to €20,000,000 or 4% of the violator's total worldwide annual turnover of the preceding financial year, whichever is higher.[138] AI implementers will need to ensure that they stay up-to-date on evolving data privacy regimes throughout the world and abide by the regimes' increasingly rigorous requirements.

---

[128] *Id.*

[129] *See generally* Herbert Hovenkamp, *The Intellectual Property-Antitrust Interface*, *in* 3 ISSUES IN COMPETITION LAW AND POLICY 1979 (ABA Section of Antitrust Law 2008) (providing a historical overview of the tense relationship between antitrust law and intellectual property law in the United States).

[130] *Id.* at 1979.

[131] *See* Orson Lucas, *What's Next: Data Privacy Trends and Insights*, KPMG, https://advisory.kpmg.us/articles/2019/data-privacy-trends-and-insights.html (last visited Mar. 26, 2021).

[132] GDPR, *supra* note 86.

[133] California Consumer Privacy Act of 2018 (CCPA), CAL. CIV. CODE § 1798.198(a) (West 2018).

[134] For example, the United States House Task Force on Artificial Intelligence recently heard testimony from AI-industry experts about how the development of new AI technologies necessitated the creation of a comprehensive federal data privacy law. *The Future of Identity in Financial Services: Threats, Challenges, And Opportunities Before the Task Force on Artificial Intelligence of the Comm. on Fin. Servs.*, 116th Cong. (2019).

[135] *See, e.g.*, GDPR, *supra* note 86, arts. 5, 7, 17.

[136] *Id.* art. 4. Whether an AI implementer is a controller or processor depends on the level of control they have in determining the purposes of the data processing.

[137] *Id.* art. 83.

[138] *Id.* art. 83(5).

Similarly, IP practitioners must account for the risk that AI will infringe others' IP, breach the terms and conditions of certain websites, or otherwise perform acts that lead to civil liability. While it is clear that AI implementers can avoid liability for patent infringement in the first instance by ensuring the AI software they create does not perform a patented method claim,[139] it is not clear what happens when an AI applying machine learning comes to practice all of those steps of such patented method. If a patent owner can prove that the AI has practiced all of the steps of the owner's patented method,[140] then it would be clear that the patent was infringed. But it would not necessarily be clear under current law who is liable for such infringement. This problem is exacerbated in the licensing context, as licensors and licensees typically allocate the risk that the licensed IP infringes others' IP through a series of conditions, covenants, representations, warranties, and indemnification provisions. However, these traditional provisions may not properly allocate liability in AI-related transactions. For example, an AI might not infringe others' IP at the beginning of the license but might come to infringe as it evolves over time. Innovative legal solutions will be required as AI becomes more prominent to ensure risks are allocated between licensors and licensees in a commercially reasonable manner.[141]

IV. BEST PRACTICES IN OBTAINING AND PRESERVING IP PROTECTION FOR AI

As explained in this paper, there are numerous practical issues at the intersection of AI and IP that must be addressed to unlock AI's full potential. Many of these issues call into question the availability of patent protection for AI and the viability of AI-related patents. That said, nearly 340,000 patent families related to AI were granted from 1960 to 2018, and the number of patent applications related

---

[139] *See* Medgraph, Inc. v. Medtronic, Inc., 843 F.3d 942, 948 (Fed. Cir. 2016).

[140] As interpreted by the Supreme Court, Rule 8 of the Federal Rules of Civil Procedure requires that a patent owner's complaint "contain sufficient factual matter, accepted as true, to state a claim to relief that is plausible on its face." Ashcroft v. Iqbal, 556 U.S. 662, 663 (2009). While many AI algorithms are quite stable, certain AI algorithms are so dynamic that even the AI's owner does not know all of the inner workings of the AI. *See also* WIPO AI REPORT, *supra* note 2, at 17, 111–17 (analyzing patent litigation involving AI-related patents and hypothesizing that the low enforcement rate of AI-related patents may be due to the difficulty of proving infringement of AI patents).

[141] In addition to staying abreast of evolving regulatory enforcement in the AI field, AI implementers must also be increasingly cognizant of the criminal and civil liability risks unique to AI. AI is already being employed by fraudsters to carry out entirely new types of crimes. One such crime occurred in March 2019, when "[c]riminals used artificial intelligence-based software to impersonate a chief executive's voice and demand a fraudulent transfer of €220,000." Catherine Stupp, *Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case*, WALL ST. J. (Aug. 30, 2019, 12:52 PM), https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402. It is believed that by feeding snippets of the CEO's voice to AI-based software, the software was able to emulate the CEO's voice down to his slight accent and intonation. *Id.* While criminals using technology for malfeasance is not a social ill readily solved through IP law, *see* Juicy Whip, Inc. v. Orange Bang, Inc., 185 F.3d 1364, 1367 (Fed. Cir. 1999) (discrediting the theory that an invention must have a moral utility to receive patent protection in the United States while stating that intellectual property laws should not displace the police powers of the states), it is a new risk IP lawyers must account for while negotiating transactions (e.g., new representations and warranties about a party's use of AI).

to the AI field rose by a factor of 6.5 between 2011 and 2017.[142] In 2019, the USPTO granted nearly 15,000 patents that mentioned AI or machine learning, almost double the number of such patents granted in 2018.[143] When properly claimed, inventors in the AI space can both receive patent protection and subsequently enforce their patents. And when properly managed, AI data can be preserved as an IP asset, and the AI regulatory landscape can be navigated. Based on current trends in the law, some recommended best practices in managing AI IP include:

- While patent applications should describe the problem to be solved, they should contain claims directed to specific functional applications of AI exhibiting a technical character or implementations of AI motivated by technical considerations, rather than broad claims covering general problems that may be solved by AI. Patent applications should describe a contribution to a technical field, rather than to a non-technical one.
- Patent applications should claim only the inventive aspects of an AI solution. They should not seek to cover previously commoditized algorithms or collections of data and should refrain from using generic terminology (e.g., "support vector machine," "reasoning engine") in claims referring to abstract models or algorithms.
- When seeking to draft a properly enabled disclosure and adequate written description, a patentee should consider whether it is adequate to provide examples of how her AI works (where such examples do not disclose proprietary data), describe attributes of data (e.g., format, fields, ranges, hyperparameters[144]), or provide only a subset of proprietary data versus disclosing all of the patentee's proprietary data.
- If a patentee must disclose the data underlying her invention to provide "reasonable certainty" about the scope of what she is claiming, she should be mindful to not improperly disclose confidential or personal data. A patentee should consider whether confidential or personal data may be redacted or altered prior to disclosure.
- Patent applications should discuss how the functionality of the AI is achieved and explain any interdependence between software and hardware. These discussions must show the patentee has possession of the AI and implemented it as claimed, not just that there is a theoretical possibility that the patentee has an AI that can achieve these functions.
- A patentee should be mindful of global IP trends and differences in patent laws between jurisdictions around the world. A patentee may

[142] WIPO AI REPORT, *supra* note 2, at 38.

[143] *See* Al AuYeung, *Who Is Winning the AI Race?*, IPWATCHDOG (Feb. 1, 2020), https://www.ipwatchdog.com/2020/02/01/winning-ai-race/id=118431/.

[144] Hyperparameters are the aspects of an AI's structure that are selected prior to the beginning of training and do not change in value as the AI is exposed to the training datasets. *See Overview of Hyperparameter Tuning*, GOOGLE CLOUD AI PLATFORM, https://cloud.google.com/ml-engine/docs/hyperparameter-tuning-overview (last updated Apr. 29, 2021).

well need to change her prosecution approach depending on whether she is applying for a patent in China, Europe, the United States, or some other jurisdiction.

- Data owners seeking to keep their data proprietary through trade secrecy must take affirmative measures to keep their data secret and should document those measures as well as their compliance with those measures.
- Data owners should be aware that their "anonymous" data may be de-anonymized, potentially subjecting them to various privacy law regimes, and should take steps to prevent de-anonymization or to avoid disclosing data they know can be readily de-anonymized.
- Participants in the AI field should track court decisions affecting what legal and technical measures are available to protect data.
- Participants in the AI field should track where they, and their algorithms, source their code and data to avoid "opening" their proprietary algorithms and data under emerging open-source licenses. Where one seeks to use open-source code and data, one should comply with the actual terms of the pertinent open-source licenses.
- Participants in the AI field should stay abreast of evolving regulatory enforcement and emerging criminal and civil liability theories in the AI field.
- Until regulators and legislatures make concrete policy decisions regarding whether AIs and other machines can be inventors, applicants should ensure a natural person can ethically claim to be the inventor of an invention produced using AI.

By following the above best practices, as well as staying current on new trends at the intersection of AI and IP, AI innovators will be better able to navigate this important enabler of the fourth industrial revolution.