Missouri Law Review

Volume 89 | Issue 3 Article 5

Summer 2024

Using Intellectual Property to Regulate Artificial Intelligence

Dennis Crouch

Using Intellectual Property to Regulate Artificial Intelligence

*Dennis Crouch**

ABSTRACT

This Article examines the complex relationship between intellectual property ("IP") rights and the regulation of artificial intelligence ("AI"). It advances two primary claims: First, while IP plays a role in guiding innovative behaviors in AI development, it does not serve as an effective mechanism for direct regulation of AI. This claim is based on the observation that IP rights, such as patents and copyrights, are primarily designed to incentivize innovation and protect creative works, while lacking the levers necessary to address the broader societal implications of AI technology. The narrow focus of IP rights on rewarding creators makes them ill-suited for managing the more complex ethical, safety, and societal challenges posed by AI systems. Furthermore, it contends that relying on IP for AI regulation could lead to unintended consequences, such as stifling important research or exacerbating existing power imbalances in the tech industry.

The Article's second primary claim is that the relationship between IP rights and AI regulation can be pernicious, as IP rights may hinder AI regulation and development in several ways. This analysis is done largely through the lens of copyright and trade secrecy. The Article analyzes how copyright law impacts AI development, particularly regarding the use of copyrighted works for training AI models and the protection of AI-generated outputs. The discussion also examines the tension between trade secret protection and the regulatory goals of transparency and explainability in AI systems.

Ultimately, the Article concludes that IP should play a supporting role in AI governance rather than serve as the primary legal and regulatory lever.

---

*C.A. Leedy Professor and Director of the Center for Intellectual Property and Entrepreneurship at the University of Missouri School of Law. I would like to thank the *Missouri Law Review* editors and staff both for their work on this Article and for their support in organizing the highly successful 2024 symposium.

TABLE OF CONTENTS

I. INTRODUCTION

This Article makes two primary claims that relate to governmental regulation of artificial intelligence ("AI"), both of which relate to intellectual property ("IP") rights. First, IP does not serve as an effective mechanism for direct regulation of AI—although, a primary role of IP is guiding innovative behaviors, even within the AI landscape. The second claim is slightly broader and creates some potential tension: the relationship between IP rights and AI regulation may be pernicious, raising concern of ways that IP rights may hinder AI regulation and development. Examples include stringent copyright on training data that substantially alters AI creation; overprotection of AI company trade secrets that undermine the transparency efforts necessary for effective regulatory oversight; nationalistic differences that may create international loopholes in the regulatory net; and the absence of rights associated with AI outputs. Further, IP regimes are unlikely to remain static during this transition period, creating an opportunity to align them more closely with regulatory needs and ethical standards of AI development. This dynamic, transitional period offers a pivotal chance for stakeholders to collaboratively refine IP frameworks, ensuring they not only foster innovation but also enhance transparency, accountability, and global cooperation in AI technologies.

Policymakers and legal scholars are calling for the regulation of AI to address existing and emerging risks such as privacy violations,[1] the perpetuation of bias and discrimination,[2] the lack of transparency and

[1] Karl Manheim & Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J.L. & TECH. 106, 160–61 (2019) ("Two things are missing from that regulatory landscape. First is adequate protection of privacy interests and democratic values. Second is an appreciation of the unique challenges that AI presents.").

[2] Kristin M. Kostick-Quenet et. al., *Mitigating Racial Bias in Machine Learning*, 50 J. L. MED. & ETHICS 92, 96 (2022) ("The US and the European Union (EU) have independently proposed initiatives for regulatory guidelines to ensure diversity, nondiscrimination, fairness and equity in ML from design to execution. Numerous other governments, private companies and institutions and non-governmental organizations have similarly proposed high level standards to improve algorithmic fairness and accuracy, many revolving around improving data quality.").

explainability,[3] job displacement,[4] safety and control issues,[5] the concentration of power in a few large tech companies,[6] national security and international relations conflicts,[7] and sweeping, but hard-to-predict, societal impacts.[8] Folks interested in low-intervention regulatory schemes often favor traditional common law approaches of property, torts, and contracts with the state's role as arbiter of rights serving as a quasi-regulator.[9] In this vein, IP

---

[3] Shlomit Yanisky-Ravid & Sean K. Hallisey, *"Equality and Privacy by Design": A New Model of Artificial Intelligence Data Transparency Via Auditing, Certification, and Safe Harbor Regimes*, 46 FORDHAM URB. L.J. 428, 473 (2019) ("The AI Data Transparency Model is a first step towards ensuring that the data used to train AI systems complies with all relevant regulations and societal expectations, which may otherwise limit the AI's use.").

[4] Kierra Burda Martin, Note, *The Robots Are Coming! And Maybe We Should Let Them: How Increased Use of Artificial Intelligence in the Workforce Could Pave the Way for A Shorter Work Week*, 57 NEW ENG. L. REV. 277, 297 (2023) ("To this extent, government regulation of AI is necessary to minimize job loss without inhibiting innovation.").

[5] Gordon Unzen, Note, *Artificial Intelligence and the Administrative State: Regulating the Government Use of Decision-Making Technology*, 25 MINN. J.L. SCI. & TECH. 209, 240 (2023) (identifying "several high-risk implementations that could impact public rights and safety"); Matthew D. Kohel & Erik J. VanderWeyden, *President Biden Issues Executive Order on AI Technology*, SAUL EWING LLP (Nov. 2, 2023), https://www.saul.com/insights/alert/president-biden-issues-executive-order-ai-technology ("The Order establishes new standards for AI safety and security because this technology has the potential to exacerbate societal harms such as fraud, discrimination, bias, and disinformation.").

[6] Anat Lior, *AI Strict Liability Vis-à-Vis AI Monopolization*, 22 COLUM. SCI. & TECH. L. REV. 90, 104–05 (2020) ("[B]ig AI companies are purchasing smaller AI startups and projects in early stages of their development, and thus eliminating any future potential competition. These acquisitions demonstrate the desire of these companies to accumulate intellectual property, datasets, and highly trained human resources in the field of AI. The effect is that competition is stifled.").

[7] *See* JAMES E. BAKER, THE CENTAUR'S DILEMMA: NATIONAL SECURITY LAW FOR THE COMING AI REGULATION 5 (Brookings Institution Press 2021).

[8] Thomas Giacobbe, *Adapting to Challenges Posed by the Fourth Industrial Revolution: A Regulatory Call to Action Concerning Cybernetic Technology*, 15 WASH. U. JURIS. REV. 141, 145–46 (2022) ("AI is so different and versatile that the only way to regulate it is to see where new issues unfold or dilemmas emerge and then adapt rules best suited to the new domain. Along this basis, establishing foundational guidelines that provide a categorical framework to easily develop regulation is a proactive approach that will allow for a smoother transition for dealing with issues when they arise."). Some of the ideas contemplated by this article are autonomous vehicle product liability, copyright issues, brain chip interactions, and threats to civil liberties. *See generally id.*

[9] Stefan Heiss, *Towards Optimal Liability for Artificial Intelligence: Lessons from the European Union's Proposals of 2020*, 12 HASTINGS SCI. & TECH. L.J. 186, 199–206 (2021) (noting tort law is an ineffective way to regulate AI because the difficulty of proving causation shifts incentives); Shyamkrishna Balganesh, *The Pragmatic Incrementalism of Common Law Intellectual Property*, 63 VAND. L. REV.

laws serve important channeling functions,[10] and they take advantage of the incentives inherent in private property systems,[11] allowing creators and inventors to reap the rewards of their efforts.[12] IP rights naturally align with AI development in several ways, including, but not limited to: copyright on AI inputs and outputs;[13] patents covering AI operations and AI-generated outputs;[14] and trade secrets for proprietary AI systems, algorithms, and training data.[15]

The pace of AI advancement also raises new challenges for IP law, such as questions around authorship and inventorship for AI-generated works,[16] the

1543, 1544–45 (2010) (explaining many intellectual property issues are a mix of common law principles of tort, contract and property combined with pragmatic caselaw applied to specific areas of law); Ian Ayres & Eric Talley, *Solomonic Bargaining: Dividing A Legal Entitlement to Facilitate Coasean Trade*, 104 YALE L.J. 1027, 1047 (1995) (asserting adherents to the Coase theorem state that if property rights are clearly defined, private parties will engage in bargaining to reach efficient outcomes despite who is originally allocated the property rights; as such, the State would play only a very minimal role in regulation).

[10] Lucas S. Osborn, *Intellectual Property Channeling for Digital Works*, 39 CARDOZO L. REV. 1303, 1338 (2018) (noting the basic view of channeling intellectual property—especially as it relates to digital works and copyrights—which is broadly applicable to AI use in patent, copyright, and trade secrets).

[11] Eric E. Johnson, *Intellectual Property and the Incentive Fallacy*, 39 FLA. ST. U. L. REV. 623, 629–30 (2012).

[12] Richard S. Gruner, *Imagination, Invention, and Patent Incentives: The Psychology of Patent Law*, U. ILL. J.L. TECH. & POL'Y 375, 425–26 (2017); *see* Sean M. O'Connor, *Patented Electric Guitar Pickups and the Creation of Modern Music Genres*, 23 GEO. MASON L. REV. 1007 (2016).

[13] Clark D. Asay, *Independent Creation in a World of AI*, 14 FIU L. REV. 201, 205 (2020) ("No AI systems are fully automated; they all require some amount of human involvement (at least for now).").

[14] Hubert Ning, *Is It Fair? Is It Competitive? Is It Human?: Artificial Intelligence and the Extent to Which We Can Patent AI-Assisted Inventions*, 49 J. LEGIS. 421, 437 (2023) (acknowledging the Supreme Court's unwillingness to take up *Thaler* and then proceeding to discuss how AI inventions have several hurdles like evergreening and product hopping similar to pharmaceuticals); *see* Ryan Abbott, *Everything is Obvious*, 66 UCLA L. REV. 2 (2019).

[15] Clark D. Asay, *Artificial Stupidity*, 61 WM. & MARY L. REV. 1187, 1220 (2020) ("[Trade secret law's] broader scope can provide legal protection to parts of AI systems that are outside the patent system's ambit, such as the training data upon which many AI systems rely . . . . [T]rade secrecy's broader scope can more readily encompass many elements of AI systems than patent law.").

[16] Shyamkrishna Balganesh, *Causing Copyright*, 117 COLUM. L. REV. 1, 33–34, 77 (2017) (arguing that a causation element should be added when determining authorship to "identify the human agent responsible for bringing the work into existence"); *but see* Victor M. Palace, *What if Artificial Intelligence Wrote This? Artificial Intelligence and Copyright Law*, 71 FLA. L. REV. 217, 241 (2019) (noting that permitting AI authorship results in legal uncertainty and, thus, concluding that AI creations should immediately enter the public domain); *see also* Thaler v. Hirshfeld,

patentability of abstract AI algorithms,[17] and potential infringement by AI systems trained on copyrighted data.[18] In the past, some have called on IP to serve a more direct regulatory role beyond the primary role of promoting generativity and its distribution.[19] Examples include the following: using IP rights to (1) control dissemination of private information;[20] (2) limit creation

558 F. Supp. 3d 238 (E.D. Va. 2021), *aff'd sub nom.* Thaler v. Vidal, 43 F.4th 1207 (Fed. Cir. 2022).

[17] Anastasia Greenberg, *Protecting Virtual Things: Patentability of Artificial Intelligence Technology for the Internet of Things*, 60 IDEA 328, 334 (2020) (discussing the subject matter eligibility, inventorship, and disclosure issues that algorithms may face in the patenting process).

[18] Jenny Quang, *Does Training AI Violate Copyright Law?*, 36 BERKELEY TECH. L.J. 1407, 1419 (2021) (discussing data training for AI models, fair use doctrine, and the potential safe harbor doctrine for training models based on Baker v. Selden, 101 U.S. 99, 100 (1879)); Brenda M. Simon & Ted Sichelman, *Data-Generating Patents*, 111 NW. U. L. REV. 377 (2017) (discussing potential use of patents to promote transparency in AI systems).

[19] Christopher A. Cotropia & James Gibson, *The Upside of Intellectual Property's Downside*, 57 UCLA L. REV. 921, 981 (2010) ("[S]uppression through privately enforced entitlements may be more efficient than top-down regulation. After all, when we want to promote innovation, we rely on intellectual property law to create a private market in information goods; direct governmental rewards for innovation play a comparatively small role."); W. Nicholson Price II, *Regulating Secrecy*, 91 WASH. L. REV. 1769, 1783–84 (2016) (discussing how IP is a form of regulation that tends to be concentrated in industries that are already heavily regulated and doesn't serve as regulation in isolation); Shubha Ghosh, *Patents and the Regulatory State: Rethinking the Patent Bargain Metaphor After Eldred*, 19 BERKELEY TECH. L.J. 1315, 1368 (2004) (thinking of patents as both regulation and deregulation); Shlomit Yanisky-Ravid & Xiaoqiong (Jackie) Liu, *When Artificial Intelligence Systems Produce Inventions: The 3A Era and an Alternative Model for Patent Law*, 39 CARDOZO L. REV. 2215, 2254 (2018); Bryan Casey & Mark A. Lemley, *You Might Be a Robot*, 105 CORNELL L. REV. 287, 355 (2020); James Grimmelman, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 51–52 (2015); s*ee* U.S. CONST. art. I, § 8, cl. 8 ("To Promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."); Christopher Buccafusco & Christopher Sprigman, *Valuing Intellectual Property: An Experiment*, 96 CORNELL L. REV. 1, 3 (2010) ("IP, perhaps more than any other substantive area of law, is grounded in the rational actor model. According to the economic account of IP, the monopolistic rights granted by copyrights and patents exist to provide economic incentives to creators."); WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 4 (The Belknap Press of Harvard University Press, 2003) ("[I]t is acknowledged that analysis and evaluation of intellectual property law are appropriately conducted within an economic framework that seeks to align that law with the dictates of economic efficiency."); *see generally* Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 WASH. L. REV. 579 (2018).

[20] Harry Surden, *Technological Cost As Law in Intellectual Property*, 27 HARV. J.L. & TECH. 135, 144 (2013) ("IP rights have a particular scope, meaning that IP law does not give private parties the right to control every possible piece of information, nor does it permit private parties to constrain every use of information by others.");

of unlicensed datasets in ways that promote less bias and more equity in those actually used;[21] and (3) provide detailed written descriptions of AI algorithms and processes as part of the patenting quid pro quo,[22] among others.[23] These proposed solutions, however, miss their mark and would likely fail in achieving their intended regulatory goals while also diverting focus away from the true purpose of IP rights. For instance, using IP to control private information would be a blunt and ineffective tool compared to additional targeted privacy regulations. IP rights are typically designed to promote innovation and creativity, not to safeguard personal data.[24] Similarly, using

see also Elizabeth Rosenblatt, *Social Justice and Copyright's Excess*, 6 TEX. A&M J. PROP. L. 5, 11–12 (2020) ("[P]romoting social justice among potential authors also promotes the creation and dissemination of works.").

[21] Levendowski, *supra* note 19, at 579. Copyright law's exclusion of access to certain copyrighted source materials may create or promote biased AI system by limiting bias mitigation techniques. *Id.* Further, the fair use doctrine is examined, illustrating how it has traditionally been used to address similar concerns in other technological fields, and asks whether it is equally capable of addressing them in the field of AI bias. *Id.*; *see* Tobias Thomas et al., *Modelling Dataset Bias In Machine-Learned Theories Of Economic Decision-Making*, 8 NATURE HUM. BEHAV. 679 (2024) (noting dataset bias can skew human decision making towards more risky behavior and socially undesirable public facing bias in datasets that are created).

[22] Mehdi Poursoltani, *Disclosing AI Inventions*, 29 TEX. INTELL. PROP. L.J. 41, 54 (2021) (suggesting that AI inventions involving the algorithms, models, or end products would likely require heightened disclosure, compared to software patents, to meet the requirements of Sections 101 and 112, but also recognizing uncertainty in the current system as to how the USPTO might handle disclosure requirements).

[23] Lateef Mtima, *An Introduction to Intellectual Property Social Justice and Entrepreneurship: Civil Rights and Economic Empowerment for the 21st Century*, *in* INTELLECTUAL PROPERTY, ENTREPRENEURSHIP, AND SOCIAL JUSTICE 5, 8 (Lateef Mtima ed., 2015) (invoking the social engineering theories of Charles Hamilton Houston to construct "IP Empowerment" as a social movement which applies IP Social Justice theory to promote grassroots IP education and social entrepreneurship and thereby delineate a 21st century Civil Rights Economic Agenda).

[24] WILLIAM FISHER, *Theories of Intellectual Property*, *in* NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY 168 (Stephen R. Munzer ed., 2001) (explaining utilitarian incentive theory of IP); Graham v. Prince, 265 F. Supp. 3d 366, 376 (S.D.N.Y. 2017) ("As embodied in the United States Constitution, the purpose of copyright is '[t]o promote the Progress of Science and useful Arts.'"); Ned Snow, *Science, Creativity, and the Copyright Clause*, 74 HASTINGS L.J. 1121, 1123 (2023) (noting that the conventional wisdom focuses on creativity as the purpose of IP law, but arguing that "knowledge and learning" lie at the root of the Constitutional IP clause); Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1597 (2003) ("[C]ourts and commentators widely agree that the basic purpose of patent law is utilitarian: We grant patents in order to encourage invention."); Peter Lee, *Patent Law's Externality Asymmetry*, 43 CARDOZO L. REV. 1923, 1977 (2022) ("Patents exist not to reward individual inventors but to advance society-wide technological progress."); Tzipi Zipper, *Mind over Matter: Addressing Challenges of Computer-Generated Works Under Copyright Law*, 22 WAKE FOREST J. BUS. & INTELL. PROP. L. 129, 135 (2022) ("The purpose behind patent law is to incentivize

IP to limit dataset creation and use could unduly restrict AI development and hinder beneficial applications, all while relying upon private enforcement of property rights to support an important issue of global development.[25]

These IP-based regulatory approaches could create perverse incentives and unintended consequences. For example, companies might avoid patenting AI inventions altogether to escape disclosure requirements, relying instead on trade secrecy.[26] This would reduce rather than promote transparency around AI development. Channeling these analyses, Professors Lemley and Casey explained in the copyright context that "[w]hile we share some of the concerns about the uses to which [machine learning] systems may be put, copyright is not the right tool to regulate those abuses."[27]

Following the introduction, this Article unfolds in five key parts. Part II explores the calls for AI regulation, driven by concerns over privacy, bias, transparency, job displacement, and power concentration within the tech industry. Part III then examines the dual role of IP rights in both supporting AI innovation through incentive structures and acting as a quasi-regulatory mechanism, potentially mitigating some of the aforementioned concerns. The discussion progresses in Part IV to highlight the limitations and challenges of relying solely on IP rights for effective AI regulation, emphasizing the misalignments between IP's innovation incentives and broader societal and ethical objectives. Part V follows, critically examining how IP rights, particularly through copyright and trade secret protections, may pose barriers to AI regulation, stressing the conflict between proprietary interests and the necessity for transparency and open innovation. Finally, Part VI concludes with a synthesis of insights, advocating for a collaborative refinement of IP frameworks to foster innovation and enhance responsible AI development that aligns with ethical standards.

---

the disclosure of information, commercialization, and development of inventions, and that recognition of machines as inventors would facilitate the protection of moral rights of human inventors."); Kelvin W. Willoughby, *How Much Does Technology Really Matter in Patent Law? A Comparative Analysis of Doctrines of Appropriate Patentable Subject Matter in American and European Patent Law*, 18 FED. CIR. BAR J. 63, 81 (2008) (explaining that the purpose behind Chinese patent law is the same as other global nations—promoting "invention-creation").

[25] *See infra* Parts III and IV.

[26] *See infra* Part IV.

[27] Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 749 (2021); Tom Simonite, *The Generative AI Battle Has a Fundamental Flaw*, WIRED (July 25, 2023, 7:00 AM), https://www.wired.com/story/artificial-intelligence-copyright-law/ [https://perma.cc/6TD8-YRBP] (noting copyright law, particularly the concept of fair use, is not well-equipped to address the full scope of artists' concerns about AI, such as employment, compensation, privacy, and use of personal characteristics).

II. BACKGROUND ON AI AND DEMANDS FOR AI REGULATION

A. Introduction to AI and Machine Learning

Over the past several years, AI has rapidly advanced. This advancement has transformed various industries and fields, including the creative economy.[28] AI is broadly defined as "an algorithm or machine capable of completing tasks that would otherwise require cognition."[29] Machine learning ("ML"), an overlapping field with AI, has evolved into a predominant technique that empowers AI systems to acquire knowledge from data, discern patterns, and arrive at decisions with minimal human oversight.[30] By leveraging algorithms that can learn and improve from experience, ML enables AI models to adapt and perform tasks autonomously without the need for explicit programming of every potential scenario.[31]

ML algorithms, particularly those used in deep learning systems, can self-improve and automate their own refinement.[32] During the training process, these algorithms are fed large amounts of data, allowing them to adjust their internal parameters, known as weights, to improve their output's match with desired results.[33] ML's ability to learn and adapt, without much intervention, is a key element of these AI models.

As further discussed later in this Article, recent advancements and improvements in ML systems have enhanced AI functionality, driven by improved software design, greater hardware capabilities, and increased access to data used for training.[34] These advancements have led to the development of groundbreaking systems like AlphaGo and AlphaFold, which have mastered complex tasks such as playing the game of Go and determining a protein's 3-D shape from its amino acid sequence, respectively.[35]

---

[28] The creative economy encompasses a wide range of industries and activities that are based on the generation and use of innovative ideas and creative content. *The Creative Economy*, THE POLICY CIRCLE, https://www.thepolicycircle.org/minibrief/the-creative-economy/ [https://perma.cc/ZKA2-BZRY] (last visited Sept. 14, 2024). This includes sectors such as art, design, music, literature, film, television, gaming, software development, scientific research, engineering and technological creation. *Id.* The creative economy involves the production and distribution of goods and services that rely on creativity, skill, and talent, often resulting in the creation of copyrightable works and patentable innovations. *Id.*

[29] Ryan Abbott & Elizabeth Rothman, *Disrupting Creativity: Copyright Law in the Age of Generative Artificial Intelligence*, 75 FLA. L. REV. 1141, 1146 (2023).

[30] *Id*. at 1147.

[31] *Id*.

[32] *Id*.

[33] *Id*.

[34] *See infra* Parts II.A.1–4.

[35] AlphaGo and AlphaFold are two groundbreaking AI systems developed by DeepMind, a subsidiary of Alphabet Inc. (Google's parent company). *AlphaGo*, GOOGLE DEEPMIND, https://deepmind.google/technologies/alphago/ [https://perma.cc/N59Y-DPGM] (last visited July 7, 2024); *see also AlphaFold*, GOOGLE DEEPMIND,

1. Generative AI and Large Language Models

One of the most significant developments in AI has been the emergence of generative AI and large language models ("LLMs"). This subpart highlights these systems and the developments made by OpenAI and Anthropic in creating generative pre-trained transformer-based ("GPT") models and AI assistants like ChatGPT and Claude, while also discussing the potential implications and the novel training approach of constitutional AI.

Generative AI and LLMs are an important subset of ML systems. Generative AI refers to AI systems capable of creating new content, such as text, images, and audio based on the patterns and characteristics learned from training data.[36] LLMs are a class of deep-learning architectures that can generate content after training on large, text-based datasets.[37]

OpenAI, a leading AI company, has been at the forefront of developing generative AI and LLMs.[38] Its GPT series, including GPT-3 and GPT-4, has showcased the potential of these LLM models in generating human-like text.[39] Released in early 2023, GPT-4 has demonstrated remarkable capabilities in understanding and generating natural language and image analysis, which is

---

https://deepmind.google/technologies/alphafold/ [https://perma.cc/7CG3-XJVV] (last visited July 7, 2024). AlphaGo was the first AI system to defeat a world champion in the ancient Chinese game of Go, which was previously considered one of the most challenging games for AI due to its vast search space and intuitive elements. *See* Elizabeth Gibney, *Google AI Algorithm Masters Ancient Game Of Go*, NATURE (Jan. 28, 2016), https://www.nature.com/articles/529445a [https://perma.cc/JL3G-VUUZ]. It used deep neural networks and reinforcement learning to master the game at a superhuman level. *Id*. AlphaFold is an AI system designed to predict the 3D structure of proteins from their amino acid sequences with high accuracy. *AlphaFold*, *supra* note 35. Determining protein structures is crucial for understanding their functions and developing new drugs and treatments. *Id*. AlphaFold achieved a breakthrough in this long-standing challenge in computational biology. *Id*.

[36] Michael Chui et al., *The Economic Potential of Generative AI: The Next Productivity Frontier*, MCKINSEY DIGITAL (June 14, 2023), https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction [https://perma.cc/T27W-FZW8].

[37] Abbott & Rothman, *supra* note 29, at 1148.

[38] OpenAI is an artificial intelligence research laboratory consisting of the for-profit corporation OpenAI LP and its parent company, the non-profit OpenAI Inc. Greg Brockman et. al., *OpenAI LP*, OPENAI (Mar. 11, 2019), https://openai.com/index/openai-lp/ [https://perma.cc/J7VH-SBYL]. The company was founded in 2015 by Elon Musk and Sam Altman, among others, with the goal of promoting and developing friendly AI in a way that benefits humanity as a whole. Greg Brockman et al., *OpenAI and Elon Musk*, OPEN AI (Mar. 5, 2024), https://openai.com/index/openai-elon-musk// [https://perma.cc/QYA9-FSRD].

[39] *GPT-4 Is OpenAI's Most Advanced System, Producing Safer And More Useful Responses*, OPENAI, https://openai.com/index/gpt-4/ [https://perma.cc/ET9S-XYUB] (last visited June 29, 2024); Austin G. Ward, *From Telegraphs to GPT-4*, ADVOCATE (Feb. 2024), https://www.advocatemagazine.com/article/2024-february/from-telegraphs-to-gpt-4 [https://perma.cc/5G69-V8NU].

also illustrated via their chatbot known as ChatGPT.[40] ChatGPT is based on the GPT architecture and has been trained on vast amounts of textual data (mostly scraped from the Internet), allowing the chatbot to engage in conversational interactions and provide informative responses to a wide range of queries.[41] The release of ChatGPT sparked widespread interest and discussion about the potential applications and implications of generative AI.[42]

Anthropic, another AI research company, developed Claude, an AI assistant that further showcases the capabilities of generative AI.[43] According to its self-statements, Claude was uniquely trained using constitutional AI, a novel approach that aims to align AI systems with human values and goals.[44] This approach involves training AI models with explicit rules and guidelines to better ensure beneficial outputs to humanity.[45]

2. Training AI Models: Supervised and Unsupervised Learning

The training process is crucial to the development of AI models, and there are two main approaches: supervised learning and unsupervised learning. Supervised learning involves training an AI model with labeled data, where the desired output is provided alongside the input data.[46] This model learns to map input features to the corresponding output labels, enabling it to make predictions or classifications on newly provided data.[47] The supervised learning approach helps AI models quickly learn how to make outputs based on relationships already understood by humans.[48] As one might guess, unsupervised learning involves training an AI model on unlabeled data, where the model must identify patterns and structures in the data without explicit guidance.[49] This approach, alternatively, is typically used to find relationships, outputs, or solutions to data sets and problems not yet completely understood by experts.[50]

---

[40] *See generally Introducing ChatGPT*, ONLINE CHATGPT (Nov. 30, 2022), https://online-chatgpt.com [https://perma.cc/F4CM-2W2P].

[41] *Id.*

[42] *What is Generative AI*?, MCKINSEY & CO. (Apr. 2, 2023), https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai [https://perma.cc/YY2T-NZDA].

[43] *Introducing Claude*, ANTHROPIC (Mar. 14, 2024), https://www.anthropic.com/index/introducing-claude [https://perma.cc/63NK-3KMW].

[44] *Claude's Constitution*, ANTHROPIC (May 9, 2023), https://www.anthropic.com/news/claudes-constitution [https://perma.cc/CP8R-7SJW].

[45] *Id.*

[46] *What is Supervised Learning?*, IBM, https://www.ibm.com/cloud/learn/supervised-learning [https://perma.cc/4N3V-T59Z] (last visited June 29, 2024).

[47] *Id.*

[48] *Id.*

[49] *Unsupervised Learning*, IBM, https://www.ibm.com/cloud/learn/unsupervised-learning [https://perma.cc/NYW4-VFRQ] (last visited July 6, 2024).

[50] *Id.*

In the context of generative AI, both supervised and unsupervised learning techniques are employed. For example, GPT models are typically pre-trained using unsupervised learning on large amounts of unlabeled textual data, allowing them to capture the statistical properties and patterns of natural language.[51] Subsequently, these pre-trained models can be fine-tuned for specific tasks using supervised learning, where labeled data is used to guide the model towards the desired output.[52]

3. Foundational Models and Their Impact

Foundational models have emerged as a key development in AI.[53] Massive amounts of data—effectively the entire Internet—train foundational models, enabling them to adapt to and perform many different of tasks with minimal fine-tuning.[54] GPT and Bidirectional Encoder Representations from Transformers ("BERT") are two common examples of foundational AI models.[55] Foundational models have the potential to transform the marketplace by facilitating the development of further application layers atop the foundational model.[56] In short, foundational AI models serve as a foundation or launchpad for exponential AI development.

However, the widespread adoption of centralized foundational models also raises important ethical and societal questions, such as bias, fairness, and the potential for misuse. In their current technological state, both the training and deployment of foundational AI models require extensive computational resources and infrastructure.[57] The massive scale of these models, often containing trillions of parameters, necessitates the use of large computer networks and distributed computing systems.[58] As a result, the deployment of foundational model applications on individual devices typically involves a centralized system architecture, somewhat similar to a traditional client-server

---

[51] Tom B. Brown et al., *Language Models are Few-Shot Learners*, ARXIV (May 28, 2020), https://arxiv.org/abs/2005.14165 [https://perma.cc/UMB2-YVJ4].

[52] *Id.*

[53] *See* Rishi Bommasani et al., *On the Opportunities and Risks of Foundation Models*, arXiv (Aug. 16, 2021), https://arxiv.org/abs/2108.07258 [https://perma.cc/8D9Z-DG4S].

[54] *Id.* at 3.

[55] *Id.* at 11.

[56] *Id.*; Priyanka Somrahe, *Foundation Models & the Path to Enterprise Adoption*, THE DATA SOURCE (Feb. 15, 2023), https://thedatasource.substack.com/p/the-data-source-12-foundation-models [https://perma.cc/B9R9-8D7F].

[57] Bommasani et al., *supra* note 53.

[58] *Id.*; Will Knight, *OpenAI's CEO Says the Age of Giant AI Models Is Already Over*, WIRED (April 18, 2023, 7:00 AM) (stating the cost of training GPT-4 was reportedly more than $100 million); Jacob Stern, *The OpenAI Debate: GPT-4's Parameters and Power*, THE ATLANTIC (Mar. 6, 2023), https://www.theatlantic.com/technology/archive/2023/03/openai-gpt-4-parameters-power-debate/673290/ [https://perma.cc/PG25-V8RT] (stating that there are over 100 trillion parameters in GPT-4).

model.[59] The foundational model is hosted on powerful servers or cloud computing platforms, while client devices—smartphones or personal computers—interact with the model through application programming interface ("API") calls or web interfaces.[60] This centralized approach allows for efficient utilization of computational resources while enabling seamless updates and improvements to the foundational model without requiring end-users to download or maintain the model locally.[61]

Simultaneously, there has been a rise in edge computing, which relies upon comparatively more distributed network architecture.[62] Edge computing refers to the deployment of computational resources and data processing capabilities closer to the source of data generation, rather than relying solely on a centralized cloud or data center.[63] In the context of AI applications, edge computing enables AI models and algorithms to run directly on edge devices or local computing nodes, allowing for improved real-time data processing and decision-making without the need for constant communication with a remote cloud infrastructure.[64] This reduces latency and typically improves efficiency.[65] Apart from the improved performance, however, the user perspective is the same—the AI models are deployed at another network node and are effectively centralized.[66]

4. The Hardware Powering AI: GPUs and Beyond

The rapid advancements in AI and ML have not only been driven by innovations in algorithms and software but also by significant hardware developments. The computational demands of training and running complex AI models have necessitated the use of specialized hardware components, particularly graphics processing units ("GPUs").[67]

---

[59] MEENU MARY JOHN ET AL., *Architecting AI Deployment: A Systematic Review of State-of-the-Art and State-of-Practice Literature*, *in* ICSOB, SOFTWARE BUSINESS: 11TH INTERNATIONAL CONFERENCE 1 (Erik Klotins & Krzysztof Wnuk eds., 2020).

[60] *See* Bommasani et al., *supra* note 53, at 11.

[61] *But see* Yuqing Tian et al., *An Edge-Cloud Collaboration Framework for Generative AI Service Provision with Synergetic Big Cloud Model and Small Edge Models*, ARXIV (Jan. 3, 2024), https://arxiv.org/abs/2401.01666 [https://perma.cc/E5RQ-BV2E] (arguing in favor of decentralizing services by moving them partly from the cloud to the edge to enable private, timely, and personalized experiences).

[62] Guoping Rong et al., *An Edge-Cloud Collaborative Computing Platform For Building AIoT Applications Efficiently*, 10 J. CLOUD COMPUTING 36, 1 (2021); JOHN, *supra* note 59, at 1.

[63] Rong et al., *supra* note 62, at 1–2; JOHN, *supra* note 59, at 1.

[64] Rong et al., *supra* note 62, at 2.

[65] *Id.*; JOHN, *supra* note 59, at 1.

[66] Rong et al., *supra* note 62, at 12.

[67] Sparsh Mittal & Jeffrey S. Vetter, *A Survey of Methods for Analyzing and Improving GPU Energy Efficiency*, ARXIV (Apr. 17 2014), https://arxiv.org/abs/1404.4629 [https://perma.cc/Q2DT-8H46].

GPUs were originally designed for rendering graphics and video and are highly efficient at parallel processing tasks, which form the core of many AI and ML computations.[68] The massive parallel architecture of GPUs allows for the simultaneous execution of thousands of threads,[69] enabling faster training and inference of AI models compared to traditional central processing units ("CPUs").[70]

The adoption of GPUs for AI workloads has been transformative, with major tech companies like NVIDIA and AMD developing specialized GPUs tailored for AI and ML.[71] These GPUs offer high performance, energy efficiency, and scalability, enabling researchers and developers to train larger and more complex models in shorter timeframes.[72] In addition to GPUs, other hardware innovations have emerged to support the growing demands of AI. Google's tensor processing units ("TPUs"), for example, are custom-built chips designed specifically for ML workloads.[73] TPUs offer high performance and efficiency for tasks such as matrix multiplication and convolution, which are fundamental operations in deep learning models.[74]

The rise of edge computing has influenced the hardware landscape for AI.[75] Many technology researchers are developing AI accelerators and system-on-chip solutions that can perform AI computations on resource-constrained devices, such as smartphones, internet-of-things devices, and embedded systems.[76] Moreover, the increasing scale and complexity of AI models have driven the need for distributed computing infrastructure. High-performance computing clusters and cloud computing platforms are essential for training and deploying large-scale AI models.[77]

---

[68] GPUs are specialized electronic circuits designed to rapidly manipulate and alter memory to accelerate the creation of images in a frame buffer intended for output to a display device. *Id.* Their highly parallel structure makes them more efficient than general-purpose CPUs for algorithms that process large blocks of data in parallel. *Id.*

[69] In computing, a thread is a unit of execution that is a component of a process. *Id.* Multiple threads can exist within one process, executing concurrently and sharing resources such as memory. *Id.* In the context of GPUs, many threads can be executed simultaneously, allowing for significant performance improvements in tasks that can be parallelized, such as machine learning computations. *Id.*

[70] *Id.*

[71] *NVIDIA A100 Tensor Core GPU*, NVIDIA, https://www.nvidia.com/en-us/data-center/a100/ [https://perma.cc/8P3V-QZ92] (last visited July 6, 2024).

[72] *Id.*

[73] Kaz Sato, *What Makes TPUs Fine-Tuned for Deep Learning?*, GOOGLE CLOUD BLOG (Aug. 30, 2018), https://cloud.google.com/blog/products/ai-machine-learning/what-makes-tpus-fine-tuned-for-deep-learning [https://perma.cc/UWX7-H3TJ].

[74] *Id.*

[75] Weisong Shi et al., *Edge Computing: Vision and Challenges*, 3 IEEE INTERNET THINGS J. 637, 637–38 (2016).

[76] *See id.* at 638.

[77] MOHAN KUMAR K.M. ET AL., *Comprehensive Survey on High Performance Computing: Technologies, Applications and Challenges*, *in* HIGH PERFORMANCE

The increasing reliance on specialized hardware for AI also raises concerns about energy consumption, environmental impact, and national security. Training and deploying large-scale AI models requires significant computational resources, leading to high energy usage and carbon footprint.[78] Efforts are underway to develop more energy-efficient hardware solutions and optimize AI algorithms to reduce their environmental impact.[79]

As part of nationalism concerns, especially current relations with China, the United States government has taken several steps to regulate and restrict the global distribution of AI tools.[80] An effort taken to address these matters is highlighted in a White House issued Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence in October 2023.[81] This order established a White House AI Council to coordinate AI-related policies across federal agencies.[82] The order directs the Secretary of Homeland Security, in coordination with the Secretary of State, to develop a plan for multilateral engagements to encourage the adoption of AI safety and security guidelines for use in critical infrastructure globally.[83] It also requires the Secretary of Homeland Security to submit a report on priority actions to mitigate cross-border AI risks.[84]

Simultaneously, the United States tightened restrictions on the export of advanced semiconductors and chipmaking equipment to China, particularly those used for AI applications.[85] According to the United States, limiting China's access to AI technologies is necessary for national security, as AI could enhance China's military capabilities or create new, existential threats

---

COMPUTING FOR COMPUTATIONAL INTELLIGENCE 3, 4–5 (Dipankar Deb et al. eds., 2020).

[78] Emma Strubell et al., *Energy and Policy Considerations for Deep Learning in NLP*, ARXIV (June 5, 2019), https://arxiv.org/abs/1906.02243 [https://perma.cc/NA7A-WNFH].

[79] *See, e.g.*, Bill Dally, *Energy Efficiency and AI Hardware Keynote Presentation at Stanford AHA Retreat,* STANDARD AHA RETREAT (Aug. 31, 2023), https://aha.stanford.edu/sites/g/files/sbiybj20066/files/media/file/aha-retreat-2023_dally_keynote_en_eff_ai_hw_0.pdf [https://perma.cc/ZER4-VCJT].

[80] Ana Swanson, *U.S. Tightens China's Access to Advanced Chips for Artificial Intelligence*, N.Y TIMES (Oct. 17, 2023), www.nytimes.com/2023/10/17/business/economy/ai-chips-china-restrictions.html [https://perma.cc/4ZT9-KXAX]; Cecilia Kang, *A.I. Leaders Press Advantage With Congress as China Tensions Rise,* N.Y. TIMES. (Mar. 27, 2024), www.nytimes.com/2024/03/27/technology/ai-lobby-china.html [https://perma.cc/GE8F-JRZ7]; *see* Exec. Order No. 14110, 88 Fed. Reg. 75191 (Nov. 1, 2023); Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use; Updates and Corrections, 88 Fed. Reg. 73458 (Oct. 25, 2023).

[81] Exec. Order No. 14110, 88 Fed. Reg. 75191 (Oct. 30, 2023).

[82] *Id.*

[83] *Id.*

[84] *Id.*

[85] Swanson, *supra* note 80.

if not properly managed.[86] The new rules require United States companies to obtain special licenses to export certain advanced chips to China or notify the government of their plans to do so.[87] These restrictions are expected to significantly impact Chinese companies developing AI chatbots and other applications, as well as United States chipmakers that derive substantial revenue from Chinese buyers, like NVIDIA, AMD, and Intel.[88] While government officials frame these restrictions in terms of national security, they also intentionally help maintain the United States' lead in the AI race and prevent China from gaining too much ground in this critical emerging technology.[89] The tensions highlight ongoing strategic competition between the United States and China over AI supremacy, which is seen as vital by many for both economic and military dominance in the twenty-first century.[90]

B. Key Areas of Concern Driving Calls for AI Regulation

The rapid advance of AI has created a multitude of concerns regarding potentially negative impacts on society, even for those who are largely techno-optimists.[91] These concerns span across various domains, including privacy, bias and discrimination, job displacement, and the existential risks posed by advanced AI systems.[92] As AI continues to permeate our daily lives and decision-making processes, it becomes increasingly crucial to address these issues through appropriate regulation.[93] At the same time, the regulatory environment is critical to help foster development toward the greatest positive impacts. It is especially important because AI offers a high risk, high reward situation in terms of significant societal upside and downside.

---

[86] *Id.*

[87] *See* Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use; Updates and Corrections, 88 Fed. Reg. 73458 (Oct. 25, 2023).

[88] *Id.*

[89] *Id.*

[90] *Id.*

[91] Techno-optimism is the belief that technological progress will ultimately have beneficial effects for humanity and help solve at least some of our most difficult challenges. *See* Noah Smith, *Thoughts on Techno-Optimism,* NOAHPINION (Oct. 20, 2023), https://www.noahpinion.blog/p/thoughts-on-techno-optimism [https://perma.cc/2VFN-VM4M]. Proponents of techno-optimism argue that advances like artificial intelligence, though disruptive in the short-term, will eventually lead to increased productivity, economic growth, and an improved quality of life. *See id.*; *contra* Marc Andreessen, *The Techno-Optimist Manifesto*, ANDREESSEN HOROWITZ (Oct. 16, 2023), https://a16z.com/the-techno-optimist-manifesto/ [https://perma.cc/H9GB-GZPR].

[92] Luciano Floridi & Josh Cowls, *A Unified Framework of Five Principles for AI in Society*, 1.1 HARV. DATA SCI. REV. 1, 7 (2019).

[93] Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 399–405 (2017).

The rapidly evolving nature of AI presents a significant challenge for policymakers and regulators. The technology is advancing at an unprecedented rate, making it difficult to predict its future trajectory and potential consequences.[94] This uncertainty necessitates a proactive and adaptable approach to AI regulation, one that can keep pace with the ever-changing landscape of AI development.[95] The following subparts delve deeper into key areas of concern in critical focal points for AI regulation.[96] However, it is important to keep in mind that the list of concerns driving calls for AI regulation is non-exhaustive; new issues will almost certainly emerge as both technology and society evolve.[97]

1. Privacy Violations and Data Misuse

The concept of privacy as a legal right has developed significantly over the past century. The recent advancement of communications technology and AI presents new challenges and concerns in defining and regulating privacy rights. To better understand the current landscape of privacy law in the context of AI, it is helpful to examine the historical foundation and key legal precedents that have shaped our understanding of privacy rights.

One of the most influential works in the development of privacy law is the seminal article *The Right to Privacy*, written by Samuel Warren and Louis Brandeis and published by Harvard Law Review in 1890.[98] Warren and Brandeis argued for the recognition of privacy as a distinct legal right, separate from existing protections for property and reputation.[99] They defined privacy as the "right to be let alone" and emphasized the importance of protecting individuals from the unwanted intrusion and public disclosure of private matters.[100] This foundational discussion laid the groundwork for modern understandings of informational and decisional privacy, which are particularly relevant in the context of AI.

Building upon this work by Warren and Brandeis, William Prosser further developed the concept of privacy law in his 1960 article, *Privacy*.[101] Prosser categorized privacy torts into four distinct categories: intrusion upon seclusion, public disclosure of private facts, false light, and appropriation of

---

[94] Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 366 (2016).

[95] Wendell Wallach & Gary Marchant, *Toward the Agile and Comprehensive International Governance of AI and Robotics*, 107 PROCEEDINGS OF THE IEEE 505, 505–08 (2019).

[96] *See infra* Part IV.B.1.

[97] Gonenc Gurkaynak et al., *Stifling Artificial Intelligence: Human Perils*, 32 COMPUT. L, & SEC. REV. 749, 750 (2016).

[98] Samuel D. Warren & Louis D. Brandei, *The Right to Privacy*, 4 HARV. L. REV. 193, 195 (1890).

[99] *Id.*

[100] *Id.*

[101] William L. Prosser, *Privacy*, 48 CAL. L. REV. 383, 389 (1960).

name or likeness.[102] These categories are still emphasized in United States privacy law today and provide a framework for understanding how AI can infringe on privacy rights through the collection, processing, and use of personal data.[103]

As technology has advanced, courts have grappled with applying the traditional privacy principles. In *United States v. Jones*,[104] the United States Supreme Court addressed privacy concerns with regard to surveillance and the collection of personal information through GPS tracking. The Court held that the installation of a GPS tracking device on a vehicle constituted a search under the Fourth Amendment, recognizing the potential for technology to enable pervasive monitoring and intrude upon individual privacy.[105] Justice Sotomayor's concurrence warned that "GPS monitoring generates a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations."[106] While Jones dealt specifically with GPS tracking, the case foreshadowed future challenges posed by advanced AI-powered surveillance technologies. Modern AI systems go beyond simply holding data. They can also analyze the data to create exceptionally detailed profiles about individuals' behaviors, associations, and lives. This raises significant concerns about decisional privacy—the ability to make choices about fundamental matters like relationships, employment, and medical care without undue scrutiny or influence. As AI surveillance capabilities grow more powerful, the Court's caution in *Jones* about pervasive monitoring takes on heightened relevance.

These foundational works and legal precedents provide a framework for understanding the evolution of privacy law and the challenges that arise with the development of AI technologies. As AI systems become increasingly sophisticated and ubiquitous, they can collect, process, and analyze vast amounts of sensitive information, often without the knowledge or explicit consent of individuals.[107] The following subparts will explore the key privacy and data misuse issues surrounding AI, drawing upon principles established in these historical works and recent legal developments. This Article sets forth several examples of how AI can cause privacy violations and data misuse, which has led to calls for AI regulation.

---

[102] *Id.*

[103] RESTATEMENT (SECOND) OF TORTS § 652A (AM. L. INST. 1997).

[104] United States v. Jones, 565 U.S. 400, 402 (2012).

[105] *Id.* at 404.

[106] *Id.* at 415 (Sotomayor, J., concurring).

[107] Mark van Rijmenam, *Privacy in the Age of AI: Risks, Challenges and Solutions*, THE DIGITAL SPEAKER (Feb. 17, 2023), https://www.thedigitalspeaker.com/privacy-age-ai-risks-challenges-solutions/ [https://perma.cc/76MV-MAK6].

a. Data Collection and Consent

A primary concern of AI and privacy is the extensive collection of personal data required to train and operate these systems. AI relies heavily on large datasets to learn patterns, make predictions, and improve performance over time.[108] However, data collection often occurs without the full understanding or informed consent of individuals.[109] This raises questions about the adequacy of current privacy frameworks and whether individuals have meaningful control over their personal information in the context of AI.

Even in cases where individuals understand that their data is collected, they often lack the true ability to opt-out or withhold consent from AI systems' pervasive data-gathering practices. As legal scholars Manheim and Kaplan note, "AI exacerbates and exponentially multiplies the existing trends to over-collect data and use data for unintended purposes not disclosed to users at the time of collection."[110] The lack of transparency coupled with AI's ravenous appetite for data can undermine the principle of informed consent and leave individuals' personal data vulnerable to unanticipated exploitation.

b. Surveillance and Loss of Anonymity

The widespread deployment of AI systems also raises concerns about increased surveillance and the erosion of anonymity. AI-powered technologies, such as facial recognition and predictive policing, can enable pervasive monitoring of individuals in both physical and digital spaces. Manheim and Kaplan warn that "AI is being used to sow seeds of distrust of government and democratic institutions, leading to paralysis of collective action."[111] This surveillance can create a chilling effect on free speech, association, and other fundamental rights, as the fear of monitoring may cause individuals to self-censor or alter their behavior.

Moreover, AI systems can conduct mass level re-identification of individuals from supposedly anonymized data, undermining efforts to protect privacy through anonymization techniques, such as virtual private networks ("VPNs").[112] This means that, even when personal data is substantially stripped of identifying information, AI can be used to single out individuals

---

[108] *Id*.

[109] Erik Lampmann-Shaver, *Privacy's Next Act*, 19 WASH. J. L. TECH. & ARTS 97, 121 (2024). Vast amounts of personal data are required in order for AI to function properly. *Id*.

[110] Manheim & Kaplan, *supra* note 1, at 121.

[111] *Id.* at 113.

[112] Luc Rocher et al., *Estimating the Success Of Re-Identifications In Incomplete Datasets Using Generative Models*, 10 NATURE COMMC'NS 3069, July 23, 2019, at 2. This study offers a contemporary look at the limitations of anonymization techniques in the face of modern AI and machine learning algorithms, suggesting that most individuals could be re-identified in large datasets with just a few attributes.

and infer sensitive details about them, effectively eroding anonymity.[113] Similar systems can create strong data inferences about most individuals in our society.[114]

c. Data Security and Breaches

The vast amount of personal data collected and processed by AI systems raises concerns about data security and potential breaches. As AI further integrates into various sectors—healthcare, finance, government, etc.—the risks associated with data breaches also increase.[115] Data breaches involving AI systems can expose sensitive personal information, leading to identity theft, financial fraud, and other harms. A part of the concern arises from short-term unpredictability because of the rapid technological advancements and many boot-strapped applications developed without thorough privacy and security implementation.[116] But, as AI further integrates into various industries and sectors, including healthcare, finance, and government, the risks associated with data breaches significantly increase.[117] Another concern stems from the scale and complexity of modern AI architecture, which further increases the potential attack surface for malicious actors.[118]

d. Privacy Violations and Data Misuse

To address privacy violations and data misuse, many are calling for comprehensive legal and regulatory frameworks that prioritize privacy

---

[113] *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701 (2010). Paul Ohm's work critically analyzes the fallacy of data anonymization and argues for a reevaluation of privacy protections in the era of big data and AI, highlighting legal and policy implications. *Id.* at 1706.

[114] Yonathan Arbel et al., *Systemic Regulation of Artificial Intelligence*, ARIZ. ST. L.J. (forthcoming 2024) (manuscript at 14) ("AI will allow attackers to cast a much wider net by cutting the cost of interacting with each potential mark. This will allow scammers to vastly expand and disguise their operations, increasing the scope and effectiveness of fraud.").

[115] *See* Mason Marks & Claudia E. Haupt, *AI Chatbots, Health Privacy, and Challenges to HIPAA Compliance*, 330 JAMA 309, 309 (2023) ("If users had searched for answers to medical questions, health information could have been publicly revealed.").

[116] *Artificial Intelligence Risks To Privacy Demand Urgent Action – Bachelet*, OHCHR (Sept. 9, 2021), https://www.ohchr.org/en/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet [https://perma.cc/6U3B-SX44].

[117] Blake Murdoch, *Privacy and Artificial Intelligence: Challenges For Protecting Health Information In a New Era*, 22 BMC MED. ETHICS, Sept. 15, 2021, at 3.

[118] Microsoft Threat Intelligence, *Staying Ahead of Threat Actors in the Age of AI, Microsoft Security Blog*, MICROSOFT (Feb. 14, 2024), https://www.microsoft.com/en-us/security/blog/2024/02/14/staying-ahead-of-threat-actors-in-the-age-of-ai/ [https://perma.cc/FFC6-H83L].

protection in the context of AI.[119] This may include measures such as mandating transparency and explainability in AI systems, requiring explicit consent for data collection and use, implementing strict data security standards, or establishing oversight mechanisms to monitor and mitigate the risks associated with AI.[120]

Well before 2022, the year ChatGPT was released, critics were already highlighting limitations of the current patchwork of data privacy laws in the United States.[121] While sector-specific laws, like the Health Insurance Portability and Accountability Act in the healthcare sector, provide some protections for sensitive personal information, they do not comprehensively address the challenges posed by AI across industries.[122]

The absence of a robust federal data privacy law has led to a fragmented approach, with the narrow Privacy Act of 1974 proving increasingly outdated and inadequate to regulate AI.[123] States like California have stepped in to fill the void by enacting their own comprehensive data privacy legislation with the California Consumer Privacy Act.[124] However, the rapid AI revolution has renewed calls for a strong federal data privacy law that would establish a unified framework for protecting personal information and regulating the development and use of AI systems.[125]

The need for comprehensive data privacy legislation is further underscored by the potential for AI to exacerbate existing power imbalances between individuals and the entities collecting and processing their personal information at staggering rates.[126] As Manheim warns, the dominance of

---

[119] *See* Manheim & Kaplan, *supra* note at 1, at 181–82.

[120] Lampmann-Shaver, *supra* note 109, at 129 ("Privacy law's next act is already underway—and every indication is it will fundamentally reshape the relationship between consumers and businesses engaged in the digital economy.").

[121] Manheim & Kaplan, *supra* note at 1, at 162.

[122] *Id.*

[123] U.S. Privacy Act of 1974, 5 U.S.C. § 552a.

[124] California Consumer Privacy Act of 2018, CAL CIV. CODE §§ 1798.100–1798.199.100 (West 2023). State action is growing rapidly in this area, with at least 19 states having enacted enhanced data privacy laws within the past few years. *U.S. State Privacy Laws*, LEWIS RICE, https://www.lewisrice.com/u-s-state-privacy-laws/ [https://perma.cc/GA94-NR7H] (last visited Aug. 29, 2024).

[125] Cathy McMorris Rodgers & Jay Obernolte, *AI's Rise Flags Need for Federal Privacy and Security Protection*, BLOOMBERG LAW (Nov. 6, 2023), https://news.bloomberglaw.com/us-law-week/ais-rise-flags-need-for-federal-privacy-and-security-protection [https://perma.cc/5WZN-WF3S]; Cameron F. Kerry, *How Privacy Legislation Can Help Address AI*, BROOKINGS (July 7, 2023), https://www.brookings.edu/articles/how-privacy-legislation-can-help-address-ai/ [https://perma.cc/L253-PQPK]; Brandon Pugh & Steven Ward, *What Does AI Need? A Comprehensive Federal Data Privacy and Security Law*, IAPP (July 12, 2023), https://iapp.org/news/a/what-does-ai-need-a-comprehensive-federal-data-privacy-and-security-law/ [https://perma.cc/T8K7-N7UM].

[126] Manheim & Kaplan, *supra* note 1, at 110 ("The power of these technology giants to act as 'Emergent Transnational Sovereigns' stems in part from the ability of

major technology companies stems partly from AI's ability to "subvert or displace regulatory law."[127]

2. Lack of Transparency and Explainability

As discussed above, many AI systems, especially deep learning models, are "black boxes."[128] Although the model developers may understand the training data and underlying model, the systems are opaque to outsiders. Even insiders are regularly surprised at the outputs. This lack of transparency makes it difficult to audit AI systems, fix errors, and ensure accountability.[129] The more complex and integrated these systems are in decision-making processes, the more pressing the transparency and explainability concerns are.[130] The lack of clarity surrounding how AI and ML models arrive at their decisions and predictions has led to questions about their accountability, fairness, and trustworthiness.[131]

a. Transparency Concerns

Transparency in AI systems refers to the ability to understand their inner workings, including the input data, algorithms used, and resulting outputs.[132] Many modern AI models, however, are inherently opaque.[133] This lack of

AI software ('West Coast Code') to subvert or displace regulatory law ('East Coast Code').").

[127] *Id.*

[128] *See supra* Part II.B.

[129] Yanisky-Ravid & Hallisey, *supra* note 3, at 473 ("The AI Data Transparency Model is a first step towards ensuring that the data used to train AI systems complies with all relevant regulations and societal expectations, which may otherwise limit the AI's use. In previous sections, this Article has identified some of the many risks that AI systems could pose for individuals and society as a whole."). Of course, the pre-AI legal system also had numerous transparency problems. *See* Mathilde Cohen, *Sincerity and Reason-Giving: When May Legal Decision-Makers Lie?*, 59 DEPAUL L. REV. 1091 (discussing the distinction between motivating and normative reasons in legal decision-making and the contention that sincerity, understood as the provision of actual motivating reasons, is not always a requirement in legal contexts).

[130] Cynthia Dwork & Martha Minow, *Distrust of Artificial Intelligence: Sources & Responses from Computer Science & Law*, 151 DAEDALUS 309, 309 (2022). Distrust of AI systems may stem from their opacity, apparent objectivity, and the perceived and actual power and autonomy of AI. *Id.*

[131] FRANK PASQUALE, NEW LAWS OF ROBOTICS: DEFENDING HUMAN EXPERTISE IN THE AGE OF AI 49–54 (Harvard University Press, 2020).

[132] Paul Grimm et al., *Artificial Intelligence as Evidence*, 19 NW. J. TECH. & INTELL. PROP. 9, 61–62 (2021).

[133] Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, 3 BIG DATA & SOC'Y, Sept. 15, 2015, at 3.

transparency has proven challenging for developers, users, and regulators to comprehend the inner workings of these systems.[134]

The opacity of AI and ML systems can conceal biases present in the training data or the algorithms themselves, leading to discriminatory outcomes.[135] Moreover, the lack of transparency hinders accountability when these systems make decisions that significantly impact individuals' lives.[136] Without the ability to examine and understand the decision-making process, affected individuals may have little recourse to challenge adverse outcomes.[137]

b. Explainability Concerns

Explainability is crucial for building trust in AI systems, as it enables users to comprehend the reasons behind specific outcomes and allows them to challenge decisions that may be erroneous or biased.[138] Explainability, a related concept to transparency, refers to the ability to provide meaningful and understandable explanations for AI systems' decisions or predictions.[139] Achieving explainability in complex AI models, however, is no simple task because the decision-making process may involve intricate interactions between numerous variables and layers of abstraction. This complexity can make it difficult for even the AI system developers to provide clear and concise explanations for their outputs.[140]

---

[134] W. Nicholson Price II, *Black-Box Medicine*, 28 HARV. J.L. & TECH. 419, 462–64 (2015). The black-box nature of these algorithms raises substantial challenges for regulators because the relationships at the heart of the black-box algorithm are not specified, nor are the weight or effect of any particular input on the output known. *Id.* at 424.

[135] Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 674 (2016).

[136] Ashley Deeks, *The Judicial Demand for Explainable Artificial Intelligence*, 119 COLUM. L. REV. 1829, 1829–30 (2019); Stephen Casper et al., *Black-Box Access is Insufficient for Rigorous AI Audits* (2024) (unpublished manuscript), accessible at https://arxiv.org/pdf/2401.14446.pdf [https://perma.cc/4GRG-KN93] (arguing that white-box access to AI systems' inner workings allows for stronger audits and testing).

[137] Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147, 1157 (2017). An algorithmic system that is unreviewable due to trade secret protection would raise due process issues to the extent that it prevents affected parties from obtaining judicial review of agency action. *Id.*

[138] Lillian Edwards & Michael Veale, *Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For*, 16 DUKE L. & TECH. REV. 18, 61–64 (2017).

[139] Sarah Wachter et al., *Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*, 31 HARV. J.L. & TECH. 841, 878 (2018).

[140] Jessica Newman, *Explainability Won't Save AI*, BROOKINGS (May 19, 2021), https://www.brookings.edu/articles/explainability-wont-save-ai/ [https://perma.cc/N2LJ-BKC8]; Quy Mai, *Transparency is a Misplaced Regulatory Focus for Holding*

The lack of explainability in AI systems can erode public trust, particularly when they are deployed in sensitive domains or when their decisions have significant consequences.[141] Without the ability to understand how these systems arrive at their conclusions, individuals may be hesitant to rely on them or accept their outcomes, limiting AI's potential benefits to society.[142]

c. Addressing Transparency and Explainability

Transparency and explainability are important factors in ensuring the trustworthiness, accountability, and fairness of AI systems, particularly those that play important societal or personal roles. In this vein, the first step is to address concerns about their black box nature. One approach to addressing transparency and explainability is the development of "explainable AI" ("XAI"), which are techniques used to further human interpretation and understand AI models' decision-making processes.[143] XAI would provide clear and understandable explanations for the decisions made by AI systems by demystifying their inner workings.[144] XAI techniques can shed light on the model architecture, training data, and decision-making processes of AI systems, enabling stakeholders, including users and other affected parties, to better understand and evaluate their outputs.[145] Another approach is to incorporate transparency and explainability requirements into the development and deployment of AI systems. This would involve documenting the data used to train these systems, testing for biases and errors,

Adaptive Software as Medical Devices (SAMDS) Accountable, 2023 B.C. INTELL. PROP. & TECH. F. 1, 14 (2023) ("Reconstructing the causal decision-making process of an adaptive SaMD through explanations is appealing for trust-inducing purposes, but the fact that these are fundamentally data-centric devices suggests that a focus on the data itself could be an equally important aspect of regulation.").

[141] William Magnuson, *Artificial Financial Intelligence*, 10 HARV. BUS. L. REV. 337, 362–63 (2020).

[142] *See* Teresa M. Harrison & Luis Felipe Luna-Reyes, *Cultivating Trustworthy Artificial Intelligence in Digital Government*, 40 SOC. SCI. COMPUT. REV. 494 (2022) (discussing how AI/ML predictive performance alone is insufficient for fostering public trust absent adequate transparency and explanation).

[143] *See* Marco Tulio Ribeiro et al., *"Why Should I Trust You?": Explaining the Predictions of Any Classifier*, 22 ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING PROC. 1135, 1135 (2016) (introducing the concept of XAI and presenting a novel explanation technique called LIME).

[144] *Id.*

[145] *See* George Lawton, *AI Transparency: What Is It and Why Do We Need It?*, TECHTARGET (Mar. 3, 2024), https://www.techtarget.com/searchcio/tip/AI-transparency-What-is-it-and-why-do-we-need-it [https://perma.cc/R33V-EMCM] (providing an overview of the key aspects of AI transparency, including the need for explainability, accountability, and stakeholder communication).

and providing clear explanations for how decisions are made.[146] By building transparency and explainability into the design of AI systems, developers can help ensure that they are fair, accountable, and trustworthy.[147]

Taking a step back, there are several levels of AI transparency to consider. At the most technical level, explainability involves understanding the internal mechanics and decision-making processes of AI models, including their algorithms, data inputs, and model structure.[148] However, full transparency at this level may raise IP concerns, as it could involve disclosing entire models, model weights, training data, and turnkey approaches that allow for full reproduction. At a higher level, transparency involves governance and accountability, which entails establishing protocols for documenting decisions made about AI systems, assigning responsibility and liability, and complying with applicable regulations. This allows for later accountability against the creators of these systems.[149]

Finally, stakeholder communication is a crucial aspect of AI transparency. This involves clearly communicating the capabilities, limitations, and potential impacts of AI systems to all relevant stakeholders.[150] The level of transparency required in stakeholder communication may vary depending on the audience, ranging from regulators and national security experts to the general public.

Policymakers and regulators play a crucial role in promoting transparency and explainability in AI systems. Policymakers can develop guidelines and standards for the development and deployment of these technologies, as well as require companies to provide clear explanations for

---

[146] *See* Yavar Bathaee, *The Artificial Intelligence Black-Box and the Failure of Intent and Causation*, 31 HARV. J. L. & TECH. 889, 928 (2018) (describing the black-box challenge and how increasing AI's transparency is a popular proposed solution).

[147] *See WHO Issues First Global Report on Artificial Intelligence (AI) in Health and Six Guiding Principles for Its Design and Use*, WORLD HEALTH ORG. (June 28, 2021), https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use [https://perma.cc/MM2V-JQ7N] (asserting that "ensuring transparency, explainability, and intelligibility" are key principles in regulating AI).

[148] *See* Anna-Mari Rusanen & Jukka K. Nurminen, *What is Transparency?*, ETHICS OF AI, https://ethics-of-ai.mooc.fi/chapter-4/2-what-is-transparency/ [https://perma.cc/E2G3-NKPF] (last visited July 6, 2024) (discussing the different definitions and interpretations of transparency in the context of AI systems, highlighting the importance of comprehensibility).

[149] *See generally* Heike Felzmann et al., *Transparency You Can Trust: Transparency Requirements for Artificial Intelligence Between Legal and Contextual Concerns*, 6 BIG DATA & SOC'Y 1 (2019) (examining the legal, social, and ethical aspects of transparency requirements for AI systems).

[150] *See* Airlie Hillard, *What is AI Transparency?*, HOLISTIC AI (Feb. 6, 2024), https://www.holisticai.com/blog/ai-transparency [https://perma.cc/HW6Q-G8NC] (emphasizing the overarching goal of AI transparency in establishing trust, particularly among citizens and users who may be at risk of harm from AI systems).

how their systems operate and make decisions.[151] Regulators can also work to ensure that individuals have the right to challenge AI-made decisions and to obtain explanations for those decisions.[152] Proposed regulations, such as the European Union's recently adopted Artificial Intelligence Act ("AI Act") and the United States' AI Foundation Model Transparency Act, seek to mandate transparency around the data and processes used to train AI models.[153] While these regulations are a major step in the AI era, their effectiveness is questionable because of the competing power of trade secrecy rights as discussed in Part IV, below.

d. The Debate Over Open Access to AI Model Weights

Although there have been many calls for greater transparency, there are a number of consequences associated with taking transparency too far, particularly pertaining to IP.[154] On February 26, 2024, the National Telecommunications and Information Administration ("NTIA") issued a request for comment on *Dual Use Foundation Artificial Intelligence Models With Widely Available Model Weights*.[155] This request was issued in accordance with Executive Order 14110, which instructed the Secretary of Commerce, through the NTIA, to "conduct a public consultation process" on the potential risks, benefits, and other implications of dual-use foundation models for which the model weights are widely accessible.[156] Additionally, the Executive Order directed the NTIA to explore appropriate policy and regulatory approaches for addressing these issues.[157]

---

[151] *See* Yesha Yadav & Chris Brummer, *Fintech and the Innovation Trilemma*, 107 GEO. L.J. 235, 242 (2019) (discussing the role of regulators in balancing innovation, transparency, and consumer protection in the context of financial technology).

[152] *See* Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 190, 206 (2019) (noting that GDPR guidelines contemplate at least internal audits of algorithms "to prevent errors, inaccuracies, and discrimination on the basis of sensitive . . . data" in individual automated decision-making).

[153] *See Beyer, Eshoo Introduce Landmark AI Regulation Bill*, CONGRESSMAN DON BEYER (Dec. 22, 2023), https://beyer.house.gov/news/documentsingle.aspx?DocumentID=6052 [https://perma.cc/7FHY-8RWH] (explaining the rationale and key provisions of the proposed legislation); H.R. 6881, 118th Cong. (2024).

[154] *See* Sayash Kapoor et al., On the Societal Impact of Open Foundation Models (Ctr. for Rsch. on Found. Models, Stanford Univ., Working Paper, 2024), https://crfm.stanford.edu/open-fms/paper.pdf [https://perma.cc/QND5-LWVQ] (discussing the importance of evaluating the risks of open AI models in comparison to the risks and benefits from closed models and pre-existing technologies).

[155] Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights, 89 Fed. Reg. 14059, 14059 (Feb. 26, 2024).

[156] *Id.*; *see supra* Part II.B.2.d.

[157] Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights, 89 Fed. Reg. 14059, 14059 (Feb. 26, 2024).

The request for comment notes that open foundation models, where the model weights are widely available, have the potential to support growth and expand access to AI usage.[158] However, it also acknowledges potential risks due to misuse, lack of accountability, and weak oversight.[159] The NTIA sought input on defining "open" and "widely available" models, comparing risks of open versus closed models, assessing benefits of open models, and considering potential regulatory mechanisms.[160]

In response to this request, two competing comments were submitted by groups of legal scholars and organizations. First, the "Accountable Tech" letter was signed by numerous civil society organizations and academics, exhibiting a broad consensus promoting "openness and transparency in AI models."[161] The letter argues that "open models can provide significant benefits to society" including "[a]dvancing innovation, competition, and research," "[p]rotecting civil rights and human rights," and "[e]nsuring safety and security."[162] While restrictive approaches, such as secrecy requirements or export controls, "come with significant negative consequences[.]"[163]

In contrast, a comment by the directors of the Center for Law & AI Risk supports "sensible" restrictions on open-sourcing generative AI systems.[164] This comment argues that generative AI has an imminent potential to cause "harm[] to human life, limb, and freedom" through misuse, accidents, or autonomous action of these systems.[165] It points to studies showing advanced AI systems' ability to autonomously find vulnerabilities and hack websites,[166] as well as the potential for AI to aid in producing explosives, identifying deadly viruses, and inventing harmful molecules, recognizing the dire potential of lethal autonomous weapons systems ("LAWS").[167] The comment

---

[158] *Id.* at 14060.

[159] *Id.* at 14061.

[160] *Id.*

[161] Letter from Accountable Tech et al. to Sec'y Gina Raimondo, Dep't of Commerce (Mar. 25, 2024) (on file with author).

[162] *Id.*

[163] *Id.*; Carrick Flynn, *Recommendations on Export Controls for AI*, CSET (Feb. 2020), https://cset.georgetown.edu/publication/recommendations-on-export-controls-for-artificial-intelligence/ [https://perma.cc/LU88-MQHD] ("New export control regulations on general purpose AI software . . . are unlikely to succeed and should not be implemented.").

[164] Ctr. for Law & AI Risk, *Comments on Dual Use Foundation Artificial Intelligence Models* 1 (NTIA 2024) (on file with author) [hereinafter *Comments on Dual Use Foundation Artificial Intelligence Models*].

[165] *Id.*

[166] Richard Fang et al., *LLM Agents Can Autonomously Hack Websites*, ARXIV (Feb. 16, 2024), https://arxiv.org/pdf/2402.06664.pdf [https://perma.cc/9SX6-42PM].

[167] Andres M. Bran et al., *ChemCrow: Augmenting Large-Language Models With Chemistry Tools*, ARXIV (Apr. 11, 2023), https://arxiv.org/abs/2304.05376 [https://perma.cc/E3TE-JREL]; Emily H. Soice et al., *Can Large Language Models Democratize Access To Dual-Use Biotechnology?*, ARXIV (Jun. 6, 2023), https://arxiv.org/abs/2306.03809 [https://perma.cc/257F-NKLD]; Fabio Urbina et al., *Dual Use Of*

contends that while closed systems can mitigate risks through alignment methods, system prompts, output filtering, and centralized control, "none of this is possible with open-sourced systems."[168] Open systems "can have their alignment stripped away at trivial cost," "be merged and fine-tuned to produce novel behaviors and capabilities," and are "effectively insulated from any enforcement action."[169] The drafters of this comment believe these risks outweigh the potential benefits of expanded AI access.

The Center's comment does not oppose open-sourcing for select AI models, acknowledging benefits for equity, research, safety and interpretability.[170] Rather it argues for restrictions on open-sourcing "powerful frontier generative AI systems," suggesting a more cautionary approach that first requires a full understanding of potential negative uses of each system.[171]

These competing comments reflect a fundamental tension in the AI governance debate. Proponents of open-source models emphasize myriad benefits of open models for innovation, competition, scientific progress, civil rights, and even safety and security.[172] They caution against broad restrictions absent clear evidence of marginal risk compared to closed models.[173] Critics highlight the unique and potentially catastrophic dangers posed by advanced AI systems in the wrong hands, which may be difficult to control once openly disseminated.[174] They advocate for a cautious, case-by-case approach to open-sourcing frontier models as part of a comprehensive regulatory framework.[175]

Artificial Intelligence Powered Drug Discovery, 4 NATURE MACH. INTEL. 189–91 (2022).

[168] *Comments on Dual Use Foundation Artificial Intelligence Models*, *supra* note 164, at 1.

[169] Xiangyu Qi et al., *Fine-Tuning Aligned Language Models Comprises Safety, Even When Users Do Not Intend To!*, ARXIV (Oct. 5, 2023), https://arxiv.org/pdf/2310.03693.pdf [https://perma.cc/4G2Q-C2CG].

[170] *Comments on Dual Use Foundation Artificial Intelligence Models*, *supra* note 164, at 1.

[171] *Id*.

[172] Elizabeth Seger et al., *Open-Sourcing Highly Capable Foundation Models: An Evaluation of Risks, Benefits, and Alternative Methods for Pursuing Open-Source Objectives*, ARXIV (Sept. 29, 2023), https://arxiv.org/pdf/2311.09227 [https://perma.cc/3XPX-Z2GC].

[173] *Id*.

[174] *Risks from Artificial Intelligence*, CTR. FOR THE STUDY OF EXISTENTIAL RISK, https://www.cser.ac.uk/research/risks-from-artificial-intelligence/ [https://perma.cc/SP3X-5U9J] (last visited July 7, 2024); Michael Littman et al., *Gathering Strength, Gathering Storms: One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report*, STAN. UNIV. 9–10 (2021), https://ai100.stanford.edu/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence-ai100-2021-1-0 [https://perma.cc/Z2RC-L926].

[175] Littman et al., *supra* note 174, at 42–44.

3. Bias, Discrimination, and Fairness Issues

Potential biases and discriminatory outcomes from AI technologies pose an urgent threat to society.[176] As AI systems become pervasive in decision-making processes that impact individuals' daily lives (such as employment, credit lending, healthcare, and criminal justice), it is important to address the issue of bias to ensure systemic fairness and equality. This risk has prompted demands for both caution and regulatory reforms.

One major fear is that AI systems can perpetuate and even amplify pre-existing societal biases in the training data—typically available on the public Internet.[177] If the training data contains historical biases or underrepresents certain groups, the resulting AI model may accordingly make decisions that discriminate against those groups.[178] For example, in the context of hiring, if an AI system is trained on historical hiring data that favors male candidates, it may learn to discriminate against female applicants.[179] Similarly, in the

[176] *See, e.g.*, Ryan Calo & Danielle Keats Citron, *The Automated Administrative State: A Crisis of Legitimacy*, 70 EMORY L.J. 797 (2021); Cary Coglianese, *Administrative Law in the Automated State*, 150 DAEDALUS 104 (2021) (ensuring that an automated state is also an empathic one.); Yonathan Arbel et al., *Systemic Regulation of Artificial Intelligence*, ARIZ. ST. L. J. (forthcoming 2024), accessible at https://ssrn.com/abstract=4666854 [https://perma.cc/2XRU-GNYE] ("As scholars have explored, these models tend to have discriminatory effects with regard to race, gender, class, ethnicity, religion, disability status, and more, especially for groups with a history of suffering discrimination or disadvantage."); Rebecca Slaughter et al., *Algorithms and Economic Justice: A Taxonomy of Harms and a Path Forward for the Federal Trade Commission*, 23 YALE J.L. & TECH. 1, 6–36 (2021) (discussing algorithmic harms and the need for regulatory solutions); Ifeoma Ajunwa, *Automated Video Interviewing as the New Phrenology*, 36 BERKELEY TECH. L.J. 1173 (2021) (examining how automated video interviewing technologies can lead to unlawful discrimination in employment); Colin Clemente Jones, *Systematizing Discrimination: AI Vendors & Title VII Enforcement*, 171 U. PA. L. REV. 235 (2022) (analyzing how AI vendors can systematize and scale discrimination in employment opportunities); Michele Estrin Gilman, *Expanding Civil Rights to Combat Digital Discrimination on the Basis of Poverty*, 75 SMU L. REV. 571 (2022) (discussing digital discrimination against low-income individuals due to automated decision-making systems); Maurice Dyson, *Combatting AI's Protectionism & Totalitarian-Coded Hypnosis: The Case for AI Reparations & Antitrust Remedies in the Ecology of Collective Self-Determination*, 75 SMU L. REV. 625 (2022) (arguing for reparations and antitrust remedies to address AI's protectionism, surveillance, and discrimination).

[177] Barocas & Selbst, *supra* note 135, at 674 (2016) (noting that "data mining can inherit the prejudices of prior decision makers or reflect the widespread biases that persist in society at large").

[178] ORLY LOBEL, THE EQUALITY MACHINE: HARNESSING DIGITAL TECHNOLOGY FOR A BRIGHTER, MORE INCLUSIVE FUTURE 9–11 (PublicAffairs, 2022) ("[R]eflecting the incommensurability of moral values and the difficulty of developing an algorithm that can reconcile competing moral answers and dilemmas.").

[179] Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, REUTERS (Oct. 9, 2018, 10:00 PM), https://www.reuters.com/article/idUSL2N1VB1FQ/ [https://perma.cc/M3Z4-63DQ].

criminal justice system, risk assessment algorithms used to predict recidivism have been shown to exhibit racial biases.[180]

As suggested above, AI's black box problem raises additional concerns for bias and discrimination.[181] The complex and opaque nature of some AI algorithms make it difficult to detect and mitigate bias, hold decision-makers accountable for discriminatory outcomes, and establish public trust.[182] The lack of transparency and explainability in AI systems can exacerbate issues of bias and discrimination.

Researchers have proposed various bias mitigation technologies. One approach is to improve the training data by taking steps to free it from historical biases and ensure that it is representative of the diverse populations the AI system will serve.[183] This may involve actively collecting data from underrepresented groups or using techniques such as data augmentation to balance the training dataset. Another approach is to develop AI algorithms that are inherently fair and unbiased, such as establishing fairness criteria or incorporating fairness constraints into the learning process.[184] Researchers have also explored the use of adversarial de-biasing techniques by training a separate AI model to detect and remove biases from the primary AI system.[185] The effectiveness of this approach may depend on the specific implementation, such as whether the de-biasing occurs during training or in real-time.

---

[180] *See* Jennifer L. Skeem & Christopher T. Lowenkamp, *Risk, Race, and Recidivism: Predictive Bias and Disparate Impact*, CRIMINOLOGY, Nov. 3, 2016, at 2; Lily Hu & Avi Feller, *Criminal Justice Algorithms: Being Race-Neutral Doesn't Mean Race-Blind*, THE CONVERSATION (Mar. 31, 2022), https://theconversation.com/criminal-justice-algorithms-being-race-neutral-doesnt-mean-race-blind-177120 [https://perma.cc/HK66-S8M9]; Will Douglas Heaven, *Predictive Policing Algorithms Are Racist. They Need to Be Dismantled.*, MIT TECH. REV. (July 17, 2020), https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/ [https://perma.cc/Z5AW-Z4F4].

[181] *See* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653 (2017); *see supra* Part II.B.2 (discussing transparency and explainability).

[182] Cary Coglianese & David Lehr, *Transparency and Algorithmic Governance*, 71 ADMIN. L. REV. 1, 14–29 (2019) (discussing the challenges of transparency in the context of algorithmic governance).

[183] Alexander Amini et al., *Uncovering and Mitigating Algorithmic Bias Through Learned Latent Structure*, PROC. 2019 CONF. ON A.I., ETHICS, & SOC'Y 289, 291 (2019) (arguing for the collection of more representative data to fight bias).

[184] Cynthia Dwork et al., *Fairness Through Awareness*, PROC. 3D INNOVATIONS IN THEORETICAL COMPUT. SCI. CONF. 214, 216–23 (2012) (proposing a framework for incorporating fairness into AI algorithms).

[185] Brian Hu Zhang et al., *Mitigating Unwanted Biases with Adversarial Learning*, PROC. 2018 AAAI/ACM CONF. ON A.I., ETHICS, & SOC'Y 335, 335–40 (2018) (presenting an adversarial approach to mitigating bias in AI systems).

4. Job Displacement and Technological Unemployment

While AI will likely create new jobs, advances have renewed worries among economists, policymakers, and the public at large about the potential for these technologies to displace human workers and exacerbate technological unemployment.[186] This, in turn, could deepen the income disparity, especially if the benefits of AI productivity are captured by owners rather than workers.[187]

While low-skilled and routine jobs have been at the highest risk of automation,[188] the rapid advancements in AI technology suggest that not even demanding, high-skilled, and creative jobs are immune from future displacement. For example, AI systems are already generating news articles, composing music, and creating artwork, demonstrating their potential to encroach upon fields viewed as the pinnacle of human creativity.[189] Similarly, in the realm of professional services, such as law and medicine, AI tools are being developed to form legal arguments, make medical diagnoses, and provide financial services.[190]

The prospect of widespread technological unemployment and growing inequality has led some commentators to warn about potential social unrest and political instability.[191] In light of these concerns, there is growing demand for regulation and policy interventions to mitigate these potential negative

---

[186] Carl Benedikt Frey & Michael A. Osborne, *The Future of Employment: How Susceptible Are Jobs to Computerisation?*, 114 TECH. FORECASTING & SOC. CHANGE 254, 265 (2016) ("According to our estimates, about 47% of total US employment is at risk [of automation]."); MARTIN FORD, RISE OF THE ROBOTS: TECHNOLOGY AND THE THREAT OF A JOBLESS FUTURE 29–34 (2015).

[187] ERIK BRYNJOLFSSON & ANDREW MCAFEE, RACE AGAINST THE MACHINE: HOW THE DIGITAL REVOLUTION IS ACCELERATING INNOVATION, DRIVING PRODUCTIVITY, AND IRREVERSIBLY TRANSFORMING EMPLOYMENT AND THE ECONOMY 36–47 (2011) (discussing the various "winners" and "losers" of the new machine age).

[188] David H. Autor, *Why Are There Still So Many Jobs? The History and Future of Workplace Automation*, 29 J. ECON. PERSPS. 3, 11 (2015). Routine, codifiable tasks are more readily substituted with machines—they are the jobs that have been most clearly and negatively impacted by computerization. *Id.*

[189] Bernard Marr, *13 Mind-Blowing Things Artificial Intelligence Can Already Do Today*, FORBES (Nov. 11, 2019, 12:31 AM), https://www.forbes.com/sites/bernardmarr/2019/11/11/13-mind-blowing-things-artificial-intelligence-can-already-do-today/ [https://perma.cc/MED2-UYV2].

[190] DANIEL SUSSKIND & RICHARD SUSSKIND, THE FUTURE OF THE PROFESSIONS: HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS 109–35 (2015) (discussing the potential impact of AI on various professional fields).

[191] *See* Cynthia Estlund, *What Should We Do After Work? Automation and Employment Law*, 128 YALE L.J. 254, 288–78 (2018). The economic and social disruption that automation threatens to bring about could also lead to political disruption. *See id.*

impacts.[192] Some have proposed implementing a "robot tax" or a universal basic income to help cushion the blow of technological unemployment and ensure an equitable distribution of AI benefits.[193] Others have emphasized the need for education and mass re-training to assist the working class in adapting to the changing market.[194] There have also been proposals for more direct regulation of AI development and deployment, such as requiring algorithmic impact assessments and establishing governance frameworks to address these job displacement concerns.[195]

5. Safety, Control, and Robustness Challenges

Ensuring safety, robustness, and ongoing human control of AI is also of paramount importance. The potential risks posed by AI range from localized harms, such as autonomous vehicle accidents, to existential threats like highly advanced AI superseding human control.[196] The fear of AI surpassing and potentially overwhelming human agency has been a recurring theme in popular culture for decades, appearing in iconic science fiction stories like *Maximum Overdrive*, *Terminator*, *The Matrix*, *2001: A Space Odyssey*, and *Ex Machina*.[197] While these depictions may have primed the public's imagination and invoked fear about AI's risks, the underlying anxieties around new, powerful, and potentially uncontrollable technologies are primal.

There have been calls for a multifaceted regulatory approach that would set safety standards, maintain meaningful human oversight, and endeavor to align AI systems with human and societal values.[198] This is particularly

---

[192] Yoshua Bengio, *Government Interventions to Avert Future Catastrophic AI Risks,* HDSR (Apr. 15, 2024), https://hdsr.mitpress.mit.edu/pub/w974bwb0/release/2 [https://perma.cc/7DSQ-BHTJ].

[193] Ryan Abbott & Bret Bogenschneider, *Should Robots Pay Taxes? Tax Policy in the Age of Automation*, 12 HARV. L. & POL'Y REV. 145, 169–173 (2018) (discussing various proposals for a robot tax); Estlund, *supra* note 191, at 258.

[194] JOSEPH E. AOUN, ROBOT-PROOF: HIGHER EDUCATION IN THE AGE OF ARTIFICIAL INTELLIGENCE 6–18 (The MIT Press, 2017) (arguing that education needs to adapt to teach people to work with AI).

[195] Bruno Bastit et al., *The AI Governance Challenge*, S&P GLOBAL (Nov. 29, 2023), https://www.spglobal.com/en/research-insights/special-reports/the-ai-governance-challenge [https://perma.cc/79LL-8LMM].

[196] Calo, *supra* note 93, at 417–18.

[197] Justin Mark & Tucker Harris, *Could 'The Terminator' Really Happen? Experts Assess Hollywood's Visions of AI*, WASH. POST (Sept. 29, 2023, 6:00 AM), https://www.washingtonpost.com/technology/interactive/2023/artificial-intelligence-ai-holloywood-movies-characters/ [https://perma.cc/P93J-QFX6]; Ian Banks, *Killer Robots & Malicious Machines: 20 Films That Will Have You Pulling the Plug!*, NIGHTMARE ON FILM ST. (Oct. 19, 2023), https://nofspodcast.com/killer-robots-malicious-machines-20-films-that-will-have-you-pulling-the-plug [https://perma.cc/7KDV-CYES].

[198] Matthew R. Gaske, *Regulation Priorities for Artificial Intelligence Foundation Models*, 26 VAND. J. ENT. & TECH. L. 1, 62 (2023).

important in high-stakes contexts such as military applications, where the use of fully autonomous LAWS raises significant ethical and legal concern.[199] The increased deployment of LAWS underscores the urgent need for robust governance frameworks to ensure responsible, human-controlled deployment of AI systems.

6. Concentration of Power in Large Tech Companies

AI is dominated by a small number of technology companies, leading to a concerning concentration of AI talent, capabilities, and power among a few influential entities.[200] This consolidation raises significant economic, social, and political risks.

Economically, the high fixed costs associated with developing advanced AI systems create barriers to entry that advantage incumbent technology giants.[201] Specialized computing hardware, large training datasets, and top AI engineering talent are all assets factoring toward this barrier of entry.[202] The barrier of entry is exceptionally high for foundational models, which include large, internet-wide aspirations for training and implementation. As a result, computational power, data, and talent are highly concentrated, making it difficult for new entrants and smaller players to compete.[203] Moreover, the reliance on large, internet-scale datasets for training and incumbents' control over proprietary datasets needed for AI development may impede competition.[204] These dynamic factors entrench the market power in dominant firms.

---

[199] Calo, *supra* note 93, at 415–16.

[200] Lina M. Khan, *Remarks of Chair Lina M. Khan at Economic Club of New York*, FTC (July 24, 2023), https://www.ftc.gov/news-events/news/speeches/remarks-chair-lina-m-khan-economic-club-new-york/ [https://perma.cc/CKB3-GWZX]. Although not a complete list, major players include OpenAI (ChatGPT), Microsoft (major investor in OpenAI and now largest AI company), NVIDIA (GPU creator), Anthropic (Claude), Alphabet (Google), Meta (Facebook), and IBM (holder of a significant number of AI related patents). Reinhardt Krause, *AI Stocks: Tech Giants, Cloud Titans, Chipmakers Battle for an Edge*, INV'S. BUS. DAILY (June 11, 2024, 8:10 AM), https://www.investors.com/news/technology/artificial-intelligence-stocks/ [https://perma.cc/BKJ7-QFWA]; Team Stash, *15 Largest AI Companies in 2024*, LEARN (June 11, 2024), https://www.stash.com/learn/top-ai-companies/ [https://perma.cc/W4MX-58UT].

[201] Tejas N. Narechania, *Machine Learning as Natural Monopoly*, 107 IOWA L. REV. 1543, 1569–88 (2022) (finding that "the fixed costs that attend to developing a machine-learning-based application are significantly greater than those associated with the average software development project . . . .").

[202] *Id.* at 1574–76.

[203] Jai Vipra & Sarah Myers West, *Computational Power and AI*, AINOW (Sept. 27, 2023), https://ainowinstitute.org/publication/policy/compute-and-ai [https://perma.cc/9AW2-97AW].

[204] Narechania, *supra* note 201, at 1575 n.136 ("[I]ntellectual property rights over these data can mimic the effects of rivalry, limiting their use to particular rightsholders.").

III. COPYRIGHT'S ROLE IN AI REGULATION

This Part looks at copyright's role in AI regulation. Part III.A provides an overview of the copyright principles in digital works. Part III.B analyzes copyright law's effect and role in AI development. Part III.C then highlights the differences in the treatment of AI training sets and outputs under copyright law. Finally, Part III.D explores the intersection of privacy rights and copyright.

A. Copyright Basics in the Digital Space

Copyright law in the United States has roots reaching back to the British Statute of Anne, which was enacted in 1710 and granted authors exclusive rights in their works for a limited period of time.[205] This concept of providing authors with temporary monopoly rights over their creative works was later enshrined in the United States Constitution, empowering Congress "to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."[206]

The Copyright Act of 1976 represents the most significant revision of United States copyright law in the past century, displacing all previous copyright statutes.[207] The 1976 Act extended copyright protection to all "original works of authorship" fixed in a tangible medium, including literary, musical, dramatic, and artistic works.[208] The 1976 Act also incorporated the concept of "fair use" for the first time, providing exceptions to exclusive copyright rights for purposes like news reporting, commentary, and parody.[209] Importantly, the 1976 Act shifted the basis for copyright duration from a fixed term to the life of the author plus fifty years (later extended to seventy years).[210]

---

[205] *History of Copyright*, US LEGAL, https://copyright.uslegal.com/history-of-copyright/ [https://perma.cc/KNW8-WGU7] (last visited July 7, 2024) (providing an overview of the history of copyright law in the United States, including the key developments leading up to the Copyright Act of 1976).

[206] U.S. CONST. art. I, § 8, cl. 8.

[207] Copyright Act of 1976, Pub. L. No. 94-553, 90 Stat. 2541; Pamela Samuelson et al., *The Copyright Principles Project: Directions for Reform*, 25 BERKELEY TECH. L.J. 1175, 1178–79 (2010); *Copyright Law of the United States (Title 17) and Related Laws Contained in Title 17 of the United States Code*, U.S. COPYRIGHT OFFICE (Dec. 23, 2022), https://www.copyright.gov/title17/ [https://perma.cc/78KN-3GRW]; *see* Apple Comput, Inc. v. Franklin Comput. Corp., 714 F.2d 1240, 1247 (3d Cir. 1983).

[208] 17 U.S.C. § 102.

[209] 17 U.S.C. § 107.

[210] Jane C. Ginsburg, *Fifty Years of U.S. Copyright: Toward a Law of Authors' Rights?*, 50 AIPLA Q.J. 635 (2023).

In the digital age, copyright law has faced new challenges in protecting original works of authorship, particularly in the digital realm.[211] However, the originality requirement mandates that a work must possess a minimal degree of creativity to qualify for copyright protection.[212] This requirement has led to debates over the copyrightability of certain types of data and software. For instance, while creative elements of software can be protected, purely functional aspects may be ineligible for copyright and are rather within the province of patents.[213] Similarly, the arrangement and selection of data in a database may be copyrightable, but the underlying facts themselves are not.[214] The United States Supreme Court has also suggested an expanded notion of fair use of computer software in certain situations.[215] However, these issues leave plenty of room for interpretation, especially with regard to new AI developments.

B. AI Training and the Potential Role of Copyright as a Blockade to AI Development

1. Copyright as it Relates to Training of AI Systems and the DMCA Safe Harbor

The advent of web scraping and other methods of bulk content extraction have posed significant challenges to copyright law in the digital age, and several cases have developed the law prior to the arrival of generative AI. Web scraping, which involves the automated collection of data from websites and other online sources, can potentially infringe upon the rights of copyright holders when the scraped content includes protected works.[216] Courts have, however, recognized that certain uses of scraped content may fall under the fair use doctrine, which permits the limited use of copyrighted material "for purposes such as criticism, comment, news reporting, teaching . . . , scholarship, or research[.]"[217] In *Perfect 10, Inc. v. Amazon.com, Inc.*, the United States Court of Appeals for the Ninth Circuit held that Google's use of thumbnails containing copyrighted images in its search results constituted fair use, as the thumbnails were highly transformative and benefitted the public by facilitating access to information.[218] The Ninth Circuit emphasized

---

[211] Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240 (3d Cir. 1983); Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 355 (1991).

[212] *Feist Publ'ns, Inc.*, 499 U.S. at 345.

[213] Comput. Assocs. Int'l, Inc. v. Altai, Inc., 982 F.2d 693, 711, 714 (2d Cir. 1992).

[214] *Feist Publ'ns, Inc.*, 499 U.S. at 349.

[215] Google LLC v. Oracle Am., Inc., 593 U.S. 1, 20–21 (2021).

[216] Jeffrey K. Hirschey, *Symbiotic Relationships: Pragmatic Acceptance of Data Scraping*, 29 BERKELEY TECH. L.J. 897, 897, 910–11 (2014).

[217] 17 U.S.C. § 107.

[218] Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1165–66 (9th Cir. 2007). Fair use in copyright law allows limited use of copyrighted material without

the transformative nature of Google's use and its effect on the potential market for the original work.[219] Similarly, as seen in *Authors Guild v. Google, Inc.* ("the Google Books case") and *Authors Guild, Inc. v. HathiTrust*, courts have grappled with the application of fair use to mass digitization projects.[220] In the Google Books case, Google scanned millions of books, including copyrighted works, to create a searchable database.[221] The Second Circuit held that Google's use of the scanned books was highly transformative and did not substitute for the original works, thus favoring a finding of fair use.[222] In *HathiTrust*, a consortium of universities created a digital library of scanned books, including copyrighted works, for preservation, text searching, and accessibility for visually impaired users.[223] The court there found that these uses were transformative and served important public interests, which, again, supported a finding of fair use.[224]

With regard to the use of unlicensed works for AI training, a Copyright Registration Service ("CRS") report highlights the argument that "the use of copyrighted works as training sets for AI is merely a transitory and non-consumptive use that does not materially interfere with owners' content or copyrights protecting it."[225] In this view, ingesting copyrighted works to develop AI is a non-consumptive intermediate use that ultimately facilitates socially beneficial applications. Some individuals in the copyright field contend this view stretches fair use too far, while others assert that the use of copyrighted works in AI training is an unfair use and, therefore, violates the owners' IP rights in the work.[226] It is easy to see the possibility that generative AI tools could use their scanning of copyrighted materials to create market substitutes that undermine creators' livelihoods. As one article notes, because AI algorithms mimic human creativity, the use of protected creative works in

permission from the rights holder, guided by four factors: the purpose and character of the use, the nature of the copyrighted work, the amount used, and the effect on the work's market value. 17 U.S.C. § 107.

[219] *Perfect 10, Inc.*, 508 F.3d at 1146.

[220] Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015); Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014).

[221] *Google, Inc.*, 804 F.3d at 207.

[222] *Id.* at 215–16.

[223] *HathiTrust*, 755 F.3d at 91–92.

[224] *Id.* at 101–02.

[225] Andrew W. Torrance & Bill Tomlinson, *Training Is Everything: Artificial Intelligence, Copyright, and "Fair Training"*, 128 DICK. L. REV. 233, 236 (2023); *see* Christopher T. Zirpoli, CONG. RSCH. SERV., LSB10922, GENERATIVE ARTIFICIAL INTELLIGENCE AND COPYRIGHT LAW (2023).

[226] Torrance & Tomlinson, *supra* note 225, at 236; James Vincent, *The Scary Truth About AI Copyright is Nobody Knows What Will Happen Next*, THE VERGE (Nov. 15, 2022, 9:00 AM), https://tinyurl.com/56xtcusr [https://perma.cc/C849-MUF9] (discussing the arguments in favor of the fair use defense for AI).

training can lead to significantly similar or identical works, which would constitute infringement.[227]

Scholars Mark Lemley and Bryan Casey provide a strong position for the fair training argument—that use of copyrighted work for AI training should generally be considered fair use:

> ML systems should generally be able to use databases for training, whether or not the contents of that database are copyrighted. There are good policy reasons to do so. First, we need to encourage people to compile new databases and to open them up for public scrutiny or innovation . . . . Second, an ML system's use of the data often is transformative as that term has come to be understood in copyright law, because even though it doesn't change the underlying work, it changes the purpose for which the work is used. And because training sets are likely to contain millions of different works with thousands of different owners, there is no plausible option simply to license all of the underlying photographs, videos, audio files, or texts for the new use. So allowing a copyright claim is tantamount to saying, not that copyright owners will get paid, but that the use won't be permitted at all, at least without legislative intervention. [228]

Some observers have proposed a new copyright exception explicitly permitting fair training. As Torrance and Tomlinson argue:

> An overriding purpose of fair use or fair dealing is to ensure that society benefits from the copyright system . . . . We propose that AI offers tremendous potential benefits for society. These benefits may be maximized by exposing AI to vast training sets that include works protected by copyright.[229]

In this view, a flexible fair training analysis, akin to current fair use tests (but customized for AI), would help strike the right balance.

The Digital Millennium Copyright Act ("DMCA") provides safe harbor provisions that may protect certain web scraping or caching activities.[230] In

---

[227] David Newhoff, *Training AI with Protected Works: Is Copyright Law Designed to Respond?*, THE ILLUSION OF MORE (July 15, 2023), https://illusionofmore.com/training-ai-with-protected-works-is-copyright-law-designed-to-respond/ [https://perma.cc/WP35-ZJ4U].

[228] Lemley & Casey, *supra* note 27, at 748–49.

[229] Torrance & Tomlinson, *supra* note 225, at 252–53.

[230] Digital Millennium Copyright Act of 1998, Pub. L. No. 105-304, 112 Stat. 2860.

particular, Section 512 of the DMCA, known as the Online Copyright Infringement Liability Limitation Act, provides safe harbors for online service providers that limit their liability for copyright infringement under specific circumstances.[231]

One relevant safe harbor is found in Section 512(b), which covers system caching.[232] This provision offers protection for intermediate and temporary storage of material on a system or network, as long as certain conditions are met.[233] These conditions require that the material is made available online by a person other than the service provider, transmitted at the direction of that person, and stored through an automatic technical process for the purpose of making the material available to users of the system or network.[234] Additionally, the service provider must comply with rules concerning the refreshing, reloading, or updating of the material, and must not modify the content nor interfere with the ability of technology to return data to the original site.[235]

While the system caching safe harbor does not explicitly mention web scraping, the process appears clearly protected so long as the other elements of the caching safe harbor are found. In *Field v. Google Inc.*, the District of Nevada held that Google's caching of web pages for use in its search engine qualified for the Section 512(b) safe harbor.[236] The court noted that Google's caching met the statutory requirements and served important public interests, such as enabling access to otherwise unavailable web pages and reducing internet congestion.[237]

However, the training of generative AI models, also known as GenAI, using scraped web data typically goes beyond the limitations set by the Section 512(b) safe harbor. In particular, the safe harbor requires that the storage of copyrighted work be done for a particular purpose: "for the purpose of making the material available to users of the system" who are directed toward requesting access to the material directly from its source.[238] Congress designed this portion of DMCA to protect search engines, as some uses of AI are for this search engine purpose, such as Google's Gemini and Perplexity.AI, which both provide links to the original sources of the information they present.[239]

---

[231] *Id.* § 512.

[232] *Id.*

[233] *Id.*

[234] *Id.*

[235] *Id.*

[236] Field v. Google Inc., 412 F. Supp. 2d 1106, 1109 (D. Nev. 2006).

[237] *Id.* at 1124.

[238] 17 U.S.C. § 512(b)(1)(C).

[239] GEMINI, https://gemini.google.com/ [https://perma.cc/QY52-RFDW] (last visited June 14, 2024). Gemini is a suite of large language models from Google AI, designed to understand and generate text, code, and other formats to help with writing, planning, and learning. *Id.* In some of its forms, it provides answers with links to relevant web sources. *Id.* Gemini provided the following response on May 28, 2024,

Most generative AI models, such as OpenAI's GPT series, DeepMind's Chinchilla, and Anthropic's Claude, are trained on vast amounts of web-scraped data without the explicit purpose of directing users back to the original sources.[240] Instead, these models aim to generate coherent and contextually

---

to the following prompt: "Briefly explain how Gemini obtains real time data from the web and incorporates that data into responses." *Id.* Response: Gemini can access real-time data through connections with other Google apps like Maps, Flights, and YouTube. *Id.* When you ask a question that might be relevant to one of these extensions, Gemini can pull information in real-time to give you the most up-to-date answer. *Id.* It essentially acts like a bridge between your question and the relevant app's data, delivering the information directly in its response. *Id.* It is important to note that Gemini itself doesn't directly crawl the web for real-time data. *Id.* It relies on pre-built connections with these specific Google apps. *Id.* Perplexity.AI is an AI-powered search engine that uses natural language processing to provide comprehensive answers to your questions, drawing from various sources and allowing follow-up questions for deeper exploration. PERPLEXITY, https://perplexity.ai [https://perma.cc/5REX-VJEV] (last visited June 14, 2024). Perplexity provided the following response on May 28, 2024, to the following prompt: "Briefly explain how Perplexity.AI obtains real time data from the web and incorporates that data into responses." *Id.* Response: Perplexity.AI utilizes a unique approach to obtain real-time data from the web and incorporate it into its responses. *Id.* It combines conversational AI capabilities with live web search to provide up-to-date and contextually relevant information. *Id.* Perplexity.AI employs web crawlers that continuously scan and index websites, capturing the latest content and updates from across the Internet. Jonathan Gillham, *Perplexity AI, The Game-Changer in Conversational AI and Web Search*, ORIGINALITY.AI (May 10, 2024), https://originality.ai/blog/perplexity-ai-statistics [https://perma.cc/EJ9D-2XHA]. When a user submits a query, Perplexity.AI performs a live search on the indexed web data, retrieving the most recent and relevant information related to the query. *Perplexity AI Integration for Make*, SYNERGETIC, https://www.go-synergetic.com/apps/perplexity-ai [https://perma.cc/6KMV-PQY6] (last visited June 15, 2024). Perplexity.AI utilizes advanced natural language processing techniques, similar to those used by ChatGPT, to understand and interpret the user's query. Gillham, *supra* note 239. The AI model analyzes the search results, extracts key information, and synthesizes a coherent and contextual response that incorporates the real-time web data. *Centralize Data and Automate Business Processes with Perplexity AI*, BOOST.SPACE AI, https://boost.space/integrations/perplexity-ai/ [https://perma.cc/DM9E-SJM6] (last visited June 15, 2024); *Perplexity AI Integration for Make*, *supra* note 239. Perplexity.AI presents the response in a conversational format, allowing users to engage in a dialogue and ask follow-up questions based on the provided information. Gillham, *supra* note 239. By seamlessly integrating web search capabilities with conversational AI, Perplexity.AI can deliver up-to-the-minute information and insights, ensuring that users receive the most current and relevant data available on the Internet. *Centralize Data and Automate Business Processes With Perplexity AI*, *supra* note 239; *Perplexity AI Integration for Make*, *supra* note 239; Gillham, *supra* note 239.

[240] Shennind Awat-Ranai, *A Modern Marriage: How AI Powered By Blockchain Could Protect IP Rights*, CHARLES RUSSELL SPEECHLYS (Mar. 12, 2024), https://www.charlesrussellspeechlys.com/en/insights/expert-insights/intellectual-property/2024/a-

relevant text based on the patterns and knowledge gleaned from the training data.[241] The generated text is often presented as the model's own output, without direct attribution or links to the original sources.[242]

Consequently, the use of copyrighted material in the training of these generative AI models is unlikely to fit under the Section 512(b) safe harbor. The primary purpose of storing and processing data in the training of generative AI models is not to direct users to the original sources but to enable the model to generate new text based on the patterns it has learned.[243] Thus, generative AI developers and providers can rely only upon the Section 512(b) safe harbor in limited circumstances, such as when the AI model is specifically designed and used as a search engine to direct users to the original sources of the information it presents.

2.The Potential Impact of Successful Web Scraping Lawsuits

Recent copyright infringement lawsuits filed against companies like Stability AI, Midjourney, and OpenAI allege that unauthorized web scraping of copyrighted material coupled with publicly available content to train AI models could have profound implications for future AI development if the plaintiffs prevail.[244] In these lawsuits, the AI companies assert that the web data is transformed in the AI model training process and the model outputs are distinct from the training data.[245] At the same time the copyright owners contend this is a massive misappropriation of copyrighted works.[246] An injunction barring the AI companies' use of web-scraped data to train AI could fundamentally alter the trajectory of AI development going forward.[247]

If the courts rule against fair use and issue an injunction barring web scraping for AI training, it would force a significant change in the AI field. With the leading AI models relying heavily on massive web-scraped datasets,

---

modern-marriage-how-ai-powered-by-blockchain-could-protect-ip-rights/ [https://perma.cc/MDW2-KFRU].

[241] *Id.*

[242] *Id.*

[243] *What is Generative AI*, NVIDIA, https://www.nvidia.com/en-us/glossary/generative-ai/ [https://perma.cc/FDB2-PNSP] (last visited June 15, 2024).

[244] *Status Of All 24 Copyright Lawsuits v. AI Companies*, CHATGPT IS EATING THE WORLD (May 31, 2024), https://Chatgptiseatingtheworld.Com/2024/05/31/Status-Of-All-24-Copyright-Lawsuits-V-Ai-Companies-May-31-2024-Nyt-Is-Willing-To-Give-Up-Use-Of-Exhibit-J-At-Trial-What/ [https://perma.cc/DU7B-GV6J].

[245] Benjamin L. W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45, 61–64 (2017) (analyzing whether training AI on copyrighted works is fair use).

[246] *Id.*

[247] Nilay Patel, *How AI Copyright Lawsuits Could Make the Whole Industry Go Extinct*, THE VERGE (Feb. 15, 2024), https://www.theverge.com/24062159/ai-copyright-fair-use-lawsuits-new-york-times-openai-chatgpt-decoder-podcast [https://perma.cc/6T39-TAVA] (exploring the high stakes and potential consequences of copyright battles between AI companies and content creators).

an injunction would cause such AI companies to fundamentally overhaul their data collection practices. Denying the ability of web scraping would stifle innovation and put the United States at a disadvantage in the AI race. One solution is for AI companies to obtain licenses from copyright holders to use their data for training.[248] However, this could be impracticable given the sheer volume and variety of data required, and the difficulty of identifying and negotiating with all the copyright owners.[249] This licensing solution would at a minimum add substantial friction in the form of transaction costs. Common copyright law solutions to these collective action problems include compulsory licensing schemes as well as collective rights organizations.[250] Although highly complex, this proposal could be workable.[251]

Perhaps more likely in the event of courts granting copyright owners the power to stop or tax web scraping for AI training is that AI companies may shift to relying primarily on public domain data, Creative Commons licensed works, and synthetic data.[252] But, limiting training data in this way could

---

[248] Lemley & Casey, *supra* note 27, at 776–77 (discussing potential shift to licensing training data if fair use does not apply).

[249] Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291, 300 (2019) (noting transaction costs of licensing for text data mining).

[250] Levendowski, *supra* note 19, at 617–18.

[251] *Id.* at 629–30 (discussing and noting challenges with compulsory licensing for AI training data); *Examining the Possibility of Compulsory Copyright Licensing for LLM Training*, JD SUPRA (Feb. 28, 2024), https://www.jdsupra.com/legalnews/examining-the-possibility-of-compulsory-2775732/ [https://perma.cc/93UX-GC5G] (analyzing the potential for a compulsory copyright licensing scheme to address the use of copyrighted works in generative AI). Compulsory licensing scheme for generative AI inputs could be a "sensible solution" as it would solve issues of copyright infringement and output ownership. Letter from Jessica Fjeld & Mason Kortz, Assistant Director & Clinical Instructor, Brooklyn L. Sch., to Andrei Iancu, Dir. U.S. Pat. & Trademark Office (Jan. 10, 2020). The automatic nature of such a license would eliminate the need for individual agreements. *Id.*; *c.f.* Letter from David Hansen & Rachel Brooke, Executive Director & Senior Staff Attorney, Authors Alliance, to Suzanne Wilson, Gen. Couns. & Assoc. Reg. Copyrights, U.S. Copyright Office (Oct. 30, 2023) (arguing that neither compulsory licensing nor collective licensing schemes are logistically feasible or sound solutions for authorizing the use of copyrighted works in AI training datasets). The scale and complexity of the datasets, as well as the "orphan works problem," make such schemes unworkable. *Id*.

[252] Levendowski, *supra* note 19, at 617. Here, I use public domain data to refer to works that are not protected by intellectual property laws – particularly copyright. These works are considered to be owned by the public at large and can be freely used without permission or attribution. *Id.* Creative Commons licenses are a set of standardized copyright licenses that allow creators to communicate which rights they reserve and which rights they waive for the benefit of other creators or the public. *See id.* These licenses range from very open (allowing any use with attribution) to more restrictive (allowing only non-commercial use with no modifications). Synthetic data, in the context of AI training, refers to artificially generated data that mimics the statistical properties and patterns of real-world data. Abid Ali Awan, *What is Synthetic Data?*, DATACAMP (Jul. 2023), https://www.datacamp.com/blog/what-is-synthetic-

significantly curtail the capabilities of AI models. Much of the knowledge and culture of the past century is still protected under copyright law and would be off limits. The models would have a narrower and antiquated understanding of the world based on (much) older or limited data. While synthetic data avoids copyright issues, it may produce models with reduced capabilities across various tasks and that are more detached from the real world.[253] Much of the current AI power comes from ingesting the collective intelligence embedded across the web, thus moving to synthetic data loses a vast amount of actual human knowledge and culture.

An injunction could also chill AI research and concentrate power among a few large players able to overcome the legal hurdles. Training data could become a scarce resource as leading companies may monopolize licenses from major copyright holders, effectively shutting smaller firms and researchers out. This could slow the overall pace of AI innovation. An injunction could also encourage some AI development to move to countries with weaker copyright protections or enforcement. China and Japan, for example, have more permissive fair use and data mining exceptions than the United States.[254] Thus, companies may find it easier to train models overseas.[255] On the other hand, successful lawsuits against these AI companies for impermissibly using scraped content would help protect authors' rights, prohibit the unauthorized use of their content by AI companies, and help provide them with compensation for their work.

The bottom line here is that a strong rights-protective approach in the copyright space has the potential of creating a stifling regulatory hurdle by cutting off access to a major underlying resource used for AI development—training data.

---

data [https://perma.cc/E8X2-G52P]. This data is created by algorithms rather than collected from real-world sources, potentially avoiding copyright issues. *Id.* However, the legal status of synthetic data is still evolving, particularly when it is generated based on copyrighted works. *See id.*

[253] Sag, *supra* note 249, at 344 (discussing limitations of synthetic data).

[254] *International Copyright Issues and Artificial Intelligence* (International AI Copyright Webinar July 26, 2023), accessible at https://www.copyright.gov/events/international-ai-copyright-webinar/International-Copyright-Issues-and-Artificial-Intelligence.pdf [https://perma.cc/4Q7V-VEM9].

[255] Brian Nussbaum, *Offshore: The Coming Global Archipelago of Corrosive AI*, LAWFARE (June 14, 2023, 4:00 AM), https://www.lawfaremedia.org/article/offshore-the-coming-global-archipelago-of-corrosive-ai [https://perma.cc/5YAT-ZW94] (arguing that regulating artificial intelligence effectively will be extremely challenging due to the potential for "jurisdictional arbitrage" and "offshoring" of AI operations, similar to the difficulties in controlling money laundering and financial crime).

C. Copyright's Legitimate Role at Channeling Creative Behavior and Ensuring Just Rewards

Copyright law can serve to encourage and channel creative behavior in the context of AI-generated works. By granting exclusive rights to the creators of original works, copyright incentivizes individuals and organizations to invest time and resources into developing innovative AI systems that can produce novel and valuable outputs.[256] In addition to incentivizing creativity, copyright law can ensure that the creators of AI-generated works receive a fair share of the economic benefits derived from their creations. This is particularly important given the potential for AI systems to generate outputs with significant commercial value. While the outputs generated by AI systems may not always fit neatly into traditional copyright categories, there are important ways in which copyright law can channel creative behavior and guarantee just rewards in the AI space.

1. Protecting the Building Blocks of AI Systems

One way that copyright law can support AI development is by protecting the various components that make up an AI system, including the software code. In addition to the software itself, compilations of data used to train an AI model may be eligible for copyright protection.[257] A crucial question is whether the selection and arrangement of the data reflects a sufficient degree of creativity to merit protection.[258] While facts themselves are not copyrightable, a curated dataset that involves creative choices in what to include and how to organize the information could potentially qualify for protection as a compilation.[259]

2. Protecting AI-Generated Outputs and its Unavoidable Challenges

While copyright can protect the building blocks of an AI system, AI-generated outputs present a more complex question. A string of recent court decisions and administrative rulings in the United States, Europe, and China have grappled with whether AI-generated works are original works of authorship entitled to copyright protection.[260]

---

[256] Levendowski, *supra* note 19, at 625 (discussing how copyright law can incentivize the creation of new AI systems).

[257] *See* Feist Publ'ns, Inc. v. Rural Tel. Serv., 499 U.S. 340, 345 (1991).

[258] Monique W. Macek & Meena Seralathan, *Benefits of Using Copyrights to Protect Artificial Intelligence and Machine Learning Inventions*, MINTZ (July 11, 2022), https://www.mintz.com/insights-center/viewpoints/2231/2022-07-11-benefits-using-copyrights-protect-artificial [https://perma.cc/NS5U-K8PP].

[259] *Id.*

[260] Patrick Zurth, *Artificial Creativity? A Case Against Copyright Protection for AI-Generated Works*, 25 UCLA J.L. & TECH. 1, 1, 4–5 (2021).

The consensus among these rulings is that AI systems, on their own, cannot be recognized as the "authors" of copyrightable works.[261] However, this does not necessarily mean that all AI-generated outputs are categorically excluded from copyright protection. The degree of human involvement and creative direction in the development and use of an AI system can affect the copyrightability of the system's outputs. Thus, true human-AI collaboration will likely be eligible for copyright protection if there is a sufficient degree of human creativity and decision-making involved in the process.

D. The Limited Role of Copyright in Protecting Privacy Rights

While copyright can play a role in channeling creative behavior and ensuring just rewards for AI-generated works, it is a poor fit for addressing other concerns, such as privacy issues or job displacement. This Part explores the limitations of using copyright law to regulate AI through the lens of two pressing privacy concerns—revenge porn and deepfakes.

1. The Example of Revenge Porn

As discussed above, a fear of AI is that it may be used to violate individual privacy rights. A prime example of this, and an issue predating AI, is the phenomenon of nonconsensual pornography distribution, also known as revenge porn, where intimate images or videos are shared online without the consent of the individual(s) depicted.[262] In an attempt to combat revenge porn, some victims have turned to copyright law, arguing that the unauthorized sharing of their intimate images constitutes copyright infringement.[263] However, this approach has several limitations.[264]

In many cases, the victim of revenge porn may not own the copyright in the shared images or videos, as they may have been created by another person, such as an ex-partner.[265] Even when the victim does own the copyright, the protection offered by copyright law is limited to the specific expression of the work and does not extend to the underlying facts or ideas.[266] This means that

---

[261] *Id*. at 5.

[262] Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345, 346 (2014) (defining revenge porn as "the distribution of sexually graphic images of individuals without their consent").

[263] Amanda Levendowski, *Using Copyright to Combat Revenge Porn*, 3 N.Y.U. J. INTELL. PROP. & ENT. L. 422, 442–43 (2014) (discussing the use of copyright law to remove nonconsensual pornography from the Internet).

[264] Eric Goldman & Jessica Silbey, *Copyright's Memory Hole*, 2019 B.Y.U. L. REV. 929, 983 (2019) ("[C]opyright may be ill-designed to redress the depicted person's paramount privacy interests.").

[265] Derek E. Bambauer, *Exposed*, 98 MINN. L. REV. 2025, 2028–29 (2014) (noting that in many cases, the creator of the intimate content, rather than the subject, holds the copyright).

[266] Cathay Y. N. Smith, *Weaponizing Copyright*, 35 HARV. J.L. & TECH. 193, 213 (2022) (discussing the idea/expression dichotomy in copyright law).

while copyright may be used to remove specific images or videos from unauthorized use, it cannot prevent the spread of the private information contained within them. Fundamentally, copyright law is designed to promote creativity and innovation, not protect privacy rights.[267] As a result, using copyright to address revenge porn often requires victims to navigate complex legal doctrines that are not tailored to their specific needs and concerns.

2. Deepfakes and the Limits of Copyright

The rise of AI technologies further exacerbates the problem of nonconsensual pornography, as AI can creatively generate new images known as deepfakes, which are AI-generated images, videos, or audio recordings that depict individuals engaging in fictitious activities.[268] Deepfakes raise significant privacy concerns, as they can create highly realistic and damaging false depictions of individuals without their consent. However, using copyright law to combat deepfakes would be similarly misguided for several reasons.

In many cases, deepfakes may not contain any copyrightable subject matter, as they are entirely generated by AI algorithms and do not involve any original creative expression by a human author.[269] Even if a deepfake does contain copyrightable elements, the use of those elements may be protected under the fair use doctrine, which allows for the use of copyrighted material for purposes such as criticism, commentary, or parody.[270] Given the transformative nature of deepfakes and their potential for social commentary, many uses of these technologies would fall within the scope of fair use, as discussed above in Part III. There are, arguably, many positive uses of deepfake technology that do not involve identity theft of a living individual or nonconsensual pornography. For instance, the Dalí Museum used deepfakes to recreate Salvador Dalí, enabling visitors to interact with the artist and take selfies with him—an arguably positive use of deepfake technology.[271] Similarly, Snoop Dogg used deepfakes to bring Tupac back to "life" as a tribute in a music video.[272] Deepfake technology also brought back Anthony

---

[267] Andrew Gilden, *Sex, Death, and Intellectual Property*, 32 HARV. J.L. & TECH. 67, 68 (2018) (arguing that copyright law is not well-suited to protect privacy interests).

[268] Elizabeth A. Rowe, *Regulating Facial Recognition Technology in the Private Sector*, 24 STAN. TECH. L. REV. 1, 38–39 (2020) (discussing the emergence of deepfake technology).

[269] Lindsey Barrett, *Rejecting Test Surveillance in Higher Education*, 2022 MICH. ST. L. REV. 675, 702–03 (2022) (noting that AI-generated content may not be eligible for copyright protection); *see* Thaler v. Vidal, 43 F.4th 1207, 1210 (Fed. Cir. 2022) (holding that the Patent Act requires individuals to be human beings).

[270] 17 U.S.C. § 107 (codifying the fair use doctrine).

[271] Millie Chow, *What Are The Positive Applications of Deepfakes?*, JUMPSTART (June 9, 2022), https://www.jumpstartmag.com/what-are-the-positive-applications-of-deepfakes/ [https://perma.cc/9HPD-GETA].

[272] *Id.*

Bourdain's voice to add context and depth to the biography of his life.[273] Deepfakes are also useful for providing personalized recommendations; for virtual try-on experiences tailored to individual customers; and for interactive training materials.[274] Just as the United States Supreme Court held in *Sony Corp. of America v. Universal City Studios, Inc.* that the manufacturers of videotape recorders were not liable for copyright infringement because the devices were "capable of substantial non-infringing uses,"[275] the creators of deepfake algorithms and tools would likely be shielded from liability as long as their technologies' primary purpose is a lawful one. However, the Court's later decision in *MGM Studios, Inc. v. Grokster, Ltd*. suggests that the creators of deepfake tools could potentially face contributory liability if they actively encourage users to engage in copyright infringement, even if the technologies are capable of substantially lawful uses.[276] Despite their difficulties, there are strong supporters of copyright use as an AI regulatory tool, including in the deepfakes arena.[277]

3. Encouraging AI Innovation

The desire for regulations not to inhibit or stifle AI innovation, but instead encourage beneficial AI development, is a critical consideration for policymakers. Regarding the previously discussed concerns for transparency, many have noted that full model explainability may come at the cost of innovation.[278] Certainly, categorical bans on certain technologies (such as

---

[273] Helen Rosner, *The Ethics of a Deepfake Anthony Bourdain Voice*, THE NEW YORKER (June 17, 2021), https://www.newyorker.com/culture/annals-of-gastronomy/the-ethics-of-a-deepfake-anthony-bourdain-voice [https://perma.cc/2Q5W-KEXN].

[274] Chow, *supra* note 271; *Video Personalization Using Deepfake Technology: Are Deepfakes All Evil?*, MAVERICK (Jan. 17, 2024 https://www.trymaverick.com/blog-posts/are-deep-fakes-all-evil-when-can-they-be-used-for-good [https://perma.cc/H993-V8SC]; *Deepfakes for Good? How Synthetic Media Is Transforming Business*, TECH INFORMED (Oct. 5, 2023), https://techinformed.com/deepfakes-for-good-how-synthetic-media-is-transforming-business/ [https://perma.cc/8BNX-FT99].

[275] Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 442 (1984).

[276] MGM Studios Inc. v. Grokster, Ltd., 545 U.S. 913, 936-37 (2005) ("[O]ne who distributes a device with the object of promoting its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third parties.").

[277] Levendowski, *supra* note 19, at 579 (2018) (asserting that copyright can create fairer artificial intelligence for women, queer people, people of color, and other marginalized people).

[278] As Sara Gerke, Assistant Professor of Law at Pennsylvania State University, Dickinson, notes, "[I]f there is sufficient proof that a black-box model performs better than a white-box model and is reasonably safe and effective, and the accuracy increase outweighs the loss of model interpretability, then regulators should generally permit marketing of the black-box AI/ML model as such (without requiring explainable AI/ML) to facilitate innovations." Sara Gerke, *Health AI for Good Rather Than Evil? The Need for A New Regulatory Framework for AI-Based Medical Devices*, 20 YALE J. HEALTH POL'Y, L. & ETHICS 432, 491 (2021) (arguing that regulators should allow

non-explainable AI) can inhibit innovation within that market space.[279] But, innovation investment is not like a light switch. Rather, because innovation is subject to comparative advantage analysis, bans in one area may channel investment to other areas.

The theory that a burdensome regulatory approach inhibits innovation is a significant challenge in implementing AI governance frameworks.[280] This fear of "stifling innovation" is a common theme in many legislative and regulatory debates, often framed as a tradeoff between innovation and other policy objectives, like consumer protection.[281] But, sometimes legal certainty and predictability are the more critical aspects needed to stimulate investment in innovation. Professor Anat Lior suggests that "[a] strict liability regime may even encourage innovation, if it provides manufacturers with a certain, predictable legal framework to operate in."[282] In this model, regulators should strive to create legal frameworks that provide legal clarity and certainty for AI developers and users while preserving business flexibility to adapt to new technological capabilities and challenges.

When properly calibrated, IP rights can promote innovation by providing incentives for investment and enabling follow-on development. Since the United States' founding, the theory of IP rights has been to "promote the progress" of science and the useful arts.[283] But, it is not entirely clear what

black-box AI models if they outperform explainable models and are safe and effective, to promote innovation). Similarly, Ashley Deeks cautions that "xAI may also stifle innovation, force developers to reveal trade secrets, and impose high monetary costs because xAI can be expensive to build." Ashley Deeks, *The Judicial Demand for Explainable Artificial Intelligence*, 119 COLUM. L. REV. 1829, 1834 (2019) (warning that requiring explainable AI could hinder innovation and be costly for developers).

[279] Regulators should avoid imposing categorical bans or prohibitions that might stifle AI innovation. Dr. Asress Adimi Gikay argues that "[a]ny regulatory authority that is anxious about pervasive machine learning decisions in the credit industry should reject the temptation to impose a categorical ban or a prohibition that might stifle innovation." Dr. Asress Adimi Gikay, *The American Way—Until Machine Learning Algorithm Beats the Law?*, 12 CASE W. RES. J.L. TECH. & INTERNET 1, 50 (2021) (cautioning against overly restrictive regulations of AI in the credit industry).

[280] The Office of Management and Budget recently issued a proposed memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. Proposed Memorandum from Sholanda D. Young for the Heads of Executive Departments and Agencies, https://www.whitehouse.gov/wp-content/uploads/2023/11/AI-in-Government-Memo-draft-for-public-review.pdf [https://perma.cc/KJ79-KBGN] (proposing a framework for responsible AI use by federal agencies).

[281] Lior, *supra* note 6, at 94 (arguing that a strict liability regime for AI could promote innovation by providing legal certainty for manufacturers).

[282] *Id.*

[283] U.S. CONST. art. I, § 8, cl. 8 (granting Congress the power to enact patent and copyright laws to promote progress).

impact IP rights will have on AI innovation, or vice versa.[284] While patents, trade secrets, and copyright may protect AI systems themselves, the training inputs are often owned by a distributed group of rights holders, and the outputs are largely unprotectable by IP rights.[285] Depending on the ultimate distribution of rights, the inputs may either foster concentration or widen the distribution of rights in the AI field.

IP rights are a form of soft regulation because they provide incentives for creators and innovators without directly mandating specific behaviors. Even once distributed, IP rights are subject to marketplace negotiation and are rarely broad enough to capture entire market segments, effectively preventing monopolies.[286] Looking at the financial innovation marketplace, Chris Brummer and Yesha Yadav suggest that similar "informal methods can offer a means of facilitating innovation."[287]

The global competition in AI development is also a significant factor driving policy discussions around AI regulation and innovation. National regulators may be focused on promoting AI innovation and development within their national sphere—while deterring such advancements in competing nations. As Senator Cruz recently stated, "Look, there is a global race for AI, and it is a race we are engaged in with China . . . we need to make sure America is leading the AI revolution."[288] The concern in the United States is that overly burdensome regulations could reduce America's competitive advantage and stifle local innovation—shifting investments offshore without the same regulatory burden.

E. Overview of Proposed AI Regulations and Frameworks

As the deployment of AI systems continues to accelerate, governments and international organizations have recognized the need for comprehensive regulations and frameworks to address the various challenges and risks associated with AI.[289] In this Part, two significant proposals for AI regulation

---

[284] Mauritz Kop, *AI & Intellectual Property: Towards an Articulated Public Domain*, 28 TEX. INTELL. PROP. L.J. 297, 321–22 (2020) (proposing limits on IP protection for datasets to promote AI innovation and investment).

[285] *See supra* Part III.C.2.

[286] *See generally* Cong. Rsch. Serv., RL34292, Intellectual Property Rights and International Trade (2020).

[287] Yadav & Brummer, *supra* note 149, at 283 (proposing that informal regulatory approaches could facilitate financial innovation).

[288] Renee Henson, *Bridging the Divide: Does the EU's AI Act Offer Code for Regulating Emergent Technologies in America?*, 89 MO. L. REV. 847, 866 (2024) (citing 169 CONG. REC. S5931-32 (daily ed. Dec. 13, 2023) (statement of Sen. Cruz)).

[289] *See generally id.* In her article, Professor Henson provides a comprehensive analysis of the European Union's Artificial Intelligence Act (AI Act) and its potential implications for AI regulation in the United States. *See id.* She examines the AI Act's risk-based approach, its key provisions for different risk categories, and the challenges and concerns raised by various stakeholders. *See id.* Professor Henson also explores

are analyzed: (1) Biden Administration's Executive Order on Promoting the Use of Trustworthy AI in the Federal Government,[290] and (2) the European Union's recently adopted AI Act.[291] This Part will also discuss relevant leading legislative proposals.

In the United States, Biden's Administration took a significant step toward AI regulation with the issuance of Executive Order 14110 on Safe, Secure, and Trustworthy AI in October 2023.[292] This Executive Order aims to promote the responsible development and use of AI within the federal government, focusing on issues such as safety, security, privacy, transparency, and non-discrimination.[293] It directs all federal agencies to assess the risks associated with their AI systems, implement appropriate safeguards, and ensure public transparency regarding their AI use.[294] Specifically, the Executive Order directs federal agencies to assess AI's potential impacts on privacy, civil rights, and civil liberties, and to ensure that AI systems are subject to appropriate oversight and accountability mechanisms.[295]

While the Executive Order is a significant development, its scope is limited to federal agencies, leaving the private sector's use of AI untouched.[296] Moreover, because it is an executive action, it lacks the force and permanence of legislation passed by Congress or even that developed through the rulemaking process of a federal agency. To create a comprehensive and lasting regulatory framework, legislative action is necessary.

That said, there have been notable bipartisan efforts in Congress to address AI regulation. In September 2023, Senators Josh Hawley (R-MO) and Richard Blumenthal (D-CT) introduced the "Bipartisan Framework," which outlines five key principles for future AI legislation.[297] These principles include establishing a licensing system for high-risk AI systems, creating legal accountability for AI-related injuries, ensuring national security and competitiveness, developing transparency requirements, and implementing safety measures for consumers and children.[298] The Bipartisan

---

recent efforts by U.S. policymakers to establish AI regulatory frameworks, such as the Bipartisan Framework for U.S. Act and the No Section 230 Immunity Act. *See id.*

[290] Exec. Order No. 14110, 88 Fed. Reg. 75191 (Nov. 1, 2023).

[291] Regulation (EU) 2023/1689 of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, art. 113 [hereinafter AI Act].

[292] Exec. Order No. 14110, 88 Fed. Reg. 75191 (Nov. 1, 2023).

[293] *Id.*

[294] *Id.*

[295] *Id.*

[296] Henson, *supra* note 288, at 868.

[297] *Blumenthal & Hawley Announce Bipartisan Framework on Artificial Intelligence Legislation*, RICHARD BLUMENTHAL: U.S. SEN. FOR CONN. (Sept. 8, 2023), https://www.blumenthal.senate.gov/newsroom/press/release/blumenthal-and-hawley-announce-bipartisan-framework-on-artificial-intelligence-legislation [https://perma.cc/B27D-JAMT].

[298] *Id.*

Framework, however, is not a fully developed bill, but rather a self-described roadmap for future legislation. It reflects a growing recognition among lawmakers for the need for comprehensive AI regulation. Translating these principles into actionable legislation will require navigating complex issues and balancing competing interests in a highly political environment.

Another notable legislative proposal is the No Section 230 Immunity for AI Act, which was also introduced by Senators Hawley and Blumenthal in June 2023.[299] This bill seeks to amend Section 230 of the Communications Decency Act to remove legal immunity for claims and charges related to generative AI.[300] The bill aims to hold AI companies accountable for harms caused by their systems by allowing private actions based upon state law.[301] In our techno-optimist environment, this legislation—like other AI regulatory approaches—faces opposition from those who argue that it could stifle innovation and put domestic companies at a disadvantage in the global AI race.

In contrast to the nascent efforts in the United States, the European Union has made significant progress toward comprehensive AI regulation with its AI Act.[302] The AI Act, recently adopted, takes a risk-based approach to AI regulation, categorizing AI systems into four risk-based tiers: unacceptable risk, high risk, general-purpose and generative AI, and limited risk.[303] The AI Act applies to all AI systems developed or deployed within the European Union, as well as those that affect European citizens.[304] Therefore, the AI Act's approach provides a global reach.

The AI Act imposes strict prohibitions on AI systems deemed to pose unacceptable risks, such as those that use subliminal techniques to distort human behavior or exploit vulnerabilities based on age, disability, or social or economic status.[305] For high-risk AI systems, such as those used in critical infrastructure, education, employment, and law enforcement, the AI Act establishes extensive requirements for risk management, data governance, transparency, human oversight, and accountability.[306] One of the more contentious aspects of the AI Act has been its treatment of general-purpose and generative AI systems, such as LLMs. The final compromise requires providers of these systems to maintain technical documentation, conduct risk assessments, and report serious incidents to authorities.[307] However, some

---

[299] S. 1993, 118th Congress (2023–2024).

[300] *Id.*

[301] *Id.*

[302] *Id*.

[303] Henson, *supra* note 288, at 851.

[304] Alex Engler, *The EU AI Act Will Have Global Impact, But a Limited Brussels Effect*, BROOKINGS (June 8, 2022), https://www.brookings.edu/blog/techtank/2021/07/12/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/ [https://perma.cc/ARW2-6NMW].

[305] AI Act, *supra* note 291, at 43–45.

[306] *Id.* at 65.

[307] *Id.* at 137.

stakeholders argue that these requirements are too burdensome and could hinder innovation and competitiveness.[308]

The AI Act also establishes a new European regulatory body, the AI Office, to oversee compliance and impose potentially steep financial penalties for violations.[309] While the Act has been praised for its comprehensive approach, concerns have surfaced about the compliance costs, potential impact on small businesses and startups, and potential stifling of AI innovation.[310]

As the United States continues to grapple with the challenges of AI regulation, it can look to the European Union's AI Act as a potential model, while also learning from its limitations and controversies. The key will be to strike a balance between promoting innovation and competitiveness while ensuring the safe, responsible, and trustworthy development and deployment of AI systems. Throughout this process of creating a regulatory system, it is important to recognize that any regulation is hampered by a major lack of AI expertise within the public sector, which is partly due to the concentration of AI talent and capabilities among a few major technology companies like OpenAI, Microsoft, NVIDIA, Anthropic, Alphabet, Meta, and IBM.[311] With few AI specialists working in government compared to these dominant firms, regulators may struggle to keep pace with rapid technological developments powered by specialized hardware, large datasets, and top engineering talent, and they may also struggle to design rules that effectively balance innovation with other policy goals.

IV. TRANSPARENCY AND TRADE SECRECY

AI systems are notoriously opaque in their operation but maintain large economic value. This suggests that trade secrets will play a major role in protecting the underlying rights of such systems. This conclusion is bolstered

---

[308] Henson, *supra* note 288, at 861.

[309] *See* Michael Veale & Frederik Zuiderveen Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, 22 COMPU. L. REV. INT. 97, 105 (2021) (noting that standardization will be a key aspect of the AI Act rulemaking); Liane Colonna, *Artificial Intelligence in Higher Education: Towards A More Relational Approach*, 8 LOY. U. CHI. J. REG. COMPLIANCE 18, 30 (2022).

[310] BENJAMIN MUELLER, HOW MUCH WILL THE ARTIFICIAL INTELLIGENCE ACT COST EUROPE? 3 (2021) (estimating compliance costs up to €400,000 per high risk product); Airlie Hilliard & Ayesha Gulley, *What Considerations Have Been Made for SMEs Under the EU AI Act?*, HOLISTIC AI (Aug. 22, 2023), https://www.holisticai.com/blog/how-will-smes-be-supported-under-the-eu-ai-act#:~:text=Indeed%2C%20to%20support%20innovation%20from,being%20more%20lenient%20regarding%20documentation [https://perma.cc/8MNM-ZDCG] (discussing European Commission impact assessment estimates for compliance costs of high-risk AI systems); Reid Blackman & Ingrid Vasiliu-Feltes, *The EU's AI Act and How Companies Can Achieve Compliance*, HARV. BUS. REV. (Feb. 22, 2024).

[311] *See supra* Part II.B.6; Lina M. Khan, Chair, Fed. Trade Comm'n, Remarks at the Economic Club of New York 7 (July 24, 2023).

by the comparative advantage that trade secrets, unlike copyright law, have by not requiring a human author or inventor to be legally protected.[312] Thus, AI-created information does not have a subject matter difficulty in the trade secrets space and is protectable as a trade secret. This Part provides a background on trade secrecy rights, discusses ways trade secrets can be used in the AI space to channel efforts, and delves into the tension between the push for transparency, explainability, and the private right of trade secrets.

A. The History and Purpose of Trade Secrecy Rights

Trade secret law has a long history in the United States, dating back to the nineteenth century. The earliest reported trade secret case was *Vickery v. Welch*, decided by the Massachusetts Supreme Judicial Court in 1837.[313] In *Vickery*, the court held that a contract for the sale of a chocolate mill included an implied promise by the seller not to disclose the secret method of making chocolate.[314] This case established the principle that trade secrets could be protected through contract law, laying the foundation for the development of trade secret doctrine.[315]

Courts continued to refine and expand the concept of trade secret protection throughout the late nineteenth and early twentieth centuries. In 1868, the Pennsylvania Supreme Court recognized that trade secrets were a form of property that could be protected from misappropriation, even in the absence of a contract.[316] This property-based view of trade secrets was further reinforced by the influential Restatement (First) of Torts, published in 1939, which defined a trade secret as "any formula, pattern, device or compilation of information which is used in one's business, and which gives him an opportunity to obtain an advantage over competitors who do not know or use it."[317] Under the modern rule, information constitutes a trade secret if it (1) derives independent economic value from not being generally known or readily ascertainable, and (2) is subject to reasonable efforts to maintain its secrecy.[318]

The United States Supreme Court affirmed the property-based conception of trade secrets in *Ruckelshaus v. Monsanto Co.*[319] The Court there held that trade secrets were a form of intangible property protected by the Takings Clause of the Fifth Amendment.[320] The Court reasoned that the

---

[312] Dennis Crouch, *Reattribution, The Poison Pill and Inventorship*, 5 BUS. ENTREPRENEURSHIP & TAX L. REV. 138, 142 (2021).

[313] Vickery v. Welch, 36 Mass. 523 (1837).

[314] *Id.*

[315] *Id.* at 524.

[316] Peabody v. Norfolk, 98 Mass. 452, 458 (1868).

[317] RESTATEMENT (FIRST) OF TORTS § 757 cmt. b (AM. LAW INST. 1939).

[318] As defined in the Define Trade Secrets Act of 2016, 18 U.S.C. § 1839(3).

[319] Ruckelshaus v. Monsanto Co., 467 U.S. 986 (1984).

[320] *Id.* at 1003–04.

owner of a trade secret has a "property right" in the information, which derives its value from the competitive advantage it provides.[321]

The *Ruckelshaus* decision solidified the IP status of trade secrets, distinct from contract-based or tort-based theories of protection. This property-based view of trade secrets has important implications for their use and enforcement. As property, trade secrets are bought, sold, and licensed like other forms of IP.[322]

However, the property analogy is not perfect, and trade secrets have some unique characteristics that distinguish them from other forms of IP. Unlike patents and copyrights, which are granted by the government and have a fixed term of protection, trade secrets were traditionally a creature of state law that can theoretically last forever, as long as the information remains secret.[323] Trade secret protection also does not require any formal registration or examination process; it arises automatically when the criteria for protection are met.[324] Additionally, trade secrets are vulnerable to reverse engineering, meaning that others may legally discover the secret by examining the product or information and working backwards to determine how it was created or developed.[325] A trade secret misappropriation action requires a showing of improper acquisition, use, or disclosure of the trade secret by another party.[326]

The main purpose of trade secret law is to encourage innovation and promote commercial ethics, such as preventing unfair competition. By providing legal protection for confidential business information, trade secret law allows companies to invest in research and development without fear that competitors will steal their developed knowledge.[327] This protection, in turn, promotes the development of new products, processes, and services that benefit consumers and drive economic growth.[328]

At the same time, trade secret law helps to maintain standards of commercial ethics by discouraging the theft of proprietary information through improper means such as bribery, espionage, or breach of a duty of confidentiality.[329] By imposing liability for misappropriating trade secrets,

---

[321] *Id.* at 1002.

[322] Mark. A. Lemley, *The Surprising Virtues of Treating Trade Secrets as IP Rights*, 61 STAN. L. REV. 311, 324–26 (2008).

[323] *Id.* at n.48.

[324] UNIFORM TRADE SECRETS ACT § 1 cmt. (UNIF. LAW COMM'N 1985) (stating trade secret rights arise when information meets the statutory definition, without any procedural formalities).

[325] *Id.*

[326] *Id.* § 1(2).

[327] Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 480–81 (1974).

[328] *Id.* at 480.

[329] E.I. duPont deNemours & Co. v. Christopher, 431 F.2d 1012, 1016 (5th Cir. 1970).

the law deters unethical behavior and promotes fair competition in the marketplace.[330]

In recent years, the importance of trade secret protection has significantly grown, as businesses increasingly rely on confidential information and strategize to gain a competitive edge. The rapid pace of technological change and the globalization of markets have made it easier than ever for trade secrets to be stolen and exploited by rivals.[331] In response, Congress passed the Defend Trade Secrets Act of 2016 ("DTSA"), which created a federal civil cause of action for trade secret misappropriation.[332] The DTSA was intended to provide a uniform and predictable framework for trade secret litigation across the United States, and to provide trade secret owners access to the federal courts.[333]

Despite its importance, trade secret law has been the subject of criticism and debate. Commentators argue that the lack of formal registry and the potential for indefinite secrecy of crucial information actually hinders innovation and competition.[334] Others raise concerns about the potential for trade secret law spurring concealed wrongdoing or stifled whistleblowing and public discourse.[335]

These concerns have led to calls for reform and efforts to strike a better balance between the interests of trade secret owners and the public interest in disclosure and accountability. For example, the recent federally enacted DTSA includes a whistleblower immunity provision that protects individuals who disclose trade secrets to the government or an attorney for the purpose of reporting suspected illegal conduct.[336] Some states have also enacted laws that limit the use of nondisclosure agreements to silence employees from disclosing information about sexual harassment or other misconduct.[337]

B. Trade Secrecy Rights and AI

Many aspects of AI systems, particularly those utilizing complex ML algorithms, meet the two-step test for trade secrecy protection: economic

---

[330] RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 39 cmt. a (AM. LAW INST. 1995).

[331] H.R. REP. NO. 114-529, at 3 (2016) (discussing the growing importance of trade secrets in the modern economy and the need for stronger protection).

[332] 18 U.S.C. § 1836(b).

[333] S. REP. NO. 114-220, at 3 (2016) (explaining the purpose of the DTSA as providing a federal cause of action for trade secret misappropriation).

[334] Robert G. Bone, *A New Look at Trade Secret Law: Doctrine in Search of Justification*, 86 CAL. L. REV. 241, 266–67 (1998) (critiquing the justifications for trade secret law and arguing that it can impede the dissemination of valuable information).

[335] David S. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure*, 59 FLA. L. REV. 135, 177–81 (2007) (calling for transparency to outweigh rights of secrecy).

[336] 18 U.S.C. § 1833(b).

[337] *See, e.g.*, CAL. CIV. PROC. CODE § 1001.

value from secrecy criteria and reasonable efforts to maintain secrecy.[338] The proprietary algorithms, particular training data, and system architecture used in AI often provide a competitive advantage and are carefully guarded by companies.[339]

One reason AI is well-suited for trade secret protection is the inherent complexity and opacity of many ML models. Unlike traditional software, where the underlying code can be examined and understood by skilled programmers, deep learning systems often make decisions based on complex neural networks that are essentially black boxes.[340] This difficulty in the transparency arena is converse to trade secrecy protections—the black box of AI operation provides built in measures to protect trade secrets. Even the developers of these systems may not fully understand how the AI arrives at its outputs, making reverse engineering extremely difficult.[341] Further, unlike the copyright and patent law arenas, trade secrecy has never required a human innovator or creator of the information.[342] This "natural" opacity of AI algorithms provides a strong argument for trade secret protection, as it inherently makes the technology difficult to copy or reverse engineer.[343] Companies can, and typically do, bolster their trade secret rights by implementing confidentiality measures, such as employee non-disclosure agreements as well as physical and cyber access controls. However, the very features that make AI well-suited for trade secret protection also create significant challenges for regulatory goals of transparency and explainability in AI systems.

In many ways, patent rights and trade secret rights represent opposing approaches to IP protection. Disclosure is the key *quid pro quo* requirement for obtaining a patent, while trade secret law requires efforts to maintain secrecy. There are many reasons for an innovator to choose one IP scheme over the other. Recent developments in patent subject matter eligibility jurisprudence have made trade secret protection increasingly appealing for many AI innovations. In particular, the United States Supreme Court's decisions in *Alice Corp. v. CLS Bank* and *Mayo Collaborative Services v. Prometheus Laboratories* have significantly narrowed the scope of patent

---

[338] *See supra* Part IV.A.

[339] Jeanne C. Fromer, *Machines as the New Oompa-Loompas: Trade Secrecy, the Cloud, Machine Learning, and Automation*, 94 N.Y.U. L. REV. 706, 717–18, 722–23, 727–28 (2019); *see generally* Gregory Gerard Greer, *Artificial Intelligence and Trade Secret Law*, 21 UIC REV. INTELL. PROP. L. 252 (2022).

[340] Burrell, *supra* note 133, at 3 (discussing the challenges of understanding the decision-making processes of complex deep learning models).

[341] Will Knight, *The Dark Secret at the Heart of AI*, MIT TECH. REV. (Apr. 11, 2017), https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/ [https://perma.cc/W34W-GJFA] (noting that a system's creators may not understand how it makes decisions).

[342] *See* Crouch, *supra* note 312.

[343] W. Nicholson Price II & Arti K. Rai, *Clearing Opacity Through Machine Learning*, 106 IOWA L. REV. 775, 791 (2021) (discussing how complex algorithms create natural trade secrets).

eligible subject matter available under patent law.[344] This has created challenges and uncertainty for patenting AI-related inventions, as many fall within the judicial exceptions to patentability carved out by the *Alice*/*Mayo* framework—laws of nature, natural phenomena, and particularly abstract ideas.[345]

Under the two-step test articulated in *Alice*/*Mayo*, courts first analyze a patent claim to determine it is directed to one of those ineligible concepts.[346] If so, courts must then consider the claim's elements to determine whether they contain an "inventive concept" sufficient to transform the ineligible concept into a patent-eligible application.[347] However, the contours of what constitutes an "abstract idea" remain somewhat nebulous, and determining whether there is an inventive concept has proven highly unpredictable in practice, especially for software and AI inventions.[348] Certainly since *Alice*, computer-implemented AI patents have faced strict scrutiny from the patent office and the courts.[349]

Consequently, many AI innovators are finding it increasingly difficult to obtain patent protection, opting instead to maintain their inventions as trade secrets. While trade secrets lack the powerful exclusionary rights of patents because of the reverse engineering potential, they can protect a wider range of subject matter (including abstract ideas and discoveries).[350] Trade secrets are also not subject to the stringent disclosure requirements that exist when obtaining a patent, and they last as long as secrecy is maintained, which could be indefinitely.[351] Therefore, despite the lack of a "right to exclude" as with a patent, many AI innovators will rationally choose trade secret protection as a more viable and predictable path for securing their IP in light of the *Alice*/*Mayo* obstacles to patentability.[352]

---

[344] Alice Corp. Pty. Ltd v. CLS Bank Int'l, 573 U.S. 208 (2014); Mayo Collaborative Servs. v. Prometheus Lab'ies, Inc., 566 U.S. 66 (2012); 35 U.S.C. § 101. The U.S. Patent & Trademark Office is also bound to follow this same test when determining whether to grant or deny a patent application. *In re* Rudy, 956 F.3d 1379, 1382–83 (Fed. Cir. 2020); Bilski v. Kappos, 561 U.S. 593, 601 (2010) (applying eligibility doctrine to patent examination).

[345] *Alice*, 573 U.S. at 217.

[346] *Id.*

[347] *Id.*

[348] Nikola L. Datzov, *The Role of Patent (In)eligibility in Promoting Artificial Intelligence Innovation*, 92 UMKC L. REV. 1, 1 (2023); Kelley Sheehan, *How the U.S. Can Make Subject Matter Eligibility More Predictable and Free Artificial Intelligence Innovation*, 23 SAN DIEGO INT'L L.J. 341, 346–47 (2022); Ben Hattenbach & Gavin Snyder, *Rethinking the Mental Steps Doctrine and Other Barriers to Patentability of Artificial Intelligence*, 19 COLUM. SCI. & TECH. L. REV. 313, 317–18 (2018).

[349] Datzov, *supra* note 348, at 1.

[350] *Frequently Asked Questions: Trade Secrets*, WIPO, https://www.wipo.int/tradesecrets/en/tradesecrets_faqs.html [https://perma.cc/UG44-RLQC] (last visited Aug. 30, 2024).

[351] Datzov, *supra* note 348, at 49–51.

[352] *Id.*

C. Trade Secrecy as a Major Hurdle to Regulatory Goals of AI Transparency and Explainability

1. Tension Between Private Property Rights and Public Interest in Transparency

The tension between the AI developers' private property rights in their trade secrets and the public interest in governmental transparency presents a complex issue that has the potential to significantly hinder regulatory goals of transparency and explainability. As scholar Christopher Morten recently explained, "Trade secrecy law now seriously hinders federal regulators from disseminating to the public reliable information about the spheres of activity that they regulate, especially those that are technology-intensive."[353] As discussed above, many have called for requiring meaningful transparency into how these systems work to ensure fairness, accountability and even due process.[354] However, large system designers and software developers often rely heavily on trade secret law to shield details about their AI systems from disclosure, citing the need to protect their business information against both unfair competition and fraudulent behavior, such as hacking.[355]

Requiring trade secret disclosure to regulators who do not further disclose the information to interested parties or the public at large is an intermediary, likely unsatisfactory approach. Other industries already use this approach, such as the pharmaceutical and energy sectors. For pharmaceuticals, drug companies must disclose detailed information to the FDA for approval without full public disclosure.[356] For the energy sector, oil and gas companies provide environmental impact assessments and safety records to the EPA, which may not be fully transparent to the public due to trade secrets and confidentiality considerations.[357] While this approach

---

[353] Christopher J. Morten, *Publicizing Corporate Secrets*, 171 U. PA. L. REV. 1319, 1326 (2023) ("Trade secrecy law now seriously hinders federal regulators from disseminating to the public reliable information about the spheres of activity that they regulate, especially those that are technology-intensive.").

[354] *See, e.g.*, Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1349–53 (2018) (arguing that trade secret law should not be a barrier to disclosure about forensic AI evidence in criminal cases).

[355] *See* Joseph G. Milner, *Sunlight and Other Disinfectants: Disclosure Obligations Under the Federal Securities and Drug Regulatory Regimes*, 72 FOOD & DRUG L.J. 141 (2017); Stefan Scheuerer, *Artificial Intelligence and Unfair Competition: Unveiling an Underestimated Building Block of the AI Regulation Landscape*, 70 GRUR INT'L 834, 840 (2021).

[356] Christopher J. Morten & Amy Kapczynski, *The Big Data Regulator, Rebooted: Why and How the FDA Can and Should Disclose Confidential Data on Prescription Drugs and Vaccines*, 109 CAL. L. REV. 493, 541 (2021).

[357] *See* P. Stephen Gidiere, *Protecting Confidential Business Information in the Hands of Environmental Regulators*, 14 NAT. RESOURCES & ENV'T 262, 262 (2000) (discussing the issue generally of required disclosure of trade secret materials).

clearly allows for some regulatory oversight, it does not fully satisfy the transparency and explainability goals sought by many stakeholders in the AI context.

The government disclosure approach is problematic, *inter alia*, because regulators may lack the technical capacity or resources to adequately assess whether an AI system is sufficiently accurate, unbiased, and trustworthy to deploy in all but the most egregious cases.[358] Moreover, even if regulators were sufficiently competent, the public may not trust the government to make this critical determination without any visibility into the decision-making process.[359] Keeping the public in the dark about the existence and nature of AI systems making important decisions undermines democratic accountability and prevents those affected from raising concerns.[360]

As a baseline though, it is important to recognize that trade secret owners have due process rights against arbitrary deprivation of their property interest by the government. Compelling disclosure thus requires more than just an agency demand—it requires adequate justification and process. That said, trade secrets are not an absolute form of property, and the government can require disclosure in appropriate circumstances, subject to constitutional constraints.[361] The key is that the disclosure must be justified by a legitimate regulatory purpose and not go further than necessary to achieve that purpose.

2. Trade Secret Protection and Government Disclosure

The Freedom of Information Act ("FOIA") is a federal law that provides the public with the right to request access to records from any federal agency.[362] Under FOIA, federal agencies are required to disclose requested information unless the information fits within one of the limited statutory exceptions.[363] Under FOIA Exemption 4, an agency can withhold trade

---

[358] *See id.*

[359] *But see* Robin Feldman, *Artificial Intelligence and Cracks in the Foundation of Intellectual Property*, UC San Francisco L.J. (2024) (forthcoming 2024) (suggesting that the federal government could provide certification services for AI).

[360] Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1533–34 (2013).

[361] See Elizabeth A. Rowe, *Striking a Balance: When Should Trade-Secret Law Shield Disclosures to the Government?*, 96 IOWA L. REV. 791, 840 (2010) ("[T]rade secret law is not designed to foster absolute secrecy.").

[362] 5 U.S.C. § 552; Justin Cox, *Maximizing Information's Freedom: The Nuts, Bolts, and Levers of FOIA*, 13 N.Y.C. L. REV. 387, 387 (2010) ("As written, the FOIA erects a relatively simple process for gaining access to the wealth of information possessed by the Executive Branch of the federal government.").

[363] FREEDOM OF INFORMATION ACT (FOIA), *FOIA.Gov (Freedom of Information Act) Frequently Asked Questions (FAQ)*, https://www.foia.gov/faq.html [https://perma.cc/J6UK-EX9U] (last visited June 24, 2024) ("Federal agencies are required to disclose any information requested under the FOIA unless it falls under one of nine exemptions which protect interests such as personal privacy, national security, and law enforcement.").

secrets from public disclosure if they are "commercial or financial information obtained from a person and privileged or confidential."[364]

When information is submitted voluntarily to the government, it is easier for agencies to subsequently disclose it to the general public. Such information will be deemed "confidential" and thus exempt from FOIA disclosure only if it "is of a kind that would customarily *not* be released to the public by the person from whom it was obtained."[365] By contrast, if the submission is compulsory, the agency must show that disclosure will impair its ability to obtain necessary information in the future or substantially harm the submitter's competitive position.[366]

To show that information is "confidential" under FOIA Exemption 4, the agency must prove that disclosure is likely "to cause substantial harm to the competitive position of the person from whom the information was obtained."[367] This creates a balancing test between the competitive harm to the submitter and the public's interest in disclosure. Agencies who wish to make a disclosure of this received information must, therefore, articulate a sufficient justification for why disclosure is necessary despite potential commercial harm.[368]

Even if information is exempt from disclosure under FOIA, an agency can still choose to release it unless another statute prohibits disclosure.[369] One relevant statute is the Trade Secrets Act, which makes it a crime for federal officials to disclose trade secrets and other confidential business information "to any extent not authorized by law."[370] An agency's decision to disclose exempt information is thus constrained only by the Trade Secrets Act, and any other applicable statute, not simply FOIA itself.[371]

3. Disclosure Requirements for Voluntary and Compulsory Submissions

In the context of government AI systems, agencies should have the authority to require disclosure of certain information, including source code, data, and other technical details necessary to assess whether the systems are

---

[364] 5 U.S.C. § 552(b)(4).

[365] Critical Mass Energy Project v. Nuclear Regulatory Comm'n, 975 F.2d 871, 873 (D.C. Cir. 1992) (en banc) (emphasis added).

[366] McDonnell Douglas Corp. v. U.S. Dep't of the Air Force, 375 F.3d 1182, 1187 (D.C. Cir. 2004) (distinguishing between information that is "required" versus "voluntarily submitted").

[367] Nat'l Parks & Conservation Ass'n v. Morton, 498 F.2d 765, 770 (D.C. Cir. 1974).

[368] *Id.*

[369] Chrysler Corp. v. Brown, 441 U.S. 281, 294 (1979).

[370] 18 U.S.C. § 1905.

[371] *See* U.S. DEP'T OF JUSTICE, DEPARTMENT OF JUSTICE GUIDE TO THE FREEDOM OF INFORMATION ACT 686 (2009) ("FOIA provides only a right of access to agency records; it imposes no limits on the agency's discretion to disclose information.").

accurate, reliable, fair, and comport with due process. Such disclosure would serve the important public interest in ensuring these systems are trustworthy and not producing arbitrary or discriminatory results.[372]

However, the legal challenge against the Corporate Transparency Act ("CTA") in *National Small Business United v. Yellen* illustrates how businesses may resist broad disclosure requirements.[373] In the context of AI systems, this resistance includes transparency requirements not tied to a narrow justification.[374] The plaintiffs in this case, representing small businesses, argued that the CTA's reporting requirements exceeded Congress's constitutional authority and violated various due process and other fundamental rights.[375] The district court agreed, finding that the CTA was not sufficiently tied to Congress's enumerated powers, such as the Commerce Clause or taxing power, and thus unconstitutional.[376] This case highlights the potential for businesses to challenge disclosure laws on constitutional grounds, arguing that they exceed the scope of legislative authority.[377] As with the CTA, businesses developing AI systems may argue that broad disclosure mandates are unconstitutional because they are not sufficiently related to commerce, national security, or other areas of congressional power.[378] They may also invoke constitutional rights, such as freedom of speech or protection against uncompensated takings, to resist disclosure.

4. Transparency and Secrecy in Other Contexts

The balance between transparency and trade secrecy remains a contentious and crucial issue in various high stakes areas, particularly when AI systems are deployed with significant societal impacts. One example in the criminal justice context involves pretrial risk assessment tools where a court found no due process violation when the developer invoked trade secret protection to withhold how the algorithm weighed various factors.[379] These

---

[372] *See* Danielle Keats Citron, *Open Code Governance*, 2008 U. CHI. LEGAL F. 355, 358 (2008).

[373] Nat'l Small Bus. United v. Yellen, No. 5:22-CV-1448-LCB, 2024 WL 899372, at *21 (N.D. Ala. Mar. 1, 2024).

[374] *Id.*

[375] *Id.*

[376] *Id.*

[377] *Id.*

[378] *See* Brent Mittelstadt et al., *Transparency Requirements for Artificial Intelligence Between Legal and Ethical Aspects*, 6 BIG DATA & SOC'Y 1, 7–8 (2019) (discussing potential legal challenges to AI transparency requirements based on trade secrets and intellectual property rights).

[379] State v. Loomis, 881 N.W.2d 749, 761–64 (Wis. 2016); *see* Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265 (2020) (discussing government use of algorithms and the implications for transparency and accountability); Elizabeth A. Rowe & Nyja Prior, *Procuring Algorithmic Transparency*, 74 ALA. L. REV. 303 (2022) (examining the challenges and legal considerations in ensuring algorithmic transparency); Lena Chan, *The Weaponization*

risk assessment tools are used to make fundamental decisions about individual liberty; those affected have a strong interest in understanding and challenging their accuracy and potential for bias.[380] Professor Rebecca Wexler argues against trade secret protections being privileged in criminal proceedings, as they can impede defendants' rights to due process and to confront evidence used against them.[381] The key in such cases is to publicly disclose at least what is necessary to serve the public interest and permit accountability. In litigation, protective orders strictly limiting access and use are commonplace.[382] This approach provides information to defense counsel, independent experts, or special masters under seal—but not the general public.

While this Article's discussion of AI transparency and disclosure has primarily focused on government use, especially in the criminal justice system, these issues arise in several other contexts where AI systems can significantly impact individuals and society. For example, in the healthcare system, AI is increasingly being used to simulate drug effects, design medical devices, and support clinical decision-making.[383] Patients and providers have a strong interest in understanding how these systems work to assess their safety, efficacy, fairness, and reliability. Flawed or biased AI could lead to improper denial of care, disparate outcomes, or unsafe treatments.[384] However, developers of healthcare AI tools often resist disclosure by invoking trade secrecy and IP protections, citing the high cost of development and major reverse engineering concerns.[385]

---

of Trade Secret Law, 124 COLUM. L. REV. 703, 706 (April 2024) (analyzing the impact of trade secret law on transparency in criminal proceedings); Natalie Ram, *Innovating Criminal Justice*, 112 NW. U. L. REV. 659 (2018) (discussing the role of technology in criminal justice and the need for transparency and oversight); Simon Chesterman, *Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity*, 69 AM. J. COMP. L. 271 (2021) (discussing the challenges posed by AI's opacity in legal contexts).

[380] *Loomis*, 881 N.W.2d at 761–64; Bloch-Wehba, *supra* note 379; Rowe & Prior, *supra* note 379; Chan, *supra* note 379, at 706; Ram, *supra* note 379; Chesterman, *supra* note 379.

[381] Wexler, *supra* note 354, at 1353.

[382] *See* People v. Superior Court (Chubbs), B258569, 2015 WL 139069, at *6–7 (Cal. Ct. App. 2015). Disclosure of source code was ordered in this criminal case, but only under strict protective order and showing of particularized need. *Id.*

[383] *See* W. Nicholson Price II, *Artificial Intelligence in Health Care: Applications and Legal Issues*, 14 SCI. TECH. LAW. 10, 10–11 (2017) (discussing various healthcare AI applications).

[384] *See* Sharona Hoffman & Andy Podgurski, *Artificial Intelligence and Discrimination in Health Care*, 19 YALE J. HEALTH POL'Y L. & ETHICS 1, 17–29 (2020) (detailing different forms of bias in medical AI).

[385] *See* Price, *supra* note 383, at 16–17 (noting that FDA generally protects confidential business information); Arti K. Rai, *Machine Learning at the Patent Office: Lessons for Patents and Administrative Law*, 104 IOWA L. REV. 2617, 2638–39 (2019) (discussing FDA's position that it needs access to proprietary data to assess medical software products).

In the context of creative works, AI systems are used to generate content like novels, articles, and artwork, which raise difficult questions about authorship, ownership, and IP rights.[386] As AI-generated works proliferate and become indistinguishable from human-created works, there may be a stronger public interest in knowing which works were developed by AI and what data the AI was trained on.[387] In this context, the government's claim of necessity for public disclosure of AI-generated work will likely carry less weight than FDA disclosure internal requirements—especially when balanced against the creator's trade secrecy rights. The common thread in all these examples is that AI systems are being deployed in socially impactful domains where the public has an interest in transparency, accountability, and oversight. While developers' trade secrecy interests are relevant, they may merit less weight in the balance against government AI use cases implicating individual liberty and due process rights.[388]

While transparency and public disclosure are often viewed as inherent public goods, there are compelling arguments that this principle should not be universally applied, particularly in the context of certain advanced AI systems. As discussed in Part II.B.2.d, the Center for Law & AI Risk contends that "near-future AI systems threaten to cause—via intentional misuse, accidents, or autonomous action—large-scale harms to human life, limb, and freedom."[389] Open-sourcing of frontier AI models strips away the ability to mitigate certain risks. The idea here is easy to understand in the national security context. Some national security secrets are likely maintained for the benefit of society—largely because the information could be used against the public and national interest. In the AI foundational model context, there are serious concerns that nefarious actors could learn the exact buttons to push to reach the negative AI result they seek and thus disrupt AI alignment.

5. Achieving Transparency through Government Contracting

In situations where AI developers and users participate in government contracting, agencies themselves can proactively seek to achieve transparency through contract negotiations to ensure disclosure rights.[390] Inserting terms into service agreements that provide for access and auditing by the agency, as well as controlled disclosure to litigants with particularized needs, would

---

[386] 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. ST. L. REV. 659, 662 (2017).

[387] *Id*. at 707–11 (proposing disclosure obligations for AI-generated works to prevent deception and support accountability).

[388] *See* Lemley & Casey, *supra* note 27, at 770.

[389] *Comments on Dual Use Foundation Artificial Intelligence Models*, *supra* note 164, at 1.

[390] *See* Sonia K. Katyal, *The Paradox of Source Code Secrecy*, 104 CORNELL L. REV. 1183, 1250–57 (2019) (suggesting procurement contracts could be used as a transparency-forcing mechanism).

prevent an AI developer from blocking all transparency and oversight. Moving forward, the government will likely be a major user and investor in AI technology and thus have substantial bargaining power as a market participant to negotiate terms when acquiring AI systems.

V. CONCLUSIONS: LIMITATIONS AND RISKS OF RELYING ON IP FOR AI REGULATION

AI is a transformative technology that holds immense potential to benefit society, but it also poses significant risks and challenges. As AI systems become more sophisticated and ubiquitous, there is a growing need for effective governance frameworks to ensure safe, responsible, and ethical development and deployment. While IP rights, such as copyright, patents, and trade secrets, play an important role in incentivizing innovation and shaping the AI landscape, this Article argues that relying too heavily on IP frameworks as the primary mechanism for AI regulation would be misguided and potentially counterproductive.

As discussed in Part I, the core thesis of this Article is twofold. First, despite its recognized role in guiding innovative behaviors within the AI landscape, IP does not serve as an effective mechanism for directly regulating AI. Second, the relationship between IP rights and AI regulation can be pernicious, as IP rights may hinder AI regulation and development in several ways, such as through stringent copyright on training data, overprotection of trade secrets, nationalistic differences creating international loopholes, and the absence of rights associated with AI outputs.

Part II of this Article explored the various reasons why AI regulation is necessary, including concerns over privacy violations, bias and discrimination, job displacement, safety and control issues, and the concentration of power in a few large tech companies. While these issues are pressing and may require regulatory or legislative attention, we should resist the temptation to extend the role of IP as a means to address these non-IP concerns. IP rights are fundamentally designed to incentivize innovation and creativity, not to directly regulate the complex societal implications of emerging technologies like AI.

This Article also focused on the roles of copyright and trade secrecy in the context of AI. As discussed in Part III, copyright law has faced challenges in adapting to the use of copyrighted works as training data for AI systems. The fair use doctrine, which allows for limited use of copyrighted material without permission from the copyright holder, is likely to play a crucial role in determining the legality of such practices. On the other hand, an overreliance on copyright to regulate AI training data could lead to unintended consequences, such as stifling AI innovation and development.

Similarly, trade secret law, as examined in Part IV, has the potential to hinder meaningful transparency, accountability, and oversight of AI systems. While companies have legitimate interests in protecting their proprietary information, overly broad assertions of trade secrecy could create significant barriers to much-needed scrutiny of AI systems, particularly in high-stakes

contexts where AI makes decisions that could be materially harmful. The tension between the goals of AI governance and the private property rights of AI developers is a complex issue that requires careful balancing.

It is important to recognize that IP law is not inherently antithetical to sound AI governance. In some ways, IP rights can complement regulatory goals by providing economic incentives for innovation, promoting knowledge dissemination, and offering malleable policy tools that can be adapted to better align with AI governance principles. However, this Article contends that IP should play a supporting role in AI governance, not be the primary legal and regulatory lever. While IP rights can help incentivize beneficial AI development, they are not well-suited to address the full range of risks and challenges AI poses.